UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

REGRESSION MODELING OF GROUND-WATER FLOW

OPEN-FILE REPORT 85-180

REGRESSION MODELING OF GROUND-WATER FLOW

By Richard L. Cooley and Richard L. Naff

U.S. GEOLOGICAL SURVEY

Open-File Report 85-180

Lakewood, Colorado

1985

UNITED STATES DEPARTMENT OF THE INTERIOR

DONALD PAUL HODEL, Secretary

GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

Ground-Water Branch
U.S. Geological Survey
National Center, MS 411
Reston, VA 22092

Copies of this report can
be purchased from:

Open-File Services Section
Western Distribution Branch
U.S. Geological Survey
Box 25046, Mail Stop 306
Denver Federal Center
Denver, CO 80225
Telephone: (303) 236-7476

PREFACE

Scientists and engineers have been using ground-water flow models to study ground-water flow systems for more than twenty years. The basic modeling process seems to be relatively obvious and straightforward: Initially, a sound conceptual model is formed and is translated into a tractable mathematical model. Contributing to (and following) this conceptualization process is the collection of field information such as 1) location and extent of hydrostratigraphic units, recharge areas, discharge areas, and system boundaries; 2) hydraulic head measurements; and 3) pumping discharges. These data form the basis for input to the flow model. Finally, the model is run and the desired information such as head distribution or flux rates is extracted. However, it has usually been observed by people engaged in modeling that two pervasive problems considerably complicate the situation. One problem is that good, general methods of measuring (or computing) some of the variables that characterize the flow system and its geologic framework do not exist. One good example is measurement of ground-water recharge. No direct ways of measuring recharge exist, and the accuracy of indirect methods is often unknown. Furthermore, many indirect methods are applicable only to unique situations. The other problem relates to errors in the measurements and their propagation into model results. No error-free measurement (or computation) methods for obtaining data on the flow system exist. Thus, even the variables that can be estimated will contribute error, so that model results will always be unreliable to some (usually unknown) extent. As a consequence of these two problems, generally it is not possible to simply measure (or compute) the necessary input variables, apply them to an adequate model, and calculate the desired results to an acceptable

accuracy. Other methods that recognize and deal with the problems of incompleteness and (or) inaccuracy of data must also be applied. The present text has been designed to teach these methods to scientists and engineers engaged in ground-water modeling.

The basic methodology is multiple nonlinear regression, in which the regression model is some type of ground-water flow model. As will be seen subsequently, this methodology is consistent with known aspects of the physical systems to be analyzed and requires relatively few assumptions. Even though the present text is directed specifically toward ground-water modeling, the procedures to be discussed are applicable to a number of different types of modeling problems. Thus, the methods are usually discussed in a general context (in other words, without reference to any specific model).

Material in the present text evolved from notes developed for training courses in parameter estimation for ground-water flow models taught by the writers and others at the U.S. Geological Survey National Training Center, Denver Federal Center, Lakewood, Colorado. The philosophy of these courses and of this text is to teach general methods that are applicable to a wide range of problems, and to teach these methods in sufficient depth that students can apply them to many problem situations not considered in the courses or this text.

The main body of the text is organized into six major sections. The first section is an introduction that discusses the general topic of modeling ground-water flow. This section shows that ground-water modeling problems are an incomplete combination of direct-type problems (solution for hydraulic head given values of flow system and framework variables) and inverse-type problems (solution for flow system and framework variables

given values of hydraulic head) that, in general, require solution by optimization procedures, which give the best fit between observed and calculated results. Because the specific optimization approach employed here is regression, and regression procedures are based on statistical concepts, the second section is included to provide the student with the necessary statistical background material. This section is not designed to be an exhaustive review of basic statistics; rather, it presents material essential to understanding the following sections. The third section presents detailed material on linear and nonlinear regression. Although most of the material on linear regression is fairly standard, some of the material on nonlinear regression is not. In particular, specific modifications presented to induce convergence of the Gauss-Newton procedure for nonlinear regression have not, to the writers' knowledge, been presented elsewhere in the form given here. The fourth section applies the nonlinear regression method to the specific problem of developing a general finite-difference model of steady-state ground-water flow. In the fifth section statistical procedures are given to analyze and use general linear and nonlinear regression models. The tests and analytical procedures presented are not exhaustive; they are the ones that the writers have found to be most useful for analyzing the real systems examined to date. The sixth section is designed to be supplemental to the preceding sections. Specialized procedures presented include nonlinear regression for models that cannot be solved directly for the dependent variable, a measure of model nonlinearity called Beale's measure, and a statistical test for compatibility of prior information on parameters and parameter estimates derived from sample (observed head) information.

A number of exercises have been included, and a complete discussion of the answers can be found in the seventh major section at the end of the text. These problems exercise the student on nearly all methods presented. In addition, three computer programs are documented and listed. These three are the nonlinear-regression ground-water flow model of section four, a program to calculate Beale's measure, and a program to calculate simulated errors in computed dependent variables such as hydraulic head.

The mathematical background necessary to use this text includes basic mathematics through differential and integral calculus, including partial derivatives, and matrix algebra. A background in elementary statistics would be useful, but is not essential. In addition, a sound knowledge of ground-water hydrology and ground-water flow modeling are needed to effectively apply the methods presented.

References for cited material are given at the end of each major section. Good supplemental sources for the unreferenced material not peculiar to this text are presented as Additional Reading at the end of each reference list. It is expected that students who have difficulty with the material in this text will consult the more expanded developments in the sources indicated as Additional Reading.

Several people contributed extensively to this text in addition to the writers. Charles R. Faust wrote earlier sections on statistics review and basic regression and contributed several exercises; Steven P. Larson wrote an earlier version and documentation of the nonlinear-regression flow program of section four and contributed earlier versions of several exercises; James V. Tracy contributed to the documentation of the nonlinear-regression flow program; and Thomas Maddock III wrote the first version of the statistics review section. In addition, all of these people helped

teach the training courses from which the present text evolved. Finally, the writers would like to thank the technical reviewers Brent M. Troutman and Allan L. Gutjahr for their many hours of review work, and secretaries Anita Egelhoff, Evelyn R. Warren, and Patricia A. Griffith for their patience and care in typing the manuscript.

METRIC CONVERSION FACTORS

For those readers who prefer to use metric units, conversion factors for terms used in this report are listed below:

Multiply	*By*	*To obtain*
foot (ft)	0.3048	meter (m)
square foot (ft^2)	0.09290	square meter (m^2)
foot per day (ft/day)	0.3048	meter per day (m/day)
square foot per second (ft^2/sec)	0.09290	square meter per second (m^2/sec)
square foot per day (ft^2/day)	0.09290	square meter per day (m^2/day)
gallons per day per foot (gpd/ft)	0.01242	square meter per day (m^2/day)
cubic foot per second (ft^3/sec)	0.02832	cubic meter per second (m^3/sec)
cubic foot per day (ft^3/day)	0.02832	cubic meter per day (m^2/day)
gallons per minute (gpm)	0.00006309	cubic meters per second (m^3/sec)
ounces per ton	31.25	grams per metric ton[1]

[1] 1 metric ton = 1 megagram

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

Page

1. INTRODUCTION

1.1. FLOW EQUATION AND BOUNDARY CONDITIONS

The most general form of the ground-water flow equation that we will consider here is given as

$$\frac{\partial}{\partial x}\left(T_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T_{yy}\frac{\partial h}{\partial y}\right) + R(H-h) + W$$

$$+ \sum_{\ell=1}^{NW} \delta(x-a_\ell)\delta(y-b_\ell)Q_\ell = S\frac{\partial h}{\partial t} \qquad (1.1\text{-}1)$$

where

$T_{\xi\xi}(x,y)$ = transmissivity ($K_{\xi\xi}b$) in the ξ = x or y direction;

$K_{\xi\xi}(x,y)$ = hydraulic conductivity of the aquifer in the ξ direction;

$b(x,y)$ = thickness of the aquifer;

$R(x,y)$ = hydraulic conductance (hydraulic conductivity divided by thickness) of sediments underlying a stream or of an aquitard underlying or overlying the aquifer;

$W(x,y,t)$ = source-sink term (positive for a source), distributed areally;

$\sum_{\ell=1}^{NW} \delta(x-a_\ell)\delta(y-b_\ell)Q_\ell$ = Dirac delta designation for NW wells, each one pumping at rate $Q_\ell(t)$ (positive for injection) and located at (a_ℓ, b_ℓ);

$S(x,y)$ = storage coefficient;

$h(x,y,t)$ = hydraulic head in the aquifer;

$H(x,y,t)$ = head at the stream bottom or at the distal side of the aquitard;

x,y = Cartesian coordinates;

t = time;

and T_{xx} and T_{yy} are continuous functions of x and y.

With suitable internal boundary conditions, the region can be zoned with respect to $T_{\xi\xi}$. Such boundary conditions involve head and specific discharge times thickness normal to the boundary (q_n) and can be stated for a boundary between $T_{\xi\xi}$ zones k and ℓ as

$$(h)_k = (h)_\ell \tag{1.1-2}$$

$$(q_n)_k = (q_n)_\ell \tag{1.1-3}$$

where $(\ )_k$ indicates that the quantity in parentheses is evaluated just within the k side of the boundary and similarly for ℓ. Zonation with respect to R, W, or S requires no internal boundary conditions.

External boundary conditions applying at the periphery of the domain being modeled are given as

$$T_{xx} \frac{\partial h}{\partial x} n_x + T_{yy} \frac{\partial h}{\partial y} n_y + \alpha h = \beta \tag{1.1-4}$$

where $\alpha(x,y,t)$ and $\beta(x,y,t)$ are given functions, and $\{n_x(x,y), n_y(x,y)\}$ is the outward-pointing unit normal at the boundary. The sum of the first two terms is the flux q_B normal to the boundary (positive for outflow). Equation (1.1-4) incorporates the standard boundary conditions of specified flux (q_B) and specified head (h_B) but also allows for linear combinations to be given.

1.2. TYPES OF SOLUTIONS

1.2.1. Direct Solution for Head

The classical problem of mathematical physics (and, by assumption, of ground-water hydrology) is to directly solve (1.1-1) through (1.1-4) for $h = h(x,y,t)$. Given that any specific problem is properly posed, such a solution will always exist. The conditions for properly posing a problem are:

1) The positions of all internal boundaries are known exactly. Examples of internal boundaries are abrupt changes in $T_{\xi\xi}$, R, S, or W; internal known flux (q_n) boundaries; and internal known head boundaries. Note that a river is often treated as either an internal known head boundary where the river is assumed to have no width, or a zone of differing R where each bank is a zone boundary.

2) The positions and types of all external boundaries are known exactly. External boundaries are frequently known flux (q_B) types or known head (h_B) types. Sometimes some linear combination is known.

3) Hydrogeologic variables $T_{\xi\xi}$, R, and S and hydrologic variables W and Q_ℓ are known at all points in the region.

4) All boundary-condition variables H, α, and β are known. The initial head (at $t = 0$) is a boundary condition and must also be known.

It is obvious that ground-water flow problems are not actually of the classical type because none of the conditions cited above ever are met exactly. Conditions 1 and 2 are often most closely fulfilled, but estimates (often crude) usually must suffice for the variables in

conditions 3 and 4. Any errors in these input variables are propagated directly into the solution. However, it can be shown that reasonable (but incorrect) estimates of the variables will yield errors in predicted h(x,y,t) that have the characteristic of being bounded (that is, they do not tend to plus or minus infinity). Also, as the errors in the input variables tend to zero, the errors in computed head do also.

1.2.2. Inverse Solution for Parameters

An inverse solution involves solving (1.1-1) through (1.1-4) for one or more of the variables $T_{\xi\xi}$, R, S, W, Q_ℓ, α, or β, over the region; these variables are termed <u>parameters</u> here. Because R, S, W, Q_ℓ, α, and β are not involved in derivatives, theoretically they may be solved for algebraically. Unless $T_{\xi\xi}$ is constant it is involved in derivatives and thus must be obtained by solving a differential equation. To understand this, note that (1.1-1) may be rearranged to give

$$a\frac{\partial T_{xx}}{\partial x} + b\frac{\partial T_{yy}}{\partial y} + cT_{xx} + dT_{yy} - F = 0 \qquad (1.2\text{-}1)$$

where

$$a(x,y,t) = \frac{\partial h}{\partial x} \qquad b(x,y,t) = \frac{\partial h}{\partial y}$$

$$c(x,y,t) = \frac{\partial^2 h}{\partial x^2} \qquad d(x,y,t) = \frac{\partial^2 h}{\partial y^2}$$

$$F(x,y,t) = S\frac{\partial h}{\partial t} - R(H-h) - W - \sum_{\ell=1}^{NW} \delta(x-a_\ell)\delta(y-b_\ell)Q_\ell$$

In general, conditions for finding R, S, W, Q_ℓ, α, or β are:

1) Conditions 1 and 2 for the classical direct solution are met.

2) Head distribution h(x,y,t) is known exactly.

3) The desired combination of parameters to be obtained is unique.

The latter condition is completely problem dependent. Because solution involves only algebraic manipulations, the condition reduces to the requirement that the system of algebraic equations involving the desired parameters has a unique solution. Generally, solution involves picking the required number of points spatially and through time to yield the necessary number of equations.

In order to find T_{xx} and T_{yy}, more conditions are required than for finding R, S, W, Q_ℓ, α, or β. These conditions are:

1) Conditions 1 and 2 for finding R, S, W, Q_ℓ, α, or β must be met.

2) The direction of the velocity vector must be known everywhere, or T_{xx}/T_{yy} must be known everywhere, or quantities a, b, c, d, and F in (1.2-1) must be known at two (or more) points in time to give a unique solution to (1.2-1) written in the form of a pair of simultaneous differential equations. These requirements result because (1.2-1) is one equation in two unknowns. Hence, an additional relationship is required. If the velocity direction is known everywhere, then by employing Darcy's law the additional relationship is derived as

$$\frac{T_{xx}}{T_{yy}} = \frac{q_x b}{q_y a} \qquad (1.2\text{-}2)$$

where a, b, and q_x/q_y (the ratio of the x and y direction fluxes) are known.

3) Either T_{xx} or T_{yy} must be known on a curve crossing all flowlines. This is an extension of the Cauchy boundary condition for a first-order differential equation involving a single

dependent variable and is the required boundary condition for (1.2-1) and (1.2-2). Note that the curve does not have to be continuous.

4) The function F in (1.2-1) must be known everywhere. This means that all quantities in F must be known. If a mathematical form for F can be assumed, then the solution can be found to within a constant multiple of the correct solution. Sometimes additional relationships for the quantities in F can be used to find R, S, W, and (or) Q_ℓ.

Ground-water flow modeling does not fit into the category of inverse solutions, although a significant part of most model studies is to find values of the parameters that allow values of calculated head to match those observed in the field. The difficulty is that the required conditions are almost never met. Head distribution is never known exactly because measurements do not exist at all points and, where these measurements do exist, they are not exact. Furthermore, some measure of $T_{\xi\xi}$ is virtually never available for a curve crossing all flowlines, and information on directions of flow vectors for even scattered locations usually is nonexistent. Assumptions concerning zonations in which T_{xx}/T_{yy} and (or) T_{xx} and T_{yy} may be considered constant simplify the problem, but the fact that h must be known still remains.

Because the head distribution is not known exactly, coefficients a, b, c, d, and F in (1.2-1) are in error. Furthermore head in all of these quantities appears as a derivative. Hence, any error in h is propagated into the inverse solution as a derivative of error. The effects of this propagation are often disastrous because, if ε_h is defined as error in head, $\varepsilon_h \to 0$ does not imply that $\partial\varepsilon_h/\partial\xi \to 0$. Also it is common that

$|\partial\varepsilon_h/\partial\xi| >> |\varepsilon_h|$, and it can happen that $|\partial\varepsilon_h/\partial\xi| \to \infty$ even if ε_h is bounded. Therefore, the error in computed $T_{\xi\xi}$ (or other parameter) may not approach zero as $\varepsilon_h \to 0$, and may, in fact, be quite large (see Neuman, 1980, p. 342-344).

1.2.3. Solution Using Real Data

In the previous section it was argued that problems involving ground-water flow modeling of real field systems are neither of the classical nor inverse type because the data necessary for the problems to be classified as either type are usually lacking. There usually exists an estimate of the hydraulic head distribution based on measurements (that are in error with respect to the model) taken at selected points. Estimates of the parameters are usually either completely unknown or have been obtained by spot measurements, few of which are directly useful for construction of appropriate effective values for use in (1.1-1). It is evident that modeling problems in ground-water hydrology involve an incomplete combination of several types of data in which error and error propagation are important considerations.

1.3. SOURCES OF ERROR IN GROUND-WATER DATA

Uncertainty (or errors) in ground-water data may have many sources, and enumeration of all possible sources would be a nearly impossible task. However, a consideration of some of the more important sources of error serves to illustrate the importance of the error component.

1.3.1. Sources of Error in Head Data

Some major potential sources of random-appearing error in head data with respect to the model ((1.1-1) through (1.1-4)) are:

1) Areal ground-water models assume that the head used is the average over the vertical. However, wells may not be open over the entire interval modeled, and if they are, they may not measure the average. Strong flows into or out of a well distort the hydraulic head field in the vicinity of the well so that the recorded water level does not represent the average head.

2) Permeability varies from point to point, which causes water levels to vary from values they would have if permeability were uniform. However, models usually do not take this detailed variation into account. Bakr and others (1978), Gutjahr and others (1978), and Smith and Freeze (1979a, 1979b) have recently studied this phenomenon.

3) Water levels measured in wells in use may contain unknown amounts of residual drawdown. In addition, unused wells may be near wells that are in use, with resulting unknown drawdown in the unused well.

4) Measurement of well-head elevation may be in error.

Actual total error from the above sources is highly problem dependent, but it is easy to imagine errors of several feet. It should be noted that measurement error in water levels was not mentioned as a major source of error because it usually amounts to one- or two-tenths of one foot or less. Finally, major model error in (1.1-1) through (1.1-4) (for example, head

dependence in one or more parameters or three dimensional flow) was also not mentioned because error resulting from this source is bias and should be detected and eliminated by analysis of model results.

1.3.2. Sources of Error in Parameter Data

Because there are several different parameters to be considered and each can be estimated or measured in several different ways, a large number of sources of error exist in parameter data. Model error is not considered here, but other types of bias are potentially important and are often difficult to detect. Some examples of errors in parameter data illustrate the nature of the problem:

1) Too few estimates of parameters are available to compute stable estimates of statistics, such as mean and variance.

2) Results of point sampling are often biased because a large amount of data does not necessarily allow computation of nearly true or effective values of a parameter and its variance. For example, permeability values from core analyses often are not representative of regional values because flow through large fractures is not reproduced by core analyses. Also, effective values of a parameter and its variability are usually not directly given by standard mean and variance formulas.

3) Transmissivities estimated from specific capacity data collected by drillers are subject to numerous sources of error. Common sources include (1) mismeasuring water levels or pumping rates, (2) allowing the water level to recover after bailing, (3) clogging of the slots or screen, and (4) inaccurate reporting. There are so many sources of error that the errors may often

almost be random. A persistent source of bias results because drillers drill wells in favorable locations and only screen (or slot) the most productive zones.

4) Transmissivities and storage coefficients estimated from pumping test analysis are subject to many of the same errors as above, but the more carefully controlled tests should reduce their frequency and magnitude. In addition, a single test may not be representative of an entire hydrostratigraphic unit.

5) Transmissivities estimated from lithological data are usually biased to an unknown degree.

1.4. MODEL CONSTRUCTION

Ground-water models are constructed by using the types of data alluded to in the previous section. Hence, it is rare that measured or estimated parameter data are either reliable or complete enough to employ directly in a model to reproduce measured head data with an acceptable model fit. As a result, adjustment of parameter values, and sometimes basic model structure, is used to improve model fit. There are two basic groups of methods currently in use to accomplish this: (1) trial and error procedures, (2) optimization methods that minimize a formal objective function.

1.4.1. Trial and Error Methods

Trial and error is the method of repeated simulation until the calculated head distribution obtained with a reasonable set of parameters fits closely enough to satisfy the analyst. Sometimes an objective measure of goodness of fit, such as $\sum (h^{calc} - h^{obs})^2$, is used to aid the analyst in deciding whether or not a change in parameters (or model

structure) has improved the overall model fit. However, no matter how it is applied, the method has several inherent critical deficiencies:

1) No methodology exists to guarantee that the simulations will proceed in a direction that could lead to the best set of parameters.
2) It is difficult to determine when that best set has been reached.
3) No practical way of determining how many other sets of parameters could yield similar correspondence between h^{calc} and h^{obs} exists.
4) It is difficult to decide whether or not additional parameters or a more refined model would significantly improve model fit.
5) No way of quantitatively assessing the predictive reliability of the model exists.

A method of model construction that addresses these deficiencies would allow construction and use of a model with a much greater degree of confidence than that provided by trial and error methods. Hence, attention is turned to formal optimization procedures.

1.4.2. Formal Optimization Procedures

Optimization procedures utilize a formal criterion of goodness of fit, often called an objective function. This function is minimized (or sometimes maximized, depending on the form of the function) with respect to the parameters to yield an optimum or best-fit solution. Minimization (or maximization) sometimes is subject to certain other criteria regarding values that the parameters, or pertinent functions of the parameters related to the model, may take on. These criteria are called constraints.

Examples of objective functions are

$$\sum_{\ell=1}^{n_s} (h_\ell - \hat{h}_\ell)^2$$

$$\sum_{\ell=1}^{n_s} |h_\ell - \hat{h}_\ell|$$

$$\max_{\ell} |h_\ell - \hat{h}_\ell|$$

$$\sum_{\ell=1}^{n_s} w_\ell (h_\ell - \hat{h}_\ell)^2 + \sum_{m=1}^{n_p} k_m (p_m - \hat{p}_m)^2$$

where

h_ℓ = observed head,

$\hat{h}_\ell$ = calculated head,

p_m = observed or prior estimate of a parameter,

$\hat{p}_m$ = calculated parameter value that when used in the model produces $\hat{h}_\ell$,

w_ℓ = weight related to the reliability of the observation h_ℓ,

k_m = similar weight applied to p_m,

n_s = number of observations of head, and

n_p = number of observations of parameters.

The last example is called a compound objective function because it contains both head and parameters explicitly. Note that minimization of each of the functions with respect to the parameters of the model produces a solution that is overall a best fit to the data, according to the objective function. If the signs of the functions were changed, maximization would produce the same result.

Examples of constraints are

$$p_m^L < \hat{p}_m < p_m^U$$

$$a\hat{p}_k + b\hat{p}_m + c\hat{p}_n = f$$

$$a\hat{p}_k + b\hat{p}_m + c\hat{p}_n < f$$

where a, b, c, and f are constants or known functions; superscript L refers to a lower limit; and superscript U refers to an upper limit. The best-fit solution obtained by minimizing (or maximizing) the appropriate objective function must simultaneously satisfy the appropriate constraints.

Because the solution obtained by an optimization procedure has known properties, it may be analyzed. The exact procedures used and the extent to which the model may be analyzed depend on the type of optimization method selected for use. Statistical regression procedures handle, on a probabilistic basis, the propagation of data errors (with respect to the model) into the estimates of parameters and predictive capability of the model. Methods have been developed for estimating parameters, testing assumptions made during development of techniques, testing model fit, determining the reliability and significance of the model and the parameters contained in it, effecting corrective measures for violation of some assumptions, and estimating the reliability of predictions to be made with the model. These procedures and the statistical background necessary to apply them are detailed in the remainder of the text.

References Cited

Bakr, A. A., Gelhar, L. W., Gutjahr, A. L., and MacMillan, J. W., 1978, Stochastic analysis of spatial variability in subsurface flows, 1--Comparison of one- and three-dimensional flows: Water Resources Research, v. 14, no. 2, p. 263-271.

Gutjahr, A. L., Gelhar, L. W., Bakr, A. A., and MacMillan, J. W., 1978, Stochastic analysis of spatial variability in subsurface flows, 2--Evaluation and application: Water Resources Research, v. 14, no. 5, p. 953-959.

Neuman, S. P., 1980, A statistical approach to the inverse problem of aquifer hydrology, 3--Improved solution method and added perspective: Water Resources Research, v. 16, no. 2, p. 331-346.

Smith, L., and Freeze, R. A., 1979a, Stochastic analysis of steady-state groundwater flow in a bounded domain, 1--One-dimensional simulations: Water Resources Research, v. 15, no. 3, p. 521-528.

_____ 1979b, Stochastic analysis of steady-state groundwater flow in a bounded domain, 2--Two-dimensional simulations: Water Resources Research, v. 15, no. 6, p. 1543-1559.

Additional Reading

Freeze, R. A., and Cherry, J. A., 1979, Groundwater: Englewood Cliffs, N.Y., Prentice-Hall, p. 15-77, 356-359.

2. REVIEW OF PROBABILITY AND STATISTICS

Casual observation of our environment indicates that many phenomena are not strictly predictable. We cannot, for instance, exactly say what the maximum air temperature at any particular location will be tomorrow, although we might be able to give a probable range. This probable range might be based on our past experience, which would enable us to say that tomorrow's high, considering the location and season, will probably fall within a specified interval. A more sophisticated forecasting model may enable us to reduce the range within which we think tomorrow's high will fall, but random elements in the forecasting procedure would preclude giving an exact answer. As another example of randomness, consider the toss of a coin. Prior to the toss, we can only give the possible outcomes, either a head or a tail, and, if the coin is fair, say that either have equal likelihood of occurring. However, this ability to state precisely that any future outcome of this experiment can, with equal probability, result in either a head or a tail, is an important advantage over that offered for predicting tomorrow's maximum temperature; in this latter case, because of the complex nature of the processes resulting in tomorrow's maximum, it is unlikely that we could give a precise statement concerning the probability that it will fall in our predicted interval. Instead of attempting to untangle these complexities, we might opt to study the past history of maximum temperatures at the location and annual date in question. By assuming that this past history will extrapolate into the future (that is, that weather dynamics in future years will remain essentially unchanged from those in previous years), we could give an

estimate of the likelihood that tomorrow's maximum will fall in a particular interval. However, tools need to be developed to carry out this investigation.

2.1. BASIC CONCEPTS

Randomness itself can be considered to be centered around an experiment; the outcome of the experiment will have a random quality attached to it. For example, in a coin-toss experiment, the outcome is dominated by the random element (either a head or a tail). On the other hand, many experiments have a large deterministic factor. For example, in a chemical titration experiment we measure the unknown and the amount of titrant used, then calculate the amount of a specific substance in the unknown. However, measurement error creeps into our experiment and we see that, from realization to realization of this same experiment, the results vary. Some experiments, such as annual peak river flows, are not ours to perform but only to observe. This experiment is an example of an event in nature that has a large random component which nature provides; as we attempt to measure these flows, we introduce additional randomness, which we generally ignore. Hydraulic conductivities measured from core samples are similar to peak flows; nature has already provided for randomness, which is constrained by certain deterministic factors such as type of source material, distance of transport, climate and diagenesis. Again, for every realization of this experiment, measurement error is introduced, which may not be small.

All possible outcomes of an experiment are known as its sample space. The sample space of a coin-toss experiment consists of either a

head (H) or tail (T):

$$S = \{H,T\}$$

If the experiment consists of the toss of two coins, then the sample space consists of

$$S_1 = \{H,H),\ (H,T),\ (T,H),\ (T,T)\}$$

On the other hand, if we are only interested in the total number of heads which might result from a single toss of two coins, we could define the experiment as this sum, which would result in the sample space

$$S_2 = \{0,1,2\}$$

In the case of S_1, every member of the sample space is equally likely to occur, while for S_2, a one is twice as likely to occur as either zero or two (provided that the coin is fair, of course).

The sample space for a hydraulic-conductivity experiment could be defined as all positive real numbers; that is, measurements from cores might result in values (outcomes) which could be as small as zero, or, if we stretch our imaginations, infinitely large. This space could be considered to be a continuous equivalent of the S_2 space for the two-coin experiment. That is, a porous medium is an extremely complex random process itself. By conducting hydraulic conductivity measurements on cores, we quantify this randomness in much the same way that counting heads quantifies an outcome of the two-coin experiment. However, by quantifying the randomness of the porous medium in this manner, we have never investigated the possible existence of more basic, perhaps nonnumeric sample spaces similar to S_1 of the two-coin experiment for a porous medium. Even if we were to discover the existence of such a space, we

would then need to find a rule, or algorithm, which would allow us to connect the two spaces. Suffice it to say, we shall not worry unduly about the possibilities of an S_1-like space for may processes; however, when they are available, they provide an excellent mechanism for investigating the characteristics of S_2-like spaces.

An event is defined as any subset of the sample space. The investigator is usually interested in the relative frequency of occurrence of an event. In the case of the S_2 space and the two-coin experiment, it is apparent that half the time a realization of the experiment should result in a one. This event is equivalent to the event in the S_1 space corresponding to the union of (H,T) and (T,H), which occurs with a relative frequency of one half. Thus the relative frequency of a head occurrence for the two-coin experiment is not dependent upon the definition of the sample space, but on the basic randomness controlling the experiment.

The investigator is frequently confronted with the problem that he needs a numerical result for the outcome of a random, not necessarily numerical, experiment. In the case of coin-toss experiments, it is seen that the basic outcome is a particular arrangement of heads and(or) tails. By assigning a head a value of one and a tail a value of zero and then summing, it is possible to translate these basic results into something measurable. It is precisely this process of assigning a numerical value to a nonnumerical outcome that leads to the definition of a random variable.

Definition: A random variable is a function whose value is a real number determined by each element in a sample space.

When the outcome of the experiment is numerical, then this result can be considered to be the random variable (this statement is merely a special case of the above definition). From the above definition it should also be clear that a mathematical transformation of a random variable is also a random variable. (Throughout this review, a random variable will be indicated by an upper case English or Greek letter, while a value that it may take on will be indicated by another letter, usually a lower case of the same type as used for the random variable.)

The concepts of a random experiment, sample space, and random variable are flexible. For instance, if in the case of the toss of two coins, the experiment is defined as the total number of heads appearing, then the S_2 sample space is an automatic result and the random variable can also be considered to be this result. However, if the experiment is defined to be the arrangement of heads and(or) tails resulting from a toss (that is, the S_1 space), then the same effect can be obtained by letting the random variable over the S_1 space be a function that assigns a one to a head and a zero to a tail and then sums the result. The investigator usually defines the sample space, or experiment, to suit a particular objective. As a matter of convenience, the space is usually selected such that the relative frequencies of occurrence of events within the space are definable. Access to such basic sample spaces as S_1 for the two-coin experiment allow for the calculation of relative frequencies for events in both S_1 and S_2; without the existence of a space like S_1, it is difficult if not impossible to determine the true relative frequency of occurrence for an event in S_2. This is also evident from the hydraulic-conductivity experiment, where only an S_2-like sample space is available to the investigator.

A random variable can also be described as either being discrete, as in the coin-toss experiment, or continuous, as represented by the hydraulic-conductivity experiment. A discrete random variable is defined over a sample space whose elements are discrete, although they may be as many as there are whole numbers (mathematicians refer to this phenomenon as being countably infinite). A continuous random variable is defined over a continuous sample space whose elements are infinite in number (therefore these elements are uncountably infinite).

2.2. FREQUENCIES AND DISTRIBUTIONS

2.2.1. Discrete Random Variables

Although frequencies of occurrence are usually associated with events in a sample space, they are also associated with values of random variables, since random variables are functions of the elements in a sample space. That is, particular values of a random variable correspond to particular events in the sample space and, therefore, have frequencies of occurrence. Even though we will speak of the relative frequency of occurrence for particular values of a random variable, we are in reality speaking of a corresponding event in the sample space. In fact, we frequently use a range of values of a random variable to define an event in a sample space, thus avoiding the task of describing which elements of the sample space compose the event.

Frequencies of occurrence for events in many discrete sample spaces can be deduced from the following axiomatic premise:

If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the relative frequency of occurrence of A is n/N.

As a simple example of employment of this premise, consider an experiment consisting of a toss of a die. The sample space consists of the integers one through six and, for any realization of the experiment, each element of the sample space has equal likelihood of occurrence. By considering each element of the sample space to be an event, it is possible to calculate the frequency of occurrence, $f(x_i)$, with which a random variable takes on the value x_i. For this experiment, only the integer values one through six of x_i have frequencies of occurrence other than zero; $f(x_i)$ can be graphically represented as shown in figure 2.2-1.

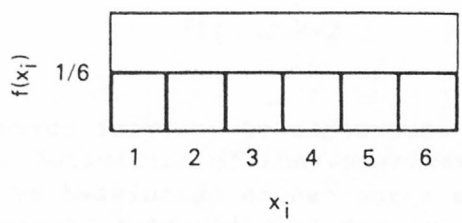


fig. 2.2-1

In this case, $f(x_i)$ is referred to as the discrete density function of the discrete random variable consisting of the outcome of a toss of a single die.

When two dice are cast, the experiment can either be defined as the sum that results from the toss, or simply as all possible arrangements that could appear on the dice. If the sum is chosen, then the sample space consists of the integers two through twelve, which would also be the range

of values that the random variable could take on. The elements of this space, however, are not equally likely to occur. The sample space consisting of all arrangements of the numbers appearing on the two dice, presented graphically in table 2.2-1, has elements which are equally likely to occur.

Table 2.2-1

		First die					
		1	2	3	4	5	6
Second die	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

The relative frequency of occurrence of an event corresponding to any subset of elements in this space can be calculated by using the premise concerning equally likely outcomes.

A random variable, consisting of the sum that results from any outcome of the two-dice experiment, takes on the integer values two through twelve over the sample space represented by table 2.2-1. The discrete density function for this random variable can now be derived from the basic premise concerning outcomes that are equally likely, since each value for this discrete random variable corresponds to a particular event consisting of a particular subset of elements in the sample space indicated by table 2.2-1. Thus, for example, the value of x_i equal three corresponds to the event

containing the elements (2,1) and (1,2), and has a relative frequence of occurrence of 2/36. Letting x_i represent the integer values that this random variable can obtain, its density function $f(x_i)$ can be represented as shown in figure 2.2-2.

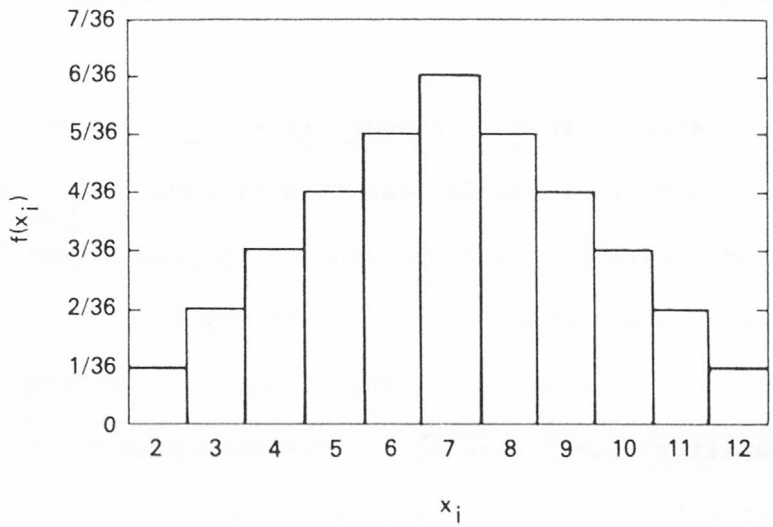


fig. 2.2-2

Note that had the first definition of the experiment been used, then every element of the sample space, consisting of the integers two through twelve, would have frequencies of occurrence, when considering each element as an event, equivalent to those shown in figure 2.2-2.

Frequencies of occurrence, or deduced frequencies of occurrence as indicated in figures 2.2-1 and 2.2-2, are indications of the future. We can make probability statements concerning the possibility of a random variable taking on future values from such knowledge. In a craps (two-dice) game, we know that the probability of rolling a natural, an outcome of seven or eleven on the first cast, is 2/9 simply because these values of the random variable for the two dice experiment correspond to

elements in the sample space which occur with a relative frequency of 2/9. Formally, the statement that this discrete random variable X take on the values of seven or eleven with a probability of 2/9 is written

$$P(X = 7 \text{ or } X = 11) = 2/9$$

The probability that this random variable takes on any integer value between two and twelve is obtainable directly from its frequency density, figure 2.2-2.

A probability statement that is frequently encountered concerns the probability that a random variable is less than or equal to a specific value. For the random variable corresponding to the sum of outcomes of the cast of two dice, we may ask what is the probability that the random variable X is less than or equal to 5? The probability of this event is equal to the probability that X take on any integer value two through five:

$$P(X \leq 5) = P(X = 2 \text{ or } X = 3 \text{ or } X = 4 \text{ or } X = 5)$$

This probability is the sum of the probabilities of the individual events that X take on the integer values two through five:

$$P(X \leq 5) = 1/36 + 2/36 + 3/36 + 4/36 = 5/18$$

(If the student is not convinced of this relationship, he or she should examine the elements of the sample space represented by table 2.2-1 to ascertain that it holds.) Note that $P(X \leq 12)$ is unity; that is, an event which occurs with a probability of one will, without doubt, take place. A probability of zero indicates, on the other hand, that the event of concern cannot possibly occur.

The probability statement $P(X \leq a)$, where a is any real number, is given a special definition for both discrete and continuous random variables. That is, $F(a) = P(X \leq a)$ is known as the cumulative distribution function of the random variable X. For the case of the sum of outcomes for two dice, F(a) appears as illustrated in figure 2.2-3.

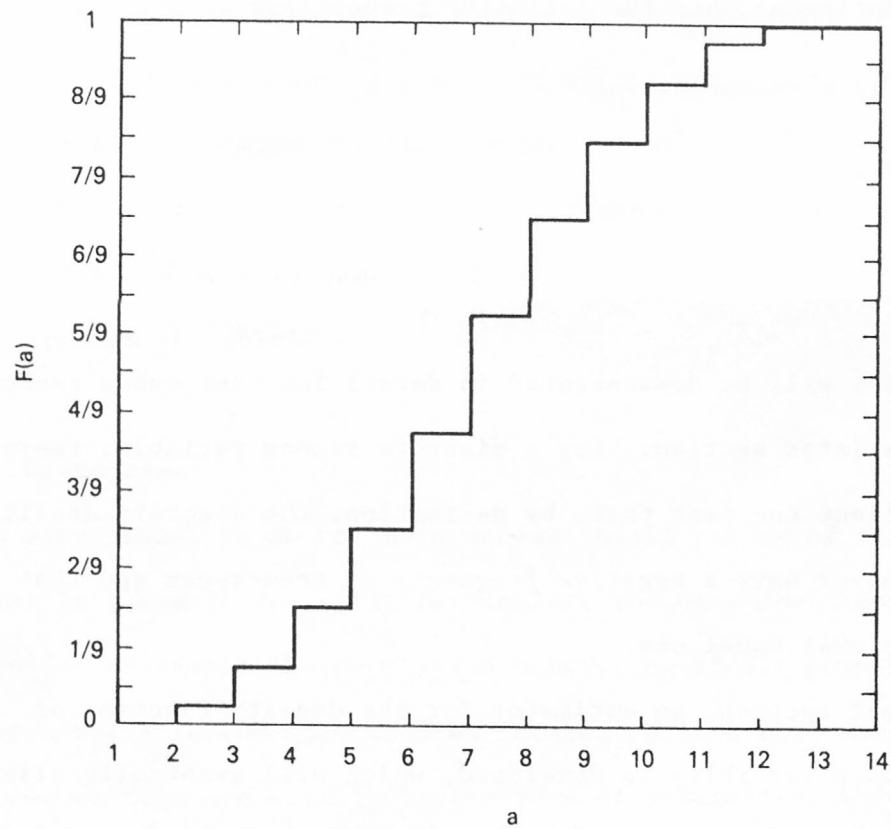


fig. 2.2-3

Because a random variable represents a functional mapping from the sample space to the real number space, we can be assured that the probability of the event $X \leq a$ exists and is equal to the sum of the probabilities of all

events corresponding to values of the random variable which are less than or equal to a. In general, for discrete random variables, the cumulative distribution function can be evaluated by summing the appropriate relative frequencies of occurrence:

$$F(a) = \sum_{x_I \leq a} f(x_I) \qquad (2.2\text{-}1)$$

The cumulative distribution function for all random variables, discrete or continuous, has the following properties:

1) $F(a)$ is a nondecreasing function of a,

2) $\lim_{a \to \infty} F(a) = 1$,

3) $\lim_{a \to -\infty} F(a) = 0$.

These properties will be demonstrated in detail for continuous random variables in a later section. For a discrete random variable, these properties reflect the fact that, by definition, the discrete density function can never have a negative frequency of occurrence and that the sum of frequencies must equal one.

In the next section, an estimator for the density function of continuous random variables is developed, which will eventually allow us to explore the nature of density and cumulative distribution functions of continuous random variables.

Problem 2.2-1

An urn contains one red, one white and two blue balls, all of equal dimensions. A ball is drawn from the urn, replaced, and then another draw is made.

a) What possible arrangements (red, white and(or) blue) of the two balls, considering order of selection, could occur (see, for example, table 2.2-1 for two dice)?

b) What is the frequency of occurrence of any of the above events? (Hint: let the balls be represented by the symbols R, W, B1, and B2.)

c) A value of one is assigned to a blue ball, two to a red ball and three to a white. A random variable consists of the sum of any outcome consisting of two draws with replacement. Develop a discrete density function for this random variable.

d) What is the probability that this random variable takes on a value of 4? What arrangements of balls correspond to this value of the random variable?

2.2.2. Histograms

In many cases, we do not have access to all values of random variables in a sample space (in particular, for many continuous random variables). We sample the population consisting of all possible values of the random variable and hope to draw inferences from this sample. The inferences we draw are usually in the form of statistics, which we refer to as sample statistics. We like to think that sample statistics estimate values of population parameters, which are constants reflecting the true frequency distribution of the random variable. This is frequently the case if the observations composing samples are made randomly and without bias. Samples composed of such observations are referred to as random samples, and are expected to be representative of the population.

Estimates of density functions for random variables are frequently made from random samples. While certain experiments, such as a coin toss, allow for the deduction of frequencies of occurrence of events, other experiments defy a theoretical calculation, forcing us to estimate from a random sample. These estimates, known as histograms, are generally constructed by repeating the experiment a large number of times (thus, sampling the population of all possible outcomes), dividing the range of these outcomes into class intervals and calculating the relative number of points that fall in each interval. We might imagine, for example, that we could watch a craps game and note the outcome of each roll of dice. After a thousand rolls, we would go away and calculate the relative percentage of each integer, two through twelve, which occurred. If these sample frequencies of occurrence were not close to that shown previously for the theoretical result, we would suspect that the dice had been tampered with.

As an example of a histogram constructed from observed values of a continuous random variable, consider the transmissivity data shown in table 2.2-2. Figure 2.2-4 represents a histogram constructed directly from these data, which constitute a random sample from the population of transmissivities as determined from specific capacities of wells in carbonate rocks of central Pennsylvania. A second histogram, figure 2.2-5, was constructed from a logarithmic transformation of these data as shown in table 2.2-3. The first histogram was constructed by using a class interval of 50,000 gpd/ft, and the second is based upon an interval of one-half a $\log_{10}$ cycle. The first histogram is not very illustrative because most of the wells have transmissivities less than 50,000 gpd/ft (the underlying population frequency is probably heavily skewed to the right). By logarithmically transforming of the random variable, we scale the abscissa so

as to remove the skewness in the histogram, causing it to be more bell-shaped. This type of transformation is frequently used on random variabl that have a zero lower bound, causing the transformed variable to have tails that tend to infinity in both directions. The transformation also tends to remove any right skewness in the frequency distribution of these random variables. With regard to the transformed variate, the histogram figure 2.2-5 suggests a bell-shaped population frequency distribution. More data and smaller class intervals, as suggested in the following paragraphs, should cause the histogram shown in figure 2.2-5 to approach its population shape, which we may suspect to be a normal distribution; the untransformed random variable would then result from a log-normal distribution.

Table 2.2-2. [1/]

Transmissivity gpd/ft	Log_{10} T	Transmissivity gpd/ft	Log_{10} T
15.0	1.176	2370.0	3.375
18.0	1.255	2440.0	3.387
21.0	1.322	2540.0	3.405
29.0	1.462	2800.0	3.447
32.0	1.505	2820.0	3.450
35.0	1.544	3380.0	3.529
50.0	1.699	4410.0	3.644
52.0	1.716	4520.0	3.655
56.0	1.748	5500.0	3.740
62.0	1.792	5650.0	3.752
84.0	1.924	6030.0	3.780
92.0	1.964	6240.0	3.795
106.0	2.025	6340.0	3.802
118.0	2.072	7290.0	3.863
142.0	2.152	8130.0	3.910
160.0	2.204	11000.0	4.041
175.0	2.243	13100.0	4.117
184.0	2.265	13700.0	4.137
202.0	2.305	14500.0	4.161
264.0	2.422	17200.0	4.236
354.0	2.549	17700.0	4.248

Table 2.2-2.--continued

Transmissivity gpd/ft	Log $_{10}$ T	Transmissivity gpd/ft	Log $_{10}$ T
370.0	2.568	19700.0	4.294
374.0	2.573	23100.0	4.364
455.0	2.658	24200.0	4.384
463.0	2.666	26400.0	4.422
515.0	2.712	33400.0	4.524
528.0	2.723	34700.0	4.540
615.0	2.789	42400.0	4.627
705.0	2.848	46300.0	4.666
753.0	2.877	52000.0	4.716
800.0	2.903	66500.0	4.823
984.0	2.993	68400.0	4.835
1150.0	3.059	132000.0	5.121
1290.0	3.111	152000.0	5.182
1500.0	3.176	423000.0	5.626
1580.0	3.199	423000.0	5.626
1670.0	3.223	528000.0	5.723
1850.0	3.267	528000.0	5.723
2310.0	3.364	528000.0	5.723

1/Adopted from Siddiqui, 1969, p. 433-436.

Table 2.2-3.

Class interval*	Number of occurrences	Relative frequency	Cumulative frequency
1.0 - 1.5	4	.051	.051
1.5 - 2.0	8	.103	.154
2.0 - 2.5	8	.103	.257
2.5 - 3.0	12	.154	.411
3.0 - 3.5	12	.154	.565
3.5 - 4.0	10	.128	.693
4.0 - 4.5	10	.128	.821
4.5 - 5.0	7	.090	.911
5.0 - 5.5	2	.026	.937
5.5 - 6.0	5	.064	1.001
	78		

*Based on $\log_{10} T$, table 2.2-2.

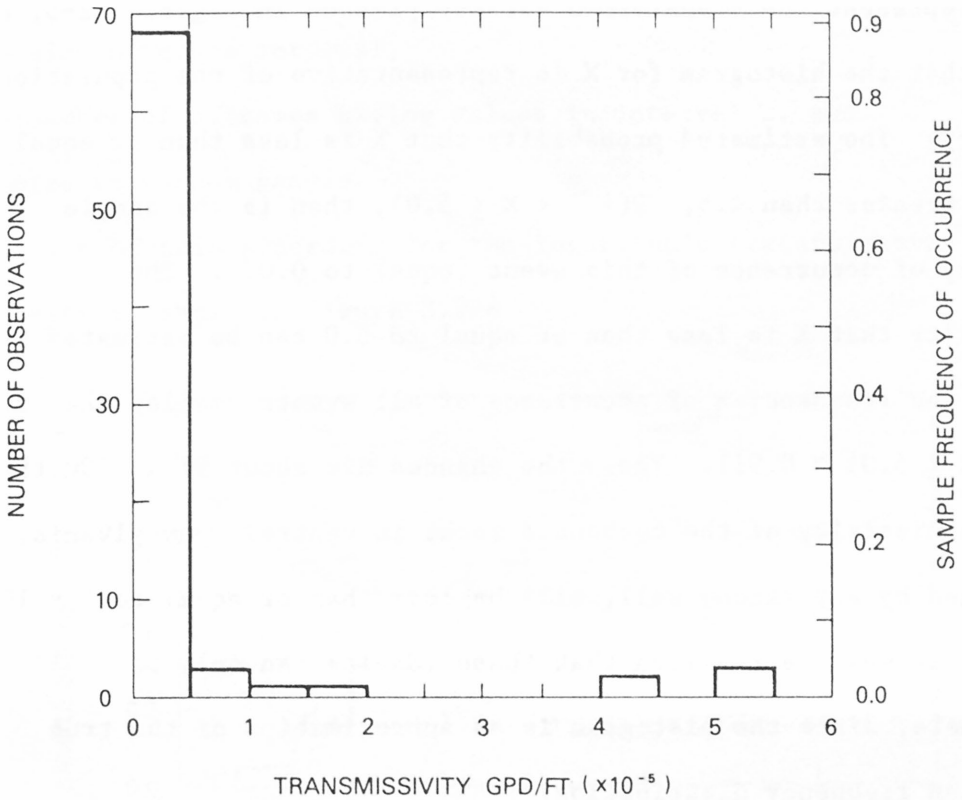


fig. 2.2-4

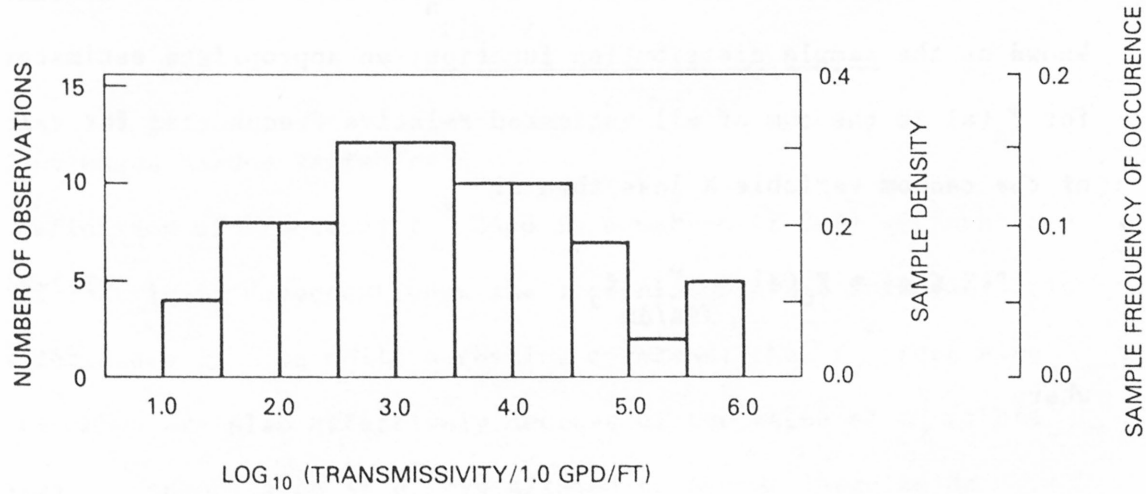


fig. 2.2-5

We are now in position to estimate the probability of occurrence of an event associated with the log-transformed random variable. Let $X = \log_{10} T$ represent the transformed variate plotted in figure 2.2-5, and assume that the histogram for X is representative of the population frequency. The estimated probability that X is less than or equal to 5.0 but greater than 4.5, $P(4.5 < X \leq 5.0)$, then is the sample frequency of occurrence of this event (equal to 0.09). The probability that X is less than or equal to 5.0 can be estimated by summing the frequencies of occurrence of all events smaller than 5.0; thus $P(X \leq 5.0) \approx 0.911$. Thus, the chances are about 91 in 100 that the transmissivity of the carbonate rocks in central Pennsylvania, as determined by any random well, will be less than or equal to 1×10^5 gpd/ft. It must be realized that these results can only be approximate, since the histogram is an approximation of the true population frequency distribution.

An estimate of the cumulative distribution function can also be constructed from a random sample. Let $F_n(a)$ represent this estimate, known as the sample distribution function; an appropriate estimator for $F_n(a)$ is the sum of all estimated relative frequencies for values of the random variable X less than a:

$$P(X \leq a) \approx F_n(a) = \sum_{i \leq a/\Delta x} f_i^* \qquad (2.2\text{-}2)$$

where

f_i^* = n_i/n = sample frequency of occurrence of an event represented by the ith class interval,

Δx = size of class interval,

n_i = number of outcomes having values in interval i, and

n = size of random sample.

An application of this procedure for the logarithmic transformation of transmissivity is shown in figure 2.2-6.

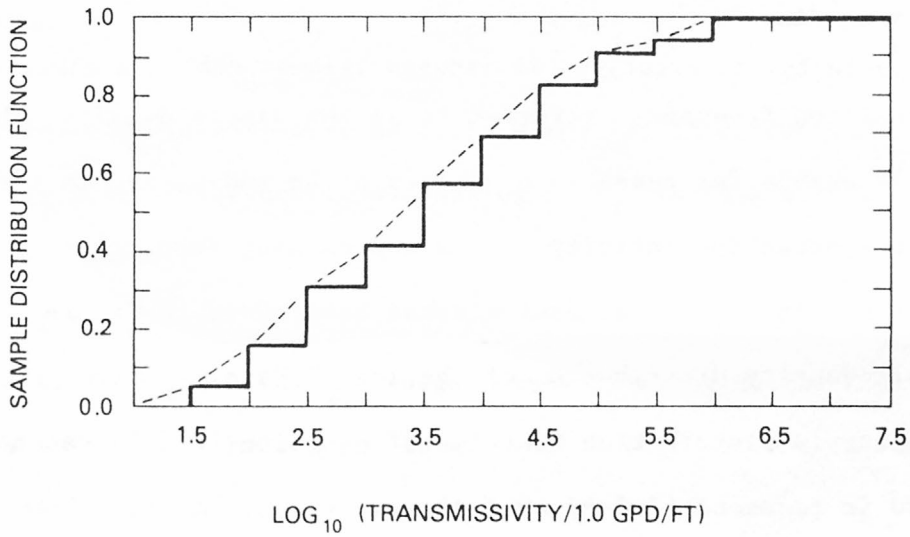


fig. 2.2-6

2.2.3. Continuous Random Variables

The definition of frequency f_i^* used in equation (2.2-2) suffers from the deficit that it is dependent upon the size of the class interval; that is, if Δx decreases in size while n remains constant, then f_i^* must also decrease since we are also effectively decreasing the value of n_i within this interval. Indeed, even if n were allowed to become large as Δx decreases, thus causing n_i for any arbitrary interval to be large, f_i^*

could still be made arbitrarily small by decreasing the interval size sufficiently. However, this phenomenon would prevent us from defining a frequency for a single point in a continuous random variable, unless we are content to associate it with some arbitrary class interval. To overcome this problem, probabilists have defined a different measure of frequency for continuous random variables that consists of the frequency of occurrence f_i^* scaled by its class interval:

$$f_i = f_i^*/\Delta x \tag{2.2-3}$$

This normalized frequency, referred to as the sample density, should be relatively stable for reasonable choices of Δx and n, and in the limiting case of n approaching infinity and Δx approaching zero, f_i should be constant. An additional ordinate has been added to figure 2.2-5 to show the sample-density distribution of the $\log_{10} T$ data.

The sample distribution function of equation (2.2-2) can now be redefined in terms of (2.2-3) as follows:

$$F_n(a) = \sum_{i \le a/\Delta x} f_i \, \Delta x \tag{2.2-4}$$

This definition lends itself to an exploration of the population equivalents of $F_n(a)$ and f_i, for if the random sample is of sufficient size to sample every member of the sample space and Δx is taken infinitely small, then the population equivalents of $F_n(a)$ and f_i should be approached. By letting n become large and Δx small, we see that

$$\sum_{i \le a/\Delta x} f_i \, \Delta x \approx \int_{-\infty}^{a} f(x) \, dx \tag{2.2-5}$$

where f(x), the population equivalent of f_i, is known as the probability density function. Since f(x) is the population equivalent of f_i, then the integral representation in (2.2-5) of summing these scaled frequencies must be the population equivalent of $F_n(a)$, which of course is the same cumulative distribution function defined earlier in section 2.2.1:

$$F(a) = \int_{-\infty}^{a} f(x)\, dx = P(X \leq a) \qquad (2.2\text{-}6)$$

However, because a random sample, whether it be finite or infinite, is countable, equation (2.2-5) must be given a special interpretation. Note that, because f(x) is the continuous analog of f_i, it is always a non-negative quantity.

A stronger statement than (2.2-5) can be made concerning the equivalence of F(a) and $F_n(a)$ for large sample sizes by noting that $F_n(a)$, prior to sampling, is a random variable. That is, if we were to collect different samples of the same size n from the same population, we would not expect that $F_n(a)$, computed from each random sampling, would have the same value. We would only hope that, as n becomes large, these different values would approach some constant. Indeed, probabilists have shown that, with a probability of one, $F_n(a)$ becomes the constant F(a) as n goes to infinity. This result is particularly remarkable if we first consider that $F_n(a)$ can only take on a countable number of values k/n, $0 \leq k \leq n$, where k is an integer (see equation (2.2-2)). Thus, while the values of F(a) are uncountably infinite (continuous), $F_n(a)$ can only be, in the case that the random sample is infinitely large, at most countably infinite. We will use this result loosely by allowing (2.2-5) to take on the indicated limits,

$$\lim_{\substack{\Delta x \to 0 \\ n \to \infty}} F_n(a) = F(a) \qquad (2.2\text{-}7)$$

and noting that this result only can occur with a probability of one.

Both f(x) and F(x) are continuous functions of values of the random variable X. For the previously illustrated case of $X = \log_{10} T$, the density function might appear as in figure 2.2-7.

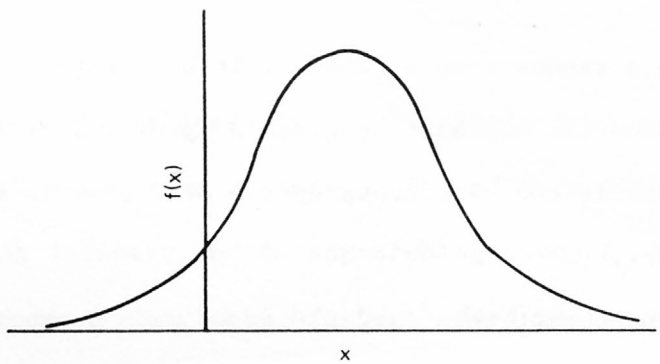


fig. 2.2-7

Figure 2.2-7 represents the population equivalent of figure 2.2-5, as if all possible outcomes of the random variable were available to us. Similarly, the cumulative frequency distribution, the population equivalent of figure 2.2-6, for this random variable might appear as in figure 2.2-8.

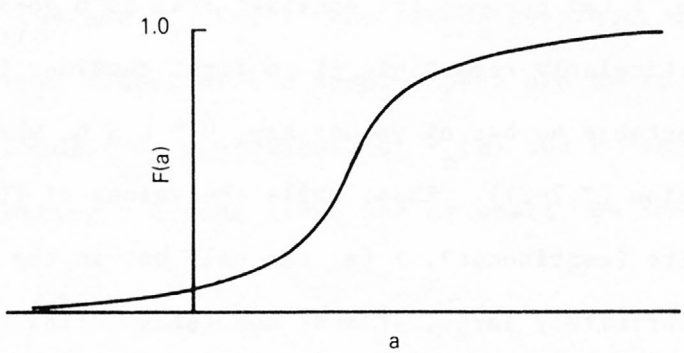


fig. 2.2-8.

Because of the relationship (2.2-6), the density function f(x) can be defined in terms of the cumulative distribution function F(x) by differentiation:

$$f(a) = \frac{dF(a)}{da} = \frac{d}{da}\int_{-\infty}^{a} f(x)\,dx \qquad (2.2\text{-}8)$$

This result follows directly from the fundamental theorem of integral calculus, and is applicable only to density functions of continuous random variables. Equation (2.2-8) is one of three concepts which define density functions of continuous random variables. The other two are that f(x) be greater than or equal zero for any possible value of the random variable and, as will be demonstrated in the next section, that the total mass under the frequency curve be unity. All density functions of continuous random variable have these concepts in common.

Problem 2.2-2

a) Construct histograms for the following specific conductance data using class intervals of 100 and 200 µmho/cm such that the abscissa and ordinate of both histograms are scaled equally. What is the effect of changing the class interval?

b) Construct a cumulative frequency distribution from your 100 µho/cm class-interval results. Let X represent the specific-conductance random variable; what is

$P(X \leq 600)$?

$P(X > 400)$?

$P(400 < X \leq 600)$?

$P(X \leq 1300)$?

Ordered specific conductance data.

Data in μmho/cm for wells in carbonate rocks of Maryland.[1/]

63	423	501	582	685	836
76	433	504	596	697	839
168	439	509	598	700	876
278	440	512	600	704	882
301	440	518	604	710	895
304	440	518	617	721	897
310	444	527	620	723	904
315	452	529	627	724	906
319	452	533	629	726	915
323	452	537	632	728	948
332	456	538	636	740	968
347	462	542	641	750	969
357	469	552	647	750	982
359	471	562	659	764	997
363	473	564	659	765	1030
389	477	564	661	779	1080
407	487	565	664	783	1106
408	490	566	665	789	1120
411	492	570	670	808	1170
413	493	575	673	808	1230
417	493	578	675	815	
418	499	582	677	820	

[1/] Adopted from Nutter, 1973, p. 63-68.

2.2.4. Properties of Cumulative Distribution Functions

In the previous section, it was noted that the cumulative distribution function F(a), defined by the probability statement $P(X \leq a)$, has the integral form (2.2-6) for continuous random variables. It is useful to have all manner of probability statements defined in terms of the cumulative distribution function, since this is a standard form. For this purpose, properties of cumulative distribution functions, with applications to other probability statements, are developed in this section.

The probability that a random variable X takes on a value in the interval (a,b] can be expressed in terms of cumulative distribution functions as

$$P(a < X \leq b) = \int_a^b f(x)\, dx = F(b) - F(a) \qquad (2.2\text{-}9)$$

This statement is a direct result of integral calculus, whereby integration is used to sum all the frequencies of occurrences of values of the random variable between a and b. It also follows from (2.2-9) that the cumulative distribution function is a nondecreasing function of x, since $0 \leq P(a < X \leq b) \leq 1$.

From equation (2.2-4) it should be fairly obvious that the total mass under the sample density curve f_i is unity; that is,

$$\lim_{a \to \infty} F_n(a) = \lim_{a \to \infty} \sum_{i \leq a/\Delta x} f_i \, \Delta x = 1 \qquad (2.2\text{-}10)$$

Since the probability density function f(x) of a continuous random variable X is a limiting form of the sample density f_i, the mass under its curve is also unity:

$$\lim_{a\to\infty} F(a) = \lim_{a\to\infty} \int_{-\infty}^{a} f(x)\, dx = 1 \qquad (2.2\text{-}11)$$

Equation (2.2-11) is a property of all cumulative distribution functions. Similarly,

$$\lim_{a\to-\infty} F(a) = \lim_{a\to-\infty} \int_{-\infty}^{a} f(x)\, dx = 0 \qquad (2.2\text{-}12)$$

which follows from integral calculus, is also a property of cumulative distribution functions.

Property (2.2-11) allows one to express $P(X > a)$ as

$$P(X > a) = \int_{a}^{\infty} f(x)\, dx = 1 - \int_{-\infty}^{a} f(x)\, dx = 1 - F(a) \qquad (2.2\text{-}13)$$

which is also a result of Riemannian integration. An alternate statement of (2.2-13) is that $P(X > a) = 1 - P(X \leq a)$.

By considering equation (2.2-9) in a limit form, we can also find the probability that $X = a$:

$$P(X = a) = \lim_{\Delta x\to 0} P(a < X \leq a+\Delta x)$$

$$= \lim_{\Delta x\to 0} [F(a+\Delta x) - F(a)] = 0 \qquad (2.2\text{-}14)$$

This result is unique to continuous random variables, in contradistinction to discrete random variables. It also follows from (2.2-14) that $P(X \leq a)$ is equivalent to $P(X < a)$ for continuous random variables, since the endpoint, a, of the semi-infinite interval does not contribute mass to the probability statement.

Probability statements (2.2-9), (2.2-11), and (2.2-13) can be demonstrated for discrete random variables by using the summation form (2.2-1) of the cumulative distribution function. In contradistinction to continuous random variables, the endpoint in $P(X \leq a)$ for a discrete random variable can contribute significant mass to the statement.

A number of frequency densities that result from randomness in nature, or probabilistic models of random events, have been investigated and published. Cumulative distributions of these densities are frequently tabulated and are to be found in many reference books on probability and statistics. Equations (2.2-9) and (2.2-13) are especially useful in evaluating probability statements of tabulated random variables.

2.2.5. An Example: The Normal Distribution

Let the random variable Y represent the amount of titrant used in a titration experiment to neutralize measured amounts of the unknown x. A scatter diagram of titrant versus unknown might appear as in figure 2.2-9.

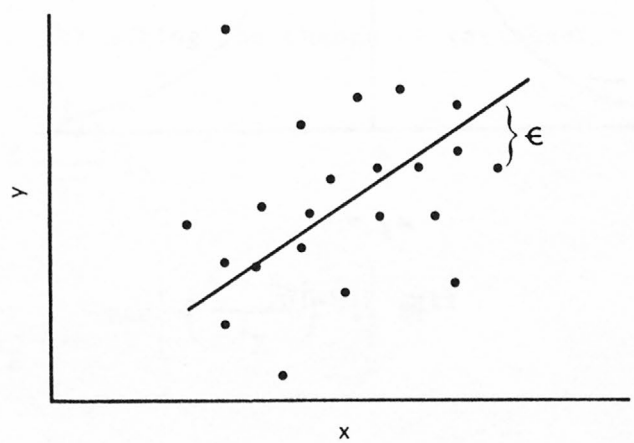
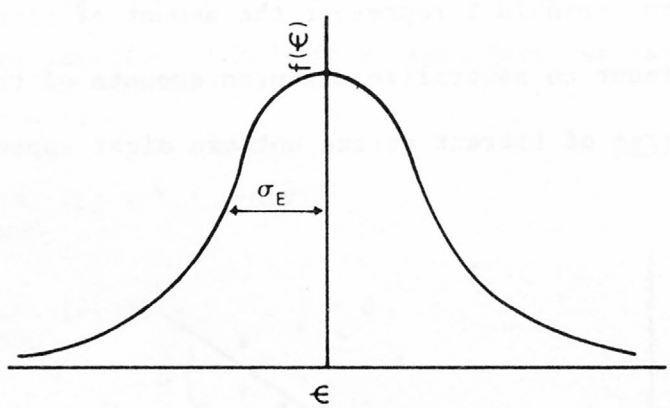


fig. 2.2-9.

The solid line represents the true stoichiometric balance between titrant and unknown. The dots, representing repetitions of the experiment, deviate from this line by an amount ε, which represents a value of the measurement error E. These errors represent a continuous random variable that could theoretically vary from $-\infty$ to $+\infty$ (the graphed points only represent a random sample from the population). If the experimental apparatus is functioning properly, however, we would expect these dots to be concentrated in the general vicinity of the solid line.

A distribution that is frequently used to model errors that are symmetrically distributed about some common point is the normal distribution. The density of the normal distribution is a bell-shaped curve, symmetric about its mean μ_E, and with most of the mass concentrated within one standard deviation σ_E of the mean (see figure 2.2-10).



$\mu_E = 0$

fig. 2.2-10.

In the case of the titration experiment, we would hope that the most frequently found value of the error would be near-zero, and expect that μ_E would equal zero. The standard deviation σ_E is a measure of the dispersion, or spread, of the errors about the mean and is equal to the distance from the mean to an inflection point on the curve $f(\varepsilon)$. The mean and standard deviation will be formally defined in a later section.

A normal random variable is frequently standardized with its mean and standard deviation by the following transformation:

$$Z = (E - \mu_E)/\sigma_E \tag{2.2-15}$$

The cumulative distribution for this standard normal random variable is tabulated (table 2.10-1) for use by the investigator since its probability density function, $f_Z(z)$, is parameter free:

$$f_Z(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \tag{2.2-16}$$

Given the density function for the standard normal random variable, it is natural to inquire about the form of density, $f_E(\varepsilon)$, of the unnormalized random variable E. Consider the cumulative frequency distribution for Z: By making the change of variables $z = (s - \mu_E)/\sigma_E$,

$$\begin{aligned} F_Z(a) &= \int_{-\infty}^{a} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz \\ &= \frac{1}{\sqrt{2\pi}\sigma_E} \int_{-\infty}^{\varepsilon} \exp\left[-\left(\frac{s - \mu_E}{\sigma_E}\right)^2/2\right] ds \end{aligned} \tag{2.2-17}$$

results, where $\varepsilon = a\sigma_E + \mu_E$ is a value of the unnormalized random variable. Since differentiation is the inverse operator of integration, equation (2.2-17) is differentiated with respect to ε to find $f_E(\varepsilon)$ (see also equation (2.2-8)).

$$f_E(\varepsilon) = \frac{d}{d\varepsilon} F_Z(a(\varepsilon)) = \frac{1}{\sqrt{2\pi}\sigma_E} \exp\left[-\left(\frac{\varepsilon - \mu_E}{\sigma_E}\right)^2 /2\right] \qquad (2.2\text{-}18)$$

Note that (2.2-18) is not parameter free, as this density is a function of the parameters μ_E and σ_E.

2.3. EXPECTATION AND THE CONTINUOUS RANDOM VARIABLE

The discussion in this section is largely presented with continuous random variables in mind. All the results, however, are applicable to discrete random variables; whenever a quantity is defined by an integration over a probability density function for the continuous case, this same quantity can almost invariably be defined by a summation over the discrete density function for the discrete case. It is left to the reader to demonstrate this fact.

2.3.1. The Mean

The mean is a measure of central tendency of a population. As an estimator of this central tendency, consider a finite random sample consisting of n values x_i of the random variable X. If the sample frequency of occurrence f_i^* is estimated from this random sample, then a logical estimator of the central tendency is to sum the product of the central value $\bar{x}_i$ of each class interval and the frequency of occurrence

for that interval:

$$\bar{x} = \sum_{i \le x_m/\Delta x} f_i^* \bar{x}_i = \sum_{i \le x_m/\Delta x} f_i \bar{x}_i \, \Delta x \qquad (2.3\text{-}1)$$

where x_m is the upper limit of the largest class interval necessary to construct f_i^*. The frequencies of occurrence f_i^* in equation (2.3-1) can be looked upon as weights that sum to one, and the quantities $\bar{x}_i$ as equally spaced values of the random variable. The values of the random variable that occur more frequently, as indicated by the random sample, receive larger weights through equation (2.3-1) and will have a greater influence on $\bar{x}$.

Equation (2.3-1) should be recognized by the reader as also being the definition for the center of mass of physical weights distributed along a line. That is, if mf_i^* represents the mass of a weight located at $\bar{x}_i$, where m is the total mass of all the weights, then equation (2.3-1) would give us the center of mass of the line with respect to the origin. In the case of a histogram, the role of the weights is played by the sample frequency of occurrence for an interval, which gives us the approximate relative likelihood that any future value of the random variable will occur in that interval. For calculation purposes, this distributed weight over any interval i is replaced by a point weight having the same mass as the distributed weight, but located at the center $\bar{x}_i$ of the interval. The sum of the products of the relative masses of these point weights, f_i, with their relative distances from the origin, $\bar{x}_i$, gives us the center of mass, which is also a measure of the central tendency. Of course, if

the sample size n were to become very large, then Δx could be made very small, refining (2.3-1) as an estimator of the central tendency of a random variable.

Reasoning similar to that leading to $\bar{x}$ as an estimator of the population mean can be applied directly to defining this parameter. First given that the density function f(x) is known, then the approximate frequency of occurrence of an event corresponding to an interval of size Δx that has as its central value $\bar{x}_i$ is $f(\bar{x}_i)\Delta x$. Thus, assuming that these relative frequencies are centered at each $\bar{x}_i$, an approximate measure of the population central tendency, μ_X, is

$$\mu_X \approx \sum_{(\text{all } i)} \bar{x}_i \, f(\bar{x}_i)\Delta x \qquad (2.3\text{-}2)$$

where the values $\bar{x}_i$ are equally spaced by Δx from each other. Of course, by letting Δx become smaller, a more accurate measure of μ_X is developed, until μ_X, also known as the expected value, $E(X)$, of the random variable X, is defined by the following integral expression:

$$\mu_X = E[X] = \int_{-\infty}^{\infty} xf(x)\, dx \qquad (2.3\text{-}3)$$

This equation is the standard form for the expected value of a univariate random variable.

Equation (2.3-3) can also be developed directly from (2.3-1) by letting $n \to \infty$ and $\Delta x \to 0$:

$$\mu_X = \lim_{\substack{n\to\infty \\ \Delta x\to 0}} \bar{x} = \lim_{\substack{n\to\infty \\ \Delta x\to 0}} \sum_{i \le x_m/\Delta x} \bar{x}_i \; f_i \, \Delta x = \int_{-\infty}^{\infty} xf(x)\, dx \qquad (2.3\text{-}4)$$

That is, as Δx becomes smaller and as the number of observations becomes very large, $\bar{x}_i$ becomes a unique continuous value of X, and f_i becomes the

continuous function f(x), and the summation can be replaced by an integration. As in the case of equation (2.2-7), we can only say that the limit indicated in (2.3-4) is reached with a very high probability as n becomes large; however, this probability should be unity as n becomes infinite.

It should be noted that μ_X is a population parameter that is characteristic of the random variable X, while $\bar{x}$, being derived from values of a finite random sample from the population of X, is only an estimate for μ_X. Estimators such as $\bar{x}$ will be developed in greater detail in a later section.

Problem 2.3-1

a) Find $\bar{x}$ from 100 μmho/cm histogram of problem 2.2-2.

b) Find μ_X for the random variable of problem 2.2-1.

2.3.2. Generalization and Application of the Expectation Operator

The operation of finding an expected value can be generalized by considering a function g(X) of continuous random variable X. If we wish to find the average effect of the function g(x) over the outcomes of a random sampling, we would again resort to the approximation

$$\overline{g(x)} = \sum_{i \leq x_m/\Delta x} f_i^* \, g(\bar{x}_i) \qquad (2.3\text{-}5)$$

That is, we weight g(x), where g(x) is evaluated at the center of every class interval, by the frequency of occurrence of that interval and sum all the weighted values of $g(\bar{x}_i)$. To obtain the population equivalent of $\overline{g(x)}$, n is taken to be large while Δx is taken to be small; this equivalent is denoted by the expectation symbol $E[g(X)]$:

$$E[g(X)] = \lim_{\substack{n\to\infty \\ \Delta x\to 0}} \overline{g(x)} = \int_{-\infty}^{\infty} g(x)f(x)\,dx \qquad (2.3\text{-}6)$$

Equation (2.3-6) represents the general form of the expectation operator for a univariate distribution when the random variable is continuous. A similar form exists for discrete random variables, in which the integration has been replaced by summation.

A trivial but useful property of the expectation operator is that the expected value of any constant c is that constant; for the continuous case, this is easily demonstrated as

$$E[c] = \int_{-\infty}^{\infty} cf(x)\,dx = c\int_{-\infty}^{\infty} f(x)\,dx = c \qquad (2.3\text{-}7)$$

where equation (2.2-11) has been invoked. A more important property of E is that it is a linear operator; that is,

$$E[ag_1(X) + bg_2(X)] = aE[g_1(X)] + bE[g_2(X)] \qquad (2.3\text{-}8)$$

For the continuous case, it is easily seen that this property results because integration itself is a linear operator:

$$\int_{-\infty}^{\infty} (ag_1(x) + bg_2(x))f(x)\,dx = a\int_{-\infty}^{\infty} g_1(x)f(x)\,dx + b\int_{-\infty}^{\infty} g_2(x)f(x)\,dx \qquad (2.3\text{-}9)$$

As a practical example of finding the expected value of a random variable, consider the problem of finding the mean of X where X is a normal random variable with mean μ_X and standard deviation σ_X:

$$E[X] = \frac{1}{\sqrt{2\pi}\sigma_X} \int_{-\infty}^{\infty} xe^{-\left(\frac{x - \mu_X}{\sigma_X}\right)^2/2}\,dx \qquad (2.3\text{-}10)$$

By a change of variable $z = (x - \mu_X)/\sigma_X$, we see that (2.3-10) becomes

$$E[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu_X + \sigma_X z)\, e^{-z^2/2}\, dz = \mu_X \qquad (2.3\text{-}11)$$

since $z.\exp(-z^2/2)$ is an odd function of z in the interval $(-\infty,\infty)$, and (2.3-7) holds (μ_X, being a population parameter, is constant). Result (2.3-11) is the reason why μ_X is defined to be $E[X]$.

2.3.3. The Variance, Standard Deviation and Coefficient of Variation

While the mean μ_X is a measure of the central tendency of a random variable, it gives no information as to how frequently a random variable will be encountered in its vicinity. The variance σ_X^2, defined as the expected value of the function $g(X) = (X - \mu_X)^2$, is a population parameter that quantifies this concept. The variance can also be looked upon as an operator that is defined in terms of another operator (the expectation operator) as follows:

$$\sigma_X^2 = Var[X] = E[(X - \mu_X)^2]$$

$$= \int_{-\infty}^{\infty} (x - \mu_X)^2\, f(x)\, dx \qquad (2.3\text{-}12)$$

where $Var[X]$ represents an operator that operates on X. Intuitively, it can be seen that σ_X^2 is the sum of the frequency weighted deviations, which have been squared, from the mean. As such, it represents the amount of dispersion of the random variable about the mean: When σ_X^2 is relatively large, then a random variable is less likely to have values in the immediate vicinity of the mean. The standard deviation σ_X is simply the square root of the variance: $\sigma_X = (Var[X])^{\frac{1}{2}}$.

By exercising the linear property of the expectation operator, definition (2.3-12) cam be expressed in an alternate form:

$$\sigma_X^2 = Var[X] = E[(X - \mu_X)^2] = E[X^2 - 2X\mu_X + \mu_X^2]$$

$$= E[X^2] - \mu_X^2 \qquad (2.3\text{-}13)$$

The variance operator, like the expectation operator, can be generalized to operate on any function g(X):

$$Var[g(X)] = E[g^2(X)] - (E[g(X)])^2 \qquad (2.3\text{-}14)$$

The variance operator, however, is <u>not</u> a linear operator, as demonstrated with the function g(X) = a + bX:

$$Var[g(X)] = E[(a + bX)^2] - (E[a + bX])^2$$

$$= b^2E[X^2] - b^2\mu_X^2 = b^2\sigma_X^2 \qquad (2.3\text{-}15)$$

since $E[a + bX] = a + b\mu_X$. From the above example it can also be demonstrated, by letting b = 0, that the variance of a constant, as expected, is zero.

When the standard deviation is normalized by the mean of the random variable ($\mu_X \neq 0$), it is referred to as the <u>coefficient of variation</u> V_x:

$$V_X = \sigma_X/\mu_X \qquad (2.3\text{-}16)$$

Estimators for this population parameter, as well as the variance and standard deviation, will be discussed in a later section.

As an example of an application of the variance operator, consider an application on the standard normal random variable $Z = (X - \mu_X)/\sigma_X$:

$$Var[(X - \mu_X)/\sigma_X] = Var[X]/\sigma_X^2 = 1 \qquad (2.3\text{-}17)$$

which results by analogy with equation (2.3-15). Thus, if X is a normal

random variable with mean μ_X and variance σ_X^2, then Z is a zero-mean random variable with a variance of unity, which is commonly denoted N(0,1).

Problem 2.3-2

a) For $f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$

(*i*) Plot f(x).

(*ii*) Derive and plot F(x).

(*iii*) Calculate $E[X]$ and $Var[X]$.

b) An estimator of the variance σ_X^2 of the random variable X can be developed directly from (2.3-12). First f(x)dx is estimated by f_i^* of (2.2-2). Then x is replaced by $\bar{x}_i$, the center of each class interval corresponding to f_i^*. Finally, μ_X is estimated by $\bar{x}$ from (2.3-2). Then

$$s_X^{*2} = \sum_{i \le x_m/\Delta x} (\bar{x}_i - \bar{x})^2 f_i^*$$

gives an estimate of σ_X^2. Apply this estimator to the log-transmissivity date of Table 2.2-3.

2.4. JOINTLY DISTRIBUTED RANDOM VARIABLES

The investigator frequently encounters the problem that he or she has to deal with two (or more) random variables in the same probability statement. As an example, in the case of random variables X and Y, where X

and Y are possibly correlated, one might desire the probability that X is less than or equal to a and Y is less than or equal to b. If the investigator should know the form of the joint probability density function f(x,y) for these two random variables, then this probability statement is definable:

$$P(X \leq a \text{ and } Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\, dydx = F(a,b) \qquad (2.4\text{-}1)$$

where F(a,b) is the equivalent cumulative distribution function. (The statement $P(X \leq a \text{ and } Y \leq b)$ is also denoted frequently as $P(X \leq a, Y \leq b)$; the more explicit form will be used in this discussion.) As in the univariate case it is required that the mass under the joint probability density function equal unity:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dxdy = 1 \qquad (2.4\text{-}2)$$

The concept of joint probability density functions applies to any number of multiple random variables; the following discussion, however, will be largely restricted to the bivariate case.

As an example of an experiment yielding jointly distributed random variables, consider the results from a simple nonsteady-state pumping test of a confined aquifer: When the Theis equation is used to evaluate data from these tests, information concerning the storativity and transmissivity of the aquifer results. Indeed, we can easily imagine that these quantities are random variables, varying from location to location in response to the local distribution of materials composing the aquifer. More importantly, however, would be the manner in which they vary with

regard to each other: Should the clay content of the aquifer increase at some point, it might be expected that the transmissivity will decrease while the storativity, reflecting the compressibility of the aquifer, would increase. Thus, it is quite possible that these quantities, with regard to the aquifer in question, could be treated as jointly distributed random variables which are, in some manner, interdependent.

Assume for the moment that we have determined the form of the joint density function of storativity and transmissivity. For argument's sake, we may let X represent the transmissivity random variable (or its logarithmic transformation) and Y represent the storativity (or a functional transformation thereof) and then denote the joint density as $f(x,y)$. Now assume that we are interested in the probability that X is less than or equal to a, regardless of the value of Y; that is, we wish to evaluate the probability that our measure of the transmissivity will take on a specific range of values, while the exact value of storativity is unimportant to us. In order that our probability statement regarding X be meaningfull, all values of Y which influence the joint density function must be taken into consideration, for different values of Y would surely influence a statement on X alone. To obtain the total contribution of Y to the joint density function, we allow that Y may take on any value in the interval $(-\infty, \infty)$ and write our probability statement as

$$P(X < a \text{ and } -\infty < Y < \infty) = \int_{-\infty}^{a} \int_{-\infty}^{\infty} f(x,y)\, dydx$$

$$= \int_{-\infty}^{a} f_X(x)\, dx \qquad (2.4\text{-}3)$$

in which the evaluation of the inner integral with respect to y results in a function $f_X(x)$ which meets all requirements to be a probability density function. Thus, we see that, in general, univariate density functions can be recovered from joint density functions by integration, and that this integration has the effect of summing the total contribution of one random variable in the bivariate joint density onto the axis of the other variate, the second variate giving the relative frequency of occurrence of the event in question. These densities are referred to as marginal probability density functions and, with respect to the bivariate joint density f(x,y), they are defined as

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy \qquad (2.4\text{-}4)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\ dx \qquad (2.4\text{-}5)$$

where $f_X(x)$ is the marginal density for the X random variable and similarly $f_Y(y)$ for the Y random variable. The marginal density concept is easily extended to multiple random variables when they are jointly distributed.

2.4.1. Expectation of Jointly Distributed Random Variables

The expectation operator for jointly distributed random variables is defined in the same manner as in the univariate case. Thus, if X and Y are jointly distributed and g(X,Y) is a function of these two random variables, then a general definition of the expectation operator is

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\ dydx \qquad (2.4\text{-}6)$$

If, on the other hand, we desire the expected value of h(X), which is a function of X only, we set g(x,y) equal to h(x) and proceed as in (2.4-6). The result,

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)f(x,y)\,dydx = \int_{-\infty}^{\infty} h(x)f_X(x)\,dx \qquad (2.4\text{-}7)$$

shows that, in such cases, finding the expected value reduces to finding the marginal density and integrating. By letting h(X) equal X, one realizes that the mean μ_X is equal to the integral of the product of x and the marginal density $f_X(x)$, as might be expected.

Consider the case where g(X,Y) equals the product $(X - \mu_X)(Y - \mu_Y)$: The expected value of this product gives one an indication of how X and Y vary together. If the absolute value of the expected value of this product is exceptionally large, then one would expect that X and Y are highly correlated. This expected value of X and Y is referred to as the covariance of X and Y, and is denoted $Cov[X,Y]$ or σ_{XY}:

$$Cov[X,Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_Y\mu_X \qquad (2.4\text{-}8)$$

Note that the covariance of X with itself is $Cov[X,X] = Var[X]$.

Returning to the example of transmissivities and storativities of the previous section, we see that the covariance provides a measure of the degree of interdependence between random variables. That is, since X and Y are both random, we would not expect observations of X and Y to show a perfect relationship; rather, the relationship will be clouded with noise. Since the expected value implies a frequency weighted average of the function in question, and since the frequency distribution will reflect the amount of relationship between X and Y, then summing the product of these

frequency weights with $(X - \mu_X)(Y - \mu_Y)$ over the total variate space will give the average relationship between X and Y. It will be demonstrated in the next section that if X, the measure of transmissivity, and Y, the measure of storativity, were independent then the covariance would theoretically be zero. However, if our intuition is correct, we would not expect this; rather we might expect, should the aquifer have a rather high clay content, that the two variables will be negatively correlated.

If the covariance is normalized with the standard deviations of the two random variables, then it is referred to as the correlation coefficient ρ_{XY}:

$$\rho_{XY} = Cov[X,Y]/(\sigma_X\sigma_Y) \qquad (2.4\text{-}9)$$

The correlation coefficient, as a measure of the linear relationship between X and Y, has the property that its absolute value is less than or equal to unity:

$$|\rho_{XY}| \leq 1 \qquad (2.4\text{-}10)$$

That is, when X and Y are precisely linearly related, then $|\rho_{XY}|$ will equal unity. If there is no relationship between X and Y, as shown in the next section, $Cov[X,Y]$ and therefore ρ_{XY} will be zero. Property (2.4-10) is demonstrated in Appendix 2.11.1, but this Appendix requires some knowledge of the next section.

2.4.2. Independent Random Variables

Two random variables X and Y are said to be independent if, for all a and b,

$$P(X \leq a \text{ and } Y \leq b) = P(X \leq a)P(Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_X(x)f_Y(y)\, dydx \qquad (2.4\text{-}11)$$

where $f_X(x)$ and $f_Y(y)$ are the densities of X and Y, respectively. Equation (2.4-11) implies that the joint density function of two independent random variables is the product of their individual densities. That is,

$$f(x,y) = f_X(x)f_Y(y) \qquad (2.4\text{-}12)$$

Of course, in general it would be expected that an event corresponding to $X \leq a$ and $Y \leq b$ would occur with equal or less frequency than an event corresponding to either $X \leq a$ or $Y \leq b$ separately. Only in the case of a complete lack of dependence between these events can we say that $P(X \leq a \text{ and } Y \leq b) = P(X \leq a)P(Y \leq b)$. This is a somewhat intuitive result that has already been used in connection with the two-dice experiment; if X is an outcome of the first die and Y the second, then $P(X = 1 \text{ and } Y = 2) = P(X = 1)P(Y = 2) = 1/36$.

A random sample is ideally a collection of independent random variables. That is, prior to their observation, each element of a random sample is a random variable; its value is not known until after the observation process is completed. It is desirable that these outcomes not have any interdependence which might affect the sample density. This generally requires careful design of the experiment from which the observations result so that all X_i, $i = 1, \ldots, n$, are independent.

The question of independence of two random variables X and Y has important implications on their covariance, for if X and Y are independent, then

$$Cov[X,Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - \mu_X)(y - \mu_X)f_X(x)f_Y(y)\,dxdy$$

$$= \int_{-\infty}^{\infty}(x - \mu_X)f_X(x)\,dx \int_{-\infty}^{\infty}(y - \mu_Y)f_Y(y)\,dy$$

$$= E[X - \mu_X]E[Y - \mu_X] = 0 \qquad (2.4\text{-}13)$$

However, if the covariance of two random variables is zero, it does not necessarily follow that they are independent. One may only suspect that independence is the cause of a zero covariance.

2.4.3. Conditional Probabilities

The marginal probability density function, as developed in equation (2.4-3), can be considered to be a special case of a more general concept referred to as conditioning. Generally speaking, a multivariate probability statement is subject to conditioning when a subset of the random variables pertaining to an experiment falls under some restriction, thus causing the remaining variables to be conditioned by this restriction. In the case of the marginal density function, we examined the probability that X is less than a, given that Y can take on any value in the interval $(-\infty, \infty)$. Thus the restriction that Y take on a specific set of values conditions the probability that X is less than a. Formally, we state this as

$$P(X < a \mid -\infty < Y < \infty)$$

In general, the restriction can be applied to any interval (b, c), where $b < c$, and need not be limited to the interval $(-\infty, \infty)$. However, as in the case of the marginal density, the variable or variables subject to restriction are effectively removed from the probability statement; the variable or variables being conditioned are the ones over which the frequency of occurrence of a specific event may be questioned.

When an experiment which results in a bivariate random variable is conditioned over a range other than $(-\infty, \infty)$, a reduction of the potential sample space available to the experiment results. In the previous example of jointly varying transmissivities X and storativities Y, if we were interested in the conditioned results that X is less than a, given that we are only interested in a specific range of values (b, c) for storativities, then the specific value which Y takes on does not interest us, as long as it falls between b and c. One could proceed as in (2.4-3) to evaluate this probability, except for an obvious pitfall: The resulting probability statement over X, where X can take on any value less than a, would not necessarily have property (2.2-11) of cumulative distribution functions. That is, as X is the remaining active random variable in the probability statement, its probability of occurrence over the interval $(-\infty, \infty)$ should be unity:

$$\lim_{a \to \infty} P(X < a | b < y < c) = 1$$

However, by restricting Y to a specific interval (b, c), it is possible and probable that, as a goes to infinity, an integral of the form (2.4-3) will have a lesser value than unity when the inner integral over y is restricted to a range of something less than $(-\infty, \infty)$. Thus, an integration of the form (2.4-3) alone will not produce a form suitable

to serve as a cumulative distribution function for the conditioned variable X.

In order that a probability statement resulting from conditioning have the limiting value of unity, these statements must be appropriately normalized. If, as in the bivariate case, we desire $P(X < a \mid b < Y < c)$, then we must normalize by $P(-\infty < X < \infty \text{ and } b < Y < c)$; that is

$$P(X < a \mid b < Y < c) = \frac{P(X < a \text{ and } b < Y < c)}{P(-\infty < X < \infty \text{ and } b < Y < c)}$$

$$= \frac{\int_{-\infty}^{a} \int_{b}^{c} f(x,y)\, dydx}{\int_{-\infty}^{\infty} \int_{b}^{c} f(x,y)\, dydx}$$

$$= \frac{\int_{-\infty}^{a} \int_{b}^{c} f(x,y)\, dydx}{\int_{b}^{c} f_Y(y)\, dy} \qquad (2.4\text{-}14)$$

Thus, the conditional probability density function, $f(x|b < Y < c)$, for the conditioned random variable X may be defined as

$$f(x|b < Y < c) = \int_{b}^{c} f(x,y)\, dy \Big/ \int_{b}^{c} f_Y(y)\, dy \qquad (2.4\text{-}15)$$

which of course gives the limiting value of unity when integrated with respect to x over the interval $(-\infty, \infty)$. Note that when $b = -\infty$ and $c = \infty$, then $f(x|b < Y < c) = f_X(x)$, as indicated by the previous discussion of marginal densities.

Remarkably, the conditional density exists even when the restriction is that, in the example of the bivariate case, Y take on a specific value. In order to see this easily, consider $P(X < a \mid Y = c)$; then (2.4-15) may be written as

$$f(x|Y = c) = \lim_{\delta \to 0} \frac{\int_c^{c+\delta} f(x,y)\, dy}{\int_c^{c+\delta} f_Y(y)\, dy}$$

$$= \lim_{\delta \to 0} \frac{[F(x,c + \delta) - F(x,c)]/\delta}{[F_Y(c + \delta) - F_Y(c)]/\delta}$$

$$= \frac{dF(x,y)/dy}{dF_Y(y)/dy}\Bigg|_{y=c}$$

$$= \frac{f(x,c)}{f_Y(c)} \qquad (2.4\text{-}16)$$

Thus, we may recover the density function for X for any particular slice, Y = c, through the joint density function f(x,y). If f(x,y) were defined for the example of transmissivity and storativity random variables, equation (2.4-16) would enable us to predict the probability of events concerning transmissivity X for any given value of storativity Y.

The student should also note that some remarkable simplifications result if X and Y are independent random variables. That is, if X and Y are independent, then from (2.4-12) and (2.4-14) we see that

$$P(X < a \mid b < Y < c) = P(X < a) \qquad (2.4\text{-}17)$$

Indeed, this is yet another way in which we can define independence of random variables.

The following problem is intended to familiarize the student with the concept of conditioning; it is not intended to be rigorous. The key to understanding conditioning, especially for discrete random variables, is to understand how it restricts the sample space and realize that the probability of occurrence of an event which contains the entire remaining sample space must be unity.

Problem 2.4-1

a) Given two dice that are thrown sequentially, what is the probability that the first is a three and the second is a two? That is,

$$P(X = 3, Y = 2)?$$

b) What is the probability that the sum of the dice is five? That is,

$$P(X + Y = 5)?$$

c) Given that the first die is three, what is the probability that the second is two? That is,

$$P(Y = 2|X = 3)?$$

d) Given that the first die is three, what is the probability that the sum of the two dice is five? That is,

$$P(X + Y = 5|X = 3)?$$

e) Given that the first die is three, what is the probability that the sum of the two dice is less than or equal to five? That is,

$$P(X + Y \leq 5|X = 3)?$$

Parts c, d and e are conditional probability statements; that is, the probability statement is conditioned by prior information.

2.4.4. Variance of a Column Vector

Our purpose in this section is to develop a representation for the variance of a column vector. As a vehicle to this end, consider the linear equation

$$Y = a_1X_1 + a_2X_2 + a_3X_3 \tag{2.4-18}$$

where Y, X_1, X_2, and X_3 are random variables and a_1, a_2 and a_3 are constants. The variance of Y is

$$Var[Y] = E\left\{[(a_1X_1 + a_2X_2 + a_3X_3) - E(a_1X_1 + a_2X_2 + a_3X_3)]^2\right\}$$

$$= a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + a_3^2\sigma_{X_3}^2$$

$$+ 2a_1a_2\sigma_{X_1X_2} + 2a_1a_3\sigma_{X_1X_3} + 2a_2a_3\sigma_{X_2X_3} \tag{2.4-19}$$

where correlations between X_1, X_2, and X_3 have been allowed for. A vector representation for equation (2.4-19) is

$$Var[Y] = E[Y^2] - (E[Y])^2 = E[\underline{a}\ \underline{X}\ \underline{X}^T\ \underline{a}^T] - E[\underline{a}\ \underline{X}]E[\underline{a}\ \underline{X}] \tag{2.4-20}$$

where

$$\underline{a} = [a_1\ a_2\ a_3] \text{ and } \underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

Since expectation is a linear operator, the right hand side of (2.4-20) can be expressed as

$$E[\underline{a}\ \underline{X}\ \underline{X}^T\ \underline{a}^T] - E[\underline{a}\ \underline{X}]E[\underline{X}^T\ \underline{a}^T]$$

$$= \underline{a}\ E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T]\ \underline{a}^T \tag{2.4-21}$$

where $\underline{a}\ \underline{X} = \underline{X}^T\ \underline{a}^T$. The expected value of a matrix is the matrix of expected values of each element. Thus

$$E[(\underline{X} - E(\underline{X}))(\underline{X} - E(\underline{X}))^T] = \begin{bmatrix} \sigma^2_{X_1} & \sigma_{X_1X_2} & \sigma_{X_1X_3} \\ \sigma_{X_1X_2} & \sigma^2_{X_2} & \sigma_{X_2X_3} \\ \sigma_{X_1X_3} & \sigma_{X_2X_3} & \sigma^2_{X_3} \end{bmatrix} \qquad (2.4\text{-}22)$$

This matrix is defined to be the variance of a 3×1 column vector $\underline{X}$, and allows one to express (2.4-19) in matrix notation as

$$Var[Y] = \underline{a}\ Var[\underline{X}]\ \underline{a}^T \qquad (2.4\text{-}23)$$

If the variances $\sigma^2_{X_1}$, $\sigma^2_{X_2}$ and $\sigma^2_{X_3}$ are all equal, the matrix (2.4-22) becomes

$$Var[\underline{X}] = \begin{bmatrix} 1 & \rho_{X_1X_2} & \rho_{X_1X_3} \\ \rho_{X_1X_2} & 1 & \rho_{X_2X_3} \\ \rho_{X_1X_3} & \rho_{X_2X_3} & 1 \end{bmatrix} \sigma^2 \qquad (2.4\text{-}24)$$

where $\rho_{X_iX_j}$ is the correlation coefficient for X_i and X_j and σ^2 is the common variance. A further reduction in (2.4-22) occurs if X_1, X_2, and X_3 are uncorrelated, causing the correlation coefficients in (2.4-24) to be zero. In this case,

$$Var[\underline{X}] = \underline{I}\sigma^2 \qquad (2.4\text{-}25)$$

where $\underline{I}$ is a 3 × 3 identity matrix. These forms have practical importance in regression.

Problem 2.4-2

a) Carry out the expectation indicated and show that (2.4-19) holds.

b) Demonstrate that relationship (2.4-21) holds and that

$$Var[\underline{a}\ \underline{X}] = \underline{a}\ Var[\underline{X}]\ \underline{a}^T$$

c) Let $Y_i = \underline{a}_i\ \underline{X}$, where $\underline{a}_i = [a_{i1}\ a_{i2}\ a_{i3}]$, and $\underline{X}$, defined as in (2.4-20), is a column vector of random variables. Further, let $\underline{Y} = \underline{A}\ \underline{X}$ where

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_p \end{bmatrix} \qquad \text{and} \qquad \underline{A} = \begin{bmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \cdot \\ \cdot \\ \cdot \\ \underline{a}_p \end{bmatrix}$$

That is, $\underline{A}$ is a p × 3 matrix composed of the row vectors $\underline{a}_i$, $i = 1, \cdots, p$. Show that $Var[\underline{Y}] = Var[\underline{A}\ \underline{X}] = \underline{A}\ Var[\underline{X}]\ \underline{A}^T$. (Hint: (2.4-22) still defines the variance of a column vector; $\sigma_{Y_iY_j} = E[Y_iY_j - E(Y_i)E(Y_j)] = E[\underline{a}_i\ \underline{X}\ \underline{X}^T\ \underline{a}_j^T - E(\underline{a}_i\ \underline{X})E(\underline{X}^T\ \underline{a}_j^T)]$.)

2.5. ESTIMATORS OF POPULATION PARAMETERS

A statistic is defined as any computation from a random sample resulting in a specific value. As such, a statistic is considered to be a random variable, since it is highly probable that the computed value would change from random sampling to random sampling. Note that this definition

precludes that a statistic contain any unknown parameters. Estimators of population parameters are considered to be statistics and therefore random variables. Consider the estimator (2.3-1) for the mean:

$$\bar{x} = \sum_{i \le x_m/\Delta x} f_i^* \bar{x}_i \qquad (2.5\text{-}1)$$

The estimated frequency f_i^* is computed from values of observations x_i originating from a random sampling of the sample space. However, prior to sampling, a random sample is merely an abstract collection of random variables X_i, $i = 1, \cdots, n$. Any function of random variables, as (2.5-1) would be prior to sampling, is also a random variable, perhaps having a completely different distribution than those individuals composing the collection.

Our discussion of statistics will largely be from the a priori viewpoint; that is, in the case of (2.5-1), $\bar{x}$ is the value of the random variable $\bar{X}$, which is an estimator for the population mean as developed from some arbitrary random sample.

2.5.1. Mean Estimator

As an estimator for the population mean, equation (2.5-1) in addition to being cumbersome to compute has the debility that it is dependent upon an arbitrary selection of a class interval. That is, since f_i^* is dependent upon Δx, the value of $\bar{X}$ will depend upon the choice of Δx used in the computation. As a means of pursuing this problem, assume that we have at our disposal a random sample consisting of n observations, and at some point their distribution appears as in figure 2.5-1.

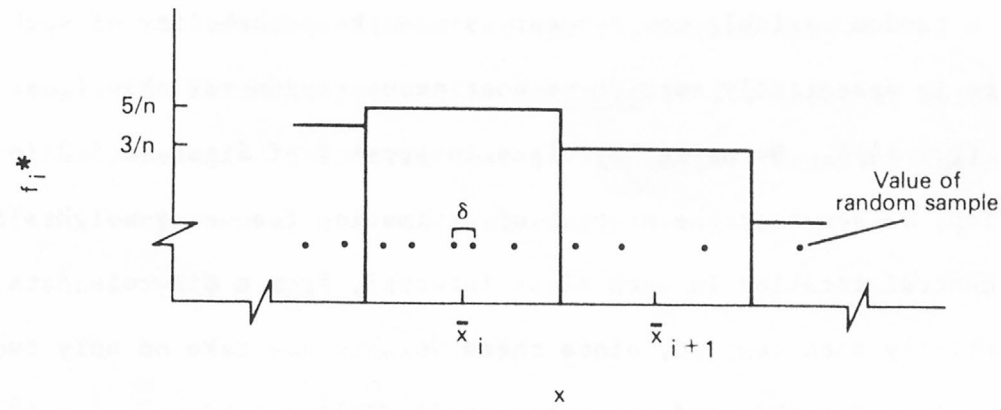


fig. 2.5-1

Since Δx is arbitrary, it can be reduced to δ, the minimum of all differences in neighboring values of the random variable. Then f_i^* would take on only two values, 1/n or 0, and would have the ragged saw-toothed shape shown in figure 2.5-2.

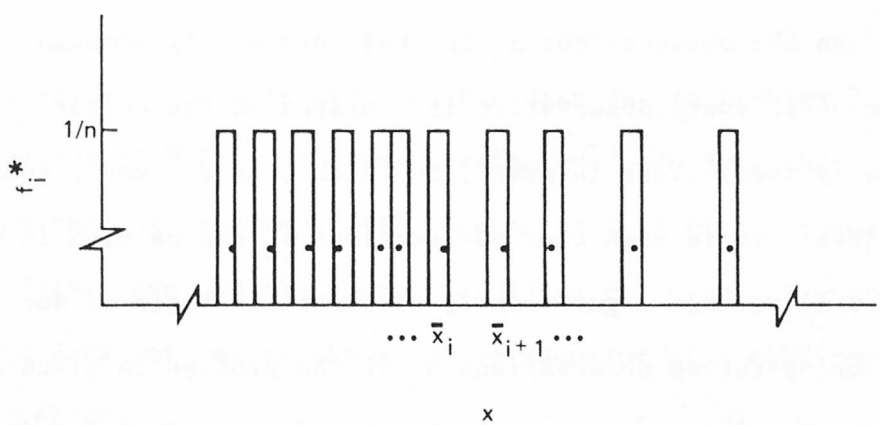


fig. 2.5-2

The teeth in figure 2.5-2 will be concentrated in regions where f_i^* in figure 2.5-1 is larger. Of course, it is highly unlikely that a repeat value of a random variable could occur, since the probability of such an occurrence is essentially zero for a continuous random variable (see equation (2.2-14)). By using the class interval δ of figure 2.5-2 in our computation, we see that the problem of estimating frequency weights for $\bar{x}_i$, the central location in each class interval, from a discrete data set has essentially been removed, since these weights now take on only two specific values for this and any other smaller class interval.

It is to be expected that an estimator of the population mean, if anything, will be better if the smaller class interval δ is used, since a value of $\bar{X}$ would contain less measurement error associated with the arbitrary selection of the class interval. It is still cumbersome, however, to calculate the central value $\bar{x}_i$ of these possibly very small class intervals, especially when one considers that many are not going to contribute to the estimator. We ask ourself if it is not possible to use the observations x_i in their place. By reducing Δx even further until every observation is isolated in the center of its own infinitesimally small class interval, in which case f_i^* would remain at the 1/n level, it is seen that observations x_i can be used in place of $\bar{x}_i$ in (2.5-1) without significantly altering the basis of our estimator. Using future observations X_i of the process in place of central-interval values, $\bar{x}_i$, an estimator based on an infinitesimally small interval would appear as

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (2.5\text{-}2)$$

where n is the size of the random sample and X_i, $i = 1, \cdots, n$, is the collection of random variables from the random sample. Equation (2.5-2) is the preferred estimator for the population mean.

A sample statistic is said to be unbiased if its expected value is equal to the population parameter that it estimates. Consider the expected value of the sample mean, derived from random variables X_i, $i = 1, \cdots, n$. Since $E[X_i] = \mu_X$,

$$E[\bar{X}] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mu_X = \mu_X \qquad (2.5\text{-}3)$$

Hence $\bar{X}$ is an unbiased estimator of μ_X. While examining an estimator for unbiased qualities is important, it does not necessarily insure that the estimator is the most efficient (or best) in the sense that the variance of the estimator is the smallest. It is, however, an important quality, and the variance estimator is examined for this quality in the next section.

Problem 2.5-1

a) Recompute the sample mean for the data set in problem 2.2-2 using equation (2.5-2) as the mean estimator. How does this result vary from that of problem 2.3-1? How do you, in light of equation (2.2-14), explain the repeat values in the data set (note that this data set represents a random sampling of a continuous random variable)?

b) With regard to a large regional aquifer, well data such as that in table 2.2-2 represent point estimates of transmissivities. It is generally considered that the best estimate of the effective transmissivity (the one to use in modeling the flow field) is

the geometric mean of these point estimates. The statistic for the geometric mean is defined as

$$\bar{T}_g = (T_1 \cdot T_2 \cdot T_3 \cdot \cdots \cdot T_n)^{\frac{1}{n}}$$

where T_i, $i = 1, \cdots, n$, is a random sample from the sample space of the T random variable. Letting $X_i = \log_{10} T_i$, we see that

$$\log_{10}[\bar{T}_g] = \frac{1}{n} \log_{10} (T_1 \cdot T_2 \cdot T_3 \cdot \cdots \cdot T_n)$$

$$= \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$$

Therefore

$$\bar{T}_g = 10^{\bar{X}}$$

What is the geometric mean of the transmissivity data in Table 2.2-2?

As a measure of the dispersion about the geometric mean, one could use the estimator

$$D_g = 10^{(\bar{X} \pm S_X)}$$

What is the dispersion D_g about the geometric mean? Use results of problem 2.3-2, part b) as values for $\bar{X}$ and S_X^2. Considering the dispersion, how do you feel about $\bar{t}_g$ being the effective transmissivity of the carbonate rocks of central Pennsylvania?

2.5.2. Variance Estimator

As an estimator of the variance σ_X^2, consider using an estimator S_X^{*2} whose value s_X^{*2} is calculated from the equation

$$s_X^{*2} = \sum_{i \le x_m/\Delta x} f_i^* (\bar{x}_i - \mu_X)^2 \qquad (2.5\text{-}4)$$

which is analogous to equation (2.3-12) for the population parameter. If the class interval is taken to be small enough so as to isolate every future observation X_i in a class interval, then (2.5-4) can be rewritten in terms of random observations X_i, $i = 1, \cdots, n$, as

$$S_X^{*2} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_X)^2 \qquad (2.5\text{-}5)$$

since $f_i^* = 1/n$ and $X_i = x_i$ prior to sampling. When the underlying population X_i, $i = 1, \cdots, n$, is normally distributed, it can be shown that equation (2.5-5) is the most efficient, unbiased estimator of the variance σ_X^2 in the sense that its variance is the least of all possible unbiased estimators for σ_X^2.

On occasion, the population mean μ_X can be determined from other considerations, as was done in the titration experiment in section 2.2.5. However, usually μ_X is also unknown, requiring that μ_X be replaced by $\bar{X}$ in (2.5-5):

$$S_X^{*2} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \qquad (2.5\text{-}6)$$

To test whether (2.5-6) is an unbiased estimator of σ_X^2, the expected value of S_X^{*2} is determined. The actual mechanics of this operation are presented in Appendix 2.11.2; only the result is presented here:

$$E(S_X^{*2}) = \frac{n-1}{n} \sigma_X^2 \qquad (2.5\text{-}7)$$

Thus S_X^{*2} is a biased estimator of σ_X^2. To produce an unbiased estimator of

σ_X^2, S_X^{*2} is multiplied by the ratio n/(n -1):

$$S_X^2 = \frac{n}{n-1} S_X^{*2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \qquad (2.5\text{-}8)$$

This estimator is unbiased, but less efficient than (2.5-6). However, it is the preferred estimator for small sample sizes.

Heuristically, one can argue that this adjustment to the estimator is necessary because the population mean μ_X is being estimated by the sample statistic $\bar{X}$. The sample mean will be located at the centroid of the random sample, regardless of whether its value is near that of the population mean. Thus, an equation that estimates the variance about this centroid will produce a smaller value than if the estimate were made about the population mean. The adjustment, then, merely compensates for the smaller deviates produced by using $\bar{X}$ in place of μ_X.

Equation (2.5-8) can be rewritten, with the aid of some algebraic manipulation, to produce a slightly more useful form for hand calculations:

$$\begin{aligned} S_X^2 &= \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \\ &= \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - 2\bar{X} \sum_{i=1}^{n} X_i + n\bar{X}^2 \right] \\ &= \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right] \end{aligned} \qquad (2.5\text{-}9)$$

The estimator for the standard deviation is taken to be the square root of the variance estimator. Values x_i, $i = 1, \cdots, n$, obtained by sampling the population of X randomly, are used in place of X_i in equation (2.5-9) in order to obtain a value s_X^2 for the sample statistic S_X^2.

2.5.3. Correlation Coefficient Estimator

In a manner analogous to the variance, an estimator for the covariance, and therefore the correlation coefficient, can be derived. Let R_{XY} represent the estimator for the correlation coefficient ρ_{XY}; then for paired data,

$$R_{XY} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum_{i=1}^{n} (X_i - \bar{X})^2 \sum_{i=1}^{n} (Y_i - \bar{Y})^2\right]^{\frac{1}{2}}} \tag{2.5-10}$$

or provided that S_X and S_Y originate from the paired data,

$$R_{XY} = \frac{1}{(n-1)S_X S_Y} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) \tag{2.5-11}$$

which can be written, for purposes of hand calculation, as

$$R_{XY} = \frac{1}{(n-1)S_X S_Y} \left(\sum_{i=1}^{n} X_i Y_i - n\bar{Y}\bar{X} \right) \tag{2.5-12}$$

where S_X and S_Y are calculated by taking the square root of either (2.5-8) or (2.5-9). The actual value r_{XY} of R_{XY} is obtained by using values x_i, $i = 1, \cdots, n$, from a random sample in place of X_i.

2.5.4. Summary

In summary, population parameters and equivalent sample statistics can be tabulated as follows:

Population Parameter	Sample Statistic
μ	$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
σ_X^2	$S_X^2 = \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right]$
V_X	$C_X = S_X/\bar{X}$
ρ_{XY}	$R_{XY} = \frac{1}{(n-1)S_X S_Y} \left(\sum_{i=1}^{n} X_i Y_i - n\bar{Y}\bar{X} \right)$

These estimators can also be stated in matrix form. For instance, let $d_i = x_i - \bar{x}$, then a value for S_X^2 is

$$s_X^2 = \frac{1}{n-1} \underline{d}^T \underline{d} \qquad (2.5\text{-}13)$$

where $\underline{d}$ is a column vector of deviates and $\underline{d}^T$ is its transpose.

If $e_i = y_i - \bar{y}$, then a value for R_{XY} is

$$r_{XY} = \frac{\underline{e}^T \underline{d}}{(n-1)s_X s_Y} \qquad (2.5\text{-}14)$$

Forms (2.5-13) and (2.5-14) are commonly encountered in linear regression.

Problem 2.5-2

Using the following data set, calculate the sample mean, variance and standard deviation of both dissolved solids and specific conductance; then calculate their correlation coefficient.

Specific conductance and dissolved solids data for wells in carbonate rocks of Maryland.[1/]

Specific conductance µmho/cm	Dissolved solids ppm
278	257
1120	610
533	338
723	458
462	264
1030	562
357	231
304	175
469	268
641	388
969	638
876	532
721	405
895	610
501	304
323	171
310	201
1230	736
504	290
319	208
704	464
1130	688
600	342

1/Adopted from Nutter, 1973, p. 63-68.

2.6. TRANSFORMATION OF RANDOM VARIABLES

As we have noted previously, statistics are combinations of random variables and, as such, must be random variables themselves. If the population from which the random sample is selected can be identified, then it is frequently possible to identify the probability density functions of statistics, which are estimators of the population parameters. If a density function is identified, then it should be possible to develop criteria for testing the accuracy of these estimators. With these objectives in mind, we proceed to identify density functions that result from the several types of transformations that produce statistics.

Before proceeding with this identification process, we make note of two general results from expectation which are applicable to all random variables, regardless of their distribution. In general, if X_1, X_2, ..., X_n are independent variables with identical mean μ_X and identical standard deviation σ_X, then

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{2.6-1}$$

is also a random variable with mean

$$\mu_{\bar{X}} = E[\bar{X}] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu_X \tag{2.6-2}$$

and variance

$$\sigma_{\bar{X}}^2 = Var[\bar{X}] = \frac{1}{n^2} \sum_{i=1}^{n} Var[X_i] = \sigma_X^2/n \tag{2.6-3}$$

Equation (2.6-2) was used previously to show that $\bar{X}$ is an unbiased estimator of μ_X, and (2.6-3) is demonstrated more fully in Appendix 2.11.2.

Note that (2.6-3) only succeeds because $Cov[X_i, X_j] = 0$, $i \neq j$; that is, the X_i's are independent.

The square root of (2.6-3), $\sigma_{\bar{X}}$, is also known as the standard error of $\bar{X}$. The standard deviation of any statistical measure is referred to as the standard error of that statistic.

2.6.1. Sum of Independent Normal Random Variables

Let X_1 and X_2 be independent normal random variables, X_1 with mean zero and variance one (N(0,1)) and X_2 with mean zero and variance k (N(0,k)). How, then, is their sum distributed? To answer this question, consider

$$P(Y \leq y) = P(X_1 + X_2 \leq y)$$

By noting that $P(X_1 + X_2 \leq y) = P(X_1 \leq y - X_2 \text{ and } -\infty \leq X_2 \leq \infty)$, comparison with (2.4-11) shows that

$$F_Y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y-x_2} \frac{e^{-x_1^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} F_{X_1}(y - x_2) \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2 \qquad (2.6\text{-}4)$$

To find the probability density function of Y, we differentiate $F_Y(y)$ with respect to y; that is

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \int_{-\infty}^{\infty} \frac{d}{dy} F_{X_1}(y - x_2) \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2$$

$$= \int_{-\infty}^{\infty} \frac{e^{-(y-x_2)^2/2}}{\sqrt{2\pi}} \frac{e^{-x_2^2/(2k)}}{\sqrt{2\pi k}} dx_2$$

$$= \frac{1}{2\pi\sqrt{k}} \int_{-\infty}^{\infty} \exp[-(y^2 - 2yx_2 + (k+1)x_2^2/k)/2]\, dx_2 \qquad (2.6\text{-}5)$$

which, after some algebraic manipulation, yields

$$f_Y(y) = \frac{\exp[-y^2/(2k+2)]}{2\pi\sqrt{k}} \int_{-\infty}^{\infty} \exp\left[-\left(x_2\sqrt{\frac{k+1}{2k}} - y\sqrt{\frac{k}{2k+2}}\right)^2\right] dx_2 \qquad (2.6\text{-}6)$$

By letting $u = \sqrt{2}\left(x_2\sqrt{\frac{k+1}{2k}} - y\sqrt{\frac{k}{2k+2}}\right)$, then

$$f_Y(y) = \frac{\exp[-y^2/(2k+2)]}{\sqrt{2\pi(k+1)}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2}\, du$$

$$= \frac{\exp[-y^2/(2k+2)]}{\sqrt{2\pi(k+1)}} \qquad (2.6\text{-}7)$$

which follows from (2.2-11). Thus, the sum of two independent zero-mean normal random variables, one with variance unity and the other with variance k, is a normal random variable with variance k + 1 (N(0,k + 1)).

If, in the previous problem, k were to equal one, then we see that the sum of two N(0,1) independent random variables is a N(0,2) random variable. By adding yet another independent N(0,1) random variable to the previous two N(0,1) random variables, induction tells us that a N(0,3) random variable results. Thus, in general, the sum of n independent N(0,1) random variables results in a N(0,n) random variable.

We are now in a position to determine the distribution of the statistic $\bar{X}$, as shown in (2.6-1), if $\bar{X}$ is determined from a random sample in which all the observations X_i, $i = 1, \ldots, n$, are

independent normal random variables with common mean μ_X and variance σ_X^2; that is, $N(\mu_X,\sigma_X^2)$. We note from (2.6-3) that $\bar{X}$ has the standard deviation $\sigma_X/\sqrt{n}$. If we standarize $\bar{X}$ by its mean and standard deviation, and multiply this result by $\sqrt{n}$, then

$$\sqrt{n}\left(\frac{\bar{X}-\mu_X}{\sigma_X/\sqrt{n}}\right)=\sqrt{n}\left(\frac{\sum_{i=1}^{n}X_i-n\mu_X}{\sqrt{n}\sigma_X}\right)=\sum_{i=1}^{n}\left(\frac{X_i-\mu_X}{\sigma_X}\right) \qquad (2.6\text{-}8)$$

results. We see that this new statistic is the sum of n normal random variables with mean zero and variance one. From the previous paragraph, (2.6-8) must be a normal random variable with mean zero and variance n. To obtain a random variable with mean zero and variance unity, one would divide (2.6-8) by the square root of n. By inspection, then, the quantity $(\bar{X}-\mu_X)/(\sigma_X/\sqrt{n})$ must be a standard normal random variable, and $\bar{X}$ must be normal with mean μ_X and variance σ_X^2/n. Thus, if it were known that a random sample were composed of normal random variables with a particular mean and variance, then one could investigate the probability that a future determination of the sample mean could take on a particular range of values.

2.6.2. The Chi-Square Distribution

We are frequently concerned with the square of a random variable, and may wish to know its density function. Assuming that the random variable X is normally distributed with mean zero and variance one (N(0,1)), we may inquire as to the nature of the distribution of its square, $Y = X^2$.

Proceeding as in the previous section, we find the cumulative distribution of Y:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx \qquad (2.6\text{-}9)$$

By taking the derivative of $F_Y(y)$, one finds the density function of Y:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{\sqrt{2\pi}} \frac{e^{-y/2}}{\sqrt{y}}, \quad y \geq 0 \qquad (2.6\text{-}10)$$

which is the chi-square density function with one degree of freedom.

The chi square has a useful additive property similar to that exhibited by independent normals. Namely, if Y_1 and Y_2 are independent chi-square random variables with degrees of freedom ν_1 and ν_2, then $Y_1 + Y_2$ is a chi-square random variable with degrees of freedom $\nu_1 + \nu_2$. Consequently, if $Y_1, Y_2, \ldots, Y_n$ are independent chi-square random variables each with 1 degree of freedom, then $Y_1 + Y_2$ is a chi square with 2 degrees of freedom, $(Y_1 + Y_2) + Y_3$ is a chi square with 3 degrees of freedom and in general $\sum_{i=1}^{n} Y_i$ is a chi square with n degrees of freedom. Values for the cumulative distribution function of the chi-square distribution with ν degrees of freedom are to be found in table 2.10-2.

If X_i, $i = 1, \ldots, n$, are independent normal random variables, each with mean μ_X and variance σ_X^2, then $\sum_{i=1}^{n} ((X_i - \mu_X)/\sigma_X)^2$

must be a chi-square random variable with n degrees of freedom. This follows from the previous argument by letting $Y_i = (X_i - \mu_X)^2/\sigma_X^2$ and noting that Y_i is the square of a N(0,1) random variable. Furthermore, because

$$\sum_{i=1}^{n} \frac{(X_i - \mu_X)^2}{\sigma_X^2} = n \frac{(\bar{X} - \mu_X)^2}{\sigma_X^2} + \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma_X^2} \qquad (2.6\text{-}11)$$

the statistic S_X^2 can be written in terms of this sum as

$$\frac{(n-1)S_X^2}{\sigma_X^2} = \sum_{i=1}^{n} \frac{(X_i - \mu_X)^2}{\sigma_X^2} - \frac{(\bar{X} - \mu_X)^2}{(\sigma_X/\sqrt{n})^2} \qquad (2.6\text{-}12)$$

Under the condition that the underlying population is independent normal, it was demonstrated in section 2.6.1 that $\bar{X}$ is normal with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$. Thus, $(\bar{X} - \mu_X)^2/(\sigma_X^2/n)$, under this condition, is chi square with one degree of freedom. One might reasonably expect, then, that

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(\nu) \qquad (2.6\text{-}13)$$

is a chi-square random variable with $\nu = n - 1$ degrees of freedom, which is indeed the case when the underlying population of X_i's are independent normal random variables.

2.6.3. The F Distribution

The density function for the ratio of two independent chi-square random variables can be calculated rather easily by the method used in the

previous sections. However, since we have little need of the actual form of this density function, known as the F distribution, we relieve the student of working through the actual calculation if he will accept the following statement: If X_1 is a chi-square random variable with ν_1 degrees of freedom, and X_2 is a chi-square random variable with ν_2 degrees of freedom, and X_1 and X_2 are independent, then

$$\frac{X_1/\nu_1}{X_2/\nu_2} \sim F(\nu_1, \nu_2) \tag{2.6-14}$$

defines the F distribution with ν_1 and ν_2 degrees of freedom.

Table 2.10-3 is a tabulation of ratio (2.6-14) that satisfies the probability statement

$$P(F(\nu_1, \nu_2) \leq F_\alpha(\nu_1, \nu_2)) = 1 - \alpha \tag{2.6-15}$$

where α equals 0.05; the meaning of (2.6-15) is illustrated in figure 2.6-1.

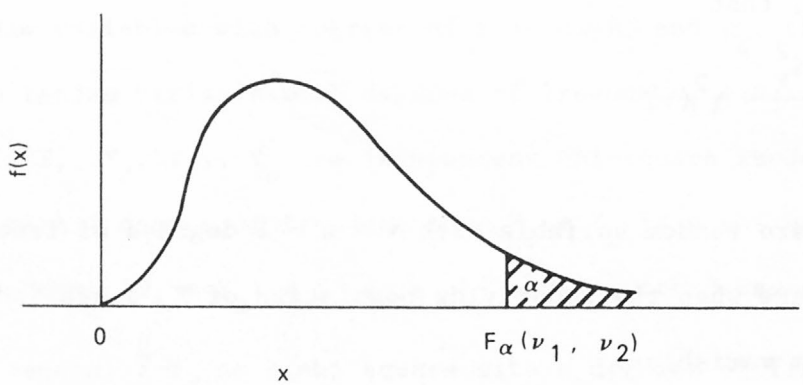


fig. 2.6-1

Note that the reciprocal of an entry $F_\alpha(\nu_1, \nu_2)$ in table 2.10-3 is equal to $F_{1-\alpha}(\nu_2, \nu_1)$. That is, if (2.6-15) holds, then

$$P(F(\nu_2, \nu_1) > F_\beta(\nu_2, \nu_1))$$

$$= P(1/F(\nu_2, \nu_1) \leq 1/F_\beta(\nu_2, \nu_1))$$

$$= P(F(\nu_1, \nu_2) \leq 1/F_\beta(\nu_2, \nu_1))$$

$$= \beta \qquad (2.6\text{-}16)$$

since by equation (2.6-14), $1/F(\nu_2, \nu_1)$ is an $F(\nu_1, \nu_2)$ random variable. By comparing (2.6-15) with the third line in (2.6-16), we see that, when β equals $1 - \alpha$,

$$F_\alpha(\nu_1, \nu_2) = \frac{1}{F_{1-\alpha}(\nu_2, \nu_1)} \qquad (2.6\text{-}17)$$

Thus, if we wish to evaluate $F_{1-\alpha}(n_1, n_2)$ for the statement

$$P(F(n_1, n_2) \leq F_{1-\alpha}(n_1, n_2)) = \alpha \qquad (2.6\text{-}18)$$

where α is the relative mass indicated in figure 2.6-2, then we need only find $F_\alpha(\nu_1, \nu_2)$, where $\nu_1 = n_2$ and $\nu_2 = n_1$, in a table of values for the F distribution and calculate its reciprocal.

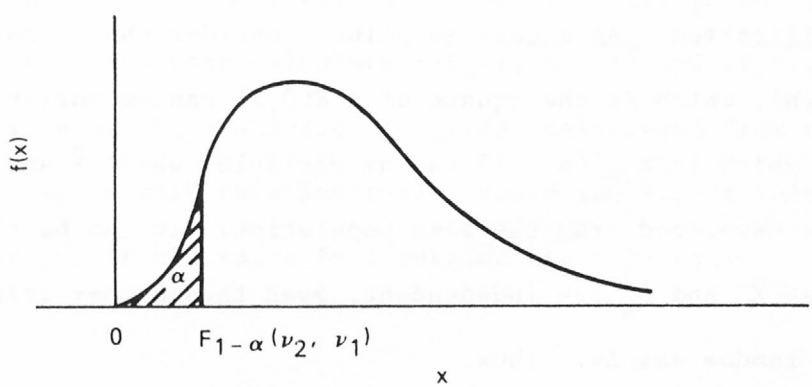


fig. 2.6-2

As an example of a practical statistic associated with the $F(\nu_1,\nu_2)$ random variable, consider two random samples of size n_1 and n_2, which have been selected from two normal populations with variances σ_1^2 and σ_2^2, respectively. Let

$$X_1 = (n_1 - 1)S_1^2/\sigma_1^2 \tag{2.6-19}$$

and

$$X_2 = (n_2 - 1)S_2^2/\sigma_2^2 \tag{2.6-20}$$

where S_1^2 and S_2^2 are sample variances that are independent, since they originate from separate random samples. From (2.6-13) and (2.6-14) it is seen that

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(\nu_1,\ \nu_2) \tag{2.6-21}$$

is an $F(\nu_1,\ \nu_2)$ random variable with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom.

If σ_1^2 were to equal σ_2^2, then equation (2.6-21) would undergo an obvious simplification. As a case in point, consider the ratio of $(\bar{X} - \mu_X)^2/(\sigma_X^2/n)$, which is the square of a N(0,1) random variable, and $(n - 1)S_X^2/\sigma_X^2$, which is a $\chi^2(n - 1)$ random variable, where $\bar{X}$ and S_X^2 are statistics developed from the same population. It can be shown (rather arduously) that $\bar{X}$ and S_X^2 are independent, even though they originate from the same random sample. Thus,

$$\frac{(\bar{X} - \mu_X)^2/(\sigma_X^2/n)}{S_X^2/\sigma_X^2} = \frac{(\bar{X} - \mu_X)^2}{S_X^2/n} \sim F(1,\ n - 1) \tag{2.6-22}$$

The square root of (2.6-22) is also known as a T random variable with n - 1 degrees of freedom. However, as the T random variable is, in general, equal to the square root of an $F(\nu_1, \nu_2)$ random variable with $\nu_1 = 1$, no additional time will be devoted to it.

Problem 2.6-1

Residuals E from a titration experiment (see section 2.2.4) have the following values in moles of acid.

-.011, +.003, +.004, -.01, +.005,
+.014, +.004, +.001, -.01, +.003

calculate $\bar{\epsilon}$ and s_E^2 from this random sample. Assume $\mu_E = 0$; from equation (2.6-22) derive the probability statement

$$P\left(-\sqrt{F_\alpha(1, n-1)} \leq \bar{E}/(S_E/\sqrt{n}) \leq \sqrt{F_\alpha(1, n-1)}\right) = 1 - \alpha$$

(hint: $a^2 \leq b$ is equivalent to $-\sqrt{b} \leq a \leq \sqrt{b}$). Find the interval corresponding to this statement when $\alpha = 0.05$ (that is, go to table 2.10-3 and find $F_\alpha(1, n-1)$ and then calculate $-\sqrt{F_\alpha(1, n-1)}$ and $\sqrt{F_\alpha(1, n-1)}$). How does the value of the statistic $\bar{E}/(S_E/\sqrt{n})$, calculated from the above random sample compare with this interval? Would you expect this result? Would it bother you if the value fell outside the interval?

2.7. CENTRAL LIMIT THEOREM

An interesting and difficult to prove theorem of statistics and probability, known as the Central Limit Theorem, concerns the sum of random variables:

Let $X_1, X_2, \ldots, X_n$ be a sequence of identically distributed, independent random variables each with mean μ_X and variance σ_X^2. Then the distribution of

$$\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}}$$

tends to a standard normal random variable as n goes to infinity. That is,

$$\lim_{n\to\infty} P\left(\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \qquad (2.7\text{-}1)$$

regardless of the distribution of X_i, $i = 1, \ldots, n$.

The key to understanding the impact of this theorem is to realize that the underlying distribution of the X_i random variables can be that of any random variable. Thus, for instance, a chi-square random variable with n degrees of freedom is the sum of other chi-square random variables, and if n becomes large enough, the chi-square random variable approaches the normal random variable (in fact, the normal tables are used to approximate the chi-square random variable for large values of n).

Another result of this theorem concerns the robustness of some of the distributions developed in the previous section. In particular, if the underlying distribution of X_i's were not normal in equation (2.6-22), this statistic would still be approximately an F(1, n - 1) random variable,

provided the sample size n were large. The argument for this statement proceeds as follows: Since the numerator of (2.6-22) is, when n is large, the square of an approximately normal random variable (by the central limit theorem), it will tend to be a chi-square distributed random variable with one degree of freedom. The denominator, on the other hand, will approach unity for large n, since another law of probability dictates that, as n becomes large, S_X^2 approaches σ_x^2 (this phenomenon occurs regardless of the underlying distribution). The net result is that, regardless of the distribution of the X_i random variables, equation (2.6-22) tends, for large n, to be a chi-square distributed random variable with one degree of freedom. However, it can be shown that, for large ν_2, the $F(\nu_1, \nu_2)$ random variable (equation (2.6-14)) tends to a chi-square random variable whose value has been diminished by a factor of $1/\nu_1$. Thus, regardless of the underlying distribution of the X_i random variables, we say that equation (2.6-22) behaves asymptotically as an F(1, n - 1) random variable when n is large, as both (2.6-22) and (2.6-14) have the same distribution for the limiting case where the degrees of freedom in the denominator become large.

2.8. CONFIDENCE LIMITS

We have already noted that statistics are random variables themselves. Now we wish to use the information developed in the previous sections concerning the form of these random variables to make a statement about the reliability of these statistics as estimators. We attempt to define an interval, based upon the statistic, such that a certain percentage of all such intervals, as constructed from different random samples, contain the population parameter that the statistic is thought to estimate. For example, if 5/6 of all possible intervals constructed from repeated

sampling contain the population parameter θ, then there is a probability of 5/6 that the interval we construct from any given random sample actually contains θ (see figure 2.8-1).

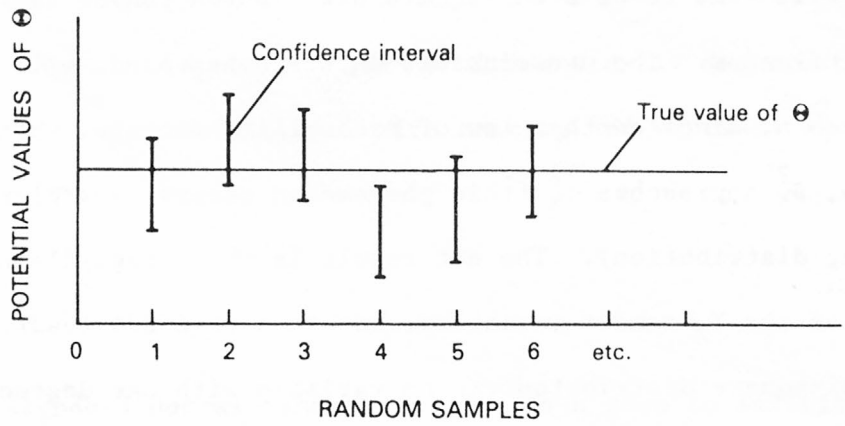


fig. 2.8-1

As an example of the interval-construction process, consider the statistic $\bar{X}$ and the population parameter μ_X. We know, from the central limit theorem, that this statistic is approximately normally distributed with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$ and that

$$\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}} \sim N(0,1) \qquad (2.8\text{-}1)$$

is approximately true. Of course, when $\bar{X}$ is based on a random sampling of a normal population, then (2.8-1) is exactly true. This standard normal random variable will be used to devise a $(1 - \alpha)100\%$ confidence interval for μ_X. This objective is achieved by first looking at the probability statement

$$P(-N_{\alpha/2}(0,1) \leq N(0,1) \leq N_{\alpha/2}(0,1)) = 1 - \alpha \qquad (2.8\text{-}2)$$

and finding the values $\pm N_{\alpha/2}(0,1)$ which correspond to $1 - \alpha$. This probability statement says that, $(1 - \alpha)100\%$ of the time, a value of $N(0,1)$, obtained from a repetition of the experiment, will fall between $-N_{\alpha/2}(0,1)$ and $N_{\alpha/2}(0,1)$. Assuming that σ_X is known, and with a little help from equation (2.8-1), equation (2.8-2) can be rewritten as

$$P\left(\bar{X} - N_{\alpha/2}(0,1)\frac{\sigma_X}{\sqrt{n}} \leq \mu_X \leq \bar{X} + N_{\alpha/2}(0,1)\frac{\sigma_X}{\sqrt{n}}\right) = 1 - \alpha \qquad (2.8\text{-}3)$$

This probability statement says that, $(1 - \alpha)100\%$ of the time, the interval $(\bar{X} - N_{\alpha/2}(0,1)\sigma_X/\sqrt{n},\ \bar{X} + N_{\alpha/2}(0,1)\sigma_X/\sqrt{n})$ constructed with a value of $\bar{X} = \bar{x}$ from a particular random sample will contain μ_X. Thus,

$$\bar{x} - N_{\alpha/2}(0,1)\frac{\sigma_X}{\sqrt{n}} \leq \mu_X \leq \bar{x} + N_{\alpha/2}(0,1)\frac{\sigma_X}{\sqrt{n}} \qquad (2.8\text{-}4)$$

is a $(1 - \alpha)100\%$ confidence interval for a random sample of size n, whose variance is known and whose sample mean $\bar{x}$ can be calculated. The investigator would be able to say that the probability is $1 - \alpha$ that this interval contains μ_X; however, in interpreting this statement, it must be realized that the interval is random, not μ_X.

As an actual numeric example of an application of (2.8-4) consider the following data: $\sigma_X = 0.3$, $\bar{x} = 2.6$, $n = 36$. Find a 95% confidence interval. From Table 2.10-1, for $\alpha/2 = 0.025$, we see that $N_{\alpha/2}(0,1) = 1.96$. Hence a 95% confidence interval is

$$2.6 - (1.96)(0.3/\sqrt{36}) \leq \mu_X \leq 2.6 + (1.96)(0.3)/\sqrt{36})$$

or

$$2.50 \leq \mu_X \leq 2.70$$

Thus, since 95 out of 100 intervals so constructed contain the mean, there is a 95% probability that this one contains μ_X.

If σ_X is not known, (2.8-4) cannot be used. However, if the population values are nearly normal, then σ_X^2 can be estimated by S_X^2 as discussed in section 2.6-2 and one can use either the T distribution (with n - 1 degrees of freedom) or the F distribution as given in (2.6-22) to make an appropriate probability statement that can be converted to an interval on μ_X.

Problem 2.8-1

a) Seven gold assays from stockpiled ore are, 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 grams per metric ton. Find a 95% confidence interval for the mean grade of the ore assuming an approximate normal distribution (hint: use equation (2.6-22): why?).

b) Write an interpretation of this interval.

2.9. HYPOTHESIS TESTING

Assume that you have determined by a method which you consider to be very good that a population parameter θ should take on a particular range of values. On the other hand, another independent source suggests that the parameter θ should take on a value b, which lies outside this range. This discrepancy is disconcerting, and you need some method of testing this independent estimate of θ. You construct a hypothesis, referred to as the null hypothesis, H_0, that b is the true value of the parameter; symbolically this may be stated:

$$H_0: \quad \theta = b$$

Ideally, of course, you wish to reject this hypothesis, but a procedure is needed whereby you can approach the problem objectively. When a random sample is available to the investigator, hypothesis tests can provide this procedure.

If, from a random sample X_1, X_2, ..., X_n, a test statistic ψ can be constructed which, in some manner, is a measure of θ, then often a statistical method can be devised to test the probable veracity of the null hypothesis. It is assumed that the distribution of the test statistic is known under the assumption of the null hypothesis, or at least can be approximated. Of course, by definition a statistic cannot contain any unknown parameters. Any population parameters which it may contain must be known either by hypothesis or some other means; otherwise ψ would cease to be a statistic. The statistical test will consist of finding a critical interval with a low probability of occurrence under the null hypothesis such that, should a value of ψ as determined from a random sample fall into this interval, the null hypothesis would be rejected and the alternate hypothesis H_1, which usually consists of one of the following, would be accepted:

H_1: $\theta > b$

H_1: $\theta < b$

H_1: $\theta \neq b$

The alternate hypothesis chosen depends on the nature of the test. A method of intelligently selecting critical intervals must be devised before the test can be completed, for not any arbitrary interval with a small probability of occurrence will do.

2.9.1. Type I Error

In hypothesis-testing procedures, one is ultimately concerned with the possibility of rejecting the null hypothesis when it is true. The objective of hypothesis testing is to make as small as possible the probability of committing this error, referred to as a type I error. That is, the probability statement

$$P(\text{reject } H_0 | H_0 \text{ true}) = \alpha \qquad (2.9\text{-}1)$$

is constructed and α, the level of significance of the test, is chosen as small as the investigator deems reasonable. For continuous random variables, the probability α must be associated with some interval about the test statistic ψ, the statistic fulfilling the requirements of the null hypothesis. Generally speaking, the test statistic ψ will contain an estimator $\hat{\theta}$ of the population parameter θ. If $\hat{\theta}$, as evaluated from some arbitrary random sampling of the experiment, were to have a value close to b, the assumed value of θ under the null hypothesis, we would not expect to reject the null hypothesis. Rather, only when this value of $\hat{\theta}$ were distant from b would the null hypothesis be rejected. Thus, the logical choice of an interval in ψ would be one in which all possible values of $\hat{\theta}$ used in the calculation of ψ would be as distant as possible from b. When the distribution of ψ has infinite tails, then this procedure will cause the interval to include one or both tails, depending on the nature of the test. This interval, whose exact starting and(or) ending point(s) will be determined by the significance level α of the test, will correspond to the critical region where H_0 will be rejected should a calculated value of ψ fall into this region. In most cases, this procedure will cause the critical interval to obtain its maximum length at the chosen significance level α.

When transforming (2.9-1) into a probability statement over the test statistic ψ, it is often preferable to first consider the impact of the alternate hypothesis on $\hat{\theta}$. Consider again the null hypothesis where $\theta = b$; then should $\theta > b$ properly represent the alternate hypothesis, it is useful to consider that, heuristically if not exactly, the probability of committing a type I error can be stated $P(\hat{\theta} > a|H_0)$, where a is some value of $\hat{\theta}$ such that $b < a < \infty$, thus giving one an understanding that $\hat{\theta}$ must take on, relatively, a large positive value in order for H_0 to be rejected. For this alternate hypothesis, a more accurate statement of (2.9-1) usually takes the form

$$P(\psi > c|\theta = b) = \alpha \qquad (2.9\text{-}2)$$

since the distribution of ψ is always assumed to be known.

In order to complete the above test, a value for c corresponding to α is obtained from a table of cumulative probabilities. Values of ψ less than c correspond to a region where the probability of committing a type I error may not be small. Therefore, if a value of ψ is less than c, we are forced to accept the null hypothesis to avoid committing a type I error. If this value is larger than c, then the probability of committing this error is considered small, and we can confidently reject H_0 at the α significance level.

Examples of hypothesis testing, which should clarify the actual mechanics of the procedure, are presented subsequently; however, before proceeding to these examples, it should be noted that we are frequently required to play the role of the devil's advocate in hypothesis testing. It often happens that we really desire to test the acceptability of a hypothesized value of a parameter. To accomplish this task, we first attempt to reject this value by making it the subject of the null

hypothesis. If we cannot reject the null hypothesis, then we must admit that the hypothesized value is indeed a candidate for the true value of the parameter in question.

2.9.2. One-Tailed Test

As an example of developing the probability statement associated with equation (2.9-2), assume that we wish to test the hypothesis that the mean of a population is μ_0, versus the alternate hypothesis that the population mean is greater than μ_0 (assume that the standard deviation σ_X is known):

$$H_0: \ \mu_X = \mu_0$$

versus

$$H_1: \ \mu_X > \mu_0$$

This test is referred to as a <u>one-tailed test</u> since the alternate hypothesis only allows for a mean greater than that indicated by H_0.

A random variable is needed whereby we may build a probability statement around the type I error. Assume that data in the form of a random sample $X_1, X_2, \ldots, X_n$ from the population exist; a natural random variable for this purpose would be the estimator of the mean $\bar{X}$. Since $\bar{X}$ is an estimator of μ_X, and since the alternate hypothesis presupposes that μ_X is large, it would seem reasonable to reject H_0 if a value of $\bar{X}$, as determined from a random sample, were significantly larger than μ_0. From the central limit theorem, it is assumed that $\bar{X}$ is approximately normally distributed with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$. Probability statement (2.9-2) can be represented in terms of the

statistic $\bar{X}$ as

$$P(\bar{X} > a|H_0) = \alpha \qquad (2.9\text{-}3)$$

Although the distribution of $\bar{X}$ is known, a statistic which will allow us to incorporate the null hypothesis that $\mu_X = \mu_0$ is needed. A statistic meeting this requirement and for which values of all the parameters can be supplied is $(\bar{X} - \mu_X)/(\sigma_X/\sqrt{n})$. With this test statistic, probability statement (2.9-3) can be restated as

$$P\left(\frac{\bar{X} - \mu_0}{\sigma_X/\sqrt{n}} > c|\mu_X = \mu_0\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma_X/\sqrt{n}} > N_\alpha(0,1)\right) = \alpha \qquad (2.9\text{-}4)$$

where μ_0 is used in place of μ_X in order to satisfy the null hypothesis. Note that, under the null hypothesis, $(\bar{X} - \mu_0)/(\sigma_X/\sqrt{n})$ is a normal random variable with mean zero and variance unity; thus, $N_\alpha(0,1)$ becomes the lower limit c of the critical region for this test.

All possible values of the test statistic greater than $N_\alpha(0,1)$, where α is the level of significance of the test, constitute the critical region where H_0 would be rejected. In other words, $N_\alpha(0,1)$ is the critical value, corresponding to the limit c in (2.9-2), which determines whether we accept or reject the null hypothesis. If a value of the test statistic is greater than $N_\alpha(0,1)$, we would reject H_0 at the α significance level. If the value were less, then we would be forced to accept the null hypothesis for fear of making a type I error.

As a sample application of this procedure, consider the data used to construct the confidence interval at the end of section 2.8: $\sigma_X = 0.3$, $\bar{x} = 2.6$ and $n = 36$. We are told that the population mean is really zero, a statement that seems rather dubious to us as we believe it to be

some positive real number. We set the null hypothesis that μ_X is indeed zero, H_0: $\mu_X = 0$, and hope that we can confidently disallow it. Our alternate hypothesis consists of our own belief; H_1: $\mu_X > 0$. As we wish to be very sure that we do not commit a type I error, we set the level of significance of our test at $\alpha = 0.025$. We determine the critical value of our test statistic from table 2.10-1: $N_\alpha(0,1) = 1.96$. We evaluate the test statistic under the assumption of the null hypothesis: $\bar{x}/(\sigma_X/\sqrt{n}) = 52$. Since this value of the test statistic is considerably larger than the critical value, we reject H_0 at the 0.025 significance level, realizing that, while we may have committed a type I error, it is highly unlikely.

2.9.3. Two-Tailed Test

Suppose that σ_X^2 is unknown, but we are given a random sample X_1, X_2, $\cdots$, X_n from a normal population. We wish to test the hypothesis

$$H_0: \mu_X = \mu_0$$

versus

$$H_1: \mu_X \neq \mu_0$$

at a signifance level α. This is referred to as a <u>two-tailed test</u>: We reject H_0 if a measure of μ_X is either significantly greater or less than μ_0.

To construct this test, recall the statistic from equation (2.6-22):

$$\frac{(\bar{X} - \mu_X)^2}{S_X^2/n} \sim F(1, n-1) \qquad (2.9\text{-}5)$$

which is the F(1, n - 1) random variable. This statistic fulfills our requirement for a test statistic: It can be used to satisfy the null hypothesis, and the remaining statistics or parameters are either known or can be evaluated from a random sample. Now consider the probability of a type I error:

$$
\begin{aligned}
&P(\text{reject } H_0 | H_0 \text{true}) \\
&\quad = P(\bar{X} < a | H_0) + P(\bar{X} > b | H_0) \\
&\quad = \alpha
\end{aligned}
\tag{2.9-6}
$$

where two critical values are necessary since it is possible to reject the null hypothesis if a value of $\bar{X}$ is either larger or smaller than μ_0. In terms of the test statistic, under the condition that the null hypothesis holds, we see that

$$
\begin{aligned}
&P\left(\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}}\right)^2 > F_\alpha(1, n-1)\right) \\
&\quad = P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(1, n-1)} \text{ or } \frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(1, n-1)}\right) \\
&\quad = P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(1, n-1)}\right) + P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(1, n-1)}\right) \\
&\quad = \alpha
\end{aligned}
\tag{2.9-7}
$$

which is equivalent to statement (2.9-6).

To complete the test, we need only to evaluate the test statistic with a random sample. If $(\bar{x} - \mu_0)^2/(s_X^2/n)$ be greater than $F_\alpha(1, n - 1)$, we would reject the null hypothesis at the α significance level.

2.9.4. Type II Error

A test statistic is usually selected for its ability to determine the probability of committing a type II error, as well as a type I error. A type II error is committed by accepting the null hypothesis when the alternate is true. By calculating the probability that the test statistic does not fall in the critical region, given that θ takes on any value other than that assumed under the null hypothesis, the probability β of committing this error can be evaluated. Thus, for tests indicated previously, β is a continuous function of possible values of the population parameter θ, other than the value b assumed under the null hypothesis. For a critical region corresponding to a given α, a good test statistic should produce small values of β for hypothetical values of θ rather distant from b. However, β should increase sharply in value as possible values of θ approach b, and obtain a value as close to one as feasible in the immediate vicinity of b. Means are available for determining test statistics which, for certain tests, excel at the above characteristics, but a presentation of these methods is beyond the scope of this course. In most cases, a statistic which contains an estimator of the population parameter being tested and for which all other parameters are either known, or estimators of said parameters are contained in the statistic, will suffice as a test statistic; however, it may not be the best test statistic.

Note that if α, the probability of committing a type I error, were made extremely small, then the null hypothesis would almost always be accepted. At first glance, one would assume that something was amiss in the hypothesis testing procedure, as it is apparently possible to bias the test by selecting an extreme value for α. However, when the value of α is decreased, the probability of committing a type II error, $\beta(\theta)$, is

increased for all values θ. Thus, an investigator who seeks to avoid committing a type I error by intentionally selecting a small value for α runs an increased risk of committing a type II error, which is equally as damaging. If need be, a plot of $\beta(\theta)$ can be made for various hypothetical values of α and θ; this can often be a rather complicated task. A rule-of-thumb value for α is 0.05, which appears to serve hypothesis test users well in most cases.

2.9.5. Summary of Method

To summarize, the steps for testing a hypothesis concerning a population parameter θ are:

1) Define the null hypothesis H_0: $\theta = \theta_0$.
2) Decide upon the nature of the test; that is, H_1: $\theta < \theta_0$, H_1: $\theta > \theta_0$ or H_1: $\theta \neq \theta_0$.
3) Choose a level of significance α.
4) Select an appropriate test statistic and establish the critical region.
5) Compute the value of the statistic from a random sample of size n.
6) Draw conclusion of test: Reject H_0 if the statistic has a value in the critical region; otherwise accept H_0.

Problem 2.9-1

a) Set up problem 2.6-1 as a hypothesis test (do not complete the test).

b) Given two random samples from independent normal populations with the following sample statistics,

Statistic	Random sample 1	Random sample 2
n	25	16
$\bar{x}$	82	78
s_x	8	7

test the following hypotheses at a significance level of $\alpha = 0.05$:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

$$H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

c) An outside source informs you that the stockpiled ore of problem 2.8-1 actually only assays an average of 9.8 grams per metric ton. Can you refute this claim at a significance level of 0.05? (Construct a hypothesis test for this purpose.)

2.10 TABLES OF PROBABILITY DISTRIBUTIONS [1/]

[1/] All tables modified from Walpole and Myers (1972), with permission from the publisher.

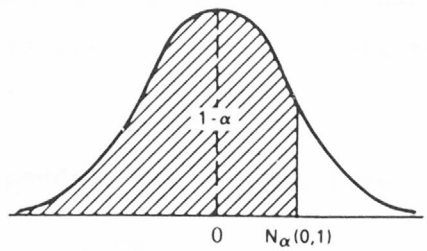


Table 2.10-1

Areas, $1-\alpha$, Under the Normal Curve

$N_\alpha(0,1)$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

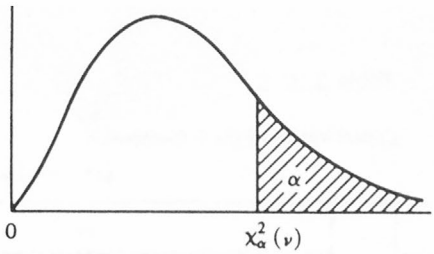


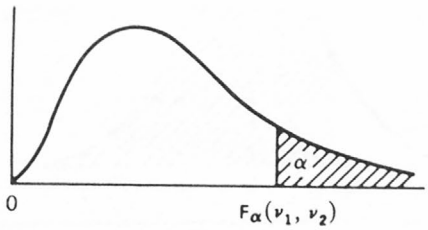
Table 2.10-2

Critical Values of the Chi-Square Distribution

	α							
ν	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	3.841	5.024	6.635	7.879
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.832	15.086	16.750
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672

Abridged from Table 8 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E. S. Pearson and the Biometrika Trustees.

Table 2.10-3



Critical Values of the F Distribution

$F_{0.05}(\nu_1, \nu_2)$

ν_2	ν_1								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Reproduced from Table 18 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E. S. Pearson and the Biometrika Trustees.

Table 2.10-3 continued

$F_{0.05}(\nu_1, \nu_2)$

ν_2	ν_1									
	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

2.11. APPENDICES

2.11.1. Correlation of Two Linearly Related Random Variables

Consider a linear relationship between two random variables X and Y such that

$$Y = a + bX + E \tag{2.11-1}$$

where E represents a zero-mean random error (independent of X). Then

$$\sigma_Y^2 = b^2\sigma_X^2 + \sigma_E^2 \tag{2.11-2}$$

since $\sigma_{XE} = E[(X - \mu_X)E] = 0$. By direct calculation of σ_{XY} from (2.11-1), one obtains

$$\begin{aligned}\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[(X - \mu_X)(b(X - \mu_X) + E)] \\ &= b\sigma_X^2 \end{aligned} \tag{2.11-3}$$

Upon squaring both sides of (2.11-3) and dividing by $\sigma_X^2\sigma_Y^2$, one obtains

$$\rho_{XY}^2 = b^2 \frac{\sigma_X^2}{\sigma_Y^2} \tag{2.11-4}$$

which, from (2.11-2) can be put in the form

$$\rho_{XY}^2 = 1 - \frac{\sigma_E^2}{\sigma_Y^2} \tag{2.11-5}$$

Again from (2.11-2), it is seen that

$$b^2\sigma_X^2 = \sigma_Y^2 - \sigma_E^2 \tag{2.11-6}$$

and, since $b^2\sigma_X^2$ is a non-negative quantity, σ_Y^2 must be greater than or equal to σ_E^2. This shows that, for a linear relationship, ρ_{XY}^2 is either less than unity, or equal to one if σ_E^2 is equal to zero.

2.11.2. Expected Value of Variance Estimator

The sample statistic S_X^{*2} is defined

$$S_X^{*2} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{2.11-7}$$

By taking expected values of both sides, one sees that

$$E\left[S_X^{*2}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} (X_i - \bar{X})^2\right]$$

$$= \frac{1}{n} E\left[\sum_{i=1}^{n} (X_i - \mu_X - (\bar{X} - \mu_X))^2\right]$$

$$= \frac{1}{n} \left[\left(\sum_{i=1}^{n} E[(X_i - \mu_X)^2]\right) - E[n(\bar{X} - \mu_X)^2]\right] \tag{2.11-8}$$

where use is made of the fact that $\sum_{i=1}^{n} (X_i - \mu_X) = n(\bar{X} - \mu_X)$. Now the second expected value in (2.11-8) becomes

$$E[n(\bar{X} - \mu_X)^2] = \frac{1}{n} E\left[\sum_{i=1}^{n} (X_i - \mu_X) \sum_{j=1}^{n} (X_j - \mu_X)\right]$$

$$= \frac{1}{n} E\left[\sum_{i=1}^{n} \sum_{j=1}^{n} (X_i - \mu_X)(X_j - \mu_X)\right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[X_i, X_j] \tag{2.11-9}$$

Since X_i and X_j are randomly selected, and therefore independent,

$$Cov[X_i, X_j] = \begin{cases} \sigma_X^2 & i = j \\ 0 & i \neq j \end{cases} \qquad (2.11\text{-}10)$$

Thus, equation (2.11-9) becomes

$$E[n(\bar{X} - \mu_X)^2] = \sigma_X^2 \qquad (2.11\text{-}11)$$

which allows us to write equation (2.11-8) as

$$E\left[S_X^{*2}\right] = \frac{1}{n}\left[\left(\sum_{i=1}^{n} \sigma_X^2\right) - \sigma_X^2\right]$$

$$= \frac{n-1}{n}\,\sigma_X^2 \qquad (2.11\text{-}12)$$

and demonstrates the desired result. Note that equation (2.11-11) also demonstrates that

$$\sigma_{\bar{X}}^2 = \sigma_X^2/n \qquad (2.11\text{-}13)$$

since $\sigma_{\bar{X}}^2 = E[(\bar{X} - \mu_X)^2]$.

References Cited

Nutter, Lawrence, 1973, Hydrogeology of the carbonate rocks, Fredrick and Hagerstown Valleys, Maryland: Maryland Geological Survey Report of Investigations No. 19, 70 p.

Siddiqui, S. H., 1969, Hydrogeologic factors influencing well yields and aquifer hydraulic properties of folded and faulted carbonate rocks in Central Pennsylvania: State College, Pennsylvania State University, unpublished Ph.D. thesis, 502 p.

Walpole, R. E. and Myers, R. H., 1972, Probability and statistics for engineers and scientists: New York, Macmillan, 506 p.

Additional Reading

Benjamin, J. R., and Cornell, A., 1970, Probability, statistics, and decision for civil engineers: New York, McGraw-Hill, 684 p.

3. REGRESSION SOLUTION OF MODELING PROBLEMS

3.1. INTRODUCTION AND BACKGROUND

Ground-water flow models are members of a class of models known as mathematical models, in which the physical model of the flow system is replaced by mathematical expressions containing mathematical variables, parameters, and constants (Krumbein and Graybill, 1965, p. 15). Mathematical models always involve simplification of the actual (true) physical system. Krumbein and Graybill (1965, p. 15) argue that mathematical models can be classified into several types, including deterministic models, statistical models, and stochastic-process models.

A deterministic model is one in which the dependent variable(s) can be exactly computed from an expression involving independent variables, parameters, and constants. Note that deterministic models do not have to be physically based, but may instead be completely empirical. The classical and inverse flow models discussed in section 1 are of the physically based deterministic type. In contrast, a statistical model is a deterministic model that has one or more random components added. These random components frequently involve measurement or other errors, but may involve separate sources of random variability as well. Incorporation of the errors in both observed heads and estimated parameters discussed in section 1 converts the deterministic flow model into a statistical model.

The term stochastic model can be considered to be synonymous with the term statistical model (Krumbein and Graybill, 1965, p. 19). A stochastic-process model may consider random effects such as those contained in the statistical model, but in addition has a stochastic process built into it. Generation of a spatially varying permeability field in an aquifer has been considered to be a stochastic process by Bakr

and others (1978), Gutjahr and others (1978), and Smith and Freeze (1979a, 1979b). Recently this type of process has been incorporated into a parameter estimation scheme for a steady-state ground-water flow model (Kitanidis and Vomvoris, 1983). Stochastic-process models will not be considered further here.

3.1.1. Assumed Model Structure

Consider an experiment where two variables, ξ and Y, are measured repeatedly. The independent variable, ξ, is considered to be a precisely defined quantity, whereas the dependent variable, Y, whose values depend upon values of the independent variable, contains some error resulting from the experimental process. A scatter diagram of the data might appear as in figure 3.1-1.

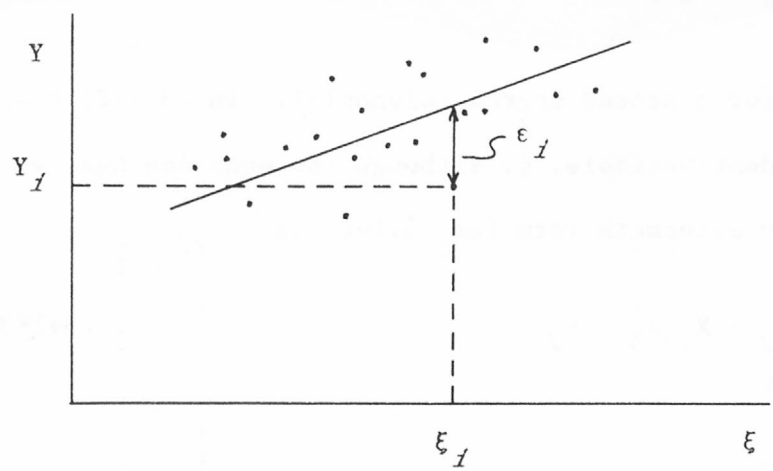


fig. 3.1-1

From the scatter diagram, or from physical considerations, the experimentalist may decide that an appropriate model equation for the data is

$$Y_i = \beta_1 + \beta_2 \xi_i + \varepsilon_i \qquad (3.1\text{-}1)$$

where β_1 and β_2 are the intercept and slope of the equation for a straight line, subscript i represents the ith observation of (ξ, Y), and ε_i is the true error in Y for observation i. The quantity $\beta_1 + \beta_2 \xi$ is the deterministic part of the equation (the computed value of the dependent variable), and, because ε_i is the true error, <u>parameters</u> β_1 and β_2 are the true parameters representing the deterministic part of the model response. True error ε_i, often called a <u>disturbance</u>, is a random variable and, thus, represents the stochastic part of the model response. Note that if the model is correct and no other source of bias in ε_i exists, $E(\varepsilon_i) = 0$.

Equation (3.1-1) is linear in parameters β_1 and β_2. Another example of a model equation that is linear in the parameters is

$$Y_i = \beta_1 + \beta_2 \xi_i + \beta_3 \xi_i^2 + \varepsilon_i \qquad (3.1\text{-}2)$$

which is the equation for a second degree polynomial. In (3.1-2) there is still only one independent variable, ξ, although the equation has two terms containing ξ. An alternate form for (3.1-2) is

$$Y_i = X_{i1}\beta_1 + X_{i2}\beta_2 + X_{i3}\beta_3 + \varepsilon_i \qquad (3.1\text{-}3)$$

where

$$\left.\begin{aligned} X_1 &= 1 \\ X_2 &= \xi \\ X_3 &= \xi^2 \end{aligned}\right\} \qquad (3.1\text{-}4)$$

In general, any equation that is linear in parameters β_1, β_2, . . ., β_p, where there are p parameters for the system, can be written in the form

$$Y_i = X_{i1}\beta_1 + X_{i2}\beta_2 + \ldots + X_{ip}\beta_p + \varepsilon_i \tag{3.1-5}$$

where

$$X_{ij} = X_{ij}(\xi_{i1}, \xi_{i2}, \ldots, \xi_{ik}) \tag{3.1-6}$$

is a function of k independent variables that multiplies the jth parameter and does not contain the parameters. Because

$$\frac{\partial}{\partial \beta_j}\left(X_{i1}\beta_1 + X_{i2}\beta_2 + \ldots + X_{ip}\beta_p\right) = X_{ij} \tag{3.1-7}$$

the X terms are often called <u>sensitivity coefficients</u> or, simply, sensitivities. They indicate the change in the model response (the computed value of the dependent variable) at observation point i for a unit change in parameter β_j. Equation (3.1-5) can be written compactly in matrix form as

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon} \tag{3.1-8}$$

where

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \tag{3.1-9}$$

$$\underline{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & & & \\ \vdots & & & \\ X_{n1} & \cdots & & X_{np} \end{bmatrix} \tag{3.1-10}$$

$$\underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \tag{3.1-11}$$

$$\underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{3.1-12}$$

and there are assumed to be n observations.

Often we are faced with models where the equation is no longer linear in the parameters, $\underline{\beta}$. For example, suppose that the model equation is

$$Y_i = \sqrt{\beta_1 + \beta_2} \quad \tan\left(\frac{\sqrt{\beta_1 + \beta_2}}{\beta_1} \xi_i \right) + \varepsilon_i \tag{3.1-13}$$

Equation (3.1-13) cannot be reduced to the form (3.1-5), and, thus, is not linear in the parameters. Equations of this type are written in the general vector form

$$\underline{Y} = \underline{f}(\underline{\xi}_1, \underline{\xi}_2, \ldots, \underline{\xi}_k; \beta_1, \beta_2, \ldots, \beta_p) + \underline{\varepsilon} \tag{3.1-14}$$

or, in more compact form,

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{\beta}) + \underline{\varepsilon} \tag{3.1-15}$$

where $\underline{f}$ is an n-vector, each element f_i of which is a general function of

the k independent variables, $\xi_{i\ell}$ ($\ell = 1, 2, \ldots, k$), and p parameters, β_j ($j = 1, 2, \ldots, p$). Equation (3.1-15) incorporates (3.1-8) because (3.1-8) is simply the special case where $\underline{f}(\underline{\xi}, \underline{\beta}) = \underline{X}(\underline{\xi})\underline{\beta}$.

Some equations may be nonlinear in $\underline{\beta}$ but linear in some transformation of $\underline{\beta}$. For example, the model

$$f_i = \beta_1 (\beta_2)^{X_i} \varepsilon_i \tag{3.1-16}$$

in which the error ε_i is multiplicative, is nonlinear in β_1 and β_2. However, (3.1-16) may be written as

$$\log f_i = \log \beta_1 + X_i \log \beta_2 + \log \varepsilon_i \tag{3.1-17}$$

which is linear in $\log \beta_1$ and $\log \beta_2$, and has an additive error term. Thus, it is of the standard linear form. Equations such as (3.1-16) are frequently best utilized in their transformed, thus linearized, form. However, all model analyses (to be discussed further on) would probably be made in terms of the transformed variables, and this would have to be remembered when results were interpreted.

Types of models other than the linear and nonlinear ones discussed above also exist. Some types involve a complex model equation that cannot be solved explicitly for the dependent variable. In other cases the function $\underline{f}(\underline{\xi}, \underline{\beta})$, which is assumed to be a known function of $\underline{\xi}$ and $\underline{\beta}$, cannot be obtained and, therefore, must be replaced by a numerical formulation. However, the basic model structure of (3.1-15), where the error $\underline{\varepsilon}$ in $\underline{Y}$ is assumed to be additive to a deterministic dependent variable vector, is always assumed. Additional complexities of the other types of models are handled by auxiliary equations appended to (3.1-15). The other models are introduced at appropriate places further on.

3.1.2. Least Squares Estimation

Because the true parameter set $\underline{\beta}$ and true error set $\underline{\varepsilon}$ are generally unknown, the true model (3.1-15) must be regarded as unknown, even though the form of the model is known (or, at least assumed). We do, however, have measurements to make up the independent variable set $\underline{\xi}$ and observation set $\underline{Y}$. We would like to use these measurements and the form of the model to obtain estimates of $\underline{\beta}$ and $\underline{\varepsilon}$. The method to be explained in the following paragraphs is based on the idea that, if estimates of $\underline{\beta}$ and $\underline{\varepsilon}$ can be found such that the error structure of the true model is duplicated as closely as possible, then the resulting model should, in some sense, be the best possible approximation of the true model.

Assume that all ε_i ($i = 1, 2, \ldots, n$) as random variables, have finite common variance σ^2 and that ε_i and ε_j, $i \neq j$, are uncorrelated. Then

$$Var(\underline{\varepsilon}) = \underline{I}\sigma^2 \tag{3.1-18}$$

The scalar variance σ^2 can be solved for by taking the trace of both sides of (3.1-18):

$$tr[Var(\underline{\varepsilon})] = tr(\underline{I})\sigma^2$$

or

$$tr[E((\underline{\varepsilon} - E(\underline{\varepsilon}))(\underline{\varepsilon} - E(\underline{\varepsilon}))^T)] = n\sigma^2$$

or

$$E[(\underline{\varepsilon} - E(\underline{\varepsilon}))^T(\underline{\varepsilon} - E(\underline{\varepsilon}))] = n\sigma^2$$

from which

$$\sigma^2 = \frac{E[(\underline{\varepsilon} - E(\underline{\varepsilon}))^T(\underline{\varepsilon} - E(\underline{\varepsilon}))]}{n} \tag{3.1-19}$$

Ordinarily the assumptions would be made that the model being used is the correct one and that no other source of bias in $\underline{\varepsilon}$ exists, so that $E(\underline{\varepsilon}) = \underline{0}$ and

$$\sigma^2 = \frac{E(\underline{\varepsilon}^T \underline{\varepsilon})}{n} \tag{3.1-20}$$

Equation (3.1-20) indicates that the sum of squared disturbances over all observations, averaged over many sets of observations, divided by n yields σ^2.

As indicated previously, the investigator only has available the data and the form of the model, so that $\underline{\varepsilon}$, σ^2, and $\underline{\beta}$ must all be considered as unknowns. However, a good approximation of the true model would produce estimates of ε that, for many observations, would yield a variance approaching σ^2. Let $\underline{b}$ be an estimator of $\underline{\beta}$. Then a linear model incorporating $\underline{b}$ is

$$\underline{Y} = \underline{X}\,\underline{b} + \underline{e} \tag{3.1-21}$$

where the vector $\underline{e}$ is an estimate of $\underline{\varepsilon}$ called the residual vector. From (3.1-20) and (3.1-21) an estimate of σ^2 is

$$\tilde{\sigma}^2 = \frac{\underline{e}^T \underline{e}}{n} \tag{3.1-22}$$

It would be expected that most arbitrary parameter sets would yield values of $\tilde{\sigma}^2$ that are larger than σ^2 because they would yield models that do not fit the data well. Of all possible parameter sets $\underline{b}$, the one that fits the data the best and at the same time minimizes $\tilde{\sigma}^2$ is the set that minimizes the sum of squares function $S(\underline{b})$,

$$S(\underline{b}) = \underline{e}^T \underline{e} \tag{3.1-23}$$

with respect to $\underline{b}$. The process of finding estimates of σ^2 and $\underline{\beta}$ by minimizing $S(\underline{b})$ is termed least squares estimation. It is developed fully later on.

Recall that (3.1-18), and, hence, (3.1-22) and (3.1-23) assume that the ε_i all come from the same distribution having variance σ^2 and that ε_i and ε_j, $i \neq j$, are uncorrelated. In some instances ε_i and ε_j have different variances, say σ_i^2 and σ_j^2, and they may even be correlated so that they have nonzero covariance σ_{ij}. In this case (3.1-18) must be written in the more general form

$$Var(\underline{\varepsilon}) = \underline{V}\sigma^2 \tag{3.1-24}$$

where $\underline{V}\sigma^2$ is a symmetric, positive definite variance-covariance matrix defined as

$$\underline{V}\sigma^2 = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ & & \cdots & & \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{bmatrix} \tag{3.1-25}$$

In (3.1-24) σ^2 is no longer the common variance of all ε_i but is instead another type of common variance. Its exact meaning can be discerned as follows: Define $\underline{V}^{\frac{1}{2}}$ as the nonsingular symmetric matrix such that $\underline{V}^{\frac{1}{2}}\underline{V}^{\frac{1}{2}} = \underline{V}$. Then, from (3.1-24),

$$\underline{V}^{-\frac{1}{2}} Var(\underline{\varepsilon}) \underline{V}^{-\frac{1}{2}} = \underline{V}^{-\frac{1}{2}} \underline{V}\, \underline{V}^{-\frac{1}{2}} \sigma^2$$

or

$$Var(\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}) = \underline{I}\sigma^2 \tag{3.1-26}$$

from which it can be seen that σ^2 is the constant or common variance of the transformed disturbances $\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}$. From (3.1-26) it can be seen that these disturbances are uncorrelated.

The more general conditions represented by (3.1-24) may easily be incorporated into the least squares procedure. All developments through (3.1-20) are repeated using $\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}$ instead of $\underline{\varepsilon}$. The result is

$$\sigma^2 = \frac{E[(\underline{V}^{-\frac{1}{2}}\underline{\varepsilon})^T(\underline{V}^{-\frac{1}{2}}\underline{\varepsilon})]}{n}$$

$$= \frac{E(\underline{\varepsilon}^T\underline{V}^{-1}\underline{\varepsilon})}{n} \qquad (3.1\text{-}27)$$

which suggests that

$$S(\underline{b}) = \underline{e}^T\underline{V}^{-1}\underline{e} \qquad (3.1\text{-}28)$$

be minimized instead of (3.1-23).

If (3.1-23) is minimized to find the parameter estimates when the more general error structure given by (3.1-24) is correct, then the incorrect error structure will be reflected in parameter estimates that are less precise than if (3.1-28) were used. The proper sum of squares function to minimize when (3.1-24) represents the correct error structure is (3.1-28).

Even more general cases can be postulated to yield $S(\underline{b})$ in the form

$$S(\underline{b}) = \underline{e}^T\underline{\omega}\underline{e} \qquad (3.1\text{-}29)$$

where $\underline{\omega}$ is a general symmetric positive definite weight matrix that

subsumes $\underline{V}^{-1}$. To apply (3.1-29), $\underline{\omega}$ does not necessarily reflect the error structure of $\underline{\varepsilon}$. Instead it may reflect the investigator's desire to emphasize (or de-emphasize) certain components of $S(\underline{b})$. Equation (3.1-29) is used in all developments to follow in which the general form is applicable.

As a final note, least squares estimation should be viewed as more than simply a parameter estimation procedure. The development given in this section is intended to show that the procedure is an attempt to reproduce the true model structure: the variance, σ^2, the distribution of ε_i, and $\underline{\beta}$. Although it is possible to use least squares estimation as just an algebraic process, making no assumptions about structure, considerably more information can be gained by taking the more general approach to make model analysis, including analysis of assumptions initially made concerning model structure, an integral part of the regression process.

3.1.3. Inclusion of Prior Information

The model structure given by (3.1-15) is general. Nothing is implied about the nature of $\underline{Y}$ except that it is a dependent variable, in error by the amount $\underline{\varepsilon}$. A very general interpretation of (3.1-15) is to assume that $\underline{Y}$, $\underline{f}$, and $\underline{\varepsilon}$ are each composed of two partitions, one giving sample information and one giving prior information on parameters. This viewpoint amounts to an expansion of the original formulation given in the previous section where only sample information was considered (Theil, 1963).

For example, suppose that an investigator collects data (ξ, Y) on a process for which the model equation is given by (3.1-2). However, suppose that he also has developed methods to collect some data directly on parameters β_1 and β_3 and suppose that these data can be represented by the equations

$$\left.\begin{aligned} P_1 &= a_{11}\beta_1 + u_1 \\ P_2 &= a_{21}\beta_1 + a_{23}\beta_3 + u_2 \end{aligned}\right\} \qquad (3.1\text{-}30)$$

where a_{ij} is a constant and u_i is a random error. If $a_{11} = 1$, then P_1 is a direct observation, subject to error u_1, of β_1. The entire set of equations representing the system, then, can be written

$$\left.\begin{aligned} Y_1 &= \beta_1 + \xi_1\beta_2 + \xi_1^2\beta_3 + \varepsilon_1 \\ Y_2 &= \beta_1 + \xi_2\beta_2 + \xi_2^2\beta_3 + \varepsilon_2 \\ &\cdots \\ Y_n &= \beta_1 + \xi_n\beta_2 + \xi_n^2\beta_3 + \varepsilon_n \\ P_1 &= a_{11}\beta_1 + u_1 \\ P_2 &= a_{21}\beta_1 + a_{23}\beta_3 + u_2 \end{aligned}\right\} \qquad (3.1\text{-}31)$$

If $\underline{Y}$, $\underline{X}$, and $\underline{\varepsilon}$ are augmented to include the prior information, then (3.1-31) is of the form (3.1-8) where

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \\ P_1 \\ P_2 \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \\ u_1 \\ u_2 \end{bmatrix}$$

$$\underline{X} = \begin{bmatrix} 1 & \xi_1 & \xi_1^2 \\ 1 & \xi_2 & \xi_2^2 \\ & \cdots & \\ 1 & \xi_n & \xi_n^2 \\ a_{11} & 0 & 0 \\ a_{21} & 0 & a_{23} \end{bmatrix}$$

Note that the number of observations is now the number of equations giving sample information (n) plus the number of equations giving prior information (2).

Although the equations giving prior information are often linear, they do not need to be. Hence, a general form of (3.1-15) to include the prior information may be assumed:

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{\beta}) + \underline{\varepsilon} \qquad (3.1\text{-}32)$$

where

$$\underline{Y} = \begin{bmatrix} \underline{Y}_s \\ \underline{Y}_p \end{bmatrix} \qquad \underline{f} = \begin{bmatrix} \underline{f}_s \\ \underline{f}_p \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \underline{\varepsilon}_s \\ \underline{\varepsilon}_p \end{bmatrix} \qquad (3.1\text{-}33)$$

and subscripts s and p indicate partitions of the respective vectors pertaining to sample and prior information, respectively. Corresponding to these partitions, it is convenient to redefine n as the total number of observations, $n = n_s + n_p$, where n_s is the number of pieces (or equations) of sample information and n_p is the number of pieces (or equations) of prior information.

To apply the least squares procedure to (3.1-15) as augmented by the prior information, it is assumed for now that

$$Var(\underline{\varepsilon}_s) = \underline{V}_s\sigma^2 \tag{3.1-34}$$

$$Var(\underline{\varepsilon}_p) = \underline{V}_p\sigma^2 \tag{3.1-35}$$

$$Cov(\underline{\varepsilon}_s, \underline{\varepsilon}_p) = \underline{0} \tag{3.1-36}$$

where $\underline{V}_s$ is symmetric, positive definite and of order n_s, and $\underline{V}_p$ is symmetric, positive definite and of order n_p. Equation (3.1-36) indicates that sample disturbances $\underline{\varepsilon}_s$ and prior information disturbances $\underline{\varepsilon}_p$ are not correlated with each other. With use of (3.1-34) through (3.1-36), (3.1-24) becomes

$$Var\begin{bmatrix} \underline{\varepsilon}_s \\ \underline{\varepsilon}_p \end{bmatrix} = \begin{bmatrix} Var(\underline{\varepsilon}_s) & \underline{0} \\ \underline{0} & Var(\underline{\varepsilon}_p) \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} \underline{V}_s & \underline{0} \\ \underline{0} & \underline{V}_p \end{bmatrix}$$

$$= \underline{V}\sigma^2 \tag{3.1-37}$$

With use of (3.1-37), (3.1-28) becomes

$$S(\underline{b}) = \underline{e}^T\underline{V}^{-1}\underline{e}$$

$$= \begin{bmatrix} \underline{e}_s^T & \underline{e}_p^T \end{bmatrix} \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{V}_p^{-1} \end{bmatrix} \begin{bmatrix} \underline{e}_s \\ \underline{e}_p \end{bmatrix}$$

$$= \underline{e}_s^T \underline{V}_s^{-1}\underline{e}_s + \underline{e}_p^T \underline{V}_p^{-1}\underline{e}_p \tag{3.1-38}$$

where the residual vector is defined as

$$\underline{e} = \begin{bmatrix} \underline{e}_s \\ \underline{e}_p \end{bmatrix} \tag{3.1-39}$$

and

$$\underline{V}^{-1} = \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{V}_p^{-1} \end{bmatrix} \tag{3.1-40}$$

The least squares procedure may be generalized even further by using (3.1-29) instead of (3.1-28) to define $S(\underline{b})$. In this case the weight matrix $\underline{\omega}$ is defined by

$$\underline{\omega} = \begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix} \tag{3.1-41}$$

where $\underline{\omega}_s$ is a symmetric positive definite submatrix of order n_s that pertains to the sample information and $\underline{\omega}_p$ is a symmetric positive definite submatrix of order n_p that pertains to the prior information. Because (3.1-41) is of block diagonal form like (3.1-37), zero correlation of sample and prior information is again assumed. Thus, $S(\underline{b})$ may be written in the same form as (3.1-38), or

$$S(\underline{b}) = \underline{e}_s^T \underline{\omega}_s \underline{e}_s + \underline{e}_p^T \underline{\omega}_p \underline{e}_p \tag{3.1-42}$$

Problem 3.1-1

You are charged with a ground-water study in the vicinity of Lake Ohpupu, Ecko County, Nevada (see figure 1). Estimates of transmissivity and recharge for the confined aquifer surrounding the lake are necessary for the completion of your report. Taking advantage of the unusually colinear equipotential contours on the west side of the lake (constructed from an unbiased source, of course), you decide that estimates based on a uniform stream tube will suffice. Recharge to the aquifer is largely from precipitation and is uniform over the region. Assume that the boundary heads at the range front and the lake are imprecisely known; estimates of these parameters will also be necessary. Your project has limited funds to bore n_s holes along the stream tube and obtain measurements of head at n_s locations of distance, s, from the range front.

The steady state flow equation for a stream tube is

$$\frac{d}{ds}\left(TD\,\frac{dh}{ds}\right) + WD = 0 \qquad (1)$$

where

T = transmissivity (ft^2/day);

W = recharge (ft/day);

D = width of stream tube (ft);

h = hydraulic head (ft); and

s = distance along tube from the range front (ft).

The boundary conditions are taken to be

$$\left.\begin{aligned} h &= h_o \text{ at } s = 0 \\ h &= h_b \text{ at } s = s_b \end{aligned}\right\} \qquad (2)$$

POTENTIOMETRIC SKETCH MAP FOR THE VICINITY OF
LAKE OHPUPU, ECKO COUNTY, NEVADA

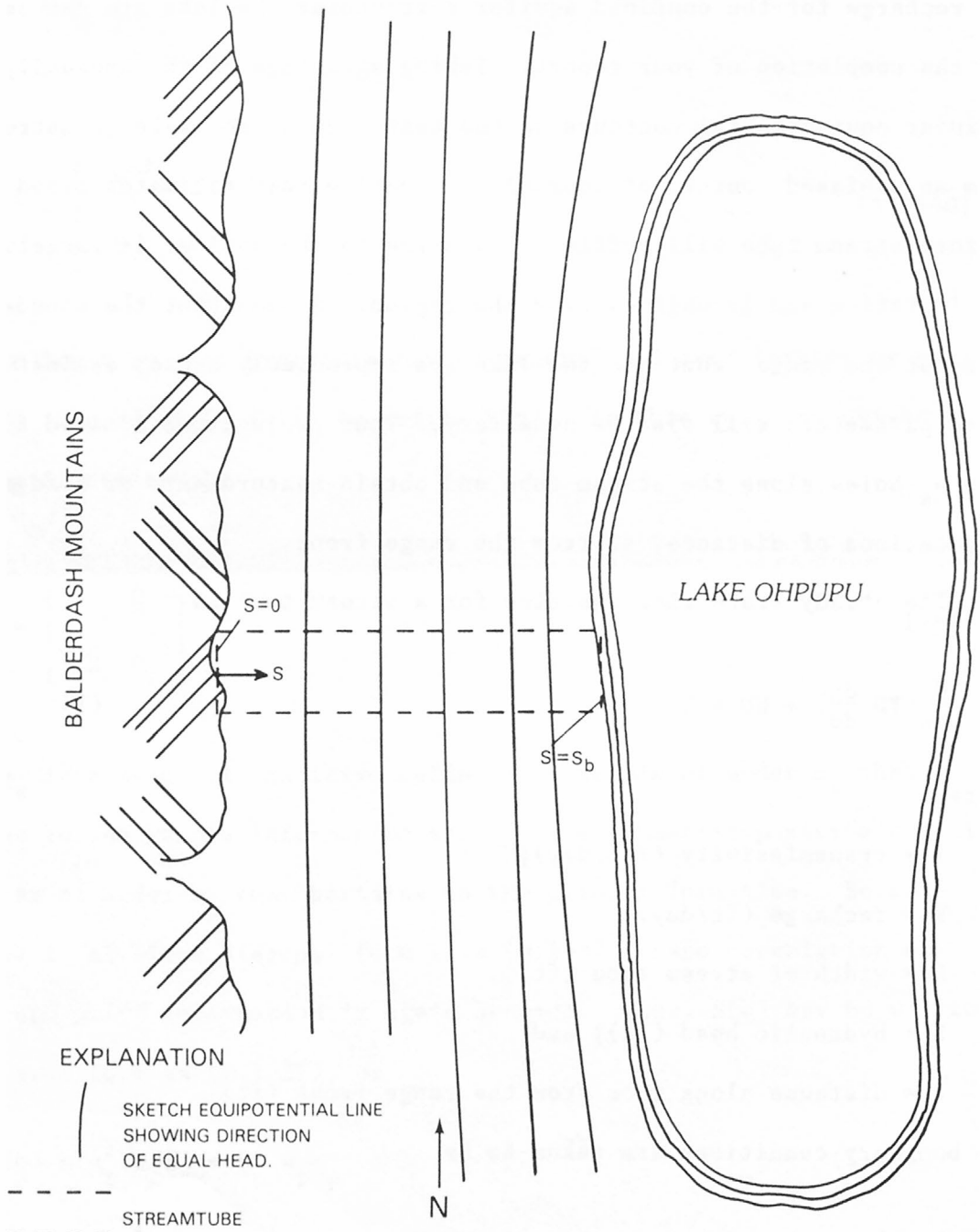


fig. 1

By integrating the flow equation twice with respect to s (assuming T, D, and W to be constant) and using the above boundary conditions, the solution for hydraulic head h along the stream tube is found to be:

$$h = \frac{W}{2T}(s_b - s)s + h_b \frac{s}{s_b} + h_o\left(\frac{s_b - s}{s_b}\right) \tag{3}$$

Let

$$\left.\begin{aligned} \beta_1 &= h_o & \beta_3 &= W/T & X_2 &= \frac{s}{s_b} \\ \beta_2 &= h_b & X_1 &= \frac{s_b - s}{s_b} & X_3 &= \frac{(s_b - s)s}{2} \end{aligned}\right\} \tag{4}$$

and write the above solution (3) to the flow equation using the definitions of β_i and X_i (i = 1, 2, 3). Then write the system of n_s linear regression equations in the three unknown parameters using matrix notation and indicate the contents of each matrix. Identify dependent variable(s), independent variable(s), sensitivities, and parameters.

Let b_i be an estimate of β_i, and:

a) Assume that $Var(\underline{\varepsilon}) = \underline{I}\sigma^2$. Write $S(\underline{b})$ using the matrix form of the model equation (3.1-21) with the estimated parameters b_i. Write a few terms of $S(\underline{b})$ using algebraic notation.

b) Assume that a unique estimate of the variance of the error associated with every head observation is available and that these errors are uncorrelated. Indicate the contents of the resulting weight matrix $\underline{\omega} = \underline{V}^{-1}$. Write $S(\underline{b})$ using the matrix form of the model equation with parameters b_i. Write a few terms of $S(\underline{b})$ using weights ω_{jj} and algebraic notation.

c) Assume case a above, except that there is a prior estimate of h_b having a standard deviation of σ_{h_b}. Indicate the contents of the resulting weight matrix $\underline{\omega} = \underline{V}^{-1}$. Write $S(\underline{b})$ using the matrix form of the model equation with parameters b_i. Write a few terms of $S(\underline{b})$ using algebraic notation. Include the term involving the prior information.

3.2. REGRESSION WHEN THE MODEL IS LINEAR

3.2.1. Derivation of Solution

The linear model assumed is

$$Y = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p + \varepsilon \qquad (3.2\text{-}1)$$

where the X_i are not functions of the parameters. If n observations are used, then an equation of the form of (3.2-1) is written for each observation, so that the system can be written in matrix form as

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon} \qquad (3.2\text{-}2)$$

To find estimates of $\underline{\beta}$ and $\underline{\varepsilon}$, the weighted error sum of squares $S(\underline{b})$,

$$\begin{aligned} S(\underline{b}) &= \underline{e}^T\underline{\omega}\,\underline{e} \\ &= (\underline{Y} - \underline{X}\,\underline{b})^T\underline{\omega}(\underline{Y} - \underline{X}\,\underline{b}) \end{aligned} \qquad (3.2\text{-}3)$$

is minimized with respect to $\underline{b}$.

Before proceeding with the minimization process, some results of matrix calculus are needed. Define the derivative with respect to a vector as the operator vector

$$\frac{\partial}{\partial \underline{b}} = \begin{bmatrix} \frac{\partial}{\partial b_1} \\ \frac{\partial}{\partial b_2} \\ \vdots \\ \frac{\partial}{\partial b_p} \end{bmatrix} \tag{3.2-4}$$

which is applied to a vector or matrix using the rules of matrix multiplication followed by evaluation of the indicated derivatives. Thus, for any two n vectors $\underline{u}$ and $\underline{v}$ that are functions of $\underline{b}$

$$\frac{\partial}{\partial \underline{b}} (\underline{u}^T \underline{v}) = (\frac{\partial}{\partial \underline{b}} \underline{u}^T)\underline{v} + (\frac{\partial}{\partial \underline{b}} \underline{v}^T)\underline{u} \tag{3.2-5}$$

Also, for any n vector $\underline{u}$ that is a function of $\underline{b}$ and any (n×m) matrix $\underline{A}$ that is not a function of $\underline{b}$

$$\frac{\partial}{\partial \underline{b}} (\underline{u}^T \underline{A}) = (\frac{\partial}{\partial \underline{b}} \underline{u}^T)\underline{A} \tag{3.2-6}$$

In particular, if $\underline{u} \equiv \underline{b}$, then, because $\frac{\partial}{\partial \underline{b}} \underline{b}^T = \underline{I}$,

$$\frac{\partial}{\partial \underline{b}} (\underline{b}^T \underline{A}) = \underline{A} \tag{3.2-7}$$

Minimization of (3.2-3) with respect to $\underline{b}$ means

$$\frac{\partial}{\partial \underline{b}} [(\underline{Y} - \underline{X}\,\underline{b})^T \underline{\omega} (\underline{Y} - \underline{X}\,\underline{b})]\Big|_{\underline{b}=\hat{\underline{b}}} = \underline{0} \tag{3.2-8}$$

where the symbol $|_{\underline{b}=\hat{\underline{b}}}$ signifies that $\hat{\underline{b}}$ is that set of parameters that causes the derivatives of $S(\underline{b})$ to be zero. To carry out the operations indicated in (3.2-8), the following procedure is employed. Let $\underline{u} = \underline{Y} - \underline{X}\,\underline{b}$ and $\underline{v} = \underline{\omega}(\underline{Y} - \underline{X}\,\underline{b})$. Then, by using (3.2-5)

$$\frac{\partial}{\partial \underline{b}} [(\underline{Y} - \underline{X}\ \underline{b})^T \underline{\omega} (\underline{Y} - \underline{X}\ \underline{b})]$$

$$= [\frac{\partial}{\partial \underline{b}} (\underline{Y} - \underline{X}\ \underline{b})^T] \underline{\omega} (\underline{Y} - \underline{X}\ \underline{b})$$

$$+ [\frac{\partial}{\partial \underline{b}} (\underline{\omega}(\underline{Y} - \underline{X}\ \underline{b}))^T] (\underline{Y} - \underline{X}\ \underline{b})$$

$$= 2[\frac{\partial}{\partial \underline{b}} (\underline{Y} - \underline{X}\ \underline{b}))^T] \underline{\omega} (\underline{Y} - \underline{X}\ \underline{b}) \qquad (3.2\text{-}9)$$

where the fact that $\underline{\omega} = \underline{\omega}^T$ was employed. With the help of (3.2-7), the derivative is evaluated as

$$\frac{\partial}{\partial \underline{b}} (\underline{Y} - \underline{X}\ \underline{b})^T$$

$$= - \frac{\partial}{\partial \underline{b}} \underline{b}^T \underline{X}^T$$

$$= - \underline{X}^T \qquad (3.2\text{-}10)$$

Hence, (3.2-9) becomes

$$\frac{\partial}{\partial \underline{b}} [(\underline{Y} - \underline{X}\ \underline{b})^T \underline{\omega} (\underline{Y} - \underline{X}\ \underline{b})]$$

$$= - 2\underline{X}^T \underline{\omega} (\underline{Y} - \underline{X}\ \underline{b}) \qquad (3.2\text{-}11)$$

or, by using (3.2-8),

$$\underline{X}^T \underline{\omega}\ \underline{X}\ \hat{\underline{b}} = \underline{X}^T \underline{\omega}\ \underline{Y} \qquad (3.2\text{-}12)$$

Equations (3.2-12) are called the normal equations, and parameters $\hat{\underline{b}}$ are called the estimates of $\underline{\beta}$. The estimates are found from

$$\hat{\underline{b}} = (\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \underline{X}^T \underline{\omega}\ \underline{Y} \qquad (3.2\text{-}13)$$

Students not comfortable with the preceding development should read Draper and Smith (1981, p. 5-17, 70-80, 85-87). This material covers fitting a straight line without and with matrix nomenclature, and then extends the results to the general linear situation. Weighted least squares (where $\underline{\omega} \neq \underline{I}$) is covered in Draper and Smith (1981, p. 108-116).

Elements of $\underline{X}$ are often of vastly differing magnitudes. Thus, when working with a calculator or computer, round-off error can cause serious errors to develop when solving (3.2-13). It is often useful to scale (3.2-12) with respect to a matrix $\underline{C}$, which is a diagonal matrix defined as follows: Let $\underline{X}^T \underline{\omega}\ \underline{X} \equiv \underline{A}$. Then $\underline{C} = diag\ \{1/A_{11}^{\frac{1}{2}}, 1/A_{22}^{\frac{1}{2}}, \ldots, 1/A_{pp}^{\frac{1}{2}}\}$, where A_{ii} is a diagonal entry of $\underline{A}$. Thus, (3.2-12) can be transformed to become

$$\underline{C}^T \underline{X}^T \underline{\omega}\ \underline{X}\ \underline{C}\ \underline{C}^{-1} \hat{\underline{b}} = \underline{C}^T \underline{X}^T \underline{\omega}\ \underline{Y} \qquad (3.2\text{-}14)$$

or

$$\underline{S}^T \underline{\omega}\ \underline{S}\ \hat{\underline{a}} = \underline{S}^T \underline{\omega}\ \underline{Y} \qquad (3.2\text{-}15)$$

where

$$\underline{S} = \underline{X}\ \underline{C} \qquad (3.2\text{-}16)$$

$$\hat{\underline{a}} = \underline{C}^{-1} \hat{\underline{b}} \qquad (3.2\text{-}17)$$

The effect of the scaling is to preserve the symmetry of $\underline{X}^T \underline{\omega}\ \underline{X}$ while at the same time to produce a matrix having all diagonal entries equal to unity. Thus, variability from entry to entry of the $\underline{S}^T \underline{\omega}\ \underline{S}$ matrix is usually reduced considerably over that of $\underline{X}^T \underline{\omega}\ \underline{X}$.

It is also sometimes useful to transform (3.2-2) and resulting normal equations (3.2-12) to incremental form. Let $\underline{Y}_b$ be defined by

$$\underline{Y}_b = \underline{X}\ \underline{b} \qquad (3.2\text{-}18)$$

Then, by subtracting (3.2-18) from (3.2-2) there results

$$\underline{Y} - \underline{Y}_b = \underline{X}(\underline{\beta} - \underline{b}) + \underline{\varepsilon} \qquad (3.2\text{-}19)$$

which is an incremental linear model. To obtain the analog to (3.2-12), premultiply (3.2-18) by $\underline{X}^T\underline{\omega}$ and subtract the result from (3.2-12) to obtain

$$\underline{X}^T\underline{\omega}\ \underline{X}(\hat{\underline{b}} - \underline{b}) = \underline{X}^T\underline{\omega}(\underline{Y} - \underline{Y}_b) \qquad (3.2\text{-}20)$$

Equation (3.2-20) can be transformed to obtain a result analogous to (3.2-15):

$$\underline{S}^T\underline{\omega}\ \underline{S}\ \hat{\underline{\delta}} = \underline{S}^T\underline{\omega}(\underline{Y} - \underline{Y}_b) \qquad (3.2\text{-}21)$$

where

$$\hat{\underline{\delta}} = \underline{C}^{-1}(\hat{\underline{b}} - \underline{b}) \qquad (3.2\text{-}22)$$

If $\hat{\underline{b}}$ as calculated initially using (3.2-15) and (3.2-17) is in error because of round-off, then $\hat{\underline{b}}$ can be used to calculate $\underline{Y}_b$, which then can be substituted into (3.2-21) to calculate $\hat{\underline{\delta}}$. By using (3.2-22), a new improved estimate of $\hat{\underline{b}}$ can be obtained. Writing (3.2-2) in incremental form also provides a basis for procedures, involving statistical analysis of the model, that apply for both linear and (with restrictions) nonlinear models. These are discussed later on.

3.2.2. Solution Algorithm

Sequential steps to follow are:

1) Form $\underline{X}^T\underline{\omega}\ \underline{X}$ and $\underline{X}^T\underline{\omega}\ \underline{Y}$.

2) Transform (3.2-12) to (3.2-15).

3) Solve (3.2-15) (or (3.2-21)) for $\hat{\underline{a}}$.

4) Solve (3.2-17) for $\hat{\underline{b}}$.

Problem 3.2-1

a) By using $\underline{\omega}$ from case c, problem 3.1-1, write out the normal equations used to estimate parameters $\underline{\beta}$ explicitly in sum-of-product (algebraic) form. (Use (3.2-12)).

b) By using either data set 1 (table 1) or 2 (table 2), generate the least squares coefficient matrix $(\underline{X}^T \underline{\omega}\ \underline{X})$ and then compute its inverse. Do not round-off any intermediate calculations or the final inverse. To aid in the calculations, table 3 gives the sums of products for the sample information from the two data sets. You must add the prior information to complete the sums of products.

c) Find the vector $\underline{\hat{b}}$. Do not round-off the results.

Table 1. Data Set 1

s_j (ft)	X_{j1}	X_{j2}	X_{j3}	Observed head, Y_j (ft)
50	0.95	0.05	23750	48.33
150	0.85	0.15	63750	45.76
250	0.75	0.25	93750	42.08
350	0.65	0.35	113750	38.34
450	0.55	0.45	123750	35.30
550	0.45	0.55	123750	31.00
650	0.35	0.65	113750	25.85
750	0.25	0.75	93750	21.76
850	0.15	0.85	63750	16.11
950	0.05	0.95	23750	12.48

Assume $\sigma^2 = 0.25\ ft^2$, $s_b = 1000$ ft, and prior information as follows: $h_b = 11$ ft and $\sigma_{h_b} = 1.1$ ft.

Table 2. Data Set 2

s_j (ft)	X_{j1}	X_{j2}	X_{j3}	Observed head, Y_j (ft)
100	0.9	0.1	45000	47.13
200	0.8	0.2	80000	44.14
300	0.7	0.3	105000	39.89
400	0.6	0.4	120000	36.36
500	0.5	0.5	125000	32.48
600	0.4	0.6	120000	29.70
700	0.3	0.7	105000	24.33
800	0.2	0.8	80000	19.10
900	0.1	0.9	45000	14.96

Assume $\sigma^2 = 0.25\ ft^2$, $s_b = 1000$ ft, and prior information as follows: $h_b = 9.5$ ft and $\sigma_{h_b} = 0.95$ ft.

Table 3. Sums of Products of Sample Information

	Data Set 1	Data Set 2
$\sum_j X_{j1} X_{j1}$	3.3250	2.8500
$\sum_j X_{j1} X_{j2}$	1.6750	1.6500
$\sum_j X_{j1} X_{j3}$	418750	412500
$\sum_j X_{j2} X_{j2}$	3.3250	2.8500
$\sum_j X_{j2} X_{j3}$	418750	412500
$\sum_j X_{j3} X_{j3}$	83340625000	83325000000
$\sum_j X_{j1} Y_j$	192.18350	168.2030
$\sum_j X_{j2} Y_j$	124.82650	119.8870
$\sum_j X_{j3} Y_j$	26879687.5	26583550

3.2.3. Singularity and Conditioning

Singularity of the least squares coefficient matrix occurs whenever columns of the sensitivity matrix, $\underline{X}$, are linearly dependent because this causes rows (or columns) of the coefficient matrix $\underline{X}^T \underline{\omega}\ \underline{X}$ to be linearly dependent. Linear dependence in $\underline{X}$ may be stated as

$$\underline{X}\ \underline{c} = \underline{0} \tag{3.2-23}$$

where not all components of the vector $\underline{c}$ of order p are zero. By pre-multiplying (3.2-23) by $\underline{X}^T \underline{\omega}$, there results

$$\underline{X}^T \underline{\omega}\ \underline{X}\ \underline{c} = \underline{0} \tag{3.2-24}$$

which shows that columns of $\underline{X}^T \underline{\omega}\ \underline{X}$ (or rows since $\underline{X}^T \underline{\omega}\ \underline{X}$ is symmetric) are linearly dependent. It should be noted that transformation of $\underline{X}^T \underline{\omega}\ \underline{X}$ to $\underline{S}^T \underline{\omega}\ \underline{S}$ alters only the form of $\underline{c}$.

Near-singularity, also referred to as ill-conditioning, occurs whenever the columns of $\underline{X}$ (or $\underline{S}$) are almost linearly dependent. Often this condition is indicated by a high degree of correlation among two or more parameter estimates. This correlation reflects the redundancy in the problem. As a result of ill-conditioning, computed parameters are likely to be affected greatly by accumulation of round-off error generated by solving the normal equations. Also, computed variances of the parameters, which are proportional to the diagonal elements of $(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}$, will be large.

A common form of ill-conditioning results if a column of $\underline{X}$ approaches zero so that $\underline{c} = [0, 0, \ldots, 1, 0, \ldots, 0]^T$, where the one appears in the row corresponding to the zero column in $\underline{X}$. This condition indicates that the model is insensitive to the parameter corresponding to the zero

column in $\underline{X}$ and that the parameter should be eliminated from the model. The problem is readily detected by examining the $\underline{X}$ matrix. Another readily detected form of ill-conditioning results if two columns of $\underline{X}$ are nearly proportional, or

$$\alpha \underline{X}_i \simeq \underline{X}_j \tag{3.2-25}$$

so that $\underline{c} = [0, 0, \ldots, \alpha, 0, \ldots, -1, 0, \ldots, 0]^T$, where α appears in row i and -1 appears in row j of $\underline{c}$. In this case

$$\underline{X}_i^T \underline{\omega} \underline{X}_j \simeq \alpha \underline{X}_i^T \underline{\omega} \underline{X}_i$$

$$\simeq \frac{1}{\alpha} \underline{X}_j^T \underline{\omega} \underline{X}_j$$

so that

$$\underline{S}_i^T \underline{\omega} \underline{S}_j = \frac{\underline{X}_i^T \underline{\omega} \underline{X}_j}{\sqrt{(\underline{X}_i^T \underline{\omega} \underline{X}_i)(\underline{X}_j^T \underline{\omega} \underline{X}_j)}}$$

$$\simeq \pm 1 \tag{3.2-26}$$

Thus, one need only examine the $\underline{S}^T \underline{\omega} \underline{S}$ matrix for an off-diagonal entry nearly equal to ± 1. This type of linear dependence indicates that parameters b_i and b_j should be combined because the model can be written as

$$\underline{Y} = \underline{X}_1 b_1 + \ldots + \underline{X}_i b_i + \ldots + \underline{X}_j b_j + \ldots + \underline{X}_p b_p$$
$$\simeq \underline{X}_1 b_1 + \ldots + \underline{X}_i (b_i + \alpha b_j) + \ldots + \underline{X}_p b_p$$

$$= \underline{X}_1 b_1 + \ldots + \underline{X}_i b_i^* + \ldots + \underline{X}_{p-1} b_{p-1} \tag{3.2-27}$$

where b_i^* replaces $b_i + \alpha b_j$ and all subsequent variables are shifted by one so that the last variable number is p-1.

An excellent way to detect general ill-conditioned (or completely linearly dependent) problems is to orthogonalize the columns of the scaled sensitivity matrix, $\underline{S}$. If the columns are all linearly independent, then they can all be transformed so as to be orthogonal to one another, that is, such that

$$\underline{Q}^T\underline{Q} = \underline{D} \tag{3.2-28}$$

where $\underline{Q}$ is the nonzero transformation of $\underline{S}$ and $\underline{D}$ is a full-rank diagonal matrix. If linear dependence exists in $\underline{S}$, then (3.2-28) is replaced by a similar diagonal form except that one or more diagonal entries will be zero. The technique is to successively transform columns such that each new column is orthogonal to all of the previously transformed columns. If column dependence exists, then eventually a column will be calculated that exists entirely of very small numbers (theoretically all zeros for a linearly dependent problem). This column, then, is almost (or completely) linearly dependent on one or more of the previous columns.

The transformation procedure is called Gram-Schmidt orthogonalization and takes the following form:

$$\left.\begin{aligned} \underline{Q}_1 &= \underline{S}_1 \\ \underline{Q}_j &= \underline{S}_j - \sum_{i=1}^{j-1} c_{ij}\underline{Q}_i \quad , \; j = 2, 3, \ldots, p \end{aligned}\right\} \tag{3.2-29}$$

where

$$c_{ij} = \frac{\underline{Q}_i^T\underline{S}_j}{\underline{Q}_i^T\underline{Q}_i} \; ,$$

$\underline{Q}_i$ = the transformed vector orthogonal to vectors already in $\underline{Q}$,

$\underline{S}_j$ = the next column vector of $\underline{S}$ to be transformed.

3.3. REGRESSION WHEN THE MODEL IS NONLINEAR

3.3.1. Modified Gauss-Newton Method

If the model is nonlinear in the parameters but is linear in the dependent variable, then the model may be written in the standard form for nonlinear regression:

$$Y = f(\xi_1, \xi_2, \ldots, \xi_k; \beta_1, \beta_2, \ldots \beta_p) + \varepsilon \qquad (3.3\text{-}1)$$

Because of the nonlinearity, f cannot be written in the form $f = X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p$. The case more complicated than (3.3-1) where the model is nonlinear in both the parameters and the dependent variable is treated in section 6.1. When there are n observations, (3.3-1) may be written in matrix form as

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{\beta}) + \underline{\varepsilon} \qquad (3.3\text{-}2)$$

or, in terms of general estimate $\underline{b}$ of $\underline{\beta}$ and estimate $\underline{e}$ of $\underline{\varepsilon}$

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{b}) + \underline{e} \qquad (3.3\text{-}3)$$

As for the linear case, the regression solution of (3.3-3) is obtained by minimizing the weighted error sum of squares:

$$S(\underline{b}) = \underline{e}^T \underline{\omega}\, \underline{e}$$

$$= (\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}))^T \underline{\omega} (\underline{Y} - \underline{f}(\underline{\xi}, \underline{b})) \qquad (3.3\text{-}4)$$

However, because (3.3-3) is nonlinear, solution of the problem is not as direct as it was for the linear case.

A convenient and robust solution method is obtained by linearizing (3.3-3) around an initial estimate of parameters then proceeding as if the problem were linear. This yields a new set of parameters that minimizes (3.3-4) where $\underline{f}$ is replaced by the linear approximation. The new parameters are then substituted for the initial set and the process is repeated to yield a better set of parameters. The iterative process stops whenever the change in calculated parameters from one iteration to the next is small. At that point the minimum of (3.3-4) has been found.

To derive the method, first $\underline{f}(\underline{\xi}, \underline{b})$ is expanded about the initial set of parameters $\underline{b}_0$ by using a truncated Taylor series to obtain a linear approximation for $\underline{f}(\underline{\xi}, \underline{b})$:

$$\underline{f}(\underline{\xi}, \underline{b}) \simeq \underline{f}(\underline{\xi}, \underline{b}_0) + \underline{X}_0(\underline{b} - \underline{b}_0) \tag{3.3-5}$$

where

$$\underline{X}_0 = \{X^0_{ij}\} = \left\{\left.\frac{\partial f_i}{\partial b_j}\right|_{\underline{b}=\underline{b}_0}\right\}_{(n \times p)} \tag{3.3-6}$$

and f_i is f calculated at the i th observation point. The components of $\underline{X}_0$ are recognized as sensitivity coefficients, or simply sensitivities.

By using (3.3-5), (3.3-3) may be written as

$$\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_0) \simeq \underline{X}_0(\underline{b} - \underline{b}_0) + \underline{e} \tag{3.3-7}$$

which is of the form of the incremental linear model of section 3.2. It is important to note that if the model is linear so that $\underline{f}(\underline{\xi}, \underline{b}) = \underline{X}(\underline{\xi})\underline{b}$, then the truncated Taylor series and, hence, (3.3-7) are exact. In this case expansion in a Taylor series is another way of deriving the incremental linear model. If the model is nonlinear, equation (3.3-7) is the approximate (linearized) model for parameters in the vicinity of $\underline{b}_0$, as illustrated by a simple one-parameter example in figure 3.3-1.

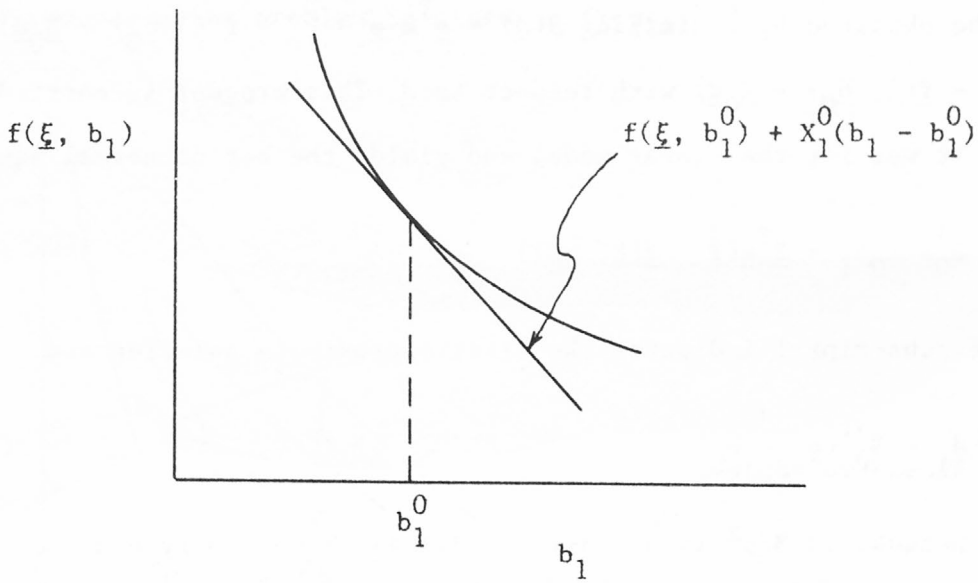


fig. 3.3-1

An operation that often reduces the number of arithmetic operations and reduces round-off error in nonlinear regression is to scale (3.3-7) using $\underline{b}_0$:

$$\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_0) \simeq \underline{Z}_0 \, \underline{d} + \underline{e} \qquad (3.3\text{-}8)$$

where

$$\underline{Z}_0 = \underline{X}_0 \underline{B}_0 \tag{3.3-9}$$

$$\underline{d} = \underline{B}_0^{-1}(\underline{b} - \underline{b}_0) \tag{3.3-10}$$

$$\underline{B}_0 = \mathit{diag}\ \{b_1^0, b_2^0, \ldots, b_p^0\} \tag{3.3-11}$$

and b_i^0 is a component of $\underline{b}_0$. Note that this operation does not change $\underline{f}(\underline{\xi}, \underline{b}_0)$. Thus, $S(\underline{b})$ and, hence, the position of the minimum are unaffected by the transformation.

An approximate best estimate of $\underline{\beta}$ (which is exact for a linear model) can be obtained by minimizing $S(\underline{d}) = \underline{e}^T \underline{\omega}\ \underline{e} = (\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_0) - \underline{Z}_0\underline{d})^T \cdot \underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_0) - \underline{Z}_0\underline{d})$ with respect to $\underline{d}$. This process is carried out exactly like it was for the linear model and yields the set of normal equations

$$\underline{Z}_0^T\underline{\omega}\ \underline{Z}_0\underline{d}_1 = \underline{Z}_0^T\underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_0)) \tag{3.3-12}$$

where subscript 1 indicates the first approximate solution and

$$\underline{d}_1 = \underline{B}_0^{-1}(\underline{b}_1 - \underline{b}_0) \tag{3.3-13}$$

Because (3.3-8) is not exact, (3.3-4) is not truly minimized, and $\underline{b}_1$ is not actually the optimal set. Hence, $\underline{b}_1$ is substituted for $\underline{b}_0$, and the entire process is repeated to yield another, hopefully improved, estimate. As a general iteration equation, (3.3-12) may be written in the form

$$\underline{Z}_r^T\underline{\omega}\ \underline{Z}_r\ \underline{d}_{r+1} = \underline{Z}_r^T\underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_r)) \tag{3.3-14}$$

where

$$\underline{Z}_r = \underline{X}_r\ \underline{B}_r \tag{3.3-15}$$

$$\underline{d}_{r+1} = \underline{B}_r^{-1}(\underline{b}_{r+1} - \underline{b}_r) \tag{3.3-16}$$

$$\underline{B}_r = diag\left\{b_1^r,\ b_2^r,\ \ldots,\ b_p^r\right\} \tag{3.3-17}$$

and b_i^r is a component of $\underline{b}_r$. As the process converges, $\underline{d}_{r+1} \to \underline{0}$ and (3.3-8) becomes $\underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}}) = \hat{\underline{e}}$. At the same point, $\hat{\underline{b}}$ minimizes $S(\underline{b})$ in (3.3-4), or $S(\hat{\underline{b}}) = (\underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}}))^T \underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}}))$, which is a minimum for the nonlinear equation. This process for finding the minimum of $S(\underline{b})$ is known as the Gauss-Newton method.

A sketch of progression of the iterations to the minimum for a hypothetical two-parameter problem is given in figure 3.3-2.

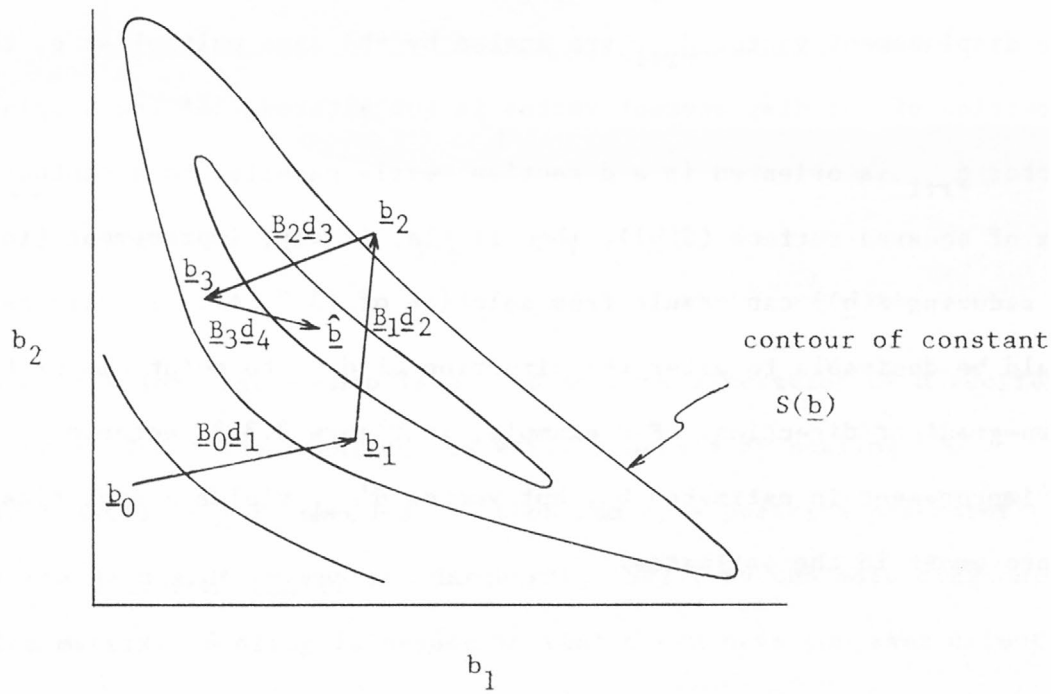


fig. 3.3-2

As diagrammed in figure 3.3-2, the solution often does not progress directly toward the minimum.

Modifications to the Basic Procedure. It is well-known that the Gauss-Newton method, as defined by (3.3-14), does not always converge. To help induce convergence, a damping parameter, ρ, is introduced by modifying (3.3-16) to give

$$\underline{b}_{r+1} = \rho \underline{B}_r \underline{d}_{r+1} + \underline{b}_r \qquad (3.3\text{-}18)$$

If $0<\rho<1$, the changes in computed parameters are less than would result for $\rho = 1$; thus the method is an interpolation method. Similarly, if $\rho>1$, the method is an extrapolation method.

By inspection of (3.3-18) it can be seen that ρ changes the magnitude of the displacement from $\underline{b}_r$ to $\underline{b}_{r+1}$. However, because all components of the displacement vector $\underline{d}_{r+1}$ are scaled by the same multiplier ρ, the direction of the displacement vector is not altered. If the displacement vector $\underline{d}_{r+1}$ is oriented in a direction nearly parallel to a contour in the sum of squares surface ($S(\underline{b})$), then little, if any, improvement (in terms of reducing $S(\underline{b})$) can result from solution of (3.3-14). In this case it would be desirable to alter the direction of $\underline{d}_{r+1}$ to point closer to a down-gradient direction. For example, in figure 3.3-3 vector $\underline{d}_{r+1}$ yields no improvement in estimates $\underline{b}_r$, but vector $\underline{d}'_{r+1}$ yields a significant improvement in the estimates.

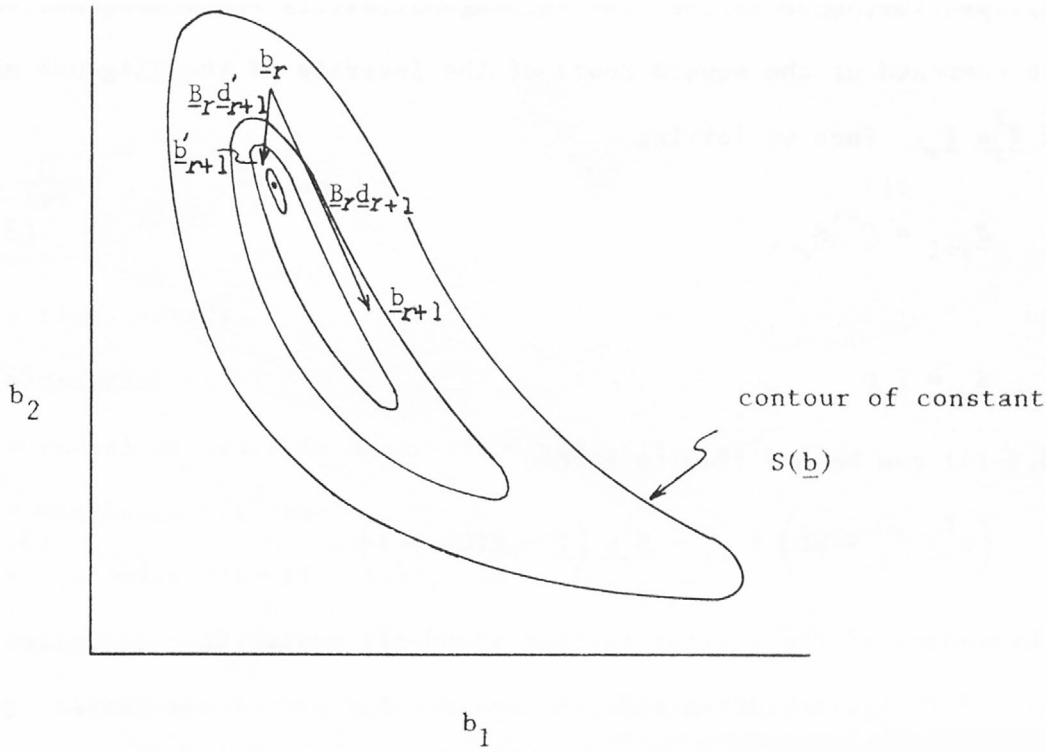


fig. 3.3-3

A modification that accomplishes the desired alteration of direction of $\underline{d}_{r+1}$, and reduces its magnitude as well, consists of scaling the coefficient matrix $\underline{Z}_r^T \underline{\omega}\ \underline{Z}_r$ of (3.3-14) then adding a positive parameter μ, known as the Marquardt parameter (Marquardt, 1963), to the main diagonal of the scaled matrix. Scaling is needed so that μ can have the same effect on each entry of the main diagonal of the coefficient matrix. The scaling process is accomplished in the same manner as for (3.2-14), in other words such that each entry of the main diagonal of the scaled matrix is unity but symmetry of the original matrix is preserved.

Mathematically, the Marquardt modification can be developed as follows. Let $\underline{C}$ be defined as the diagonal matrix whose diagonal elements are composed of the square roots of the inverses of the diagonal elements of $\underline{Z}_r^T \underline{\omega}\ \underline{Z}_r$. Then by letting

$$\underline{\delta}_{r+1} = \underline{C}^{-1} \underline{d}_{r+1} \tag{3.3-19}$$

and

$$\underline{S}_r = \underline{Z}_r \underline{C} \tag{3.3-20}$$

(3.3-14) can be modified to become

$$\left(\underline{S}_r^T \underline{\omega}\ \underline{S}_r + \mu \underline{I}\right) \underline{\delta}_{r+1} = \underline{S}_r^T \underline{\omega} \left(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_r)\right) \tag{3.3-21}$$

A byproduct of the scaling is that round-off errors in calculations to solve for $\underline{\delta}_{r+1}$ are often reduced considerably over those resulting from making the calculations without scaling.

Solution algorithm. The sequential steps implementing the modified Gauss-Newton procedure are:

1) Calculate $\underline{f}(\underline{\xi}, \underline{b}_r)$ and $\underline{S}_r$ using initial parameters $\underline{b}_r$ and the combination of (3.3-6) (with index r replacing 0), (3.3-15), and (3.3-20).
2) Solve (3.3-21) for $\underline{\delta}_{r+1}$.
3) Solve (3.3-19) for $\underline{d}_{r+1}$.
4) Solve (3.3-18) for $\underline{b}_{r+1}$.
5) Test to determine if $|d_i^{r+1}| > \varepsilon$, where ε is a small number such as 0.01.
6) If $|d_i^{r+1}| > \varepsilon$, increment r by one and return to 1. If not, then the process has converged.

Problem 3.3-1

The Theis equation for flow to a well in a confined, nonleaky aquifer is

$$s = \frac{Q}{4\pi T} \int_{r^2S/4Tt}^{\infty} \frac{e^{-z}}{z}\, dz \qquad (1)$$

where t = time (days);

s = drawdown (ft);

r = radial distance to observation well (ft);

Q = discharge (ft^3/sec);

T = transmissivity (ft^2/sec); and

S = storage coefficient.

The integral can be evaluated by summing the infinite series

$$\int_u^{\infty} \frac{e^{-z}}{z}\, dz = -0.577216 - \ln u + u - \frac{u^2}{2\cdot 2!} + \frac{u^3}{3\cdot 3!} - \frac{u^4}{4\cdot 4!} + \ldots \qquad (2)$$

where $n! = n\cdot(n-1)\cdot(n-2) \ldots 3\cdot 2\cdot 1$.

The Theis equation is nonlinear in the parameters T and S. Using the information in section 3.3.1, develop an algorithm for solving this equation for T and S, given time and drawdown data.

a) Let

$$f(t,r;\ T,S) = \frac{Q}{4\pi T} W(u) \qquad (3)$$

where $u = r^2S/4Tt$ and W(u) is the integral in equation (2).

Find the sensitivities for T and S (see equation (3.3-6).

Hint:

$$\frac{d}{d\alpha}\int_{u(\alpha)}^{\infty} f(x)dx = - f(u(\alpha))\ \frac{du(\alpha)}{d\alpha}$$

b) Assume that initial estimates of transmissivity and storage coefficient, T_0 and S_0, exist. Form the diagonal matrix (3.3-17) and the scaled sensitivities (3.3-15).

c) Construct, for n time-observations at a single spatial location, the incremental linear model (see (3.3-8)). Note that solution of (3.3-14) yields scaled displacements $\underline{d}_{r+1}$. Indicate, from (3.3-16), how the unscaled increments can be recovered.

Figure 1 is a flow diagram for programming the steps indicated in section 3.3.1. Obtain a coding sheet and proceed to write code according to this flow diagram. Helpful hints, numbered on the flow diagram, follow:

(1) n = number of observations;

r = radial distance to observation well;

t_i = time of each observation;

s_i = observed drawdown for each time t_i;

T_0 = initial guess, transmissivity;

S_0 = initial guess, storage coefficient;

Q = discharge;

ρ = damping parameter;

ε = convergence criterion;

r_{mx} = maximum number of iterations.

FLOW CHART

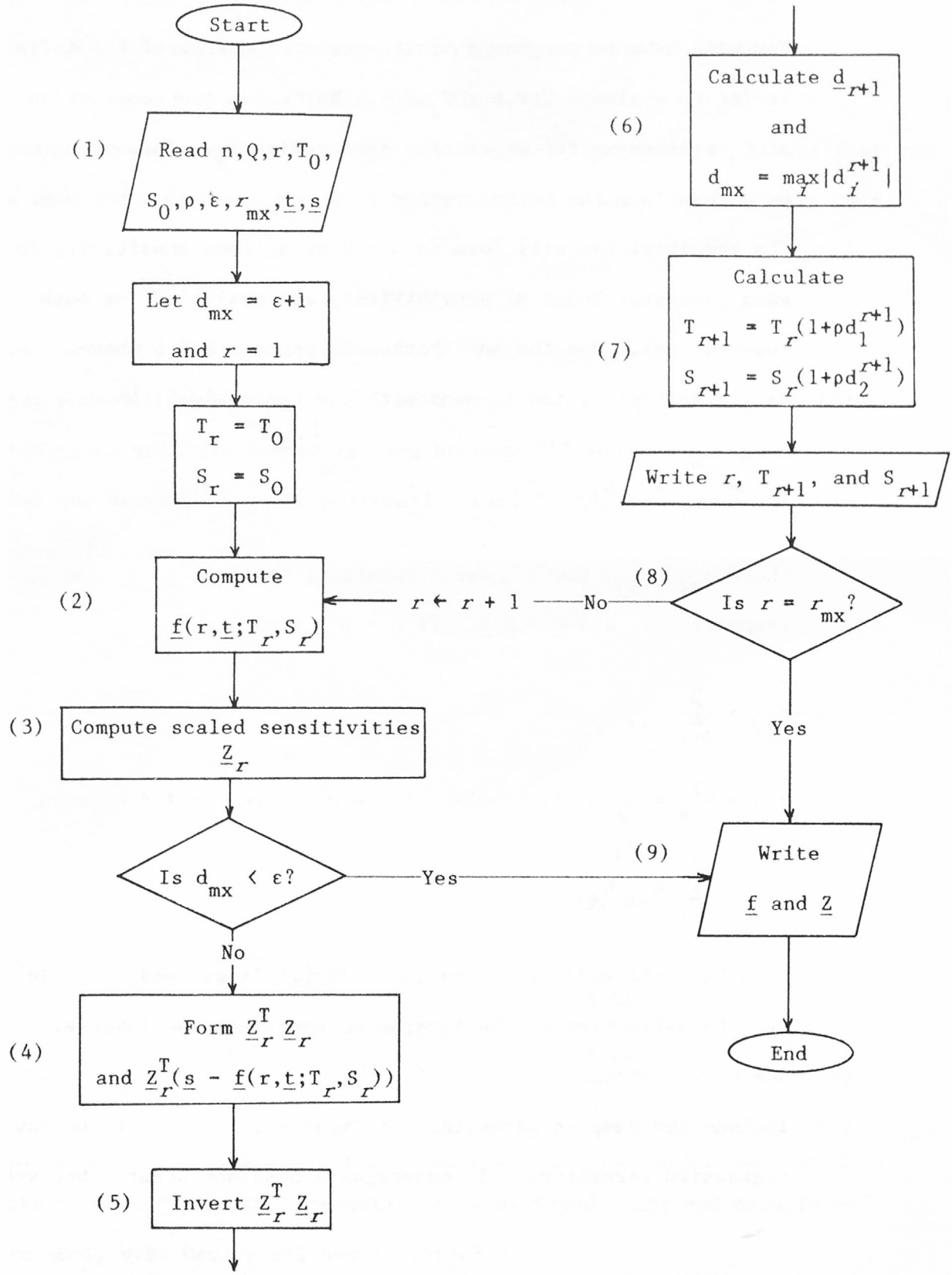


fig. 1

(2) $f_i(r,t_i;\ T_r,S_r) = \frac{Q}{4\pi T_r} W(r,t_i;\ T_r,S_r)$

You will have to program a finite number of terms of the infinite series to evaluate $W(r,t_i;\ T_r,S_r)$. Note that one computation of f_i will occur for every time observation t_i. These computations form the column vector $\underline{f}$.

(3) The sensitivities will form an n × 2 array (one sensitivity for each parameter T and S; sensitivities are evaluated for each observation). Use the information in parts a and b above.

(4) Matrix multiplication is done with the basic algorithm

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

where a_{ik}, b_{kj}, and c_{ij} are elements of matrices $\underline{A}$, $\underline{B}$, and $\underline{C}$, respectively, and $\underline{C} = \underline{A}\ \underline{B}$. If $\underline{A} = \underline{B}^T$, then

$$c_{ij} = \sum_{k=1}^{n} b_{ki} b_{kj}$$

since $b^T_{ik} = b_{ki}$. Similarly, if $\underline{A} = \underline{D}^T$, then $\underline{C} = \underline{A}\ \underline{B}$ becomes

$$c_{ij} = \sum_{k=1}^{n} d_{ki} b_{kj}$$

Note that the model error variance $Var(\underline{\varepsilon})$ is assumed to be $\underline{I}\sigma^2$.

(5) Use the definition of the inverse to construct the inverse.

(6) See part c above.

(7) Include the damping parameter ρ in your calculation of the new regression parameters. If convergence does not occur, then you

may be best advised to set ρ to a value less than one.

(8) This is the end of the iteration loop.

(9) You may wish to compute the final estimate of residuals $\underline{e} = \underline{f} - \underline{s}$ then print them also.

Because convergence problems may arise, limit the number of passes that can be made through the algorithm. Test your code using the following data.

36 HOUR PUMPING TEST

Test starts: 12 February 1976, 0805 hrs.

Test ends: 13 February 1976, 2005 hrs.

Production well was pumped varying between 517 gpm and 530 gpm.

Time and drawdown data at observation well 175 ft from production well appear in table 1:

Table 1.

Time (sec)	Drawdown (ft)
480	1.71
1020	2.23
1500	2.54
2040	2.77
2700	3.04
3720	3.25
4920	3.56

Data from S. P. Larson, presently of S. S. Papadopulos and Associates (formerly U.S. Geological Survey), 1978.

To aid in debugging your computer program, some example calculations follow. Assume that $T_0 = 0.1$ ft^2/sec, $S_0 = 0.0005$, $Q = 1.16$ ft^3/sec, and $r_i = 175$ ft.

$$u_i^0 = \frac{(r_i)^2 S_0}{4T_0 t_i} = \frac{(175)^2(.0005)}{4(.1)t_i} = 38.28125/t_i$$

$$f_i^0 = \frac{Q}{4\pi T_0} W(u_i^0) = \frac{1.16}{4\pi(.1)} W(u_i^0) = 0.923099\ W(u_i^0)$$

t_i	u_i^0	$W(u_i^0)$	f_i^0
480	0.0797526	2.02980	1.87371
1020	0.0375306	2.74256	2.53165
1500	0.0255208	3.11640	2.87675
2040	0.0187653	3.41721	3.15442
2700	0.0141782	3.69296	3.40897
3720	0.0102907	4.00957	3.70123
4920	0.00778074	4.28665	3.95700

$$Z_{iT}^0 = -f_i^0 + \frac{Q}{4\pi T_0} e^{-u_i^0} = -f_i^0 + 0.923099\ e^{-u_i^0}$$

$$Z_{iS}^0 = -\frac{Q}{4\pi T_0} e^{-u_i^0} = -0.923099\ e^{-u_i^0}$$

Note that $Z_{iT}^0 = X_{iT}^0 T_0$ and $Z_{iS}^0 = X_{iS}^0 S_0$ are calculated directly, without first calculating sensitivities X_{iT}^0 and X_{iS}^0 then formally making the multiplication. These scaled sensitivities result from the following scaled linearized model.

$$f_i^1 = f_i^0 + X_{iT}^0 \Delta T + X_{iS}^0 \Delta S$$

$$= f_i^0 + X_{iT}^0 T_0 \left(\frac{\Delta T}{T_0}\right) + X_{iS}^0 S_0 \left(\frac{\Delta S}{S_0}\right)$$

$$= f_i^0 + Z_{iT}^0 \frac{\Delta T}{T_0} + Z_{iS}^0 \frac{\Delta S}{S_0}$$

Z_{iT}^0	Z_{iS}^0
-1.02137	-0.852339
-1.64255	-0.889097
-1.97691	-0.899839
-2.24848	-0.905938
-2.49887	-0.910103
-2.78758	-0.913648
-3.04106	-0.915944

$$\sum \left(Z_{iT}^0\right)^2 = 35.96800209 \qquad \sum \left(Z_{iS}^0\right)^2 = 5.649402682$$

$$\sum Z_{iT}^0 Z_{iS}^0 = 13.75336059$$

z^0_{iT}	s_i	f^0_i	$z^0_{iT}(s_i-f^0_i)$
-1.02137	1.71	1.87371	0.1672084827
-1.64255	2.23	2.53165	0.4954752075
-1.97691	2.54	2.87675	0.6657244425
-2.24848	2.77	3.15442	0.8643606816
-2.49887	3.04	3.40897	0.9220080639
-2.78758	3.25	3.70123	1.257839723
-3.04106	3.56	3.95700	1.20730082

$\sum z^0_{iT}\left(s_i-f^0_i\right) = 5.579917421$

z^0_{iS}	s_i	f^0_i	$z^0_{iS}(s_i-f^0_i)$
-0.852339	1.71	1.87371	0.1395364177
-0.889097	2.23	2.53165	0.2681961101
-0.899839	2.54	2.87675	0.3030207833
-0.905938	2.77	3.15442	0.3482606860
-0.910103	3.04	3.40897	0.3358007039
-0.913648	3.25	3.70123	0.4122653870
-0.915944	3.56	3.95700	0.3636297680

$\sum z^0_{iS}\left(s_i-f^0_i\right) = 2.170709856$

$$\frac{\Delta S}{S_0} = \frac{2.170709856 - (13.75336059)(5.579917421)/35.96800209}{5.649402682 - (13.75336059)^2/35.96800209}$$

$$= \frac{.03707407158}{.3904247981} = 0.09495829097$$

$$\frac{\Delta T}{T_0} = (5.579917421 - (.09495829097)(13.75336059))/35.96800209$$

$$= 0.1188256660$$

$$S_1 = (1 + .0949583)(.0005) = 0.000547479$$

$$T_1^1 = (1 + .118826)(.1) = 0.111883$$

$$u_i^1 = \frac{(175)^2(.000547479)}{4(.111883)t_i} = 37.4645/t_i$$

$$f_i^1 = \frac{1.16}{4\pi(.111883)} W(u_i^1) = 0.825057\ W(u_i^1)$$

t_i	u_i^1	$W(u_i^1)$	f_i^1
480	0.0780510	2.04973	1.69114
1020	0.0367299	2.76334	2.27991
1500	0.0249763	3.13743	2.58856
2040	0.0183650	3.43838	2.83686
2700	0.0138757	3.71423	3.06445
3720	0.0100711	4.03091	3.32573
4920	0.00761474	4.30805	3.55439

$$Z^1_{iT} = - f^1_i + 0.825057\ e^{-u^1_i}$$

$$Z^1_{iS} = - 0.825057\ e^{-u^1_i}$$

Z^1_{iT}	Z^1_{iS}
-0.928031	-0.763109
-1.48461	-0.795303
-1.78385	-0.804705
-2.02682	-0.810043
-2.25076	-0.813688
-2.50894	-0.816789
-2.73559	-0.818798

3.3.2. Nonlinear Regression When the Model is Numerical

The basic model equations assumed in all previous developments have been of the closed form or analytical type where the dependent variable $\underline{f}$ is a known function of $\underline{\xi}$ and $\underline{\beta}$. In many cases such models may either not exist or be too complicated for practical use. In these cases the basic equation relating the dependent variable to the independent variables and parameters may be a numerical solution that can be stated in the general form

$$\underline{D}(\underline{h},\underline{\xi},\underline{\beta})\underline{h} = \underline{q}(\underline{h},\underline{\xi},\underline{\beta}) \qquad (3.3\text{-}22)$$

Equation (3.3-22) is a nonlinear matrix equation in which $\underline{h}$ is the solution (dependent variable) vector of order m; $\underline{D}$ is a nonsingular coefficient matrix of order m that is a function of $\underline{h}$, $\underline{\xi}$, and $\underline{\beta}$; and $\underline{q}$ is a vector of

order m that is a function of $\underline{h}$, $\underline{\xi}$, and $\underline{\beta}$. Order m is not related to the number of observations n, but instead is simply the order required to give a good numerical approximation to the solution of the problem.

If (3.3-22) is linear in $\underline{h}$ so that $\underline{D}$ and $\underline{q}$ are not functions of $\underline{h}$, then (3.3-22) may be solved directly for $\underline{h}$. In this case the Gauss-Newton method may be used to obtain the regression solution. (The nonlinear case is considered in section 6.1.2.) The procedure is as follows. First write (3.3-22) in the form

$$\underline{h} = \underline{D}^{-1}\underline{q} \tag{3.3-23}$$

which is explicit in the dependent variable $\underline{h}$. Next, note that $\underline{h}$ in (3.3-23) and $\underline{f}$ in (3.3-2) (or, as an estimate, (3.3-3)) are expressions of the same quantity, the only difference between them being that elements of $\underline{h}$ are values of the dependent variable computed at points defined by the numerical solution and elements of $\underline{f}$ are values of the dependent variable computed at observation points. If all n observation points are contained in the set of points required for the numerical solution, which implies that $m \geq n$, then $\underline{f}$ is obtained from $\underline{h}$ simply by eliminating those entries in $\underline{h}$ not corresponding to observation points. In other instances the points in m may not correspond in any way to those in n. For these instances an interpolation scheme would be used to obtain $\underline{f}$ from $\underline{h}$. In either case, the vector $\underline{f}(\underline{\xi}, \underline{b}_r)$ is obtained by using $\underline{h}$ computed from (3.3-23) in which $\underline{b}_r$ was used to evaluate $\underline{D}$ and $\underline{q}$.

The final step in forming the Gauss-Newton solution is to derive the sensitivity matrix $\underline{X}$. To accomplish this step, write (3.3-22) in terms of a general parameter set $\underline{b}$, then differentiate it with respect to $\underline{b}$ to yield

$$\underline{D}\,\frac{\partial \underline{h}}{\partial b_j} + \frac{\partial \underline{D}}{\partial b_j}\,\underline{h} = \frac{\partial \underline{q}}{\partial b_j}\,, \quad j=1,\ 2,\ \ldots,\ p \tag{3.3-24}$$

or

$$\frac{\partial \underline{h}}{\partial b_j} = \underline{D}^{-1}\left(\frac{\partial \underline{q}}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j}\,\underline{h}\right)\,, \quad j=1,\ 2,\ \ldots,\ p \tag{3.3-25}$$

The quantity $\partial \underline{h}/\partial b_j$ forms a column of the sensitivity matrix for points in m. Sensitivity matrix $\underline{X}_r$ would be found by first computing $\partial \underline{h}/\partial b_j$ using (3.3-25) written in terms of $\underline{h}_r$ and $\underline{b}_r$, then following one of the two procedures described above for obtaining $\underline{f}$ from $\underline{h}$. By incorporating the procedures to compute numerical estimates of $\underline{f}(\underline{\xi},\underline{b}_r)$ and $\underline{X}_r$, the algorithm given for the Gauss-Newton method may be followed exactly to find the regression solution of (3.3-23).

By studying the sequence of calculations in the solution algorithm carefully, it will be noted that in order to calculate $\underline{h}_r$ and $\underline{X}_r$, $\underline{D}_r$ ($\underline{D}$ computed using $\underline{b}_r$) must already have been computed. This requires one of two possible calculation schemes. Either $\underline{D}_r$ and p matrices of the form $(\partial \underline{D}/\partial b_j)_r$ must be formed at the same time and stored before $\underline{h}_r$ and $\underline{X}_r$ are calculated, or $\underline{D}_r$ must be formed before $\underline{h}_r$ is calculated and then each matrix $(\partial \underline{D}/\partial b_j)_r$ formed as needed to calculate each column (that is $\underline{X}_j^r$) of $\underline{X}_r$. The first alternative could require a considerable amount of computer memory or the use of scratch files, whereas the second alternative could require repetitive calculation because many arithmetic operations could be the same for forming both $\underline{D}_r$ and $(\partial \underline{D}/\partial b_j)_r$. Often, however, matrices $(\partial \underline{D}/\partial b_j)_r$ can be written in a condensed or decomposed form to conserve computer memory. In this way $\underline{D}_r$ and the decomposed form of

$(\partial \underline{D}/\partial b_j)_r$ can be computed together without using a significant amount of extra memory. Then each matrix $(\partial \underline{D}/\partial b_j)_r$ may be assembled as needed without performing numerous repetitive calculations.

A significant amount of computer memory can also be wasted unless care is taken when forming $\underline{X}_r$. The general procedure is to form the column vector $(\partial \underline{q}/\partial b_j)_r - (\partial \underline{D}/\partial b_j)_r \underline{h}_r$, then use (3.3-25) to form the vector $(\partial \underline{h}/\partial b_j)_r$. From this, vector $\underline{X}_j^r$ is immediately formed. The matrix composed of vectors $(\partial \underline{h}/\partial b_j)_r$ should not be stored in central computer memory because it is often large. If desired it may be stored column by column on a scratch file for later retrieval and printing.

Problem 3.3-2

Assume the finite difference representation of a flow problem shown in figure 1.

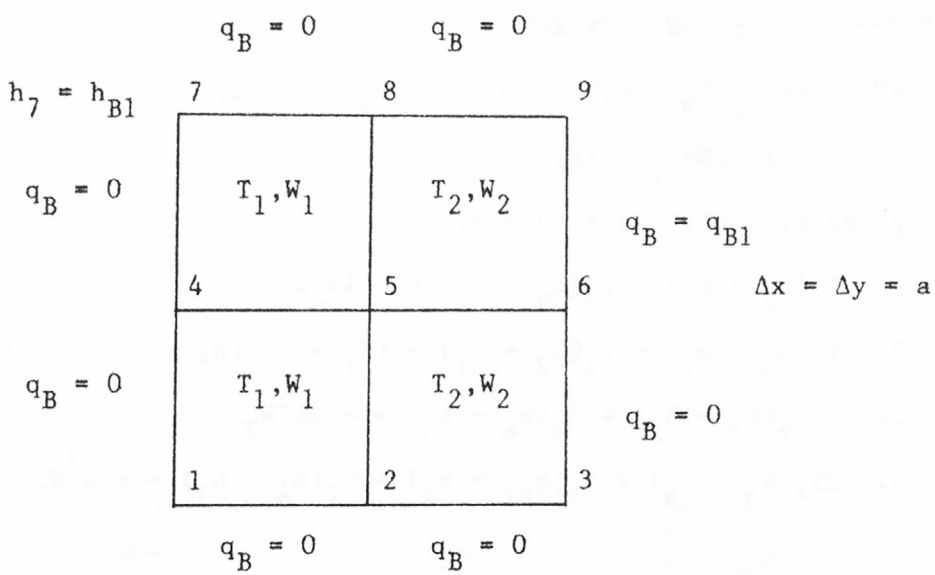


fig. 1

Finite difference equations for this problem can be written as follows:

1. $\frac{1}{2}\Delta y T_1(h_2 - h_1)/\Delta x + \frac{1}{2}\Delta x T_1(h_4 - h_1)/\Delta y = -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1$

2. $\frac{1}{2}\Delta y T_2(h_3 - h_2)/\Delta x - \frac{1}{2}\Delta y T_1(h_2 - h_1)/\Delta x + \frac{1}{2}\Delta x T_1(h_5 - h_2)/\Delta y$
$+ \frac{1}{2}\Delta x T_2(h_5 - h_2)/\Delta y = -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1 - \frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2$

3. $-\frac{1}{2}\Delta y T_2(h_3 - h_2)/\Delta x + \frac{1}{2}\Delta x T_2(h_6 - h_3)/\Delta y = -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2$

4. $\Delta y T_1(h_5 - h_4)/\Delta x + \frac{1}{2}\Delta x T_1(h_{B1} - h_4)/\Delta y - \frac{1}{2}\Delta x T_1(h_4 - h_1)/\Delta y$
$= -\frac{1}{2}\Delta x\Delta y W_1$

5. $\Delta y T_2(h_6 - h_5)/\Delta x - \Delta y T_1(h_5 - h_4)/\Delta x$
$+ \frac{1}{2}\Delta x T_1(h_8 - h_5)/\Delta y + \frac{1}{2}\Delta x T_2(h_8 - h_5)/\Delta y$
$- \frac{1}{2}\Delta x T_1(h_5 - h_2)/\Delta y - \frac{1}{2}\Delta x T_2(h_5 - h_2)/\Delta y$
$= -\frac{1}{2}\Delta x\Delta y W_1 - \frac{1}{2}\Delta x\Delta y W_2$

6. $-\Delta y T_2(h_6 - h_5)/\Delta x + \frac{1}{2}\Delta x T_2(h_9 - h_6)/\Delta y - \frac{1}{2}\Delta x T_2(h_6 - h_3)/\Delta y$
$= -\frac{1}{2}\Delta x\Delta y W_2 - \frac{1}{2}\Delta y q_{B1}$

7. $h_7 = h_{B1}$

8. $\frac{1}{2}\Delta y T_2(h_9 - h_8)/\Delta x - \frac{1}{2}\Delta y T_1(h_8 - h_{B1})/\Delta x$
$- \frac{1}{2}\Delta x T_1(h_8 - h_5)/\Delta y - \frac{1}{2}\Delta x T_2(h_8 - h_5)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_1 - \frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2$

9. $-\frac{1}{2}\Delta y T_2(h_9 - h_8)/\Delta x - \frac{1}{2}\Delta x T_2(h_9 - h_6)/\Delta y$
$= -\frac{1}{2}\Delta x\frac{1}{2}\Delta y W_2 - \frac{1}{2}\Delta y q_{B1}$

or, by assuming that $\Delta x = \Delta y = a$,

1. $T_1(h_2 - h_1) + T_1(h_4 - h_1) = -\frac{1}{2}a^2 W_1$

2. $T_2(h_3 - h_2) - T_1(h_2 - h_1) + (T_1 + T_2)(h_5 - h_2) = -\frac{1}{2}a^2(W_1 + W_2)$

3. $-T_2(h_3 - h_2) + T_2(h_6 - h_3) = -\frac{1}{2}a^2 W_2$

4. $2T_1(h_5 - h_4) + T_1(h_{B1} - h_4) - T_1(h_4 - h_1) = -a^2 W_1$

5. $2T_2(h_6 - h_5) - 2T_1(h_5 - h_4) + (T_1 + T_2)(h_8 - h_5)$
$- (T_1 + T_2)(h_5 - h_2) = - a^2(W_1 + W_2)$

6. $- 2T_2(h_6 - h_5) + T_2(h_9 - h_6) - T_2(h_6 - h_3) = - a^2W_2 - aq_{B1}$

7. $h_7 = h_{B1}$

8. $T_2(h_9 - h_8) - T_1(h_8 - h_{B1}) - (T_1 + T_2)(h_8 - h_5)$
$= - \frac{1}{2}a^2(W_1 + W_2)$

9. $- T_2(h_9 - h_8) - T_2(h_9 - h_6) = - \frac{1}{2}a^2W_2 - aq_{B1}$

a) Verify the finite difference equations so that you understand their physical basis. (Hint: read appendix, section 4.3.1.)

b) Write the equations in matrix form:

$$\underline{D}\ \underline{h} = \underline{q}$$

by writing out explicitly $\underline{D}$, $\underline{h}$, and $\underline{q}$.

c) Let vectors $(\partial \underline{q}/\partial b_j)_r - (\partial \underline{D}/\partial b_j)_r \underline{h}_r = \underline{J}_j^r$, and develop $\underline{J}_j^r$ (j = 1, 2, 3, 4) for the parameters $\beta_1 = T_1$, $\beta_2 = T_2$, $\beta_3 = W_1$, and $\beta_4 = q_{B1}$.

d) Write out vectors $\underline{J}_j b_j$ (j = 1, 2, 3, 4).

e) Using the modified Gauss-Newton method, develop the algorithm (not computer program) to solve for the parameters. Assume that all nodes except number 7 are observation points.

3.3.3. Convergence and Conditioning

It can be shown that a value of ρ, $0<\rho\leq 1$, exists for which the Gauss-Newton procedure, as modified using (3.3-18), will converge to the global minimum value of $S(\underline{b})$ provided that:

1) An initial estimate of the parameters can be found such that they lie within a parameter region R bounded by sets of parameters $\underline{b}^*$ defined by

$$S(\underline{b}_0) \leq \min_{\underline{b}^*} S(\underline{b}^*) \qquad (3.3\text{-}26)$$

and the global minimum point lies within this region.

2) For all $\underline{b}$ belonging to R, $\underline{X}$ is a continuous and unique matrix function.

3) The matrix $\underline{S}^T\underline{\omega}\,\underline{S}$ is nonsingular and is a continuous function of $\underline{b}$.

Condition 2 is almost always met. Condition 1 requires that the system be well enough understood that intelligent initial estimates of parameters can be made. Difficulties frequently arise in connection with condition 3. Unless the problem is correctly specified, the least squares coefficient matrix (for example, $\underline{S}^T\underline{\omega}\,\underline{S}$) can be singular. Moreover, problems often arise because of ill-conditioning (that is, near singularity) of the matrix. Although the addition of the Marquardt parameter, μ, is intended to help these cases, convergence can be difficult to obtain. In the following paragraphs, the general question of convergence is considered first. This is followed by discussions of singularity and ill-conditioning.

In general, the rate of convergence has been found to be related to the number of parameters being estimated, as predicted by theory. That is, the greater the number of parameters, the slower the rate, all other things being

equal. It is also related to the conditioning of the problem and to the nearness of the initial set of parameters to the optimum set, in that the rate of convergence is usually much faster near a minimum of $S(\underline{b})$. In general, one may expect convergence within a number of iterations equal to either 5 or twice the number of parameters, whichever is greater. Fewer iterations are often required for well-conditioned problems.

A problem that frequently retards the convergence rate, or even causes divergence, is overshoot. This happens when the parameter correction vector $\rho\underline{\delta}_{r+1}$ has a favorable orientation but is much longer than an ideal value. The result is that the new set of parameters $\underline{b}_{r+1}$ is almost as far as (or even further than) the old set $\underline{b}_r$ from the optimum value. A two-parameter example is illustrated in figure 3.3-4.

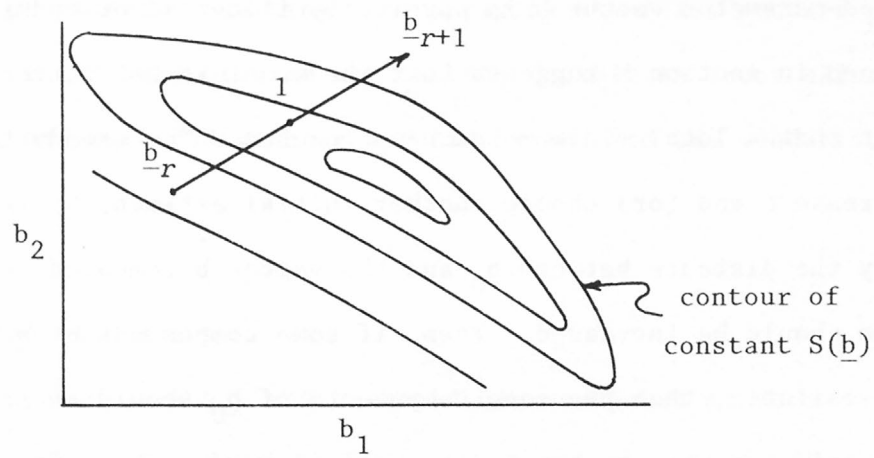


fig. 3.3-4

Overshoot is detected as large oscillations with accompanying changes in sign of components of $\rho\underline{\delta}_{r+1}$ from one iteration to the next. The remedy is to decrease the value of ρ such that $0<\rho<1$. In figure 3.3-4, a good value would be $\rho = 0.5$, which would give $\underline{b}_{r+1}$ at point 1.

Care must be taken not to make ρ too small so that undershoot becomes a problem. Undershoot occurs when $\rho\underline{\delta}_{r+1}$ is too small, and it manifests itself as small steps $\rho\underline{\delta}_{r+1}$, the components of which usually do not change sign. The remedy is to increase the value of ρ, in rare occasions such that $\rho>1$. As a practical guide, the best value of ρ is one that causes some oscillations in sign of a few components of $\rho\underline{\delta}_{r+1}$ from one iteration to the next.

Sometimes the solution may converge to a local minimum instead of the global minimum. Most commonly this can occur 1) if ρ is too large and overshoot causes the search vector $\rho\underline{\delta}_{r+1}$ to escape from the region defined by (3.3-26), or 2) if the initial estimate $\underline{b}_0$ is not in the region defined by (3.3-26). Detection of this problem is accomplished through adequate knowledge of the system so that the wrong solution can be recognized. If the computed parameter vector $\hat{\underline{b}}$ is physically illogical or model analysis discussed in section 5 suggests that the model is not correct, then one might suspect that a local minimum has been reached. The remedy to the problem is to decrease ρ and (or) choose another initial estimate $\underline{b}_0$. If $\underline{b}_0$ is changed, usually the distance between $\underline{b}_0$ and the vector $\hat{\underline{b}}$ computed for the local minimum should be increased. Thus, if some components of $\hat{\underline{b}}$ are far too large to be realistic, then the same components of $\underline{b}_0$ should be reduced in value. If several attempts at changing ρ and (or) $\underline{b}_0$ do not produce a change in $\hat{\underline{b}}$, then the cause of the poor results is probably not a local minimum.

As for the linear case, singularity of the least squares coefficient matrix occurs whenever columns of the sensitivity matrix are linearly dependent. Near singularity, caused by near linear dependence, is a more frequent occurrence. As a result of ill-conditioning, step sizes $\underline{\delta}_{r+1}$ can be highly erratic, appearing to head toward no well-defined point and can

be dominated by overshoot. In addition, some problems may start fairly well conditioned for the initial parameters but may become progressively more poorly conditioned during the iterative solution process.

The same techniques for analysis of poorly conditioned problems as are used for linear problems may be used for nonlinear problems as well. Whenever a problem is poorly conditioned, the sensitivity matrix $\underline{X}$ (or $\underline{Z}$) may be examined for a near zero column, and $\underline{S}^T\underline{\omega}\,\underline{S}$ may be examined for off-diagonal components near unity. Also, the orthogonal transformation may be used to indicate that ill-conditioning exists and to point out possible columns where linear dependency occurs.

Use of the Marquardt parameter, μ, is intended to improve conditioning by adding a small quantity to the main diagonal of the least squares coefficient matrix. Although conditioning is always artificially improved by employing μ, the parameters resulting from applying the least squares process to a very poorly conditioned problem may be considerably in error unless the actual causes of the poor conditioning are discovered and the conditioning improved without using μ.

3.3.4. Computation of μ and ρ

For best efficiency, both μ and ρ should be recomputed at each iteration, r. A number of schemes exist in the literature for making these computations, but virtually all schemes involve assuming several trial values of μ and ρ, then performing all of the calculations for iteration r for each of the trial values. The best values to use are then computed so as to minimize or substantially decrease $S(\underline{b})$. The problem with these schemes is that they require so much time that one is often much better off settling for approximate values of μ and ρ computed by using a much simpler scheme.

The scheme adopted here is derived from the considerations discussed in section 3.3.3. Parameter μ is used only when the problem is so poorly conditioned that the search direction $\underline{\delta}$ must be altered. Overshoot and undershoot are controlled primarily through use of ρ.

Because the Marquardt parameter is used with a scaled problem formulation, computations must be made using scaled quantities (see (3.3-15) and (3.3-20)). By direct computation, it can be verified that the scaled gradient $(\partial S(\underline{b})/\partial b_j)B_{jj}C_{jj}$ (j=1, 2, ..., p) of the sum of squares $S(\underline{b})$ is given by

$$\underline{C}^T\underline{B}^T \frac{\partial S(\underline{b})}{\partial \underline{b}} = -\underline{g} = -\underline{S}^T\underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b})) \tag{3.3-27}$$

Thus, by definition (see Spiegel, 1959, p. 16) the angle between $\underline{g}$, which points directly down the sum of squares surface, and displacement vector $\underline{\delta}$ is given for the rth iteration by

$$\cos\theta = \frac{\underline{\delta}_{r+1}^T\underline{g}_r}{\sqrt{(\underline{\delta}_{r+1}^T\underline{\delta}_{r+1})(\underline{g}_r^T\underline{g}_r)}} \tag{3.3-28}$$

For two parameters, b_1 and b_2, the relationships given by (3.3-28) are illustrated in figure 3.3-5. (Note the use of scaled parameters; see (3.3-16) and (3.3-19).)

If $\theta = 90°$, then as discussed in section 3.3.1, no improvement in parameters is likely to result from application of the Gauss-Newton procedure. However, application of the Marquardt parameter, $\mu > 0$, will result in $\theta < 90°$ (Marquardt, 1963) because vector $\underline{\delta}$ is shifted progressively toward $\underline{g}$ as μ increases. Thus, a viable scheme for choosing μ is to define a maximum value of θ, $\theta_{mx} < 90°$, and compute μ so that θ never exceeds θ_{mx}. This can be

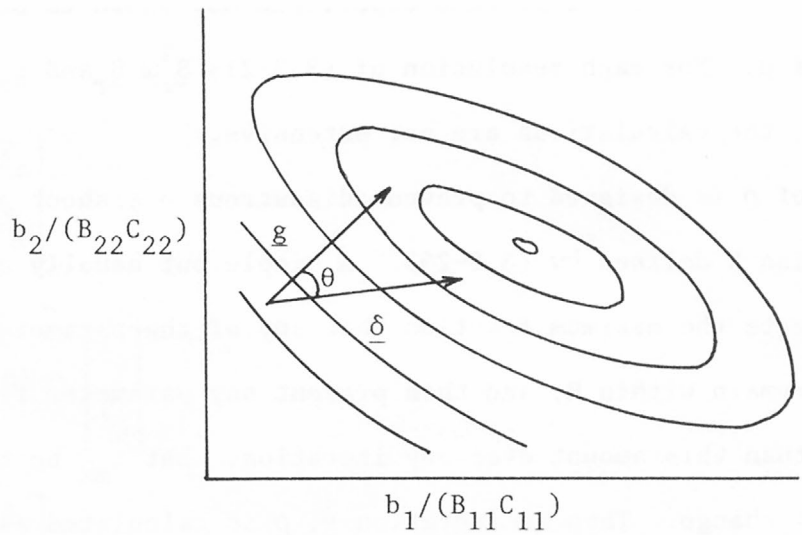


fig. 3.3-5

accomplished rather simply. At the beginning of the regression set $\mu_0 = 0$. Then at each iteration r, check and recompute μ as necessary:

$$\left.\begin{aligned} \mu_r &= \mu_\ell \text{ if } \underline{\delta}_{r+1}^T \underline{g}_r \geq \cos\theta_{mx} \sqrt{(\underline{\delta}_{r+1}^T \underline{\delta}_{r+1})(\underline{g}_r^T \underline{g}_r)} \\ \mu_{\ell+1} &= \tfrac{3}{2}\mu_\ell + .001 \text{ if } \underline{\delta}_{r+1}^T \underline{g}_r < \cos\theta_{mx} \sqrt{(\underline{\delta}_{r+1}^T \underline{\delta}_{r+1})(\underline{g}_r^T \underline{g}_r)} \end{aligned}\right\} \quad (3.3\text{-}29)$$

At the beginning of iteration r, $\ell = 1$ and $\mu_\ell = \mu_{r-1}$. Then (3.3-21) is solved and (3.3-29) is applied. If the second of (3.3-29) is employed, (3.3-21) is resolved using $\mu_{\ell+1}$, and ℓ is incremented by one. This process is continued until the first of (3.3-29) is used, at which point the appropriate value of μ for iteration r has been found. The formula for computing

$\mu_{\ell+1}$ from μ_ℓ is empirical but gives what experience has shown to be a good range in values of μ. For each resolution of (3.3-21) $\underline{S}_r^T \underline{\omega}\, \underline{S}_r$ and $\underline{g}_r$ are not recomputed. Thus, the calculations are not extensive.

Computation of ρ is designed to prevent disastrous overshoot and to keep $\rho\underline{\delta}$ within the region R defined by (3.3-26). A simple but usually effective scheme is to estimate the maximum fraction that any of the parameters could change and still remain within R, and then prevent any parameter from changing anymore than this amount over any iteration. Let t_{mx} be this maximum fractional change. Then at iteration r, ρ is calculated as follows:

$$t = \max_i |d_i^{r+1}| \tag{3.3-30}$$

$$\left.\begin{aligned} &\rho = 1 \text{ if } t \leq t_{mx} \\ &\rho = t_{mx}/t \text{ if } t > t_{mx} \end{aligned}\right\} \tag{3.3-31}$$

3.4. REGRESSION INCLUDING PRIOR INFORMATION

3.4.1. Model Structure

Recall that the standard nonlinear regression model including prior information on the parameters may be written in the form (see (3.1-32))

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{\beta}) + \underline{\varepsilon} \tag{3.4-1}$$

where

$$\underline{Y} = \begin{bmatrix} \underline{Y}_s \\ \underline{Y}_p \end{bmatrix} \tag{3.4-2}$$

$$\underline{f}(\underline{\xi}, \underline{\beta}) = \begin{bmatrix} \underline{f}_s \\ \underline{f}_p \end{bmatrix} \tag{3.4-3}$$

$$\underline{\varepsilon} = \begin{bmatrix} \underline{\varepsilon}_s \\ \underline{\varepsilon}_p \end{bmatrix} \tag{3.4-4}$$

and subscripts s and p indicate sample and prior information, respectively. To correspond with the above partitions into sample and prior information, the sensitivity matrix should be written in the form

$$\underline{X} = \begin{bmatrix} \underline{X}_s \\ \underline{X}_p \end{bmatrix} \tag{3.4-5}$$

where $\underline{X}$ is a function of $\underline{\beta}$ for a nonlinear model. The various transformations of $\underline{X}$ used to control round-off error are defined analogously. Finally, recall that the weight matrix is partitioned as

$$\underline{\omega} = \begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix} \tag{3.4-6}$$

where $\underline{\omega}_s$ and $\underline{\omega}_p$ correspond to sample and prior information, respectively.

Often partition $\underline{X}_p$ will be obtained in a different manner than $\underline{X}_s$. For example, the model for the sample information may be numerical of the type (3.3-23), whereas the model for the prior information may be of the analytical linear or nonlinear form. Thus, $\underline{X}_s$ would be obtained as described in section 3.3.2, and $\underline{X}_p$ would be obtained as described in section 3.3.1. Other types of differences are handled in a similar fashion. Obviously, if sample and prior models are the same, then $\underline{X}_s$ and $\underline{X}_p$ are obtained in the same manner.

In spite of the possibilities for combinations of linear, nonlinear, analytical, and numerical models, it should be remembered that all models have the general form of the incremental linear model when expanded in the Taylor series. Because all models resolve to the incremental linear form, for simplicity subsequent discussions in this section are based on this model only.

3.4.2. Solution Procedures

Whenever $\underline{w}_s$ and $\underline{w}_p$ are both known, solution for both linear and nonlinear models is unaltered from that given in the previous sections. However, recall that, because of the block diagonal form of (3.4-6), $S(\underline{b})$ and, hence, the normal equations can be written in a special form. By applying the standard minimization technique to $S(\underline{b})$ as given by (3.1-42):

$$S(\underline{b}) = [\underline{e}_s^T \; \underline{e}_p^T] \begin{bmatrix} \underline{w}_s & \underline{0} \\ \underline{0} & \underline{w}_p \end{bmatrix} \begin{bmatrix} \underline{e}_s \\ \underline{e}_p \end{bmatrix} \tag{3.4-7}$$

the normal equations for the incremental linear model become

$$[\underline{X}_s^T \; \underline{X}_p^T]\begin{bmatrix}\underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p\end{bmatrix}\begin{bmatrix}\underline{X}_s \\ \underline{X}_p\end{bmatrix}(\hat{\underline{b}} - \underline{b})$$

$$= [\underline{X}_s^T \; \underline{X}_p^T]\begin{bmatrix}\underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p\end{bmatrix}\begin{bmatrix}\underline{Y}_s - \underline{Y}_{bs} \\ \underline{Y}_p - \underline{Y}_{bp}\end{bmatrix} \qquad (3.4\text{-}8)$$

or

$$(\underline{X}_s^T\underline{\omega}_s\underline{X}_s + \underline{X}_p^T\underline{\omega}_p\underline{X}_p)(\hat{\underline{b}} - \underline{b})$$

$$= \underline{X}_s^T\underline{\omega}_s(\underline{Y}_s - \underline{Y}_{bs}) + \underline{X}_p^T\underline{\omega}_p(\underline{Y}_p - \underline{Y}_{bp}) \qquad (3.4\text{-}9)$$

where $\underline{Y}_{bs}$ and $\underline{Y}_{bp}$ are $\underline{Y}_b$ for sample and prior information, respectively. Equation (3.4-9) is of the same form and thus applies as the equation for each iteration of solution of a nonlinear regression problem.

Frequently the weight matrix is constructed from variance-covariance matrices for $\underline{\varepsilon}_s$ and $\underline{\varepsilon}_p$ that are given in the form

$$Var(\underline{\varepsilon}_s) = \underline{V}_s\sigma^2 \qquad (3.4\text{-}10)$$

$$Var(\underline{\varepsilon}_p) = \underline{U} \qquad (3.4\text{-}11)$$

where the usual form $\underline{V}_p\sigma^2$ for (3.4-11) cannot be used because $Var(\underline{\varepsilon}_p)$ is not known as a function of σ^2. Thus, with $\underline{\omega}$ defined as

$$\underline{\omega} = \begin{bmatrix}[Var(\underline{\varepsilon}_s)]^{-1} & \underline{0} \\ \underline{0} & [Var(\underline{\varepsilon}_p)]^{-1}\end{bmatrix}\sigma^2$$

$$= \begin{bmatrix}\underline{V}^{-1}_s & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2\end{bmatrix} \qquad (3.4\text{-}12)$$

(3.4-9) becomes

$$(\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s + \underline{X}_p^T \underline{U}^{-1} \underline{X}_p \sigma^2)(\hat{\underline{b}} - \underline{b})$$

$$= \underline{X}_s^T \underline{V}_s^{-1} (\underline{Y}_s - \underline{Y}_{bs}) + \underline{X}_p^T \underline{U}^{-1} \sigma^2 (\underline{Y}_p - \underline{Y}_{bp}) \qquad (3.4\text{-}13)$$

Hence, it would appear that σ^2 would have to be known in order to form the regression solution, whereas σ^2 is considered to be an unknown.

Theil (1963) showed that, for a linear model, σ^2 may be estimated for use in (3.4-13) by its ordinary least squares estimate (that is, the estimate obtained when prior information is not used). Bias produced by this estimate was shown by Theil (1963) to be of the order of $n_s^{-\frac{1}{2}}$. The procedure to be followed is to first solve the ordinary least squares problem by omitting all prior information; then find the estimate of σ^2 (to be given further on); finally use this estimate of σ^2 in the normal equations to solve the complete problem, including prior information.

References Cited

Bakr, A. A., Gelhar, L. W., Gutjahr, A. L., and MacMillan, J. W., 1978, Stochastic analysis of spatial variability in subsurface flows, 1--Comparison of one- and three-dimensional flows: Water Resources Research, v. 14, no. 2, p. 263-271.

Draper, N. R., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

Gutjahr, A. L., Gelhar, L. W., Bakr, A. A., and MacMillan, J. W., 1978, Stochastic analysis of spatial variability in subsurface flows, 2--Evaluation and application: Water Resources Research, v. 14, no. 5, p. 953-959.

Marquardt, D. W., 1963, An algorithm for least squares estimation of nonlinear parameters: Journal of the Society for Industrial and Applied Mathematics, v. 11, no. 2, p. 431-441.

Kitanidis, P. K., and Vomvoris, E. G., 1983, A geostatistical approach to the inverse problem in ground-water modeling (steady state) and one-dimensional simulations: Water Resources Research, v. 19, no. 3, p. 677-690.

Krumbein, W. C., and Graybill, F. A., 1965, An introduction to statistical models in geology: New York, McGraw-Hill, 475 p.

Smith, L., and Freeze, R. A., 1979a, Stochastic analysis of steady-state groundwater flow in a bounded domain, 1--One-dimensional simulations: Water Resources Research, v. 15, no. 3, p. 521-528.

_____1979b, Stochastic analysis of steady-state groundwater flow in a bounded domain, 2--Two-dimensional simulations: Water Resources Research, v. 15, no. 6, p. 1543-1559.

Spiegel, M. R., 1959, Theory and problems of vector analysis: New York, Schaum, 225 p.

Theil, H., 1963, On the use of incomplete prior information in regression analysis: American Statistical Association Journal, v. 58, no. 302, p. 401-414.

Additional Reading

Beck, J. V., and Arnold, K. J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Draper, N. R., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

4. NUMERICAL NONLINEAR REGRESSION SOLUTION OF GENERAL STEADY-STATE GROUND-WATER FLOW PROBLEMS

4.1. ASSUMED MODEL AND SOLUTION PROCEDURE

A model to solve fairly general steady-state ground-water flow problems by using the regression procedures presented in section 3 is given in this section. A complete description of the method is given first. Documentation and listing of the computer program are given in the appendix, section 4.3.4.

4.1.1. Problem Specification

The equation assumed to govern ground-water flow for the class of problems to be analyzed is derived from (1.1-1) by letting $\partial h/\partial t \to 0$, which results in

$$\frac{\partial}{\partial x}\left(T_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T_{yy}\frac{\partial h}{\partial y}\right) + R(H - h) + W + \sum_{\ell=1}^{NW} \delta(x - a_\ell)\delta(y - b_\ell)Q_\ell = 0 \qquad (4.1\text{-}1)$$

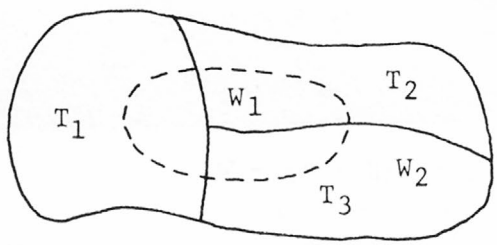
where the symbols are defined the same as for (1.1-1).

Functions $T_{\xi\xi}$ (that is, T_{xx} and T_{yy}), R, and W are each formulated within the region being modeled as the product of a parameter and a given (or known) function. To provide for spatial variability of parameters, the region is subdivided into a number of discrete zones within each of which the parameters are assumed constant. Hence, known spatial variability (often, but not necessarily, smooth or continuous) is superimposed upon the discontinuous spatial variability dictated by the parameter zonation. As an example, hydraulic conductivity $K_{\xi\xi}$ may often be considered to be constant

within particular rock types, each of which may be considered to be a discrete zone. Thus, $K_{\xi\xi}$ may be considered to be a parameter. Thickness b may be known from measurements and may vary continuously. The function $T_{\xi\xi}$ is, of course, given as $K_{\xi\xi}b$. Finally, because the controls that dictate a particular zonation may vary from parameter to parameter, zones for one type of parameter (for example, the parameter contained in $T_{\xi\xi}$) do not necessarily correspond to zones for another type (for example, the parameter contained in W). An example of zonation is given in figure 4.1-1 where the given function is unity so that the parameters are $T_{\xi\xi}$ and W.

$T = T_{xx}$

$= T_{yy}$



Dashed lines separate W zones, and solid lines separate T zones.

fig. 4.1-1

Internal boundary conditions applying at discontinuities in $T_{\xi\xi}$ are that the specific discharge normal to the boundary and the hydraulic head both remain unchanged as the boundary is crossed. External boundary conditions applying on the periphery of the region being modeled include specified specific discharge normal to the boundary, specified hydraulic head at the boundary, or a mixture of the two types along the boundary.

Specific discharge q_B normal to the boundary is assumed to vary along the boundary in a manner similar to that of $T_{\xi\xi}$. It may have discontinuities and may vary smoothly between discontinuities. It might be expected that discontinuities in q_B would often correspond to discontinuities in $T_{\xi\xi}$.

Hydraulic head variation along a specified head boundary is a continuous function, h_B, although the boundary may be subdivided into segments within each of which head h_B may vary linearly or curvilinearly with distance. Subdivision into segments is often based on the causes of the known head conditions along various segments of the boundary.

Unknown quantities to be determined are $T_{\xi\xi}$, R, W, q_B and h_B. Single values of T_{xx} or T_{yy}, R, and W (or multipliers such as K_{xx} or K_{yy}) are assumed to be parameters in each zone, although any of these parameters may be held constant (assumed known as exact prior information) in each zone. Separate zones (or segments) are specified for the values of q_B, and single multipliers for the fluxes in each flux zone are assumed to be parameters. Within each of the separate segments (zones) of a specified head boundary the heads are adjusted as a linear function of distance by the regression procedure so that the parameters are the values of head at each end of each segment. Even though the adjustment is linear, the actual shape of the head profile along the boundary may be curvilinear.

4.1.2. Matrix Form of Regression Model

For most field problems (4.1-1) with its attendant boundary conditions cannot be solved analytically. Thus, the regression solution must be based on a numerical solution of (4.1-1), which is expressed as a matrix equation. The particular numerical solution method is given in the appendix, section 4.3.1.

The matrix equation comprising the numerical solution is given as

$$\underline{D}\ \underline{h} = \underline{q} \qquad (4.1\text{-}2)$$

where

$\underline{D}$ = the square coefficient matrix of order m, the number of nodes used to discretize the modeled region;

$\underline{h}$ = the hydraulic head vector of order m; and

$\underline{q}$ = the known vector of order m.

Matrix $\underline{D}$ contains parameters for $T_{\xi\xi}$ and R, whereas vector $\underline{q}$ can contain all parameter types. To express any specified head value, say $h_j = h_{Bj}$,

$$D_{jj} = 1 \qquad D_{ji} = D_{ij} = 0, \; i \neq j \tag{4.1-3}$$

$$q_j = h_{Bj} \tag{4.1-4}$$

where D_{jj}, D_{ij}, and D_{ji} are components of $\underline{D}$, and q_j is a component of $\underline{q}$. To accomplish the condition that $D_{ij} = 0$ in equation i, $i \neq j$, the term $D_{ij} h_{Bj}$ is tranferred to the right-hand side of equation i so that q_i contains the term $-D_{ij}h_{Bj}$. Then D_{ij} in $\underline{D}$ is set to zero.

The known head value is computed from

$$h_{Bj} = \frac{h^0_{Bj} \left[L_j H_t + (1 - L_j) H_s\right]}{L_j H^0_t + (1 - L_j) H^0_s} \tag{4.1-5}$$

where

s = node at one end of the boundary segment within which node j lies;

t = node at the other end;

H_s = head (parameter) at node s;

H_t = head (parameter) at node t;

L_j = ratio of distance along the boundary from node s to node j, to distance along the boundary from node s to node t; and

superscript 0 = an initial or reference value.

Indices s and t can be equal so that $j = s = t$ for the case where only one specified head is present. Also, H_s and H_t can represent the same parameter, so that the entire specified head boundary behaves as a unit.

Because (4.1-2) is a linear matrix equation, the modified Gauss-Newton procedure is employed to solve the regression problem. Equation (4.1-2) is the same as (3.3-22), except that in (4.1-2) the coefficient matrix $\underline{D}$ and right-hand-side vector $\underline{q}$ are not functions of dependent variable $\underline{h}$. Hence, sensitivities may be calculated using (3.3-25). For simplicity, observation points are assumed to coincide with node points.

Prior information is assumed to be given (if available) on each parameter individually so that

$$\underline{X}_p = [\underline{I}_{n_p} \ \underline{0}]_{(n_p \times p)} \tag{4.1-6}$$

where $\underline{I}_{n_p}$ is the identity matrix of order n_p. Thus, in (4.1-6) direct prior information is assumed to be given on the first n_p parameters. Placement of these parameters first in the vector $\underline{\beta}$ simplifies theoretical statement of (4.1-6) but is not necessary in practice.

The linearized regression model assumed, then, is of the form of (3.3-7) partitioned as suggested by (3.4-2) through (3.4-5):

$$
\left.
\begin{aligned}
Y_{s1} &\cong f_1^r + X_{11}^r(b_1^{r+1} - b_1^r) + X_{12}^r(b_2^{r+1} - b_2^r) + \ldots + X_{1p}^r(b_p^{r+1} - b_p^r) + e_{s1} \\
Y_{s2} &\cong f_2^r + X_{21}^r(b_1^{r+1} - b_1^r) + X_{22}^r(b_2^{r+1} - b_2^r) + \ldots + X_{2p}^r(b_p^{r+1} - b_p^r) + e_{s2} \\
&\ldots \\
Y_{sn_s} &\cong f_{n_s}^r + X_{n_s 1}^r(b_1^{r+1} - b_1^r) + X_{n_s 2}^r(b_2^{r+1} - b_2^r) + \ldots + X_{n_s p}^r(b_p^{r+1} - b_p^r) \\
&\quad + e_{sn_s} \\
Y_{p1} &\cong b_1^r + (b_1^{r+1} - b_1^r) + e_{p1} \\
Y_{p2} &\cong b_2^r + (b_2^{r+1} - b_2^r) + e_{p2} \\
&\ldots \\
Y_{pn_p} &\cong b_{n_p}^r + (b_{n_p}^{r+1} - b_{n_p}^r) + e_{pn_p}
\end{aligned}
\right\} \quad (4.1\text{-}7)
$$

where $\underline{f}_r$ is the current vector of computed hydraulic heads at observation points and Y_{pi} is the prior estimate of β_i. Partitions $\underline{f}_s$ and $\underline{f}_p$ are given in (4.1-7) by $\underline{f}_s \equiv \underline{f}_r$ and $\underline{f}_p \equiv \underline{b}_p^r$, where $\underline{b}_p^r$ signifies the rth estimate of the first n_p parameters (on which there is prior information). Similarly, partitions $\underline{X}_s$ and $\underline{X}_p$ are $\underline{X}_s \equiv \underline{X}_r$ and $\underline{X}_p \equiv [\underline{I}_{n_p}\ \underline{0}]$, as given by (4.1-6).

No correlation or other coupling is assumed to exist among components of either $\underline{\varepsilon}_s$ or $\underline{\varepsilon}_p$. Matrix $\underline{\omega}$ is assumed to be of the form

$$
\underline{\omega} = \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2 \end{bmatrix} \qquad (4.1\text{-}8)
$$

where

$$\underline{V}_s^{-1} = [Var(\underline{\varepsilon}_s)]^{-1}\sigma^2 \tag{4.1-9}$$

$$\underline{U}^{-1}\sigma^2 = [Var(\underline{\varepsilon}_p)]^{-1}\sigma^2 \tag{4.1-10}$$

and both $\underline{V}_s^{-1}$ and $\underline{U}^{-1}$ are diagonal.

4.1.3. Nonlinear Regression Solution

Nonlinear regression solution for the model given in section 4.1.2 is accomplished by using the algorithm at the end of section 3.3.1. Sensitivities $\underline{X}_r$ are calculated as illustrated in the appendix, section 4.3.2.

The normal equations used are (3.3-21), as modified to include prior information (see (3.4-13)):

$$[(\underline{S}_s)_r^T \underline{V}_s^{-1}(\underline{S}_s)_r + (\underline{S}_p)_r^T \underline{U}^{-1}(\underline{S}_p)_r s^2 + \mu\underline{I}]\underline{\delta}_{r+1}$$

$$= (\underline{S}_s)_r^T \underline{V}_s^{-1}(\underline{Y}_s - (\underline{f}_s)_r) + (\underline{S}_p)_r^T \underline{U}^{-1}s^2(\underline{Y}_p - \underline{b}_p^r) \tag{4.1-11}$$

where subscripts s and p refer to sample and prior information, respectively;

$\underline{S}_p$ = the matrix $[\underline{I}_{n_p} \quad \underline{0}]$ from (4.1-6), transformed using (3.3-15) and (3.3-20);

$\underline{b}_p^r$ = a vector composed of the rth estimate of those parameters on which there is prior information; and

s^2 = the ordinary least squares estimate of σ^2 (to be developed later on).

4.2. SINGULARITY AND CONDITIONING

Singularity of the least squares coefficient matrix can occur whenever 1) there are no measured flow rates (such as well or spring discharges) in the model and 2) an attempt is made to compute all parameters. To understand how this occurs consider first the case where there are no specified head parameters and no prior information, but all other parameters are to be computed. In this case it can be shown (see the appendix, section 4.3.3) that

$$\underline{J}\,\underline{b} = \underline{0} \tag{4.2-1}$$

where subscripts r were omitted to simplify nomenclature, and $\underline{J} = \{\underline{J}_j\} = \{\partial \underline{q}/\partial b_j - (\partial \underline{D}/\partial b_j)\underline{h}\}$. By using (3.3-24) it follows that

$$\underline{J}_j = \frac{\partial \underline{q}}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j}\,\underline{h} = \underline{D}\frac{\partial \underline{h}}{\partial b_j} \tag{4.2-2}$$

so that

$$\underline{J}\,\underline{b} = \sum_{j=1}^{p} \underline{J}_j b_j = \underline{D} \sum_{j=1}^{p} \frac{\partial \underline{h}}{\partial b_j}\, b_j \tag{4.2-3}$$

Because $\underline{D}$ is nonsingular,

$$\underline{D}^{-1}\underline{D} \sum_{j=1}^{p} \frac{\partial \underline{h}}{\partial b_j}\, b_j = \underline{0}$$

or, by eliminating those nodes not corresponding to observation points from $\partial \underline{h}/\partial b_j$,

$$\sum_{j=1}^{p} \underline{Z}_{sj} = \underline{0} \tag{4.2-4}$$

Equation (4.2-4) can also be written in the form

$$\underline{Z}_s \underline{c} = \underline{0} \tag{4.2-5}$$

where

$$\underline{c} = [1, 1, \ldots, 1]^T_{(p)} \tag{4.2-6}$$

It will be recalled that (4.2-5) implies that the least squares coefficient matrix is singular.

If there is a known flow rate Q_i at node i, then (4.2-1) becomes

$$\underline{J}\ \underline{b} = \underline{Q} \tag{4.2-7}$$

where

$$\underline{Q} = [0, 0, \ldots, Q_i, \ldots, 0]^T_{(m)} \tag{4.2-8}$$

so that, all other things being equal,

$$\underline{Z}_s\ \underline{c} \neq \underline{0} \tag{4.2-9}$$

In the case where at least one parameter j is fixed, then $\underline{J}\ \underline{b}$ has column j of $\underline{J}$ and element j of $\underline{b}$ deleted. Thus, (4.2-1) no longer holds, so that (4.2-9) will hold, if no other source of singularity exists.

Whenever there are specified head parameters, $\underline{J}$ and $\underline{Z}_s$ both contain columns resulting from these parameters, and $\underline{J}\ \underline{b} \neq \underline{0}$. However, for those columns not involving the specified head parameters, $\sum_j \underline{J}_j b_j = \underline{0}$, where j denotes all parameters except specified head parameters. If $c_j = 0$ for those columns resulting from specified head parameters and $c_j = 1$ for the remaining columns, then (4.2-5) holds for the entire sensitivity matrix, which indicates that the problem is again singular. Because addition of specified head parameters has no influence on this type of singularity, it is assumed for simplicity for the remainder of this section that there are no specified head parameters.

Singularity caused by attempting to find all parameters in the absence of known flow rates can be rectified by using prior information. For the case of prior information, (4.2-5) can be written

$$\begin{bmatrix} \underline{Z}_s \\ \\ \underline{Z}_p \end{bmatrix} \underline{c} = \underline{0} \qquad (4.2\text{-}10)$$

The only way for (4.2-10) to hold is if $\underline{Z}_s\ \underline{c} = \underline{0}$ and $\underline{Z}_p\ \underline{c} = \underline{0}$. If the only cause of the singularity is given by (4.2-1), then $\underline{c} = \underline{1}$ is the only linearly independent solution of (4.2-10), and, if $\underline{Z}_p$ is derived from (4.1-6), $Z_{pii} = 0$ ($i = 1, 2, \ldots, p$) is the only way that (4.2-10) can hold. Hence, prior information on any parameter can theoretically condition the problem so that all parameters can be found.

The maximum number of parameters that can be found for any problem can also be obtained through nondimensionalization (or partial nondimensionalization) to find the smallest number of independent groups. In addition, nondimensionalization also illustrates the idea that solution is actually

often best expressed in terms of ratios of the parameters. As an example, consider the case where a region is composed of two zones where T_1, T_2, W_1, and W_2 are parameters. Then the flow equations for each zone are

$$\left.\begin{aligned} T_1\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right) + W_1 &= 0 \\ T_2\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right) + W_2 &= 0 \end{aligned}\right\} \quad (4.2\text{-}11)$$

and the boundary conditions between zones are

$$\left.\begin{aligned} T_1\left(\frac{\partial h}{\partial n}\right)_1 &= T_2\left(\frac{\partial h}{\partial n}\right)_2 \\ (h)_1 &= (h)_2 \end{aligned}\right\} \quad (4.2\text{-}12)$$

where the notation $(\ldots)_1$ indicates that the quantity in parentheses is evaluated just within the 1 side of the boundary and similarly for $(\ldots)_2$. If (4.2-11) and (4.2-12) are written in the alternate forms

$$\left.\begin{aligned} \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{W_1}{T_1} &= 0 \\ \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{W_2}{T_2} &= 0 \end{aligned}\right\} \quad (4.2\text{-}13)$$

$$\left.\begin{aligned} \frac{T_1}{T_2}\left(\frac{\partial h}{\partial n}\right)_1 &= \left(\frac{\partial h}{\partial n}\right)_2 \\ (h)_1 &= (h)_2 \end{aligned}\right\} \quad (4.2\text{-}14)$$

it is seen that instead of four independent parameters, there are only three written as ratios: T_1/W_1, T_2/W_2, T_1/T_2. A known flow rate in zone 1, Q_1, would add the term Q_1/T_1 to the first of (4.1-13). In this case T_1/W_1, T_2/W_2, T_1/T_2, and Q_1/T_1 could all be considered parameters. Knowing Q_1 (either exactly or with uncertainty) would provide unique estimates of the four original parameters. Another approach would be to find T_1, T_2, W_1, W_2 as parameters, knowing that the problem is not singular because there are four independent ratios for the problem.

Another common way for singularity to occur is if a column of $\underline{Z}_s$ (and, thus $\underline{X}_s$) is zero: $\underline{Z}_{sj} = \underline{0}$. This results if measurements are taken at points where there is no sensitivity to the parameter, b_j, corresponding to the column. If this is the only source of singularity, then $\underline{c} = [0, 0, \ldots, 1, 0, \ldots, 0]^T$, where the one appears in row j of $\underline{c}$, is the only linearly independent solution of (4.2-5). In this case if $\underline{Z}_p$ is derived from (4.1-6) and there is prior information on b_j, then $\underline{Z}_p\underline{c} \neq \underline{0}$ so that the prior information solves the singularity problem. Addition of prior information on any other parameter alone obviously will not help.

Two sources of singularity result if $\underline{Z}_{sj} = \underline{0}$ and $\underline{Jb} = \underline{0}$. In this case one solution of (4.2-5) is, as before, given by (4.2-6). However, because $\underline{Z}_{sj} = \underline{0}$, c_j can be any arbitrary value less than infinity and so can be set to zero. Hence, $\underline{c} = [1, 1, \ldots, 0, 1, \ldots, 1]^T$, where the zero appears in row j of $\underline{c}$, is another solution to (4.2-5). Addition of prior information on parameter j alone does not solve the singularity problem because, even

though Z_{pjj} does not equal zero, c_j and Z_{pii}, $i \neq j$, do equal zero so that $\underline{Z}_p\underline{c} = \underline{0}$. A third solution of (4.2-5) is $\underline{c} = [0, 0, \ldots, 1, 0, \ldots, 0]^T$, where the one appears in row j of $\underline{c}$. In this instance addition of prior information on any or all parameters except parameter j yields $\underline{Z}_p\underline{c} = \underline{0}$. Thus, it may be concluded that if $\underline{Z}_{sj} = \underline{0}$ and $\underline{Jb} = \underline{0}$, then the problem is singular even if prior information on any or all parameters is added.

If the columns of $\underline{Z}_s$ are almost linearly dependent, then the problem is ill-conditioned. Thus, if either Q_i in (4.2-8) is almost zero or $\underline{Z}_{sj} \cong \underline{0}$, then an ill-conditioned problem can result. However, there are a number of ways in which ill-conditioning can occur. The techniques given in section 3.2.3 can be used to detect conditioning problems.

Problem 4.2-1

Solve problem 3.2-1 with the numerical regression computer program (appendix 4.3.4). Assume that the stream tube is one foot wide and that transmissivities are unity. Place a row of nodes along each side of the stream tube, but specify observed heads only along one row (number of observed heads should be the same as in problem 3.2-1). Allow two iterations. What would happen if you were to attempt to estimate both W and T?

Problem 4.2-2

Figure 1 gives the zone map for a steady-state ground-water flow system in a hypothetical region. The finite difference mesh and types of boundary conditions also are shown on the map. Use the regression program (appendix 4.3.4) to construct a regression flow model for the region.

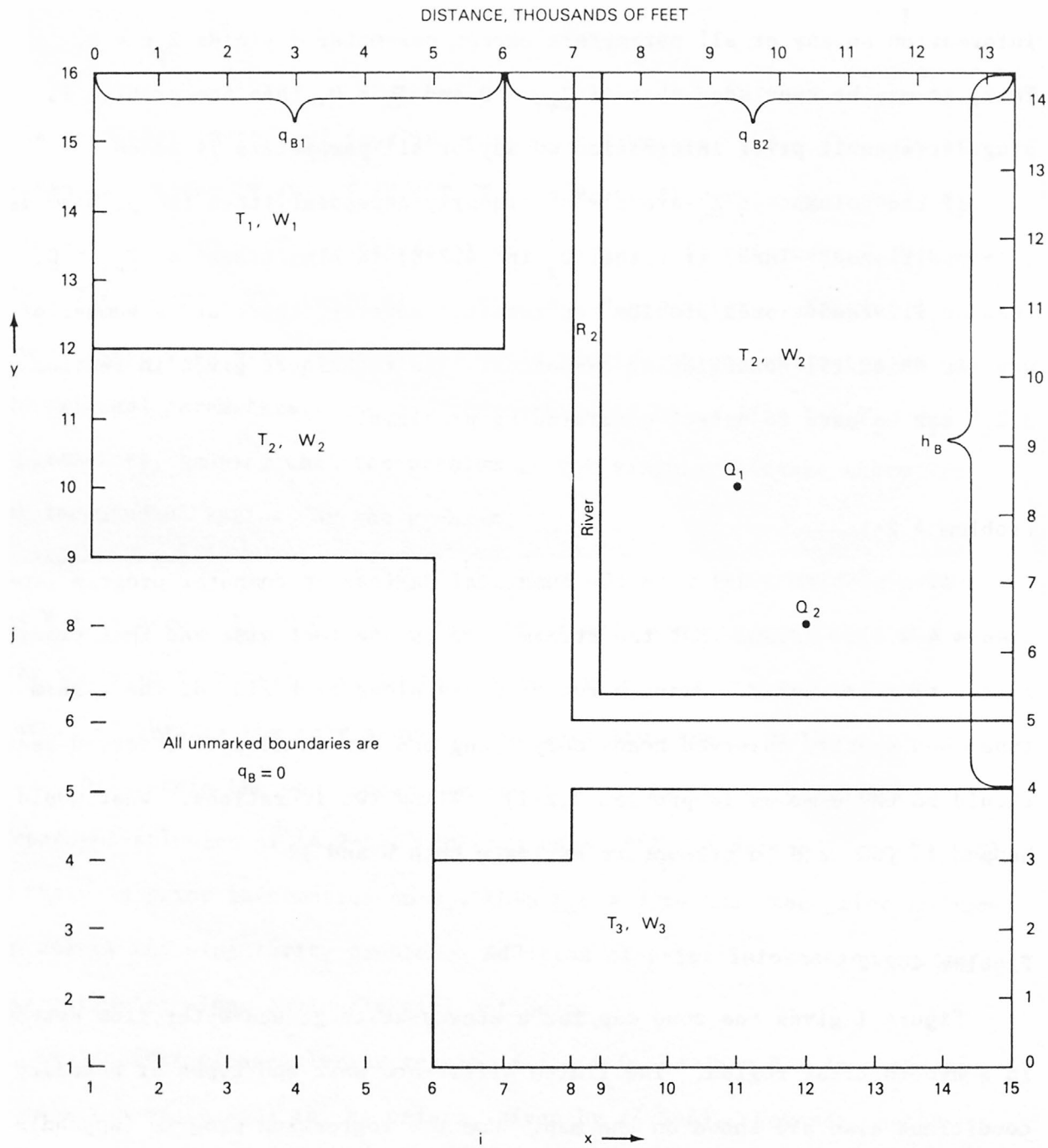

DISTANCE, THOUSANDS OF FEET
q_{B1}
q_{B2}
T_1, W_1
T_2, W_2
R_2
T_2, W_2
h_B
Q_1
River
Q_2
All unmarked boundaries are
$q_B = 0$
T_3, W_3
y
x
j
i

fig. 1

There is prior information on the following parameters:

$T_2 = 420$ ft^2/day	coefficient of variation =	0.20
$W_1 = 0.0004$ ft/day		0.30
$W_3 = 0.00017$ ft/day		0.30
$R_2 = 0.08$ day^{-1}		0.10
$Q_1 = -97000$ ft^3/day		0.02
$Q_2 = -51000$ ft^3/day		0.02
h_B at (15, 16) = 10.4 ft		0.10
h_B at (15, 7) = 4.8 ft } one parameter		0.10
h_B at (15, 6) = 4.8 ft }		0.10
h_B at (15, 5) = 5.4 ft		0.10

Assume that h_B varies linearly between the estimated values. For one reason or another, the estimates of h_B at the four nodes are not observations. (They may have been interpolated from a contour map, for example.)

There is no more prior information, but probable limits of variation for the remaining parameters are

$$30 \leq T_1 \leq 80$$

$$10 \leq T_3 \leq 40$$

$$-0.0003 \leq W_2 \leq -0.00005$$

$$0.2 \leq q_{B1} \leq 0.8$$

$$0.15 \leq q_{B2} \leq 0.4$$

From these ranges, initial estimates of the parameters may be determined.

The observed head data in table 1 were collected. They are of uniform reliability.

Table 1.

Node	Value	Node	Value
(8,2)	60.70 ft	(7,11)	6.68 ft
(14,2)	75.64	(13,11)	-15.32
(12,3)	60.27	(3,12)	16.88
(10,4)	29.67	(5,12)	15.87
(7,5)	4.22	(9,12)	4.48
(11,5)	4.37	(11,12)	-18.34
(13,5)	6.07	(13,13)	-2.47
(15,5)	5.81	(15,13)	8.10
(10,7)	4.57	(3,14)	54.12
(8,8)	5.21	(5,14)	38.27
(12,8)	-44.89	(10,14)	0.053
(15,8)	7.01	(12,14)	-2.92
(7,9)	6.95	(7,15)	8.30
(4,10)	12.21	(14,15)	4.54
(9,10)	4.04	(2,16)	85.82
(11,10)	-89.36	(11,16)	2.26

The river stage is about 4.5 ft everywhere.

Assuming that $\sigma^2 = 1$, find all possible parameters for the model. First, however, determine how many parameters you can find!

Examine the sensitivity maps. Are there data in relatively high sensitivity areas for all parameters? Do you think that there are places where new data points would improve the results?

4.3. APPENDICES

4.3.1. Integrated Finite Difference Model

The numerical solution of (4.1-1) is obtained by using integrated finite difference methods. A rectangular grid of nodes is assumed as indicated in figure 4.3-1.

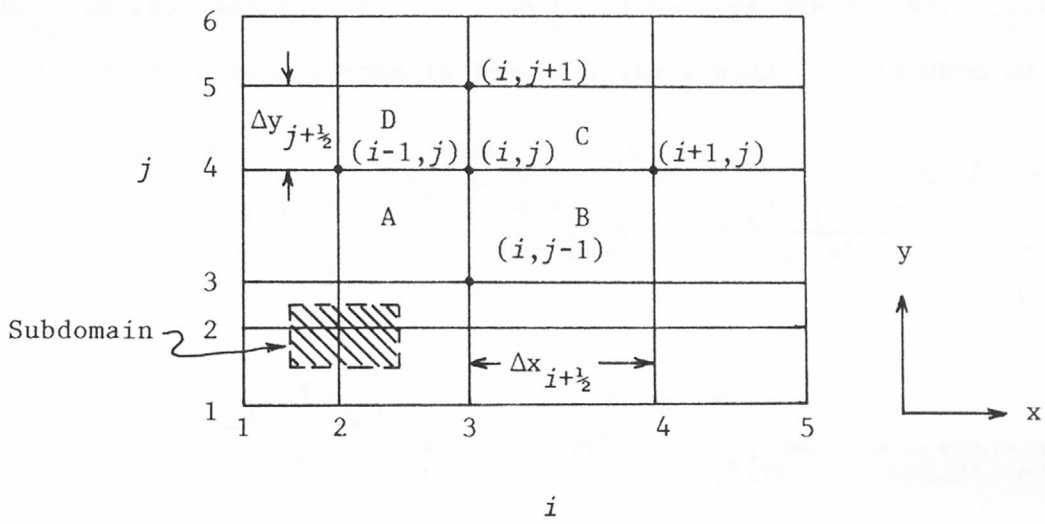


fig. 4.3-1

The coordinates of a typical node (i,j) are given as (x_i,y_j). With the nomenclature shown in figure 4.3-1, (4.1-1) can be integrated over a subdomain enclosing node (i, j) to produce

$$\int_{\Delta y_j} (T_{xx} \frac{\partial h}{\partial x})_{i+\frac{1}{2}} dy - \int_{\Delta y_j} (T_{xx} \frac{\partial h}{\partial x})_{i-\frac{1}{2}} dy$$

$$+ \int_{\Delta x_i} (T_{yy} \frac{\partial h}{\partial y})_{j+\frac{1}{2}} dx - \int_{\Delta x_i} (T_{yy} \frac{\partial h}{\partial y})_{j-\frac{1}{2}} dx$$

$$+ \int_{\Delta x_i}\int_{\Delta y_j} R(H - h)dxdy + \int_{\Delta x_i}\int_{\Delta y_j} Wdxdy + \sum_{p=1}^{p'} Q_p = 0 \qquad (4.3\text{-}1)$$

where

$$\left.\begin{aligned} \Delta x_i &= \tfrac{1}{2}(\Delta x_{i+\frac{1}{2}} + \Delta x_{i-\frac{1}{2}}) \\ \Delta y_j &= \tfrac{1}{2}(\Delta y_{j+\frac{1}{2}} + \Delta y_{j-\frac{1}{2}}) \end{aligned}\right\} \quad (4.3\text{-}2)$$

and p' is the number of pumping wells in subdomain $\Delta x_i \Delta y_j$.

If $T_{\xi\xi}$, R, and W are assumed to be constant in each cell (A, B, C, D) adjacent to node (i, j), then a valid numerical approximation of (4.3-1) is

$$T_{xxi+\frac{1}{2},j}\Delta y_j \frac{h_{i+1,j} - h_{i,j}}{\Delta x_{i+\frac{1}{2}}} - T_{xxi-\frac{1}{2},j}\Delta y_j \frac{h_{i,j} - h_{i-1,j}}{\Delta x_{i-\frac{1}{2}}}$$

$$+ T_{yyi,j+\frac{1}{2}} \Delta x_i \frac{h_{i,j+1} - h_{i,j}}{\Delta y_{j+\frac{1}{2}}} - T_{yyi,j-\frac{1}{2}} \Delta x_i \frac{h_{i,j} - h_{i,j-1}}{\Delta y_{j-\frac{1}{2}}}$$

$$+ R_{i,j}\Delta x_i \Delta y_j (H_{i,j} - h_{i,j}) + W_{i,j}\Delta x_i \Delta y_j + Q_{i,j} = 0 \quad (4.3\text{-}3)$$

where

$$T_{xxi+\frac{1}{2},j} = \frac{\Delta y_{j-\frac{1}{2}} T_{xxB} + \Delta y_{j+\frac{1}{2}} T_{xxC}}{2\Delta y_j} \quad (4.3\text{-}4)$$

$$T_{xxi-\frac{1}{2},j} = \frac{\Delta y_{j-\frac{1}{2}} T_{xxA} + \Delta y_{j+\frac{1}{2}} T_{xxD}}{2\Delta y_j} \quad (4.3\text{-}5)$$

$$T_{yyi,j+\frac{1}{2}} = \frac{\Delta x_{i-\frac{1}{2}} T_{yyD} + \Delta x_{i+\frac{1}{2}} T_{yyC}}{2\Delta x_i} \tag{4.3-6}$$

$$T_{yyi,j-\frac{1}{2}} = \frac{\Delta x_{i-\frac{1}{2}} T_{yyA} + \Delta x_{i+\frac{1}{2}} T_{yyB}}{2\Delta x_i} \tag{4.3-7}$$

$$R_{i,j} = \frac{\Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} R_A + \Delta x_{i+\frac{1}{2}} \Delta y_{j-\frac{1}{2}} R_B + \Delta x_{i+\frac{1}{2}} \Delta y_{j+\frac{1}{2}} R_C + \Delta x_{i-\frac{1}{2}} \Delta y_{j+\frac{1}{2}} R_D}{4\Delta x_i \Delta y_j} \tag{4.3-8}$$

$$W_{i,j} = \frac{\Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} W_A + \Delta x_{i+\frac{1}{2}} \Delta y_{j-\frac{1}{2}} W_B + \Delta x_{i+\frac{1}{2}} \Delta y_{j+\frac{1}{2}} W_C + \Delta x_{i-\frac{1}{2}} \Delta y_{j+\frac{1}{2}} W_D}{4\Delta x_i \Delta y_j} \tag{4.3-9}$$

$$Q_{i,j} = \sum_{p'} Q_p \tag{4.3-10}$$

Because of the way that the cells are designated, all zone boundaries are assumed to pass through node points; for example, see figure 4.3-2.

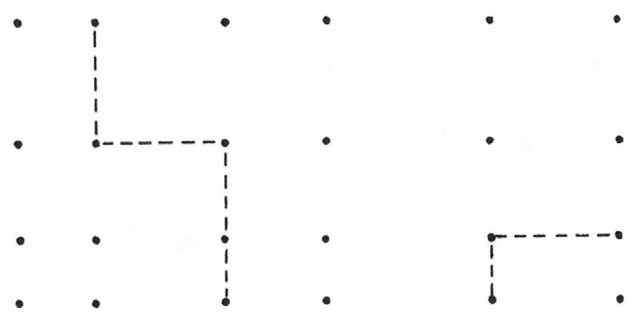


fig. 4.3-2

If the node points in (4.3-3) are designated as

$$k = i + NC\cdot(j - 1) \tag{4.3-11}$$

where NC is the number of columns (in the i direction), then the grid is renumbered as in figure 4.3-3.

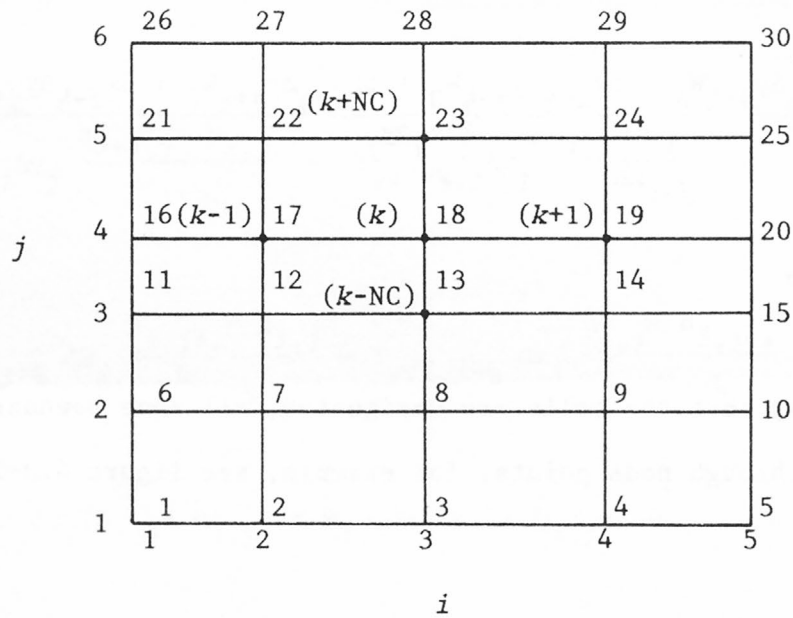


fig. 4.3-3

Equation (4.3-3) then becomes

$$T_{xxk,3}\Delta y_j \frac{h_{k+1} - h_k}{\Delta x_{i+\frac{1}{2}}} - T_{xxk,1}\Delta y_j \frac{h_k - h_{k-1}}{\Delta x_{i-\frac{1}{2}}}$$

$$+ T_{yyk,4}\Delta x_i \frac{h_{k+NC} - h_k}{\Delta y_{j+\frac{1}{2}}} - T_{yyk,2}\Delta x_i \frac{h_k - h_{k-NC}}{\Delta y_{j-\frac{1}{2}}}$$

$$+ R_k\Delta x_i\Delta y_j(H_k - h_k) + W_k\Delta x_i\Delta y_j + Q_k = 0 \qquad (4.3\text{-}12)$$

where

$$T_{xxk,1} \equiv T_{xxi-\frac{1}{2},j}\ ,\ T_{yyk,2} \equiv T_{yyi,j-\frac{1}{2}},\ T_{xxk,3} \equiv T_{xxi+\frac{1}{2},j},$$

$$T_{yyk,4} \equiv T_{yyi,j+\frac{1}{2}}.$$

In matrix form the numerical solution is

$$\underline{D}\ \underline{h} = \underline{q} \qquad (4.3\text{-}13)$$

where, from (4.3-12), for node k not on a specified head boundary

$$D_{k,k-NC} = - T_{yyk,2} \frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}} \qquad (4.3\text{-}14)$$

$$D_{k,k-1} = - T_{xxk,1} \frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} \qquad (4.3\text{-}15)$$

$$D_{k,k} = T_{xxk,3}\frac{\Delta y_j}{\Delta x_{i+\frac{1}{2}}} + T_{yyk,4}\frac{\Delta x_i}{\Delta y_{j+\frac{1}{2}}} + T_{xxk,1}\frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} + T_{yyk,2}\frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}}$$

$$+ R_k \Delta x_i \Delta y_j \qquad (4.3\text{-}16)$$

$$D_{k,k+1} = - T_{xxk,3}\frac{\Delta y_j}{\Delta x_{i+\frac{1}{2}}} \qquad (4.3\text{-}17)$$

$$D_{k,k+NC} = - T_{yyk,4}\frac{\Delta x_i}{\Delta y_{j+\frac{1}{2}}} \qquad (4.3\text{-}18)$$

$$q_k = R_k \Delta x_i \Delta y_j H_k + W_k \Delta x_i \Delta y_j + Q_k \qquad (4.3\text{-}19)$$

For node k on a specified head boundary, $D_{k,k-NC} = D_{k,k-1} = D_{k,k+1} = D_{k,k+NC} = 0$, $D_{k,k} = 1$, and $q_k = h_{Bk}$, the specified head. All remaining $D_{k,\ell} = 0$ for equation k in both cases. To preserve symmetry of $\underline{D}$, equations ℓ, $\ell \neq k$, are modified as indicated just after (4.1-4).

The flow across specified flow boundaries is incorporated by using the Q_k term, so that the total flow crossing the specified flow boundary of the subdomain around node k is added into Q_k. If $Q_k = 0$ on a boundary node and the head at the node is not specified, then the boundary for the node is automatically a no-flow type. When computing the total flow to add into a specified flow node, it must be remembered that nodes are on boundaries so that subdomains for boundary nodes are only fractions of the full subdomains. For example, see figure 4.3-4.

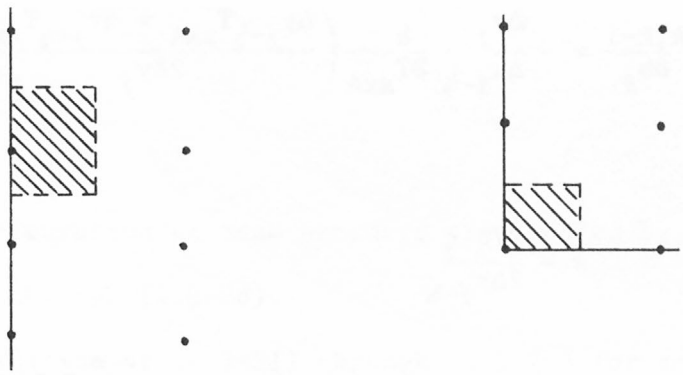


fig. 4.3-4

4.3.2. Computation of Sensitivities for the Integrated Finite Difference Model

Partial derivatives $\underline{J}_\ell$, defined by

$$\underline{J}_\ell = \frac{\partial \underline{q}}{\partial b_\ell} - \frac{\partial \underline{D}}{\partial b_\ell}\,\underline{h},\ \ell=1,\ 2,\ \ldots,\ p \qquad (4.3\text{-}20)$$

are employed in (3.3-25) to compute sensitivities for the nonlinear regression solution of the numerical model given in 4.3.1. Using the definitions of the elements of $\underline{D}$ and $\underline{q}$ given in 4.3.1, partial derivatives for the various parameters are computed as follows:

1) $b_\ell \equiv T_{xxA}$ at node k

$$\frac{\partial D_{k,k-NC}}{\partial b_\ell} = 0 \tag{4.3-21}$$

$$\frac{\partial D_{k,k-1}}{\partial b_\ell} = -\frac{\Delta y_j}{\Delta x_{i-\frac{1}{2}}} \frac{\partial}{\partial T_{xxA}} \left(\frac{\Delta y_{j-\frac{1}{2}} T_{xxA} + \Delta y_{j+\frac{1}{2}} T_{xxD}}{2\Delta y_j} \right)$$

$$= -\frac{\Delta y_{j-\frac{1}{2}}}{2\Delta x_{i-\frac{1}{2}}} \tag{4.3-22}$$

$$\frac{\partial D_{k,k}}{\partial b_\ell} = \frac{\Delta y_{j-\frac{1}{2}}}{2\Delta x_{i-\frac{1}{2}}} \tag{4.3-23}$$

$$\frac{\partial D_{k,k+1}}{\partial b_\ell} = 0 \tag{4.3-24}$$

$$\frac{\partial D_{k,k+NC}}{\partial b_\ell} = 0 \tag{4.3-25}$$

Similar expressions result for $b_\ell \equiv T_{xxB}$, T_{xxC}, and T_{xxD}. If more than one of the cells A, ..., D lie in the same zone, then derivatives for the individual expressions are summed to form the final value. For example, if $b_\ell \equiv T_{yyA} = T_{yyB} = T_{yyC} = T_{yyD}$, then

$$\frac{\partial D_{k,k\text{-NC}}}{\partial b_{\ell}} = \frac{\partial D_{k,k\text{-NC}}}{\partial T_{yyA}} + \frac{\partial D_{k,k\text{-NC}}}{\partial T_{yyB}} + \frac{\partial D_{k,k\text{-NC}}}{\partial T_{yyC}} + \frac{\partial D_{k,k\text{-NC}}}{\partial T_{yyD}}$$

$$= -\frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}}\left(\frac{\Delta x_{i-\frac{1}{2}}}{2\Delta x_i} + \frac{\Delta x_{i+\frac{1}{2}}}{2\Delta x_i} + 0 + 0\right)$$

$$= -\frac{\Delta x_i}{\Delta y_{j-\frac{1}{2}}} \qquad (4.3\text{-}26)$$

Derivatives for any configuration of zone boundary are handled by combinations of the form of (4.3-21) through (4.3-26).

An example of application of (4.3-21) through (4.3-26) for an irregular zone boundary is given in figure 4.3-5. Let $T_{xx} = T_{yy} = T$ for simplicity. Then for a two zone problem, the two transmissivities are T_1 and T_2.

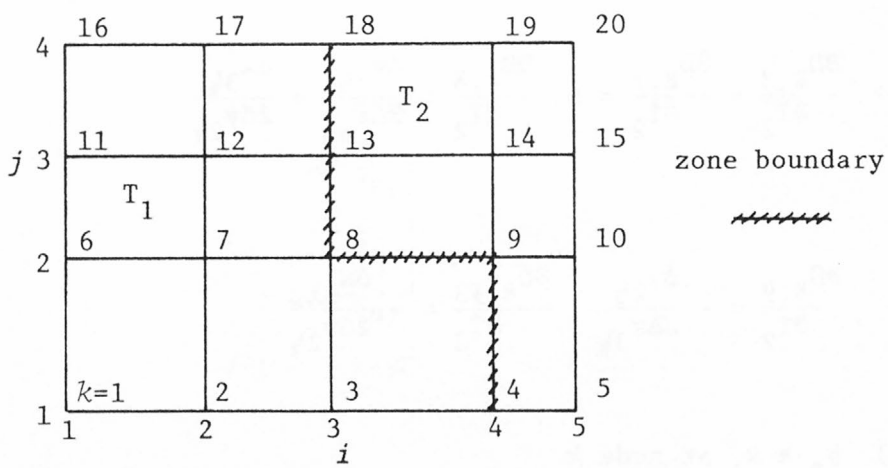


fig. 4.3-5

For node 7:

$$\frac{\partial D_{7,2}}{\partial T_1} = -\frac{\Delta x_2}{\Delta y_{1\frac{1}{2}}}, \quad \frac{\partial D_{7,6}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{1\frac{1}{2}}}, \quad \frac{\partial D_{7,7}}{\partial T_1} = \frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} + \frac{\Delta x_2}{\Delta y_{2\frac{1}{2}}}$$

$$+ \frac{\Delta y_2}{\Delta x_{1\frac{1}{2}}} + \frac{\Delta x_2}{\Delta y_{1\frac{1}{2}}}, \quad \frac{\partial D_{7,8}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}}, \quad \frac{\partial D_{7,12}}{\partial T_1} = -\frac{\Delta x_2}{\Delta y_{2\frac{1}{2}}}$$

For node 7 the derivatives of the D_{ij}'s with respect to T_2 are all zero because T_2 does not appear in any of the D_{ij}'s. For node 8:

$$\frac{\partial D_{8,3}}{\partial T_1} = -\frac{\Delta x_3}{\Delta y_{1\frac{1}{2}}}, \quad \frac{\partial D_{8,7}}{\partial T_1} = -\frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}}, \quad \frac{\partial D_{8,8}}{\partial T_1} = \frac{\Delta y_{1\frac{1}{2}}}{2\Delta x_{3\frac{1}{2}}} + \frac{\Delta x_{2\frac{1}{2}}}{2\Delta y_{2\frac{1}{2}}}$$

$$+ \frac{\Delta y_2}{\Delta x_{2\frac{1}{2}}} + \frac{\Delta x_3}{\Delta y_{1\frac{1}{2}}}, \quad \frac{\partial D_{8,9}}{\partial T_1} = -\frac{\Delta y_{1\frac{1}{2}}}{2\Delta x_{3\frac{1}{2}}}, \quad \frac{\partial D_{8,13}}{\partial T_1} = -\frac{\Delta x_{2\frac{1}{2}}}{2\Delta y_{2\frac{1}{2}}}$$

$$\frac{\partial D_{8,3}}{\partial T_2} = \frac{\partial D_{8,7}}{\partial T_2} = 0, \quad \frac{\partial D_{8,8}}{\partial T_2} = \frac{\Delta y_{2\frac{1}{2}}}{2\Delta x_{3\frac{1}{2}}} + \frac{\Delta x_{3\frac{1}{2}}}{2\Delta y_{2\frac{1}{2}}}$$

$$\frac{\partial D_{8,9}}{\partial T_2} = -\frac{\Delta y_{2\frac{1}{2}}}{2\Delta x_{3\frac{1}{2}}}, \quad \frac{\partial D_{8,13}}{\partial T_2} = -\frac{\Delta x_{3\frac{1}{2}}}{2\Delta y_{2\frac{1}{2}}}$$

2) $b_\ell \equiv R_A$ at node k

$$\frac{\partial D_{k,k-NC}}{\partial b_\ell} = \frac{\partial D_{k,k-1}}{\partial b_\ell} = \frac{\partial D_{k,k+1}}{\partial b_\ell} = \frac{\partial D_{k,k+NC}}{\partial b_\ell} = 0 \qquad (4.3\text{-}27)$$

$$\frac{\partial D_{k,k}}{\partial b_\ell} = \Delta x_i \Delta y_j \left(\frac{\Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}}}{4 \Delta x_i \Delta y_j} \right) = \tfrac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} \tag{4.3-28}$$

$$\frac{\partial q_k}{\partial b_\ell} = \tfrac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} \, H_k \tag{4.3-29}$$

Derivatives for R_B, R_C, and R_D are similar.

3) $b_\ell \equiv W_A$ at node k

$$\frac{\partial q_k}{\partial b_\ell} = \tfrac{1}{4} \Delta x_{i-\frac{1}{2}} \Delta y_{j-\frac{1}{2}} \tag{4.3-30}$$

All derivatives of D_{ij} are zero.

4) $b_\ell \equiv q_{B1}$, boundary flux 1 in Q_k, where, for example, $Q_k = q_{B2} \cdot \frac{1}{2} \Delta y_{j-\frac{1}{2}}$ + $q_{B1} \cdot \frac{1}{2} \Delta y_{j+\frac{1}{2}}$ (see figure 4.3-6).

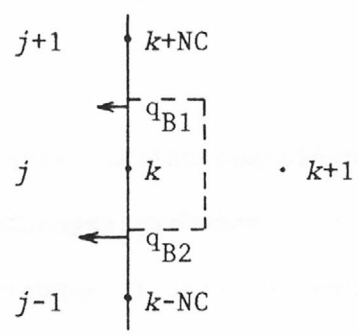


fig. 4.3-6

$$\frac{\partial q_k}{\partial b_\ell} = \tfrac{1}{2}\Delta y_{j+\frac{1}{2}} \tag{4.3-31}$$

A similar expression results for $b_\ell \equiv q_{B2}$, and if $q_{B1} = q_{B2}$, the derivative is the sum of two parts, (4.3-31) and its analog for q_{B2}.

5) $b_\ell \equiv H_s$ or H_t

The specified head at any node k along a specifed head boundary is given as

$$h_{Bk} = A_k[L_k H_t + (1 - L_k)H_s] \tag{4.3-32}$$

where

$$A_k = \frac{h^0_{Bk}}{L_k H^0_t + (1 - L_k)H^0_s} \tag{4.3-33}$$

and the meanings of the symbols are defined after (4.1-5).

If node m is adjacent to a boundary segment bounded by nodes s and t, and node k lies in the segment so that it appears in equation m, then for $b_\ell \equiv H_s$,

$$\frac{\partial q_m}{\partial b_\ell} = - D_{m,k} \frac{\partial h_{Bk}}{\partial H_s} = - D_{m,k} A_k (1 - L_k) \tag{4.3-34}$$

Similarly, for $b_\ell \equiv H_t$,

$$\frac{\partial q_m}{\partial b_\ell} = - D_{m,k} \frac{\partial h_{Bk}}{\partial H_t} = - D_{m,k} A_k L_k \tag{4.3-35}$$

If node k lies on the boundary, then the equation for node k in (4.3-13) becomes

$$h_k = h_{Bk}$$

$$= A_k[L_k H_t + (1 - L_k)H_s] \tag{4.3-36}$$

and, for $b_\ell \equiv H_s$,

$$\frac{\partial q_k}{\partial b_\ell} = A_k(1 - L_k) \tag{4.3-37}$$

and similarly for $b_\ell \equiv H_t$.

4.3.3. Derivation of (4.2-1)

By careful examination of (4.3-12) through (4.3-19) it will be seen that, if there are no specified head parameters, $\underline{q} - \underline{D}\,\underline{h} = \underline{0}$ can be written in the form

$$a_{i1}b_1 + a_{i2}b_2 + \ldots + a_{ip}b_p - Q_i = 0, \; i = 1, 2, \ldots, m \tag{4.3-38}$$

where

a_{ij} = coefficient containing Δx, Δy, and head differences;

b_j = any parameter <u>except</u> a specified head parameter; and

Q_i = term not containing parameters in $\underline{b}$.

Define

$$\underline{J}_j = \frac{\partial \underline{q}}{\partial b_j} - \frac{\partial \underline{D}}{\partial b_j} \underline{h} \tag{4.3-39}$$

Then, by carrying out the differentiations indicated in (4.3-39) and comparing the result with (4.3-38) it will be seen that

$$J_{ij} b_j = a_{ij} b_j \tag{4.3-40}$$

so that

$$\sum_{j=1}^{p} J_{ij} b_j = \sum_{j=1}^{p} a_{ij} b_j = Q_i \tag{4.3-41}$$

If $\underline{b}$ contains all possible parameters (except specified head parameters) and there are no known fluxes, then $Q_i = 0$ and

$$\sum_{j=1}^{p} J_{ij} b_j = 0 \tag{4.3-42}$$

or

$$\underline{Jb} = \underline{0} \tag{4.3-43}$$

4.3.4. Documentation of Program for Nonlinear Regression Solution of Steady-State Ground-Water Flow Problems

Introduction. This program is designed to obtain a nonlinear regression solution to the finite-difference model of steady-state ground-water flow given in section 4.3.1. Basic calculation methods are given in sections 4.1 and 4.3.2.

The computer program is composed of a main program and seven subroutines. The main program controls input-output and performs all computations that cannot be accomplished more effectively with subroutines. The seven subroutines (D4SOLV, COEF, LSTSQ, PRTOT, ORDER, ARRAY, ARRAYI) perform the following specialized tasks:

D4SOLV - Obtains an LDU factorization solution of the set of linear algebraic equations resulting from application of the finite difference methods, assuming the equations are ordered in an alternating diagonal fashion.

COEF - Computes coefficients necessary for the determination of scaled sensitivities and heads.

LSTSQ - Computes the coefficients of the normal equations and solves the system of equations to determine the vector of parameter changes.

PRTOT - Prints matrices or vectors in a column configuration.

ORDER - Computes equation numbers at grid points corresponding to the alternating diagonal ordering scheme.

ARRAY - Loads and (or) prints 1- and 2-dimensional real array variables.

ARRAYI - Loads and (or) prints 1- and 2-dimensional integer array variables.

The basic flow of the program can be described as follows:

A) Data are input and variables are initialized.

B) Using coefficients generated in COEF, an initial solution corresponding to the initial parameter estimates is computed by D4SOLV.

C) In an iterative fashion, the following four steps are taken until the regression technique converges or until the number of iterations exceeds the maximum allowed.

   1) Scaled sensitivities are calculated using coefficients computed in COEF and in the main program.

   2) LSTSQ is employed to form and solve the normal equations.

   3) Parameters are updated using the parameter change vector generated by LSTSQ.

   4) Various coefficients involving the updated parameters are computed in COEF, and current estimates of head are computed using D4SOLV.

D) Various statistics associated with the regression analysis are computed.

Aquifer Property Zonation and Variable Definition. Basic model geometry is defined by the finite difference grid that is constructed over the region to be modeled. Nodes, consisting of grid intersections, are numbered from the lower left-hand corner of the grid (columns from left to right and rows from bottom to top). Cells, consisting of intragrid areas bounded by four adjacent nodes, are numbered similarly (see figure 4.3-7).

The finite difference grid is divided into aquifer property zones, which define zonal values for the aquifer properties T_{xx}, T_{yy}, R, and W. Each zonal

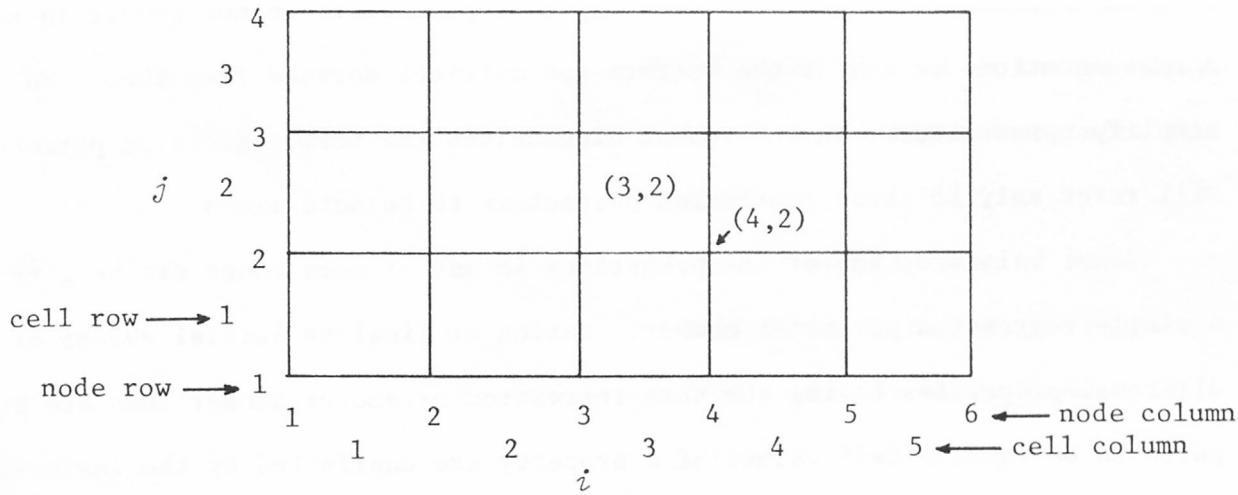


fig. 4.3-7

value is constant within the zone. Variation of a property within a zone is accomplished by assigning cell values. The aquifer property at any particular cell is computed as the product of the zonal value and the cell value of the property. Thus, if all cell values for a property within a zone are given a value of unity, the zonal value becomes the value of the property for each cell within that zone.

Zones are created by subdividing the grid into groups of cells having distinct combinations of zonal properties (TRANX, TRANY, VLEAK, and QDIST). Cells belonging to these zones are accordingly assigned distinct zone numbers (IZN), reserving IZN=0 to indicate groups of nodes outside of the model area.

All zonal values are either regarded as regression parameters to be determined by the procedure or are held constant and, thus, are not regarded as regression parameters to be determined, as specified in the input. From a conceptual viewpoint there is no difference between these two designations because zonal values that are held constant can be regarded as regression

parameters having exact prior information. However, from a computational viewpoint it is most efficient to eliminate parameters being held constant from the calculations. Hence, these types of parameters do not appear in the normal equations or any of the vectors and matrices derived from them. To simplify nomenclature, in subsequent discussions the term regression parameter will refer only to those regression parameters to be determined.

Zonal values of any of the properties in one or more zones can be given a single regression parameter number. Ratios of final to initial values of all zonal properties having the same regression parameter number then are computed to be equal. Cell values of a property are unaffected by the regression procedure. A high degree of flexibility for distributing aquifer properties while, at the same time, keeping the number of regression parameters to a minimum is achieved with the above schemes.

Cell values for an aquifer property are input using rectangular blocks of cells, which are defined for each property for convenience of input only. These blocks need not bear any relationship to the zones. Cell values may be constant or variable within each block.

The definition of some of the more important variables related to aquifer properties in the computer program are given below.

Variable name	Definition
TRANX, TRANY, VLEAK, QDIST	Zonal value for transmissivity (T_{xx} and T_{yy}), hydraulic conductance (R), and distributed recharge (W), respectively.
IZN	An integer array that indicates the zone number of each cell.

T, SL, QRE	Cell value for transmissivity, hydraulic conductance, and distributed recharge, respectively.
HR	Nodal value of head on the boundary of the confining bed opposite the aquifer.
WELL	Nodal (or point) value of known volumetric discharge (or recharge) from a well or other known-rate source-sink phenomena.
HO	Array of observed heads and known (or fixed) specified heads.
QBND	Zonal values of specified flow.
PLA, PLB	Arrays, the sum of which gives the specified head at any point along a segment where the specified heads at one (or both) end(s) of the segment is a (or are) regression parameters(s).

Boundary Conditions and Boundary Parameters. Two types of boundary conditions may be used: specified flow and (or) specified head. Non-zero specified-flow boundaries where the flow rate is known can be imposed by assigning the appropriate value of the specified flow rate to the nodel value of WELL. Because the no-flow boundary is the default condition, zeros do not have to be assigned to WELL to simulate this condition. Specified-head boundaries where the head is known can be imposed by assigning the value of the known head at the boundary nodes to variable HO, which also describes

nodal values of observed head. A negative one must be assigned to variable IN corresponding to each node that is to be considered a specified head node. Segments of the boundary that will be considered as either flow or head regression parameters must not be entered into WELL or HO arrays.

Different variables are used if specified-flow or specified-head boundary conditions are to be considered as regression parameters. Variable QBND is used to indicate the zonal value for discharge across groups of nodes that form a specified-flow segment (or zone). The specified volumetric flow per unit width for each cell boundary within a zone is the product of QBND for that zone and a multiplier for the cell boundary. By restricting a flow-boundary zone to a single node, point recharge and discharge can be simulated. QBND can be a regression parameter and, therefore, can be modified by the regression procedure; the multiplier is unaffected. If the variable IZ that specifies the regression parameter number for the boundary zone is set to zero, then QBND is held constant and the segment is treated as a known-flow boundary, thus giving two possible ways (via WELL and via QBND) to designate known-flow boundaries.

Specified-head boundary nodes that are to be considered regression parameters are defined by segments composed of a sequence of nodes (variables ILOC and JLOC) along portions of the boundary. The specified head of the first and last nodes in the sequence can be either different regression parameters or a single regression parameter or held constant, depending on the nature of the problem. Adjustments to these nodes computed by the regression procedure are apportioned to other nodes in the sequence. The proportion is the ratio of the distance (along the sequence of nodes) between the end node and the node of interest, to the distance between the two end nodes. These

factors are computed by the program for a given segment. As in the case with flow-boundary parameters, if the regression parameter number is set to zero, then that parameter is held constant. If the parameter numbers at both ends of the segment are set to zero, then the segment is treated as a known-head boundary, thus giving two possible ways to designate known-head boundaries. A negative one must be entered into IN for all nodes on the specified-head boundary, whether or not the boundary involves regression parameters.

Prior Information on Regression Parameters. If estimates of the regression parameters and their (less than infinite) reliability are available from other sources (for example, aquifer tests), it may be desirable to introduce this information into the regression analysis. For this case, initial values for the parameters are taken to be the prior information. The reliability of each estimate is represented by a coefficient of variation. Array WP is used to store these values for both aquifer regression parameters (TRANX, TRANY, VLEAK, and QDIST), specified-head regression parameters, and flow-boundary regression parameters (QBND). However, only the coefficients of variation for the aquifer regression parameters are read directly into WP. The coefficients of variation for boundary regression parameters are read in through temporary variables (CVQB for specified-flow regression parameters, and CVHA and CVHB for specified-head regression parameters) and are only subsequently placed into the WP array.

The use of prior information of known reliability requires an estimate of the error variance of the heads (variable, EV) computed using ordinary least squares. If the estimate differs substantially from the value computed by the analysis using prior information, the problem should be resolved using the latter computed value as the estimate of error variance.

In some instances prior information of unknown reliability may be available. Use of this type of information is an advanced topic and is not covered in this report. The papers by Cooley (1982, 1983) cover the method in detail. Variables RP and BP are used to input the additional information needed for this method.

Solution-Only Mode. To facilitate the calculation of certain statistical measures, the program is capable of bypassing the regression analysis and computing only head distributions for various combinations of parameter values. This is accomplished by specifying the solution only option (variable ISO) and providing the various combinations of parameter values for which solutions are desired. These solutions can be used to test the assumption of model linearity in the vicinity of the optimum parameter estimates.

Using the Program. The computer code has been designed to be as machine independent as possible. Also, to minimize confusion, all arrays have been dimensioned explicitly. The following list summarizes the minimum dimensions required for the program to operate properly for a specific problem. If

N_g is the number of grid points ($N_x \times N_y$),

N_x is the number of grid columns,

N_y is the number of grid rows,

N_e is the number of active nodes in the grid,

N_o is the number of observed heads at active nodes,

N_{oT} is the <u>total</u> number of observed heads plus the number of parameters on which there is prior information,

N_z is the number of aquifer property zones in the model grid,

N_p is the number of regression parameters for aquifer property zones,

N_q is the number of regression parameters for specified-flow boundaries,

N_{hs} is the number of specified-head zones (or segments),

N_h is the number of regression parameters for specified-head boundaries,

N_{qp} is the total number of nodes on a boundary where flow is a regression parameter,

N_{hp} is the total number of nodes along a boundary where head is composed of one or more specified-head regression parameters,

N_{mh} is the maximum number of nodes in any specified-head zone, and

N_R is the total number of regression parameters, $N_p + N_q + N_h$,

then the array variables should be dimensioned as in table 4.3-1.

Note that array variables that have a single dimension (TRANX, TRANY, T, HO, etc.) and are passed to subroutines are dimensioned as unity within the subroutines (only the initial address of an array is actually passed to a subroutine). This unit dimension should <u>not</u> be changed in subroutines when the dimension of the variable is changed in the main program. A similar system is used for multidimensional arrays not in common, and their dimensions within subroutines should not be changed either. Arrays AU and AL are passed in COMMON and must be dimensioned in all subroutines. To accompany any change in program dimensions, variable NVE defined near the beginning of the program must be set equal to N_R.

Table 4.3-1

Variable name	Dimension
WELL,HR,W,HO,HC,ILOC,JLOC,IN	N_g
T,SL,QRE,CXS,CYS,VLS,QRS,IZN	$(N_x-1)(N_y-1)$
DX	N_x
DY,JPOS	N_y
TRANX,TRANY,VLEAK,QDIST	N_z
IPRM	$4,N_z$
QBND	N_q
QBF	N_{qp}
PLA,PLB,IBHN	N_{hp} or N_{mh} whichever is larger
CXHR,CXHL,CYHT,CYHB	N_{hp}
IHSN	$2N_{hp}$
IBPA, IBPB	N_{hs}
IBZN,IBNA,IBNB	$N_{qp}+N_{hs}$
P,WP,NCB,NCE	N_R
S	N_R,N_o
X	N_R,N_{oT}
A	N_R,N_R
V	N_e or $4N_R$ whichever is larger
IC,AU	$5,N_e/2$
AL	$N_m,N_e/2$[1/]

[1/] These dimensions are approximate. The exact sizes required are calculated and printed in subroutine ORDER. N_m is N_x+3 or N_y+3 whichever is smaller.

Input Data.

Card Set A.

Three title cards of user's choice (format, 20A4).

Card Set B.

Problem size information; one card (format, 16I5).

Card columns	Variable	Definition
1-5	ID	Number of node columms.
6-10	JD	Number of node rows.
11-15	NZNS	Number of aquifer property zones.
16-20	NPAR	Number of regression parameters associated with aquifer property zones.
21-25	NWELS	Number of known point flows.
26-30	NQBND	Number of specified-flow boundary zones.
31-35	NBQP	Number of regression parameters associated with specified-flow zones.
36-40	NBHZ	Number of specified-head boundary zones.
41-45	NBHP	Number of regression parameters associated with specified-head zones.
46-50	NUM	Maximum number of iterations allowed for the regression analysis.

51-55	IPRX	Additional print sensitivites and orthogonalize sensitivities option. Code 1 to select the option.
56-60	IPO	Additional printout option. Code 1 to select the option.
61-65	ISO	Head-solutions only option. Code 1 to select the option.

Card Set C.

Special input parameters; one card (format, 8F10.0).

Card columns	Variable	Definition
1-10	BMX	Maximum fractional change, t_{mx}, allowed any regression parameter over any iteration.
11-20	CSA	Cosine, $\cos\theta_{mx}$, of the maximum angle allowed between the gradient direction and the search direction (normally set to 0.08).
21-30	RP	Ridge parameter for regression analysis using prior information of unknown reliability. Code 0.0 if not used.
31-40	BP	Bias parameter for regression using prior information of unknown reliability. Code 0.0 if not used.

41-50	EV	Estimated error variance for problems using prior information of known reliability. Code 0.0 if not used.

Card sets D through K.

A number of variables are input into the code by first subdividing the grid into rectangular regions (blocks) and then reading the variables for each block. Blocking can be applied to either cells or nodes, depending upon the variable being input. Blocking allows considerable flexibility in the input of certain variables and, once understood, can speed the construction of a model. Block and zone boundaries do not necessarily have to coincide; blocking is basically a convenient way of assigning variable values to every node or cell in the grid.

Card sets D through K represent the real (floating-point) variables subject to blocking. These variables, in the order they must appear, are listed below.

Card Set	Variable	Type of Variable	Definition
D	DX	Cell array	Distance between grid points in x or I direction.
E	DY	Cell array	Distance between grid points in y or J direction.
F	HO	Nodal array	Observed head.
G	W	Nodal array	Weighting values for observed heads. At all nodes where there is no observation W = 0; W ≠ 0

			not allowed on constant head boundaries where head is not a regression parameter.
H	T	Cell array	Multiplier (cell value) for transmissivity.
I	SL	Cell array	Multiplier (cell value) for hydraulic conductance of confining bed.
J	HR	Nodal array	Head on boundary of confining bed opposite the aquifer.
K	QRE	Cell array	Multiplier (cell value) for recharge rate per unit area.

Each card set D through K consists of an initial card defining the number of blocks (NOBL) into which the grid has been subdivided, and then a subsequent card or set of cards that define the blocks and the value or values of the variable to be input. The initial card, read with a 2I5 format, has the following form:

Card columns	Variable	Definition
1-5	NOBL	Number of rectangular input blocks into which the variable has been subdivided.
6-10	IPRN	Print option for full array. Set to 0 for print. Set to 1 for no output.

The initial card directs the program to seek NOBL blocks of information for a particular variable. If the variable is uniform over the block, then a single card suffices to define the block location and the uniform value to be assigned to every node or cell. If the variable is nonuniform over the block, then by specifying a value of IVAR equal unity on this card the program can be directed to seek additional cards specifying values of the variable for each node or cell in the block This information is input through the format 4I5, F10.0, I5 as follows:

Card columns	Variable	Definition
1-5	IB	Beginning column of the rectangular input block.
6-10	IE	Final column of the rectangular input block.
11-15	JB	Beginning row of the rectangular input block.
16-20	JE	Final row of the rectangular input block.
21-30	FACT	If the array set is uniform for the entire block, FACT is the cell or nodal value that is assigned to each element. If the array set is not uniform, each cell or nodal value on the subsequent data cards will be multiplied by FACT.

31-35	IVAR	Code 0 if the array set is uniform. Code 1 if it is not uniform.

If a value of IVAR equal to unity is specified, then the program will seek subsequent node or cell data sufficient to define the variable at every node or cell in the block. This information is input through the temporary variable A(I,J) with an 8F10.0 format in the following manner:

Card columns	Variable	Definition
1-10	A(IB,JB)	Temporary variable specifying nodal or cell values for the arrays in card sets D through K. Note that the information is read-in row by row for the grid, each new row beginning a new card.
11-20	A(IB+1,JB)	
⋮	⋮	
	A(IE,JB)	
1-10	A(IB,JB+1)	
⋮	A(IB+1,JB+1)	
	⋮	
	A(IE,JB+1)	
	⋮	
	A(IE,JE)	

Any variable that is not defined by blocking over a particular part of the grid will be automatically set to zero on that part.

Card sets D and E, representing the internal spatial dimensions of the grid, are read in by blocking for convenience to the programmer. Because both the horizontal spacing DX and vertical spacing DY (as measured from the lower left-hand corner of the grid) are, in reality, singly dimensioned arrays, it is necessary to set JB and JE equal to unity for both variables. The variable IE then equals ID-1 in the case of DX, and JD-1 in the case of DY

(IB equals one, of course, in both cases). Variable grid spacing can be input by specifying a value of IVAR equal to unity and following this with the necessary array information in an 8F10.0 format.

Note that only cell values of transmissivity, hydraulic conductance, and recharge (Card sets H, I, and K) are read in through the blocking scheme. Card M contains the zonal values by which these cell values are multiplied. A typical example of usage would be to form transmissivity as the product of hydraulic conductivity and thickness. Card set H would contain the variable thickness of the aquifer, and the variable on card M would represent the hydraulic conductivity zone by zone. Their product would be the transmissivity.

Card Set L.

Two integer variables also are input by blocking. Both are defined below, as read in by a 16I5 format, although only that variable associated with card set L is input at this location.

Card Set	Variable	Type of Variable	Definition
L	IZN	Cell array	Zone number of each cell. Each cell having a nonzero zone number must have T > 0.
R	IN	Nodal array	Denotes specified head. Set to -1 at nodes where head is specified, including nodes in segments involving specified-head regression parameters, and leave as zero at all remaining nodes.

The initial card for these data is identical to that of the real variable case. The card defining blocks into which integer variables are divided is similar to that of the real variable case with the exception of the variable IFACT, as noted subsequently (format, 6I5):

Card columns	Variable	Definition
1-5	IB	Beginning column of the rectangular input block.
6-10	IE	Final column of the rectangular input block.
11-15	JB	Beginning row of the rectangular input block.
16-20	JE	Final row of the rectangular input block.
21-25	IFACT	If the array set is uniform for the entire block, IFACT is the cell or nodal value that is assigned to each grid point. If the array set is not uniform, each cell or nodal value on the subsequent data cards will not be multiplied by IFACT.
26-30	IVAR	Code 0 if the array set is uniform. Code 1 if it is not uniform.

The nonuniform integer input is identical to the real variable case, except that the temporary variable INT(I,J) input with format 16I5 is used in place of A(I,J).

Card Set M.

Zonal aquifer property values; set contains NZNS cards (format, I5, 4F10.0).

Card columns	Variable	Definitions
1-5	I	Zone number.
6-15	TRANX(I)	Zonal x-transmissivity value for zone I.
16-25	TRANY(I)	Zonal y-transmissivity value for zone I.
26-35	VLEAK(I)	Zonal hyraulic conductance value for zone I.
36-45	QDIST(I)	Zonal distributed recharge value for zone I. (QDIST*QRE has units of volumetric rate per unit area.)

Cards may appear in any order with respect to zone number, but there must be NZNS cards.

Card Set N.

Aquifer regression parameter numbers; set contains NZNS cards (format, 16I5).

Card columns	Variable	Definition
1-5	I	Zone number.
6-10	IPRM(1,I)	Parameter number of x-transmissivity in zone I. Code 0 if it is not a regression parameter.

11-15	IPRM(2,I)	Parameter number of y-transmissivity in zone I. Code 0 if it is not a regression parameter.
16-20	IPRM(3,I)	Parameter number for hydraulic conductance in zone I. Code 0 if it is not a regression parameter.
21-25	IPRM(4,I)	Parameter number for distributed recharge in zone I. Code 0 if it is not a regression parameter.

Cards may appear in any order with respect to zone number, but there must be NZNS cards. Parameters must be numbered 1 through NPAR. Note that parameters of the same (or even different) property in different zones may have identical parameter numbers. Omit this card set if NPAR equals zero.

Card Set 0.

Coefficients of variation for aquifer regression parameters; eight values per card, for a total of NPAR values (format, 8F10.0).

Card columns	Variable	Definition
1-10	WP(1)	Coefficient of variation of each aquifer regression parameter. Enter values for aquifer regression parameters defined by IPRM in increasing order. Code 0.0 if no prior information exists for the parameter.
11-20	WP(2)	
.	.	
.	.	
.	.	
	WP(NPAR)	

Omit card set if NPAR equals zero.

Card Set P.

Known point flow rates; set contains NWELS cards (format, 2I5, F10.0).

Card columns	Variable	Definition
1-5	I	Column location of point flow.
6-10	J	Row location of point flow.
11-20	WELL(I,J)	Total volumetric flow to or from node, negative for withdrawal.

Omit card set if NWELS equals zero.

Card Set Q.

Specified boundary-flow zones and flow-zone parameters; set contains NQBND cards, one card for each zone (format, 5I5, 3F10.0).

Card columns	Variable	Format	Definition
1-5	IA	I5	Column location of the A end of the segment (zone).
6-10	JA	I5	Row location of the A end of the segment (zone).
11-15	IB	I5	Column location of the B end of segment (zone).
16-20	JB	I5	Row location of the B end of segment (zone).
21-25	IZ	I5	Regression parameter number. Set equal to zero if QBND is not a regression parameter.
26-35	QBND	F10.0	Zonal flow value.

36-45	CVQB	F10.0	Coefficient of variation for regression parameter. Code as 0.0 if no prior information exists on the parameter.
46-55	QBM	F10.0	Multiplier for zonal flow value.

If IA equals IB and JA equals JB, the flow is restricted to a single node. In this case, the product QBND*QBM equals total volumetric flow into or out of the node. Otherwise, the product is a volumetric rate per unit cell width. When card set is used to model flow boundary conditions, zone must follow either a row or column. Regression parameters are numbered from 1 to NBQP. Note that by setting IZ equal to zero, a fixed specified-flow condition is simulated. Omit card set when NQBND equals zero.

Card Set R.

Specified-head boundary designation; see card set L. This card set, when used in conjunction with card set F or S (following), can be used to construct specified-head boundaries. In particular, nodes designated in this card set by -1 are forced to take on values specified in card set F or S. In addition to peripheral boundary conditions, this card set, in conjunction with card set F or S, can be used to model other constant head conditions such as bodies of open water.

Card Set S.

This card set defines specified-head boundary parameter zones. Each zone is associated with a subset of the S cards and the number of subsets will

equal NBHZ. The initial card in each subset contains size and descriptive information about the zone and appears as follows (format, 4I5, 2F10.0):

Card columns	Variable	Definition
1-5	IZ	Segment or zone number.
6-10	NN	Number of nodes in the segment.
11-15	IBPA	Regression parameter number for A end of segment. Set equal to zero if head at the A end is not a regression parameter.
16-20	IBPB	Regression parameter number for B end of segment. Set equal to zero if head at the B end is not a regression parameter.
21-30	CVHA	Coefficient of variation of head at A end of segment. Code as 0.0 if no prior information exists on the parameter.
31-40	CVHB	Coefficient of variation of head at B end of segment. Code as 0.0 if no prior information exists on the parameter.

Segments are numbered from 1 through NBHZ and regression parameters are numbered 1 through NBHP. A single head change can be found for all the intermediate head nodes by allowing IBPA to equal IBPB.

After reading the initial card, the program then seeks NN subsequent cards in each subset that define the heads along the boundary segment (format, 2I5, F10.0).

Card columns	Variable	Definition
1-5	ILOC	Column location of node.
6-10	JLOC	Row location of node.
11-20	V	Estimated head at node ILOC, JLOC.

Note that any shape of head surface can be input along a boundary segment. Since heads at the segment ends are the only regression parameters in each segment, their influence is distributed to the intermediate nodes by linear interpolation. The linear interpolation is based upon distance from the parameter in question, with a weight of one assigned at the end node occupied by the parameter and zero at the node of the opposite end of the segment. If zeros are assigned to the head variables V at intermediate nodes in a boundary segment, then the program automatically assumes that the head surface along the segment is simply a straight line between the heads specified at the end nodes A and B. Omit the entire card set if NBHZ equals zero.

The following card sets are required only if the solution-only option (ISO, Card set B) is specified. In the following descriptions input variables may be loosely termed parameters for convenience, but it must be realized that they are not actually regression parameters because no regression is performed. Also, card sets V and W require that the parameters being varied have nonzero numbers corresponding to regression parameter numbers defined for the initial solution. Hence, the input for the initial solution must be coded as if it were to be a regression for these parameters although no regression will actually be performed.

Card Set T.

Additional solution specification, one card (format, I5).

Card columns	Variable	Definition
1-5	N	Number of solutions required using alternative parameter sets. Code 0 if a solution is desired only for the initial set of parameters.

Card Set U.

Zonal aquifer properties; the set contains NZNS cards (format, I5, 4F10.0).

Card columns	Variable	Definition
1-5	I	Zone number.
6-15	TRANX(I)	Zonal x-transmissivity value for zone I.
16-25	TRANY(I)	Zonal y-transmissivity value for zone I.
26-35	VLEAK(I)	Zonal hydraulic conductance value for zone I.
36-45	QDIST(I)	Zonal distributed recharge value for zone I.

Note that the zones may appear in any order but all zones must be defined.

Card Set V.

Specified-flow boundary values; the set contains NBQP cards (format, I5, F10.0).

Card columns	Variable	Definition
1-5	I	Specified-flow boundary zone number. This number must be nonzero and must correspond to IZ in card set Q.
6-15	QBND(I)	Zonal specified-flow boundary value for zone I.

Note that the zones may appear in any order but all zones must be defined if NBQP is not zero. Omit card set if NBQP is zero.

Card Set W.

Specified-head boundary parameters; set contains NBHZ cards (format, I5, F10.0).

Card columns	Variable	Definition
1-5	I	Specified-head boundary zone number. This number must correspond to IZ in card set S.
6-15	HBA	Specified head at the A end of the zone.
16-25	HBB	Specified head at the B end of the zone.

Note that the parameters may appear in any order but all parameters must be defined if NBHZ is not zero. Omit card set if NBHZ is zero.

Card sets U, V, and W will be repeated N times. The program will compute and print the solution corresponding to each set of parameters.

Output. The following discussion gives the content and order of the output obtained from the program. It should be noted that some of the output is only obtained under certain specified conditions. All output is clearly labeled. However, order numbers in the following discussion do not appear in the output; they are for convenience in listing the order of output only. The statistical measures cited below are described in section 5.

1. Three title cards (card set A).
2. Problem size information (card set B).
3. Special input parameters (card set C).
4. Array sets (card sets D through L). For each variable, input information (block number, IB, IE, JB, JE, variable name and value of FACT or IFACT) for each block is listed block by block. If specified, this information is then followed by the values of all entries in the array.
5. Zonal aquifer properties (card set M).
6. Aquifer regression parameter numbers (card set N). This is printed only if NPAR is greater than zero.
7. Coefficients of variation for aquifer regression parameters (card set O). This is printed only if NPAR is greater than zero.
8. Known point-flow rates (card set P). This is printed only if NWELS is greater than zero.
9. Specified boundary-flow information (card set Q). This is printed only if NQBND is greater than zero.
10. Specified boundary-head distribution (card set R). This is printed in the same manner as indicated under 4.
11. Specified boundary-head information (card set S). For each segment, input information (segment number, number of nodes in the segment, IBPA, IBPB, CVHA, and CVHB) is printed followed by a listing of node

locations and input values of specified heads. This is printed only if NBHZ is greater than zero.

12. Error message. If IZN>1 and T≤0 occur at the same cell, then the following error message is printed: "AT CELL *i*,*j*, IZN>0 AND T = 0."

13. Error message. If the error in 12 happens at one or more cells, then the following message is printed: "PROGRAM ABORTED BECAUSE OF CONFLICT BETWEEN IZN AND T." Execution then terminates.

14. Error message. If an active node is isolated from other active nodes, then the following message is printed: "NODE *i*,*j* ISOLATED, SOLUTION SINGULAR."

15. Information on matrix solution procedure used to compute heads. Message that solution is by LDU factorization is followed by the computed minimum required dimensions of arrays used in the solution.

16. If the error in 14 occurred, then execution terminates.

17. Initial solution for heads.

18. Program branch. If the solution-only option is specified, then output skips to 36.

19. Number of observations, number of parameters having prior information, and initial sum of weighted, squared deviations of computed from observed heads.

20. Scaled sensitivity matrix $\underline{Z}$. This is printed only if optional print-out was selected.

21. Error message. If any diagonal term of the coefficient matrix $\underline{S}_s^T \underline{V}_s^{-1} \underline{S}_s + \underline{S}_p^T \underline{U}^{-1} \underline{S}_p s^2$ of the normal equations is smaller than 10^{-10}, then the following message is printed: "SENSITIVITIES FOR PARAMETER *i* EFFECTIVELY ZERO." If this error occurs and if the sensitivity print

and orthogonalization option was selected, then the sensitivities and orthogonalized sensitivities are printed in the forms given in 34 and 35. Execution then terminates.

22. Coefficient matrix $\underline{S}_s^T \underline{V}_s^{-1} \underline{S}_s + \underline{S}_p^T \underline{U}_p^{-1} \underline{S}_p s^2 + \mu \underline{I}$ and the gradient vector $\underline{S}_s^T \underline{V}_s^{-1} (\underline{Y}_s - \underline{f}_s) + \underline{S}_p^T \underline{U}_p^{-1} s^2 (\underline{Y}_p - \underline{b}_p)$. This is printed only if the optional print-out was selected. (Note: If prior information of unknown relibility was used, then the matrix and vector will be modified to include this information.)

23. Error message. If, during solution of the normal equations, it becomes evident that the problem is singular, then the following message is printed: "LEAST SQUARES COEFFICIENT MATRIX SINGULAR: SOLUTION FOR PARAMETERS NOT UNIQUE." The course of action is then the same as that given under 21.

24. Iteration number, current sum of weighted, squared deviations of computed from observed heads, determinant of the least squares coefficient matrix defined in 22 above, followed by the current scaled parameter change or displacement vector $\underline{d}$ defined by (3.3-10). The ordering of entries in the vector is as follows: Boundary-flow regression parameters, listed in sequential order; boundary-head regression parameters, listed in sequential order; and zonal aquifer regression parameters, listed in sequential order.

25. Current estimates of parameters. This is printed only if optional print-out was selected.

26. Error message. If a parameter is more than one thousand times smaller in magnitude than initially specified, then, the message "PARAMETER *i*

EFFECTIVELY ZERO" is printed. If this problem occurs, then further iterations are aborted, and the current solution is taken as the final one. This action does not imply that the minimum sum of squares has been reached.

27. Solution converged message and final number of iterations. If the solution did not converge in the allotted number of iterations then a message to this effect is written instead.

28. Error message. If the coefficient matrix of the normal equations (see 22) is singular when $\mu = 0$, and this has not been detected because $\mu > 0$ has been computed and used by the program, then the message given in 23 is printed. In this case the subsequent course of action is the same as in 21.

29. Error variance (s^2) (see section 5.4.1), final total sum of squares (sum of weighted, squared deviations of computed from observed heads plus sum of weighted, squared deviations of computed from prior estimates of parameters), and correlation coefficient (R_y) (see section 5.4.2).

30. Final parameter estimates and their estimated standard errors (see section 5.4.3).

31. Estimated, scaled variance-covariance matrix, $(\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s + \underline{Z}_p^T \underline{U}^{-1} \underline{Z}_p s^2)^{-1} s^2$ (see section 5.4.3). The ordering of rows and columns matches that for $\underline{d}$ in 24.

32. Correlation matrix for scaled parameters, $\{r_{ij}\} = \{Cov(d_i,d_j)/(Var(d_i) \cdot Var(d_j)^{\frac{1}{2}}\}$ (see section 5.4.4). Again, the ordering matches that for $\underline{d}$. Entries of the correlation matrix for scaled parameters may have different signs than entries of an analogous correlation matrix for unscaled parameters because of the differing signs of scaled and unscaled parameters.

33. Computed and observed heads, and weighted residuals, $\hat{\underline{f}}$ (see section 5.5.1). Listed entries for weighted residuals not at observation points are set to zero to indicate that they cannot be computed at those points.

34. Scaled sensitivities, $\underline{Z}$, printed parameter by parameter. These are printed only if the sensitivity print and orthogonalization option was selected or if the solution did not converge.

35. Orthogonalized, scaled sensitivities, $\underline{Q}$. These are printed only if $\underline{Z}$ in 34 above is printed. This completes the output for regression solutions.

36. Zonal aquifer properties (card set T), boundary-flow parameters (card set U), and boundary-head parameters (card set V), all in sequential order; the solution number, and the solution for heads. These are printed for all solutions when the solution-only option is invoked. Boundary-flow parameters and boundary-head parameters are printed only if NBQP and NBHP are each nonzero.

Example Problem. The following example problem illustrates use of most of the program options. As illustrated in figure 4.3-8, the modeled area consists of three aquifer zones bounded by three specified boundary-flow

zones containing three boundary-flow regression parameters, a no-flow boundary, and two specified boundary-head segments containing three boundary-head regression parameters. Initial values for the regression parameters are:

$q_{B1} = 8$

$q_{B2} = 0.8$

$q_{B3} = 1$

$h_{B1} = 40$

$h_{B2} = 10$

$h_{B3} = 16$

$T_1 = 4000$

$T_2 = 400$

$T_3 = 1000$

$W_1 = 0.0001$

$W_2 = 0.0005$

$W_3 = -0.0001$

$R_1 = 0.001$

$R_2 = 0.0007$

$R_3 = 0.0015$

Parameter q_{B1} has prior information with a coefficient of variation of 0.1 on it. Because all aquifer properties and boundary flows are constant within their respective zones, multipliers for these parameters may be assigned values of unity. Assume that the estimated error variance for use with the prior information has a value of unity, $t_{mx} = 1.5$, and $\cos\theta_{mx} = 0.08$.

LINEAR VARIATION IN HEAD
LINEAR VARIATION IN HEAD
h_{B1}
h_{B2}
h_{B3}
q_{B1}
q_{B2}
q_{B3}
$q_B = 0$
1
2
3
1000
1000
i
j
x
y

fig. 4.3-8

Locations of observed heads are indicated by the small open circles on figure 4.3-8. Values corresponding to these locations are given in table 4.3-2.

Table 4.3-2

Node	Value	Node	Value
(1,1)	29.51	(10,4)	22.46
(3,1)	26.38	(12,4)	17.16
(5,1)	25.16	(14,4)	15.74
(7,1)	26.81	(16,4)	14.69
(9,1)	23.82	(1,5)	36.05
(11,1)	23.59	(3,5)	31.64
(13,1)	16.68	(5,5)	32.91
(15,1)	15.31	(7,5)	25.05
(2,2)	28.27	(9,5)	25.47
(4,2)	25.17	(11,5)	20.39
(6,2)	26.94	(13,5)	14.40
(8,2)	23.59	(15,5)	14.37
(10,2)	22.53	(2,6)	38.66
(12,2)	17.89	(4,6)	34.51
(14,2)	15.87	(6,6)	27.81
(16,2)	15.98	(8,6)	25.98
(1,3)	31.25	(10,6)	22.16
(3,3)	27.50	(12,6)	15.73
(5,3)	26.22	(14,6)	14.79
(7,3)	25.65	(16,6)	10.48
(9,3)	24.35	(1,7)	41.28
(11,3)	22.48	(3,7)	34.82
(13,3)	15.85	(5,7)	32.00

(15,3)	15.71	(7,7)	26.97
(2,4)	29.36	(9,7)	25.67
(4,4)	27.76	(11,7)	18.12
(6,4)	25.51	(13,7)	18.40
(8,4)	24.43	(15,7)	11.90

All observations have a weight w_{ii} of unity.

Heads H on the distal side of the aquitard are computed by first assigning constant values to all cells in an aquifer property zone, then computing nodal values as the average of all adjacent cell values. Cell values are 25 for zone 1, 35 for zone 2, and 15 for zone 3.

Input data for the example problem are coded on figure 4.3-9. These data should be compared with the card input instructions given above.

Output is given in figure 4.3-10. Statistical measures listed are described in section 5.

PROGRAMMER ______ DIVISION ______
LOCATION ______ PHONE ______

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ______
PROJECT ______

SHEET 1 OF 4

IDENTIFICATION: EXAMPLE PRØBLEM FØR CARD INPUT
INFORMATION:

```
               EXAMPLE PRØBLEM FØR CARD INPUT


   16    7    3    9    0    3    3    2    3   20    1    1    0
       1.5       .08         0         0         1
    1
    1   15    1    1      1000
    1
    1    6    1    1      1000
    1
    1   16    1    7         1    1
     29.51         0     26.38         0     25.16         0     26.81         0
     23.82         0     23.59         0     16.68         0     15.31         0
         0     28.27         0     25.17         0     26.94         0     23.59
         0     22.53         0     17.89         0     15.87         0     15.98
     31.25         0     27.50         0     26.22         0     25.65         0
     24.35         0     22.48         0     15.85         0     15.71         0
         0     29.36         0     27.76         0     25.51         0     24.43
         0     22.46         0     17.16         0     15.74         0     14.69
     36.05         0     31.64         0     32.91         0     25.05         0
     25.47         0     20.39         0     14.40         0     14.37         0
         0     38.66         0     34.51         0     27.81         0     25.98
         0     22.16         0     15.73         0     14.79         0     10.48
     41.28         0     34.82         0     32.00         0     26.97         0
     25.67         0     18.12         0     18.40         0     11.90         0
    1
    1   16    1    7         1    1
         1         0         1         0         1         0         1         0
         1         0         1         0         1         0         1         0
         0         1         0         1         0         1         0         1
```

FORM 9—1633 A
JULY 1967

0 = ZERO   Ø = ALPHA O   1 = ONE   I = ALPHA I   2 = TWO   Z = ALPHA Z   / = SLASH | = VERT. BAR   - = MINUS   _ = HORZ. BAR

DATA CODING FORM
GPO 16—88265 p-1

fig. 4.3-9

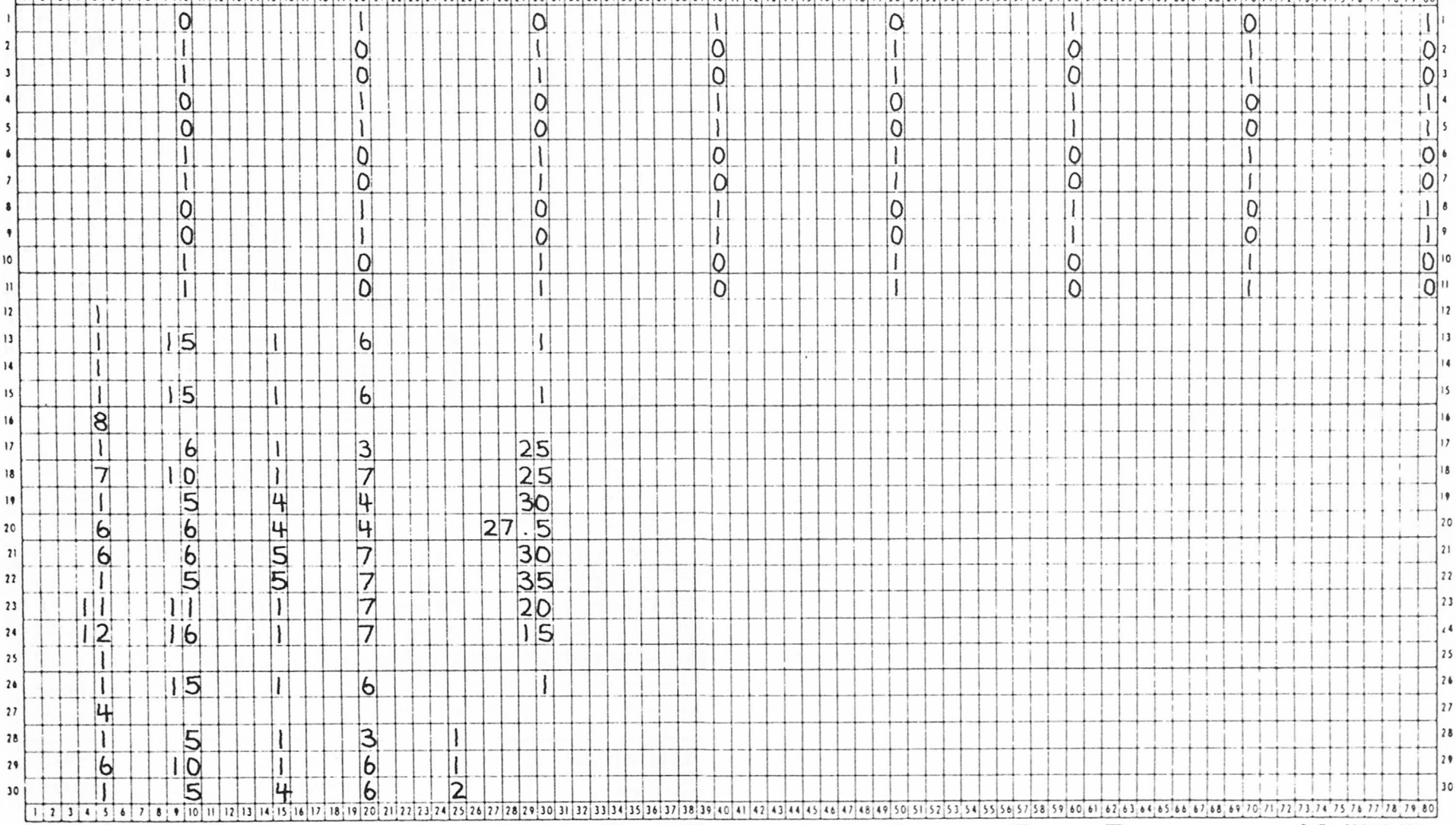
PROGRAMMER _____ DIVISION _____
LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

NUMBER _____
PROJECT _____

SHEET 2 OF 4

IDENTIFICATION: _____
INFORMATION: _____



FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Z = ALPHA Z / = SLASH | = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1

fig. 4.3-9 (Continued)

PROGRAMMER	DIVISION	U. S. DEPARTMENT OF THE INTERIOR GEOLOGICAL SURVEY **COMPUTER CODING FORM**	NUMBER	SHEET
LOCATION	PHONE		PROJECT	3 OF 4

IDENTIFICATION:

INFORMATION:

```
   11   15    1    6    3
    1      4000      4000      .001     .0001
    2       400       400     .0007     .0005
    3      1000      1000     .0015    -.0001
    1    1    1    4    7
    2    2    2    5    8
    3    3    3    6    9
         0         0         0         0         0         0         0         0
         0
    1    1    1    4    1         8        .1         1
    1    4    1    7    2        .8         0         1
    8    1   16    1    3         1         0         1
    2
   16   16    1    7   -1
    1   15    7    7   -1
    1   16    1    2         0         0
    1    7        40
    2    7
    3    7
    4    7
    5    7
    6    7
    7    7
    8    7
    9    7
   10    7
   11    7
   12    7
   13    7
   14    7
```

FORM 9—1633 A
JULY 1967

0 = ZERO   Ø = ALPHA O   1 = ONE   I = ALPHA I   2 = TWO   Ƶ = ALPHA Z   / = SLASH | = VERT. BAR   - = MINUS   _ = HORZ. BAR

DATA CODING FORM
GPO 16—88265 p-1

fig. 4.3-9 (Continued)

PROGRAMMER _____ DIVISION _____

LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

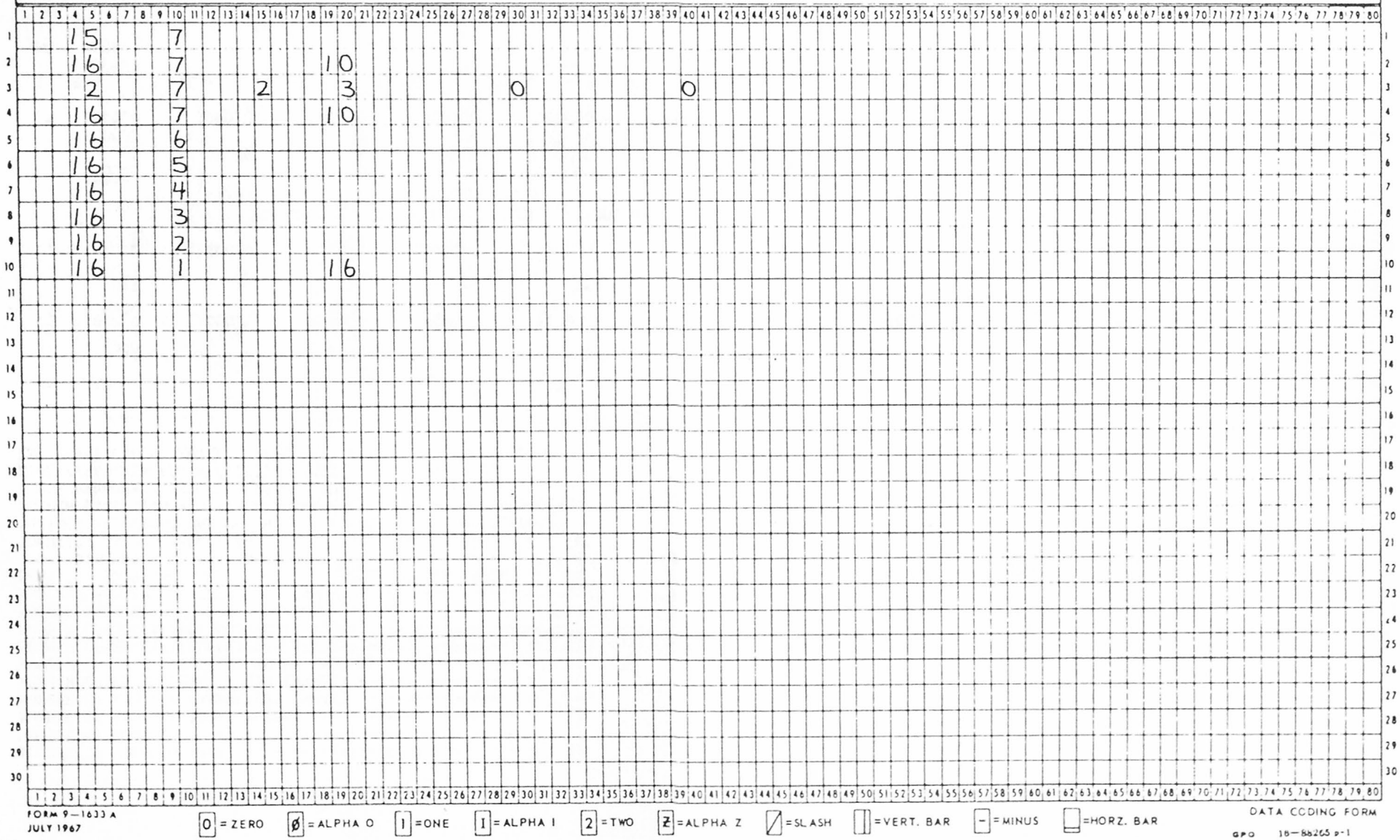
NUMBER _____

PROJECT _____

SHEET 4 OF 4

IDENTIFICATION: _____

INFORMATION: _____



FORM 9-1633 A
JULY 1967

0 = ZERO ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Z = ALPHA Z / = SLASH || = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM

GPO 16-88265 p-1

fig. 4.3-9 (Continued)

EXAMPLE PROBLEM FOR CARD INPUT

```
NUMBER OF COLUMNS (ID) ------------------------------ =   16
NUMBER OF ROWS (JD) --------------------------------- =    7
NUMBER OF AQUIFER ZONES (NZNS) ---------------------- =    3
NUMBER OF AQUIFER PARAMETERS (NPAR) ----------------- =    9
NUMBER OF KNOWN POINT FLOWS (NWELS) ----------------- =    0
NUMBER OF SPECIFIED FLOW ZONES (NQBND) -------------- =    3
NUMBER OF SPECIFIED FLOW PARAMETERS (NBQP) ---------- =    3
NUMBER OF SPECIFIED HEAD ZONES (NBHZ) --------------- =    2
NUMBER OF SPECIFIED HEAD PARAMETERS (NBHP) ---------- =    3
MAXIMUM NUMBER OF ITERATIONS (NUM) ------------------ =   20
SENSITIVITY PRINT AND ORTHOGONALIZATION OPTION (IPRX) =    1
ADDITIONAL PRINTOUT OPTION (IPO) -------------------- =    1
SOLUTION ONLY OPTION (ISO) -------------------------- =    0
MAXIMUM ALLOWABLE PARAMETER CORRECTION (BMX) - =  1.5000
SEARCH DIRECTION ADJUSTMENT PARAMETER (CSA) -- =  .80000E-01
RIDGE PARAMETER FOR REGRESSION (RP) ---------- = 0.
BIAS PARAMETER FOR REGRESSION (BP) ----------- = 0.
ESTIMATED ERROR VARIANCE (EV) ---------------- =  1.0000

 1  IB =     1  IE =    15  JB =     1  JE =     1  DX   = 1000.0

 DX   ARRAY:
          1          2          3          4          5          6          7          8          9         10

 1   1000.0     1000.0     1000.0     1000.0     1000.0     1000.0     1000.0     1000.0     1000.0     1000.0

         11         12         13         14         15

 1   1000.0     1000.0     1000.0     1000.0     1000.0

 1  IB =     1  IE =     6  JB =     1  JE =     1  DY   = 1000.0

 DY   ARRAY:
          1          2          3          4          5          6

 1   1000.0     1000.0     1000.0     1000.0     1000.0     1000.0

 1  IB =     1  IE =    16  JB =     1  JE =     7  HO   = 1.0000

 HO   ARRAY:
          1          2          3          4          5          6          7          8          9         10

 7   41.280     0.         34.820     0.         32.000     0.         26.970     0.         25.670     0.
 6   0.         38.660     0.         34.510     0.         27.810     0.         25.980     0.         22.160
 5   36.050     0.         31.640     0.         32.910     0.         25.050     0.         25.470     0.
 4   0.         29.360     0.         27.760     0.         25.510     0.         24.430     0.         22.460
 3   31.250     0.         27.500     0.         26.220     0.         25.650     0.         24.350     0.
 2   0.         28.270     0.         25.170     0.         26.940     0.         23.590     0.         22.530
 1   29.510     0.         26.380     0.         25.160     0.         26.810     0.         23.820     0.
```

Fig. 4.3-10

	11	12	13	14	15	16
7	18.120	0.	18.400	0.	11.900	0.
6	0.	15.730	0.	14.790	0.	10.480
5	20.390	0.	14.400	0.	14.370	0.
4	0.	17.160	0.	15.740	0.	14.690
3	22.480	0.	15.850	0.	15.710	0.
2	0.	17.890	0.	15.870	0.	15.980
1	23.590	0.	16.680	0.	15.310	0.

1 IB = 1 IE = 16 JB = 1 JE = 7 W = 1.0000

W ARRAY:

	1	2	3	4	5	6	7	8	9	10
7	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.
6	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000
5	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.
4	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000
3	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.
2	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000
1	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.	1.0000	0.

	11	12	13	14	15	16
7	1.0000	0.	1.0000	0.	1.0000	0.
6	0.	1.0000	0.	1.0000	0.	1.0000
5	1.0000	0.	1.0000	0.	1.0000	0.
4	0.	1.0000	0.	1.0000	0.	1.0000
3	1.0000	0.	1.0000	0.	1.0000	0.
2	0.	1.0000	0.	1.0000	0.	1.0000
1	1.0000	0.	1.0000	0.	1.0000	0.

1 IB = 1 IE = 15 JB = 1 JE = 6 T = 1.0000

T ARRAY:

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Fig. 4.3-10 (Continued)

```
          11          12          13          14          15

 6   1.0000      1.0000      1.0000      1.0000      1.0000
 5   1.0000      1.0000      1.0000      1.0000      1.0000
 4   1.0000      1.0000      1.0000      1.0000      1.0000
 3   1.0000      1.0000      1.0000      1.0000      1.0000
 2   1.0000      1.0000      1.0000      1.0000      1.0000
 1   1.0000      1.0000      1.0000      1.0000      1.0000

1  IB =     1  IE =    15  JB =     1  JE =     6  SL   = 1.0000
SL   ARRAY:

          1           2           3           4           5           6           7           8           9          10

 6   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000
 5   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000
 4   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000
 3   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000
 2   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000
 1   1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000      1.0000

          11          12          13          14          15

 6   1.0000      1.0000      1.0000      1.0000      1.0000
 5   1.0000      1.0000      1.0000      1.0000      1.0000
 4   1.0000      1.0000      1.0000      1.0000      1.0000
 3   1.0000      1.0000      1.0000      1.0000      1.0000
 2   1.0000      1.0000      1.0000      1.0000      1.0000
 1   1.0000      1.0000      1.0000      1.0000      1.0000

1  IB =     1  IE =     6  JB =     1  JE =     3  HR   = 25.000
2  IB =     7  IE =    10  JB =     1  JE =     7  HR   = 25.000
3  IB =     1  IE =     5  JB =     4  JE =     4  HR   = 30.000
4  IB =     6  IE =     6  JB =     4  JE =     4  HR   = 27.500
5  IB =     6  IE =     6  JB =     5  JE =     7  HR   = 30.000
6  IB =     1  IE =     5  JB =     5  JE =     7  HR   = 35.000
7  IB =    11  IE =    11  JB =     1  JE =     7  HR   = 20.000
8  IB =    12  IE =    16  JB =     1  JE =     7  HR   = 15.000

HR   ARRAY:

          1           2           3           4           5           6           7           8           9          10

 7   35.000      35.000      35.000      35.000      35.000      30.000      25.000      25.000      25.000      25.000
 6   35.000      35.000      35.000      35.000      35.000      30.000      25.000      25.000      25.000      25.000
 5   35.000      35.000      35.000      35.000      35.000      30.000      25.000      25.000      25.000      25.000
 4   30.000      30.000      30.000      30.000      30.000      27.500      25.000      25.000      25.000      25.000
 3   25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000
 2   25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000
 1   25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000      25.000
```

Fig. 4.3-10 (Continued)

	11	12	13	14	15	16
7	20.000	15.000	15.000	15.000	15.000	15.000
6	20.000	15.000	15.000	15.000	15.000	15.000
5	20.000	15.000	15.000	15.000	15.000	15.000
4	20.000	15.000	15.000	15.000	15.000	15.000
3	20.000	15.000	15.000	15.000	15.000	15.000
2	20.000	15.000	15.000	15.000	15.000	15.000
1	20.000	15.000	15.000	15.000	15.000	15.000

1 IB = 1 IE = 15 JB = 1 JE = 6 QRE = 1.0000

QRE ARRAY:

	1	2	3	4	5	6	7	8	9	10
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

	11	12	13	14	15
6	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000	1.0000
3	1.0000	1.0000	1.0000	1.0000	1.0000
2	1.0000	1.0000	1.0000	1.0000	1.0000
1	1.0000	1.0000	1.0000	1.0000	1.0000

1 IB = 1 IE = 5 JB = 1 JE = 3 IZN = 1
2 IB = 6 IE = 10 JB = 1 JE = 6 IZN = 1
3 IB = 1 IE = 5 JB = 4 JE = 6 IZN = 2
4 IB = 11 IE = 15 JB = 1 JE = 6 IZN = 3

IZN ARRAY:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
6	2	2	2	2	2	1	1	1	1	1	3	3	3	3	3
5	2	2	2	2	2	1	1	1	1	1	3	3	3	3	3
4	2	2	2	2	2	1	1	1	1	1	3	3	3	3	3
3	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3
2	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3
1	1	1	1	1	1	1	1	1	1	1	3	3	3	3	3

INITIAL AQUIFER PARAMETERS BY ZONE

ZONE	TRANX	TRANY	VLEAK	QDIST
1	4000.0	4000.0	.10000E-02	.10000E-03
2	400.00	400.00	.70000E-03	.50000E-03
3	1000.0	1000.0	.15000E-02	-.10000E-03

Fig. 4.3-10 (Continued)

AQUIFER PARAMETER NUMBERS

ZONE	TRANX	TRANY	VLEAK	QDIST
1	1	1	4	7
2	2	2	5	8
3	3	3	6	9

COEFFICIENTS OF VARIATION FOR AQUIFER PARAMETERS

PAR. NO.	COEF. VAR.	PAR. NO.	COEF. VAR.	PAR. NO.	COEF. VAR.
1	0.	4	0.	7	0.
2	0.	5	0.	8	0.
3	0.	6	0.	9	0.

INITIAL SPECIFIED FLOW DATA

NODE NO.S IA	JA	IB	JB	PAR. NO.	FLOW PARAMETER	COEF. VAR.	MULTIPLIER
1	1	1	4	1	8.0000	.10000	1.0000
1	4	1	7	2	.80000	0.	1.0000
8	1	16	1	3	1.0000	0.	1.0000

1 IB = 16 IE = 16 JB = 1 JE = 7 IBZN = -1
2 IB = 1 IE = 15 JB = 7 JE = 7 IBZN = -1

IBZN ARRAY:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
7	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1

INITIAL SPECIFIED HEAD DATA

NO. OF NODES IN ZONE 1 = 16
NO. PAR. A = 1 NO. PAR. B = 2
COEF. VAR. PAR. A = 0. COEF. VAR. PAR. B = 0.

Fig. 4.3-10 (Continued)

INITIAL VALUES OF HEAD

I	J	HEAD
1	7	40.000
2	7	0.
3	7	0.
4	7	0.
5	7	0.
6	7	0.
7	7	0.
8	7	0.
9	7	0.
10	7	0.
11	7	0.
12	7	0.
13	7	0.
14	7	0.
15	7	0.
16	7	10.000

NO. OF NODES IN ZONE 2 = 7
NO. PAR. A = 2 NO. PAR. B = 3
COEF. VAR. PAR. A = 0. COEF. VAR. PAR. B = 0.

INITIAL VALUES OF HEAD

I	J	HEAD
16	7	10.000
16	6	0.
16	5	0.
16	4	0.
16	3	0.
16	2	0.
16	1	16.000

SOLUTION BY LDU FACTORIZATION ASSUMING D4 ORDERING
****WARNING*****MINIMUM DIMENSIONS FOR ARRAYS USED BY THIS METHOD ARE AS FOLLOWS:
AU: 5 BY 45
AL: 9 BY 45
IC: 5 BY 45
B: 90

INITIAL SOLUTION:

	1	2	3	4	5	6	7	8	9	10
7	40.000	38.000	36.000	34.000	32.000	30.000	28.000	26.000	24.000	22.000
6	37.229	36.064	35.263	34.409	32.890	28.278	26.913	25.554	24.162	22.673
5	35.441	34.373	33.789	33.201	31.928	27.398	26.273	25.255	24.188	22.940
4	31.311	29.853	28.947	28.280	27.600	26.598	25.818	25.043	24.168	23.046
3	30.504	28.992	28.034	27.353	26.759	26.128	25.538	24.917	24.161	23.113
2	30.074	28.548	27.576	26.905	26.369	25.875	25.397	24.882	24.212	23.216
1	29.940	28.411	27.438	26.771	26.255	25.800	25.368	24.947	24.366	23.409

Fig. 4.3-10 (Continued)

```
          11          12          13          14          15          16

   7   20.000      18.000      16.000      14.000      12.000      10.000
   6   20.982      17.072      15.556      14.585      13.453      11.000
   5   21.315      16.957      15.500      14.807      14.007      12.000
   4   21.447      16.977      15.532      14.946      14.378      13.000
   3   21.534      17.036      15.605      15.088      14.724      14.000
   2   21.660      17.180      15.772      15.308      15.118      15.000
   1   21.916      17.623      16.251      15.816      15.719      16.000

NO. OF OBSERVATIONS =   56
NO. OF PARAMETERS HAVING PRIOR INFORMATION =    1

ESTIMATED SUM OF SQUARED ERRORS FOR INITIAL SOLUTION =  62.212

     NODAL LOCATION AND SCALED SENSITIVITIES FOR EACH PARAMETER
    1    1  3.6727      .37761E-01  .63299E-02  .41467      .57848E-01  .24491E-03 -1.5291
            .20248     -.10080E-01 -2.5865      .58764E-01 -.65671E-02  .88772E-01  .65710E-01
           -.27446E-03
    3    1  1.2772      .24314E-01  .10512E-01  .53104      .85986E-01  .39764E-03  .24127
            .19229     -.16301E-01 -1.9461      .78204E-01 -.10637E-01  .88592E-01  .61088E-01
           -.44396E-03
    5    1  .43351      .10195E-01  .29967E-01  .91338      .17803      .10474E-02  .33822
            .14256     -.42379E-01 -1.1040      .87190E-01 -.27802E-01  .87771E-01  .45904E-01
           -.11550E-02
    7    1  .14158      .35816E-02  .10882      1.3596      .30295      .30394E-02  .83801E-01
            .76474E-01 -.11934     -.37748      .55089E-01 -.79277E-01  .85192E-01  .24826E-01
           -.32568E-02
    9    1  .46480E-01  .12104E-02  .37073      1.3709      .35670      .92087E-02 -.17611
            .33341E-01 -.34274      .19590      .25592E-01 -.23324      .77403E-01  .10797E-01
           -.93554E-02
   11    1  .18022E-01  .47446E-03  .49558      .94666      .27636      .28646E-01  1.0522
            .14656E-01 -.98740      .64792E-01  .11508E-01 -.70222      .54222E-01  .47335E-02
           -.26538E-01
   13    1  .16816E-02  .44456E-04  .67776      .12394      .94553E-01  .41097      .11223
            .14394E-02  .30068      .77542E-02  .11396E-02 -1.0478      .52278E-02  .46424E-03
           -.60589E-01
   15    1  .14265E-03  .37813E-05  .53296      .16022E-01  .24008      4.6514      .10152E-01
            .12647E-03 -.12566      .74317E-03  .10068E-03 -.37371      .45450E-03  .40746E-04
           -.45352E-01
    2    2  2.1391      .38953E-01  .70743E-02  .46794      .65538E-01  .27480E-03 -.33264
            .23027     -.11317E-01 -2.2948      .69856E-01 -.73715E-02  .87371E-01  .73856E-01
           -.30815E-03
    4    2  .73245      .17283E-01  .16239E-01  .74019      .13061      .61718E-03  .29789
            .19882     -.25302E-01 -1.4751      .10484     -.16511E-01  .86967E-01  .63138E-01
           -.68912E-03
    6    2  .24060      .61098E-02  .47836E-01  1.3474      .27683      .17263E-02  .19050
            .12045     -.69809E-01 -.69732      .84646E-01 -.45804E-01  .85515E-01  .39165E-01
           -.19029E-02
    8    2  .77440E-01  .20182E-02  .14360      1.6709      .40022      .50712E-02 -.31424E-01
            .54070E-01 -.19833     -.50049E-01  .41252E-01 -.13190      .81207E-01  .17539E-01
           -.54179E-02
```

Fig. 4.3-10 (Continued)

```
10   2  .26564E-01  .70071E-03  .23450      1.3981      .39677      .15484E-01  .26637
       .21945E-01 -.56959      .33128      .17275E-01 -.38847      .67982E-01  .70873E-02
     -.15644E-01
12   2  .52898E-02  .14007E-03  .21476      .39631      .14989      .11391      .37583
       .46091E-02  .78272      .26881E-01  .36595E-02 -1.3787      .16618E-01  .14860E-02
     -.53315E-01
14   2  .49057E-03  .13021E-04  .19256      .57841E-01  .23067      1.2535      .36057E-01
       .44198E-03  .25398      .27048E-02  .35270E-03 -.42958      .15800E-02  .14233E-03
     -.58743E-01
 1   3  3.4904      .85382E-01  .56983E-02  .51285      .58605E-01  .22517E-03 -1.8050
       .31034     -.93026E-02 -2.3153      .55556E-01 -.60516E-02  .82978E-01  .10551
     -.25325E-03
 3   3  1.1843      .34507E-01  .91359E-02  .65361      .94350E-01  .36581E-03 -.12232
       .32132     -.15146E-01 -1.7116      .12750     -.98444E-02  .82839E-01  .99733E-01
     -.41228E-03
 5   3  .39259      .10951E-01  .22888E-01  1.3018      .24242      .95887E-03  .11560E-01
       .22944     -.39933E-01 -.92531      .16744     -.25894E-01  .82158E-01  .75197E-01
     -.10867E-02
 7   3  .11940      .32247E-02  .56894E-01  2.3810      .50865      .26473E-02  .29811E-01
       .10071     -.11028     -.31708      .79158E-01 -.71477E-01  .79848E-01  .32803E-01
     -.30027E-02
 9   3  .39003E-01  .10447E-02  .11083      2.3483      .62438      .77958E-02  .39077E-02
       .36832E-01 -.32071      .23185      .29592E-01 -.20851      .73034E-01  .11883E-01
     -.87594E-02
11   3  .15235E-01  .40729E-03  .11689      1.6060      .50880      .23550E-01  1.2530
       .14804E-01 -.93635      .10513      .11944E-01 -.61136      .51348E-01  .47607E-02
     -.25823E-01
13   3  .14286E-02  .38176E-04  .66737E-01  .23707      .22167      .31913      .11805
       .13985E-02  .55627      .98336E-02  .11292E-02 -.70005      .48803E-02  .44929E-03
     -.60169E-01
15   3  .12169E-03  .32511E-05  .37527E-01  .40644E-01  1.0958      3.3291      .10089E-01
       .11952E-03 -.48766E-01  .83892E-03  .96539E-04  .44652E-01  .41862E-03  .38380E-04
     -.45140E-01
 2   4  1.8209      .88655E-01  .57427E-02  .67995      .66367E-01  .23055E-03 -1.2067
       .52594     -.95542E-02 -1.5543      .97748E-01 -.62077E-02  .74334E-01  .16365
     -.26005E-03
 4   4  .61266      .20763E-01  .12207E-01  1.0539      .16217      .52142E-03 -.73901
       .47520     -.21865E-01 -.86149      .29426     -.14140E-01  .74146E-01  .14787
     -.59472E-03
 6   4  .15885      .46024E-02  .27894E-01  3.2771      .62115      .14758E-02 -.35192
       .22218     -.64379E-01 -.28246      .18184     -.41003E-01  .72682E-01  .73456E-01
     -.17470E-02
 8   4  .51470E-01  .14292E-02  .48881E-01  4.1607      .96892      .36502E-02 -.36747E-01
       .69181E-01 -.16514     -.11250E-01  .58076E-01 -.10372      .70061E-01  .22234E-01
     -.44753E-02
10   4  .18344E-01  .49923E-03  .64502E-01  3.3261      1.0344      .10740E-01  .28630
       .22000E-01 -.51144      .36406      .18222E-01 -.31516      .59494E-01  .70352E-02
     -.13863E-01
12   4  .37106E-02  .10038E-03  .31186E-01  1.0228      .47757      .74329E-01  .36930
       .42348E-02  .76168      .34853E-01  .34853E-02 -1.1724      .14317E-01  .13540E-02
     -.51797E-01
```

Fig. 4.3-10 (Continued)

```
14    4  .34785E-03  .93766E-05  .17367E-01  .22884      .70359      .78321      .33566E-01
        .38315E-03  .84438E-01  .30762E-02  .31394E-03 -.83497E-01  .13157E-02  .12257E-03
       -.57443E-01
 1    5  .76147      .97428      .15875E-02  3.1230      .53671E-01  .64121E-04 -.54339
       -1.0239     -.26627E-02 -.47725     -.21092     -.17287E-02  .21374E-01  .50123
       -.72464E-04
 3    5  .32546      .59137E-01  .25531E-02  2.9394      .14555      .11015E-03 -.26549
       -.91638     -.46499E-02 -.34596      .63139     -.29995E-02  .21649E-01  .49540
       -.12640E-03
 5    5  .11047      .57205E-02  .75221E-02  3.7707      .48924      .41907E-03 -.31903
       -1.4682     -.18796E-01 -.13957      1.4015     -.11849E-01  .28290E-01  .40446
       -.50881E-03
 7    5  .49128E-01  .14511E-02  .21123E-01  7.5615      1.4489      .15869E-02 -.15133
        .14169     -.75545E-01 -.16917      .12940     -.46606E-01  .57298E-01  .44601E-01
       -.20368E-02
 9    5  .19731E-01  .55337E-03  .31053E-01  6.5529      1.8057      .43329E-02 -.26101E-01
        .34708E-01 -.23158      .21988      .29798E-01 -.13751      .54707E-01  .10947E-01
       -.61963E-02
11    5  .82014E-02  .22543E-03  .28030E-01  4.4665      1.6798      .12606E-01  1.0865
        .11549E-01 -.81642      .11284      .96994E-02 -.46166      .38990E-01  .36624E-02
       -.21636E-01
13    5  .78562E-03  .21425E-04  .79262E-02  1.0409      .99611      .16360      .94063E-01
        .10083E-02  .43777      .93753E-02  .83922E-03 -.50028      .34509E-02  .32068E-03
       -.55279E-01
15    5  .67812E-04  .18396E-05  .30619E-02  .28110      2.5312      1.6728      .74868E-02
        .81628E-04 -.38846      .72193E-03  .67500E-04  .41883      .28109E-03  .26016E-04
       -.42164E-01
 2    6  .13934      .21222      .52938E-03  10.725      .21579      .22462E-04 -.10017
       -.18028     -.94596E-03 -.11233     -.42942     -.61078E-03  .57617E-02  .46593
       -.25718E-04
 4    6  .54859E-01  .12094E-01  .13611E-02  9.5456      .64002      .73716E-04 -.86686E-01
       -.91363     -.33126E-02 -.63908E-01  .54767     -.20868E-02  .70690E-02  .44666
       -.89634E-04
 6    6  .25323E-01  .88032E-03  .80304E-02  13.767      2.0931      .61376E-03 -.59779
        .27089     -.29884E-01  .23624E-01  .20876     -.18287E-01  .32726E-01  .76541E-01
       -.80378E-03
 8    6  .14647E-01  .43084E-03  .11594E-01  12.548      2.7945      .13189E-02 -.11414E-01
        .46736E-01 -.71883E-01 -.35333E-01  .39477E-01 -.42323E-01  .35820E-01  .14163E-01
       -.19149E-02
10    6  .61083E-02  .17101E-03  .13371E-01  9.5185      3.3550      .37626E-02  .77341E-01
        .11088E-01 -.27348      .27539      .93992E-02 -.14725      .31497E-01  .34667E-02
       -.70983E-02
12    6  .12815E-02  .35372E-04  .49386E-02  4.1905      2.6050      .25187E-01  .20930
        .19366E-02  .67366      .21772E-01  .16295E-02 -.88469      .70055E-02  .61113E-03
       -.37484E-01
14    6  .12243E-03  .33496E-05  .16250E-02  1.7271      3.0714      .26231      .16366E-01
        .16512E-03 -.12829      .16529E-02  .13787E-03  .14941      .57928E-03  .52370E-04
       -.41828E-01
```

Fig. 4.3-10 (Continued)

SCALED LEAST SQUARES MATRIX AND GRADIENT VECTOR

	1	2	3	4	5	6	7	8	9	10
1	1.0000	.13317	.17099E-01	.20165E-01	.90366E-02	.71068E-04	-.35370	.11366	-.13671E-01	-.48758
2	.13317	1.0000	.64661E-02	.77897E-01	.82024E-02	.25505E-04	-.24478	-.39631	-.52388E-02	-.20846
3	.17099E-01	.64661E-02	1.0000	.47384E-01	.59304E-01	.16373	.20587	.32808E-01	-.21300	-.34027E-02
4	.20165E-01	.77897E-01	.47384E-01	1.0000	.43839	.17691E-02	-.27537E-01	-.66174E-01	-.73124E-01	-.29440E-01
5	.90366E-02	.82024E-02	.59304E-01	.43839	1.0000	.27922	.27664E-01	.45383E-02	-.77518E-01	-.33928E-02
6	.71068E-04	.25505E-04	.16373	.17691E-02	.27922	1.0000	.67823E-02	.21096E-03	-.14278E-01	.37787E-03
7	-.35370	-.24478	.20587	-.27537E-01	.27664E-01	.67823E-02	1.0000	-.69544E-01	-.26635	.53102
8	.11366	-.39631	.32808E-01	-.66174E-01	.45383E-02	.21096E-03	-.69544E-01	1.0000	-.30813E-01	-.25996
9	-.13671E-01	-.52388E-02	-.21300	-.73124E-01	-.77518E-01	-.14278E-01	-.26635	-.30813E-01	1.0000	-.25211E-01
10	-.48758	-.20846	-.34027E-02	-.29440E-01	-.33928E-02	.37787E-03	.53102	-.25996	-.25211E-01	1.0000
11	.76453E-01	-.11478	.43106E-01	.10554	.65180E-01	.26532E-03	-.17273	-.53051	-.44226E-01	-.20068
12	-.84692E-02	-.31911E-02	-.63606	-.78825E-01	-.10046	-.52616E-01	-.34007	-.20677E-01	-.20498	-.23735E-01
13	.31876	.14497	.33531	.17203	.14428	.35636E-02	-.17912	.23497	-.34146	-.67194
14	.16830	.60353	.32631E-01	.20337	.55433E-01	.18356E-03	-.31368	-.63216	-.31616E-01	-.33069
15	-.54464E-02	-.20517E-02	-.58178	-.63744E-01	-.18631	-.20016	-.23649	-.13554E-01	-.16613	-.16667E-01

	11	12	13	14	15
1	.76453E-01	-.84692E-02	.31876	.16830	-.54464E-02
2	-.11478	-.31911E-02	.14497	.60353	-.20517E-02
3	.43106E-01	-.63606	.33531	.32631E-01	-.58178
4	.10554	-.78825E-01	.17203	.20337	-.63744E-01
5	.65180E-01	-.10046	.14428	.55433E-01	-.18631
6	.26532E-03	-.52616E-01	.35636E-02	.18356E-03	-.20016
7	-.17273	-.34007	-.17912	-.31368	-.23649
8	-.53051	-.20677E-01	.23497	-.63216	-.13554E-01
9	-.44226E-01	-.20498	-.34146	-.31616E-01	-.16613
10	-.20068	-.23735E-01	-.67194	-.33069	-.16667E-01
11	1.0000	-.28414E-01	.30415	.50312	-.18664E-01
12	-.28414E-01	1.0000	-.26777	-.20134E-01	.73701
13	.30415	-.26777	1.0000	.33113	-.17797
14	.50312	-.20134E-01	.33113	1.0000	-.13167E-01
15	-.18664E-01	.73701	-.17797	-.13167E-01	1.0000

-.37087 .90839 .69633 .16964 .49968 1.6963 .51601 -.99079
-.45346 1.2969 -1.2398 -.22907 -1.3453 .30868 -.67924

ITERATION NO. 1
YSQ = 62.212 DET(C) = .31948E-03 AMP = 0.
SCALED REGRESSION COEFFICIENTS:
.29713E-13 -1.0581 .81262E-01 .20822E-02 .41406E-02 .88810E-01 -.32389 -.96997
-.64503 -.24799E-01 -1.7053 -.31984 -3.2429 .78322 -2.5992

Fig. 4.3-10 (Continued)

UPDATED PARAMETERS

1	8.0000			
2	.40847			
3	1.0376			
1	40.039	10.019		
2	10.019	16.657		
1	3400.8	3400.8	.98853E-03	-.50000E-04
2	220.54	220.54	.14786E-03	.68114E-03
3	701.64	701.64	.12781E-02	.20223E-04

* * * * * * * *

Output for iterations 2 through 6 is analogous to output for iteration 1, and, thus, is omitted. The scaled least squares matrix and gradient vector following is for iteration 6.

* * * * * * * *

SCALED LEAST SQUARES MATRIX AND GRADIENT VECTOR

	1	2	3	4	5	6	7	8	9	10
1	1.0000	.14295	.98485E-02	.28550E-01	.77942E-02	.14932E-04	-.45891	-.56820E-01	-.72582E-02	-.57010
2	.14295	1.0000	.23673E-02	.13444	.12050E-01	.29220E-05	-.26481	-.63943	-.18290E-02	-.17598
3	.98485E-02	.23673E-02	1.0000	.25953E-01	.33840E-01	.14324	.12516	.34937E-03	-.11514	.29284E-01
4	.28550E-01	.13444	.25953E-01	1.0000	.42776	.70059E-03	-.38769E-01	-.22608	-.41526E-01	-.97069E-02
5	.77942E-02	.12050E-01	.33840E-01	.42776	1.0000	.27776	.18259E-01	-.40783E-01	-.44719E-01	.18788E-01
6	.14932E-04	.29220E-05	.14324	.70059E-03	.27776	1.0000	.20587E-02	.63503E-05	-.58981E-02	.14509E-03
7	-.45891	-.26481	.12516	-.38769E-01	.18259E-01	.20587E-02	1.0000	.17180	-.17219	.58498
8	-.56820E-01	-.63943	.34937E-03	-.22608	-.40783E-01	.63503E-05	.17180	1.0000	.19293E-03	.43378E-01
9	-.72582E-02	-.18290E-02	-.11514	-.41526E-01	-.44719E-01	-.58981E-02	-.17219	.19293E-03	1.0000	-.51165E-01
10	-.57010	-.17598	.29284E-01	-.97069E-02	.18788E-01	.14509E-03	.58498	.43378E-01	-.51165E-01	1.0000
11	-.10614E-01	-.46230	.11169E-01	-.52273E-02	.36389E-01	.25097E-04	.82455E-02	-.17720	-.11965E-01	.86505E-02
12	-.44164E-02	-.10535E-02	-.67278	-.53213E-01	-.84249E-01	-.88041E-01	-.22676	-.38005E-03	-.38104	-.34320E-01
13	-.35646	-.12938	-.25625	-.16347	-.12543	-.12579E-02	.19756	.83920E-01	.25862	.55596
14	.13920	.64480	.93450E-02	.26989	.53228E-01	.17694E-04	-.25747	-.95324	-.91209E-02	-.18514
15	.27229E-02	.65011E-03	.54316	.43188E-01	.14376	.16090	.15229	.24789E-03	.30206	.22358E-01

	11	12	13	14	15
1	-.10614E-01	-.44164E-02	-.35646	.13920	.27229E-02
2	-.46230	-.10535E-02	-.12938	.64480	.65011E-03
3	.11169E-01	-.67278	-.25625	.93450E-02	.54316
4	-.52273E-02	-.53213E-01	-.16347	.26989	.43188E-01
5	.36389E-01	-.84249E-01	-.12543	.53228E-01	.14376
6	.25097E-04	-.88041E-01	-.12579E-02	.17694E-04	.16090
7	.82455E-02	-.22676	.19756	-.25747	.15229
8	-.17720	-.38005E-03	.83920E-01	-.95324	.24789E-03
9	-.11965E-01	-.38104	.25862	-.91209E-02	.30206
10	.86505E-02	-.34320E-01	.55596	-.18514	.22358E-01
11	1.0000	-.72451E-02	-.12575	.45636E-01	.45959E-02
12	-.72451E-02	1.0000	.20760	-.54336E-02	-.80292
13	-.12575	.20760	1.0000	-.22633	-.13295
14	.45636E-01	-.54336E-02	-.22633	1.0000	.34231E-02
15	.45959E-02	-.80292	-.13295	.34231E-02	1.0000

Fig. 4.3-10 (Continued)

```
 .19812E-03  -.25707E-03  -.15761E-02  -.30471E-03   .33699E-03   .27447E-03  -.51657E-03   .11084E-03
-.20590E-03  -.28675E-03  -.93586E-03   .11319E-02  -.11511E-03   .50789E-03  -.47620E-03

ITERATION NO.   6
YSQ =  51.970      DET(C) =  .45800E-04  AMP = 0.
SCALED REGRESSION COEFFICIENTS:
 .55367E-14  -.14731E-02  -.13799E-02  -.13195E-04   .24670E-04   .16216E-04  -.14841E-03   .13487E-02
-.30958E-03   .16237E-03  -.90764E-03   .86110E-04  -.13674E-03   .21318E-02   .27212E-02

UPDATED PARAMETERS
    1    8.0000
    2    .22321
    3    .89599
    1    40.109          10.042
    2    10.042          17.454
   1     2865.3         2865.3         .99782E-03  -.17358E-03
   2     117.67         117.67         .96559E-04   .33062E-03
   3     497.85         497.85         .10785E-02   .14220E-03


SOLUTION CONVERGED IN   6 ITERATIONS

ERROR VARIANCE =  1.2374
ESTIMATED SUM OF SQUARED ERRORS =  51.970
CORRELATION COEFFICIENT =  .99033

 ESTIMATED SPECIFIED FLOW PARAMETERS
  ZONE      QBND          STD. ER.
    1     8.0000          .88990
    2     .22321          .59589
    3     .89599          1.1272

                  ESTIMATED SPECIFIED HEADS
           HEAD          STD. ER.          HEAD          STD. ER.
 ZONE        A             A               B             B
    1     40.109          .68266         10.042          .66622
    2     10.042          .66622         17.454          1.1536

            ESTIMATED AQUIFER PARAMETERS BY ZONE
 ZONE      TRANX          TRANY          VLEAK          QDIST
    1     2865.3         2865.3         .99782E-03  -.17358E-03
    2     117.67         117.67         .96559E-04   .33062E-03
    3     497.85         497.85         .10785E-02   .14220E-03

               ESTIMATED STANDARD ERRORS BY ZONE
          STD. ER.       STD. ER.       STD. ER.       STD. ER.
 ZONE      TRANX          TRANY          VLEAK          QDIST
    1     1293.4         1293.4         .34445E-03  -.50156E-03
    2     267.19         267.19         .22622E-03   .73743E-03
    3     426.60         426.60         .99513E-03   .71226E-03
```

Fig. 4.3-10 (Continued)

SCALED VARIANCE-COVARIANCE MATRIX

	1	2	3	4	5	6	7	8	9	10
1	.12374E-01	.12374E-01	.12374E-01	.14037E-16	-.18460E-16	-.42223E-17	.12374E-01	.12374E-01	.12374E-01	.12374E-01
2	.12374E-01	7.1270	.17926	.14479E-02	-.89790E-03	-.19144E-02	.36771	4.9253	-.13660	.13294
3	.12374E-01	.17926	1.5828	.16405E-03	.88865E-02	-.10674E-01	.17028	.89348E-01	.61356	-.14275E-01
4	.14037E-16	.14479E-02	.16405E-03	.28968E-03	-.48668E-03	.14243E-03	.19402E-03	-.11679E-02	.56782E-04	-.48957E-03
5	-.18460E-16	-.89790E-03	.88865E-02	-.48668E-03	.44012E-02	-.13178E-02	.93779E-03	.15683E-02	.32202E-02	-.12911E-02
6	-.42223E-17	-.19144E-02	-.10674E-01	.14243E-03	-.13178E-02	.43685E-02	-.46185E-03	-.26240E-02	-.12360E-02	.69670E-03
7	.12374E-01	.36771	.17028	.19402E-03	.93779E-03	-.46185E-03	.20377	.27356	.13842	-.44172E-01
8	.12374E-01	4.9253	.89348E-01	-.11679E-02	.15683E-02	-.26240E-02	.27356	5.1554	-.22096	.16192
9	.12374E-01	-.13660	.61356	.56782E-04	.32202E-02	-.12360E-02	.13842	-.22096	.73426	.35543E-01
10	.12374E-01	.13294	-.14275E-01	-.48957E-03	-.12911E-02	.69670E-03	-.44172E-01	.16192	.35543E-01	.11917
11	.12374E-01	5.5910	.10210	.95276E-03	-.16798E-02	-.17314E-02	.29319	4.6513	-.21847	.10755
12	.12374E-01	-.48073E-01	.75242	.45102E-03	.93419E-03	-.52882E-02	.15233	-.13844	.51406	.18831E-01
13	.12374E-01	4.0784	-.16586	.25418E-02	.12522E-01	-.95589E-02	.20222	4.1394	-1.0247	-.27200
14	.12374E-01	4.1118	.84191E-01	-.32992E-02	.41988E-02	-.33516E-02	.26700	4.9213	-.21898	.15537
15	.12374E-01	-.23149	.18966	.31674E-02	-.35068E-01	-.36029E-01	.13377	-.33276	.31282	.53647E-01

	11	12	13	14	15
1	.12374E-01	.12374E-01	.12374E-01	.12374E-01	.12374E-01
2	5.5910	-.48073E-01	4.0784	4.1118	-.23149
3	.10210	.75242	-.16586	.84191E-01	.18966
4	.95276E-03	.45102E-03	.25418E-02	-.32992E-02	.31674E-02
5	-.16798E-02	.93419E-03	.12522E-01	.41988E-02	-.35068E-01
6	-.17314E-02	-.52882E-02	-.95589E-02	-.33516E-02	-.36029E-01
7	.29319	.15233	.20222	.26700	.13377
8	4.6513	-.13844	4.1394	4.9213	-.33276
9	-.21847	.51406	-1.0247	-.21898	.31282
10	.10755	.18831E-01	-.27200	.15537	.53647E-01
11	5.4889	-.12915	4.1340	4.1499	-.30918
12	-.12915	.85135	-.79237	-.13995	2.6821
13	4.1340	-.79237	8.3492	4.0154	-1.3863
14	4.1499	-.13995	4.0154	4.9748	-.34255
15	-.30918	2.6821	-1.3863	-.34255	25.090

CORRELATION MATRIX FOR SCALED PARAMETERS

	1	2	3	4	5	6	7	8	9	10
1	1.0000	.41668E-01	.88418E-01	.74144E-14	-.25015E-14	-.57429E-15	.24642	.48992E-01	.12982	.32224
2	.41668E-01	1.0000	.53372E-01	.31867E-01	-.50698E-02	-.10850E-01	.30513	.81254	-.59713E-01	.14425
3	.88418E-01	.53372E-01	1.0000	.76614E-02	.10647	-.12836	.29984	.31278E-01	.56915	-.32869E-01
4	.74144E-14	.31867E-01	.76614E-02	1.0000	-.43102	.12662	.25253E-01	-.30222E-01	.38934E-02	-.83327E-01
5	-.25015E-14	-.50698E-02	.10647	-.43102	1.0000	-.30054	.31315E-01	.10412E-01	.56647E-01	-.56376E-01
6	-.57429E-15	-.10850E-01	-.12836	.12662	-.30054	1.0000	-.15480E-01	-.17485E-01	-.21823E-01	.30535E-01
7	.24642	.30513	.29984	.25253E-01	.31315E-01	-.15480E-01	1.0000	.26690	.35786	-.28346

Fig. 4.3-10 (Continued)

8	.48992E-01	.81254	.31278E-01	-.30222E-01	.10412E-01	-.17485E-01	.26690	1.0000	-.11357	.20659
9	.12982	-.59713E-01	.56915	.38934E-02	.56647E-01	-.21823E-01	.35786	-.11357	1.0000	.12016
10	.32224	.14425	-.32869E-01	-.83327E-01	-.56376E-01	.30535E-01	-.28346	.20659	.12016	1.0000
11	.47480E-01	.89391	.34638E-01	.23894E-01	-.10808E-01	-.11181E-01	.27722	.87438	-.10882	.13298
12	.12056	-.19516E-01	.64818	.28720E-01	.15261E-01	-.86713E-01	.36572	-.66079E-01	.65018	.59123E-01
13	.38497E-01	.52871	-.45625E-01	.51685E-01	.65321E-01	-.50051E-01	.15503	.63094	-.41385	-.27269
14	.49873E-01	.69054	.30003E-01	-.86908E-01	.28376E-01	-.22735E-01	.26519	.97176	-.11458	.20179
15	.22208E-01	-.17311E-01	.30096E-01	.37153E-01	-.10553	-.10883	.59160E-01	-.29259E-01	.72883E-01	.31026E-01

	11	12	13	14	15
1	.47480E-01	.12056	.38497E-01	.49873E-01	.22208E-01
2	.89391	-.19516E-01	.52871	.69054	-.17311E-01
3	.34638E-01	.64818	-.45625E-01	.30003E-01	.30096E-01
4	.23894E-01	.28720E-01	.51685E-01	-.86908E-01	.37153E-01
5	-.10808E-01	.15261E-01	.65321E-01	.28376E-01	-.10553
6	-.11181E-01	-.86713E-01	-.50051E-01	-.22735E-01	-.10883
7	.27722	.36572	.15503	.26519	.59160E-01
8	.87438	-.66079E-01	.63094	.97176	-.29259E-01
9	-.10882	.65018	-.41385	-.11458	.72883E-01
10	.13298	.59123E-01	-.27269	.20179	.31026E-01
11	1.0000	-.59744E-01	.61067	.79416	-.26347E-01
12	-.59744E-01	1.0000	-.29720	-.68002E-01	.58033
13	.61067	-.29720	1.0000	.62304	-.95784E-01
14	.79416	-.68002E-01	.62304	1.0000	-.30661E-01
15	-.26347E-01	.58033	-.95784E-01	-.30661E-01	1.0000

SAMPLE RESIDUALS

I	J	PREDICTED VALUE	OBSERVED VALUE	WEIGHTED RESIDUAL
1	1	29.988	29.510	.47777
2	1	27.977	0.	0.
3	1	26.826	26.380	.44644
4	1	26.139	0.	0.
5	1	25.687	25.160	.52684
6	1	25.345	0.	0.
7	1	25.055	26.810	-1.7551
8	1	24.806	0.	0.
9	1	24.409	23.820	.58942
10	1	23.621	0.	0.
11	1	22.257	23.590	-1.3331
12	1	17.722	0.	0.
13	1	16.544	16.680	-.13645
14	1	16.282	0.	0.
15	1	16.426	15.310	1.1157
16	1	17.454	0.	0.
1	2	30.105	0.	0.
2	2	28.095	28.270	-.17495
3	2	26.943	0.	0.

Fig. 4.3-10 (Continued)

4	2	26.249	25.170	1.0794
5	2	25.782	0.	0.
6	2	25.409	26.940	-1.5305
7	2	25.074	0.	0.
8	2	24.720	23.590	1.1296
9	2	24.220	0.	0.
10	2	23.386	22.530	.85609
11	2	21.935	0.	0.
12	2	17.048	17.890	-.84160
13	2	15.814	0.	0.
14	2	15.527	15.870	-.34305
15	2	15.585	0.	0.
16	2	16.218	15.980	.23841
1	3	30.499	31.250	-.75148
2	3	28.493	0.	0.
3	3	27.340	27.500	-.15980
4	3	26.630	0.	0.
5	3	26.114	26.220	-.10621
6	3	25.640	0.	0.
7	3	25.199	25.650	-.45050
8	3	24.741	0.	0.
9	3	24.155	24.350	-.19520
10	3	23.267	0.	0.
11	3	21.790	22.480	-.69007
12	3	16.875	0.	0.
13	3	15.618	15.850	-.23236
14	3	15.282	0.	0.
15	3	15.152	15.710	-.55827
16	3	14.983	0.	0.
1	4	31.294	0.	0.
2	4	29.315	29.360	-.45113E-01
3	4	28.170	0.	0.
4	4	27.443	27.760	-.31699
5	4	26.852	0.	0.
6	4	26.121	25.510	.61129
7	4	25.473	0.	0.
8	4	24.861	24.430	.43100
9	4	24.157	0.	0.
10	4	23.194	22.460	.73447
11	4	21.698	0.	0.
12	4	16.819	17.160	-.34134
13	4	15.552	0.	0.
14	4	15.154	15.740	-.58562
15	4	14.800	0.	0.
16	4	13.748	14.690	-.94206
1	5	36.304	36.050	.25364
2	5	34.897	0.	0.
3	5	33.923	31.640	2.2829
4	5	32.961	0.	0.
5	5	31.303	32.910	-1.6070
6	5	26.879	0.	0.
7	5	25.933	25.050	.88337
8	5	25.086	0.	0.

Fig. 4.3-10 (Continued)

9	5	24.185	25.470	-1.2851
10	5	23.088	0.	0.
11	5	21.558	20.390	1.1678
12	5	16.804	0.	0.
13	5	15.529	14.400	1.1286
14	5	15.033	0.	0.
15	5	14.428	14.370	.57542E-01
16	5	12.513	0.	0.
1	6	38.593	0.	0.
2	6	37.152	38.660	-1.5078
3	6	35.970	0.	0.
4	6	34.692	34.510	.18165
5	6	32.677	0.	0.
6	6	27.906	27.810	.96020E-01
7	6	26.682	0.	0.
8	6	25.453	25.980	-.52672
9	6	24.186	0.	0.
10	6	22.809	22.160	.64860
11	6	21.189	0.	0.
12	6	16.934	15.730	1.2037
13	6	15.585	0.	0.
14	6	14.804	14.790	.14347E-01
15	6	13.839	0.	0.
16	6	11.277	10.480	.79746
1	7	40.109	41.280	-1.1707
2	7	38.105	0.	0.
3	7	36.100	34.820	1.2804
4	7	34.096	0.	0.
5	7	32.091	32.000	.91448E-01
6	7	30.087	0.	0.
7	7	28.082	26.970	1.1125
8	7	26.078	0.	0.
9	7	24.074	25.670	-1.5965
10	7	22.069	0.	0.
11	7	20.065	18.120	1.9446
12	7	18.060	0.	0.
13	7	16.056	18.400	-2.3444
14	7	14.051	0.	0.
15	7	12.047	11.900	.14670
16	7	10.042	0.	0.

SCALED SENSITIVITY ARRAYS

PARAMETER NUMBER 1

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	.42460	.33427	.21781	.12842	.67639E-01	.23221E-01	.17658E-01	.11438E-01	.69301E-02	.42070E-02
5	1.3790	.96960	.58770	.33383	.17455	.62148E-01	.42120E-01	.25147E-01	.14488E-01	.85369E-02
4	4.2864	2.3748	1.3129	.71922	.37811	.15349	.78191E-01	.41296E-01	.22382E-01	.12761E-01
3	4.4225	2.4436	1.3495	.74089	.39801	.20081	.10307	.53841E-01	.28774E-01	.16210E-01
2	4.4727	2.4781	1.3706	.75474	.41081	.21855	.11535	.60964E-01	.32682E-01	.18396E-01
1	4.4859	2.4883	1.3774	.75936	.41496	.22331	.11895	.63211E-01	.33971E-01	.19136E-01

Fig. 4.3-10 (Continued)

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.28255E-02	.72087E-03	.18336E-03	.46369E-04	.11044E-04	0.
5	.56642E-02	.14357E-02	.36327E-03	.91483E-04	.21722E-04	0.
4	.83561E-02	.21032E-02	.52921E-03	.13268E-03	.31396E-04	0.
3	.10552E-01	.26463E-02	.66371E-03	.16593E-03	.39175E-04	0.
2	.11961E-01	.29965E-02	.75068E-03	.18745E-03	.44209E-04	0.
1	.12443E-01	.31168E-02	.78064E-03	.19487E-03	.45948E-04	0.

PARAMETER NUMBER 2

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	1.1764	.33627	.96966E-01	.28226E-01	.78876E-02	.48413E-03	.27892E-03	.15768E-03	.88360E-04	.51112E-04
5	1.1964	.34823	.10311	.31266E-01	.93267E-02	.95680E-03	.57096E-03	.31835E-03	.17541E-03	.10027E-03
4	.92323E-01	.43500E-01	.20778E-01	.10114E-01	.48670E-02	.18623E-02	.92854E-03	.48019E-03	.25571E-03	.14385E-03
3	.37494E-01	.25723E-01	.15094E-01	.83583E-02	.44461E-02	.22055E-02	.11240E-02	.58532E-03	.31241E-03	.17590E-03
2	.19261E-01	.15760E-01	.10769E-01	.66892E-02	.39015E-02	.21573E-02	.11678E-02	.62850E-03	.34148E-03	.19410E-03
1	.14734E-01	.12773E-01	.92839E-02	.60563E-02	.36716E-02	.21053E-02	.11681E-02	.63818E-03	.34979E-03	.19980E-03

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.33616E-04	.84735E-05	.21332E-05	.53473E-06	.12648E-06	0.
5	.65611E-04	.16494E-04	.41439E-05	.10372E-05	.24506E-06	0.
4	.93595E-04	.23467E-04	.58846E-05	.14709E-05	.34725E-06	0.
3	.11448E-03	.28709E-04	.72001E-05	.17999E-05	.42495E-06	0.
2	.12679E-03	.31851E-04	.79986E-05	.20015E-05	.47286E-06	0.
1	.13076E-03	.32880E-04	.82632E-05	.20688E-05	.48894E-06	0.

PARAMETER NUMBER 3

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	.68486E-03	.79644E-03	.11680E-02	.19370E-02	.34704E-02	.67933E-02	.76750E-02	.93089E-02	.10784E-01	.11154E-01
5	.17098E-02	.19879E-02	.28992E-02	.47024E-02	.80054E-02	.14646E-01	.17270E-01	.22017E-01	.26428E-01	.28013E-01
4	.35846E-02	.41811E-02	.61227E-02	.98353E-02	.15786E-01	.23355E-01	.30753E-01	.42725E-01	.54100E-01	.59421E-01
3	.37249E-02	.43691E-02	.65022E-02	.10758E-01	.18278E-01	.30724E-01	.50369E-01	.78906E-01	.10666	.12197
2	.38736E-02	.45894E-02	.70221E-02	.12164E-01	.22205E-01	.41591E-01	.78630E-01	.14334	.20879	.24703
1	.39392E-02	.46905E-02	.72771E-02	.12907E-01	.24516E-01	.49283E-01	.10659	.25694	.41083	.48563

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.97031E-02	.33537E-02	.15306E-02	.90783E-03	.48640E-03	0.
5	.24802E-01	.94432E-02	.51754E-02	.35802E-02	.20910E-02	0.
4	.54286E-01	.24890E-01	.17355E-01	.13899E-01	.88256E-02	0.
3	.11723	.72375E-01	.63036E-01	.55936E-01	.38424E-01	0.
2	.25774	.24107	.24298	.22951	.17214	0.
1	.58050	.91322	.96448	.94399	.79338	0.

Fig. 4.3-10 (Continued)

PARAMETER NUMBER 4

	1	2	3	4	5	6	7	8	9	10
7	40.109	37.435	34.761	32.087	29.414	26.740	24.066	21.392	18.718	16.044
6	15.757	15.115	14.198	13.319	12.802	13.597	12.914	11.722	10.335	8.8437
5	5.6477	5.5016	5.2713	5.1404	5.4069	6.8161	6.7679	6.3265	5.6554	4.8450
4	.47510	.49659	.58044	.79231	1.3159	2.7394	3.3709	3.3635	3.0838	2.6605
3	.33132	.35371	.43226	.60302	.93099	1.4651	1.7864	1.8439	1.7293	1.5065
2	.25812	.27780	.34237	.46645	.66402	.91391	1.0873	1.1385	1.0849	.95386
1	.23542	.25371	.31214	.41879	.57591	.75731	.88916	.93417	.89575	.79083

	11	12	13	14	15	16
7	13.370	10.696	8.0219	5.3479	2.6740	0.
6	7.2297	3.4964	2.2081	1.4007	.69068	0.
5	3.9070	1.4234	.69458	.38932	.18364	0.
4	2.1328	.67784	.26167	.12143	.52241E-01	0.
3	1.2068	.36134	.11948	.45460E-01	.17016E-01	0.
2	.76528	.22371	.68174E-01	.22351E-01	.72107E-02	0.
1	.63516	.18447	.54786E-01	.16962E-01	.50902E-02	0.

PARAMETER NUMBER 5

	1	2	3	4	5	6	7	8	9	10
7	0.	.66948	1.3390	2.0084	2.6779	3.3474	4.0169	4.6864	5.3559	6.0253
6	.15437	.32167	.57947	.88758	1.3001	2.0642	2.2908	2.6018	2.9072	3.1263
5	.10108	.14793	.24626	.39234	.63976	1.1630	1.2781	1.4285	1.5572	1.6031
4	.37238E-01	.44371E-01	.67813E-01	.11845	.22974	.50871	.67498	.77432	.83211	.82871
3	.34303E-01	.40108E-01	.59287E-01	.98265E-01	.16950	.28416	.37378	.43126	.45790	.44514
2	.31703E-01	.36432E-01	.51603E-01	.80030E-01	.12486	.18360	.23483	.26917	.28251	.26991
1	.30680E-01	.34997E-01	.48630E-01	.73247E-01	.10980	.15444	.19454	.22177	.23143	.21949

	11	12	13	14	15	16
7	6.6948	7.3643	8.0338	8.7033	9.3727	10.042
6	3.0581	2.1623	2.1660	2.4977	3.6137	8.3685
5	1.4582	.74320	.66071	.91660	2.0414	6.6948
4	.72303	.30102	.24784	.45143	1.3615	5.0211
3	.37901	.14187	.11486	.25732	.88064	3.3474
2	.22647	.79798E-01	.61139E-01	.13957	.46338	1.6737
1	.18310	.62514E-01	.42729E-01	.78653E-01	.16307	0.

Fig. 4.3-10 (Continued)

PARAMETER NUMBER 6

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	.79871E-05	.96627E-05	.15741E-04	.30536E-04	.66682E-04	.16052E-03	.20843E-03	.34846E-03	.62894E-03	.11526E-02
5	.19191E-04	.22869E-04	.35713E-04	.64832E-04	.13051E-03	.29528E-03	.39729E-03	.67777E-03	.12336E-02	.22731E-02
4	.38824E-04	.45717E-04	.68780E-04	.11589E-03	.20258E-03	.35538E-03	.54598E-03	.96766E-03	.17842E-02	.33159E-02
3	.39865E-04	.46966E-04	.70755E-04	.11950E-03	.20982E-03	.37137E-03	.65366E-03	.11995E-02	.22409E-02	.42079E-02
2	.40584E-04	.47877E-04	.72407E-04	.12314E-03	.21885E-03	.39593E-03	.72534E-03	.13534E-02	.25521E-02	.48379E-02
1	.40847E-04	.48220E-04	.73065E-04	.12466E-03	.22271E-03	.40602E-03	.75083E-03	.14080E-02	.26645E-02	.50736E-02

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.21098E-02	.12143E-01	.48541E-01	.19019	.74394	2.9089
5	.41815E-02	.24218E-01	.96940E-01	.38015	1.4876	5.8179
4	.61529E-02	.36053E-01	.14477	.56904	2.2301	8.7268
3	.79018E-02	.47142E-01	.19056	.75340	2.9661	11.636
2	.92024E-02	.56141E-01	.22957	.91942	3.6681	14.545
1	.97204E-02	.60220E-01	.24928	1.0176	4.1858	17.454

PARAMETER NUMBER 7

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.25257	-.18965	-.13125	-.11414	-.15863	-.36382	.82468E-01	.68072E-01	-.39117E-01	.14326E-01
5	-.83871	-.53074	-.32916	-.26055	-.28702	-.49265	.81000E-02	.34236E-01	-.29604E-01	.12174
4	-2.7306	-1.2019	-.66483	-.52614	-.47229	-.22372	.25675E-01	.11950E-01	-.22345E-01	.18612
3	-2.1466	-.53856	.22986E-01	.15137	.13667	.12147	.90259E-01	.22222E-02	-.32703E-01	.19942
2	-1.9174	-.29281	.27671	.39680	.34542	.24154	.11307	-.30293E-01	-.87818E-01	.14636
1	-1.8595	-.23208	.33901	.45625	.39418	.26716	.10369	-.12219	-.25436	-.37065E-01

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.68203	.15438	.35602E-01	.83070E-02	.18596E-02	0.
5	.96368	.23417	.56822E-01	.13755E-01	.31584E-02	0.
4	1.0785	.26886	.66809E-01	.16518E-01	.38585E-02	0.
3	1.1101	.27817	.69712E-01	.17421E-01	.41129E-02	0.
2	1.0631	.26639	.67407E-01	.17064E-01	.40783E-02	0.
1	.84605	.23376	.62425E-01	.16304E-01	.39673E-02	0.

Fig. 4.3-10 (Continued)

PARAMETER NUMBER 8

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-1.6907	-1.1838	-1.3760	-1.6891	-1.7057	.10911	.40204E-01	.15971E-01	.68175E-02	.32154E-02
5	-2.1290	-1.6034	-1.7519	-2.0024	-1.9284	.11218	.49736E-01	.22421E-01	.10457E-01	.52506E-02
4	.16976	.22510	.22699	.19715	.14347	.77847E-01	.41456E-01	.21324E-01	.10982E-01	.59377E-02
3	.11451	.12487	.12354	.10830	.82331E-01	.52680E-01	.31349E-01	.17860E-01	.10031E-01	.58094E-02
2	.78391E-01	.79803E-01	.77026E-01	.67868E-01	.53541E-01	.37533E-01	.24314E-01	.14952E-01	.89667E-02	.54683E-02
1	.66740E-01	.66706E-01	.63703E-01	.56233E-01	.45072E-01	.32663E-01	.21886E-01	.13874E-01	.85366E-02	.53106E-02

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.19130E-02	.45399E-03	.10842E-03	.25956E-04	.59165E-05	0.
5	.32197E-02	.77764E-03	.18849E-03	.45701E-04	.10522E-04	0.
4	.37762E-02	.93232E-03	.23037E-03	.56799E-04	.13255E-04	0.
3	.38227E-02	.96392E-03	.24274E-03	.60866E-04	.14400E-04	0.
2	.36901E-02	.94522E-03	.24146E-03	.61325E-04	.14661E-04	0.
1	.36180E-02	.93227E-03	.23945E-03	.61109E-04	.14667E-04	0.

PARAMETER NUMBER 9

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.75090E-03	-.91882E-03	-.15408E-02	-.31109E-02	-.71128E-02	-.17921E-01	-.24159E-01	-.44026E-01	-.89803E-01	-.19529
5	-.17835E-02	-.21392E-02	-.34007E-02	-.63482E-02	-.13269E-01	-.31281E-01	-.43101E-01	-.77470E-01	-.15116	-.30372
4	-.35714E-02	-.42132E-02	-.63714E-02	-.10833E-01	-.19247E-01	-.34829E-01	-.54497E-01	-.98561E-01	-.18628	-.35584
3	-.36585E-02	-.43134E-02	-.65108E-02	-.11028E-01	-.19435E-01	-.34506E-01	-.60468E-01	-.11031	-.20442	-.37908
2	-.37097E-02	-.43725E-02	-.65968E-02	-.11173E-01	-.19726E-01	-.35305E-01	-.63610E-01	-.11619	-.21318	-.38976
1	-.37268E-02	-.43925E-02	-.66275E-02	-.11231E-01	-.19860E-01	-.35669E-01	-.64625E-01	-.11811	-.21655	-.39711

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	-.45562	.74064	.31604	-.15858	-.53942	0.
5	-.61831	.80305	.38449	-.45127E-01	-.36751	0.
4	-.67817	.82227	.43703	.78442E-01	-.15591	0.
3	-.69829	.85395	.49857	.19870	.46273E-01	0.
2	-.70536	.86706	.53157	.27715	.19924	0.
1	-.73064	.51366	.15513	-.77320E-01	-.77489E-01	0.

PARAMETER NUMBER 10

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.21461	-.18829	-.13097	-.60536E-01	.28358E-01	.19023	-.86879E-01	.20201E-01	.22395	.34362
5	-.65836	-.56241	-.38278	-.18932	.70600E-02	.29502	-.25476E-02	.10859	.32648	.43432
4	-1.8354	-1.4828	-.96322	-.47670	-.10002	.58512E-01	-.28963E-02	.15778	.36891	.45367
3	-2.3761	-2.0098	-1.4723	-.96752	-.55797	-.26495	-.61871E-01	.16305	.37265	.44181
2	-2.5626	-2.1918	-1.6466	-1.1327	-.70591	-.36786	-.94794E-01	.15029	.35228	.41003
1	-2.6054	-2.2335	-1.6864	-1.1699	-.73874	-.39131	-.10687	.13537	.32736	.37799

Fig. 4.3-10 (Continued)

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.72892E-01	.16227E-01	.36702E-02	.83840E-03	.18392E-03	0.
5	.98942E-01	.23482E-01	.55634E-02	.13150E-02	.29557E-03	0.
4	.98954E-01	.24048E-01	.58335E-02	.14105E-02	.32334E-03	0.
3	.81382E-01	.19996E-01	.49446E-02	.12245E-02	.28744E-03	0.
2	.48973E-01	.12910E-01	.34317E-02	.90732E-03	.22436E-03	0.
1	.16148E-01	.71955E-02	.23960E-02	.71351E-03	.18851E-03	0.

PARAMETER NUMBER 11

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.98050	-.64065	-.18252	.29247	.66344	.54779E-01	.21941E-01	.90993E-02	.39595E-02	.18787E-02
5	-.49212	-.15779	.26482	.67563	.94213	.76941E-01	.31524E-01	.13664E-01	.62384E-02	.30912E-02
4	-.43643E-02	.21622E-01	.55541E-01	.81925E-01	.87700E-01	.45292E-01	.24525E-01	.12553E-01	.64104E-02	.34397E-02
3	.95634E-02	.17416E-01	.30623E-01	.40822E-01	.41399E-01	.28094E-01	.17269E-01	.99828E-02	.56420E-02	.32757E-02
2	.11114E-01	.13921E-01	.19373E-01	.23551E-01	.23393E-01	.18195E-01	.12486E-01	.79427E-02	.48633E-02	.30043E-02
1	.10920E-01	.12627E-01	.16142E-01	.18816E-01	.18570E-01	.15143E-01	.10882E-01	.72043E-02	.45575E-02	.28855E-02

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.11184E-02	.26518E-03	.63230E-04	.15109E-04	.34374E-05	0.
5	.18840E-02	.45339E-03	.10955E-03	.26487E-04	.60841E-05	0.
4	.21790E-02	.53667E-03	.13232E-03	.32562E-04	.75869E-05	0.
3	.21571E-02	.54410E-03	.13704E-03	.34366E-04	.81307E-05	0.
2	.20384E-02	.52377E-03	.13415E-03	.34145E-04	.81771E-05	0.
1	.19807E-02	.51259E-03	.13214E-03	.33829E-04	.81394E-05	0.

PARAMETER NUMBER 12

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.60847E-03	-.74129E-03	-.12295E-02	-.24460E-02	-.55014E-02	-.13645E-01	-.18167E-01	-.32215E-01	-.63372E-01	-.13153
5	-.14517E-02	-.17368E-02	-.27420E-02	-.50645E-02	-.10439E-01	-.24248E-01	-.33135E-01	-.58534E-01	-.11180	-.21946
4	-.29187E-02	-.34406E-02	-.51923E-02	-.87959E-02	-.15527E-01	-.27761E-01	-.43126E-01	-.77363E-01	-.14477	-.27346
3	-.29928E-02	-.35274E-02	-.53201E-02	-.90004E-02	-.15837E-01	-.28076E-01	-.49257E-01	-.89960E-01	-.16684	-.30942
2	-.30396E-02	-.35840E-02	-.54124E-02	-.91825E-02	-.16256E-01	-.29223E-01	-.53014E-01	-.97696E-01	-.18130	-.33646
1	-.30559E-02	-.36042E-02	-.54476E-02	-.92577E-02	-.16441E-01	-.29720E-01	-.54337E-01	-.10052	-.18734	-.35144

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	-.28909	-.98605	-.44765	.56873E-01	.46083	0.
5	-.43745	-1.1550	-.56407	-.86971E-01	.26945	0.
4	-.51538	-1.2268	-.64376	-.22879	.46768E-01	0.
3	-.56863	-1.3110	-.75393	-.39294	-.18604	0.
2	-.62083	-1.4737	-.96320	-.64454	-.47279	0.
1	-.66785	-1.7539	-1.3022	-1.0036	-.81719	0.

Fig. 4.3-10 (Continued)

PARAMETER NUMBER 13

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	-.23325E-01	-.23463E-01	-.24177E-01	-.26940E-01	-.37113E-01	-.73570E-01	-.74709E-01	-.73885E-01	-.71380E-01	-.65980E-01
5	-.65557E-01	-.65645E-01	-.66189E-01	-.68624E-01	-.78465E-01	-.11662	-.11681	-.11460	-.10993	-.10074
4	-.16153	-.16136	-.16074	-.15934	-.15603	-.14689	-.14140	-.13708	-.13067	-.11910
3	-.16620	-.16603	-.16544	-.16418	-.16176	-.15754	-.15346	-.14878	-.14151	-.12861
2	-.16849	-.16833	-.16781	-.16677	-.16500	-.16232	-.15896	-.15429	-.14665	-.13312
1	-.16918	-.16903	-.16853	-.16758	-.16602	-.16370	-.16055	-.15590	-.14816	-.13444

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	-.54188E-01	-.12535E-01	-.29455E-02	-.69848E-03	-.15843E-03	0.
5	-.82450E-01	-.20152E-01	-.49267E-02	-.12025E-02	-.27829E-03	0.
4	-.97270E-01	-.24332E-01	-.60760E-02	-.15108E-02	-.35479E-03	0.
3	-.10491	-.26517E-01	-.66926E-02	-.16814E-02	-.39837E-03	0.
2	-.10852	-.27555E-01	-.69887E-02	-.17645E-02	-.41998E-03	0.
1	-.10959	-.27860E-01	-.70760E-02	-.17892E-02	-.42644E-03	0.

PARAMETER NUMBER 14

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	1.5613	1.5561	1.5316	1.4450	1.1429	.83253E-01	.32228E-01	.13225E-01	.57657E-02	.27561E-02
5	1.6097	1.6040	1.5775	1.4865	1.1757	.10100	.43652E-01	.19513E-01	.90885E-02	.45685E-02
4	.18607	.18422	.17767	.16290	.13201	.70715E-01	.37063E-01	.18879E-01	.96710E-02	.52155E-02
3	.10759	.10605	.10091	.90562E-01	.72666E-01	.46979E-01	.27910E-01	.15840E-01	.88680E-02	.51241E-02
2	.69628E-01	.68402E-01	.64488E-01	.57296E-01	.46415E-01	.32981E-01	.21473E-01	.13219E-01	.79239E-02	.48282E-02
1	.58363E-01	.57257E-01	.53793E-01	.47667E-01	.38876E-01	.28540E-01	.19256E-01	.12243E-01	.75388E-02	.46891E-02

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.16497E-02	.39277E-03	.94037E-04	.22560E-04	.51501E-05	0.
5	.28045E-02	.67791E-03	.16445E-03	.39904E-04	.91932E-05	0.
4	.33141E-02	.81794E-03	.20208E-03	.49821E-04	.11627E-04	0.
3	.33684E-02	.84891E-03	.21369E-03	.53563E-04	.12669E-04	0.
2	.32565E-02	.83385E-03	.21294E-03	.54063E-04	.12921E-04	0.
1	.31938E-02	.82277E-03	.21127E-03	.53901E-04	.12934E-04	0.

PARAMETER NUMBER 15

	1	2	3	4	5	6	7	8	9	10
7	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
6	.45887E-04	.56120E-04	.93989E-04	.18942E-03	.43221E-03	.10867E-02	.14622E-02	.26529E-02	.53772E-02	.11591E-01
5	.10905E-03	.13076E-03	.20771E-03	.38728E-03	.80820E-03	.19021E-02	.26182E-02	.46959E-02	.91371E-02	.18300E-01
4	.21846E-03	.25770E-03	.38963E-03	.66227E-03	.11759E-02	.21255E-02	.33241E-02	.60100E-02	.11356E-01	.21693E-01
3	.22381E-03	.26386E-03	.39826E-03	.67453E-03	.11886E-02	.21104E-02	.37001E-02	.67559E-02	.12539E-01	.23312E-01
2	.22697E-03	.26753E-03	.40365E-03	.68378E-03	.12075E-02	.21620E-02	.38980E-02	.71269E-02	.13096E-01	.24021E-01
1	.22803E-03	.26877E-03	.40558E-03	.68745E-03	.12160E-02	.21849E-02	.39601E-02	.72382E-02	.13259E-01	.24220E-01

Fig. 4.3-10 (Continued)

	11	12	13	14	15	16
7	0.	0.	0.	0.	0.	0.
6	.26722E-01	.82142E-01	.93302E-01	.92239E-01	.76189E-01	0.
5	.37149E-01	.10166	.11611	.11445	.92746E-01	0.
4	.41358E-01	.10664	.12171	.11976	.96428E-01	0.
3	.43109E-01	.10803	.12310	.12103	.97263E-01	0.
2	.43818E-01	.10844	.12346	.12134	.97453E-01	0.
1	.44010E-01	.10854	.12353	.12139	.97487E-01	0.

ORTHOGONALIZED SENSITIVITY MATRIX
NODAL LOCATION AND VALUES FOR EACH PARAMETER
PARAMETER NOS.:

I	J	1	2	3	4	5	6	7	8
1	1	4.4859	-.48767E-01	-.13305E-02	-.14391	.91228E-03	.18274E-02	-1.2636	.74327E-01
3	1	1.3774	-.10214E-01	.56534E-02	.14887	.14289E-01	-.13827E-01	.52283	.56649E-01
5	1	.41496	-.22024E-02	.24026E-01	.49075	.39369E-01	-.52312E-01	.44241	.41208E-01
7	1	.11895	-.51572E-03	.10645	.73529	.63577E-01	-.19915	.83762E-01	.25175E-01
9	1	.33971E-01	-.13109E-03	.41079	.36070	.90364E-02	-.69833	-.38814	.16825E-01
11	1	.12443E-01	-.45378E-04	.58049	-.11588	-.60512E-01	-.95893	.65255	-.15985E-01
13	1	.78064E-03	-.27873E-05	.96448	-1.1902	-.24979	-1.3153	-.25962	-.81228E-02
15	1	.45948E-04	-.16149E-06	.79338	-1.0189	-.73071E-01	2.8583	-.23948	.21935E-02
2	2	2.4781	-.19320E-01	.16692E-02	.27465E-02	.83353E-02	-.53861E-02	.40264E-01	.83587E-01
4	2	.75474	-.39947E-02	.11273E-01	.35332	.26805E-01	-.26995E-01	.49568	.64397E-01
6	2	.21855	-.93643E-03	.41332E-01	.83021	.69598E-01	-.90550E-01	.25717	.39043E-01
8	2	.60964E-01	-.23448E-03	.14327	.94490	.99069E-01	-.27184	-.70745E-01	.22226E-01
10	2	.18396E-01	-.66313E-04	.24701	.63238	.89557E-01	-.44207	.65154E-01	.62039E-02
12	2	.29965E-02	-.10567E-04	.24107	-.87876E-01	-.16883E-01	-.34867	.18598	-.55409E-02
14	2	.18745E-03	-.65192E-06	.22951	-.27391	.68910E-01	.50794	-.55573E-01	-.84891E-03
16	2	0.	0.	0.	0.	1.6737	14.034	.70854E-01	.21577E-01
1	3	4.4225	-.25109E-01	-.14977E-02	-.23047	.10899E-01	-.99879E-03	-1.5448	.17498
3	3	1.3495	-.40100E-02	.49041E-02	.22400	.16080E-01	-.13048E-01	.20704	.13631
5	3	.39801	-.11881E-02	.17806E-01	.84850	.61686E-01	-.48309E-01	.18516	.88533E-01
7	3	.10307	-.33513E-03	.50247E-01	1.7063	.15879	-.13206	.86670E-01	.41583E-01
9	3	.28774E-01	-.94908E-04	.10663	1.5874	.23223	-.24884	-.66003E-01	.20964E-01
11	3	.10552E-01	-.34887E-04	.11722	1.0539	.20877	-.25441	1.0708	-.14754E-01
13	3	.66371E-03	-.21952E-05	.63035E-01	.38020E-01	.82732E-01	.58143E-01	.49089E-01	-.12344E-02
15	3	.39175E-04	-.12959E-06	.38424E-01	-.32584E-01	.86732	2.6359	.21273E-02	.19757E-03
2	4	2.3748	.98830E-02	.13487E-02	.81151E-03	.12708E-01	-.61779E-02	-.86493	.30706
4	4	.71922	-.66836E-04	.89812E-02	.65615	.31975E-01	-.24469E-01	-.42898	.22402
6	4	.15349	-.31050E-03	.23173E-01	2.6853	.19514	-.96729E-01	-.21125	.10289
8	4	.41296E-01	-.10439E-03	.42677E-01	3.3021	.38423	-.18657	.11510E-02	.43015E-01
10	4	.12761E-01	-.36786E-04	.59406E-01	2.5819	.51243	-.25232	.16481	.17175E-01
12	4	.21032E-02	-.63057E-05	.24888E-01	.64540	.21754	-.72226E-01	.25961	-.18296E-02
14	4	.13268E-03	-.40721E-06	.13899E-01	.10347	.43367	.41307	.12198E-01	-.14141E-02
16	4	0.	0.	0.	0.	5.0211	7.1938	.15892E-01	-.16538E-01
1	5	1.3790	1.1769	-.13320E-02	-4.2271	.34988	-.10758	.86093E-01	.26936
3	5	.58770	.94790E-01	.20881E-02	4.3902	-.27167	.82530E-01	-.18632	-1.5087
5	5	.17455	.68559E-02	.77900E-02	5.3104	.36723E-01	-.20569E-01	-.26056	-1.8665
7	5	.42120E-01	-.25278E-04	.17220E-01	6.7385	.51341	-.18086	.56323E-02	.96153E-01
9	5	.14488E-01	-.29682E-04	.26411E-01	5.6190	.91448	-.31891	-.41826E-01	.46080E-01
11	5	.56642E-02	-.14569E-04	.24796E-01	3.8741	1.0122	-.34428	.94993	.28864E-02
13	5	.36327E-03	-.99838E-06	.51750E-02	.68785	.58119	-.88823E-01	.51911E-01	-.40364E-04

Fig. 4.3-10 (Continued)

15	5	.21722E-04	-.62426E-07	.20910E-02	.18094	2.0202	.86739	-.25266E-02	-.10089E-01
2	6	.33427	.33154	.39297E-05	12.342	-1.1272	.35290	.73067E-01	-.40102
4	6	.12842	.26409E-01	.17529E-02	13.078	-.58780	.18578	-.74730E-01	-1.5300
6	6	.23221E-01	.15542E-03	.67655E-02	13.583	.53586	-.16533	-.36455	.21437
8	6	.11438E-01	-.42329E-05	.92953E-02	11.708	1.2830	-.39889	.59552E-01	.91718E-01
10	6	.42070E-02	-.84406E-05	.11149E-01	8.8287	2.1301	-.66191	-.18134E-02	.52555E-01
12	6	.72087E-03	-.17310E-05	.33528E-02	3.4920	1.7687	-.53111	.14215	.98569E-02
14	6	.46369E-04	-.12166E-06	.90778E-03	1.3995	2.3401	-.52485	-.62475E-02	-.65497E-02
16	6	0.	0.	0.	0.	8.3685	.35384	-.39070E-01	-.54652E-01
1	7	0.	0.	0.	40.109	-4.5033	1.4034	.35052E-01	.32951
3	7	0.	0.	0.	34.761	-2.5639	.80745	.21505E-01	.27575
5	7	0.	0.	0.	29.414	-.62450	.21152	.79581E-02	.22198
7	7	0.	0.	0.	24.066	1.3149	-.38442	-.55890E-02	.16822
9	7	0.	0.	0.	18.718	3.2543	-.98035	-.19136E-01	.11446
11	7	0.	0.	0.	13.370	5.1937	-1.5763	-.32683E-01	.60697E-01
13	7	0.	0.	0.	8.0219	7.1331	-2.1722	-.46230E-01	.69344E-02
15	7	0.	0.	0.	2.6740	9.0725	-2.7682	-.59777E-01	-.46828E-01
1	0	10.000	-.14156	-.11708E-01	-.56538	-.29190E-01	.28208E-01	1.3109	-.14430

PARAMETER NOS.:

I	J	9	10	11	12	13	14	15
1	1	-.15164	-.80640	-.63818E-01	-.27523	.46256E-01	-.16129	-.60053E-02
3	1	.59612E-01	-1.6521	.13613E-01	.22952	-.28859E-01	-.28362E-01	.60963E-02
5	1	.38947E-01	-.89178	.27706E-01	.17332	-.98966E-01	-.31008E-01	.78237E-02
7	1	-.35690E-01	-.13642	.15372E-01	.91525E-01	-.13652	-.70253E-01	.65468E-02
9	1	-.19530	.51341	-.46033E-02	.14389	-.13279	-.82443E-01	.69498E-03
11	1	-.55619	-.40643	-.43089E-03	-.16103	.58231E-01	.22013E-01	-.21806E-01
13	1	.27994	.18116E-01	-.16463E-01	.20066	.68297E-01	.35344E-01	-.61504E-02
15	1	.12700E-01	.50078E-01	-.13578E-01	.23021	.60914E-01	.30983E-01	.89668E-02
2	2	.39432E-02	-1.6140	-.54294E-02	.10411	.12875E-01	-.41604E-01	.18152E-02
4	2	.52232E-01	-1.2304	.33546E-01	.20197	-.68673E-01	.39383E-02	.68483E-02
6	2	.41589E-02	-.46396	.27710E-01	.10488	-.12097	-.38989E-01	.71561E-02
8	2	-.99607E-01	.19188	.10658E-01	.72889E-02	-.13889	-.76664E-01	.43161E-02
10	2	-.34005	.35873	.56451E-02	-.26053	-.72561E-01	-.45572E-01	-.63595E-02
12	2	.92906	-.16105	.18723E-01	-.39069	-.18886E-01	-.78442E-02	-.33683E-01
14	2	.30599	-.63268E-02	.13088E-02	-.10485	.12385E-01	.44911E-02	.56300E-01
16	2	-.20744E-01	.12924E-02	.14465E-01	-.52624E-01	-.21807E-01	-.99948E-02	-.23225E-01
1	3	-.18562	-.40804	-.29986E-01	-.36313	.36591E-01	-.53287E-01	-.87320E-02
3	3	.21529E-01	-1.2504	.47410E-01	.13234	-.36819E-01	.86791E-01	.30174E-02
5	3	.77043E-02	-.56052	.59368E-01	.87751E-01	-.10299	.54129E-01	.53063E-02
7	3	-.39031E-01	-.89190E-01	.26936E-01	.17900E-01	-.12917	-.42317E-01	.56225E-02
9	3	-.19214	.41047	.11047E-01	-.19288	-.11926	-.66651E-01	.66487E-03
11	3	-.54729	-.52379	.19193E-01	-.61880	.62681E-01	.30833E-01	-.14117E-01
13	3	.51481	-.51777E-01	.98968E-02	-.26212	-.86129E-02	-.57829E-02	.49523E-01
15	3	.48914E-01	-.31607E-02	.10291E-02	-.99609E-01	.72433E-02	.83913E-03	.72799E-01
2	4	-.10535	-.36889	.64264E-01	-.18224	-.23925E-01	.22560	-.57904E-02
4	4	-.59280E-01	-.29563E-01	.12210	-.85607E-01	-.96979E-01	.26529	-.10118E-02
6	4	-.52601E-01	.22359	.64924E-01	-.65418E-01	-.12976	.60129E-01	.30855E-02
8	4	-.86204E-01	.16402	.22029E-01	-.80875E-01	-.11900	-.48688E-01	.34707E-02
10	4	-.32125	.36393	.13362E-01	-.42006	-.63526E-01	-.36375E-01	-.39981E-02
12	4	.85928	-.16344	.22827E-01	-.48133	-.26056E-01	-.11884E-01	-.11648E-01
14	4	.82619E-01	-.11664E-01	.90981E-03	-.14188	.10032E-01	.12865E-02	.95124E-01
16	4	.49299E-04	-.12526E-01	-.67027E-02	.13409E-01	-.70561E-02	-.29903E-02	-.19549E-01
1	5	.89444E-02	.14888	.17337	.35074E-01	.27653E-01	-.36604E-01	.22414E-03
3	5	-.23319E-01	-.29353	-.11449	-.42987E-01	-.30378E-01	.65931E-01	-.92757E-03

Fig. 4.3-10 (Continued)

5	5	-.40189E-01	-.46636E-01	.42302	.50351E-02	.14549E-01	-.18482E-01	.10996E-03
7	5	-.30112E-01	.40174E-02	.49748E-01	-.85513E-02	-.90908E-01	.25535E-01	.26000E-02
9	5	-.14215	.34932	.14963E-01	-.19025	-.89538E-01	-.48143E-01	.31608E-03
11	5	-.49100	-.43511	.18624E-01	-.58473	.64733E-01	.31331E-01	-.88849E-02
13	5	.39442	-.44547E-01	.73454E-02	-.23815	-.49626E-02	-.50952E-02	.60250E-01
15	5	-.36303	.73253E-02	-.12098E-01	.73024E-02	.31129E-01	.10550E-01	.11205
2	6	.17084E-01	-.74005E-01	-.59363	-.56037E-01	-.79480E-01	.84858E-01	.72463E-03
4	6	-.21361E-02	-.18682	-.12711	-.87791E-02	-.12313E-01	.41270E-02	-.32952E-04
6	6	-.43890E-01	.41142	.91514E-01	-.78664E-01	-.54148E-01	.15829	-.24325E-02
8	6	-.17484E-01	-.22462E-01	.22126E-01	.11889E-01	-.46893E-01	-.21319E-01	.56178E-03
10	6	-.17572	.33388	.82454E-02	-.23343	-.32060E-01	-.26019E-01	-.37203E-02
12	6	.76910	-.10984	.15195E-01	-.34151	-.20030E-01	-.12267E-01	-.15070E-01
14	6	-.14897	-.58071E-02	-.99005E-02	-.32281E-01	.23954E-01	.64641E-02	.92845E-01
16	6	.20842E-01	-.26344E-01	-.27871E-01	.79442E-01	.76946E-02	.40141E-02	-.15873E-01
1	7	.30262E-01	-.33679E-02	.58950E-01	.51030E-01	.42298E-01	-.26276E-01	.50633E-02
3	7	.30392E-01	-.73526E-02	.45963E-01	.59221E-01	.38668E-01	-.21771E-01	.25168E-02
5	7	.30522E-01	-.11337E-01	.32975E-01	.67411E-01	.35037E-01	-.17265E-01	-.29660E-04
7	7	.30653E-01	-.15322E-01	.19988E-01	.75601E-01	.31407E-01	-.12759E-01	-.25762E-02
9	7	.30783E-01	-.19307E-01	.70009E-02	.83792E-01	.27777E-01	-.82536E-02	-.51227E-02
11	7	.30913E-01	-.23292E-01	-.59863E-02	.91982E-01	.24146E-01	-.37479E-02	-.76691E-02
13	7	.31044E-01	-.27276E-01	-.18974E-01	.10017	.20516E-01	.75778E-03	-.10216E-01
15	7	.31174E-01	-.31261E-01	-.31961E-01	.10836	.16885E-01	.52635E-02	-.12762E-01
1	0	.16798	1.5859	-.98925E-03	.23271	.58746E-02	.25236E-01	.49318E-02

Fig. 4.3-10 (Continued)

Program Listing.

```
C       FINITE DIFFERENCE PROGRAM FOR NONLINEAR REGRESSION SOLUTION
C       OF TWO-DIMENSIONAL, STEADY-STATE, GROUND-WATER FLOW PROBLEMS
C       BY R. L. COOLEY, USGS, DENVER, COLO.
      DIMENSION TITLE(20),DX(30),DY(30),T(500),SL(500),QRE(500)
     1,WELL(500),HR(500),W(500),HO(500),HC(500),TRANX(20),TRANY(20)
     2,VLEAK(20),QDIST(20),CXS(500),CYS(500),VLS(500),QRS(500),P(20)
     3,WP(20),QBND(10),QBF(50),PLA(50),PLB(50),CXHR(50),CXHL(50)
     4,CYHT(50),CYHB(50),S(20,70),X(20,110),Z(500),A(20,20),V(500)
      DIMENSION AU(5,250),AL(20,250)
      DIMENSION JPOS(30),IZN(500),IBZN(50),IPRM(4,20),IBPA(10),IBPB(10)
     1,IBNA(50),IBNB(50),IBHN(50),IHSN(100),NCB(20),NCE(20)
     2,ILOC(500),JLOC(500),IN(500),IC(5,250)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NQSD,NBH,NVAR
     1,NVX2,NVX3,NBQP,IPO,KOUNT,INDT,IBHZ
      COMMON/TNME/IIN,IOUT
      COMMON/FLT/AU,AL
      COMMON/REG/BMX,AP,CSA,AMP,RP,BP,YSQ
      EQUIVALENCE (TITLE(1),A(1,1),X(1,1),AU(1,1)),(ILOC(1),CXS(1))
     1,(JLOC(1),CYS(1)),(HO(1),Z(1))
C**DEFINE INPUT, OUTPUT, AND SCRATCH FILES, AND ARRAY DIMENSIONS
      IIN=5
      IOUT=6
      ITA=7
      NVD=20
C**READ THREE TITLE CARDS
      WRITE(IOUT,804)
      DO 5 I=1,3
      READ(IIN,801) (TITLE(J),J=1,20)                                  CARD A
    5 WRITE(IOUT,803) (TITLE(J),J=1,20)
C**READ JOB SPECIFICATION DATA
      READ(IIN,800) ID,JD,NZNS,NPAR,NWELS,NQBND,NBQP,NBHZ,NBHP,NUM,IPRX
     1,IPO,ISO                                                         CARD B
      WRITE(IOUT,802) ID,JD,NZNS,NPAR,NWELS,NQBND,NBQP,NBHZ,NBHP,NUM
     1,IPRX,IPO,ISO
      READ(IIN,820) BMX,CSA,RP,BP,EV                                   CARD C
      WRITE(IOUT,806) BMX,CSA,RP,BP,EV
C**READ INITIAL ARRAY DATA
      IM=ID-1
      JM=JD-1
      CALL ARRAY(DX,IM,1,1,0)                                          CARD D
      CALL ARRAY(DY,JM,1,2,0)                                          CARD E
      CALL ARRAY(HO,ID,JD,3,0)                                         CARD F
      NIJ=ID*JD
      DO 10 N=1,NIJ
   10 HC(N)=HO(N)
      CALL ARRAY(W,ID,JD,4,0)                                          CARD G
      CALL ARRAY(T,IM,JM,5,0)                                          CARD H
      CALL ARRAY(SL,IM,JM,6,0)                                         CARD I
      CALL ARRAY(HR,ID,JD,7,0)                                         CARD J
      CALL ARRAY(QRE,IM,JM,8,0)                                        CARD K
C**READ GRID ZONATION
      CALL ARRAYI(IZN,IM,JM,1,0)                                       CARD L
C**READ INITIAL AQUIFER PARAMETERS BY ZONE
```

```
      WRITE(IOUT,810)
      DO 30 J=1,NZNS
      READ(IIN,812) I,TRANX(I),TRANY(I),VLEAK(I),QDIST(I)          CARD M
      WRITE(IOUT,814) I,TRANX(I),TRANY(I),VLEAK(I),QDIST(I)
      DO 20 K=1,4
   20 IPRM(K,J)=0
   30 CONTINUE
C**READ AQUIFER PARAMETER NO'S
      NBPAR=NBHP+NBQP
      IF(NPAR.LE.0) GO TO 45
      WRITE(IOUT,816)
      DO 40 J=1,NZNS
      READ(IIN,800) I,(IPRM(K,I),K=1,4)                            CARD N
      WRITE(IOUT,818) I,(IPRM(K,I),K=1,4)
      DO 35 K=1,4
   35 IPRM(K,I)=IPRM(K,I)+NBPAR
   40 CONTINUE
C**READ AQUIFER PARAMETER COEFFICIENTS OF VARIATION
      READ(IIN,820) (WP(K+NBPAR),K=1,NPAR)                         CARD O
      WRITE(IOUT,822)
      CALL PRTOT(WP(NBPAR+1),NPAR,0,1)
   45 JPOS(1)=0
      DO 50 J=2,JD
   50 JPOS(J)=JPOS(J-1)+ID
C**READ POINT FLOW DATA
      DO 55 N=1,NIJ
   55 WELL(N)=0.
      IF(NWELS.LE.0) GO TO 61
      WRITE(IOUT,824)
      DO 60 K=1,NWELS
      READ(IIN,826) I,J,WELL(I+JPOS(J))                            CARD P
      WRITE(IOUT,828) I,J,WELL(I+JPOS(J))
   60 CONTINUE
C**READ AND FORM ARRAYS FOR SPECIFIED POINT OR LINE FLOWS
   61 NQSD=0
      IF(NQBND.LE.0) GO TO 85
      WRITE(IOUT,830)
      N=0
      DO 80 J=1,NQBND
      READ(IIN,832) IA,JA,IB,JB,IZ,QB,CVQB,QBM                     CARD Q
      WRITE(IOUT,831) IA,JA,IB,JB,IZ,QB,CVQB,QBM
      M=1
      K=IA-1
      IF(JA.EQ.JB) GO TO 62
      M=ID
      K=JA-1
   62 MA=IA+JPOS(JA)
      MB=IB+JPOS(JB)-M
      IF(MB.GE.MA) GO TO 64
      IF(IZ.LE.0) GO TO 63
      N=N+1
      IBNA(N)=MA
      IBNB(N)=MA
      QBF(N)=.5*QBM
      IBZN(N)=IZ
```

```
      GO TO 68
   63 WELL(MA)=QB*QBM
      GO TO 80
   64 QBM=.5*QBM
      IF(IZ.LE.0) GO TO 70
      DO 66 L=MA,MB,M
      N=N+1
      IBNA(N)=L
      IBNB(N)=L+M
      K=K+1
      TEMP=DX(K)
      IF(M.EQ.ID) TEMP=DY(K)
      QBF(N)=QBM*TEMP
   66 IBZN(N)=IZ
   68 QBND(IZ)=QB
      WP(IZ)=CVQB
      GO TO 80
   70 TMP=QB*QBM
      DO 75 L=MA,MB,M
      K=K+1
      TEMP=DX(K)
      IF(M.EQ.ID) TEMP=DY(K)
      TEMP=TMP*TEMP
      WELL(L)=WELL(L)+TEMP
   75 WELL(L+M)=WELL(L+M)+TEMP
   80 CONTINUE
      NQSD=N
C**READ SPECIFIED BOUNDARY HEAD POSITIONS AS -1'S
   85 CALL ARRAYI(IN,ID,JD,2,0)                                    CARD R
C**READ DATA AND FORM ARRAYS FOR SPECIFIED HEADS AND PARAMETERS
      IBHZ=0
      IF(NBHZ.LE.0) GO TO 110
      WRITE(IOUT,833)
      NBH=0
      DO 108 KK=1,NBHZ
      READ(IIN,834) IZ,NN,M,N,CVHA,CVHB                            CARD S
      WRITE(IOUT,836) IZ,NN,M,N,CVHA,CVHB
      DO 95 J=1,NN
      READ(IIN,826) ILOC(J),JLOC(J),V(J)                           CARD S
   95 WRITE(IOUT,840) ILOC(J),JLOC(J),V(J)
      IBPA(IZ)=M+NBQP
      IBPB(IZ)=N+NBQP
      J=JLOC(1)
      K=ILOC(1)+JPOS(J)
      IBNA(IZ+NQSD)=K
      J=JLOC(NN)
      IBNB(IZ+NQSD)=ILOC(NN)+JPOS(J)
      IF(M.GT.0) WP(M+NBQP)=CVHA
      IF(N.GT.0) WP(N+NBQP)=CVHB
      M=M+N
      NBHS=NBH
      IF(IN(K).LT.-1) GO TO 100
      NBH=NBH+1
      IF(M.GT.0) IN(K)=-NBH-1
      IBZN(NBH+NQSD)=IZ
```

```
      IBHN(NBH)=K
      PLA(NBH)=V(1)
      PLB(NBH)=0.
  100 IF(W(K).GT.0.) IBHZ=1
      IF(NN.LE.1) GO TO 107
      DIST=0.
      DO 102 KNT=2,NN
      J=JLOC(KNT)
      L=ILOC(KNT)+JPOS(J)
      IF(W(L).GT.0.) IBHZ=1
      NBH=NBH+1
      IF(M.GT.0) IN(L)=-NBH-1
      IBZN(NBH+NQSD)=IZ
      IBHN(NBH)=L
      JM1=JLOC(KNT-1)
      IF(J.EQ.JM1) GO TO 101
      J=MIN0(J,JM1)
      DIST=DIST+DY(J)
      GO TO 102
  101 I=MIN0(ILOC(KNT),ILOC(KNT-1))
      DIST=DIST+DX(I)
  102 PLB(NBH)=DIST
      N=NBH-NN+1
      DO 106 KNT=2,NN
      J=JLOC(KNT)
      L=ILOC(KNT)+JPOS(J)
      N=N+1
      TMP=PLB(N)/DIST
      TMPA=TMP*V(NN)
      TMPB=(1.-TMP)*V(1)
      TMPC=TMPA+TMPB
      IF(ABS(V(KNT)).LE.0.) GO TO 104
      TMP=V(KNT)/TMPC
      TMPA=TMPA*TMP
      TMPB=TMPB*TMP
      TMPC=V(KNT)
  104 PLA(N)=TMPB
      PLB(N)=TMPA
  106 HC(L)=TMPC
  107 IF(M.LE.0) NBH=NBHS
  108 HC(K)=V(1)
      IF(NBHP.LE.0) IBHZ=0
C**COMPARE T WITH IZN FOR CONFLICT
  110 IER=0
      N=0
      DO 115 J=1,JM
      DO 115 I=1,IM
      N=N+1
      IF(IZN(N).LE.0) GO TO 115
      IF(T(N).GT.0.) GO TO 115
      IER=1
      WRITE(IOUT,842) I,J
  115 CONTINUE
      IF(IER.LT.1) GO TO 120
      WRITE(IOUT,844)
```

```
      STOP
C**TRANSFER DOMAIN GEOMETRY TO IN(M)
  120 N=0
      DO 122 J=1,JM
      DO 122 I=1,IM
      N=N+1
      IF(IZN(N).LE.0) GO TO 122
      M=N+J
      IF(IN(M).GE.0) IN(M)=1
      IF(IN(M-1).GE.0) IN(M-1)=1
      IF(IN(M+ID-1).GE.0) IN(M+ID-1)=1
      IF(IN(M+ID).GE.0) IN(M+ID)=1
  122 CONTINUE
C**SET UP D4 ORDERING
      CALL ORDER(JPOS,IN,IC)
C**ADJUST DX AND DY
      DO 130 I=1,IM
  130 DX(I)=.5*DX(I)
      DO 135 J=1,JM
  135 DY(J)=.5*DY(J)
C**COMPUTE INITIAL SOLUTION
      CALL COEF(DX,DY,T,SL,QRE,WELL,HR,HC,TRANX,TRANY,VLEAK,QDIST,CXS
     1,CYS,VLS,QRS,CXHR,CXHL,CYHT,CYHB,QBND,QBF,V,IZN,IBZN,IBNA,IBNB,IN
     2,IC)
      CALL D4SOLV(HC,V,IN,IC)
      WRITE(IOUT,846)
      CALL ARRAY(HC,ID,JD,0,1)
      IF(ISO.EQ.1) GO TO 640
      NVAR=NPAR+NBPAR
      NVX2=NVAR+NVAR
      NVX3=NVX2+NVAR
C**COMPUTE AND COUNT PRIOR INFORMATION DATA
      NPRIR=0
      DO 137 I=1,NVAR
      P(I)=1.
      IF(WP(I).LE.0.) GO TO 137
      WP(I)=EV/(WP(I)*WP(I))
      NPRIR=NPRIR+1
  137 CONTINUE
C**COMPUTE INITIAL SUM SQUARE ERRORS
      NOBS=0
      YSQ=0.
      DO 140 N=1,NIJ
      IF (W(N).LE.0.) GO TO 140
      IF(IN(N).EQ.-1) GO TO 138
      NOBS=NOBS+1
      TMP=HO(N)-HC(N)
      YSQ=YSQ+TMP*W(N)*TMP
      GO TO 140
  138 W(N)=0.
  140 CONTINUE
      WRITE(IOUT,848) NOBS,NPRIR,YSQ
C**ORDER IBZN AND CORRESPONDING ARRAYS FOR LINE FLOW PARAMETERS
C  FROM SMALLEST TO LARGEST
      IF(NQSD.LT.1) GO TO 158
```

```
      DO 150 I=1,NQSD
      DO 148 J=I,NQSD
      IF(IBZN(J).GE.IBZN(I)) GO TO 148
      ITMP=IBZN(I)
      IBZN(I)=IBZN(J)
      IBZN(J)=ITMP
      ITMP=IBNA(I)
      IBNA(I)=IBNA(J)
      IBNA(J)=ITMP
      ITMP=IBNB(I)
      IBNB(I)=IBNB(J)
      IBNB(J)=ITMP
      TMP=QBF(I)
      QBF(I)=QBF(J)
      QBF(J)=TMP
  148 CONTINUE
  150 CONTINUE
C**DEFINE BEGINNING AND END POINT ARRAYS FOR LINE FLOW PARAMETERS
      NCB(1)=1
      IB=1
      DO 156 J=1,NBQP
      DO 152 I=IB,NQSD
      II=I
      IF(IBZN(I).GT.J) GO TO 154
  152 CONTINUE
      NCE(J)=II
      GO TO 156
  154 NCE(J)=II-1
      NCB(J+1)=II
      IB=II
  156 CONTINUE
C**DEFINE SEQUENCE NO.S, AND BEGINNING AND END POINT ARRAYS
C  FOR SPECIFIED HEAD PARAMETERS
  158 IF(NBHP.LT.1) GO TO 166
      NTMP=NBQP+1
      NCB(NTMP)=1
      L=0
      DO 164 IP=NTMP,NBPAR
      DO 162 J=1,NBH
      I=IBZN(J+NQSD)
      K=IBPA(I)
      IF(K.EQ.IP) GO TO 160
      K=IBPB(I)
      IF(K.EQ.IP) GO TO 160
      GO TO 162
  160 L=L+1
      IHSN(L)=J
  162 CONTINUE
      NCE(IP)=L
  164 NCB(IP+1)=L+1
C**DEFINE BEGINNING AND END POINT ARRAYS FOR AQUIFER PARAMETERS
  166 IF(NPAR.LT.1) GO TO 174
      DO 168 I=1,NPAR
  168 NCB(I+NBPAR)=0
      N=0
```

```
      DO 172 J=1,JM
      DO 172 I=1,IM
      N=N+1
      L=IZN(N)
      IF(L.LT.1) GO TO 172
      DO 170 M=1,4
      K=IPRM(M,L)
      IF(K.LE.NBPAR) GO TO 170
      NCE(K)=N
      IF(NCB(K).LE.0) NCB(K)=N
  170 CONTINUE
  172 CONTINUE
C**BEGIN ITERATIONS
  174 INDT=0
      ER=.01
      ERP=1000.
      AMP=0.
      KOUNT=0
  176 KOUNT=KOUNT+1
      REWIND ITA
C**SOLVE FOR SCALED SENSITIVITIES:
      DO 260 N=1,NVAR
      DO 178 I=1,NEQ
  178 V(I)=0.
      LB=NCB(N)
      LE=NCE(N)
C**ASSEMBLE R.H.S. FOR SPECIFIED LINE FLOW PARAMETERS
      IF(N.GT.NBQP) GO TO 182
      DO 180 L=LB,LE
      I=IBNA(L)
      J=IN(I)
      TMP=QBND(N)*QBF(L)
      IF(J.GT.0) V(J)=V(J)+TMP
      I=IBNB(L)
      J=IN(I)
      IF(J.GT.0) V(J)=V(J)+TMP
  180 CONTINUE
      GO TO 216
C**ASSEMBLE R.H.S. FOR SPECIFIED HEAD PARAMETERS
  182 IF(N.GT.NBPAR) GO TO 200
      DO 190 L=LB,LE
      K=IHSN(L)
      I=IBZN(K+NQSD)
      TMP=0.
      IF(IBPA(I).EQ.N) TMP=PLA(K)
      IF(IBPB(I).EQ.N) TMP=TMP+PLB(K)
      I=IBHN(K)+1
      IF(I.LE.NIJ) J=IN(I)
      IF(J.GT.0) V(J)=V(J)+CXHR(K)*TMP
      I=IBHN(K)-1
      IF(I.GT.0) J=IN(I)
      IF(J.GT.0) V(J)=V(J)+CXHL(K)*TMP
      I=IBHN(K)+ID
      IF(I.LE.NIJ) J=IN(I)
      IF(J.GT.0) V(J)=V(J)+CYHT(K)*TMP
```

```
      I=IBHN(K)-ID
      IF(I.GT.0) J=IN(I)
      IF(J.GT.0) V(J)=V(J)+CYHB(K)*TMP
  190 CONTINUE
      GO TO 216
C**ASSEMBLE R.H.S. FOR X-DIRECTION TRANSMISSIVITY PARAMETERS
  200 IZ=IZN(LB)
      IF(IPRM(1,IZ).NE.N) GO TO 204
      DO 202 L=LB,LE
      J=IZN(L)
      K=IPRM(1,J)
      IF(K.NE.N) GO TO 202
      NA=L+(L-1)/(ID-1)
      NB=NA+1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      IF(INA.GT.0) V(INA)=V(INA)+CXS(L)*(HC(NB)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+CXS(L)*(HC(NA)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+CXS(L)*(HC(ND)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+CXS(L)*(HC(NC)-HC(ND))
  202 CONTINUE
C**ASSEMBLE R.H.S. FOR Y-DIRECTION TRANSMISSIVITY PARAMETERS
  204 IF(IPRM(2,IZ).NE.N) GO TO 208
      DO 206 L=LB,LE
      J=IZN(L)
      K=IPRM(2,J)
      IF(K.NE.N) GO TO 206
      NA=L+(L-1)/(ID-1)
      NB=NA+1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      IF(INA.GT.0) V(INA)=V(INA)+CYS(L)*(HC(ND)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+CYS(L)*(HC(NC)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+CYS(L)*(HC(NB)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+CYS(L)*(HC(NA)-HC(ND))
  206 CONTINUE
C**ASSEMBLE R.H.S. FOR VERTICAL LEAKAGE PARAMETERS
  208 IF(IPRM(3,IZ).NE.N) GO TO 212
      DO 210 L=LB,LE
      J=IZN(L)
      K=IPRM(3,J)
      IF(K.NE.N) GO TO 210
      NA=L+(L-1)/(ID-1)
      NB=NA+1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
```

```
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      IF(INA.GT.0) V(INA)=V(INA)+VLS(L)*(HR(NA)-HC(NA))
      IF(INB.GT.0) V(INB)=V(INB)+VLS(L)*(HR(NB)-HC(NB))
      IF(INC.GT.0) V(INC)=V(INC)+VLS(L)*(HR(NC)-HC(NC))
      IF(IND.GT.0) V(IND)=V(IND)+VLS(L)*(HR(ND)-HC(ND))
  210 CONTINUE
C**ASSEMBLE R.H.S. FOR RECHARGE PARAMETERS
  212 IF(IPRM(4,IZ).NE.N) GO TO 216
      DO 214 L=LB,LE
      J=IZN(L)
      K=IPRM(4,J)
      IF(K.NE.N) GO TO 214
      NA=L+(L-1)/(ID-1)
      NB=NA+1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      IF(INA.GT.0) V(INA)=V(INA)+QRS(L)
      IF(INB.GT.0) V(INB)=V(INB)+QRS(L)
      IF(INC.GT.0) V(INC)=V(INC)+QRS(L)
      IF(IND.GT.0) V(IND)=V(IND)+QRS(L)
  214 CONTINUE
C**MODIFY R.H.S.--UPPER HALF
  216 DO 220 J=1,ICR1
      II=IC(1,J)
      DO 218 I=2,II
      LR=IC(I,J)
      V(LR)=V(LR)-AU(I,J)*V(J)
  218 CONTINUE
  220 V(J)=V(J)/AU(1,J)
C**MODIFY R.H.S.--LOWER HALF
      JJ=NEQ-ICR
      DO 224 J=1,JJ
      JR=J+ICR1
      LR=JR
      DO 222 I=2,IB1
      LR=LR+1
      IF (AL(I,J).NE.0.) V(LR)=V(LR)-AL(I,J)*V(JR)
  222 CONTINUE
  224 V(JR)=V(JR)/AL(1,J)
C**BACK SOLVE--LOWER HALF
      V(NEQ)=V(NEQ)/AL(1,NEQ-ICR1)
      DO 230 J=1,JJ
      KK=NEQ-J
      KL=KK-ICR1
      L=KK
      DO 226 I=2,IB1
      L=L+1
      IF (AL(I,KL).NE.0.) V(KK)=V(KK)-AL(I,KL)*V(L)
  226 CONTINUE
```

```
  230 CONTINUE
C**BACK SOLVE--UPPER HALF
      DO 250 J=1,ICR1
      KK=ICR-J
      II=IC(1,KK)
      DO 240 I=2,II
      L=IC(I,KK)
      V(KK)=V(KK)-AU(I,KK)*V(L)
  240 CONTINUE
  250 CONTINUE
      WRITE(ITA) (V(I),I=1,NEQ)
      L=0
      DO 255 I=1,NIJ
      K=IN(I)
      IF(W(I).LE.0..OR.K.LT.1) GO TO 255
      L=L+1
      S(N,L)=V(K)
  255 CONTINUE
  260 CONTINUE
      IF(IPO.NE.1) GO TO 270
      WRITE(IOUT,850)
      N=0
      L=0
      DO 265 J=1,JD
      DO 265 I=1,ID
      N=N+1
      IF(W(N).LE.0..OR.IN(N).LT.1) GO TO 265
      L=L+1
      WRITE(IOUT,852) I,J,(S(K,L),K=1,NVAR)
  265 CONTINUE
C**CALL LEAST SQUARES
  270 CALL LSTSQ(W,HO,HC,P,WP,PLA,PLB,S,A,V,IBZN,IBPA,IBPB,IBHN,IN,NVD)
      IF(INDT.GT.0) GO TO 521
      IF(IPO.EQ.1) WRITE(IOUT,854)
C**COMPUTE NEW SPECIFIED FLOW PARAMETERS
      IF(NBQP.LE.0) GO TO 282
      DO 280 K=1,NBQP
      QBND(K)=(V(K)+1.)*QBND(K)
      IF(IPO.EQ.1) WRITE(IOUT,856) K,QBND(K)
  280 CONTINUE
C**COMPUTE NEW SPECIFIED HEAD PARAMETERS
  282 IF(NBHP.LE.0) GO TO 290
      DO 286 N=1,NBH
      M=IBZN(N+NQSD)
      K=IBPA(M)
      L=IBPB(M)
      IF(K.GT.NBQP) PLA(N)=PLA(N)*(V(K)+1.)
      IF(L.GT.NBQP) PLB(N)=PLB(N)*(V(L)+1.)
      J=IBHN(N)
      HC(J)=PLA(N)+PLB(N)
  286 CONTINUE
      IF(IPO.NE.1) GO TO 290
      DO 288 J=1,NBHZ
      K=IBNA(J+NQSD)
      L=IBNB(J+NQSD)
```

```
      WRITE(IOUT,856) J,HC(K),HC(L)
  288 CONTINUE
C**COMPUTE NEW AQUIFER PARAMETERS
  290 DO 300 K=1,NZNS
      L=IPRM(1,K)
      IF(L.GT.NBPAR) TRANX(K)=TRANX(K)*(V(L)+1.)
      L=IPRM(2,K)
      IF(L.GT.NBPAR) TRANY(K)=TRANY(K)*(V(L)+1.)
      L=IPRM(3,K)
      IF(L.GT.NBPAR) VLEAK(K)=VLEAK(K)*(V(L)+1.)
      L=IPRM(4,K)
      IF(L.GT.NBPAR) QDIST(K)=QDIST(K)*(V(L)+1.)
      IF(IPO.EQ.1) WRITE(IOUT,814) K,TRANX(K),TRANY(K),VLEAK(K),QDIST(K)
  300 CONTINUE
C**CHECK FOR CONVERGENCE
      TMP=0.
      IF(AP.GT.1.E-10) TMP=(1.+AMP)/AP
      DO 330 K=1,NVAR
      IF(ABS(V(K)*TMP).GT.ER) GO TO 335
  330 CONTINUE
      GO TO 350
  335 IF(KOUNT.EQ.NUM) GO TO 340
C**CALCULATE NEW SCALED PRIOR INFORMATION PARAMETERS
      IND=0
      DO 337 I=1,NVAR
      TEMP=V(I)+1.
      P(I)=P(I)/TEMP
      IF(ABS(P(I)).LT.ERP) GO TO 337
      WRITE(IOUT,858) I
      IND=1
  337 WP(I)=WP(I)*TEMP*TEMP
      IF(IND.GT.0) GO TO 360
C**COMPUTE NEW HEADS AT GRID POINTS
      CALL COEF(DX,DY,T,SL,QRE,WELL,HR,HC,TRANX,TRANY,VLEAK,QDIST,CXS
     1,CYS,VLS,QRS,CXHR,CXHL,CYHT,CYHB,QBND,QBF,V,IZN,IBZN,IBNA,IBNB,IN
     2,IC)
      CALL D4SOLV(HC,V,IN,IC)
      GO TO 176
  340 WRITE(IOUT,860) NUM
      GO TO 360
  350 WRITE(IOUT,862) KOUNT
C**COMPUTE FINAL ESTIMATES OF HEAD
  360 REWIND ITA
      DO 366 K=1,NVAR
      READ(ITA) (CXS(I),I=1,NEQ)
      DO 364 J=1,NIJ
      L=IN(J)
      IF(L.GT.0) HC(J)=HC(J)+CXS(L)*V(K)
  364 CONTINUE
  366 CONTINUE
C**COMPUTE SUM OF SQUARED ERRORS
      TEMP=AP*(2.-AP)
      TMPA=2.*RP/(2.-AP)
      SUM=0.
      DO 380 I=1,NVAR
```

```
      TMP=P(I)-1.
  380 SUM=SUM+TEMP*V(I+NVAR)*(V(I+NVX2)-TMPA*V(I+NVX3)
     1*(BP*P(I)-1.-.5*V(I)))-TMP*WP(I)*TMP
      YSQ=YSQ-SUM
C**CORRECT A FOR MARQUARDT PARAMETER
      IF (NVAR.EQ.1) GO TO 440
      IF(AMP.LE.0.) GO TO 386
      DO 384 I=1,NVAR
      A(I,I)=1.+RP
      DO 382 J=I,NVAR
  382 A(J,I)=A(I,J)
  384 CONTINUE
      AMP=-1.
      CALL LSTSQ(W,HO,HC,P,WP,PLA,PLB,S,A,V,IBZN,IBPA,IBPB,IBHN,IN,NVD)
      IF(INDT.GT.0) GO TO 521
C**COMPUTE A-INVERSE
  386 A(NVAR,NVAR)=1./A(NVAR,NVAR)
      NM1=NVAR-1
      DO 430 K=1,NM1
      KP1=K+1
      DO 400 I=KP1,NVAR
      SUM=0.
      IM1=I-1
      DO 390 J=K,IM1
  390 SUM=SUM+A(I,J)*A(J,K)
      A(K,I)=-SUM
  400 A(I,K)=-SUM*A(I,I)
      DO 420 I=1,K
      SUM=A(K,I)
      DO 410 J=KP1,NVAR
  410 SUM=SUM+A(J,K)*A(I,J)
      A(K,I)=SUM
  420 A(I,K)=A(K,I)
  430 CONTINUE
      DO 435 J=1,NVAR
  435 A(J,NVAR)=A(NVAR,J)
      GO TO 450
  440 A(1,1)=1./(1.+RP)
C**COMPUTE TR((A-INVERSE)**2) AND A-INVERSE - RP*(A-INVERSE)**2
  450 TRACE=0.
      IF(RP.LE.0.) GO TO 459
      DO 457 N=1,NVAR
      DO 452 J=1,NVAR
  452 V(J)=A(J,N)
      SUMA=0.
      DO 456 J=N,NVAR
      SUM=0.
      DO 454 I=1,NVAR
  454 SUM=SUM+V(I)*A(I,J)
      V(J+NVAR)=SUM
  456 A(J,N)=A(J,N)-RP*SUM
  457 TRACE=TRACE+V(N+NVAR)
C**COMPUTE ERROR VARIANCE
  459 TEMP=NPRIR-NVAR
      OBS=NOBS
```

```
      VAR=YSQ/(OBS+TEMP+RP*RP*TRACE)
C**COMPUTE CORRELATION COEFFICIENT
      SUMA=0.
      SUMB=0.
      SUMC=0.
      SUMD=0.
      SUM=0.
      DO 460 N=1,NIJ
      IF (W(N).LE.0.) GO TO 460
      TMP=W(N)**.5
      W(N)=TMP
      TEMP=TMP*HO(N)
      TMP=TMP*HC(N)
      SUMA=SUMA+TEMP
      SUMB=SUMB+TMP
      SUMC=SUMC+TEMP*TEMP
      SUMD=SUMD+TMP*TMP
      SUM=SUM+TEMP*TMP
  460 CONTINUE
      R=(OBS*SUM-SUMA*SUMB)/((OBS*SUMC-SUMA*SUMA)*(OBS*SUMD-SUMB*SUMB))
     1**.5
C**PRINT ERROR VARIANCE, ESTIMATED SUM OF SQUARED ERRORS, AND CORRELATION
C**COEFFICIENT
      WRITE(IOUT,864) VAR,YSQ,R
C**COMPUTE SCALED VARIANCE-COVARIANCE MATRIX
      DO 463 J=1,NVAR
      TEMP=V(J+NVX3)
      DO 462 I=J,NVAR
      A(I,J)=VAR*A(I,J)/(V(I+NVX3)*TEMP)
  462 A(J,I)=A(I,J)
  463 V(J)=A(J,J)**.5
C**PRINT PARAMETERS AND STANDARD ERRORS
      IF(NBQP.LE.0) GO TO 466
      WRITE(IOUT,866)
      DO 464 J=1,NBQP
      STDER=QBND(J)*V(J)
  464 WRITE(IOUT,856) J,QBND(J),STDER
  466 IF(NBHP.LE.0) GO TO 470
      WRITE(IOUT,868)
      DO 468 J=1,NBHZ
      K=IBNA(J+NQSD)
      L=IBPA(J)
      STERA=0.
      IF(L.GT.NBQP) STERA=HC(K)*V(L)
      M=IBNB(J+NQSD)
      N=IBPB(J)
      STERB=0.
      IF(N.GT.NBQP) STERB=HC(M)*V(N)
  468 WRITE(IOUT,856) J,HC(K),STERA,HC(M),STERB
  470 WRITE(IOUT,870)
      DO 480 J=1,NZNS
      WRITE(IOUT,814) J,TRANX(J),TRANY(J),VLEAK(J),QDIST(J)
  480 CONTINUE
      WRITE(IOUT,872)
      DO 490 J=1,NZNS
```

```
      K=IPRM(1,J)
      STERX=0.
      IF (K.GT.NBPAR) STERX=TRANX(J)*V(K)
      K=IPRM(2,J)
      STERY=0.
      IF (K.GT.NBPAR) STERY=TRANY(J)*V(K)
      K=IPRM(3,J)
      STERV=0.
      IF (K.GT.NBPAR) STERV=VLEAK(J)*V(K)
      K=IPRM(4,J)
      STERQ=0.
      IF (K.GT.NBPAR) STERQ=QDIST(J)*V(K)
      WRITE(IOUT,814) J,STERX,STERY,STERV,STERQ
  490 CONTINUE
C**PRINT SCALED VARIANCE-COVARIANCE MATRIX
      WRITE(IOUT,874)
      CALL PRTOT(A,NVAR,NVD,0)
C**COMPUTE AND PRINT CORRELATION MATRIX FOR SCALED PARAMETERS
      DO 510 J=1,NVAR
      TEMP=V(J)
      DO 500 I=J,NVAR
      A(I,J)=A(I,J)/(V(I)*TEMP)
  500 A(J,I)=A(I,J)
  510 CONTINUE
      WRITE(IOUT,876)
      CALL PRTOT(A,NVAR,NVD,0)
C**COMPUTE AND PRINT RESIDUALS
      WRITE(IOUT,878)
      N=0
      DO 520 J=1,JD
      DO 520 I=1,ID
      N=N+1
      IF(IN(N).EQ.0) GO TO 520
      RES=W(N)*(HC(N)-HO(N))
      WRITE(IOUT,880) I,J,HC(N),HO(N),RES
  520 CONTINUE
C**PRINT SCALED SENSITIVITIES FOR EACH NODE
  521 IF(IPRX.LE.0.AND.KOUNT.LT.NUM) STOP
      WRITE(IOUT,882)
      REWIND ITA
      DO 530 KK=1,NVAR
      READ(ITA) (V(I),I=1,NEQ)
      WRITE(IOUT,884) KK
      DO 525 N=1,NIJ
      Z(N)=0.
      L=IN(N)
      IF(L.GT.0) GO TO 523
      IF(L.GE.-1) GO TO 525
      L=-L-1
      IZ=IBZN(L+NQSD)
      TMP=0.
      IF(IBPA(IZ).EQ.KK) TMP=PLA(L)
      IF(IBPB(IZ).EQ.KK) TMP=TMP+PLB(L)
      Z(N)=TMP
      GO TO 525
```

```
  523 Z(N)=V(L)
  525 CONTINUE
  530 CALL ARRAY(Z,ID,JD,0,1)
      IF(NVAR.LT.2) STOP
C**ORTHOGONALIZE COLUMNS OF SENSITIVITY MATRIX (S):
C**DEFINE I AND J POINTERS AND SCALE S USING W
      K=0
      M=0
      N=0
      DO 535 J=1,JD
      DO 535 I=1,ID
      K=K+1
      IF(W(K).LT.1.E-10) GO TO 535
      N=N+1
      ILOC(N)=I
      JLOC(N)=J
      L=IN(K)
      IF(L.GT.0) GO TO 533
      DO 532 KK=1,NVAR
  532 X(KK,N)=0.
      L=-L-1
      IZ=IBZN(L+NQSD)
      KK=IBPA(IZ)
      IF(KK.GT.NBQP) X(KK,N)=PLA(L)*W(K)
      KK=IBPB(IZ)
      IF(KK.GT.NBQP) X(KK,N)=X(KK,N)+PLB(L)*W(K)
      GO TO 535
  533 M=M+1
      DO 534 KK=1,NVAR
  534 X(KK,N)=S(KK,M)*W(K)
  535 CONTINUE
      IF(NPRIR.LE.0) GO TO 539
      DO 538 I=1,NVAR
      IF(WP(I).LT.1.E-10) GO TO 538
      N=N+1
      ILOC(N)=I
      JLOC(N)=0
      DO 537 J=1,NVAR
  537 X(J,N)=0.
      X(I,N)=WP(I)**.5
  538 CONTINUE
C**ORTHOGONALIZE S
  539 NTMP=NOBS+NPRIR
      DO 540 I=1,NTMP
  540 W(I)=X(1,I)
      DO 600 N=2,NVAR
      NM1=N-1
      SUM=0.
      DO 550 I=1,NTMP
      SUM=SUM+W(I)*W(I)
      X(NM1,I)=W(I)
  550 CONTINUE
      IF(SUM.LT.1.E-20) GO TO 610
      V(NM1)=1./SUM
      DO 570 J=1,NM1
```

```
      SUM=0.
      DO 560 K=1,NTMP
  560 SUM=SUM+V(J)*X(J,K)*X(N,K)
  570 T(J)=SUM
      DO 590 K=1,NTMP
      SUM=0.
      DO 580 I=1,NM1
  580 SUM=SUM+X(I,K)*T(I)
  590 W(K)=X(N,K)-SUM
  600 CONTINUE
C**PRINT ORTHOGONALIZED S
  610 WRITE(IOUT,886)
      K=1
      L=8
      DO 630 M=1,NVAR,8
      IF(L.GT.NVAR) L=NVAR
      WRITE(IOUT,888) (I,I=K,L)
      DO 620 J=1,NTMP
      X(NVAR,J)=W(J)
      WRITE(IOUT,890) ILOC(J),JLOC(J),(X(I,J),I=K,L)
  620 CONTINUE
      K=K+8
      L=L+8
  630 CONTINUE
      STOP
C**READ, PRINT, AND EXECUTE FOR ALTERNATE SOLUTIONS
  640 READ(IIN,800) N                                              CARD T
      IF(N.LE.0) STOP
      DO 690 K=1,N
      WRITE(IOUT,892) K
      DO 650 L=1,NZNS
      READ(IIN,812) I,TRANX(I),TRANY(I),VLEAK(I),QDIST(I)          CARD U
  650 WRITE(IOUT,814) I,TRANX(I),TRANY(I),VLEAK(I),QDIST(I)
      IF(NBQP.LE.0) GO TO 665
      DO 660 L=1,NBQP
  660 READ(IIN,812) I,QBND(I)                                      CARD V
      WRITE(IOUT,894)
      CALL PRTOT(QBND,NBQP,0,1)
  665 IF(NBHZ.LE.0) GO TO 685
      WRITE(IOUT,896)
      DO 670 M=1,NBHZ
      READ(IIN,812) I,HBA,HBB                                      CARD W
      J=IBNA(I+NQSD)
      L=IBNB(I+NQSD)
      Z(J)=HBA/HC(J)
      Z(L)=HBB/HC(L)
  670 WRITE(IOUT,856) I,HBA,HBB
      DO 680 NN=1,NBH
      M=IBZN(NN+NQSD)+NQSD
      J=IBNA(M)
      L=IBNB(M)
      KK=IBHN(NN)
      PLA(NN)=Z(J)*PLA(NN)
      PLB(NN)=Z(L)*PLB(NN)
      HC(KK)=PLA(NN)+PLB(NN)
```

```
  680 CONTINUE
  685 CALL COEF(DX,DY,T,SL,QRE,WELL,HR,HC,TRANX,TRANY,VLEAK,QDIST,CXS
     1,CYS,VLS,QRS,CXHR,CXHL,CYHT,CYHB,QBND,QBF,V,IZN,IBZN,IBNA,IBNB,IN
     2,IC)
      CALL D4SOLV(HC,V,IN,IC)
      WRITE(IOUT,898) K
      CALL ARRAY(HC,ID,JD,0,1)
  690 CONTINUE
      STOP
C
  800 FORMAT (16I5)
  801 FORMAT (20A4)
  802 FORMAT(56HONUMBER OF COLUMNS (ID) ------------------------------ =
     $,I5
     $/56H NUMBER OF ROWS (JD) ---------------------------------- =,I5
     $/56H NUMBER OF AQUIFER ZONES (NZNS) ----------------------- =,I5
     $/56H NUMBER OF AQUIFER PARAMETERS (NPAR) ------------------ =,I5
     $/56H NUMBER OF KNOWN POINT FLOWS (NWELS) ------------------ =,I5
     $/56H NUMBER OF SPECIFIED FLOW ZONES (NQBND) --------------- =,I5
     $/56H NUMBER OF SPECIFIED FLOW PARAMETERS (NBQP) ----------- =,I5
     $/56H NUMBER OF SPECIFIED HEAD ZONES (NBHZ) ---------------- =,I5
     $/56H NUMBER OF SPECIFIED HEAD PARAMETERS (NBHP) ----------- =,I5
     $/56H MAXIMUM NUMBER OF ITERATIONS (NUM) ------------------- =,I5
     $/56H SENSITIVITY PRINT AND ORTHOGONALIZATION OPTION (IPRX) =,I5
     $/56H ADDITIONAL PRINTOUT OPTION (IPO) --------------------- =,I5
     $/56H SOLUTION ONLY OPTION (ISO) --------------------------- =,I5)
  803 FORMAT (1H ,20A4)
  804 FORMAT (1H1)
  806 FORMAT (50H MAXIMUM ALLOWABLE PARAMETER CORRECTION (BMX) - =
     $,G11.5
     $/50H SEARCH DIRECTION ADJUSTMENT PARAMETER (CSA) -- = ,G11.5
     $/50H RIDGE PARAMETER FOR REGRESSION (RP) ---------- = ,G11.5
     $/50H BIAS PARAMETER FOR REGRESSION (BP) ----------- = ,G11.5
     $/50H ESTIMATED ERROR VARIANCE (EV) ---------------- = ,G11.5)
  810 FORMAT (1HO,12X,34HINITIAL AQUIFER PARAMETERS BY ZONE/6H  ZONE
     1,5X,5HTRANX,8X,5HTRANY,8X,5HVLEAK,8X,5HQDIST)
  812 FORMAT (I5,4F10.0)
  814 FORMAT (1H ,I4,2X,4(2X,G11.5))
  816 FORMAT (1HO,11X,25HAQUIFER PARAMETER NUMBERS/1H ,5X,4HZONE,4X
     1,5HTRANX,3X,5HTRANY,3X,5HVLEAK,3X,5HQDIST)
  818 FORMAT (1H ,8I8)
  820 FORMAT (8F10.0)
  822 FORMAT (1HO,12X,48HCOEFFICIENTS OF VARIATION FOR AQUIFER PARAMETER
     1S/1H ,3(3X,4HPAR.,8X,5HCOEF.,4X)/1H ,3(4X,3HNO.,9X,4HVAR.,4X))
  824 FORMAT (1HO,11X,11HPOINT FLOWS/1H ,7X,1HI,7X,1HJ,4X,9HVOL. RATE)
  826 FORMAT (2I5,F10.0)
  828 FORMAT (1H ,2I8,4X,G11.5)
  830 FORMAT (1HO,22X,27HINITIAL SPECIFIED FLOW DATA/1H ,6X,9HNODE NO.S
     1,7X,4HPAR.,6X,4HFLOW,9X,5HCOEF./1H ,19H  IA   JA   IB   JB,4X
     2,3HNO.,4X,9HPARAMETER,7X,4HVAR.,5X,10HMULTIPLIER)
  831 FORMAT (1H ,4(1X,I3,1X),2X,I3,3X,3(2X,G11.5))
  832 FORMAT (5I5,3F10.0)
  833 FORMAT (1HO,18X,27HINITIAL SPECIFIED HEAD DATA)
  834 FORMAT(4I5,2F10.0)
  836 FORMAT (1HO,22H  NO. OF NODES IN ZONE,I4,3H = ,I3/1H ,15H  NO. PAR
```

```
     1. A = ,I3,17X,13HNO. PAR. B = ,I3/1H ,22H  COEF. VAR. PAR. A =
     2,G11.5,22H  COEF. VAR. PAR. B = ,G11.5/1H ,20X,22HINITIAL VALUES O
     3F HEAD/1H ,21X,1HI,5X,1HJ,8X,4HHEAD)
  840 FORMAT (1H ,19X,2(I3,3X),2X,G11.5)
  842 FORMAT(9HOAT CELL ,2I5,15H, IZN>O AND T=O)
  844 FORMAT(54HOPROGRAM ABORTED BECAUSE OF CONFLICT BETWEEN IZN AND T)
  846 FORMAT (1HO,1X,17HINITIAL SOLUTION:)
  848 FORMAT (23HONO. OF OBSERVATIONS = ,I4/46H NO. OF PARAMETERS HAVING
     1 PRIOR INFORMATION = ,I4//56H ESTIMATED SUM OF SQUARED ERRORS FOR
     2INITIAL SOLUTION = ,G11.5)
  850 FORMAT(1HO,5X,58HNODAL LOCATION AND SCALED SENSITIVITIES FOR EACH
     1PARAMETER)
  852 FORMAT(1H ,2(1X,I4),7(1X,G11.5)/1H ,(10X,7(1X,G11.5)))
  854 FORMAT(19HOUPDATED PARAMETERS)
  856 FORMAT (1H ,1X,I4,3X,4(G11.5,4X))
  858 FORMAT (11HOPARAMETER ,I3,17H EFFECTIVELY ZERO)
  860 FORMAT (//32HOSOLUTION FAILED TO CONVERGE IN ,I3,11H ITERATIONS)
  862 FORMAT (//23HOSOLUTION CONVERGED IN ,I3,11H ITERATIONS)
  864 FORMAT (18HOERROR VARIANCE = ,G11.5/35H ESTIMATED SUM OF SQUARED E
     1RRORS = ,G11.5/27H CORRELATION COEFFICIENT = ,G11.5)
  866 FORMAT(1HO,1X,35HESTIMATED SPECIFIED FLOW PARAMETERS
     1/7H   ZONE,5X,4HQBND,9X,8HSTD. ER.)
  868 FORMAT (1HO,19X,25HESTIMATED SPECIFIED HEADS/1H ,11X,2(4HHEAD
     1,9X,8HSTD. ER.,9X)/1H ,6H  ZONE,2(7X,1HA,6X),2(9X,1HB,4X))
  870 FORMAT (1HO,11X,36HESTIMATED AQUIFER PARAMETERS BY ZONE/6H  ZONE
     1,5X,5HTRANX,8X,5HTRANY,8X,5HVLEAK,8X,5HQDIST)
  872 FORMAT (1HO,13X,33HESTIMATED STANDARD ERRORS BY ZONE/1H ,9X
     1,8HSTD. ER.,5X,8HSTD. ER.,5X,8HSTD. ER.,5X,8HSTD. ER./6H  ZONE,5X
     2,5HTRANX,8X,5HTRANY,8X,5HVLEAK,8X,5HQDIST)
  874 FORMAT (34HOSCALED VARIANCE-COVARIANCE MATRIX)
  876 FORMAT (41HOCORRELATION MATRIX FOR SCALED PARAMETERS)
  878 FORMAT (1HO,22X,16HSAMPLE RESIDUALS/1H ,5X,1HI,4X,1HJ,6X
     1,9HPREDICTED,7X,8HOBSERVED,8X,8HWEIGHTED/1H ,19X,5HVALUE,10X
     2,5HVALUE,10X,8HRESIDUAL)
  880 FORMAT (1H ,2X,2(I4,1X),3(5X,G11.5))
  882 FORMAT (26HOSCALED SENSITIVITY ARRAYS)
  884 FORMAT(18HOPARAMETER NUMBER ,I5)
  886 FORMAT(1HO,5X,33HORTHOGONALIZED SENSITIVITY MATRIX
     1/1H ,45H NODAL LOCATION AND VALUES FOR EACH PARAMETER)
  888 FORMAT (1H ,14X,15HPARAMETER NOS.:/1H ,4X,1HI,4X,1HJ,5X,8(I3,9X))
  890 FORMAT (1H ,2(1X,I4),8(1X,G11.5))
  892 FORMAT (1HO,4X,47HAQUIFER PARAMETERS FOR ADDITIONAL SOLUTION NO.
     1,I3/6H  ZONE,5X,5HTRANX,8X,5HTRANY,8X,5HVLEAK,8X,5HQDIST)
  894 FORMAT (1HO,23X,25HSPECIFIED FLOW PARAMETERS/1H ,3X,3(4HFLOW,5X
     1,9HSPECIFIED,6X)/1H ,3X,3(4HZONE,8X,4HFLOW,8X))
  896 FORMAT (1HO,4X,25HSPECIFIED HEAD PARAMETERS/1H ,6H  HEAD,5X
     1,2(4HPAR.,11X)/1H ,6H  ZONE,6X,1HA,14X,1HB)
  898 FORMAT(13HOSOLUTION NO.,I5)
      END
      SUBROUTINE ARRAY(A,IND,JND,N,IT)
C**IF IT=0, SUBROUTINE FOR LOADING 1 AND 2 DIMENSIONAL ARRAYS
C**IF IT=1, SUBROUTINE FOR PRINTING 2 DIMENSIONAL ARRAYS
      DIMENSION A(IND,JND),NME(8)
      COMMON/TNME/IIN,IOUT
      DATA NME/4HDX  ,4HDY  ,4HHO  ,4HW   ,4HT   ,4HSL  ,4HHR  ,4HQRE /
```

```
      IF(IT.EQ.1) GO TO 55
      DO 5 J=1,JND
      DO 5 I=1,IND
    5 A(I,J)=0.
      WRITE(IOUT,105)
      READ(IIN,70) NOBL,IPRN
      DO 50 K=1,NOBL
      READ(IIN,70) IB,IE,JB,JE,FACT,IVAR
      WRITE(IOUT,80) K,IB,IE,JB,JE,NME(N),FACT
      IF (IVAR.GT.0) GO TO 20
      DO 10 J=JB,JE
      DO 10 I=IB,IE
   10 A(I,J)=FACT
      GO TO 50
   20 DO 40 J=JB,JE
      READ(IIN,90) (A(I,J),I=IB,IE)
      DO 40 I=IB,IE
   40 A(I,J)=A(I,J)*FACT
   50 CONTINUE
      IF (IPRN.GT.0) RETURN
      WRITE(IOUT,100) NME(N)
   55 DO 60 K=1,IND,10
      I10=K+9
      IF(I10.GT.IND) I10=IND
      WRITE(IOUT,110) (I,I=K,I10)
      WRITE(IOUT,105)
      DO 60 J=1,JND
      JR=JND-J+1
   60 WRITE(IOUT,120) JR,(A(I,JR),I=K,I10)
      RETURN
C
   70 FORMAT (4I5,F10.0,3I5)
   80 FORMAT (1H ,I3,2X,5HIB = ,I5,2X,5HIE = ,I5,2X,5HJB = ,I5,2X
     1,5HJE = ,I5,2X,A4,2H = ,G11.5)
   90 FORMAT (8F10.0)
  100 FORMAT (1H0,2X,A4,7H ARRAY:)
  105 FORMAT (1H )
  110 FORMAT (1H0,10X,10(I3,9X))
  120 FORMAT (1H ,1X,I3,1X,10(1X,G11.5))
      END
      SUBROUTINE ARRAYI(INT,IND,JND,N,IT)
C**IF IT=0, SUBROUTINE FOR LOADING 1 AND 2 DIMENSIONAL INTEGER ARRAYS
C**IF IT=1, SUBROUTINE FOR PRINTING 2 DIMENSIONAL INTEGER ARRAYS
      DIMENSION INT(IND,JND),NME(2)
      COMMON/TNME/IIN,IOUT
      DATA NME/4HIZN ,4HIBZN/
      IF(IT.EQ.1) GO TO 45
      DO 5 J=1,JND
      DO 5 I=1,IND
    5 INT(I,J)=0
      WRITE(IOUT,100)
      READ(IIN,60) NOBL,IPRN
      DO 40 K=1,NOBL
      READ(IIN,60) IB,IE,JB,JE,IFACT,IVAR
      WRITE(IOUT,70) K,IB,IE,JB,JE,NME(N),IFACT
```

```
      IF (IVAR.GT.0) GO TO 20
      DO 10 J=JB,JE
      DO 10 I=IB,IE
   10 INT(I,J)=IFACT
      GO TO 40
   20 DO 30 J=JB,JE
      READ(IIN,60) (INT(I,J),I=IB,IE)
   30 CONTINUE
   40 CONTINUE
      IF (IPRN.GT.0) RETURN
      WRITE(IOUT,80) NME(N)
   45 DO 50 K=1,IND,30
      I30=K+29
      IF(I30.GT.IND) I30=IND
      WRITE(IOUT,90) (I,I=K,I30)
      WRITE(IOUT,100)
      DO 50 J=1,JND
      JR=JND-J+1
   50 WRITE(IOUT,110) JR,(INT(I,JR),I=K,I30)
      RETURN
C
   60 FORMAT (16I5)
   70 FORMAT (1H ,I3,2X,5HIB = ,I5,2X,5HIE = ,I5,2X,5HJB = ,I5,2X
     1,5HJE = ,I5,2X,A4,2H = ,I5)
   80 FORMAT (1H0,2X,A4,7H ARRAY:)
   90 FORMAT (1H0,4X,30(1X,I3))
  100 FORMAT (1H )
  110 FORMAT (1H ,31(1X,I3))
      END
      SUBROUTINE ORDER(JPOS,IN,IC)
      DIMENSION JPOS(1),IN(1),IC(5,1)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD
      COMMON/TNME/IIN,IOUT
C**COMPUTE EQUATION NUMBERS FOR D4 ORDERING:
      NXP=ID+JD-1
      K=0
C**ORDER--LEFT TO RIGHT, BOTTOM TO TOP
      DO 20 I=1,NXP,2
      DO 20 J=1,JD
      IK=I-J+1
      IF(IK.LT.1.OR.IK.GT.ID) GO TO 20
      N=IK+JPOS(J)
      IF(IN(N).LE.0) GO TO 20
      K=K+1
      IN(N)=K
   20 CONTINUE
      ICR=K+1
      DO 30 I=2,NXP,2
      DO 30 J=1,JD
      IK=I-J+1
      IF(IK.LT.1.OR.IK.GT.ID) GO TO 30
      N=IK+JPOS(J)
      IF(IN(N).LE.0) GO TO 30
      K=K+1
      IN(N)=K
```

```
   30 CONTINUE
C**COMPUTE BAND WIDTH AND DETERMINE CONNECTING EQUATION NUMBERS:
      MNO=9999
      MXO=0
      N=0
      IND=0
      DO 80 J=1,JD
      DO 80 I=1,ID
      N=N+1
      JR=IN(N)
      IF (JR.LE.0.OR.JR.GE.ICR) GO TO 80
      IU=1
C**BELOW
      IF ((J-1).LT.1) GO TO 40
      IF (IN(N-ID).LE.0) GO TO 40
      IU=IU+1
      IC(IU,JR)=IN(N-ID)
      MM=IN(N-ID)-JR
      MXO=MAXO(MM,MXO)
      MNO=MINO(MM,MNO)
C**LEFT
   40 IF ((I-1).LT.1) GO TO 50
      IF (IN(N-1).LE.0) GO TO 50
      IU=IU+1
      IC(IU,JR)=IN(N-1)
      MM=IN(N-1)-JR
      MNO=MINO(MM,MNO)
      MXO=MAXO(MM,MXO)
C**RIGHT
   50 IF ((I+1).GT.ID) GO TO 60
      IF (IN(N+1).LE.0) GO TO 60
      IU=IU+1
      IC(IU,JR)=IN(N+1)
      MM=IN(N+1)-JR
      MXO=MAXO(MM,MXO)
      MNO=MINO(MM,MNO)
C**ABOVE
   60 IF ((J+1).GT.JD) GO TO 70
      IF (IN(N+ID).LE.0) GO TO 70
      IU=IU+1
      IC(IU,JR)=IN(N+ID)
      MM=IN(N+ID)-JR
      MXO=MAXO(MM,MXO)
      MNO=MINO(MM,MNO)
   70 IC(1,JR)=IU
      IF(IU.GT.1) GO TO 80
      WRITE(IOUT,110) I,J
      IND=1
   80 CONTINUE
      NEQ=K
      ICR1=ICR-1
      IB1=MXO-MNO+1
      LH1=NEQ-ICR1
      WRITE(IOUT,90)
      WRITE(IOUT,100) ICR1,IB1,LH1,ICR1,NEQ
```

```
      IF(IND.GT.0) STOP
      RETURN
C
   90 FORMAT (51HOSOLUTION BY LDU FACTORIZATION ASSUMING D4 ORDERING)
  100 FORMAT (82H ****WARNING*****MINIMUM DIMENSIONS FOR ARRAYS USED BY
     1THIS METHOD ARE AS FOLLOWS:/1H ,12H AU:    5 BY,I5/1H ,4H AL:
     2,I5,3H BY,I5/1H ,12H IC:    5 BY,I5/1H ,4H  B:,I5)
  110 FORMAT (1HO,6HNODE (,I3,1H,,I3,30H) ISOLATED. SOLUTION SINGULAR.)
      END
      SUBROUTINE COEF(DX,DY,T,SL,QRE,WELL,HR,HC,TRANX,TRANY,VLEAK,QDIST
     1,CXS,CYS,VLS,QRS,CXHR,CXHL,CYHT,CYHB,QBND,QBF,B,IZN,IBZN,IBNA,IBNB
     2,IN,IC)
      DIMENSION DX(1),DY(1),T(1),SL(1),QRE(1),WELL(1),HR(1),HC(1)
     1,TRANX(1),TRANY(1),VLEAK(1),QDIST(1),CXS(1),CYS(1),VLS(1),QRS(1)
     2,CXHR(1),CXHL(1),CYHT(1),CYHB(1),QBND(1),QBF(1),B(1)
      DIMENSION IZN(1),IBZN(1),IBNA(1),IBNB(1),IN(1),IC(5,1)
      DIMENSION AU(5,250),AL(20,250)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NQSD,NBH
      COMMON/FLT/AU,AL
C**INITIALIZE ARRAYS
      DO 10 J=1,ICR1
      DO 10 I=1,5
   10 AU(I,J)=0.
      DO 20 J=1,LH1
      DO 20 I=1,IB1
   20 AL(I,J)=0.
      DO 40 I=1,NIJ
      N=IN(I)
      IF(N.GT.0) B(N)=WELL(I)
   40 CONTINUE
      DO 42 I=1,NBH
      CXHR(I)=0.
      CXHL(I)=0.
      CYHT(I)=0.
   42 CYHB(I)=0.
C**CALCULATE B FOR SPECIFIED FLOW PARAMETERS
      IF(NQSD.LE.0) GO TO 52
      DO 50 I=1,NQSD
      IZ=IBZN(I)
      TMP=QBND(IZ)*QBF(I)
      INA=IBNA(I)
      L=IN(INA)
      IF(L.GT.0) B(L)=B(L)+TMP
      INB=IBNB(I)
      L=IN(INB)
      IF(L.GT.0) B(L)=B(L)+TMP
   50 CONTINUE
C**BEGIN MAIN LOOP
   52 N=0
      DO 150 J=1,JM
      DYN=DY(J)
      DO 150 I=1,IM
      N=N+1
      M=IZN(N)
      IF(M.LE.0) GO TO 150
```

```
      NB=N+J
      NA=NB-1
      NC=NB+ID
      ND=NA+ID
      INA=IN(NA)
      INB=IN(NB)
      INC=IN(NC)
      IND=IN(ND)
      DXN=DX(I)
      CX=TRANX(M)*T(N)*DYN/(DXN+DXN)
      CXS(N)=CX
      CY=TRANY(M)*T(N)*DXN/(DYN+DYN)
      CYS(N)=CY
      AREA=DXN*DYN
      VL=VLEAK(M)*SL(N)*AREA
      VLS(N)=VL
      QRT=QDIST(M)*QRE(N)*AREA
      QRS(N)=QRT
      E=CX+CY+VL
C**CALCULATE AU, AL, B, AND COEFFICIENT ARRAYS FOR SPECIFIED HEAD PARAMETERS
      K=-INA-1
      IF(K) 60,75,53
   53 CXHR(K)=CXHR(K)+CX
      CYHT(K)=CYHT(K)+CY
      GO TO 75
   60 IF(INA.GE.ICR) GO TO 65
      AU(1,INA)=AU(1,INA)+E
      AU(4,INA)=AU(4,INA)-CX
      AU(5,INA)=AU(5,INA)-CY
      GO TO 70
   65 AL(1,INA-ICR1)=AL(1,INA-ICR1)+E
   70 B(INA)=B(INA)+QRT+VL*(HR(NA)-HC(NA))+CX*(HC(NB)-HC(NA))
     1+CY*(HC(ND)-HC(NA))
   75 K=-INB-1
      IF(K) 85,100,77
   77 CXHL(K)=CXHL(K)+CX
      CYHT(K)=CYHT(K)+CY
      GO TO 100
   85 IF(INB.GE.ICR) GO TO 90
      AU(1,INB)=AU(1,INB)+E
      AU(3,INB)=AU(3,INB)-CX
      AU(5,INB)=AU(5,INB)-CY
      GO TO 95
   90 AL(1,INB-ICR1)=AL(1,INB-ICR1)+E
   95 B(INB)=B(INB)+QRT+VL*(HR(NB)-HC(NB))+CX*(HC(NA)-HC(NB))
     1+CY*(HC(NC)-HC(NB))
  100 K=-INC-1
      IF(K) 110,125,102
  102 CXHL(K)=CXHL(K)+CX
      CYHB(K)=CYHB(K)+CY
      GO TO 125
  110 IF(INC.GE.ICR) GO TO 115
      AU(1,INC)=AU(1,INC)+E
      AU(2,INC)=AU(2,INC)-CY
      AU(3,INC)=AU(3,INC)-CX
```

```
      GO TO 120
  115 AL(1,INC-ICR1)=AL(1,INC-ICR1)+E
  120 B(INC)=B(INC)+QRT+VL*(HR(NC)-HC(NC))+CX*(HC(ND)-HC(NC))
     1+CY*(HC(NB)-HC(NC))
  125 K=-IND-1
      IF(K) 135,150,127
  127 CXHR(K)=CXHR(K)+CX
      CYHB(K)=CYHB(K)+CY
      GO TO 150
  135 IF(IND.GE.ICR) GO TO 140
      AU(1,IND)=AU(1,IND)+E
      AU(2,IND)=AU(2,IND)-CY
      AU(4,IND)=AU(4,IND)-CX
      GO TO 145
  140 AL(1,IND-ICR1)=AL(1,IND-ICR1)+E
  145 B(IND)=B(IND)+QRT+VL*(HR(ND)-HC(ND))+CX*(HC(NC)-HC(ND))
     1+CY*(HC(NA)-HC(ND))
  150 CONTINUE
C**COMPRESS AU
      N=0
      DO 190 J=1,JD
      DO 190 I=1,ID
      N=N+1
      K=IN(N)
      IF(K.LE.0.OR.K.GT.ICR1) GO TO 190
      IF(IC(1,K).EQ.5) GO TO 190
      IU=1
      IF((J-1).LT.1) GO TO 160
      IF(IN(N-ID).LE.0) GO TO 160
      IU=IU+1
      AU(IU,K)=AU(2,K)
  160 IF((I-1).LT.1) GO TO 170
      IF(IN(N-1).LE.0) GO TO 170
      IU=IU+1
      AU(IU,K)=AU(3,K)
  170 IF((I+1).GT.ID) GO TO 180
      IF(IN(N+1).LE.0) GO TO 180
      IU=IU+1
      AU(IU,K)=AU(4,K)
  180 IF((J+1).GT.JD) GO TO 190
      IF(IN(N+ID).LE.0) GO TO 190
      IU=IU+1
      AU(IU,K)=AU(5,K)
  190 CONTINUE
      RETURN
      END
      SUBROUTINE D4SOLV(HC,B,IN,IC)
      DIMENSION HC(1),B(1)
      DIMENSION IN(1),IC(5,1)
      DIMENSION AU(5,250),AL(20,250)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1
      COMMON/FLT/AU,AL
C**DECOMPOSE TO FILL AL
      DO 280 J=1,ICR1
      II=IC(1,J)
```

```
      DO 270 I=2,II
      LR=IC(I,J)
      L=LR-ICR1
      C=AU(I,J)/AU(1,J)
      DO 260 K=I,II
      KL=IC(K,J)-LR+1
      AL(KL,L)=AL(KL,L)-C*AU(K,J)
  260 CONTINUE
      AU(I,J)=C
      B(LR)=B(LR)-C*B(J)
  270 CONTINUE
  280 B(J)=B(J)/AU(1,J)
C**DECOMPOSE AL
      JJ=NEQ-ICR
      DO 310 J=1,JJ
      JR=J+ICR1
      L=J
      DO 300 I=2,IB1
      L=L+1
      IF (AL(I,J).EQ.0.) GO TO 300
      LR=L+ICR1
      C=AL(I,J)/AL(1,J)
      KL=0
      DO 290 K=I,IB1
      KL=KL+1
      IF (AL(K,J).NE.0.) AL(KL,L)=AL(KL,L)-C*AL(K,J)
  290 CONTINUE
      AL(I,J)=C
      B(LR)=B(LR)-C*B(JR)
  300 CONTINUE
  310 B(JR)=B(JR)/AL(1,J)
C**BACK SOLVE--LOWER HALF
      B(NEQ)=B(NEQ)/AL(1,NEQ-ICR1)
      DO 330 J=1,JJ
      K=NEQ-J
      KL=K-ICR1
      L=K
      DO 320 I=2,IB1
      L=L+1
      IF (AL(I,KL).NE.0.) B(K)=B(K)-AL(I,KL)*B(L)
  320 CONTINUE
  330 CONTINUE
C**BACK SOLVE--UPPER HALF
      DO 350 J=1,ICR1
      K=ICR-J
      II=IC(1,K)
      DO 340 I=2,II
      L=IC(I,K)
      B(K)=B(K)-AU(I,K)*B(L)
  340 CONTINUE
  350 CONTINUE
C**COMPUTE HC+DELTHC
      DO 360 N=1,NIJ
      L=IN(N)
      IF(L.LE.0) GO TO 360
```

```
      HC(N)=HC(N)+B(L)
  360 CONTINUE
      RETURN
      END
      SUBROUTINE LSTSQ(W,HO,HC,P,WP,PLA,PLB,S,C,B,IBZN,IBPA,IBPB,IBHN
     1,IN,NVD)
      DIMENSION W(1),HO(1),HC(1),P(1),WP(1),PLA(1),PLB(1),S(NVD,1)
     1,C(NVD,1),B(4)
      DIMENSION IBZN(1),IBPA(1),IBPB(1),IBHN(1),IN(1)
      COMMON/INT/NIJ,NEQ,ICR,ICR1,IB1,LH1,ID,JD,IM,JM,NQSD,NBH,NVAR
     1,NVX2,NVX3,NBQP,IPO,KOUNT,INDT,IBHZ
      COMMON/TNME/IIN,IOUT
      COMMON/REG/BMX,AP,CSA,AMP,RP,BP,YSQ
C**CHECK FOR NONZERO MARQUARDT PARAMETER
      IF(AMP.LT.-.5) GO TO 105
C**INITIALIZE
      DO 20 J=1,NVAR
      DO 10 I=1,NVAR
   10 C(I,J)=0.
   20 B(J)=0.
      YSQ=0.
C**FORM COEFFICIENT MATRIX AND RIGHT-HAND SIDE VECTOR
      K=0
      DO 70 N=1,NIJ
      IF(IN(N).LT.1.OR.W(N).LE.0.) GO TO 70
      TEMP=HO(N)-HC(N)
      K=K+1
      DO 60 J=1,NVAR
      TMP=W(N)*S(J,K)
      DO 50 I=J,NVAR
   50 C(I,J)=S(I,K)*TMP+C(I,J)
   60 B(J)=TMP*TEMP+B(J)
      YSQ=YSQ+TEMP*W(N)*TEMP
   70 CONTINUE
      IF(IBHZ.LE.0) GO TO 74
      DO 73 N=1,NBH
      J=IBHN(N)
      IF(W(J).LE.0.) GO TO 73
      M=IBZN(N+NQSD)
      K=IBPA(M)
      L=IBPB(M)
      TMPC=HO(J)-HC(J)
      YSQ=YSQ+TMPC*W(J)*TMPC
      IF(K.LE.NBQP) GO TO 72
      TMPA=PLA(N)*W(J)
      C(K,K)=C(K,K)+TMPA*PLA(N)
      B(K)=B(K)+TMPA*TMPC
      IF(L.LE.NBQP) GO TO 73
      TMP=TMPA*PLB(N)
      C(K,L)=C(K,L)+TMP
      C(L,K)=C(L,K)+TMP
   72 TMPB=PLB(N)*W(J)
      C(L,L)=C(L,L)+TMPB*PLB(N)
      B(L)=B(L)+TMPB*TMPC
   73 CONTINUE
```

```
   74 IF (NVAR.EQ.1) GO TO 190
      DO 80 I=1,NVAR
      TEMP=C(I,I)+WP(I)
      IF(TEMP.GT.1.E-10) GO TO 78
      WRITE(IOUT,260) I
      INDT=1
      GO TO 80
   78 C(I,I)=TEMP**.5
   80 CONTINUE
      IF(INDT.GT.0) RETURN
      NM1=NVAR-1
      DO 100 J=1,NM1
      TEMP=C(J,J)
      JP1=J+1
      DO 90 I=JP1,NVAR
      C(I,J)=C(I,J)/(C(I,I)*TEMP)
   90 C(J,I)=C(I,J)
      B(J)=(B(J)+WP(J)*(P(J)-1.))/TEMP+RP*TEMP*(BP*P(J)-1.)
      B(J+NVX2)=B(J)
      B(J+NVX3)=TEMP
  100 C(J,J)=1.+RP+AMP
      TEMP=C(NVAR,NVAR)
      B(NVAR)=(B(NVAR)+WP(NVAR)*(P(NVAR)-1.))/TEMP
     1+RP*TEMP*(BP*P(NVAR)-1.)
      B(NVX3)=B(NVAR)
      B(NVAR+NVX3)=TEMP
      C(NVAR,NVAR)=1.+RP+AMP
      IF(IPO.NE.1) GO TO 105
      WRITE(IOUT,250)
      CALL PRTOT(C,NVAR,NVD,0)
      WRITE(IOUT,230) (B(I),I=1,NVAR)
C**SOLVE FOR B USING LDU FACTORIZATION:
C**DECOMPOSITION AND FORWARD SUBSTITUTION
  105 DET=1.
      DO 140 K=1,NM1
      PIV=C(K,K)
      DET=DET*PIV
      IF(ABS(PIV).GT.1.E-10) GO TO 110
      WRITE(IOUT,210)
      INDT=1
      RETURN
  110 PIV=1./PIV
      KP1=K+1
      DO 130 J=KP1,NVAR
      TMP=C(J,K)*PIV
      DO 120 I=J,NVAR
  120 C(I,J)=C(I,J)-TMP*C(I,K)
  130 B(J)=B(J)-TMP*B(K)
      C(K,K)=PIV
  140 CONTINUE
      DET=DET*C(NVAR,NVAR)
      IF(ABS(C(NVAR,NVAR)).GT.1.E-10) GO TO 150
      WRITE(IOUT,210)
      INDT=1
      RETURN
```

```
  150 IF(AMP.LT.-.5) RETURN
C**BACK SUBSTITUTION
      B(NVX2)=B(NVAR)/C(NVAR,NVAR)
      B(NVAR)=B(NVX2)/B(NVAR+NVX3)
      I=NVAR
  160 I=I-1
      IF (I.LE.0) GO TO 175
      IP1=I+1
      SUM=0.
      DO 170 J=IP1,NVAR
  170 SUM=SUM+C(J,I)*B(J+NVAR)
      B(I+NVAR)=(B(I)-SUM)*C(I,I)
      B(I)=B(I+NVAR)/B(I+NVX3)
      GO TO 160
C**CHECK SOLUTION AND ADD MARQUARDT PARAMETER IF NEEDED
  175 TMPA=0.
      TMPB=0.
      TMPC=0.
      DO 176 I=1,NVAR
      TMPA=TMPA+B(I+NVAR)*B(I+NVAR)
      TMPB=TMPB+B(I+NVX2)*B(I+NVX2)
  176 TMPC=TMPC+B(I+NVAR)*B(I+NVX2)
      IF(TMPC.GT.CSA*SQRT(TMPA*TMPB)) GO TO 200
      AMP=1.5*AMP+.001
      IF(AMP.GT.1.) GO TO 200
      DO 180 I=1,NVAR
      B(I)=B(I+NVX2)
      C(I,I)=1.+RP+AMP
      DO 178 J=I,NVAR
  178 C(J,I)=C(I,J)
  180 CONTINUE
      GO TO 105
C**SOLUTION WHEN NVAR=1
  190 TEMP=C(1,1)+WP(1)
      IF(TEMP.GT.1.E-10) GO TO 195
      I=1
      WRITE(IOUT,260) I
      INDT=1
      RETURN
  195 B(4)=TEMP**.5
      B(3)=(B(1)+WP(1)*(P(1)-1.))/B(4)+RP*B(4)*(BP*P(1)-1.)
      C(1,1)=1.+RP
      DET=C(1,1)
      B(2)=B(3)/DET
      B(1)=B(2)/B(4)
C**PRINT AND ADJUST REGRESSION COEFFICIENTS
  200 WRITE(IOUT,220) KOUNT,YSQ,DET,AMP
      WRITE(IOUT,230) (B(J),J=1,NVAR)
      TEMP=0.
      DO 201 J=1,NVAR
      TMP=ABS(B(J))
      IF(TMP.GT.TEMP) TEMP=TMP
  201 CONTINUE
      AP=1.
      IF(TEMP.GT.BMX) AP=BMX/TEMP
```

```
      DO 202 J=1,NVAR
  202 B(J)=AP*B(J)
      RETURN
C
  210 FORMAT (43HOLEAST SQUARES COEFFICIENT MATRIX SINGULAR;/35H SOLUTIO
     1N FOR PARAMETERS NOT UNIQUE)
  220 FORMAT (1HO,14HITERATION NO. ,I3/1H ,6HYSQ = ,G11.5,2X
     1,9HDET(C) = ,G11.5,2X,6HAMP = ,G11.5
     2/1H ,31HSCALED REGRESSION COEFFICIENTS:)
  230 FORMAT ((1H ,8(G11.5,2X)))
  250 FORMAT(49HO SCALED LEAST SQUARES MATRIX AND GRADIENT VECTOR)
  260 FORMAT (29HOSENSITIVITIES FOR PARAMETER ,I4,17H EFFECTIVELY ZERO)
      END
      SUBROUTINE PRTOT(C,NO,NOD,IT)
C**IF IT=0, PRINT SYMMETRIC MATRIX DIVIDED VERTICALLY INTO TEN-COLUMN BLOCKS
C**IF IT=1, PRINT VECTOR IN THREE COLUMNS
      DIMENSION C(1)
      COMMON/TNME/IIN,IOUT
      IF(IT.EQ.1) GO TO 25
      DO 20 L=1,NO,10
      J10=L+9
      IF(J10.GT.NO) J10=NO
      WRITE(IOUT,30) (J,J=L,J10)
      WRITE(IOUT,50)
      K=-NOD
      DO 10 I=1,NO
      K=K+NOD
   10 WRITE(IOUT,40) I,(C(J+K),J=L,J10)
      WRITE(IOUT,60)
   20 CONTINUE
      RETURN
   25 NR=NO/3
      IF((3*NR).NE.NO) NR=NR+1
      DO 26 K=1,NR
   26 WRITE(IOUT,80) (L,C(L),L=K,NO,NR)
      RETURN
C
   30 FORMAT (1HO,8X,I3,9(9X,I3))
   40 FORMAT (1H ,I3,10(1X,G11.5))
   50 FORMAT (1H )
   60 FORMAT (1HO)
   80 FORMAT(1H ,3X,3(I3,7X,G11.5,3X))
      END
```

References Cited

Cooley, R.L., 1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1--Theory: Water Resources Research, v. 18, no. 4, p. 965-976.

_____1983, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1--Applications: Water Resources Research, v. 19, no. 3, p. 662-676.

Additional Reading

Cooley, R.L., 1977, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 1--Theory and numerical properties: Water Resources Research, v. 13, no. 2 p. 318-324.

_____1979, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 2--Application of statistical analysis: Water Resources Research, v. 15, no. 3, p. 603-617.

_____1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1--Theory: Water Resources Research, v. 18, no. 4, p. 965-976.

_____1983, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1--Applications: Water Resources Research, v. 19, no. 3, p. 662-676.

5. ELEMENTARY ANALYSIS AND USE OF THE REGRESSION MODEL

5.1. ASSUMED FORMS OF MODEL EQUATIONS

As a purely algebraic process, regression contains no assumptions other than those already mentioned. However, to statistically analyze results of, and predictions to be made by, the method, additional assumptions must be made. Based on these assumptions an effective methodology has been developed (see, for example, Draper and Smith, 1981) to analyze and use a linear regression model.

The statistical methods also may be applied to a nonlinear model, provided the model is close enough to being linear. Fortunately, whether or not the model is close enough can usually be determined. All statistics and procedures are, accordingly, derived for a linear, or effectively linear, model. To make the equations applicable for both a linear and nonlinear model, the basic types of models assumed are the incremental linear model and the nonlinear model as linearized using the Taylor series expansion.

The model assumed, then, is of the form

$$\underline{Y}_{\beta} - \underline{Y}_0 \cong \underline{X}(\underline{\beta} - \underline{b}_0) \tag{5.1-1}$$

where strict equality applies for a linear model and, for a nonlinear model, $\underline{X}$ is assumed to be evaluated at $\underline{b}_0$. By definition

$$\underline{Y}_{\beta} = \underline{f}(\underline{\xi}, \underline{\beta}) \tag{5.1-2}$$

$$\underline{Y}_0 = \underline{f}(\underline{\xi}, \underline{b}_0) \tag{5.1-3}$$

Based on (5.1-1), the true regression model is

$$\underline{Y} - \underline{Y}_0 \cong \underline{X}(\underline{\beta} - \underline{b}_0) + \underline{\varepsilon} \tag{5.1-4}$$

where strict equality only holds for a linear model because $\underline{\varepsilon}$ is the true vector of disturbances. The estimated regression model derived from (5.1-4) is

$$\underline{Y} - \underline{Y}_0 \cong \underline{X}(\underline{b} - \underline{b}_0) + \underline{e} \tag{5.1-5}$$

where, as for (5.1-4), strict equality only holds for a linear model because $\underline{e}$ is assumed to be the true vector of residuals defined by $\underline{e} = \underline{Y} - \underline{f}(\underline{\xi}, \underline{b})$.

By minimizing $S(\underline{b}) = \underline{e}^T \underline{\omega}\, \underline{e}$ with respect to $\underline{b}$ using the appropriate standard procedure, exact best-fit estimates $\hat{\underline{b}}$ of $\underline{\beta}$ and $\hat{\underline{e}} = \underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}})$ of $\underline{\varepsilon}$ are obtained. For a linear model $\hat{\underline{b}}$ is obtained exactly by using (5.1-5) as the estimated regression model. For a nonlinear model, use of the linearized model leads to an approximate relationship to find $\hat{\underline{b}}$. Thus, by minimizing $S(\underline{b})$ using (5.1-5) as the estimated regression model, normal equations

$$\underline{X}^T \underline{\omega}\, \underline{X}(\hat{\underline{b}} - \underline{b}_0) \cong \underline{X}^T \underline{\omega}(\underline{Y} - \underline{Y}_0) \tag{5.1-6}$$

that are approximate for a nonlinear model are derived. The regression model obtained by replacing general estimates $\underline{b}$ and $\underline{e}$ in (5.1-5) with best-fit estimates $\hat{\underline{b}}$ and $\hat{\underline{e}}$ is

$$\underline{Y} - \underline{Y}_0 \cong \underline{X}(\hat{\underline{b}} - \underline{b}_0) + \hat{\underline{e}} \tag{5.1-7}$$

By utilizing the definition of $\hat{\underline{e}}$ ($\hat{\underline{e}} = \underline{Y} - \hat{\underline{Y}}$) in (5.1-7) a predictive model

$$\hat{\underline{Y}} - \underline{Y}_0 \cong \underline{X}(\hat{\underline{b}} - \underline{b}_0) \tag{5.1-8}$$

is obtained.

A final point is the establishment of the general condition for a minimum in $S(\underline{b})$. If $\underline{b}_0$ is set equal to $\hat{\underline{b}}$ in (5.1-6), then $\underline{Y}_0 = \hat{\underline{Y}}$, and

$$\underline{X}^T\underline{\omega}(\underline{Y} - \hat{\underline{Y}}) = \underline{0} \tag{5.1-9}$$

Because any approximation inherent in (5.1-6) is removed as $\underline{b}_0 \to \hat{\underline{b}}$, (5.1-9) is exact for both linear and nonlinear models. The left-hand side of (5.1-9) is the negative of the gradient of $S(\underline{b})$. Thus, (5.1-9) states that the gradient of $S(\underline{b})$ is zero at a minimum point of $S(\underline{b})$.

From here on, for simplicity the approximate equality sign in regression models and normal equations (e.g., (5.1-1), (5.1-4) through (5.1-8)) is replaced by an equal sign. However, it must be remembered that all relationships derived by using the linearized model are approximate for a nonlinear model.

5.2. ASSUMPTIONS OF REGRESSION MODELING

Some of the assumptions listed below have already been mentioned; they are discussed more completely here.

1) A true model exists:

$$Y = f(\xi_1, \xi_2, \ldots, \xi_k; \underline{\beta}) + \varepsilon \tag{5.2-1}$$

The model response, Y, consists of two parts, a deterministic part, f, and an additive stochastic part, ε.

2) The disturbances, ε, have the following properties

$$E(\underline{\varepsilon}) = \underline{0} \tag{5.2-2}$$

$$Var(\underline{\varepsilon}) = \underline{V}\sigma^2 \tag{5.2-3}$$

where the structure, $\underline{V}$, of the variance-covariance matrix $\underline{V}\sigma^2$ is assumed to be symmetric positive definite and known. Alternate forms for (5.2-2) and (5.2-3) are obtained by premultiplying (5.2-2), and pre- and postmultiplying (5.2-3), by $\underline{V}^{-\frac{1}{2}}$ to obtain

$$E(\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}) = \underline{0} \tag{5.2-4}$$

$$Var(\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}) = \underline{I}\sigma^2 \tag{5.2-5}$$

The assumptions given by (5.2-2) through (5.2-5) indicate that $\underline{\varepsilon}$ is considered to be a vector of random variables with zero mean and variance-covariance matrix $\underline{V}\sigma^2$. Furthermore, weighted disturbances $\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}$ have constant variance $\underline{I}\sigma^2$ and are uncorrelated. To require the expected value of $\underline{\varepsilon}$ to be zero is to require that (5.2-1) be the true (or unbiased) model and to require in addition that $\underline{\varepsilon}$ be unbiased. Although imperfections in most physical theories prevent the former assumption from holding strictly true, a model should be constructed so that the absolute value of any $E(\varepsilon_j)$ is as small as possible. From the practical point of view, it is required that the bias not be significant. Criteria for this are developed later on.

The full form of $\underline{V}$ is usually very difficult to obtain from the type of data usually available. However, if $\underline{V}$ is assumed to be diagonal so that there is no correlation among the ε_j, then $\underline{V}$ can often be found by using graphical methods of analyzing residuals, to be discussed later on.

3) The matrices $\underline{\omega}$ and $\underline{V}^{-1}$ are equivalent. That is

$$\underline{\omega} \equiv \underline{V}^{-1} \tag{5.2-6}$$

Equation (5.2-6) is not essential to compute an unbiased estimate of $\underline{\beta}$. This fact may be demonstrated as follows. Solve equation (5.1-6) for $\hat{\underline{b}} - \underline{b}_0$ and take the expected value of it to obtain:

$$\begin{aligned} E(\hat{\underline{b}} - \underline{b}_0) &= (\underline{X}^T\underline{\omega}\,\underline{X})^{-1}\underline{X}^T\underline{\omega}\,E(\underline{Y} - \underline{Y}_0) \\ &= (\underline{X}^T\underline{\omega}\,\underline{X})^{-1}\,\underline{X}^T\underline{\omega}\,\underline{X}(\underline{\beta} - \underline{b}_0) \\ &= \underline{\beta} - \underline{b}_0 \end{aligned} \tag{5.2-7}$$

where the fact that $E(\underline{\varepsilon}) = \underline{0}$ was used. From (5.2-7) it is seen that

$$E(\hat{\underline{b}}) = \underline{\beta} \tag{5.2-8}$$

For a linear model at least, the Gauss-Markov theorem (Beck and Arnold, 1977, p. 232-234) establishes that the variance of $\hat{b}_j$, $Var(\hat{b}_j)$, is a minimum if (5.2-6) is true. Furthermore, to compute $Var(\hat{\underline{b}})$ correctly, whether or not (5.2-6) is true, $\underline{V}$ must be known. Hence, assumption of another form for $\underline{\omega}$ (such as $\underline{I}$, for example) would not avoid the problem of having to know $\underline{V}$ to analyze the model.

In case $\underline{V}$ is diagonal, then

$$\underline{V} = \begin{bmatrix} 1/\omega_1 & & & \\ & 1/\omega_2 & & \\ & & \ddots & \\ & & & 1/\omega_n \end{bmatrix} \tag{5.2-9}$$

4) The disturbances are normally distributed:

$$\underline{\varepsilon} \sim N(\underline{0}, \underline{V}\sigma^2) \tag{5.2-10}$$

or

$$\underline{V}^{-\frac{1}{2}}\underline{\varepsilon} \sim N(\underline{0}, \underline{I}\sigma^2) \tag{5.2-11}$$

Assumption of (5.2-10) or (5.2-11) is only necessary if investigations using the F distribution are to be performed.

That $\underline{\varepsilon}$ (or $\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}$) be normally distributed implies that the elements of $\underline{V}^{-\frac{1}{2}}\ \underline{\varepsilon}$ are neither systematic nor constant but are equally likely to be positive or negative. In addition, small errors are more frequent than large ones. Many types of models are subject to a number of sources of error, any one of which may or may not be normally distributed. However, in the case where a resultant error is the sum of a number of components, Central Limit Theorem implies that $\underline{\varepsilon}$ (or $\underline{V}^{-\frac{1}{2}}\underline{\varepsilon}$) could be normally distributed even if its component vectors were not.

Because $\underline{\varepsilon}$ and $\underline{\beta}$ are unknown, the assumptions (1 through 4) discussed cannot be checked directly. However, they may often be checked indirectly, which is a subject of model analysis.

5.3. RELATIONSHIPS BETWEEN RESIDUALS AND DISTURBANCES

Many of the investigations involving the regression model are based either directly or indirectly on relationships between residuals $\hat{\underline{e}}$ and disturbances $\underline{\varepsilon}$. Residuals may be written in terms of disturbances by employing (5.1-6) and (5.1-7). First equation (5.1-7) is written in the form

$$\hat{\underline{e}} = \underline{Y} - \underline{Y}_0 - \underline{X}(\hat{\underline{b}} - \underline{b}_0) \tag{5.3-1}$$

Then (5.1-6) is solved for $\hat{\underline{b}} - \underline{b}_0$ and substituted into (5.3-1) to obtain

$$\begin{aligned}\hat{\underline{e}} &= \underline{Y} - \underline{Y}_0 - \underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}(\underline{Y} - \underline{Y}_0) \\ &= (\underline{I} - \underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega})(\underline{Y} - \underline{Y}_0)\end{aligned} \tag{5.3-2}$$

If $\underline{b}_0$ is set equal to $\underline{\beta}$, then $\underline{Y} - \underline{Y}_0 = \underline{Y} - \underline{Y}_\beta = \underline{\varepsilon}$ and (5.3-2) gives

$$\hat{\underline{e}} = (\underline{I} - \underline{X}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega})\underline{\varepsilon} \tag{5.3-3}$$

It is frequently more convenient to work with weighted residuals, $\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}}$, and weighted disturbances, $\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon}$. In this case (5.3-3) becomes

$$\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}} = (\underline{I} - \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}})\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon} \tag{5.3-4}$$

An interesting and useful property of the matrix appearing in either (5.3-3) or (5.3-4) is displayed, for (5.3-4) for example, as follows:

$$[\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}]^T[\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}]$$

$$= \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}\underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}}$$

$$= \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}} \tag{5.3-5}$$

In other words, the transpose of the matrix times the matrix yields the original matrix. This result is true for both (5.3-3) and (5.3-4). For (5.3-4) it is also true that the matrix is symmetric (as can be seen in the derivation of (5.3-5)), so that the matrix times itself yields the original matrix. This type of matrix is known as a symmetric idempotent matrix.

Another useful property can be derived based on the idempotency discussed above. For convenience let

$$\underline{R} = \underline{\omega}^{\frac{1}{2}}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T\underline{\omega}^{\frac{1}{2}} \tag{5.3-6}$$

Then, because $\underline{R}$ is idempotent,

$$(\underline{I} - \underline{R})(\underline{I} - \underline{R}) = \underline{I} - \underline{R} - \underline{R} + \underline{R}$$

$$= \underline{I} - \underline{R} \qquad (5.3\text{-}7)$$

Hence, $\underline{I} - \underline{R}$ is also idempotent.

5.4. SOME STATISTICAL MEASURES

The first step in model analysis should always be to examine some statistical measures that indicate (1) goodness of fit of the model to the data and (2) model conditioning as it affects reliability of the computed parameters. With the background given above, useful statistical measures can be derived.

5.4.1. The Error Variance, s^2

For a linear model this measure is an unbiased estimate of σ^2. For a nonlinear model, s^2 is biased. A sketch of the derivation follows: The sum of squares, $S(\hat{\underline{b}})$, is defined as

$$S(\hat{\underline{b}}) = \hat{\underline{e}}^T \underline{\omega}\, \hat{\underline{e}}$$

$$= (\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}})^T(\underline{\omega}^{\frac{1}{2}}\hat{\underline{e}}) \qquad (5.4\text{-}1)$$

Using (5.3-4), (5.3-7), and the fact that tr(scaler) = scaler, (5.4-1) becomes

$$S(\hat{\underline{b}}) = (\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})^T(\underline{I} - \underline{R})^T(\underline{I} - \underline{R})(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})$$

$$= tr[(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})^T(\underline{I} - \underline{R})(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})]$$

$$= tr[(\underline{I} - \underline{R})(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})^T] \qquad (5.4\text{-}2)$$

The expected value of (5.4-2) is

$$E(S(\hat{\underline{b}})) = tr\ \{(\underline{I} - \underline{R})E[(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})^T]\}$$

$$= tr[(\underline{I} - \underline{R})Var(\underline{\omega}^{\frac{1}{2}}\underline{\varepsilon})]$$

$$= tr[(\underline{I} - \underline{R})\sigma^2]$$

$$= (n - p)\sigma^2 \qquad (5.4\text{-}3)$$

from which

$$\sigma^2 = \frac{E(S(\hat{\underline{b}}))}{n - p} \qquad (5.4\text{-}4)$$

The fact that $tr(\underline{R}) = p$ can be demonstrated by rearranging the matrices within $\underline{R}$. The estimate, s^2, of σ^2 is

$$s^2 = \frac{S(\hat{\underline{b}}))}{n - p} = \frac{(\underline{Y} - \hat{\underline{Y}})^T\underline{\omega}(\underline{Y} - \hat{\underline{Y}})}{n - p} \qquad (5.4\text{-}5)$$

A useful modification of (5.4-5) that is exact for a linear model and almost exact for a nonlinear one is obtained by choosing $\underline{b}_0$ in (5.1-7) to be very near $\hat{\underline{b}}$. Then

$$s^2 = \frac{(\underline{Y} - \underline{Y}_0 - \underline{X}(\hat{\underline{b}} - \underline{b}_0))^T \underline{\omega}(\underline{Y} - \underline{Y}_0 - \underline{X}(\hat{\underline{b}} - \underline{b}_0))}{n - p}$$

$$= \frac{(\underline{Y} - \underline{Y}_0)^T \underline{\omega}(\underline{Y} - \underline{Y}_0) - 2(\hat{\underline{b}} - \underline{b}_0)^T \underline{X}^T \underline{\omega}\ (\underline{Y} - \underline{Y}_0) + (\hat{\underline{b}} - \underline{b}_0)^T \underline{X}^T \underline{\omega}\ \underline{X}(\hat{\underline{b}} - \underline{b}_0)}{n - p}$$

$$= \frac{(\underline{Y} - \underline{Y}_0)^T \underline{\omega}(\underline{Y} - \underline{Y}_0) - (\hat{\underline{b}} - \underline{b}_0)^T \underline{X}^T \underline{\omega}(\underline{Y} - \underline{Y}_0)}{n - p} \qquad (5.4\text{-}6)$$

where use was made of (5.1-6). For a linear model $\underline{b}_0$ may be chosen to be $\underline{0}$. For a nonlinear model a bias exists in (5.4-5) or (5.4-6) that results from the fact that development is based on assumption of a linear model.

Even when biased, s^2 gives a useful measure of overall goodness of fit of the model. The standard deviation or scatter is given by s. In general, $s/\Delta Y_s$ should be small, where ΔY_s is the difference between maximum and minimum values of Y_s.

5.4.2. The Correlation, R_y, Between $\underline{\omega}^{\frac{1}{2}}\ \underline{Y}$ and $\underline{\omega}^{\frac{1}{2}}\ \hat{\underline{Y}}$

This measure is defined as

$$R_y = \frac{\underline{d}_{\underline{Y}-\underline{Y}}^T \underline{d}_{\underline{Y}-\hat{\underline{Y}}}}{[(\underline{d}_{\underline{Y}-\underline{Y}}^T \underline{d}_{\underline{Y}-\underline{Y}})(\underline{d}_{\hat{\underline{Y}}-\hat{\underline{Y}}}^T \underline{d}_{\hat{\underline{Y}}-\hat{\underline{Y}}})]^{\frac{1}{2}}} \qquad (5.4\text{-}7)$$

where

$$\underline{d}_y = \underline{w}^{\frac{1}{2}} \underline{Y} - m_y \underline{1} \tag{5.4-8}$$

$$\underline{d}_{\hat{y}} = \underline{w}^{\frac{1}{2}} \hat{\underline{Y}} - m_{\hat{y}} \underline{1} \tag{5.4-9}$$

$$m_y = \sum_{i=1}^{n} (\underline{w}^{\frac{1}{2}} \underline{Y})_i / n \tag{5.4-10}$$

$$m_{\hat{y}} = \sum_{i=1}^{n} (\underline{w}^{\frac{1}{2}} \hat{\underline{Y}})_i / n \tag{5.4-11}$$

$\underline{1}$ = vector of ones,

$(\underline{w}^{\frac{1}{2}}\underline{Y})_i$ is a component of $\underline{w}^{\frac{1}{2}}\underline{Y}$, and similarly for $(\underline{w}^{\frac{1}{2}}\hat{\underline{Y}})_i$. The correlation R_y is another measure of goodness of fit. Usually it should be greater than about 0.9 to indicate a good fit to the reliable data.

5.4.3. The Variance-Covariance Matrix for $\hat{\underline{b}}$

This measure may be derived directly from (5.1-6) and is

$$Var(\hat{\underline{b}}) = Var[(\underline{X}^T \underline{w}\ \underline{X})^{-1} \underline{X}^T \underline{w} (\underline{Y} - \underline{Y}_0) + \underline{b}_0]$$

$$= (\underline{X}^T \underline{w}\ \underline{X})^{-1} \underline{X}^T \underline{w} Var(\underline{Y}) \underline{w}\ \underline{X} (\underline{X}^T \underline{w}\ \underline{X})^{-1}$$

$$= (\underline{X}^T \underline{w}\ \underline{X})^{-1} \sigma^2 \tag{5.4-12}$$

where assumptions (5.2-3) and (5.2-6) and the fact that $Var(\underline{Y}) = Var(\underline{Y} - \underline{Y}_{\beta}) = \underline{\omega}^{-1}\sigma^2$ were employed. The estimate of $Var(\hat{\underline{b}})$ is

$$\widehat{Var}(\hat{\underline{b}}) = (\underline{X}^T \underline{\omega}\ \underline{X})^{-1} s^2 \qquad (5.4\text{-}13)$$

The standard error of the estimate for the ith parameter is given by the square root of the ith diagonal component of $(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} s^2$. This estimate is a measure of the range over which the respective parameter may be varied to produce a similar solution for the dependent variable as that obtained using $\hat{\underline{b}}$.

5.4.4. The Correlation, r_{ij}, Between any Two Parameters $\hat{b}_i$ and $\hat{b}_j$

By definition

$$r_{ij} = \frac{Cov(\hat{b}_i, \hat{b}_j)}{[Var(\hat{b}_i)\ Var(\hat{b}_j)]^{\frac{1}{2}}} \qquad (5.4\text{-}14)$$

where the variance and covariance terms are components of $(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} s^2$. This measure gives an estimate of the degree of linear dependence of one parameter on another throughout the course of repeated experiments if such experiments were to be carried out. As discussed earlier, it is an indication of the degree of linear dependency in the sensitivity matrix.

Problem 5.4-1

This problem is concerned with preliminary analysis of the linear regression solution of problem 3.2-1. The measures (except $s/\Delta Y_s$), to be computed in a, b, and c below, also are calculated by the computer model of problem 4.2-1. Check your computations against the computer-generated results.

a) Using (5.4-6), compute s^2. Compute $s/\Delta Y_s$. Would you say that the fit is very good?

b) Using (5.4-13), compute $\widehat{Var}(\underline{\hat{b}})$. Are the parameters determined very precisely?

c) Using (5.4-14), determine $\underline{r}$, the correlation matrix. Are there any evident problems with conditioning?

5.5. ANALYSIS OF RESIDUALS

Examination of the statistics discussed in the previous section should give a preliminary indication of general model conditioning and model fit to the data. However, a thorough analysis of residuals is necessary in order to examine the validity of the assumptions given in section 5.2. Interest is focused primarily on indications of nonrandomness of the residuals and on indications that the residuals are not distributed normally. The analysis should include both sample and prior information partitions of the residuals so that any incompatibility between the two partitions can be detected as differences between the two partitions. Although the techniques given in the present section are usually adequate to detect any incompatibility, a formal test given in section 6.3 also can be applied if desired.

Analytical methods used here are graphical. Draper and Smith (1981, p. 141-192) give a number of methods for examining residuals, and they emphasize that graphical procedures involving visual analysis are the most valuable tools because violations of assumptions serious enough to require corrective action generally are apparent on the various plots. However, to use the procedures effectively it is necessary to determine the properties that the residuals should be expected to exhibit under ideal conditions.

5.5.1. Distribution of Residuals

Investigation of the distribution of residuals $\hat{e}_j$ (or $\underline{\omega}_j^{\frac{1}{2}} \underline{\hat{e}}$, where $\underline{\omega}_j^{\frac{1}{2}}$ is a row of $\underline{\omega}^{\frac{1}{2}}$) made in order to infer the distribution of disturbances ε_j (or

$\underline{w}_j^{\frac{1}{2}} \underline{\varepsilon}$) is difficult because, even if assumption (5.2-5) holds so that the elements $\underline{w}_j^{\frac{1}{2}} \underline{\varepsilon}$ are uncorrelated and have equal variance, elements of $\underline{w}_j^{\frac{1}{2}} \hat{\underline{e}}$ are correlated and have unequal variance. To show this for the linearized model, (5.3-4) and (5.3-6) can be combined to give

$$\underline{w}^{\frac{1}{2}}\hat{\underline{e}} = (\underline{I} - \underline{R})\underline{w}^{\frac{1}{2}}\underline{\varepsilon} \tag{5.5-1}$$

from which

$$\begin{aligned} E(\underline{w}^{\frac{1}{2}}\hat{\underline{e}}) &= (\underline{I} - \underline{R})\underline{w}^{\frac{1}{2}}E(\underline{\varepsilon}) \\ &= \underline{0} \end{aligned} \tag{5.5-2}$$

and

$$\begin{aligned} Var(\underline{w}^{\frac{1}{2}}\hat{\underline{e}}) &= Var[(\underline{I} - \underline{R})\underline{w}^{\frac{1}{2}}\underline{\varepsilon}] \\ &= (\underline{I} - \underline{R})Var(\underline{w}^{\frac{1}{2}}\underline{\varepsilon})(\underline{I} - \underline{R})^T \\ &= (\underline{I} - \underline{R})(\underline{I} - \underline{R})^T\sigma^2 \\ &= (\underline{I} - \underline{R})\sigma^2 \end{aligned} \tag{5.5-3}$$

Hence, if assumptions (5.2-4) through (5.2-6), and (5.2-11) hold,

$$\underline{\omega}^{\frac{1}{2}} \hat{\underline{e}} = \hat{\underline{f}} \sim N(\underline{0}, (\underline{I} - \underline{R})\sigma^2) \tag{5.5-4}$$

where for convenience, the definition is made that $\underline{\omega}^{\frac{1}{2}} \hat{\underline{e}} = \hat{\underline{f}}$.

It can be shown that $\underline{I} - \underline{R} \to \underline{I}$ as $n - p \to \infty$. Whenever $n - p$ becomes small, correlation and unequal variance become significant. Most tests for distribution of residuals assume equal (or, a common) variance and no correlation because all residuals are assumed to have come from the same univariate distribution. Therefore, correlation and unequal variance of the $\hat{e}_j$ (or $\hat{f}_j$ if $\underline{\omega} \neq \underline{I}$) are serious problems with regard to testing for normality when the number of parameters is not small compared to the number of observations.

Another difficulty concerns the determination of whether or not the model fits the data. If the model fits the data and correlation of the values of $\hat{e}_j$ (or $\hat{f}_j$) is not significant, then these residuals should appear to be nearly random. However, if correlation is significant, then the correlation will be reflected in the residual values. Patterns could develop in some of the plots (to be discussed) and these patterns could be mistakenly attributed to lack of model fit.

5.5.2. Graphical Procedures

The first step in using graphical procedures is to develop a control group. Several sets of simulated residuals distributed as in (5.5-4) form the control group. These sets are then compared graphically with the true residuals $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) to help decide whether the distribution of $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) differs to a visually detectable extent from a normal distribution and whether

correlation could cause an apparently nonrandom (or non-normal) pattern of residuals to develop in the residual plots.

A set of simulated residuals may be generated by generating a set of uncorrelated random normal deviates $\underline{d}$ so that $E(\underline{d}) = \underline{0}$ and $Var(\underline{d}) = \underline{I}s^2$, then forming linear combinations of these deviates that have the covariance given by (5.5-3). The method of generating the simulated residuals from the uncorrelated random deviates can be derived as follows. Assume, as a working hypothesis, that

$$\underline{g} = \underline{\Omega}\ \underline{d} \tag{5.5-5}$$

where $\underline{g}$ is the set of simulated residuals and $\underline{\Omega}$ is a symmetric and nonstochastic matrix to be determined. Vector $\underline{g}$ must be generated so that $E(\underline{g}) = \underline{0}$ and $Var(\underline{g}) = (\underline{I} - \underline{R})s^2$ (see (5.5-2) and (5.5-3)). From (5.5-5)

$$\begin{aligned} E(\underline{g}) &= \underline{\Omega}\ E(\underline{d}) \\ &= \underline{0} \end{aligned} \tag{5.5-6}$$

as required. By definition

$$\begin{aligned} Var(\underline{g}) &= \underline{\Omega}\ Var(\underline{d})\underline{\Omega}^T \\ &= \underline{\Omega}^2 s^2 \end{aligned} \tag{5.5-7}$$

where the definition of $Var(\underline{d})$ and the symmetry of $\underline{\Omega}$ were used. Hence, $\underline{\Omega}$ must be defined so that

$$\underline{\Omega}^2 = \underline{I} - \underline{R} \tag{5.5-8}$$

However, because $\underline{I} - \underline{R}$ is idempotent, $(\underline{I} - \underline{R}) = (\underline{I} - \underline{R})^2$, and (5.5-8) may be simplified to become

$$\underline{\Omega} = \underline{I} - \underline{R} \tag{5.5-9}$$

so that (5.5-5) assumes as its final form

$$\underline{g} = (\underline{I} - \underline{R})\underline{d} \tag{5.5-10}$$

To generate a set of simulated residuals, $\underline{g}$, it is a simple matter to generate a set of uncorrelated random normal deviates, $\underline{d}$, then use (5.5-10). This procedure is followed for the number of sets (usually at least three) desired to form the control group.

Normal probability plots. These are graphs of cumulative frequency, F, versus values of the elements of vectors $\hat{\underline{e}}$, $\hat{\underline{f}}$, $\underline{d}$, or $\underline{g}$. Cumulative frequency corresponding to the ith element of one of the vectors ($\hat{\underline{e}}$, for example) is computed from the formula

$$F_i = m_i/(n + 1), \; i = 1, 2, \ldots, n \tag{5.5-11}$$

where m_i is the number of values of $\hat{e}$ (for example) smaller than or equal to $\hat{e}_i$, and n is the number of observations. Use of n + 1 in the denominator adjusts for the fact that F_n cannot be equal to 1 because the tail of the normal distribution extends to infinity. An example of a normal probability plot for $\underline{\hat{e}}$ is illustrated in figure 5.5-1.

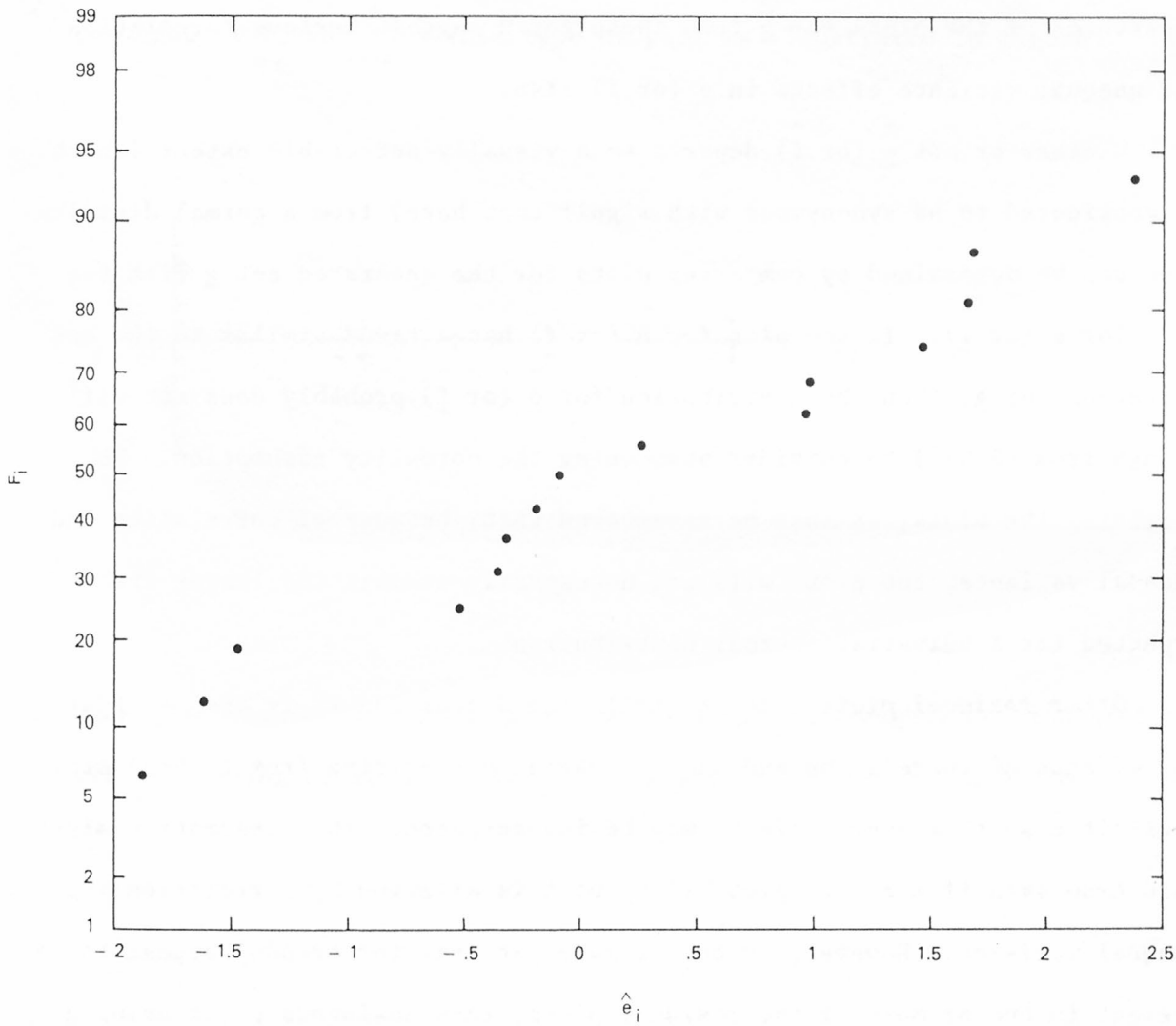


fig. 5.5-1

To determine the effects of correlation and unequal variance, normal probability plots for the sets $\underline{d}$ and the sets $\underline{g}$ may be compared. If the plots for $\underline{d}$ and $\underline{g}$ are very similar, then a normal probability plot of $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) would not be expected to be affected to a great extent by correlation and unequal variance. Other types of plots involving $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) probably would not be affected much by correlation and unequal variance either. Significant departures of the plots for $\underline{g}$ from those for $\underline{d}$ suggest serious correlation and unequal variance effects in $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) also.

Whether or not $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) departs to a visually detectable extent (which is considered to be synonymous with significant here) from a normal distribution can be determined by comparing plots for the generated set $\underline{g}$ with the plot for $\hat{\underline{e}}$ (or $\hat{\underline{f}}$). If the plot for $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) has a trend similar to the set of curves for $\underline{g}$, then the distribution for $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) probably does not differ enough from (5.5-4) to consider abandoning the normality assumption. When examining the plots, it must be remembered that, because of correlation and unequal variance, the plots will not necessarily exhibit the linear trend expected for a univariate normal distribution.

Other residual plots. In the following discussion it is assumed that the effects of correlation and unequal variance resulting from (5.5-3) are negligible so that other effects may be investigated. This assumption might hold true even if a normal probability plot is affected by correlation and unequal variance. However, if one or more patterns (or trends) appear to be present in one or more of the residual plots, then analogous plots using $\underline{g}$ instead of $\hat{\underline{e}}$ (or $\hat{\underline{f}}$) also should be prepared and examined. If the suspicious patterns also are present in the plots using $\underline{g}$, then the patterns probably result from correlation and unequal variance inherent in $(\underline{I} - \underline{R})s^2$ and not from model error.

Three types of plot are often useful: (1) Plot of $\hat{f}_j$ vs. $\hat{Y}_j$; (2) plots of $\hat{f}_j$ vs. independent variables (ξ_i); (3) plot of $\hat{f}_j$ vs. Cartesian coordinates of point j. For generality, $\hat{f}_j = \underline{\omega}_j^{\frac{1}{2}}\hat{\underline{e}}$ (where $\underline{\omega}_j^{\frac{1}{2}}$ is a row of $\underline{\omega}^{\frac{1}{2}}$) is used instead of $\hat{e}_j$ in the following discussions. If $\underline{\omega} = \underline{I}$ was employed in the regression, then $\hat{f}_j = \hat{e}_j$ is plotted. Additional discussion of the first two types of plots may be found in Draper and Smith (1981, p. 147-148).

1) Plot of $\hat{f}_j$ vs. $\hat{Y}_j$. This type of plot is illustrated in figure 5.5-2.

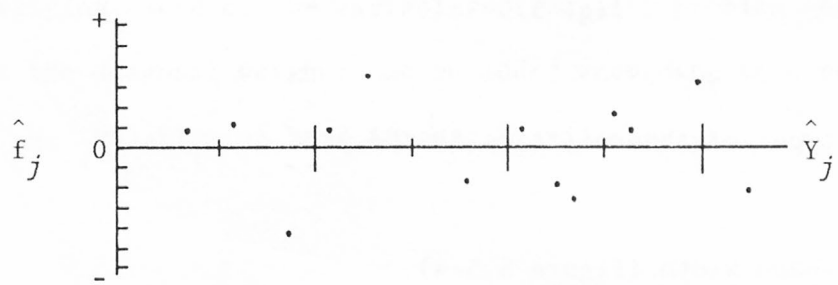


fig. 5.5-2

Under the given assumptions, the plot should display a roughly horizontal band of residuals having no apparent trend, as sketched in figure 5.5-3. A

standard runs test could be used to test for randomness of signs of $\hat{f}_j$ along the $\hat{Y}_j$ axis.

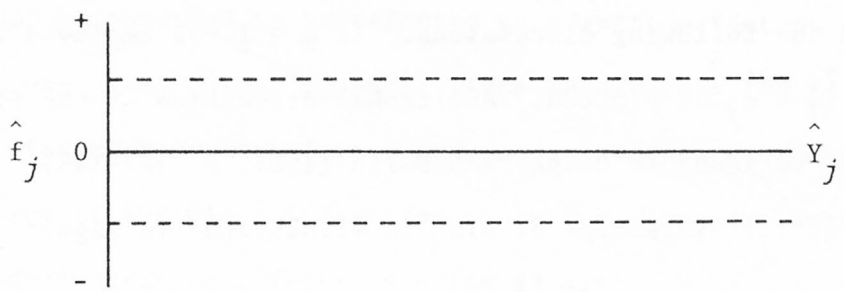


fig. 5.5-3

Three principal types of abnormalities in the plot of $\hat{f}_j$ vs. $\hat{Y}_j$ are often apparent:

a) Unequal band width (figure 5.5-4).

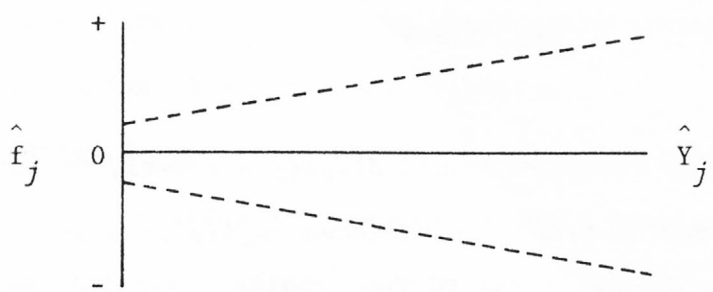


fig. 5.5-4

This type of trend (or one opposite to it) generally indicates that the variance of $\hat{f}_j$ is not constant. In figure 5.5-4 the observations would appear to be less reliable as $\hat{Y}_j$ increases.

If the abnormal plot resulted from a least squares analysis where $\underline{\omega} = \underline{I}$ had been assumed, then a diagonal form of $\underline{\omega} \neq \underline{I}$ might be indicated. In the illustration, ω_{jj} should decrease with $\hat{Y}_j$. However, if some form of $\underline{\omega} \neq \underline{I}$ had originally been assumed, then $\hat{f}_j$ should be plotted in addition to $\hat{e}_j$. An abnormal plot, of the form shown in figure 5.5-4, involving $\hat{f}_j$ would suggest the $\hat{f}_j$ are not of equal reliability and that $\underline{\omega}$ is not correct. Hence, $\underline{\omega}$ should be modified, and the regression performed again. If a full form of $\underline{\omega}$ is required by the true model, this form cannot generally be discovered by this type of analysis. Before a weighted least squares is performed, the underlying cause of the variable reliability problem should be investigated so that the diagonal weights can be added according to a rational criterion.

b) Sloping band (figure 5.5-5).

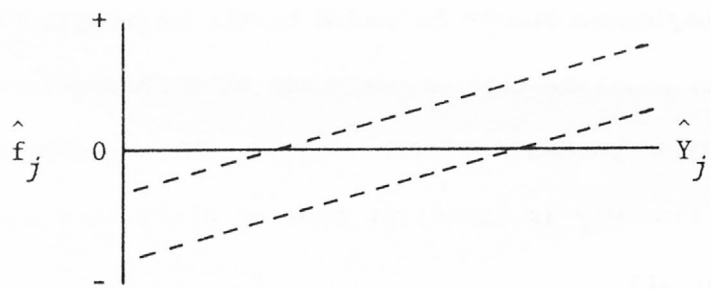


fig. 5.5-5

This type of problem often indicates model error. Typically, it is caused by omitting an intercept from the model. For a linear model having an intercept, $\Sigma \hat{f}_j \hat{Y}_j = 0$ always. Hence, in this case an overall slope such as depicted in figure 5.5-5 cannot result from correlation and unequal variance in $(\underline{I} - \underline{R})s^2$. Also, if the model is linear (or effectively so) and $E(\underline{\varepsilon}) = \underline{0}$, then $Cov(\hat{\underline{f}}, \hat{\underline{Y}}) = \underline{0}$ even if there is no intercept. Thus, it might be expected that a

plot such as shown in figure 5.5-5 would usually not result from correlation and unequal variance if the model is correct and if the observed data are adequate.

c) Curved or irregularly shaped band (figure 5.5-6).

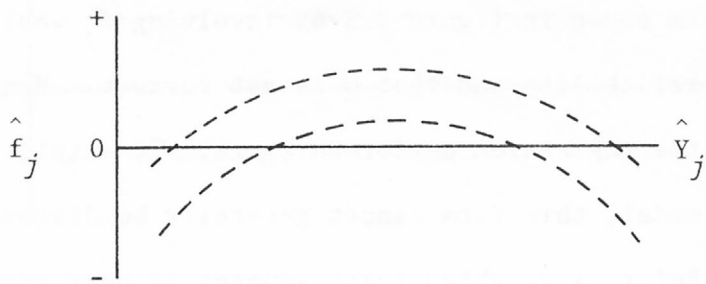


fig. 5.5-6

This is another indication of model error. The model is inadequate because it does not account for all of the sources of variability in the observed data. More, or different, terms should be added to the regression equation. For the plot shown, it is possible that a quadratic term should be added.

2) Plots of $\hat{f}_j$ vs. independent variables (ξ_i). These plots are interpreted in much the same way as the first type of plot; they simply provide a different viewpoint.

3) For trend surface types of regression: plot of $\hat{f}_j$ vs. distance (for a one-dimensional system) or $\hat{f}_j$ vs. Cartesian coordinates of point j (for two-dimensional systems). Three-dimensional systems can be reduced to two dimensions by using cross sections. Systematic highs and lows in the residual pattern usually suggest that the model does not fit the data well. In general, the easier it is to contour the residual map, the more nonrandom the residuals are likely to be. Positive and negative values of large and small magnitude should be mixed together.

Problem 5.5-1

Conduct a graphical analysis of residuals $\hat{\underline{e}}$ resulting from the Theis equation problem (problem 3.3-1). Compute five sets of random normal deviates $\underline{d}$ and simulated residuals $\underline{g}$ using the residuals analysis program (appendix 5.8.1). Plot $\underline{d}$ and $\underline{g}$ on normal probability paper. Does the plot of $\hat{\underline{e}}$ differ significantly from the plots of $\underline{g}$? What can you conclude about the distribution of $\hat{\underline{e}}$? Does the Theis model appear to be adequate?

The code given in appendix 5.8.1 is designed to read COV(I,J) (which is $(\underline{Z}^T\underline{\omega}\underline{Z})^{-1}s^2$), W(I) (which is $\underline{\omega}_s$) and S(I,J) (which is $\underline{Z}_s$) in unformatted form from a file labeled ITB. Normally, these data would be read to file ITB from the numerical nonlinear regression program of appendix 4.3.4. However, the Theis data are not obtained from the numerical nonlinear regression program. The easiest way to read the Theis data is to modify the residuals analysis program by replacing READ(ITB) in statements reading cards B, C, and D with READ(5,2). Data for the program should then be coded as explained in appendix 5.8.1, except that the data for cards B, C, and D will now be coded in format 8F10.0.

Problem 5.5-2

Use the output from the two-dimensional flow model to analyze the model results of problem 4.2-2.

a) What is the value of s^2? R_y? Compute $s/\Delta Y_s$.

b) Is there any evidence of ill-conditioning in the results? If there is, what is the problem? Which parameters are well-determined and which are not? Why?

c) Conduct a graphical analysis of residuals. Develop four sets of random normal deviates $\underline{d}$ and simulated residuals $\underline{g}$, computed by

using the residuals analysis program (appendix 5.8.1). Plot $\underline{d}$ and $\underline{g}$ on normal probability paper using (5.5-11). Are correlation effects evident? Plot $\hat{\underline{f}}$ (why $\hat{\underline{f}}$ instead of $\hat{\underline{e}}$?) on normal probability paper. Does the plot differ significantly from the plots of $\underline{g}$? Plot $\hat{f}_j$ versus $\hat{Y}_j$. Is there an abnormal pattern? Plot $\hat{e}_j$ versus Cartesian coordinate of point j (omitting the prior information). Again, is there a pattern to the residuals? What do you conclude about the adequacy of the model?

5.6. INVESTIGATION OF ALTERNATIVE PARAMETER SETS

5.6.1. Generalized W Statistic

Suppose we want to test the null hypothesis that some subset $\underline{\beta}_2$ of parameter set $\underline{\beta}$ cannot be distinguished from some corresponding given subset $\underline{\tilde{\beta}}_2$. That is, test

$$H_0: \underline{\beta}_2 = \underline{\tilde{\beta}}_2 \text{ versus } H_1: \underline{\beta}_2 \neq \underline{\tilde{\beta}}_2.$$

The linearized model assumed is

$$\underline{Y} = \underline{Y}_0 + \underline{X}(\underline{\beta} - \underline{b}_0) + \underline{\varepsilon}$$

$$= \underline{Y}_0 + \underline{X}_1(\underline{\beta}_1 - \underline{b}_{01}) + \underline{X}_2(\underline{\beta}_2 - \underline{b}_{02}) + \underline{\varepsilon} \qquad (5.6\text{-}1)$$

where $\underline{X}$, $\underline{\beta}$, and $\underline{b}_0$ are conformably partitioned as follows:

$$\underline{X} = [\underline{X}_1 \ \underline{X}_2] \qquad (5.6\text{-}2)$$

$$\left.\begin{aligned} \underline{\beta} &= \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{bmatrix} \\ \underline{b}_0 &= \begin{bmatrix} \underline{b}_{01} \\ \underline{b}_{02} \end{bmatrix} \end{aligned}\right\} \quad (5.6\text{-}3)$$

Based on (5.6-1), we may state a predictive model of the form

$$\tilde{\underline{Y}} = \underline{Y}_0 + \underline{X}_1(\tilde{\underline{b}}_1 - \underline{b}_{01}) + \underline{X}_2(\tilde{\underline{\beta}}_2 - \underline{b}_{02}) \quad (5.6\text{-}4)$$

where $\tilde{\underline{b}}_1$ is an estimate of $\underline{\beta}_1$ and, under the null hypothesis, $\underline{\beta}_2$ is assumed to be given by $\tilde{\underline{\beta}}_2$.

The W statistic, which is stated explicitly later on, is based on a comparison of the restricted sum of squares $(\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega}(\underline{Y} - \tilde{\underline{Y}})$ and the unrestricted sum of squares $(\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})$. The unrestricted sum of squares is obtained from the standard least squares analysis. The restricted sum of squares is obtained by minimizing

$$S_0(\tilde{\underline{b}}_1) = (\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega}(\underline{Y} - \tilde{\underline{Y}}) \quad (5.6\text{-}5)$$

with respect to $\tilde{\underline{b}}_1$ while holding $\tilde{\underline{\beta}}_2$ constant. For a linear model this results in the normal equations:

$$\tilde{\underline{b}}_1 - \underline{b}_{01} = (\underline{X}_1^T \underline{\omega}\, \underline{X}_1)^{-1} \underline{X}_1^T \underline{\omega}(\underline{Y} - \underline{Y}_0 - \underline{X}_2(\tilde{\underline{\beta}}_2 - \underline{b}_{02})) \quad (5.6\text{-}6)$$

where $\underline{b}_{01}$ and $\underline{b}_{02}$ can be set to zero if desired. If the model is nonlinear, then (5.6-5) is minimized with respect to $\tilde{\underline{b}}_1$ by using the standard procedures, except that $\tilde{\underline{\beta}}_2$ is held constant.

If many repeat experiments using linear normal equations (5.6-6) were performed, and if assumptions (5.2-5), (5.2-6), and (5.2-10) held true, then it would be found that

$$W = \frac{[(\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega}(\underline{Y} - \tilde{\underline{Y}}) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})]/q}{(\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})/(n - p)}$$

$$= \frac{[(\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega}(\underline{Y} - \tilde{\underline{Y}}) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})]/q}{s^2} \sim F(q, n - p) \qquad (5.6\text{-}7)$$

where q is the order of $\tilde{\underline{\beta}}_2$, which is the number of restrictions in H_0, and (5.4-5) was used. The symbol W stands for a random variable. For each experiment, a value w of the random variable W could be computed. According to (5.6-7), by repeating the experiment many times, the probability of the ratio W having a value of w or a smaller value would be found to be given by the cumulative density function F.

Because (5.6-7) is proportional to the difference between the restricted and unrestricted sums of squares divided by the unrestricted sum of squares, one might suspect the null hypothesis H_0: $\underline{\beta}_2 = \tilde{\underline{\beta}}_2$ to be true if w is small. However, if w is large, then one might suspect that H_0 is incorrect. The rejection region for the hypothesis test is determined by the probability statement $P(W > F_\alpha(q, n - p)) = \alpha$, where α is the significance level of the test and $F_\alpha(q, n - p)$ is the upper $100\alpha\%$ point of the F distribution with q and n - p degrees of freedom. If the ratio w is greater than $F_\alpha(q, n - p)$, as found in any table of critical values for the F distribution, then the null hypothesis is rejected because values of w greater than $F_\alpha(q, n - p)$ form the rejection region.

An alternate form for the numerator of (5.6-7) may be derived by manipulating the linearized models. The result, after extensive algebra, is

$$(\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega} (\underline{Y} - \tilde{\underline{Y}}) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega} (\underline{Y} - \hat{\underline{Y}})$$

$$= (\tilde{\underline{\beta}}_2 - \hat{\underline{b}}_2)^T [\underline{H} (\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \underline{H}^T]^{-1} (\tilde{\underline{\beta}}_2 - \hat{\underline{b}}_2) \qquad (5.6\text{-}8)$$

where

$$\underline{H} = [\underline{0}\ \underline{I}_q]_{(q \times p)} \qquad (5.6\text{-}9)$$

$\underline{I}_q$ = identity matrix of order q, and (5.6-10)

$\hat{\underline{b}}_2$ = the partition corresponding to $\underline{\beta}_2$ found from the standard (unrestricted) least squares analysis.

Thus, the alternate form for (5.6-7) is

$$W = \frac{(\tilde{\underline{\beta}}_2 - \hat{\underline{b}}_2)^T [\underline{H} (\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \underline{H}^T]^{-1} (\tilde{\underline{\beta}}_2 - \hat{\underline{b}}_2)/q}{s^2} \qquad (5.6\text{-}11)$$

If the model is linear, either (5.6-7) or (5.6-11) may be used to compute w. Both equations require an unrestricted regression to obtain $\hat{\underline{b}}$ and $\hat{\underline{Y}}$, but (5.6-7) requires, in addition, a restricted regression using (5.6-6) to obtain $\tilde{\underline{b}}_1$ and $\tilde{\underline{Y}}$. Hence, for a linear model, (5.6-11) is often more efficient to use

than (5.6-7) for practical computations. If (5.6-1) is a linearized equation system, derived from a nonlinear model, then neither (5.6-7) nor (5.6-8) nor (5.6-11) is exact. However, if (5.6-1) behaves in fashion that is close enough to being linear, then (5.6-7), (5.6-8), and (5.6-11) are good approximations. When working with a nonlinear model, w should be computed using both (5.6-7) and (5.6-11). If the conclusions reached by using the two different expressions for w are different, then the model may be too nonlinear for investigations using the W statistic. Further investigation of model nonlinearity may be performed by employing Beale's measure, which is discussed in section 6.2.

In summary, the procedure for testing

$$H_0: \underline{\beta}_2 = \underline{\tilde{\beta}}_2 \text{ versus } H_1: \underline{\beta}_2 \neq \underline{\tilde{\beta}}_2$$

in the model $\underline{Y} = \underline{Y}_0 + \underline{X}(\underline{\beta} - \underline{b}_0) + \underline{\varepsilon}$ is:

1) Carry out a regression without any restriction to find $\underline{\hat{b}}$ for the full model.
2) When using (5.6-7), fix $\underline{\tilde{\beta}}_2$ and use (5.6-6) fo find $\underline{\tilde{b}}$, for the restricted model. When using (5.6-11), skip this step.
3) Form the ratio w using (5.6-7) or (5.6-11).
4) Compare w with the appropriate value of $F_\alpha(q, n - p)$.

5.6.2. Joint Confidence Region for $\underline{\beta}_2$

Equations (5.6-7) or (5.6-11) also may be used to obtain a joint confidence region on $\underline{\beta}_2$. The confidence region interpretation is based on fixing a probability of occurrence, $P(W < F_\alpha(q, n - p)) = 1 - \alpha$ then finding

those vectors $\tilde{\underline{\beta}}_2$ that would yield the specified F or a smaller value. The joint confidence region may be written as

$$(\underline{Y} - \tilde{\underline{Y}})^T \underline{\omega}(\underline{Y} - \tilde{\underline{Y}}) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}}) \leq s^2 q F_\alpha(q, n - p) \tag{5.6-12}$$

or

$$(\underline{\beta}_2 - \hat{\underline{b}}_2)^T [\underline{H}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \underline{H}^T]^{-1} (\underline{\beta}_2 - \hat{\underline{b}}_2) \leq s^2 q F_\alpha(q, n - p) \tag{5.6-13}$$

Equation (5.6-13) plots as a family of q dimensional ellipsoids (which for q = 1 is a line segment bounded by two values for $\underline{\beta}_2$) in parameter space, and these ellipsoids are centered on $\hat{\underline{b}}$. All ellipsoids corresponding to probability levels smaller than 1-α lie within the outermost ellipsoid, which is defined by strict equality in (5.6-13). Hence, the specified probability is the probability that $\underline{\beta}_2$ lies within the ellipsoid. An equivalent statement is that, if many experiments were conducted, then (1 - α)100% of the ellipsoids would contain the true parameter set $\underline{\beta}_2$. Hence, the outermost ellipsoid may be considered to be a joint confidence region on $\underline{\beta}_2$.

Points (in parameter space) on the edge of the confidence region corresponding to the maximum and minimum (or extreme) values that some parameter $\tilde{\beta}_{2i}$ may attain and remain in the confidence region are given by

$$\tilde{\underline{b}} = \hat{\underline{b}} \pm \frac{\sqrt{q F_\alpha(q, n - p)}}{s_{bi}} \underline{v}_{bi} \tag{5.6-14}$$

where $\tilde{\underline{b}}^T = [\tilde{\underline{b}}_1^T \ \tilde{\underline{\beta}}_2^T]$ and $\underline{v}_{bi}$ is the ith column of $\underline{v}_b = (\underline{X}^T \underline{\omega}\ \underline{X})^{-1} s^2$. Note that s_{bi} is the square root of v_{bii}.

Equation i in (5.6-14) gives the extreme values of $\tilde{\beta}_{2i}$. The parameter vector $\underset{\sim}{\tilde{b}}$ computed using (5.6-14) is the same vector that would result if (1) $\tilde{\beta}_{2i}$ were computed using equation i in (5.6-14), (2) then the remaining values in $\underset{\sim}{\tilde{\beta}}_2$ were computed to satisfy (5.6-13) (with strict equality applying to give points on the edge of the confidence region), and (3) finally, partition $\underset{\sim}{\tilde{b}}_1$ were computed using (5.6-6).

Parameter sets computed using (5.6-14) will exactly satisfy the relationship $s^2 qF_\alpha(q, n - p) = (\underline{Y} - \underline{\tilde{Y}})^T \underline{\omega}(\underline{Y} - \underline{\tilde{Y}}) - (\underline{Y} - \underline{\hat{Y}})^T \underline{\omega}(\underline{Y} - \underline{\hat{Y}}) = (\underset{\sim}{\tilde{\beta}}_2 - \underset{\sim}{\hat{b}}_2)^T [\underline{H}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \underline{H}^T]^{-1} (\underset{\sim}{\tilde{\beta}}_2 - \underset{\sim}{\hat{b}}_2)$ if the model is linear. However, as discussed in section 5.6.1, this relationship is not exact for a nonlinear model. Therefore, if the regression model is nonlinear, the parameter sets computed using (5.6-14) should be substituted into the nonlinear model and $(\underline{Y} - \underline{\tilde{Y}})^T \underline{\omega}(\underline{Y} - \underline{\tilde{Y}}) - (\underline{Y} - \underline{\hat{Y}})^T \underline{\omega}(\underline{Y} - \underline{\hat{Y}})$ should be computed. If this value is different enough from $s^2 qF_\alpha(q, n - p)$ to change any conclusions, then the model is too nonlinear to use to generate linearized confidence regions. Beale's measure discussed in section 6.2 also can be used to gage nonlinearity.

Two end-member cases involving the W statistic are often considered separately. In one case $q = p$ so that $\underline{\beta}_2 \equiv \underline{\beta}$. All parameters are thus included in any test of H_0, and the confidence region is on all parameters simultaneously. This confidence region is called a joint confidence region

on all parameters. An example for two parameters is diagrammed in figure 5.6-1.

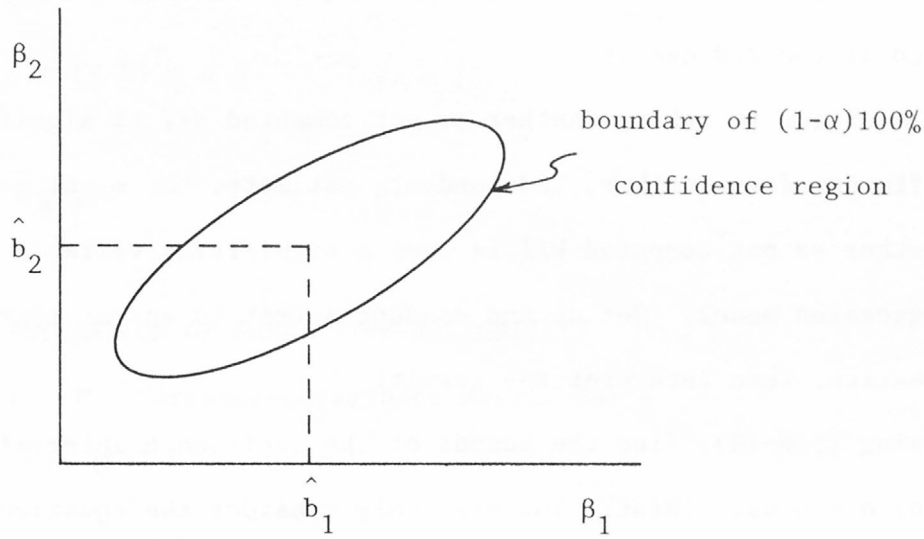


fig. 5.6-1

In the other case q = 1 and $\underline{\beta}_2 \equiv \beta_p$. Thus, only one parameter is considered in any test of H_0, and the confidence region is on only one parameter. The confidence region for this case is termed an individual confidence interval on parameter β_p.

Problem 5.6-1

a) A method for estimating recharge in Nevada is known as the Maxey-Eakin method. Using this method, W in the vicinity of Lake Ohpupu (problem 3.2-1) was estimated to be 0.0003 ft/day. Also, by using specific capacity estimates from well-log analysis, T was found to be 10 ft^2/day. Using these estimates, test the null hypothesis that there is no significant difference at α = 0.05 between W/T as

estimated above and the regression estimate of W/T. (Hint: Use equation (5.6-11).) Based on the result of the hypothesis test, would you consider using the prior estimate of W/T as prior information in the regression model? What other information would you need if you did use it?

b) In addition to asking whether or not computed W/T is significantly different from another, independent, estimate, one might ask whether or not computed W/T is even a significant variable in the regression model. Set up and conduct a test to answer this question, then interpret the result.

c) Using (5.6-14), find the bounds of the confidence interval on W/T for $\alpha = 0.05$. (Hint: You need only consider the equation corresponding to parameter W/T in the system implied by (5.6-14).)

Problem 5.6-2

To the program for computing T and S using the Theis equation (problem 3.3-1), add the code necessary to compute s^2 using (5.4-5) and $\widehat{Var}(\hat{\underline{b}})$ using (5.4-13). You will have to unscale the entries of $(\underline{Z}^T\underline{Z})^{-1}$ to obtain $\widehat{Var}(\hat{\underline{b}})$. Examination of the calculation procedures for obtaining $\underline{Z}^T\underline{Z}$ and $(\underline{Z}^T\underline{Z})^{-1}$ will indicate how the unscaling should be accomplished. Use your new program and (5.6-14) to find the sets of parameters corresponding to extreme values of T and to extreme values of S assuming q = 2.

Problem 5.6-3

Using (5.6-14) and the results of problem 4.2-2, find the sets of parameters corresponding to extreme values of T_3 and to extreme values of

q_{B1} with q = 2. You will have to unscale the entries of the variance-covariance matrix to do this. Equation (3.3-15) may be utilized for this purpose:

$$\underline{Z}^T \underline{\omega}\ \underline{Z} = (\underline{X}\ \underline{B})^T \underline{\omega}\ \underline{X}\ \underline{B}$$

Hence, $\underline{X}^T \underline{\omega}\ \underline{X}$ = ?

5.7. INVESTIGATION OF PREDICTIVE RELIABILITY

5.7.1. The Variance-Covariance Matrix for $\hat{\underline{Y}}$

Equation (5.1-8) is used to obtain

$$Var(\hat{\underline{Y}}) = Var[\underline{X}(\hat{\underline{b}} - \underline{b}_0) + \underline{Y}_0]$$

$$= \underline{X}\ Var(\hat{\underline{b}})\underline{X}^T \qquad (5.7\text{-}1)$$

An analogous measure for weighted values of $\hat{Y}$ is

$$Var(\underline{\omega}^{\frac{1}{2}}\ \hat{\underline{Y}}) = Var[\underline{\omega}^{\frac{1}{2}}\ \underline{X}(\hat{\underline{b}} - \underline{b}_0) + \underline{\omega}^{\frac{1}{2}}\ \underline{Y}_0]$$

$$= \underline{\omega}^{\frac{1}{2}}\ \underline{X}\ Var(\hat{\underline{b}})\underline{X}^T \underline{\omega}^{\frac{1}{2}} \qquad (5.7\text{-}2)$$

By using (5.3-6) and (5.4-12), (5.7-2) can be written

$$Var(\underline{\omega}^{\frac{1}{2}}\ \hat{\underline{Y}}) = \underline{R}\ \sigma^2 \qquad (5.7\text{-}3)$$

Estimates corresponding to (5.7-1), (5.7-2), and (5.7-3) are

$$\widehat{Var}(\hat{\underline{Y}}) = \underline{X}\ \widehat{Var}(\hat{\underline{b}})\underline{X}^T \tag{5.7-4}$$

$$\widehat{Var}(\underline{\omega}^{\frac{1}{2}}\ \hat{\underline{Y}}) = \underline{\omega}^{\frac{1}{2}}\ \underline{X}\ \widehat{Var}(\hat{\underline{b}})\underline{X}^T\underline{\omega}^{\frac{1}{2}} \tag{5.7-5}$$

$$= \underline{R}s^2 \tag{5.7-6}$$

The standard error of $\hat{Y}_i$ is given by the square root of the ith diagonal entry of $Var(\hat{\underline{Y}})$. This estimate gives a measure of potential variability in $\hat{Y}_i$ resulting from $Var(\hat{\underline{b}})$.

It is important to note that (5.7-1), (5.7-2), (5.7-4), and (5.7-5) are valid for prediction vectors $\hat{\underline{Y}}$ having entries that are not necessarily at observation points. This fact may be understood by observing that prediction equation (5.1-8) used to derive $Var(\hat{\underline{Y}})$ or $Var(\underline{\omega}^{\frac{1}{2}}\ \hat{\underline{Y}})$ (or their estimates) is valid for any set of points, not just observation points. However, entries in $\underline{X}$ and, if (5.7-2) or (5.7-5) is used, $\underline{\omega}$ must be available for all points in $\hat{\underline{Y}}$. Matrix $(\underline{X}\ \underline{\omega}\ \underline{X})^{-1}s^2$ used for $\widehat{Var}(\hat{\underline{b}})$ is, of course, the standard one based on entries in $\underline{X}$ only at the observation points.

5.7.2. Confidence Interval on $Y_{\beta j}$

If all parameters are allowed to vary over the confidence region (5.6-12) or (5.6-13) with q = p, then the maximum and minimum values produced for $\tilde{Y}_j$ form a corresponding confidence interval on $\tilde{Y}_j$. In this case, because q = p, $\underline{\beta}_2 \equiv \underline{\beta}$ and $\tilde{Y}_j = Y_{\beta j}$. The resulting confidence interval on $Y_{\beta j}$ is

$$Y_{\beta j} = \hat{Y}_j \pm \sqrt{pF_\alpha(p,\ n - p)}\ s_{yj} \tag{5.7-7}$$

where

$$s_{yj} = s\sqrt{\underline{X}_j(\underline{X}^T\underline{\omega}\,\underline{X})^{-1}\underline{X}_j^T} \qquad (5.7\text{-}8)$$

and $\underline{X}_j$ is a row of a sensitivity matrix corresponding to the point j. The point j need not correspond to an observation point.

Equation (5.7-8) gives a simultaneous confidence interval. That is, the probability is $1-\alpha$ that $Y_{\beta j}$ lies within the interval indicated by (5.7-7) and that Y_β for all other possible points lies within similar intervals simultaneously. If intervals on a number of consecutive points are computed and plotted, the result is called a confidence band.

5.7.3. Tolerance Interval on Predicted Observation, Y_j^{pred}

Equation (5.7-7) gives a confidence interval on a computed value of Y or, in other words, the mean of Y. In some instances, a corresponding interval on a predicted observation is desired, and this is termed a tolerance interval. Tolerance intervals on k predicted values of Y simultaneously can be readily computed if $\underline{\omega} = \underline{V}^{-1}$ is diagonal, and are given by (Lieberman, 1961):

$$Y_j^{pred} = \hat{Y}_j \pm \sqrt{kF_\alpha(k,\ n-p)}\ \sqrt{s^2/\omega_j + s_{yj}^2}\ ,\quad j = 1,\ 2,\ \ldots,\ k \qquad (5.7\text{-}9)$$

As in (5.7-7), point j need not (and, in general, probably would not) correspond to an observation point. However, ω_j for the prediction point has to be known.

The term $s^2/\omega_j + s^2_{yj}$ is the total variance in predicted observation Y_j^{pred}. This form results because $Y_j^{pred} = (Y_j^{pred} - \hat{Y}_j) + \hat{Y}_j$, where $Y_j^{pred} - \hat{Y}_j$ is statistically independent of $\hat{Y}_j$, so that, as an estimate, $\widehat{Var}(Y_j^{pred}) = \widehat{Var}(Y_j^{pred} - \hat{Y}_j) + \widehat{Var}(\hat{Y}_j)$ or $\widehat{Var}(Y_j^{pred}) = s^2/\omega_j + s^2_{yj}$.

Equation (5.7-9) does not give tolerance intervals on all $\underline{Y}_{pred}$ simultaneously. Furthermore, as k increases, the tolerance interval increases without bound. This results because the normal distribution, which the errors in Y are assumed to follow, has infinite tails. Hence, even though the probability of an error that is large in magnitude is small, as the number of values of Y considered simultaneously increases, the probability of an arbitrarily large error in at least one of them increases also. Usually the tolerance interval is computed using k = 1.

Problem 5.7-1

a) Write out explicitly the form for general entry (ij) of $\underline{X}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}\underline{X}^T$, which is used in equation (5.7-1). For this exercise let $(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} \equiv \underline{A}$ and write the result in terms of $\underline{A}$ to simplify the expression. Select and compute a diagonal entry of this matrix at an observation point used for the linear regression solution of problem 3.2-1. Note that you can replace $\underline{X}$ by $\underline{S}$ in $\underline{X}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}\underline{X}^T$, and that this leaves the result unaltered. Can you show this?

b) Using (5.7-4) and the diagonal entry computed in part a, determine $\widehat{Var}(Y_j)$, where j is the selected diagonal entry.

c) Using (5.7-7) and the results of b, find the confidence interval on $Y_{\beta j}$.

5.8. APPENDICES

5.8.1. Documentation of Program to Compute Vectors $\underline{d}$ and $\underline{g}$ of Section 5.5.2.

This program computes vector $\underline{d}$ of random normal deviates, vector $\underline{g}$ of correlated normal deviates as defined by (5.5-10), and other useful information related to the distributions of $\hat{\underline{e}}$ and $\hat{\underline{f}}$. Sample and direct prior information are assumed to be given in the form (3.4-12). The scaled sensitivity matrix for the direct prior information is theoretically of the form $\underline{Z}_p = \hat{\underline{X}}_p \underline{B} = [\underline{I}_p \ \underline{0}]\underline{B}$, where $\underline{B}$ is the diagonal matrix composed of parameter estimates $\hat{b}_j$ (j = 1, 2, ..., p). However, matrix $[\underline{I}_p \ \underline{0}]$ may be rearranged to conform with any parameter ordering.

A random (0,1) number generator is employed as a function subroutine. This routine assumes an integer computer word length equal to at least 1077109141.

As coded, the contents of card sets B, C, and D are assumed to be stored in unformatted form on file ITB = 8. This is so that these sets do not have to be input manually. The listing appended contains the code and instructions for insertion into the program of appendix 4.3.4 so that the required data are stored in the proper form and order for use in the present program. The user will have to supply the job control language necessary to store the data and retrieve them for use.

Two variables, NVD and NTD defined near the beginning of the program, must be redefined each time the dimensions of the program are changed. NVD must be set equal to dimensions of COV and the first dimension of S, all three of which are at least NVAR, and NTD must be set equal to the dimensions of R, which are at least NTOT = NOBS + NPRIR.

Input Data.

Card Set A.

Problem size information; one card (format 5I5, F10.0).

Card columns	Variable	Definition
1-5	NVAR	Number of parameters, p.
6-10	NOBS	Number of sample observations, n_s.
11-15	NPRIR	Number of regression parameters having direct prior information, n_p.
16-20	NSETS	Number of sets of $\underline{d}$ and $\underline{g}$ vectors to be computed.
21-25	NRAN	Seed for random number generator: any odd number between 1 and 1048575.
26-35	VAR	Error variance, s^2.

Card Set B.

Scaled covariance matrix, $(\underline{Z}^T\underline{\omega Z})^{-1}s^2$ (unformatted; stored in file ITB).

Variable	Definition
COV(1,1)	Scaled covariance matrix, entered sequentially from the diagonal element through NVAR for each new regression parameter number. Each new diagonal element begins a new record.
COV(2,1)	
⋮	
COV(NVAR,1)	
COV(2,2)	
⋮	

COV(NVAR,2)

COV(3,3)

⋮

COV(NVAR,3)

⋮

COV(NVAR,NVAR)

Card Set C.

Weight matrix for sample information, $\underline{V}_s^{-1}$ (unformatted; stored in file ITB).

Variable	Definition
W(1)	Diagonal weight matrix for sample information, entered sequentially from 1 through NOBS.
W(2)	
⋮	
W(NOBS)	

Card Set D.

Scaled sensitivity matrix for sample information, $\underline{Z}_s$ (unformatted; stored in file ITB).

Variable	Definition
S(1,1)	Scaled sensitivity matrix for sample information, entered sequentially 1 through NVAR for each observation. Each new observation begins a new record, for a total of NOBS observations.
S(2,1)	
⋮	
S(NVAR,1)	
S(1,2)	
⋮	

S(NVAR,2)

⋮

S(NVAR,NOBS)

Card Set E.

Parameter numbers having prior information (format 16I5).

Card columns	Variable	Definition
1-5	IPR(1)	Array subscript numbers for regression parameters in $\hat{\underline{b}}$ having prior information. For use with the regression ground-water program, the array subscript numbers must be the order numbers in the parameter vector computed by that program.
6-10	IPR(2)	
⋮	⋮	
	IPR(NPRIR)	

Omit card set if NPRIR = 0.

Card Set F.

Standard deviation matrix for prior information, $\underline{U}^{\frac{1}{2}}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	F(1)	Diagonal standard deviation matrix for prior information, entered in the same order as IPR(I) from 1 through NPRIR.
11-20	F(2)	
⋮	⋮	
	F(NPRIR)	

Omit card set if NPRIR = 0.

Card Set G.

Scaled sensitivity matrix for prior information, $\underline{Z}_p$ (format 8F10.0).

Card columns	Variable	Definition
1-10	SP(1)	Diagonal matrix condensed from $\underline{Z}_p$; composed of elements of $\underline{B}$ on which there is prior information, and ordered the same as F(I) from 1 through NPRIR.
11-20	SP(2)	
⋮	⋮	
	SP(NPRIR)	

Omit card set if NPRIR = 0.

Output. Output is all clearly labeled. It is ordered as follows:

1. Card sets A through G.
2. Set number of vectors $\underline{d}$ and $\underline{g}$. Data for numbers 2 through 5 below are printed sequentially for each set.
3. Vector $\underline{d}$. This vector is ordered from smallest to largest entry, and each entry is paired with its theoretical frequency as computed by using (5.5-11).
4. Vector $\underline{g}$. Each entry is printed in its natural position corresponding to its position in a row or column of $\underline{R}$. Rows and columns of $\underline{R}$ are ordered by first sample observation numbers followed by prior information numbers, which are the subscripts of IPR(I).
5. Vector $\underline{g}$. This vector is ordered and paired with its theoretical frequency in the same way as $\underline{d}$ is.
6. Covariance matrix $(\underline{I} - \underline{R})s^2$. This matrix, which is an estimate of the one defined by (5.5-3), is for weighted residuals $\hat{\underline{f}}$ composed of both sample and prior information.
7. Correlation matrix for $\hat{\underline{f}}$. This matrix is derived from $(\underline{I} - \underline{R})s^2$.

Program Listing.

```
C********************* RESIDUALS ANALYSIS PROGRAM **********************
      DIMENSION S(20,50),COV(20,20),W(50),SP(20),SD(20),IPR(20),R(70,70)
     1,D(70),G(70),F(70)
      EQUIVALENCE (S(1,1),D(1)),(W(1),F(1),SD(1)),(COV(1,1),G(1))
C
C         FORMAT LIST
C
    1 FORMAT (5I5,F10.0)
    2 FORMAT (8F10.0)
    3 FORMAT (9H1NVAR  = ,I4/9H NOBS  = ,I4/9H NPRIR = ,I4
     1/9H NSETS = ,I4/9H NRAN  = ,I4/9H VAR   = ,G11.5)
    4 FORMAT (1HO,14X,42HRELIABILITY WEIGHTS FOR SAMPLE INFORMATION
     1/1H ,3X,3(3HNO.,11X,1HW,9X))
    5 FORMAT (16I5)
    6 FORMAT (1HO,14X,43HNO.S OF PARAMETERS HAVING PRIOR INFORMATION
     1/1H ,3X,3(3HNO.,8X,3HIPR,10X))
    7 FORMAT (1HO,14X,40HSTANDARD DEVIATIONS OF PRIOR INFORMATION
     1/1H ,3X,3(3HNO.,10X,2HSD,9X))
    8 FORMAT (25HOSCALED COVARIANCE MATRIX)
    9 FORMAT (44HOSCALED SENSITIVITIES FOR OPTIMUM PARAMETERS)
   10 FORMAT (1HO,27X,13HSCALE FACTORS/1H ,3X,3(3HNO.,10X,2HSP,9X))
   11 FORMAT (1HO,20X,31HORDERED, RANDOM NORMAL DEVIATES/1H ,3X,2(3HNO.
     1,8X,1HD,14X,1HF,10X))
   12 FORMAT (1HO,18X,35HORDERED, CORRELATED NORMAL DEVIATES/1H ,3X
     1,2(3HNO.,8X,1HG,14X,1HF,10X))
   13 FORMAT (32HOCOVARIANCE MATRIX FOR RESIDUALS)
   14 FORMAT (43HODATA GENERATED FROM RANDOM NUMBER SET NO. ,I3)
   15 FORMAT (33HOCORRELATION MATRIX FOR RESIDUALS)
   16 FORMAT (1HO,22X,26HCORRELATED NORMAL DEVIATES/1H ,3X,3(3HNO.,11X
     1,1HG,9X))
C
C         DEFINE ARRAY DIMENSIONS FOR PRTOT
C
      NVD=20
      NTD=70
C
C         READ AND PRINT INPUT DATA THEN CONVERT IT INTO FORMS NEEDED
C         FOR CALCULATIONS
C         NOTE: NRAN MUST BE ODD AND MUST LIE BETWEEN 1 AND 1048575
C
      READ(5,1) NVAR,NOBS,NPRIR,NSETS,NRAN,VAR                    CARD A
      WRITE(6,3) NVAR,NOBS,NPRIR,NSETS,NRAN,VAR
      ITB=8
      DO 25 J=1,NVAR
      READ(ITB) (COV(I,J),I=J,NVAR)                               CARD B
      DO 20 I=J,NVAR
   20 COV(J,I)=COV(I,J)
   25 CONTINUE
      WRITE(6,8)
      CALL PRTOT(COV,NVAR,NVAR,NVD)
      READ(ITB) (W(I),I=1,NOBS)                                   CARD C
      WRITE(6,4)
      CALL PRTOT(W,NOBS,1,0)
```

```
      DO 30 I=1,NOBS
   30 W(I)=W(I)**.5
      DO 35 J=1,NOBS
      READ(ITB) (S(I,J),I=1,NVAR)                                    CARD D
   35 CONTINUE
      WRITE(6,9)
      CALL PRTOT(S,NVAR,NOBS,NVD)
      DO 45 J=1,NOBS
      DO 40 I=1,NVAR
   40 S(I,J)=S(I,J)*W(J)
   45 CONTINUE
      SIGMA=VAR**.5
      IF(NPRIR.LT.1) GO TO 55
      READ(5,5) (IPR(I),I=1,NPRIR)                                   CARD E
      WRITE(6,6)
      CALL PRTOTC(IPR,NPRIR)
      READ(5,2) (SD(I),I=1,NPRIR)                                    CARD F
      WRITE(6,7)
      CALL PRTOT(SD,NPRIR,1,0)
      READ(5,2) (SP(I),I=1,NPRIR)                                    CARD G
      WRITE(6,10)
      CALL PRTOT(SP,NPRIR,1,0)
      DO 50 I=1,NPRIR
   50 SP(I)=(SIGMA/SD(I))*SP(I)
C
C         COMPUTE (I-R)*VAR MATRIX
C
   55 DO 80 K=1,NOBS
      DO 70 J=1,NVAR
      SUM=0.
      DO 60 I=1,NVAR
   60 SUM=SUM+S(I,K)*COV(I,J)
   70 R(J,K)=SUM
   80 CONTINUE
      IF(NPRIR.LT.1) GO TO 90
      DO 84 K=1,NOBS
      DO 82 I=1,NPRIR
      J=IPR(I)
   82 R(I+NOBS,K)=-SP(I)*R(J,K)
   84 CONTINUE
      DO 88 J=1,NPRIR
      L=IPR(J)
      DO 86 I=J,NPRIR
      K=IPR(I)
   86 R(I+NOBS,J+NOBS)=-SP(I)*COV(K,L)*SP(J)
   88 CONTINUE
   90 DO 110 K=1,NOBS
      DO 100 J=K,NOBS
      SUM=0.
      DO 95 I=1,NVAR
   95 SUM=SUM+S(I,K)*R(I,J)
  100 R(J,K)=-SUM
  110 CONTINUE
      NTOT=NOBS+NPRIR
      DO 130 J=1,NTOT
```

```
      DO 120 I=J,NTOT
  120 R(J,I)=R(I,J)
  130 R(J,J)=VAR+R(J,J)
C
C         COMPUTE THEORETICAL FREQUENCIES FOR DATA SETS
C
      TMP=NTOT+1
      DO 135 I=1,NTOT
      TEMP=I
  135 F(I)=TEMP/TMP
      DO 180 K=1,NSETS
      WRITE(6,14) K
C
C         COMPUTE RANDOM NORMAL DEVIATES D AND CORRELATED NORMAL DEVIATES G
C
      DO 150 I=1,NTOT
      SUM=-6.
      DO 140 J=1,12
  140 SUM=SUM+RANUM(NRAN)
  150 D(I)=SIGMA*SUM
      DO 170 J=1,NTOT
      SUM=0.
      DO 160 I=1,NTOT
  160 SUM=SUM+R(I,J)*D(I)
  170 G(J)=SUM/VAR
C
C         ORDER AND PRINT RANDOM NORMAL DEVIATES AND CORRELATED NORMAL DEVIATES
C
      DO 174 I=1,NTOT
      DO 172 J=I,NTOT
      IF(D(J).GE.D(I)) GO TO 172
      TMP=D(I)
      D(I)=D(J)
      D(J)=TMP
  172 CONTINUE
  174 CONTINUE
      WRITE(6,11)
      CALL PRTOTA(D,F,NTOT)
      WRITE(6,16)
      CALL PRTOT(G,NTOT,1,0)
      DO 178 I=1,NTOT
      DO 176 J=I,NTOT
      IF(G(J).GE.G(I)) GO TO 176
      TMP=G(I)
      G(I)=G(J)
      G(J)=TMP
  176 CONTINUE
  178 CONTINUE
      WRITE(6,12)
      CALL PRTOTA(G,F,NTOT)
  180 CONTINUE
C
C         PRINT COVARIANCE MATRIX (I-R)*VAR
C
      WRITE(6,13)
```

```
      CALL PRTOT(R,NTOT,NTOT,NTD)
C
C          COMPUTE AND PRINT CORRELATION MATRIX
C
      DO 210 I=1,NTOT
  210 D(I)=R(I,I)**.5
      DO 230 J=1,NTOT
      TMP=D(J)
      DO 220 I=J,NTOT
      R(I,J)=R(I,J)/(TMP*D(I))
  220 R(J,I)=R(I,J)
  230 CONTINUE
      WRITE(6,15)
      CALL PRTOT(R,NTOT,NTOT,NTD)
      STOP
      END
      SUBROUTINE PRTOT(C,NR,NC,NRD)
C
C          PRINT MATRICES AND VECTORS
C
      DIMENSION C(1)
      IF(NC.EQ.1) GO TO 25
      DO 20 L=1,NC,10
      J10=L+9
      IF(J10.GT.NC) J10=NC
      WRITE(6,35) (J,J=L,J10)
      WRITE(6,50)
      KBC=(L-1)*NRD
      KEC=(J10-1)*NRD
      DO 10 I=1,NR
      KB=KBC+I
      KE=KEC+I
   10 WRITE(6,40) I,(C(K),K=KB,KE,NRD)
      WRITE(6,60)
   20 CONTINUE
      RETURN
   25 N=NR/3
      IF((3*N).NE.NR) N=N+1
      DO 30 K=1,N
   30 WRITE(6,80) (L,C(L),L=K,NR,N)
      RETURN
   35 FORMAT (1H0,8X,I3,9(9X,I3))
   40 FORMAT (1H ,I3,10(1X,G11.5))
   50 FORMAT (1H )
   60 FORMAT (1H0)
   80 FORMAT (1H ,2X,3(I3,7X,G11.5,3X))
      END
      SUBROUTINE PRTOTA(VALA,VALB,NO)
C
C          PRINT OUT VALUES IN TWO GROUPS OF THREE COLUMNS
C
      DIMENSION VALA(NO),VALB(NO)
      NR=NO/2
      IF(2*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
```

```
      WRITE(6,20) (L,VALA(L),VALB(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,2(I3,4X,G11.5,4X,G11.5,4X))
      END
      SUBROUTINE PRTOTC(IVAL,NO)
C
C         PRINT OUT INTEGERS IN THREE GROUPS OF TWO COLUMNS
C
      DIMENSION IVAL(NO)
      NR=NO/3
      IF(3*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(6,20) (L,IVAL(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,3(I3,8X,I4,9X))
      END
      FUNCTION RANUM(IRAN)
      DATA MODU,MULT,NADD/1048576,1027,221589/
      IRAN=MULT*IRAN+NADD
      IRAN=IRAN-(IRAN/MODU)*MODU
      RANUM=FLOAT(IRAN)/FLOAT(MODU)
      RETURN
      END
```

Listing of Inserts to the Regression Ground-Water Flow Program.

```
C
C**INSERT AFTER STATEMENT 490
      ITB=8
      REWIND ITB
      DO 1000 J=1,NVAR
 1000 WRITE(ITB) (A(I,J),I=J,NVAR)
      K=0
      DO 1100 I=1,NIJ
      IF(W(I).LT.1.E-10) GO TO 1100
      K=K+1
      CXS(K)=W(I)*W(I)
 1100 CONTINUE
      WRITE(ITB) (CXS(I),I=1,NOBS)
C
C
C**INSERT AFTER STATEMENT 520
      M=0
      N=0
      DO 1500 I=1,NIJ
      IF(W(I).LT.1.E-10) GO TO 1500
      N=N+1
      L=IN(I)
      IF(L.GT.0) GO TO 1300
      DO 1200 K=1,NVAR
 1200 X(K,N)=0.
      L=-L-1
      IZ=IBZN(L+NQSD)
      K=IBPA(IZ)
      IF(K.GT.NBQP) X(K,N)=PLA(L)
      K=IBPB(IZ)
      IF(K.GT.NBQP) X(K,N)=X(K,N)+PLB(L)
      GO TO 1500
 1300 M=M+1
      DO 1400 K=1,NVAR
 1400 X(K,N)=S(K,M)
 1500 CONTINUE
      DO 1600 J=1,NOBS
 1600 WRITE(ITB) (X(I,J),I=1,NVAR)
C
```

References Cited

Beck, J. V., and Arnold, K. J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Draper, N., and Smith, H., 1981, Applied regression analysis [2d ed]: New York, John Wiley, 709 p.

Lieberman, G. J., 1961, Prediction regions for several predictions from a single regression line: Technometrics, v. 3, p. 21-27.

Additional Reading

Beck, J. V., and Arnold, K. J., 1977, Parameter estimation in engineering and science: New York, John Wiley, 501 p.

Graybill, F. A., 1976, Theory and application of the linear model: North Scituate, Massachusetts, Duxbury, 704 p.

6. SOME ADVANCED TOPICS

6.1. ADVANCED MODELS

6.1.1. Regression When the Dependent Variable is Implicit

In all cases considered in previous sections, the assumption has been made that the deterministic part of the model equation can be solved explicitly for the dependent variable. However, this may not be true in some cases. An example of such a model written in terms of the true value for the dependent variable $f = f(\xi, \beta_1, \beta_2)$ is

$$\frac{\beta_1 k}{3(\beta_1 + \beta_2)} \left\{ \tfrac{1}{2} \ln\left(\frac{(k + f)^2}{k^2 - kf + f^2} \right) + \sqrt{3} \left[\tan^{-1}\left(\frac{2f-k}{k\sqrt{3}} \right) - \tan^{-1}\left(\frac{-1}{\sqrt{3}} \right) \right] \right\} - \xi = 0 \qquad (6.1\text{-}1)$$

where $k = \sqrt[3]{\beta_1 + \beta_2}$ and $f = f(\xi, \beta_1, \beta_2)$ is the exact solution of (6.1-1). As can be seen, f is implicit in the model equation and cannot be directly solved for. A general deterministic form for an exact model (that is, a model that does not contain ε) where the dependent variable is implicit is

$$g[f(\underline{\xi}, \underline{\beta}), \underline{\xi}, \underline{\beta}] = 0 \qquad (6.1\text{-}2)$$

where $\underline{\xi}$ and $\underline{\beta}$ are defined as usual.

Based on (6.1-2), a true regression model can be written in terms of observation vector $\underline{Y}$ and disturbances $\underline{\varepsilon}$ in the usual form

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{\beta}) + \underline{\varepsilon} \tag{6.1-3}$$

where $\underline{f}(\underline{\xi},\underline{\beta})$ is the vector of order n that is computed from

$$\underline{g}[\underline{f}(\underline{\xi},\underline{\beta}),\underline{\xi}, \underline{\beta}] = \underline{0} \tag{6.1-4}$$

In (6.1-4) vector $\underline{g}$ of order n represents n equations, each of which has the form of (6.1-2) written for an observation point. As an example, (6.1-1) would be written in the form

$$g_i = \frac{\beta_1 k}{3(\beta_1 + \beta_2)} \left\{ \tfrac{1}{2} \ln\left(\frac{(k + f_i)^2}{k^2 - k f_i + f_i^2} \right) + \sqrt{3} \left[\tan^{-1}\left(\frac{2f_i - k}{k\sqrt{3}} \right) - \tan^{-1}\left(\frac{-1}{\sqrt{3}} \right) \right] \right\} - \xi_i = 0, \; i = 1, 2, \ldots, n \tag{6.1-5}$$

Note that in (6.1-5) only f evaluated at point i (that is, f_i) appears in the equation to compute g_i. However, in general this equation could contain values of f evaluated at any number of the possible points $j = 1, 2, \ldots, n$. An example of this type of model is the numerical model discussed in the next section.

The estimated regression model derived from (6.1-3) and (6.1-4) is

$$\underline{Y} = \underline{f}(\underline{\xi}, \underline{b}) + \underline{e} \tag{6.1-6}$$

and

$$\underline{g}[\underline{f}(\underline{\xi}, \underline{b}), \underline{\xi}, \underline{b}] = \underline{0} \tag{6.1-7}$$

where $\underline{b}$ and $\underline{e}$ are, as usual, general estimates of $\underline{\beta}$ and $\underline{\varepsilon}$, respectively.

The general approach to solving the implicit-variable problem is very similar to that followed in section 3.3.1. First, the dependent variable values are written using a Taylor series expansion about an initial set of parameters. Then, from this, the linearized regression problem is set up and solved recursively to give the final solution to the nonlinear problem.

Taylor series expansion of $\underline{f}$ about an arbitrary initial set of parameters $\underline{b}_0$ can be written in the form of (3.3-5),

$$\underline{f}(\underline{\xi}, \underline{b}) \cong \underline{f}_0 + \underline{X}_0(\underline{b} - \underline{b}_0) \tag{6.1-8}$$

where

$$\underline{f}_0 = \underline{f}(\underline{\xi}, \underline{b}_0) \tag{6.1-9}$$

$$\underline{X}_0 = \left\{ \left. \frac{\partial f_i}{\partial b_j} \right|_{\underline{b} = \underline{b}_0} \right\} \tag{6.1-10}$$

By using (6.1-6), (6.1-8) can be modified to give the estimated linearized regression model

$$\underline{Y} - \underline{f}_0 = \underline{X}_0(\underline{b} - \underline{b}_0) + \underline{e} \tag{6.1-11}$$

which is exactly the same as the model used for the standard nonlinear case discussed in section 3.3.1.

To solve the linearized regression model based on (6.1-11) $\underline{f}_0$ can be computed from (6.1-7) with $\underline{b} = \underline{b}_0$, and $\underline{X}_0$ can be computed by implicitly differentiating (6.1-7) with respect to b_j ($j = 1, 2, \ldots, p$), setting $\underline{b} = \underline{b}_0$, then solving for $\underline{X}_0$. To accomplish this computation of $\underline{X}_0$, note that for any differential change $d\underline{b}$ in parameter vector $\underline{b}$ to produce a new solution $\underline{f}$ of $\underline{g} = \underline{0}$, the total differential $d\underline{g}$ must equal zero because $\underline{g}$ is always zero. Hence, by employing the chain rule of calculus, there results

$$d\underline{g} = \left(\frac{\partial \underline{g}}{\partial b_j} + \underline{M}\,\frac{\partial \underline{f}}{\partial b_j}\right) db_j = \underline{0}, \quad j = 1, 2, \ldots, p \tag{6.1-12}$$

where $\underline{M} = \{M_{ij}\} = \{\partial g_i / \partial f_j\}$. Note that if g_i contains only f_i, then $\underline{M}$ is diagonal. Equation (6.1-12) can be evaluated using $\underline{b} = \underline{b}_0$ and $\underline{f}_0 = \underline{f}(\underline{\xi}, \underline{b}_0)$ to give

$$\left(\frac{\partial \underline{f}}{\partial b_j}\right)_0 = \underline{X}_j^0 = -\,\underline{M}_0^{-1} \left(\frac{\partial \underline{g}}{\partial b_j}\right)_0, \quad j = 1, 2, \ldots, p \tag{6.1-13}$$

where

$$\underline{M}_0 = \underline{M}(\underline{f}_0, \underline{\xi}, \underline{b}_0) = \left\{ \frac{\partial g_i}{\partial f_j} \,\middle|\, \underline{f} = \underline{f}_0 \right\}_{(n \times n)} \tag{6.1-14}$$

and subscript (or superscript) 0 means that the quantity is evaluated using $\underline{b} = \underline{b}_0$ and $\underline{f} = \underline{f}_0$. Solution of (6.1-7) for $\underline{f}_0$ (using, for example, Newton

iteration) followed by solution of (6.1-13) for $\underline{X}_0$ provides a convenient method of obtaining initial values $\underline{f}_0$ and $\underline{X}_0$ from initial parameter estimates $\underline{b}_0$. However, for subsequent iterations this method can be time-consuming because it involves solving $\underline{g} = \underline{0}$ each time a new vector $\underline{f}$ and a new set of sensitivity vectors $\underline{X}_j$ are to be computed from an updated parameter set.

A good method for computing good approximate values of $\underline{f}$ and $\underline{X}$ corresponding to some parameter set $\underline{b}$ involves approximating $\underline{g}$ with another Taylor series expansion. If $\underline{b}$ is close enough to $\underline{b}_0$ to allow dropping all terms except linear terms, then Taylor series expansion of (6.1-7) about initial set of dependent variable values $\underline{f}_0$ can be written as

$$\underline{g}[\underline{f}(\underline{\xi}, \underline{b}), \underline{\xi}, \underline{b}] = \underline{0}$$

$$\cong \underline{g}(\underline{f}_0, \underline{\xi}, \underline{b}) + \underline{M}_b(\underline{f} - \underline{f}_0) \qquad (6.1\text{-}15)$$

where $\underline{M}_b = \underline{M}(\underline{f}_0, \underline{\xi}, \underline{b})$ and $\underline{f} \cong \underline{f}\ (\underline{\xi}, \underline{b})$. By knowing $\underline{b}$, (6.1-15) may be solved for $\underline{f}$. Corresponding values of $\underline{X}$ are obtained as follows. If (6.1-15) is implicitly differentiated with respect to b_j ($j = 1, 2, \ldots, p$), there results

$$\frac{\partial \underline{g}}{\partial b_j} \cong - \underline{M}_b \frac{\partial \underline{f}}{\partial b_j} - \frac{\partial \underline{M}_b}{\partial b_j} (\underline{f} - \underline{f}_0) \qquad (6.1\text{-}16)$$

or,

$$\frac{\partial \underline{f}}{\partial b_j} = \underline{X}_j \cong - \underline{M}_b^{-1} \left(\frac{\partial \underline{g}}{\partial b_j} + \frac{\partial \underline{M}_b}{\partial b_j} (\underline{f} - \underline{f}_0) \right), \quad j = 1, 2, \ldots, p \qquad (6.1\text{-}17)$$

By using the above results, solution of the nonlinear regression problem is obtained by a procedure analogous to the procedure followed for the standard nonlinear problem. As indicated previously, to begin the first iteration assume an initial set of parameters $\underline{b}_0$ and solve (6.1-7) for $\underline{f}_0 = \underline{f}(\underline{\xi}, \underline{b}_0)$. Then solve (6.1-13) for $\underline{X}_0$. Next, form and solve normal equations (3.3-12) by scaling (6.1-11) with respect to $\underline{B}_0 = diag\ \{b_1^0, b_2^0, \ldots, b_p^0\}$, substituting the result into $S(\underline{b})$ (given by (3.3-4)), and then minimizing $S(\underline{b})$ with respect to $\underline{b}$. That is, form and solve for $\underline{d}_1$

$$\underline{Z}_0^T \underline{\omega} \underline{Z}_0 \underline{d}_1 = \underline{Z}_0^T \underline{\omega} (\underline{Y} - \underline{f}_0) \tag{6.1-18}$$

where

$$\underline{Z}_0 = \underline{X}_0 \underline{B}_0 \tag{6.1-19}$$

$$\underline{d}_1 = \underline{B}_0^{-1} (\underline{b}_1 - \underline{b}_0) \tag{6.1-20}$$

For the second iteration, solve (6.1-15), written using $\underline{b} = \underline{b}_1$, for $\underline{f} - \underline{f}_0 = \underline{f}_1 - \underline{f}_0$. The result can be stated as

$$\underline{u}_1 = - \underline{M}_1^{-1} \underline{g}_1 \tag{6.1-21}$$

where subscript 1 on $\underline{M}$ and $\underline{g}$ indicates evaluation using the most recent values available for $\underline{f}$ and $\underline{b}$ (that is, $\underline{f}_0$ and $\underline{b}_1$), and

$$\underline{u}_1 = \underline{f}_1 - \underline{f}_0 \tag{6.1-22}$$

Next, solve (6.1-17) for $\underline{X}_1$:

$$\underline{X}_j^1 = - \underline{M}_1^{-1}\left[\left(\frac{\partial \underline{g}}{\partial b_j}\right)_1 + \left(\frac{\partial \underline{M}_b}{\partial b_j}\right)_1 \underline{u}_1\right] , \; j = 1, 2, \ldots, p \tag{6.1-23}$$

where $(\;)_1$ indicates evaluation using $\underline{f}_0$ and $\underline{b}_1$. Finally, form and solve the normal equations, written in terms of $\underline{Z}_1$ and $\underline{f}_1$, for $\underline{d}_2$.

For general iteration r, the equations to solve are

$$\underline{u}_r = - \underline{M}_r^{-1} \underline{g}_r \tag{6.1-24}$$

$$\underline{f}_r = \underline{u}_r + \underline{f}_{r-1} \tag{6.1-25}$$

$$\underline{X}_j^r = - \underline{M}_r^{-1}\left[\left(\frac{\partial \underline{g}}{\partial b_j}\right)_r + \left(\frac{\partial \underline{M}_b}{\partial b_j}\right)_r \underline{u}_r\right] , \; j = 1, 2, \ldots, p \tag{6.1-26}$$

$$\underline{Z}_r = \underline{X}_r \underline{B}_r \tag{6.1-27}$$

$$\underline{Z}_r^T \underline{\omega} \underline{Z}_r \underline{d}_{r+1} = \underline{Z}_r^T \underline{\omega} (\underline{Y} - \underline{f}_r) \tag{6.1-28}$$

$$\underline{b}_{r+1} = \underline{B}_r \underline{d}_{r+1} + \underline{b}_r \tag{6.1-29}$$

where $\underline{f}_0 = \underline{f}(\underline{\xi}, \underline{b}_0)$ so that $\underline{g}_0 = \underline{0}$, and $\underline{f}_{-1} = \underline{f}_0$.

At convergence of the solution $\underline{g}_r$, $\underline{u}_r$, and $\underline{d}_{r+1}$ all tend to zero so that $\underline{f}_r = \underline{f}(\underline{\xi}, \hat{\underline{b}})$ where $\hat{\underline{b}} = \underline{b}_{r+1} \cong \underline{b}_r$. At this point $S(\hat{\underline{b}}) = (\underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}}))^T \underline{\omega}(\underline{Y} - \underline{f}(\underline{\xi}, \hat{\underline{b}}))$ is at a minimum, and the nonlinear regression problem has been solved.

The solution procedure given by (6.1-24) through (6.1-29) can actually be considered to be a generalization of the Gauss-Newton procedure discussed in section 3.3.1, because if the standard nonlinear model is stated in the form

$$\underline{g} = \underline{f}(\underline{\xi}, \underline{b}) - \underline{f} = \underline{0} \tag{6.1-30}$$

then

$$\underline{g}_r = \underline{f}(\underline{\xi}, \underline{b}_r) - \underline{f}_{r-1} \tag{6.1-31}$$

$$\underline{M}_r = - \underline{I} \tag{6.1-32}$$

$$\left(\frac{\partial \underline{g}}{\partial b_j}\right)_r = \left(\frac{\partial \underline{f}}{\partial b_j}\right)_r = \underline{X}_j^r \tag{6.1-33}$$

and (6.1-24) through (6.1-29) become the standard Gauss-Newton algorithm.

Iteration parameters ρ and μ should be applied to the present method in the same manner as they are for the Gauss-Newton method. Use of ρ to modify step size $\underline{d}_{r+1}$ leads to (3.3-18) ($\underline{b}_{r+1} = \rho \underline{B}_r \underline{d}_{r+1} + \underline{b}_r$) to compute $\underline{b}_{r+1}$. To employ μ, definitions analogous to (3.3-19) and (3.3-20) are used, and (6.1-28) is transformed to

$$(\underline{S}_r^T \underline{\omega} \underline{S}_r + \mu \underline{I}) \underline{\delta}_{r+1} = \underline{S}_r^T \underline{\omega} (\underline{Y} - \underline{f}_r) \tag{6.1-34}$$

where

$$\underline{\delta}_{r+1} = \underline{C}^{-1} \underline{d}_{r+1} \tag{6.1-35}$$

$$\underline{S}_r = \underline{Z}_r \underline{C} \tag{6.1-36}$$

The method for solution of the implicit-variable model given here requires the same three conditions to guarantee convergence to a gobal minimum as discussed for the modified Gauss-Newton method in section 3.3.3. However, in addition, the method requires that $\underline{M}$ and $\partial \underline{g}/\partial b_j$ (j = 1, 2, ..., p) be continuous and unique for all $\underline{b}$ belonging to region R (see (3.3-26)).

Solution Algorithm.

1) Before the first iteration, solve (6.1-7) for $\underline{f}_0$ using an initial estimate $\underline{b}_0$ for $\underline{b}$, and set $\underline{f}_{-1} = \underline{f}_0$.
2) Solve (6.1-24) and the combination of (6.1-26) and (6.1-27),

$$\underline{M}_r \underline{Z}_j^r = - \left[\left(\frac{\partial \underline{g}}{\partial b_j} \right)_r + \left(\frac{\partial \underline{M}}{\partial b_j} \right)_r \underline{u}_r \right] b_j^r \tag{6.1-37}$$

for $\underline{u}_r$ and $\underline{Z}_j^r$ (j=1, 2, ..., p).

3) Solve (6.1-25) for $\underline{f}_r$.
4) Solve (6.1-34) for $\underline{\delta}_{r+1}$.
5) Solve (6.1-35) for $\underline{d}_{r+1}$.
6) Solve (3.3-18) for $\underline{b}_{r+1}$.
7) Test to determine if $|d_i^{r+1}| > \varepsilon$ (i=1, 2, ..., p).
8) If $|d_i^{r+1}| > \varepsilon$, increment r by one and return to 2. If not, then the process has converged.

6.1.2. Regression When the Implicit-Variable Model is Numerical

If the numerical model assumes the general form of (3.3-22), which for convenience is restated here as

$$\underline{D}(\underline{h},\underline{\xi},\underline{\beta})\underline{h} = \underline{q}(\underline{h},\underline{\xi},\underline{\beta}) \tag{6.1-38}$$

then the method derived in the previous section can be applied. The solution can be conceptually developed in two stages, first making the assumption that numerical solution points coincide exactly with observation points, which implies m = n, then second relaxing the assumption by following either of the two procedures described to obtain $\underline{f}$ from $\underline{h}$ for the Gauss-Newton method in section 3.3.2.

To develop the first stage of the solution, first note that because m = n, (6.1-38) may be written in the form $\underline{g} = \underline{0}$ analogous to (6.1-7):

$$\underline{g} = \underline{q} - \underline{D}\,\underline{h} = \underline{0} \tag{6.1-39}$$

where $\underline{h} \equiv \underline{f}$. Next expand (6.1-39) in a Taylor series to give equations exactly analogous to (6.1-24) and (6.1-26). Pertinent quantities in these equations are given by

$$\underline{M}_i^r = \left(\frac{\partial \underline{q}}{\partial h_i}\right)_r - \left(\frac{\partial \underline{D}}{\partial h_i}\right)\underline{h}_{r-1} - \underline{D}_i^r, \quad i = 1, 2, \ldots, m \tag{6.1-40}$$

$$\left(\frac{\partial \underline{g}}{\partial b_j}\right)_r = \left(\frac{\partial \underline{q}}{\partial b_j}\right)_r - \left(\frac{\partial \underline{D}}{\partial b_j}\right)_r \underline{h}_{r-1} \tag{6.1-41}$$

$$\left(\frac{\partial \underline{M}_{bi}}{\partial b_j}\right)_r = \left(\frac{\partial^2 \underline{q}}{\partial b_j \partial h_i}\right)_r - \left(\frac{\partial^2 \underline{D}}{\partial b_j \partial h_i}\right)_r \underline{h}_{r-1} - \left(\frac{\partial \underline{D}_i}{\partial b_j}\right)_r \tag{6.1-42}$$

where subscript i on a matrix denotes a column of the matrix. By using (6.1-41) and (6.1-42), an equation analogous to (6.1-26) can be written

$$\left(\frac{\partial \underline{h}}{\partial b_j}\right)_r = -\underline{M}_r^{-1}\left\{\left(\frac{\partial \underline{q}}{\partial b_j}\right)_r - \left(\frac{\partial \underline{D}}{\partial b_j}\right)_r \underline{h}_r + \sum_{i=1}^{m}\left[\left(\frac{\partial^2 \underline{q}}{\partial b_j \partial h_i}\right)_r - \left(\frac{\partial^2 \underline{D}}{\partial b_j \partial h_i}\right)_r \underline{h}_{r-1}\right] u_i^r\right\}, \quad j = 1, 2, \ldots, p \tag{6.1-43}$$

The second stage results from using one of the two procedures for obtaining $\underline{f}$ from $\underline{h}$ described in section 3.3.2 to obtain $\underline{f}$ from $\underline{h}$ and $\underline{X}_j$ from $\partial \underline{h}/\partial b_j$ in the present case. With $\underline{f}$ and $\underline{X}$ defined, the solution algorithm of section 6.1.1 can be applied directly.

6.2. BEALE'S MEASURE OF NONLINEARITY

Most of the methodology developed to analyze and use regression models is based on the assumption that the model is linear in the parameters. In the case that the model is nonlinear, Beale (1960) has developed a measure of degree of nonlinearity with respect to the confidence regions on parameters.

To develop the measure, consider a linearized model of the form of (5.1-1), where, for convenience here, general estimate $\underline{b}$ replaces $\underline{\beta}$ and $\underline{b}_0$ is set equal to $\hat{\underline{b}}$ to result in

$$\underline{Y}_b^o = \hat{\underline{Y}} + \underline{X}(\underline{b} - \hat{\underline{b}}) \tag{6.2-1}$$

where $\hat{\underline{Y}} \equiv \underline{f}(\underline{\xi}, \hat{\underline{b}})$. To emphasize that $\underline{Y}_b^o$ is a linearized estimate of $\underline{Y}_b$, a superscript o is appended to $\underline{Y}_b$. If $\underline{Y}_b^o$ is calculated from (6.2-1) for m sets of parameter vectors $\underline{b}_\ell$ ($\ell = 1, 2, \ldots, m$), then a measure c^2 of model nonlinearity in the region covered by the varied parameter sets is

$$c^2 = \sum_{\ell=1}^{m} (\underline{Y}_\ell - \underline{Y}_\ell^o)^T \underline{\omega} (\underline{Y}_\ell - \underline{Y}_\ell^o) \tag{6.2-2}$$

where $\underline{Y}_\ell$ is $\underline{Y}_b$ computed using parameter set $\underline{b}_\ell$ in the actual nonlinear model, and $\underline{Y}_\ell^o$ is $\underline{Y}_b^o$ computed using parameter set $\underline{b}_\ell$ in linearized model (6.2-1). Equation (6.2-2) is the sum of squared distances (i.e., squared lengths of vectors) between points $\underline{\omega}^{\frac{1}{2}}\underline{Y}_\ell^o$ and $\underline{\omega}^{\frac{1}{2}}\underline{Y}_\ell$ in observation space. (Recall that the distance between two points is the length of the vector joining the points, and that the squared length of a vector is given by the sum of squared lengths of its components.)

As explained further on, the most useful measure of nonlinearity is obtained by multiplying (6.2-2) by the quantity

$$qs^2 / \sum_{\ell=1}^{m} [(\underline{Y}_\ell - \hat{\underline{Y}})^T \underline{\omega} (\underline{Y}_\ell - \hat{\underline{Y}})]^2$$

to obtain

$$\hat{N}_b = qs^2 \frac{\sum_{\ell=1}^{m} (\underline{Y}_\ell - \underline{Y}_\ell^o)^T \underline{w} (\underline{Y}_\ell - \underline{Y}_\ell^o)}{\sum_{\ell=1}^{m} [(\underline{Y}_\ell - \hat{\underline{Y}})^T \underline{w} (\underline{Y}_\ell - \hat{\underline{Y}})]^2} \qquad (6.2\text{-}3)$$

which is an extension for $q \leq p$ of Beale's measure of nonlinearity (Beale, 1960, p. 54-55).

Equation (6.2-3) can be justified as follows (see also Guttman and Meeter, 1965). The weighted distance between $\underline{Y}_\ell^o$ and $\underline{Y}_\ell$ is designated εd so that the geometric relationships among weighted vectors $\hat{\underline{Y}}$, $\underline{Y}_\ell$, and $\underline{Y}_\ell^o$ can be diagrammed in observation space as shown in figure 6.2-1.

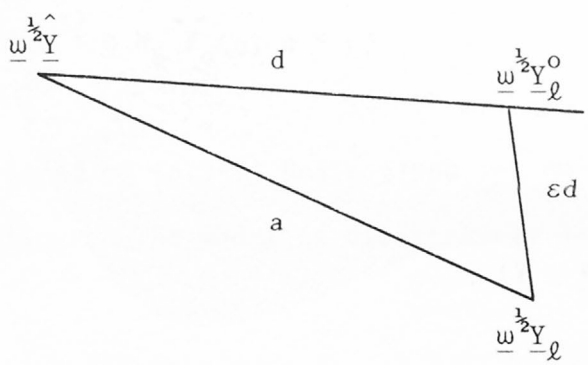


fig. 6.2-1

Now, if $\varepsilon << 1$, $d^2 \cong a^2$ and

$$\varepsilon^2 = \frac{(\varepsilon d)^2}{d^2} = \frac{d^2(\varepsilon d)^2}{d^4} \cong \frac{d^2(\varepsilon d)^2}{(a^2)^2}$$

$$= d^2 \frac{(\underline{Y}_\ell - \underline{Y}_\ell^o)^T \underline{\omega}(\underline{Y}_\ell - \underline{Y}_\ell^o)}{[(\underline{Y}_\ell - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y}_\ell - \hat{\underline{Y}})]^2} \qquad (6.2\text{-}4)$$

where, by definition,

$$d^2 = (\underline{Y}_\ell^o - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y}_\ell^o - \hat{\underline{Y}})$$

$$= [\underline{X}(\underline{b}_\ell - \hat{\underline{b}})]^T \underline{\omega}[\underline{X}(\underline{b}_\ell - \hat{\underline{b}})]$$

$$= (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega}\, \underline{X}(\underline{b}_\ell - \hat{\underline{b}}) \qquad (6.2\text{-}5)$$

Now

$$(\underline{Y} - \underline{Y}_\ell^o)^T \underline{\omega}(\underline{Y} - \underline{Y}_\ell^o) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})$$

$$= (\underline{Y} - \hat{\underline{Y}} - \underline{X}(\underline{b}_\ell - \hat{\underline{b}}))^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}} - \underline{X}(\underline{b}_\ell - \hat{\underline{b}}))$$

$$- (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}})$$

$$= -2(\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}}) + (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega}\, \underline{X}(\underline{b}_\ell - \hat{\underline{b}})$$

$$= (\underline{b}_\ell - \hat{\underline{b}})^T \underline{X}^T \underline{\omega}\, \underline{X}(\underline{b}_\ell - \hat{\underline{b}}) \qquad (6.2\text{-}6)$$

where (5.1-9) was used. Combination of (6.2-5) and (6.2-6) shows that

$$d^2 = (\underline{Y} - \underline{Y}_\ell^o)^T \underline{\omega}(\underline{Y} - \underline{Y}_\ell^o) - (\underline{Y} - \hat{\underline{Y}})^T \underline{\omega}(\underline{Y} - \hat{\underline{Y}}) \tag{6.2-7}$$

If $\underline{Y}_\ell^o$ is assumed to lie on the edge of confidence region (5.6-12) so that $\underline{b}_\ell^T = [\tilde{\underline{b}}_{1\ell}^T \ \tilde{\underline{\beta}}_2^T]$, then from (5.6-12) and (6.2-7) it can be seen that

$$d^2 = qs^2 F_\alpha(q, n - p) \tag{6.2-8}$$

Hence, if both numerator and denominator of (6.2-4) are averaged over m sets of parameters, there results

$$E(\varepsilon^2) \cong \hat{N}_b \, F_\alpha(q, n - p) \tag{6.2-9}$$

Based on (6.2-9) Beale (1960, p. 60) ranked the degrees of nonlinearity as follows: The model is disastrously nonlinear if

$$\hat{N}_b > 1/F_\alpha(q, n - p) \tag{6.2-10}$$

because in this case $E(\varepsilon^2)>1$ and the discrepancy is actually greater than d. If

$$\hat{N}_b < 0.01 \, /F_\alpha(q, n - p) \tag{6.2-11}$$

then the model is classed as being effectively linear because $E(\varepsilon^2)<0.01$. For points in between, the linear theory is adequate to give a rough idea of significance but may not bring out full implications of the analysis.

Rigorous use of (6.2-10) and (6.2-11) theoretically requires that disturbances be distributed normally. However, it would be convenient to be able to gage the degree of nonlinearity of the model irrespective of the properties of $\underline{\varepsilon}$. If the confidence region (5.6-12) were large enough to encompass virtually all physically plausible sets of parameters, then model nonlinearity as assessed using (6.2-10) and (6.2-11) would be meaningful. Based upon past experience, F values generated using $\alpha = 0.05$ has been found to be adequate to gage nonlinearity.

As a final item, Beale (1960, p. 55) stated that, whereas the value of $\hat{N}_b$ may depend on the configuration of the points defined by the vectors $\underline{b}_\ell$, $\hat{N}_b$ should not depend greatly on the number of points used or their distances from $\hat{\underline{b}}$ if these distances are not too great. Because (6.2-9) is justified by assuming that the points lie on the edge of the confidence region, a reasonable way to obtain the points is to choose them from (5.6-14). Thus, one could use $m \leq 2q$ sets of parameters $\underline{b}_\ell$. Note that, whether or not the model is linear, $\tilde{\underline{b}}_{1\ell}$ and $\tilde{\underline{\beta}}_2$ corresponding to the partition of $\underline{b}$ given in (5.6-3) are properly chosen without the necessity of performing additional least squares solutions to obtain each set $\tilde{\underline{b}}_{1\ell}$. This is true because subset $\tilde{\underline{b}}_{1\ell}$ is required to lie on the edge of the linearized confidence region.

Problem 6.2-1

Four sets of parameters that correspond to four points on the edge of the linearized confidence region (5.6-13) result from problem 5.6-2. These four sets of parameters can be subdivided into two groups of two. Pick two different parameter sets from the two groups, and compute two corresponding sets of drawdowns at the observation points using the nonlinear (Theis) model. Then, using the Beale's measure program (appendix 6.4-2), find Beale's measure. Is the model nearly linear?

Problem 6.2-2

Use the four parameter sets resulting from problem 5.6-3 in the non-linear regression flow program of appendix (4.3.4), as augmented by the inserts of appendix (6.4.1), to compute Beale's measure. Are the various statistical measures obtained from the linearized model approximately valid (at least as determined from the four parameter sets employed)?

6.3. COMPATIBILITY OF PRIOR AND REGRESSION ESTIMATES OF PARAMETERS

If the regression model contains prior information on the parameters, an important part of the analysis to determine whether or not the model is correct is to test the null hypothesis that the prior and sample information are in agreement; in other words,

$$H_0: E(\underline{Y}_p - \underline{Y}_{0p}) = \underline{X}_p(\underline{\beta} - \underline{b}_0) \text{ versus } H_1: E(\underline{Y}_p - \underline{Y}_{0p}) \neq \underline{X}_p(\underline{\beta} - \underline{b}_0).$$

As indicated in section 5.5, graphical analysis of residuals can usually detect an incompatibility between sample and prior information. However, in some cases an additional test might be desired. Theil (1963) shows that the test statistic

$$\Gamma = (\underline{Y}_p - \underline{Y}_{0p} - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0))^T[\sigma^2\underline{X}_p(\underline{X}_s^T\underline{\omega}_s\underline{X}_s)^{-1}\underline{X}_p^T + \sigma^2\underline{\omega}_p^{-1}]^{-1} \cdot(\underline{Y}_p - \underline{Y}_{0p} - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0)) \qquad (6.3\text{-}1)$$

where vector $\hat{\underline{b}}^*$ is the ordinary least squares estimate of vector $\underline{\beta}$, is Chi square distributed with n_p degrees of freedom ($\chi^2(n_p)$) provided that all of

the assumptions (5.2-1) through (5.2-3), (5.2-6), and (5.2-10) hold true, σ^2 is known, and $\underline{\omega}$ is of the form

$$\underline{\omega} = \begin{bmatrix} \underline{\omega}_s & \underline{0} \\ \underline{0} & \underline{\omega}_p \end{bmatrix} \tag{6.3-2}$$

where $\underline{\omega}_s$ and $\underline{\omega}_p$ are known and symmetric positive definite of order n_s and n_p, respectively. If $\underline{\omega}$ is of the form

$$\underline{\omega} = \begin{bmatrix} \underline{V}_s^{-1} & \underline{0} \\ \underline{0} & \underline{U}^{-1}\sigma^2 \end{bmatrix} \tag{6.3-3}$$

and σ^2 is unknown, then the test statistic

$$\hat{\Gamma} = (\underline{Y}_p - \underline{Y}_{0p} - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0))^T[s^2\underline{X}_p(\underline{X}_s^T\underline{V}_s^{-1}\underline{X}_s)^{-1}\underline{X}_p^T + \underline{U}]^{-1} \cdot(\underline{Y}_p - \underline{Y}_{0p} - \underline{X}_p(\hat{\underline{b}}^* - \underline{b}_0)) \tag{6.3-4}$$

is asymptotically $\chi^2(n_p)$ distributed. If the computed value of $\hat{\Gamma}$, $\hat{\gamma}$, is greater than $\chi^2_\alpha(n_p)$, where α indicates significance level, then the null hypothesis is rejected.

Problem 6.3-1

Using (6.3-4), test the compatibility of the prior estimate of the boundary head, Y_{p2}, and the pure regression estimate, $\hat{b}_2^*$, of problem 3.2-1. To conduct this test you will have to do an ordinary least squares regression. The model of appendix (4.3.4) may be employed for this in the

same manner as for problem 4.2-1. Use the model output to obtain the necessary quantities in (6.3-4). Hint: note that

$$s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_{ss}^{-1} \underline{X}_s)^{-1} \underline{X}_p^T = s^2 \underline{Z}_p (\underline{Z}_s^T \underline{V}_{ss}^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T$$

Can you show this?

6.4. APPENDICES

6.4.1. Documentation of Program to Compute Beale's Measure

This program performs a straightforward computation of Beale's measure, (6.2-3). Vectors $\underline{Y}_\ell$, $\underline{Y}_\ell^o$, and $\hat{\underline{Y}}$ are assumed to be composed of sample information and direct prior information on some or all parameters. The weight matrix for sample and prior information is assumed to be given in the form (3.4-12), and the sensitivity matrix for the prior information is assumed to be given by (4.1-6).

There are two versions of the program. One is for general use, and all variables needed for the calculation must be read in. The other version is designed to be an integral part of the regression ground-water model documented in appendix 4.3.4 and requires only q and the extra sets of parameters needed for the Beale's measure calculation as input in addition to input already required for the regression solution.

Input data for General Version.

Card Set A.

Problem size information; one card (format 4I5, F10.0).

Card Columns	Variable	Definition
1-5	NVAR	Number of parameters, p.
6-10	NRES	Number of restrictions, q.

11-15	NOBS	Number of sample observations, n_s.
16-20	NPRIR	Number of regression parameters having direct prior information, n_p.
21-25	NPTS	Number of data sets to compute Beale's measure, m.
26-35	VAR	Error variance, s^2.

Card Set B.

Estimated regression parameters, $\hat{\underline{b}}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	BOPT(1)	Estimated regression parameters,
11-20	BOPT(2)	entered sequentially from 1
⋮	⋮	through NVAR.
	BOPT(NVAR)	

Card Set C.

Dependent variable vector for sample information, $\hat{\underline{Y}}_s$, computed using $\hat{\underline{b}}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	YOPT(1)	Computed dependent variable
11-20	YOPT(2)	values, entered sequentially
⋮	⋮	from 1 through NOBS.
	YOPT(NOBS)	

Card Set D.

Weight matrix for sample information, $\underline{V}_s^{-1}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	W(1)	Diagonal weight matrix for
11-20	W(2)	sample information, entered

⋮	⋮	sequentially from 1 through NOBS.
	W(NOBS)	

Card Set E.

Scaled sensitivity matrix for sample information, $\underline{Z}_s$ (format 8F10.0).

Card columns	Variable	Definition
1-10	S(1,1)	Scaled sensitivity matrix for sample information, entered sequentially 1 through NVAR for each observation. Each new observation begins a new card, for a total of NOBS observations.
11-20	S(2,1)	
⋮	⋮	
	S(NVAR,1)	
1-10	S(1,2)	
⋮	⋮	
	S(NVAR,2)	
	⋮	
	S(NVAR,NOBS)	

Card Set F.

Parameter numbers having prior information (format 16I5).

Card columns	Variable	Definition
1-5	IPR(1)	Array subscript numbers for regression parameters in BOPT(I) having prior information, entered in any order from 1 through NPRIR.
6-10	IPR(2)	
⋮	⋮	
	IPR(NPRIR)	

Omit card set if NPRIR = 0.

Card Set G.

Standard deviation matrix for prior information, $\underline{U}^{\frac{1}{2}}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	WP(1)	Diagonal standard deviation matrix for prior information, entered in the same order as IPR(I) from 1 through NPRIR.
⋮	⋮	
	WP(NPRIR)	

Omit card set if NPRIR = 0.

Card Set H.

Alternate parameters sets, $\underline{b}_{\ell}$ (format 8F10.0). This card set and the next one are read in sequence (H, I, H, I, ...) a total of NPTS times.

Card columns	Variable	Definition
1-10	B(1)	Alternate parameter sets, entered sequentially 1 through NVAR. Order must be the same as for BOPT(I).
11-20	B(2)	
⋮	⋮	
	B(NVAR)	

Card Set I.

Alternate dependent variable vectors for sample information, $\underline{Y}_{s\ell}$, computed using $\underline{b}_{\ell}$ (format 8F10.0).

Card columns	Variable	Definition
1-10	YC(1)	Alternate sample dependent variable values computed using the nonlinear model, entered sequentially 1 through NOBS. Order must be the same as for YOPT(I).
11-20	YC(2)	
⋮	⋮	
	YC(NOBS)	

Output for General Version. Output is all clearly labeled. It is ordered as follows:

1. Card sets A through G.
2. Card sets H and I. Data for numbers 2 through 4 below are printed sequentially for each data set ℓ ($\ell = 1, 2, \ldots, m$).
3. Dependent variable vector , $\underline{Y}^{o}_{s\ell}$, for sample information, computed using the linearized model.
4. Total sums of squares $(\underline{Y}_{\ell} - \hat{\underline{Y}})^{T}\underline{\omega}(\underline{Y}_{\ell} - \hat{\underline{Y}})$ and $(\underline{Y}^{o}_{\ell} - \hat{\underline{Y}})^{T}\underline{\omega}(\underline{Y}^{o}_{\ell} - \hat{\underline{Y}})$, where $\underline{Y}^{T}_{\ell} = [\underline{Y}^{T}_{s\ell}\ \underline{b}^{T}_{\ell}]$, $\underline{Y}^{oT}_{\ell} = [\underline{Y}^{oT}_{s\ell}\ \underline{b}^{T}_{\ell}]$, $\hat{\underline{Y}}^{T} = [\hat{\underline{Y}}^{T}_{s}\ \hat{\underline{b}}^{T}]$, and $\underline{\omega} = \begin{bmatrix} \underline{V}^{-1}_{s} & \underline{0} \\ \underline{0} & \underline{U}^{-1}s^{2} \end{bmatrix}$.
5. Beale's measure, $\hat{N}_{b} \equiv BN$.

Use of Version Integral with the Regression Ground-Water Program. This version consists of sets of statements to be inserted into the program of appendix 4.3.4, as indicated on the appended listing. Input is the same as if a regression solution were to be obtained, except that, in addition, card sets T through W are used after card set S to enter the data for the $\ell = 1, 2, \ldots, m$ alternate solutions, and a final card to input q with format I5 is added. Output consists of the regression output (with the exception that sensitivities are not printed or orthogonalized) plus output analogous to numbers 2 through 4 of the general version of the Beale's measure program.

Program Listing for General Version.

```
C********************  BEALE'S MEASURE PROGRAM  ************************
C
      DIMENSION BOPT(20),YOPT(100),B(20),YC(100),YL(100),S(20,100)
     1,W(100),IPR(20),WP(20)
      COMMON S,NVAR,NOBS
C
C          FORMAT LIST
C
    1 FORMAT (5I5,F10.0)
    2 FORMAT (8F10.0)
    3 FORMAT (9H1NVAR  = ,I4/9H NRES  = ,I4/9H NOBS  = ,I4
     1/9H NPRIR = ,I4/9H NPTS  = ,I4/9H VAR   = ,G11.5)
    4 FORMAT (1HO,26X,18HOPTIMUM PARAMETERS
     1/1H ,3X,3(3HNO.,9X,4HBOPT,8X))
    5 FORMAT (1HO,9X,52HDEPENDENT VARIABLES COMPUTED WITH OPTIMUM PARAME
     1TERS/1H ,3X,3(3HNO.,9X,4HYOPT,8X))
    6 FORMAT (1HO,21X,26HPARAMETERS FOR SAMPLE NO. ,I3
     1/1H ,3X,3(3HNO.,11X,1HB,9X))
    7 FORMAT (1HO,12X,44HDEPENDENT VARIABLES COMPUTED FOR SAMPLE NO. ,I3
     1/1H ,3X,3(3HNO.,10X,2HYC,9X))
    8 FORMAT (1HO,43HSCALED SENSITIVITIES FOR OPTIMUM PARAMETERS)
    9 FORMAT (1HO,6X,55HLINEARIZED DEPENDENT VARIABLES COMPUTED FOR SAMP
     1LE NO. ,I3/1H ,3X,3(3HNO.,10X,2HYL,9X))
   10 FORMAT (1HO,5HBN = ,G11.5)
   11 FORMAT (23HOSS((YC-YOPT)*W**.5) = ,G11.5
     1/23H SS((YL-YOPT)*W**.5) = ,G11.5)
   12 FORMAT (1HO,14X,42HRELIABILITY WEIGHTS FOR SAMPLE INFORMATION
     1/1H ,3X,3(3HNO.,10X,1HW,10X))
   13 FORMAT (16I5)
   14 FORMAT (1HO,12X,43HNO.S OF PARAMETERS HAVING PRIOR INFORMATION
     1/1H ,3X,3(3HNO.,8X,3HIPR,10X))
   15 FORMAT (1HO,14X,40HSTANDARD DEVIATIONS OF PRIOR INFORMATION
     1/1H ,3X,3(3HNO.,10X,2HWP,9X))
C
C          READ BASE DATA
C
      READ(5,1) NVAR,NRES,NOBS,NPRIR,NPTS,VAR                    CARD A
      WRITE(6,3) NVAR,NRES,NOBS,NPRIR,NPTS,VAR
      READ(5,2) (BOPT(J),J=1,NVAR)                               CARD B
      WRITE(6,4)
      CALL PRTOTB(BOPT,NVAR)
      READ(5,2) (YOPT(I),I=1,NOBS)                               CARD C
      WRITE(6,5)
      CALL PRTOTB(YOPT,NOBS)
      READ(5,2) (W(I),I=1,NOBS)                                  CARD D
      WRITE(6,12)
      CALL PRTOTB(W,NOBS)
      DO 20 J=1,NOBS
      READ(5,2) (S(I,J),I=1,NVAR)                                CARD E
   20 CONTINUE
      WRITE(6,8)
      CALL PRTOT
      DO 30 J=1,NOBS
```

```
      DO 25 I=1,NVAR
   25 S(I,J)=S(I,J)/BOPT(I)
   30 CONTINUE
      IF(NPRIR.LT.1) GO TO 45
      READ(5,13) (IPR(I),I=1,NPRIR)                                    CARD F
      WRITE(6,14)
      CALL PRTOTC(IPR,NPRIR)
      READ(5,2) (WP(I),I=1,NPRIR)                                      CARD G
      WRITE(6,15)
      CALL PRTOTB(WP,NPRIR)
      DO 40 I=1,NPRIR
   40 WP(I)=VAR/(WP(I)*WP(I))
C
C          READ DATA FOR EACH SAMPLE AND COMPUTE BEALE'S MEASURE, BN
C
   45 SUMA=0.
      SUMB=0.
      DO 80 M=1,NPTS
      READ(5,2) (B(J),J=1,NVAR)                                        CARD H
      WRITE(6,6) M
      CALL PRTOTB(B,NVAR)
      READ(5,2) (YC(I),I=1,NOBS)                                       CARD I
      WRITE(6,7) M
      CALL PRTOTB(YC,NOBS)
      SUMC=0.
      SUMD=0.
      DO 60 J=1,NOBS
      SUM=YOPT(J)
      DO 50 I=1,NVAR
   50 SUM=SUM+S(I,J)*(B(I)-BOPT(I))
      YL(J)=SUM
      TMP=YC(J)-SUM
      SUMA=SUMA+TMP*W(J)*TMP
      TMP=YC(J)-YOPT(J)
      SUMC=SUMC+TMP*W(J)*TMP
      TMP=SUM-YOPT(J)
      SUMD=SUMD+TMP*W(J)*TMP
   60 CONTINUE
      IF(NPRIR.LT.1) GO TO 75
      DO 70 J=1,NPRIR
      I=IPR(J)
      TMP=B(I)-BOPT(I)
      TMP=TMP*WP(J)*TMP
      SUMC=SUMC+TMP
   70 SUMD=SUMD+TMP
   75 WRITE(6,9) M
      CALL PRTOTB(YL,NOBS)
      WRITE(6,11) SUMC,SUMD
   80 SUMB=SUMB+SUMC*SUMC
      TMP=NRES
      BN=TMP*VAR*SUMA/SUMB
      WRITE(6,10) BN
      STOP
      END
      SUBROUTINE PRTOTB(VAL,NO)
```

```
C
C          PRINT OUT VALUES IN THREE GROUPS OF TWO COLUMNS
C
      DIMENSION VAL(NO)
      NR=NO/3
      IF(3*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(6,20) (L,VAL(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,3(I3,7X,G11.5,3X))
      END
      SUBROUTINE PRTOTC(IVAL,NO)
C
C          PRINT OUT INTEGERS IN THREE GROUPS OF TWO COLUMNS
C
      DIMENSION IVAL(NO)
      NR=NO/3
      IF(3*NR.NE.NO) NR=NR+1
      DO 10 K=1,NR
      WRITE(6,20) (L,IVAL(L),L=K,NO,NR)
   10 CONTINUE
      RETURN
   20 FORMAT (1H ,2X,3(I3,8X,I4,9X))
      END
      SUBROUTINE PRTOT
C
C          PRINT VALUES AS MATRICES DIVIDED VERTICALLY INTO TEN-COLUMN BLOCKS
C
      DIMENSION TMP(20,100)
      COMMON TMP,NR,NC
      DO 60 K=1,NC,10
      J10=K+9
      IF(J10.GT.NC) J10=NC
      WRITE(6,70) (J,J=K,J10)
      WRITE(6,90)
      DO 30 I=1,NR
   30 WRITE(6,80) I,(TMP(I,J),J=K,J10)
      WRITE(6,100)
   60 CONTINUE
   70 FORMAT(1H0,8X,I3,9(9X,I3))
   80 FORMAT (1H ,I3,10(1X,G11.5))
   90 FORMAT (1H )
  100 FORMAT (1H0)
      RETURN
      END
```

```
C
C**INSERT JUST BEFORE EQUIVALENCE STATEMENT FOR BEALE'S MEASURE
      DIMENSION BOPT(20),HOPT(500),HL(500)
      EQUIVALENCE (HO(1),HOPT(1)),(QRS(1),HL(1))
C
C
C**INSERT AFTER STATEMENT LABEL 520 FOR BEALE'S MEASURE
      IF(NBQP.LE.0) GO TO 1200
      DO 1100 J=1,NBQP
 1100 BOPT(J)=QBND(J)
 1200 IF(NBHZ.LE.0) GO TO 1400
      DO 1300 J=1,NBHZ
      K=IBNA(J+NQSD)
      L=IBPA(J)
      IF(L.GT.NBQP) BOPT(L)=HC(K)
      K=IBNB(J+NQSD)
      L=IBPB(J)
      IF(L.GT.NBQP) BOPT(L)=HC(K)
 1300 CONTINUE
 1400 DO 1500 J=1,NZNS
      K=IPRM(1,J)
      IF(K.GT.NBPAR) BOPT(K)=TRANX(J)
      K=IPRM(2,J)
      IF(K.GT.NBPAR) BOPT(K)=TRANY(J)
      K=IPRM(3,J)
      IF(K.GT.NBPAR) BOPT(K)=VLEAK(J)
      K=IPRM(4,J)
      IF(K.GT.NBPAR) BOPT(K)=QDIST(J)
 1500 CONTINUE
      J=0
      K=0
      DO 1700 N=1,NIJ
      IF(W(N).LT.1.E-10) GO TO 1700
      K=K+1
      HOPT(K)=HC(N)
      IF(IN(N).LT.1) GO TO 1700
      J=J+1
      DO 1600 I=1,NVAR
 1600 S(I,J)=S(I,J)/BOPT(I)
 1700 CONTINUE
      IF(NPRIR.LE.0) GO TO 1900
      DO 1800 I=1,NVAR
 1800 WP(I)=WP(I)/(BOPT(I)*BOPT(I))
 1900 SUMA=0.
      SUMB=0.
      GO TO 640
C
C
C**INSERT JUST BEFORE STATEMENT LABEL 690 FOR BEALE'S MEASURE
      IF(NBQP.LE.0) GO TO 2200
      DO 2100 J=1,NBQP
 2100 V(J)=QBND(J)
 2200 IF(NBHZ.LE.0) GO TO 2400
```

```
      DO 2300 J=1,NBHZ
      L=IBNA(J+NQSD)
      M=IBPA(J)
      IF(M.GT.NBQP) V(M)=HC(L)
      L=IBNB(J+NQSD)
      M=IBPB(J)
      IF(M.GT.NBQP) V(M)=HC(L)
 2300 CONTINUE
 2400 DO 2500 J=1,NZNS
      I=IPRM(1,J)
      IF(I.GT.NBPAR) V(I)=TRANX(J)
      I=IPRM(2,J)
      IF(I.GT.NBPAR) V(I)=TRANY(J)
      I=IPRM(3,J)
      IF(I.GT.NBPAR) V(I)=VLEAK(J)
      I=IPRM(4,J)
      IF(I.GT.NBPAR) V(I)=QDIST(J)
 2500 CONTINUE
      SUMC=0.
      SUMD=0.
      M=0
      L=0
      DO 2700 J=1,NIJ
      IF(W(J).LT.1.E-10) GO TO 2700
      M=M+1
      SUM=HC(J)
      IF(IN(J).LT.1) GO TO 2650
      L=L+1
      SUM=HOPT(M)
      DO 2600 I=1,NVAR
 2600 SUM=SUM+S(I,L)*(V(I)-BOPT(I))
 2650 HL(M)=SUM
      TMP=(HC(J)-SUM)*W(J)
      SUMA=SUMA+TMP*TMP
      TMP=(HC(J)-HOPT(M))*W(J)
      SUMC=SUMC+TMP*TMP
      TMP=(SUM-HOPT(M))*W(J)
      SUMD=SUMD+TMP*TMP
 2700 CONTINUE
      IF(NPRIR.LT.1) GO TO 2900
      DO 2800 J=1,NVAR
      IF(WP(J).LT.1.E-10) GO TO 2800
      TMP=V(J)-BOPT(J)
      TMP=TMP*WP(J)*TMP
      SUMC=SUMC+TMP
      SUMD=SUMD+TMP
 2800 CONTINUE
 2900 WRITE(IOUT,3000)
 3000 FORMAT (1H0,27X,16HLINEARIZED HEADS/1H ,4X,3(3HNO.,10X,2HHL,9X))
      CALL PRTOT(HL,NOBS,0,1)
      WRITE(IOUT,3100) SUMC,SUMD
 3100 FORMAT (23HOSS((HC-HOPT)*W**.5) = ,G11.5
     1/23H SS((HL-HOPT)*W**.5) = ,G11.5)
      SUMB=SUMB+SUMC*SUMC
C
```

```
C
C**INSERT AFTER STATEMENT LABEL 690 FOR BEALE'S MEASURE
      READ(IIN,800) NRES
      WRITE(IOUT,3200) NRES
 3200 FORMAT (8H0NRES = ,I4)
      TMP=NRES
      BN=TMP*VAR*SUMA/SUMB
      WRITE(IOUT,3300) BN
 3300 FORMAT (1H0,5HBN = ,G11.5)
C
```

References Cited

Beale, E. M. L., 1960, Confidence regions in nonlinear estimation: Journal of the Royal Statistical Society, Series B, v. 22, no. 1, p. 41-76.

Guttman, I., and Meeter, D. A., 1965, On Beale's measure of nonlinearity: Technometrics, v. 7, no. 4, p. 623-637.

Theil, H., 1963, On the use of incomplete prior information in regression analysis: American Statistical Association Journal, v. 58, no. 302, p. 401-414.

Additional Reading

Cooley, R. L., 1977, A method of estimating parameters and assessing reliability for models of steady-state groundwater flow, 1--Theory and numerical properties: Water Resources Research, v. 13, no. 2, p. 318-324.

_____1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1--Theory: Water Resources Research, v. 18, no. 4, p. 965-976.

Problem 2.2-1

a) (R R), (R W), (R B), (W W), (W R), (W B), (B B), (B W), (B R)

b)

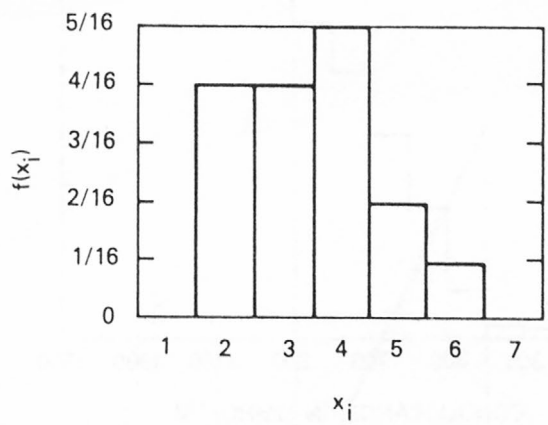
	R	W	B1	B2
R	RR	RW	RB1	RB2
W	WR	WW	WB1	WB2
B1	B1R	B1W	B1B1	B1B2
B2	B2R	B2W	B2B1	B2B2

P(RR) = 1/16 P(WW) = 1/16 P(BB) = 1/4

P(RW) = 1/16 P(WR) = 1/16 P(BW) = 1/8

P(RB) = 1/8 P(WB) = 1/8 P(BR) = 1/8

c)

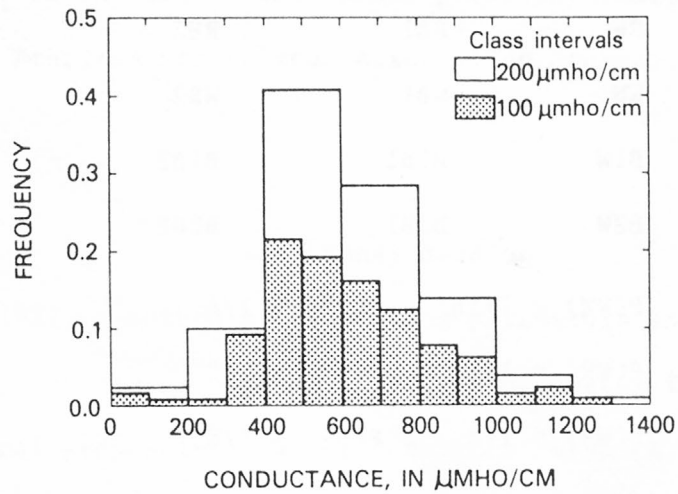


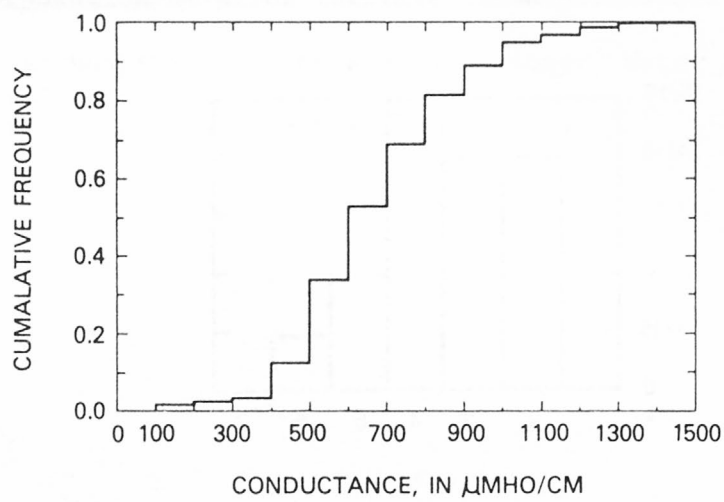
d) P(X = 4) = 5/16

(R R), (W B), (B W)

Problem 2.2-2

a) The effect of increasing the class interval size is to increase the frequency value for each class. This tends to make the histogram appear more peaked for larger intervals. Increasing the class interval also tends to smooth out irregularities.





b) $P(X \leq 600) = 0.53$

$P(X > 400) = 0.88$

$P(400 < X \leq 600) = 0.41$

$P(X \leq 1300) = 1.0$

Problem 2.3-1

a) $\bar{x} = \sum_{(\text{all } i)} \bar{x}_i f_i^* = (2 \times 50 + 1 \times 150 + 1 \times 250 + 12 \times 350$

$+ 28 \times 450 + 25 \times 550 + 21 \times 650 + 16 \times 750 + 10 \times 850$

$+ 8 \times 950 + 2 \times 1050 + 3 \times 1150 + 1 \times 1250)/130$

$= 612$

b) By replacing the integration in (2.3-3) with a summation, it is seen that the population mean μ_X for this random variable is represented by

$$\mu_X = \sum_{i=2}^{6} i f_i$$

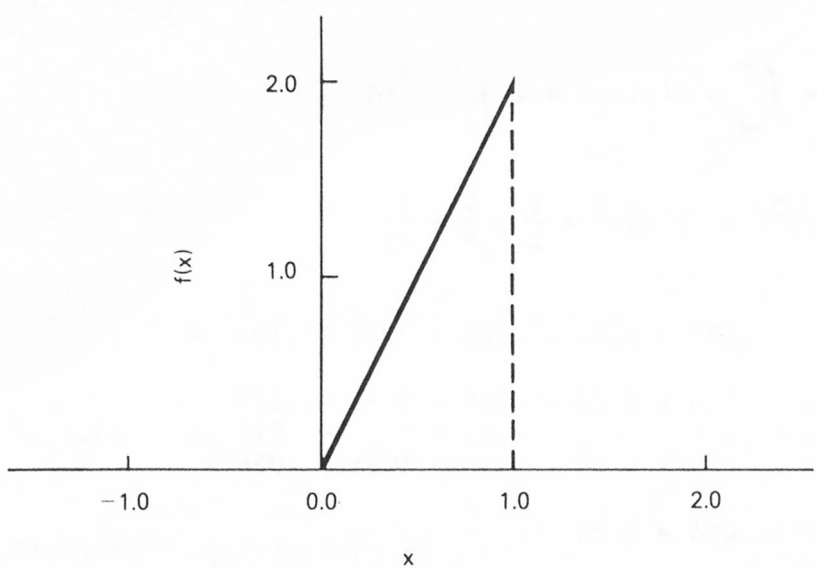
$$= (2 \times 1/4 + 3 \times 1/4 + 4 \times 5/16 + 5 \times 1/8 + 6 \times 1/16)$$

$$= 3\frac{1}{2}$$

Problem 2.3-2

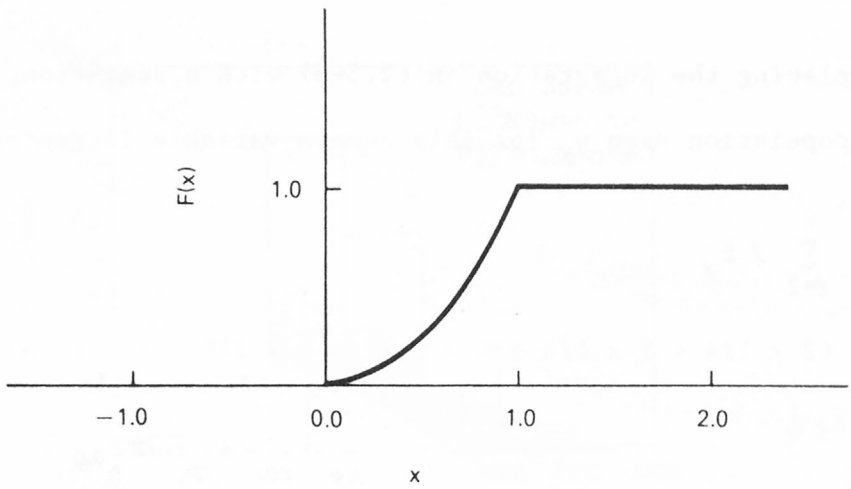
a) Triangular density problem.

(i)



(*ii*)

$$F(x) = \int_{-\infty}^{x} f(s)\,ds = \begin{cases} 1 & , x > 1 \\ 2\int_0^x s\,ds = x^2 & , 0 \le x \le 1 \\ 0 & , x < 0 \end{cases}$$



(*iii*)

$$E[X] = \int_{-\infty}^{\infty} s\,f(s)\,ds = 2\int_0^1 s^2\,ds = \frac{2}{3}$$

$$E[X^2] = \int_{-\infty}^{\infty} s^2 f(s)\,ds = 2\int_0^1 s^3\,ds = \frac{1}{2}$$

$$\sigma_X^2 = E[X^2] - (E[X])^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

b) $\bar{x} = 1.25 \times .051 + 1.75 \times .103 + 2.25 \times .103$
$+ 2.75 \times .154 + 3.25 \times .154 + 3.75 \times .128$
$+ 4.25 \times .128 + 4.75 \times .090 + 5.25 \times .026$
$+ 5.75 \times .064 = 3.36$

$$s_x^{*2} = \sum_{i=1}^{10} (\bar{x}_i - \bar{\bar{x}})^2 f_i^*$$

$$= (-2.11)^2 \times .051 + (-1.61)^2 \times .103 + (-1.11)^2 \times .103$$
$$+ (-.61)^2 \times .154 + (-.11)^2 \times .154 + (.39)^2 \times .128$$
$$+ (.89)^2 \times .128 + (1.39)^2 \times .090 + (1.89)^2 \times .026$$
$$+ (2.39)^2 \times .064 = 1.43$$

Problem 2.4-1

a) $P(X = 3 \text{ and } Y = 2) = P(X = 3)P(Y = 2) = 1/36$

b) $P(X + Y = 5) = 1/9$

c) $P(Y = 2 \mid X = 3) = 1/6$

d) $P(X + Y = 5 \mid X = 3) = 1/6$

e) $P(X + Y \leq 5 \mid X = 3) = 1/3$

Problem 2.4-2

a) $E[Y] = a_1\mu_{X_1} + a_2\mu_{X_2} + a_3\mu_{X_3}$

$$Var[Y] = E\{[Y - E(Y)]^2\} = E\{[a_1(X_1 - \mu_{X_1}) + a_2(X_2 - \mu_{X_2})$$
$$+ a_3(X_3 - \mu_{X_3})]^2\}$$
$$= E[a_1^2(X_1 - \mu_{X_1})^2 + a_2^2(X_2 - \mu_{X_2})^2 + a_3^2(X_3 - \mu_{X_3})^2$$
$$+ 2a_1a_2(X_1 - \mu_{X_1})(X_2 - \mu_{X_2}) + 2a_1a_3(X_1 - \mu_{X_1})(X_3 - \mu_{X_3})$$
$$+ 2a_2a_3(X_2 - \mu_{X_2})(X_3 - \mu_{X_3})]$$

$$= a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + a_3^2\sigma_{X_3}^2$$

$$+ 2a_1a_2\sigma_{X_1X_2} + 2a_1a_3\sigma_{X_1X_3} + 2a_2a_3\sigma_{X_2X_3}$$

since $\sigma_{X_i}^2 = E[(X_i - \mu_{X_i})^2]$ and $\sigma_{X_iX_j} = E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})]$.

b) $\underline{a}\ Var[\underline{X}]\ \underline{a}^T = \underline{a}\ E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T]\ \underline{a}^T$

$= \underline{a}\ E[\underline{X}\ \underline{X}^T - \underline{X}\ E[\underline{X}]^T - E[\underline{X}]\ \underline{X}^T + E[\underline{X}]E[\underline{X}]^T]\ \underline{a}^T$

$= \underline{a}\ (E[\underline{X}\ \underline{X}^T] - E[\underline{X}]E[\underline{X}]^T)\ \underline{a}^T$

$= E[\underline{a}\ \underline{X}\ \underline{X}^T\ \underline{a}^T] - E[\underline{a}\ \underline{X}]E[\underline{X}^T\ \underline{a}^T]$

$= Var[\underline{a}\ \underline{X}]$

c)

$$Var[\underline{Y}] = \begin{bmatrix} \sigma_{Y_1}^2 & \sigma_{Y_2Y_2} & \cdots & \sigma_{Y_1Y_p} \\ & \sigma_{Y_2}^2 & \cdots & \sigma_{Y_2Y_p} \\ \text{symmetry} & & & \vdots \\ & & & \sigma_{Y_p}^2 \end{bmatrix}$$

$$\sigma_{Y_i}^2 = Var[Y_i] = Var[\underline{a}_i\ \underline{X}] = \underline{a}_i\ Var[\underline{X}]\ \underline{a}_i^T$$

$$\sigma_{Y_iY_j} = E[Y_iY_j - E(Y_i)E(Y_j)]$$

$$= E[\underline{a}_i\ \underline{X}\ \underline{X}^T\ \underline{a}_j^T - E(\underline{a}_i\ \underline{X})E(\underline{X}^T\ \underline{a}_j^T)]$$

$$= \underline{a}_i\ E[\underline{X}\ \underline{X}^T - E(\underline{X})E(\underline{X}^T)]\ \underline{a}_j^T$$

$$= \underline{a}_i\ Var[\underline{X}]\ \underline{a}_j^T$$

$$Var[\underline{Y}] = \begin{bmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{a}_p \end{bmatrix} Var[\underline{X}] \; [\underline{a}_1^T \; \underline{a}_2^T \cdots \underline{a}_p^T] = \underline{A} \; Var[\underline{X}] \; \underline{A}^T$$

Problem 2.5-1

a) $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = 610$

The results differ very little, indicating the validity of approximation (2.3-1). The repeat values probably result from measurement errors, indicating that a value of the random variable could not be determined more precisely than one μmho/cm.

b) $\bar{x} = 3.36 \qquad s_X^2 = 1.43$

$$\bar{t}_g = 10^{\bar{x}} = 2290 \text{ gpd/ft}$$

$$d_g = 10^{(\bar{x} \pm s_X)} = \bar{t}_g \times 10^{\pm s_X} = (146, 35900) \text{ gpd/ft}$$

The large dispersion strongly suggests that $\bar{t}_g$ may not represent the true geometric mean, since there is significant scatter in T in the vicinity of $\bar{t}_g$.

Problem 2.5-2

Let X be the specific conductance random variable and Y be the dissolved solids random variable. Then

$\sum x_i = 14999$ $\qquad$ $\sum y_i = 9140$

$\sum x_i^2 = 11743379$ $\qquad$ $\sum y_i^2 = 4316966$

$\sum x_i y_i = 7095973$ $\qquad$ $n = 23$

$\bar{x} = 652$ $\qquad$ $\bar{y} = 397$

$s_X^2 = 89200$ $\qquad$ $s_Y^2 = 31300$

$s_X = 299$ $\qquad$ $s_Y = 176$

$$r_{XY} = 0.98$$

Problem 2.6-1

$\bar{x} = 3 \times 10^{-4}$ $\qquad$ $s_X^2 = 6.6 \times 10^{-5}$

$$P(\bar{X}^2/(S_X^2/n) \leq F_\alpha(1, n-1))$$

$$= P(-\sqrt{F_\alpha(1, n-1)} \leq \bar{X}/(S_X/\sqrt{n}) \leq \sqrt{F_\alpha(1, n-1)})$$

$$= P(-2.26 \leq \bar{X}/(S_X/\sqrt{n}) \leq 2.26) = 0.95$$

$$\bar{x}/(s_X/\sqrt{n}) = 0.12$$

The value of the statistic is well within the interval ($-2.26 < 0.12 < 2.26$). This result is expected, if the titration experiment is valid, since 95% of all values of the statistic $\bar{X}/(S_X/\sqrt{n})$, calculated from repeated random sampling, would be expected to fall in this interval. If the value had fallen outside the interval, one should feel uneasy, since this should occur only 5% of the time. One would then be obligated to question whether the assumption $\mu_X = 0$ inherent to the titration test is valid.

Problem 2.8-1

a) $\bar{x} = 10 \qquad s_X^2 = 0.08 \qquad n = 7$

$$P\left(\frac{(\bar{X} - \mu_X)^2}{S_X^2/n} \leq F_\alpha(1,6) \right) = P\left(-\sqrt{F_\alpha(1,6)} \leq \frac{\bar{X} - \mu_X}{S_X/\sqrt{n}} \leq \sqrt{F_\alpha(1,6)} \right)$$

Therefore an interval can be constructed from

$$-\sqrt{F_\alpha(1,6)} \leq \frac{\bar{x} - \mu_X}{s_X/\sqrt{n}} \leq \sqrt{F_\alpha(1,6)}$$

where $F_\alpha(1,6) = 5.99$. Thus

$$-s_X \sqrt{\frac{F_\alpha(1,6)}{n}} \leq (\bar{x} - \mu_X) \leq s_X \sqrt{\frac{F_\alpha(1,6)}{n}}$$

or

$$\bar{x} + s_X \sqrt{\frac{F_\alpha(1,6)}{n}} \geq \mu_X \geq \bar{x} - s_X \sqrt{\frac{F_\alpha(1,6)}{n}}$$

and

$$10.26 \geq \mu_X \geq 9.74$$

b) Since 95% of all intervals so constructed will contain μ_X, there is a 0.95 probability that this interval contains μ_X.

Problem 2.9-1

a) H_0: $\mu_X = 0$

H_1: $\mu_X \neq 0$

b) From equation (2.6-21), the statement P(reject H_0/H_0 true) = α becomes

$$P\left(\frac{S_1/\sigma_1^2}{S_2^2/\sigma_2^2} > a \;\middle|\; \frac{\sigma_1^2}{\sigma_2^2} = 1\right) = P(S_1^2/S_2^2 > c) = 0.05$$

where S_1^2/S_2^2 is an F(24,15) random variable. Since the critical region is defined by values of F(24,15) greater than or equal to c, c must be equal to $F_{.05}(24,15)$ (see equation 2.6-15)). Thus, c = 2.29 and, since $s_1^2/s_2^2 = 1.31$, we accept H_0 at 0.05 significance level.

c) Hypothesis to be tested for rejection:

$$H_0: \quad \mu_X = 9.8$$

Alternate hypothesis:

$$H_1: \quad \mu_X \neq 9.8$$

From equation (2.9-7),

$$P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} < -\sqrt{F_\alpha(\nu_1,\nu_2)}\right) + P\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}} > \sqrt{F_\alpha(\nu_1,\nu_2)}\right)$$

$$= P\left(\left(\frac{\bar{X} - \mu_0}{S_X/\sqrt{n}}\right)^2 > F_\alpha(1,\ n-1)\right) = \alpha$$

Thus, the critical value for the statistic $(\bar{X} - \mu_0)^2/(S_X^2/n)$ is $F_\alpha(1,\ n-1)$, which, at the 0.05 significance level, has a value of 5.99. Since the statistic for the random sample in question takes on the value 3.50, we are forced to accept the possibility that $\mu_X = 0$: We cannot safely reject H_0 at the 0.05 significance level.

Problem 3.1-1

Upon substitution of (4) into (3), one obtains

$$h = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

For n observations

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon}$$

where $\underline{Y} - \underline{h} = \underline{\varepsilon}$,

$$\underline{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_s} \end{bmatrix} \qquad \underline{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_{n_s} \end{bmatrix} \qquad \underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{n_s} \end{bmatrix}$$

and

$$\underline{X} = \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ \vdots & \vdots & \vdots \\ X_{n_s 1} & X_{n_s 2} & X_{n_s 3} \end{bmatrix}$$

In the regression model:

$\underline{Y}$ = observed dependent variable;

$\underline{h}$ = computed dependent variable;

s = independent variable (distance along stream tube);

$\underline{X}$ = sensitivities; and

$\underline{\beta}$ = parameters.

a) $S(\underline{b}) = [\underline{Y} - \underline{X}\,\underline{b}]^T[\underline{Y} - \underline{X}\,\underline{b}]$

where

$$\underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

That is,

$$S(\underline{b}) = \left[Y_1 - (b_1X_{11} + b_2X_{12} + b_3X_{13})\right]^2$$

$$+ \left[Y_2 - (b_1X_{21} + b_2X_{22} + b_3X_{23})\right]^2$$

$$+ \ldots$$

$$+ \left[Y_{n_s} - (b_1X_{n_s1} + b_2X_{n_s2} + b_3X_{n_s3})\right]^2$$

b) $S(\underline{b}) = [\underline{Y} - \underline{X}\,\underline{b}]^T \underline{\omega}\, [\underline{Y} - \underline{X}\,\underline{b}]$

where $\underline{\omega} = \underline{V}^{-1}$ and

$$\underline{V} = \begin{bmatrix} \sigma_1^2/\sigma^2 & & \underline{0} \\ & \sigma_2^2/\sigma^2 & \\ & & \ddots \\ \underline{0} & & \sigma_{n_s}^2/\sigma^2 \end{bmatrix}$$

so that

$$\underline{V}^{-1} = \begin{bmatrix} \sigma_1^{-2} & & \underline{0} \\ & \sigma_2^{-2} & \\ & & \ddots \\ \underline{0} & & \sigma_{n_s}^{-2} \end{bmatrix} \sigma^2$$

That is,

$$S(\underline{b}) = \sigma^2 \left\{ \left[\frac{1}{\sigma_1} (Y_1 - (b_1X_{11} + b_2X_{12} + b_3X_{13})) \right]^2 \right.$$

$$+ \left[\frac{1}{\sigma_2} (Y_2 - (b_1X_{21} + b_2X_{22} + b_3X_{23})) \right]^2$$

$$+ \ldots$$

$$\left. + \left[\frac{1}{\sigma_{n_s}} (Y_{n_s} - (b_1X_{n_s1} + b_2X_{n_s2} + b_3X_{n_s3})) \right]^2 \right\}$$

c) $S(\underline{b}) = [\underline{Y} - \underline{X}\,\underline{b}]^T \underline{\omega}\, [\underline{Y} - \underline{X}\,\underline{b}]$

where

$$\underline{Y} - \underline{X}\,\underline{b} = \begin{bmatrix} Y_1 - (b_1 X_{11} + b_2 X_{12} + b_3 X_{13}) \\ \vdots \\ Y_{n_s} - (b_1 X_{n_s 1} + b_2 X_{n_s 2} + b_3 X_{n_s 3}) \\ h_b - b_2 \end{bmatrix}$$

and $\underline{\omega} = \underline{V}^{-1}$, so that

$$\underline{V} = \begin{bmatrix} 1 & & \underline{0} \\ & \ddots & \\ & & 1 \\ \underline{0} & & \sigma^2_{h_b}/\sigma^2 \end{bmatrix}$$

and

$$\underline{V}^{-1} = \begin{bmatrix} 1 & & \underline{0} \\ & \ddots & \\ & & 1 \\ \underline{0} & & \sigma^2/\sigma^2_{h_b} \end{bmatrix}$$

That is,

$$S(b) = \left[Y_1 - (b_1 X_{11} + b_2 X_{12} + b_3 X_{13})\right]^2$$

$$+ \ldots$$

$$+ \left[Y_{n_s} - (b_1 X_{n_s 1} + b_2 X_{n_s 2} + b_3 X_{n_s 3})\right]^2$$

$$+ \sigma^2\left[h_b - b_2\right]^2/\sigma^2_{h_b}$$

Problem 3.2-1

a) Note that if $\underline{\omega}$ is diagonal

$$\underline{X}^T \underline{\omega}\, \underline{X} = \left\{ \sum_{i=1}^{n_s} X_{ik}\, \omega_{ii}\, X_{ij} + \delta_{k2}\, \delta_{2j}\, \sigma^2/\sigma^2_{h_b} \right\} = \left\{ \sum_{i=1}^{n_s+1} X_{ik}\, \omega_{ii}\, X_{ij} \right\}$$

and

$$\underline{X}^T \underline{\omega}\ \underline{Y} = \left\{ \sum_{i=1}^{n_s} X_{ik} \omega_{ii} Y_i + \delta_{k2} h_b \sigma^2 / \sigma^2_{h_b} \right\} = \left\{ \sum_{i=1}^{n_s+1} X_{ik} \omega_{ii} Y_i \right\}$$

where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\underline{X}^T = \begin{bmatrix} X_{11} & X_{21} & & X_{n_s 1} & 0 \\ X_{12} & X_{22} & \cdots & X_{n_s 2} & 1 \\ X_{13} & X_{23} & & X_{n_s 3} & 0 \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & & X_{n_s 1} & X_{n_s+1,1} \\ X_{12} & X_{22} & \cdots & X_{n_s 2} & X_{n_s+1,2} \\ X_{13} & X_{23} & & X_{n_s 3} & X_{n_s+1,3} \end{bmatrix}$$

$$\underline{\omega} = \begin{bmatrix} 1 & & & \underline{0} \\ & \ddots & & \\ & & 1 & \\ \underline{0} & & & \sigma^2/\sigma^2_{h_b} \end{bmatrix} = \begin{bmatrix} \omega_{11} & & & \underline{0} \\ & \ddots & & \\ & & \omega_{n_s n_s} & \\ \underline{0} & & & \omega_{n_s+1,n_s+1} \end{bmatrix}$$

and

$$\underline{Y} = \begin{bmatrix} \bar{Y}_1 \\ Y_2 \\ \cdots \\ Y_{n_s} \\ h_b \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_{n_s} \\ Y_{n_s+1} \end{bmatrix}$$

Thus, the normal equations

$$\underline{X}^T \underline{\omega}\ \underline{X}\ \hat{\underline{b}} = \underline{X}^T \underline{\omega}\ \underline{Y}$$

can be written

$$\sum_{j=1}^{p} \sum_{i=1}^{n_s+1} X_{ik} \omega_{ii} X_{ij}\, \hat{b}_j = \sum_{i=1}^{n_s+1} X_{ik} \omega_{ii} Y_i, \qquad k = 1, 2, \ldots, p$$

or

$$\sum_{i=1}^{n_s+1} \left(\sum_{j=1}^{p} X_{ik} \omega_{ii} X_{ij} \hat{b}_j \right) = \sum_{i=1}^{n_s+1} X_{ik} \omega_{ii} Y_i, \qquad k = 1, 2, \ldots, p$$

As a comparison, the normal equations may be derived without using matrix techniques. Let $n = n_s + 1$ so that

$$S(\underline{b}) = \sum_{i=1}^{n} e_i^2 \omega_{ii}$$

The model equations are given as

$$Y_i = \sum_{j=1}^{p} X_{ij} b_j + e_i, \qquad i = 1, 2, \ldots, n$$

Therefore

$$S(\underline{b}) = \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_{ii}$$

Take the derivative with respect to any parameter b_k ($k = 1, 2, \ldots, p$):

$$\frac{\partial S}{\partial b_k} = \frac{\partial}{\partial b_k} \left[\sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_{ii} \right]$$

Because the derivative of a sum is the sum of derivatives, look at one term:

$$\frac{\partial}{\partial b_k} \left[\left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right)^2 \omega_{ii} \right]$$

$$= -2 \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right) \omega_{ii} X_{ik}$$

Thus,

$$\frac{\partial S}{\partial b_k} = -2 \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} b_j \right) \omega_{ii} X_{ik}$$

Set $\frac{\partial S}{\partial b_k} = 0$ to find the minimum so that

$$\sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} X_{ij} \hat{b}_j \right) \omega_{ii} X_{ik} = 0, \qquad k = 1, 2, \ldots, p$$

or

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{p} X_{ik} \omega_{ii} X_{ij} \hat{b}_j \right) = \sum_{i=1}^{n} X_{ik} \omega_{ii} Y_i, \qquad k = 1, 2, \ldots, p$$

b) Consider matrix $\underline{A}$ such that

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Assume $\underline{A}$ is symmetric; its cofactor matrix $\underline{A}_c$ is

$$\underline{A}_c = \begin{bmatrix} (a_{22}a_{33} - a_{23}^2) & (a_{13}a_{23} - a_{12}a_{33}) & (a_{12}a_{23} - a_{13}a_{22}) \\ (a_{13}a_{23} - a_{12}a_{33}) & (a_{11}a_{33} - a_{13}^2) & (a_{12}a_{13} - a_{11}a_{23}) \\ (a_{12}a_{23} - a_{13}a_{22}) & (a_{12}a_{13} - a_{11}a_{23}) & (a_{11}a_{22} - a_{12}^2) \end{bmatrix}$$

The determinant of $\underline{A}$ is

$$|\underline{A}| = a_{11}(a_{22}a_{33} - a_{23}^2) + a_{12}(a_{13}a_{23} - a_{12}a_{33}) + a_{13}(a_{12}a_{23} - a_{13}a_{22})$$

The inverse of $\underline{A}$ is

$$\underline{A}^{-1} = \underline{A}_c^T / |\underline{A}|$$

Let $\underline{A} = \underline{X}^T \underline{\omega}\, \underline{X}$ where

$$\underline{\omega} = \begin{bmatrix} 1 & & & \underline{0} \\ & \ddots & & \\ & & 1 & \\ \underline{0} & & & \sigma^2/\sigma_{h_b}^2 \end{bmatrix}$$

Thus, by adding the prior information to the information in table 3, there results

	Data Set 1	Data Set 2
$\sum_j X_{j2}^2 \omega_{jj}$	3.53161157	3.12700831
$\sum_j X_{j2}\, \omega_{jj} Y_j$	127.0992273	122.5185789

and, for data set 1,

$$\underline{X}^T \underline{\omega}\, \underline{X} = \begin{bmatrix} 3.3250 & 1.6750 & 418750 \\ & 3.53161157 & 418750 \\ \text{symmetric} & & 83340625000 \end{bmatrix}$$

or, for data set 2,

$$\underline{X}^T \underline{\omega}\, \underline{X} = \begin{bmatrix} 2.8500 & 1.6500 & 412500 \\ & 3.12700831 & 412500 \\ \text{symmetric} & & 8.3325\times10^{10} \end{bmatrix}$$

Then, for data set 1,

$$\underline{A}_c = \begin{bmatrix} 1.18975153\times10^{11} & 3.575601563\times10^{10} & -777456.0949 \\ & 1.017560156\times10^{11} & -690937.5 \\ \text{symmetric} & & 8.93698347 \end{bmatrix}$$

and

$$|\underline{X}^T \underline{\omega}\, \underline{X}| = 1.299239702\times10^{11}$$

Therefore

$$[\underline{X}^T \underline{\omega}\, \underline{X}]^{-1} = \begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169\times10^{-6} \\ & 0.7831966299 & -5.318014058\times10^{-6} \\ \text{symmetric} & & 6.878625596\times10^{-11} \end{bmatrix}$$

Similarly, for data set 2

$$\underline{A}_c = \begin{bmatrix} 9.040171743\times10^{10} & 3.267\times10^{10} & -609265.9279 \\ & 6.732\times10^{10} & -495000 \\ \text{symmetric} & & 6.189473684 \end{bmatrix}$$

and

$$|\underline{X}^T \underline{\omega}\ \underline{X}| = 6.022819942\times10^{10}$$

Therefore

$$[\underline{X}^T \underline{\omega}\ \underline{X}]^{-1} = \begin{bmatrix} 1.500986553 & 0.5424369368 & -1.011595787\times10^{-5} \\ & 1.117748839 & -8.218741466\times10^{-6} \\ \text{symmetric} & & 1.027670384\times10^{-10} \end{bmatrix}$$

c) By adding the prior information to the information in table 3, we obtain for $\underline{X}^T \underline{\omega}\ \underline{Y}$:

	Data Set 1	Data Set 2
$\sum_j X_{j1}\ \omega_{jj} Y_j$	192.18350	168.2030
$\sum_j X_{j2} \omega_{jj} Y_j$	127.0992273	122.5185789
$\sum_j X_{j3} \omega_{jj} Y_j$	26879687.5	26583550

By evaluating $(\underline{X}\ \omega\ \underline{X})^{-1}\ \underline{X}^T \omega\ \underline{Y}$, estimates $\hat{\underline{b}}$ are obtained as

	Data Set 1	Data Set 2
$\hat{b}_1$	50.12043881	50.01097198
$\hat{b}_2$	9.487418691	9.701194703
$\hat{b}_3$	0.00002302475114	0.00002342971729

Problem 3.3-1

The answers to parts a, b, and c are found in the section at the end of the problem where aids in debugging the computer code are given. The authors' computer code and output follow:

```
      DIMENSION T(10),S(10),F(10),Z(2,10),C(2,2),G(2),D(2)
   10 FORMAT (I5,6F10.0,I5)
   20 FORMAT (1H1,35HNO. OF OBSERVATIONS (N) -------- = ,I7
     $/1H ,35HPUMPING RATE (Q) --------------- = ,G11.5
     $/1H ,35HDISTANCE FROM WELL CENTER (R) -- = ,G11.5
     $/1H ,35HINITIAL TRANSMISSIVITY (T0) ---- = ,G11.5
     $/1H ,35HINITIAL STORAGE COEFFICIENT (S0) = ,G11.5
     $/1H ,35HDAMPING PARAMETER (AP) --------- = ,G11.5
     $/1H ,35HCLOSURE CRITERION (ER) --------- = ,G11.5
     $/1H ,35HMAXIMUM NO. OF ITERATIONS (ITMX) = ,I7)
   30 FORMAT (8F10.0)
   32 FORMAT (10HOTIMES (T))
   34 FORMAT (23HOOBSERVED DRAWDOWNS (S))
   36 FORMAT (15HOITERATION NO. ,I4
     $/36H CURRENT ESTIMATES OF PARAMETERS (B))
   38 FORMAT (28HOSOLUTION FAILED TO CONVERGE)
   40 FORMAT ((1H ,10(G11.5,2X)))
   42 FORMAT (19HOSOLUTION CONVERGED)
   44 FORMAT (29HOFINAL COMPUTED DRAWDOWNS (F))
   46 FORMAT (16HORESIDUALS (F-S))
   48 FORMAT (18HOERROR VARIANCE = ,G11.5)
   50 FORMAT (12HOVAR(T)    = ,G11.5/12H COV(T,S) = ,G11.5
     $/12H VAR(S)    = ,G11.5)
   52 FORMAT (26HOSCALED SENSITIVITIES (Z):/5H TO T)
   54 FORMAT (5H TO S)
C          READ AND PRINT INPUT DATA
      READ(5,10) N,Q,R,T0,S0,AP,ER,ITMX
      WRITE(6,20) N,Q,R,T0,S0,AP,ER,ITMX
      READ(5,30) (T(I),I=1,N)
      WRITE(6,32)
      WRITE(6,40) (T(I),I=1,N)
      READ(5,30) (S(I),I=1,N)
      WRITE(6,34)
      WRITE(6,40) (S(I),I=1,N)
      FU=R*R/4.
      FW=Q/12.5664
      DMAX=ER+1.
      DO 140 KNT=1,ITMX
      UTMP=FU*S0/T0
      WTMP=FW/T0
      DO 80 I=1,N
C          COMPUTE NEW DRAWDOWN (F(I))
C            COMPUTE U AND W(U) FIRST
      U=UTMP/T(I)
      WU=W(U)
```

```
C           THEN F(I)
      F(I)=WTMP*WU
C           COMPUTE SCALED SENSITIVITIES (Z(I,J))
      TMP=EXP(-U)
      Z(1,I)=WTMP*(TMP-WU)
   80 Z(2,I)=-WTMP*TMP
C           CHECK FOR CONVERGENCE
      IF(DMAX.LT.ER) GO TO 150
C           ASSEMBLE COEFFICIENT MATRIX (C(I,J)) AND
C           GRADIENT VECTOR (G(J))
      DO 86 J=1,2
      DO 84 I=J,2
      C(I,J)=0.
   84 C(J,I)=0.
   86 G(J)=0.
      DO 110 K=1,N
      TMP=S(K)-F(K)
      DO 100 J=1,2
      DO 90 I=J,2
      C(I,J)=Z(I,K)*Z(J,K)+C(I,J)
   90 C(J,I)=C(I,J)
  100 G(J)=Z(J,K)*TMP+G(J)
  110 CONTINUE
C           INVERT COEFFICIENT MATRIX
      DET=C(1,1)*C(2,2)-C(1,2)*C(2,1)
      TMP=C(1,1)
      C(1,1)=C(2,2)/DET
      C(2,2)=TMP/DET
      C(1,2)=-C(1,2)/DET
      C(2,1)=C(1,2)
C           COMPUTE PARAMETER DISPLACEMENTS (D(I)), MAX. DISPLACEMENT
C           (DMAX), AND PARAMETERS (T0 AND S0)
      DMAX=0.
      DO 130 J=1,2
      D(J)=C(1,J)*G(1)+C(2,J)*G(2)
      TMP=ABS(D(J))
      IF(TMP.GT.DMAX) DMAX=TMP
  130 CONTINUE
      T0=T0*(1.+AP*D(1))
      S0=S0*(1.+AP*D(2))
C           PRINT PARAMETERS
      WRITE(6,36) KNT
      WRITE(6,40) T0,S0
  140 CONTINUE
      WRITE(6,38)
      GO TO 160
  150 WRITE(6,42)
C           PRINT DRAWDOWNS, RESIDUALS (F(I)-S(I)), AND SCALED
C           SENSITIVITIES
  160 WRITE(6,44)
      WRITE(6,40) (F(I),I=1,N)
      DO 170 I=1,N
  170 F(I)=F(I)-S(I)
      WRITE(6,46)
```

```
      WRITE(6,40) (F(I),I=1,N)
      WRITE(6,52)
      WRITE(6,40) (Z(1,I),I=1,N)
      WRITE(6,54)
      WRITE(6,40) (Z(2,I),I=1,N)
C          COMPUTE AND PRINT ERROR VARIANCE (VAR) AND
C          COVARIANCE MATRIX FOR PARAMETERS
      VAR=0.
      DO 180 I=1,N
  180 VAR=VAR+F(I)*F(I)
      VAR=VAR/(N-2.)
      WRITE(6,48) VAR
      C(1,1)=T0*C(1,1)*T0*VAR
      C(1,2)=T0*C(1,2)*S0*VAR
      C(2,2)=S0*C(2,2)*S0*VAR
      WRITE(6,50) C(1,1),C(1,2),C(2,2)
      STOP
      END
      FUNCTION W(X)
C          COMPUTE THE WELL FUNCTION OF X
      W=0.
      IF(X.GT.10.) GO TO 20
      W=-0.577216-ALOG(X)+X
      TERM=X
      DO 10 J=2,36
      RJ=J
      TERM=-TERM*X/RJ
      TMP=TERM/RJ
      W=W+TMP
      IF(ABS(TMP).LT.1.E-7) GO TO 20
   10 CONTINUE
   20 RETURN
      END
```

```
1NO. OF OBSERVATIONS (N) -------- =       7
 PUMPING RATE (Q) --------------- =  1.1600
 DISTANCE FROM WELL CENTER (R) -- =  175.00
 INITIAL TRANSMISSIVITY (TO) ---- =  .10000
 INITIAL STORAGE COEFFICIENT (SO) =  .50000E-03
 DAMPING PARAMETER (AP) --------- =  1.0000
 CLOSURE CRITERION (ER) --------- =  .10000E-01
 MAXIMUM NO. OF ITERATIONS (ITMX) =      10
OTIMES (T)
  480.00       1020.0       1500.0       2040.0       2700.0       3720.0       4920.0
OOBSERVED DRAWDOWNS (S)
  1.7100       2.2300       2.5400       2.7700       3.0400       3.2500       3.5600
OITERATION NO.    1
 CURRENT ESTIMATES OF PARAMETERS (B)
  .11188       .54748E-03
OITERATION NO.    2
 CURRENT ESTIMATES OF PARAMETERS (B)
  .11347       .55219E-03
OITERATION NO.    3
 CURRENT ESTIMATES OF PARAMETERS (B)
  .11349       .55221E-03
OSOLUTION CONVERGED
OFINAL COMPUTED DRAWDOWNS (F)
  1.6715       2.2521       2.5564       2.8012       3.0256       3.2832       3.5086
ORESIDUALS (F-S)
 -.38538E-01   .22079E-01   .16407E-01   .31217E-01  -.14393E-01   .33213E-01  -.51356E-01
OSCALED SENSITIVITIES (Z):
 TO T
 -.91882      -1.4679      -1.7630      -2.0026      -2.2234      -2.4779      -2.7014
 TO S
 -.75264      -.78421      -.79343      -.79866      -.80223      -.80527      -.80724
OERROR VARIANCE =  .14328E-02
OVAR(T)    =  .95030E-05
 COV(T,S) = -.11369E-06
 VAR(S)    =  .14595E-08
```

Problem 3.3-2

b)

$$\underline{D} = \begin{bmatrix}
2T_1 & -T_1 & & -T_1 & & & & & \\
-T_1 & 2(T_1+T_2) & -T_2 & & -(T_1+T_2) & & & & \\
 & -T_2 & 2T_2 & & & -T_2 & & & \\
-T_1 & & & 4T_1 & -2T_1 & & & & \\
 & -(T_1+T_2) & & -2T_1 & 4(T_1+T_2) & -2T_2 & -(T_1+T_2) & & \\
 & & -T_2 & & -2T_2 & 4T_2 & & & -T_2 \\
 & & & & & & 1 & & \\
 & & & & -(T_1+T_2) & & & 2(T_1+T_2) & -T_2 \\
 & & & & & -T_2 & & -T_2 & 2T_2
\end{bmatrix}$$

$$\underline{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix}$$

$$\underline{q} = \begin{bmatrix} \frac{1}{2}a^2W_1 \\ \frac{1}{2}a^2(W_1 + W_2) \\ \frac{1}{2}a^2W_2 \\ a^2W_1 + T_1h_{B1} \\ a^2(W_1 + W_2) \\ a^2W_2 + aq_{B1} \\ h_{B1} \\ \frac{1}{2}a^2(W_1 + W_2) + T_1h_{B1} \\ \frac{1}{2}a^2W_2 + aq_{B1} \end{bmatrix}$$

c)

$$\underline{J}_1 = \begin{bmatrix} h_2 + h_4 - 2h_1 \\ h_1 + h_5 - 2h_2 \\ 0 \\ h_1 + 2h_5 - 4h_4 + h_{B1} \\ h_2 + 2h_4 + h_8 - 4h_5 \\ 0 \\ 0 \\ h_5 - 2h_8 + h_{B1} \\ 0 \end{bmatrix}$$

$$\underline{J}_2 = \begin{bmatrix} 0 \\ h_3 + h_5 - 2h_2 \\ h_2 + h_6 - 2h_3 \\ 0 \\ h_2 + 2h_6 + h_8 - 4h_5 \\ h_3 + 2h_5 + h_9 - 4h_6 \\ 0 \\ h_5 + h_9 - 2h_8 \\ h_6 + h_8 - 2h_9 \end{bmatrix}$$

$$\underline{J}_3 = \begin{bmatrix} \frac{1}{2}a^2 \\ \frac{1}{2}a^2 \\ 0 \\ a^2 \\ a^2 \\ 0 \\ 0 \\ \frac{1}{2}a^2 \\ 0 \end{bmatrix}$$

$$\underline{J}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ a \\ 0 \\ 0 \\ a \end{bmatrix}$$

d)

$$\underline{J}_1 T_1 = \begin{bmatrix} T_1(h_2 - h_1) + T_1(h_4 - h_1) \\ T_1(h_1 - h_2) + T_1(h_5 - h_2) \\ 0 \\ T_1(h_1 - h_4) + 2T_1(h_5 - h_4) + T_1(h_{B1} - h_4) \\ T_1(h_2 - h_5) + 2T_1(h_4 - h_5) + T_1(h_8 - h_5) \\ 0 \\ 0 \\ T_1(h_5 - h_8) + T_1(h_{B1} - h_8) \\ 0 \end{bmatrix}$$

$$\underline{J}_2 T_2 = \begin{bmatrix} 0 \\ T_2(h_3 - h_2) + T_2(h_5 - h_2) \\ T_2(h_2 - h_3) + T_2(h_6 - h_3) \\ 0 \\ T_2(h_2 - h_5) + 2T_2(h_6 - h_5) + T_2(h_8 - h_5) \\ T_2(h_3 - h_6) + 2T_2(h_5 - h_6) + T_2(h_9 - h_6) \\ 0 \\ T_2(h_5 - h_8) + T_2(h_9 - h_8) \\ T_2(h_6 - h_9) + T_2(h_8 - h_9) \end{bmatrix}$$

$$\underline{J}_3 W_1 = \begin{bmatrix} \frac{1}{2}a^2 W_1 \\ \frac{1}{2}a^2 W_1 \\ 0 \\ a^2 W_1 \\ a^2 W_1 \\ 0 \\ 0 \\ \frac{1}{2}a^2 W_1 \\ 0 \end{bmatrix} \qquad \underline{J}_4 q_{B1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ aq_{B1} \\ 0 \\ 0 \\ aq_{B1} \end{bmatrix}$$

e) 1) Let $r = 0$.

2) Compute $\underline{D}_r$ and $\underline{q}_r$.

3) $\underline{h}_r = \underline{D}_r^{-1}\ \underline{q}_r$

4) Obtain $\underline{f}(\underline{\xi}, \underline{b}_r)$ from $\underline{h}_r$ by deleting node 7.

5)

$$\left(\frac{\partial \underline{h}}{\partial b_j}\right)_r b_j^r = \underline{D}_r^{-1}\left[\left(\frac{\partial \underline{q}}{\partial b_j}\right)_r - \left(\frac{\partial \underline{D}}{\partial b_j}\right)_r \underline{h}_r\right] b_j^r \Bigg\} j = 1, 2, \ldots, p$$

Obtain $\underline{Z}_j^r$ from $\left(\frac{\partial \underline{h}}{\partial b_j}\right)_r b_j^r$ by deleting node 7.

6) Compute $\underline{Z}_r^T\ \underline{w}\ \underline{Z}_r$ and $\underline{Z}_r^T\ \underline{w}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_r))$.

7) Define $\underline{C} = \{C_{ii}\} = \{(\underline{Z}_r^T \underline{w}\ \underline{Z}_r)_{ii}^{-\frac{1}{2}}\}$.

8) Compute $\underline{S}_r^T \underline{w}\ \underline{S}_r$ and $\underline{S}_r^T \underline{w}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_r))$

where $\underline{S}_r = \underline{Z}_r \underline{C}$.

9) $\underline{\delta}_{r+1} = (\underline{S}_r^T \underline{w}\ \underline{S}_r + \mu \underline{I})^{-1} \underline{S}_r^T \underline{w}(\underline{Y} - \underline{f}(\underline{\xi}, \underline{b}_r))$

10) $\underline{d}_{r+1} = \underline{C}\ \underline{\delta}_{r+1}$

11) $\underline{b}_{r+1} = \rho \underline{B}_r \underline{d}_{r+1} + \underline{b}_r$

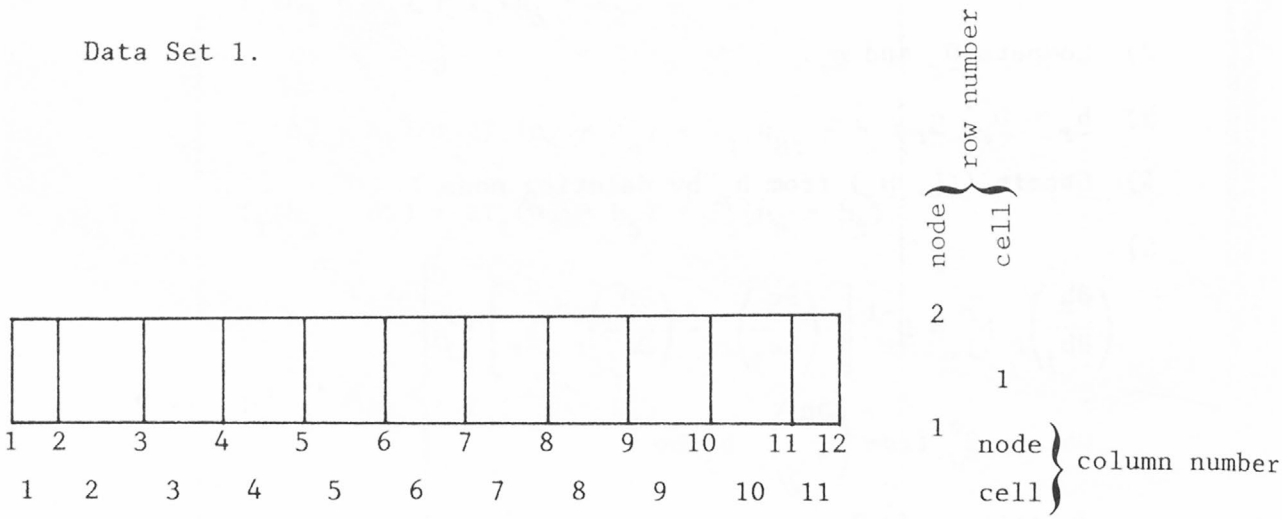
12) If $|d_i^{r+1}| > \varepsilon$ for all i=1, 2, ..., p, then increment r by one and return to 2. If not then:

13) $\underline{h}_{final} = \underline{D}_{r+1}^{-1} \underline{q}_{r+1}$

Values of μ and ρ can be computed at each iteration by using algorithms defined by (3.3-29) through (3.3-31) if desired.

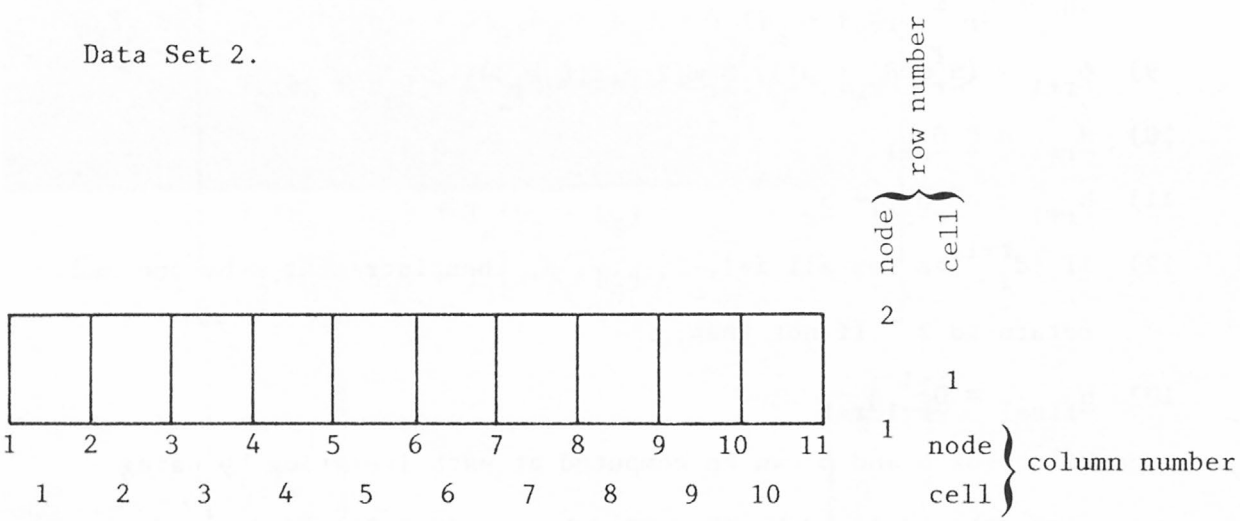
Problem 4.2-1

Data Set 1.



Nodes in columns 2 through 11 have observations. Nodes in columns 1 and 12 form specified head boundaries. Spacing: Cell row 1, 1 foot; cell columns 1 and 11, 50 feet; cell columns 2 through 10, 100 feet.

Data Set 2.



Nodes in columns 2 through 10 have observations. Nodes in columns 1 and 11 form specified head boundaries. Spacing: Cell row 1, 1 foot; cell columns 1 through 10, 100 feet.

If both T and W were estimated, the problem would be singular because the only unique parameter is W/T.

PROGRAMMER _____ DIVISION _____
LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

NUMBER _____
PROJECT _____

SHEET 1 OF 2

IDENTIFICATION: LAKE ØHPUPU, DATA SET 1. FØR PRØBLEM 4.2-1
INFORMATION: _____

```
LAKE ØHPUPU, DATA SET 1.
 12   2   1   1   0   0   0   2   2   2   1   1   0
     50       0       0       0       .25
  1   0
  1  11   1   1       1       1
     50     100     100     100     100     100     100
   100     100      50
  1   0
  1   1   1   1       1
  1   0
  2  11   1   1       1       1
    48.33   45.76   42.08   38.34   35.30   31.00   25.85   21.76
    16.11   12.48
  1   0
  2  11   1   1       1
  1   0
  1  11   1   1       1
  1   0
  1  11   1   1       0
  1   0
  1  12   1   2       0
  1   0
  1  11   1   1       1
  1   0
  1  11   1   1   1
  1       1       1       0       .001
  1   0   0   0   1
  1   0
```

FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [|] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—86265 p-1

PROGRAMMER _____ DIVISION _____

LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER _____

PROJECT _____

SHEET 2 OF 2

IDENTIFICATION: _____

INFORMATION: _____

```
2  0
1  1  1  2  -1
12 12 1  2  -1
1  2  1  1     0        0
1  1        49
1  2        49
2  2  2  2        .1       .1
12 1        11
12 2        11
```

FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [|] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—86263 p-1

410

PROGRAMMER ______ DIVISION ______
LOCATION ______ PHONE ______

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ______
PROJECT ______

SHEET 1 OF 2

IDENTIFICATION: LAKE ØHPUPU, DATA SET 2. FØR PRØBLEM 4.2-1

INFORMATION: ______

```

  LAKE ØHPUPU, DATA SET 2.

   11    2    1    1    0    0    0    2    2    2    1    1    0
        50         0         0         0       .25
    1    0
    1   10    1    1       100
    1    0
    1    1    1    1         1
    1    0
    2   10    1    1         1    1
     47.13     44.14     39.89     36.36     32.48     29.70     24.33     19.10
     14.96
    1    0
    2   10    1    1         1
    1    0
    1   10    1    1         1
    1    0
    1   10    1    1         0
    1    0
    1   11    1    2         0
    1    0
    1   10    1    1         1
    1    0
    1   10    1    1    1
    1         1         1         0      .001
    1    0    0    0    1
         0
    2    0
    1    1    1    2  -1
```

FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Ƶ = ALPHA Z / = SLASH ‖ = VERT. BAR - = MINUS ▭ = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

PROGRAMMER ______ DIVISION ______

LOCATION ______ PHONE ______

NUMBER ______

PROJECT ______

SHEET 2 OF 2

IDENTIFICATION: ______

INFORMATION: ______

```
         1         2         3         4
1234567890123456789012345678901234567890
   11   11    1    2   -1
    1    2    1    1         0         0
    1    1        49
    1    2        49
    2    2    2    2        .1        .1
   11    1       9.5
   11    2       9.5
```

FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Ƶ = ALPHA Z / = SLASH | = VERT. BAR − = MINUS ‾ = HORZ. BAR

DATA CODING FORM

GPO 18—88265 p-1

Problem 4.2-2

Because prior information is available, all parameters can be estimated. Data points are located in areas of high sensitivity for all parameters. More data points in areas of highest sensitivity might improve results for parameters having low sensitivity.

PROGRAMMER ______ DIVISION ______

LOCATION ______ PHONE ______

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ______

PROJECT ______

SHEET 1 OF 5

IDENTIFICATION: SEVERAL ZØNE, 2-D FLØW DATA SET FØR PRØBLEM 4.2-2

INFORMATION: ______

```
          CLASS PRØBLEM
          TWØ-DIMENSIØNAL FLØW
          SEVERAL ZØNES
   15   16    4    7    0    4    4    2    3   10    1    1    0
         2       .08         0         0         1
    1    0
    1   14    1    1         1    1
      1000      1000      1000      1000      1000      1000      1000       400
      1000      1000      1000      1000      1000      1000
    1    0
    1   15    1    1         1    1
      1000      1000      1000      1000      1000       400      1000      1000
      1000      1000      1000      1000      1000      1000      1000
    2    0
    6   15    2    8         1    1
         0         0     60.70         0         0         0         0         0
     75.64         0
         0         0         0         0         0         0     60.27         0
         0         0
         0         0         0         0     29.67         0         0         0
         0         0
         0      4.22         0         0         0      4.37         0      6.07
         0      5.81
         0         0         0         0         0         0         0         0
         0         0
         0         0         0         0      4.57         0         0         0
         0         0
         0         0      5.21         0         0         0    -44.89         0
         0      7.01
    1   15    9   16         1    1
```

FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Z = ALPHA Z / = SLASH | = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM
GPO 18—88265 p-1

415

PROGRAMMER ______ DIVISION ______
LOCATION ______ PHONE ______

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ______
PROJECT ______

SHEET 2 OF 5

IDENTIFICATION:
INFORMATION:

Line	Cols 1–10	Cols 11–20	Cols 21–30	Cols 31–40	Cols 41–50	Cols 51–60	Cols 61–70	Cols 71–80
1	0	0	0	0	0	0	6.95	0
2	0	0	0	0	0	0	0	
3	0	0	0	12.21	0	0	0	0
4	4.04	0	-89.36	0	0	0	0	
5	0	0	0	0	0	0	6.68	0
6	0	0	0	0	-15.32	0	0	
7	0	0	16.88	0	15.87	0	0	0
8	4.48	0	-18.34	0	0	0	0	
9	0	0	0	0	0	0	0	0
10	0	0	0	0	-2.47	0	8.10	
11	0	0	54.12	0	38.27	0	0	0
12	0	.053	0	-2.92	0	0	0	
13	0	0	0	0	0	0	8.30	0
14	0	0	0	0	0	4.54	0	
15	0	85.82	0	0	0	0	0	0
16	0	0	2.26					
17	2 0							
18	6 15	2 8	1	1				
19	0	0	1	0	0	0	0	0
20	1	0						
21	0	0	0	0	0	0	1	0
22	0	0						
23	0	0	0	0	1	0	0	0
24	0	0						
25	0	1	0	0	0	1	0	1
26	0	1						
27	0	0	0	0	0	0	0	0
28	0	0						
29	0	0	0	0	1	0	0	0
30	0	0						

FORM 9—1633 A
JULY 1967

0 = ZERO   Ø = ALPHA O   1 = ONE   I = ALPHA I   2 = TWO   Z = ALPHA Z   / = SLASH | = VERT. BAR   - = MINUS   _ = HORZ. BAR

DATA CODING FORM
GPO 16—88265 p-1

PROGRAMMER ____ DIVISION ____

LOCATION ____ PHONE ____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ____

PROJECT ____

SHEET 3 OF 5

IDENTIFICATION: ____

INFORMATION: ____

Line	Col. 5	Col. 9–10	Col. 15	Col. 19–20	Col. 28–30	Col. 35	Col. 40	Col. 50	Col. 60	Col. 70	Col. 80
1		0		0	1		0	0	0	1	0
2		0		1							
3	1	15	9	16	1	1					
4		0		0	0		0	0	0	1	0
5		0		0	0		0	0	0	0	
6		0		0	0		1	0	0	0	0
7		1		0	1		0	0	0	0	
8		0		0	0		0	0	0	1	0
9		0		0	0		0	1	0	0	
10		0		0	1		0	1	0	0	0
11		1		0	1		0	0	0	0	
12		0		0	0		0	0	0	0	0
13		0		0	0		0	1	0	1	
14		0		0	1		0	1	0	0	0
15		0		1	0		1	0	0	0	
16		0		0	0		0	0	0	1	0
17		0		0	0		0	0	1	0	
18		0		1	0		0	0	0	0	0
19		0		0	1		0	0	0	0	
20	2	0									
21	6	14	1	8	1	0					
22	1	14	9	15	1	0					
23	2	0									
24	8	14	6	6	1	0					
25	8	8	7	15	1	0					
26	2	0									
27	8	15	6	7	4.5	0					
28	8	9	7	16	4.5	0					
29	2	0									
30	6	14	1	8	1	0					

FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Ƶ = ALPHA Z / = SLASH | = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM

GPO 18—88265 p-1

PROGRAMMER ____________ DIVISION ____________

LOCATION ____________ PHONE ____________

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER ____________

PROJECT ____________

SHEET 4 OF 5

IDENTIFICATION: ____________

INFORMATION: ____________


```
         1111111111222222222233333333334444444444555555555566666666667777777778
1234567890123456789012345678901234567890123456789012345678901234567890123456789 0
    1   14    9   15         1    0
    7    0
    6   14    1    8    2    0
    1   14    9   15    2    0
    1    6   12   15    1    0
    6    7    1    3    3    0
    8   14    1    4    3    0
    8   14    6    6    4    0
    8    8    7   15    4    0
    1        70        70         0     .0004
    2       420       420         0    -.0002
    3        15        15         0    .00017
    4       420       420       .08         0
    1    1    1    0    2
    2    3    3    0    4
    3    5    5    0    6
    4    3    3    7    0
         0        .3         2         0         0        .3        .1
    1   16    7   16    1        45         0       .01
    7   16   15   16    2       420         0    .00056
   11   10   11   10    3    -97000       .02         1
   12    8   12    8    4    -51000       .02         1
    1    0
   15   15    5   16   -1    0
    1   10    1    2         1        .1
   15   16      10.4
   15   15
   15   14
   15   13
   15   12
```

FORM 9—1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Ƶ = ALPHA Z / = SLASH | = VERT. BAR – = MINUS _ = HORZ. BAR

DATA CODING FORM
GPO 16—86265 p-1

PROGRAMMER _____ DIVISION _____

LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER _____

PROJECT _____

SHEET 5 OF 5

IDENTIFICATION: _____

INFORMATION: _____

```
15  11
15  10
15   9
15   8
15   7      4.8
 2   2   2      3      .1          .1
15   6      4.8
15   5      5.4
```

FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [|] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1

Problem 5.4-1

a) Let $\underline{b}_0 = \underline{0}$. Then

$$s^2 = \frac{\underline{Y}^T \underline{\omega}\ \underline{Y} - \hat{\underline{b}}^T \underline{X}^T \underline{\omega}\ \underline{Y}}{n - p}$$

$$\underline{Y} = \begin{bmatrix} 48.33 \\ 45.76 \\ 42.08 \\ 38.34 \\ 35.30 \\ 31.00 \\ 25.85 \\ 21.76 \\ 16.11 \\ 12.48 \\ 11 \end{bmatrix} \quad \text{for data set 1, or} \quad \begin{bmatrix} 47.13 \\ 44.14 \\ 39.89 \\ 36.36 \\ 32.48 \\ 29.70 \\ 24.33 \\ 19.10 \\ 14.96 \\ 9.5 \end{bmatrix} \quad \text{for data set 2}$$

$$\underline{\omega} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & & \frac{0.25}{(1.1)^2} \end{bmatrix} \text{ or } \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & & \frac{0.25}{(0.95)^2} \end{bmatrix}$$

$$\hat{\underline{b}} = \begin{bmatrix} 50.12043881 \\ 9.487418691 \\ 0.00002302475114 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 50.01097198 \\ 9.701194703 \\ 0.00002342971729 \end{bmatrix}$$

$$\underline{X}^T \underline{\omega}\ \underline{Y} = \begin{bmatrix} 192.18350 \\ 127.0992273 \\ 26879687.5 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 168.2030 \\ 122.5185789 \\ 26583550 \end{bmatrix}$$

$\underline{Y}^T \underline{\omega}\ \underline{Y} = 11459.5411$ or 10225.4391

$\hat{\underline{b}}^T \underline{X}^T \underline{\omega}\ \underline{Y} = 11457.06305$ or 10223.41717

$s^2 = 0.30975625$ or 0.2888471429

$s/\Delta Y_s = 0.5565574993/(48.33 - 12.48) = 0.0155$ or

$= 0.5374450138/(47.13 - 14.96) = 0.0167$

The fit is fairly good.

b)

$$(\underline{X}^T \underline{\omega}\ \underline{X})^{-1} s^2 =$$

$$0.30975625 \begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169 \times 10^{-6} \\ & 0.7831966299 & -5.318014058 \times 10^{-6} \\ \text{symmetric} & & 6.878625596 \times 10^{-11} \end{bmatrix}$$

$$= \begin{bmatrix} 0.2836527946 & 0.08524715876 & -1.853560079 \times 10^{-6} \\ & 0.2426000511 & -1.647288092 \times 10^{-6} \\ \text{symmetric} & & 2.130697270 \times 10^{-11} \end{bmatrix}$$

or

$$0.2888471429 \begin{bmatrix} 1.500986553 & 0.5424369368 & -1.011595787 \times 10^{-5} \\ 1 & 1.117748839 & -8.218741466 \times 10^{-6} \\ \text{symmetric} & & 1.027670384 \times 10^{-10} \end{bmatrix}$$

$$= \begin{bmatrix} 0.4335556774 & 0.1566813594 & 2.921965528 \times 10^{-6} \\ & 0.3228585586 & 2.373959991 \times 10^{-6} \\ \text{symmetric} & & 2.968396543 \times 10^{-11} \end{bmatrix}$$

Because the standard errors (the square roots of the diagonal elements) are small, the parameters are well determined.

c)

$$\underline{r} = \begin{bmatrix} 1 & 0.3249682531 & -0.7539668635 \\ & 1 & -0.7245414573 \\ \text{symmetric} & & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} 1 & 0.4187827075 & -0.8145011290 \\ & 1 & -0.7668426954 \\ \text{symmetric} & & 1 \end{bmatrix}$$

The problem is fairly well conditioned.

Problem 5.5-1

The plots of $\underline{d}$ and $\underline{g}$ suggest that correlation manifests itself in two ways: The variability of $\underline{g}$ from set to set is smaller than the variability of $\underline{d}$ from set to set, and the plots of $\underline{g}$ do not appear to be linear on normal probability paper. The plot of $\hat{\underline{e}}$ does not appear to differ very much from either the plots of $\underline{d}$ or $\underline{g}$, although the plot of $\hat{\underline{e}}$ does show the same nonlinear trend as displayed by the plots of $\underline{g}$. Therefore, one may say that the plot of $\hat{\underline{e}}$ does not appear to differ significantly from the plot of a $N(\underline{0},(\underline{I} - \underline{R})s^2)$ random variable. Furthermore, this distribution of residuals suggests that the Theis model is adequate to describe the observed-drawdown data set.

PROGRAMMER	DIVISION	U. S. DEPARTMENT OF THE INTERIOR GEOLOGICAL SURVEY **COMPUTER CODING FORM**	NUMBER	SHEET
LOCATION	PHONE		PROJECT	1 OF 1

IDENTIFICATION: INPUT TØ RESIDUALS ANALYSIS PRØGRAM FØR PRØBLEM 5.5-1

INFORMATION: THEIS DATA SET FRØM PRØBLEM 3.3-1

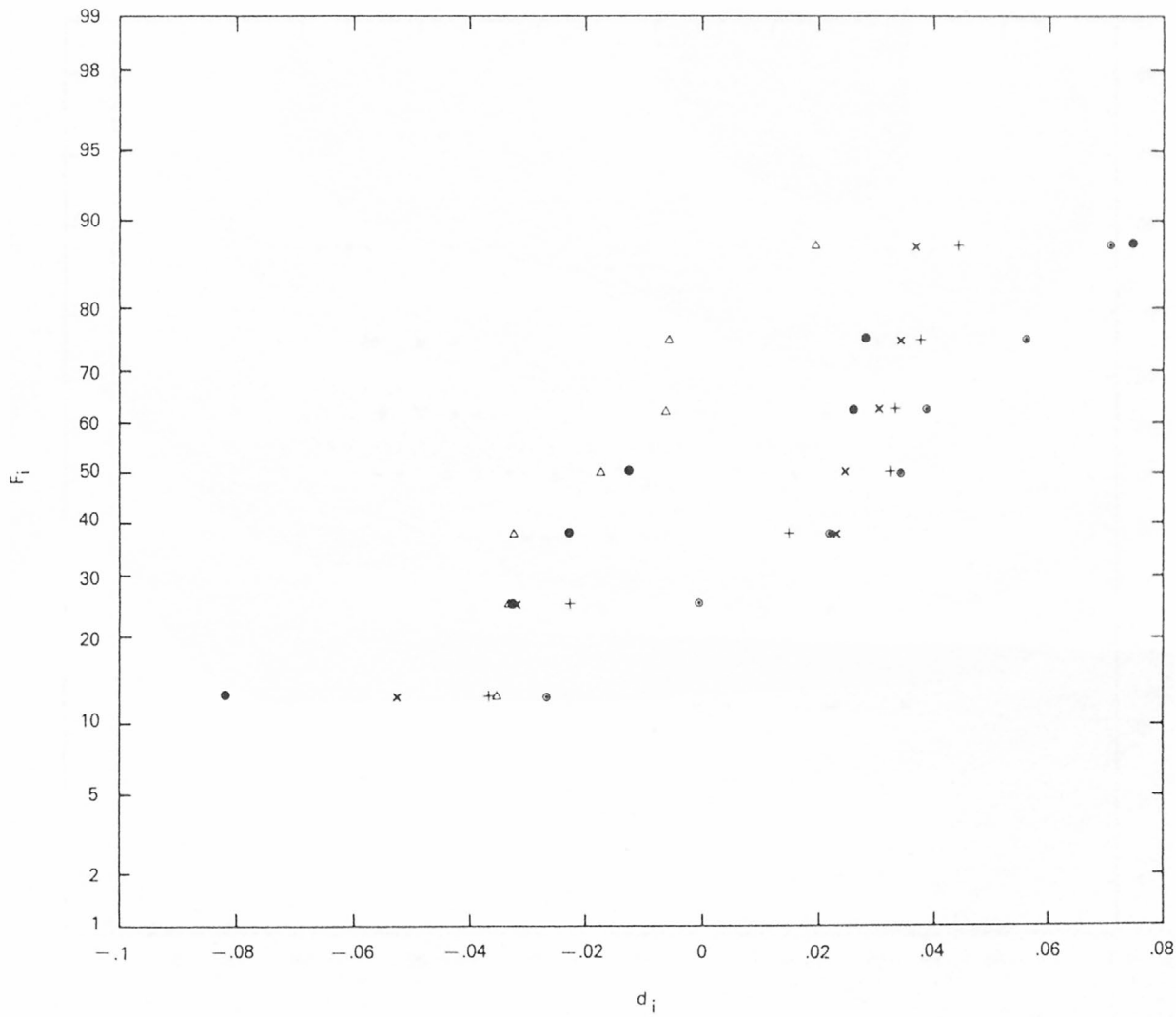
```
    2   7    0    5    1   .0014328
 .00073781 -.0018141
  .0047863
         1         1         1         1         1         1         1
   -.91882   -.75264
   -1.4679   -.78421
   -1.7630   -.79343
   -2.0026   -.79866
   -2.2234   -.80223
   -2.4779   -.80527
   -2.7014   -.80724
```

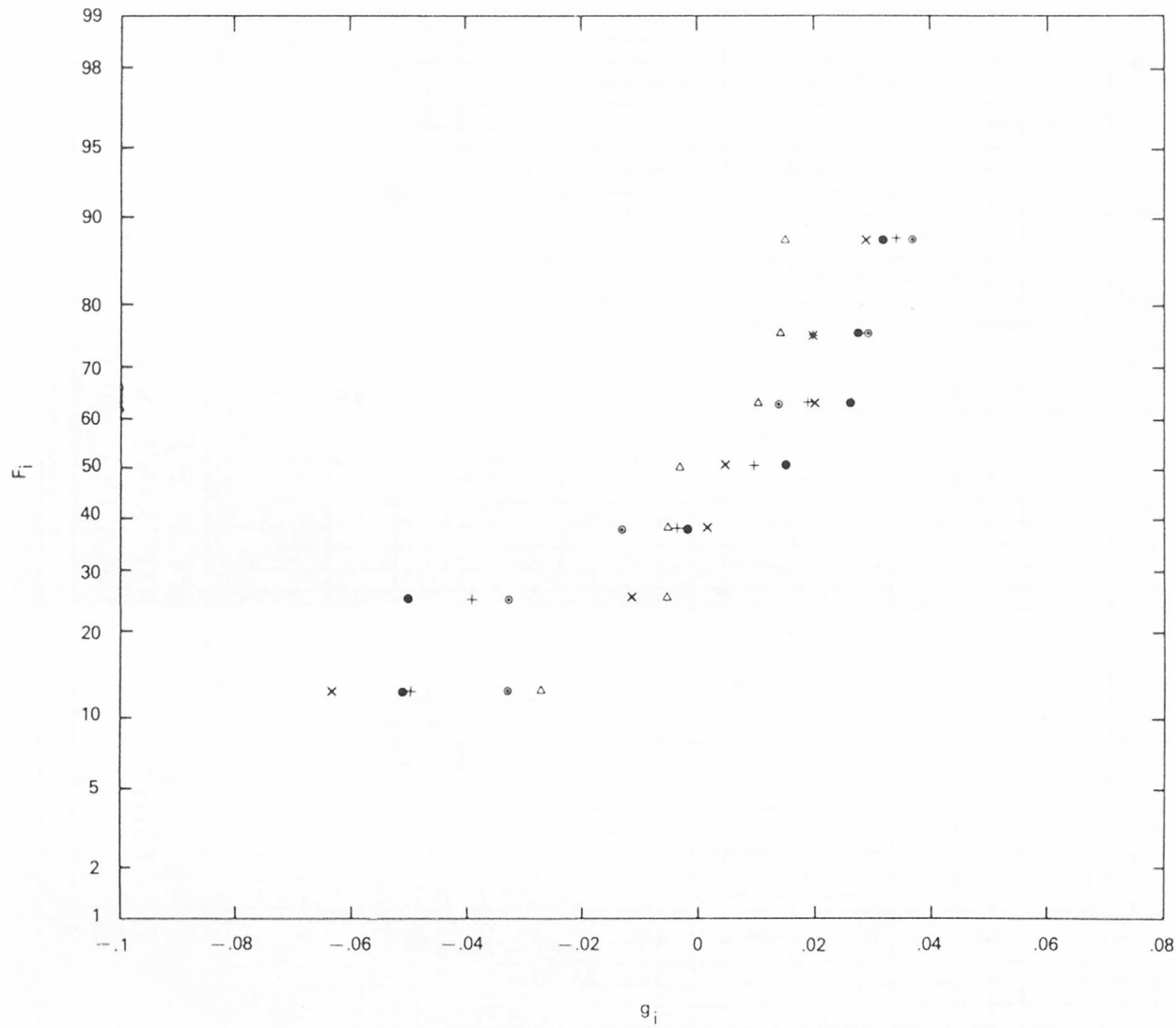
FORM 9—1633 A
JULY 1967

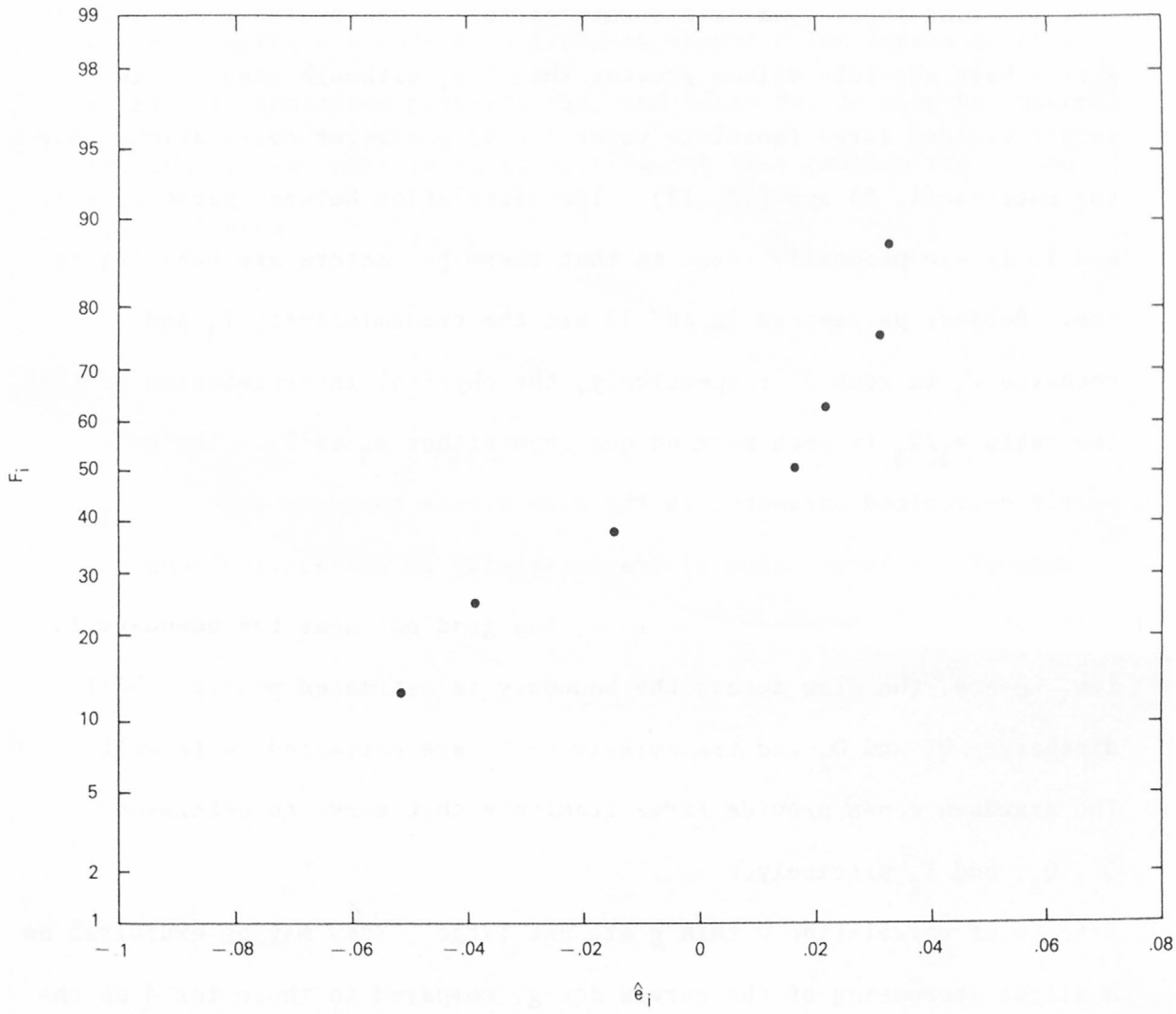
0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Z = ALPHA Z / = SLASH | = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1


99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
−.1
−.08
−.06
−.04
−.02
0
.02
.04
.06
.08
d_i


99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
−.1
−.08
−.06
−.04
−.02
0
.02
.04
.06
.08
g_i


99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
−.1
−.08
−.06
−.04
−.02
0
.02
.04
.06
.08
$\hat{e}_i$

Problem 5.5-2

a) $s^2 = 0.98677$

$R_y = 0.99964$

$s/\Delta Y_s = 0.99336297/175.18 = 0.00567$

b) The problem is not well conditioned. Entries (1, 8), (1, 9), (8, 9), and (12, 13) (and their symmetric counterparts) of the scaled least squares matrix have absolute values greater than 0.9, although these large values yielded large (absolute value > 0.9) parameter correlations only for entries (1, 8) and (12, 13). The correlation between parameters 12 and 13 is exceptionally large so that these parameters are behaving as one. Because parameters 12 and 13 are the transmissivity T_3 and recharge W_3 in zone 3, respectively, the physical interpretation is that the ratio W_3/T_3 is much more unique than either W_3 or T_3. The only poorly determined parameter is the flow across boundary zone 2, q_{B2}. Because of the large value of transmissivity in the aquifer zone adjacent to this boundary flow zone, the gradient near the boundary is low. Hence, the flow across the boundary is estimated poorly. Well discharges Q_1 and Q_2 and transmissivity T_2 are estimated quite well. The drawdown cones provide large gradients that serve to determine Q_1, Q_2, and T_2 precisely.

c) Effects of correlation within $\underline{g}$ are not large. They may be exhibited as a slight steepening of the curves for $\underline{g}$, compared to those for $\underline{d}$ on the normal probability plots. The plot of $\hat{\underline{f}}$ is very similar to those of $\underline{g}$, which suggests that the distribution of $\hat{\underline{f}}$ does not differ significantly from a $N(0,(I - R)s^2)$ distribution. Note that $\hat{\underline{e}}$ could not have been used instead of $\hat{\underline{f}}$ to make the comparison because the weight matrix $\underline{w}$ is

not equal to $\underline{I}$. The plot of $\hat{f}_j$ versus $\hat{Y}_j$ shows no pattern. However, the plot of $\hat{e}_j$ versus Cartesian coordinate shows a group of negative residuals in the upper center of the area. This sign pattern was inherited from the sign pattern of the original errors $\underline{\varepsilon}$ that were generated (recall that this exercise is based on a hypothetical problem), and the original errors are random N(0,1) deviates. Hence, the sign pattern occurred entirely by chance. The lesson is that apparently nonrandom patterns can, and often do, develop by chance, and the analyst must learn to distinguish true problem areas from apparent ones.

PROGRAMMER _____ DIVISION _____
LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

NUMBER _____
PROJECT _____

SHEET 1 OF 1

IDENTIFICATION: INPUT TØ RESIDUALS ANALYSIS PRØGRAM FØR PRØBLEM 5.5-2
INFORMATION: DATA SET FRØM TWØ-DIMENSIØNAL FLØW MØDEL ØF PRØBLEM 4.2-2

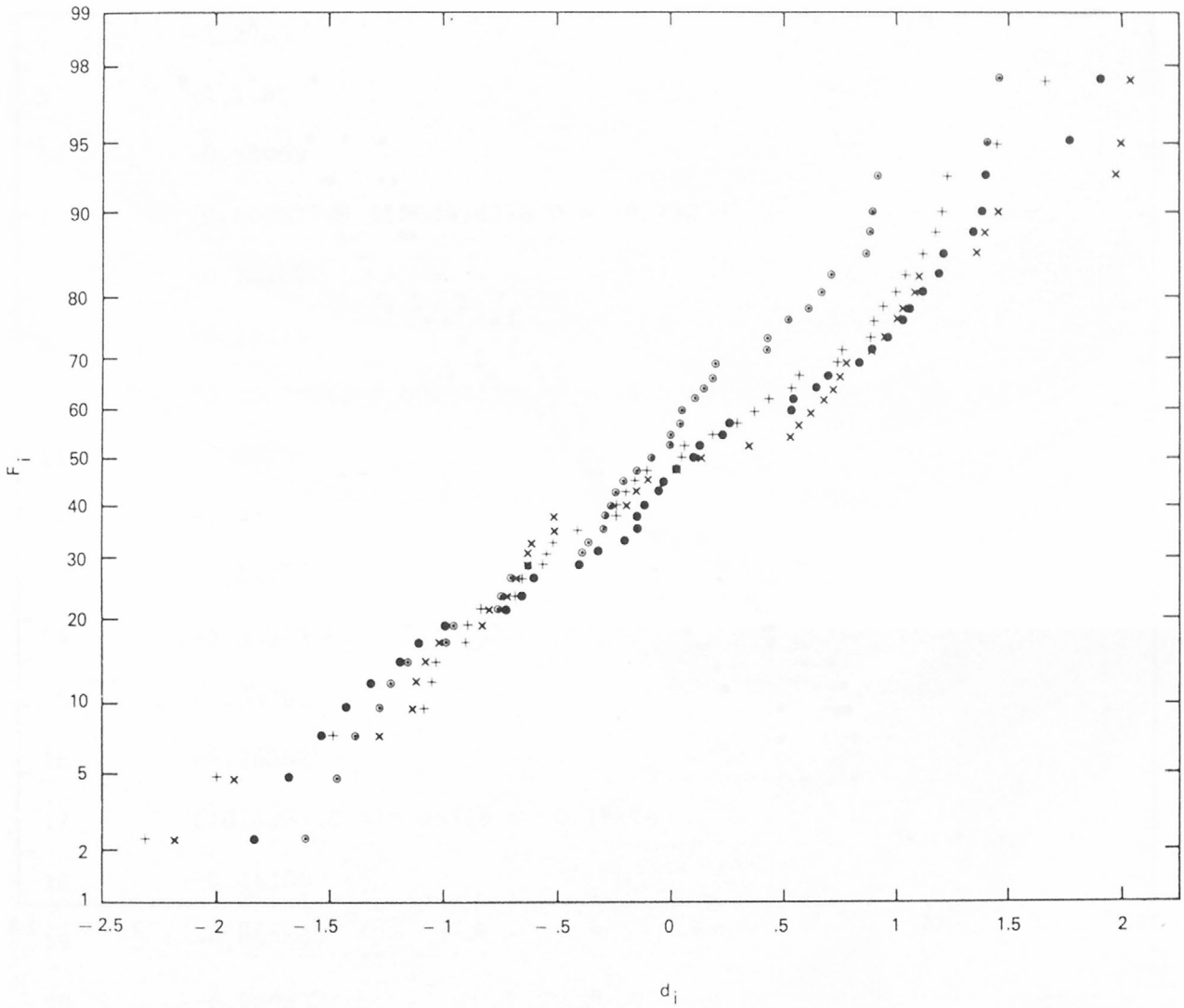
```
 14   32   9    4   437   .98677
 3    4    5    6    7    9   10   13   14
    1940      1020      1.04      .48      .54      .00012      84      .000051
    .008
   -97005     -50962     10.197     5.1209     5.4730  .00031134  487.73  .00013503
 .080718
```

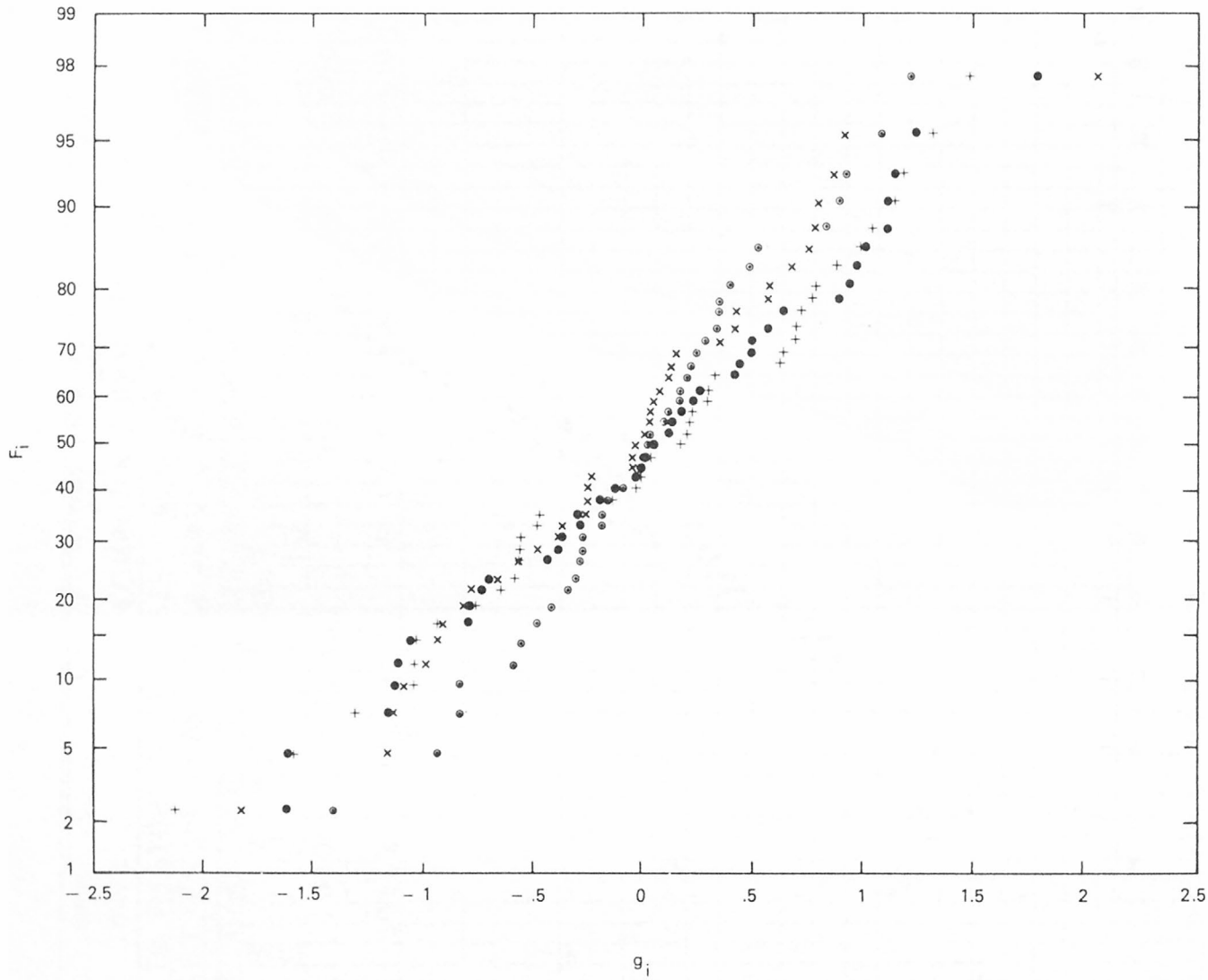
FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [|] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1

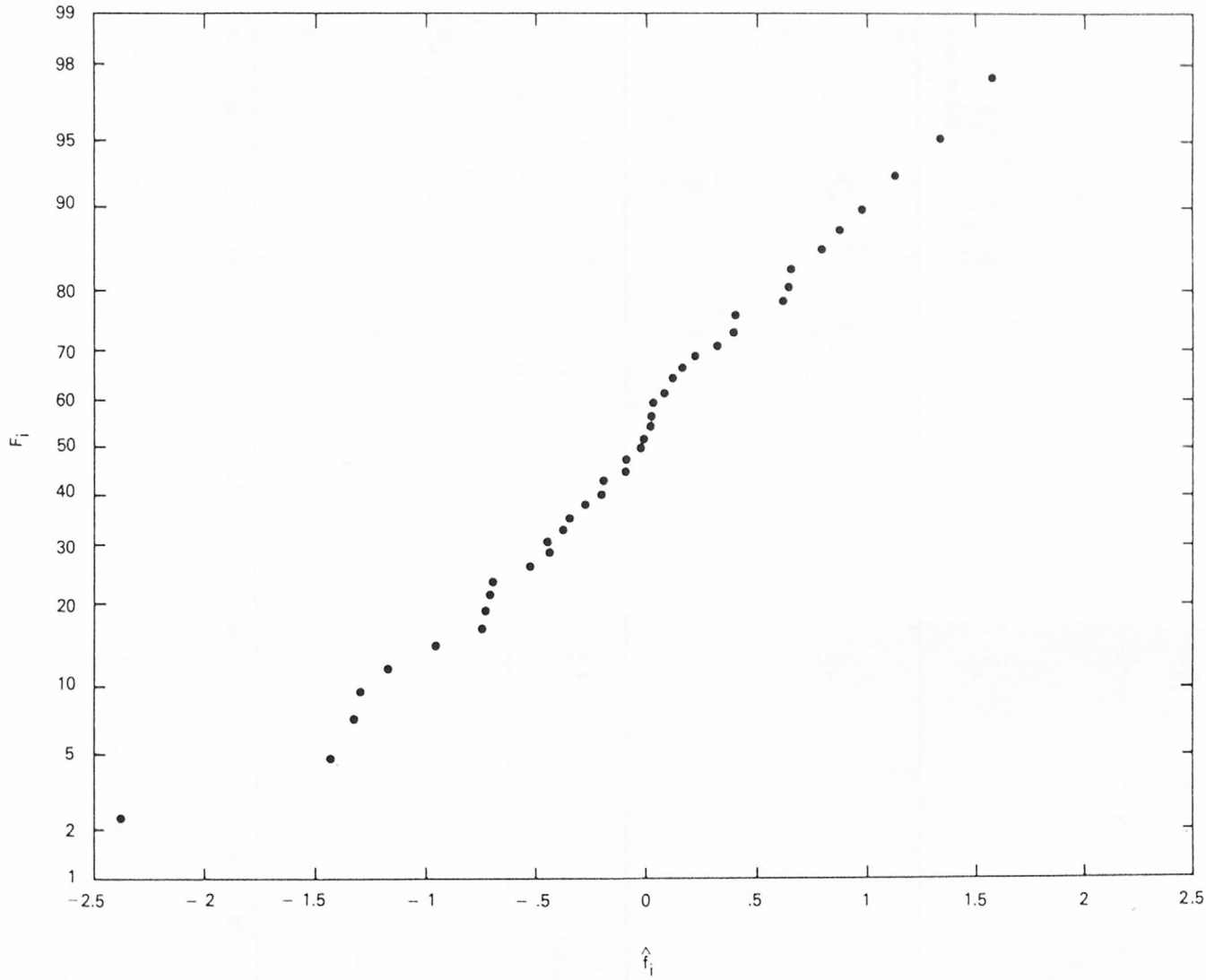

99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
−2.5
−2
−1.5
−1
−.5
0
.5
1
1.5
2
d_i

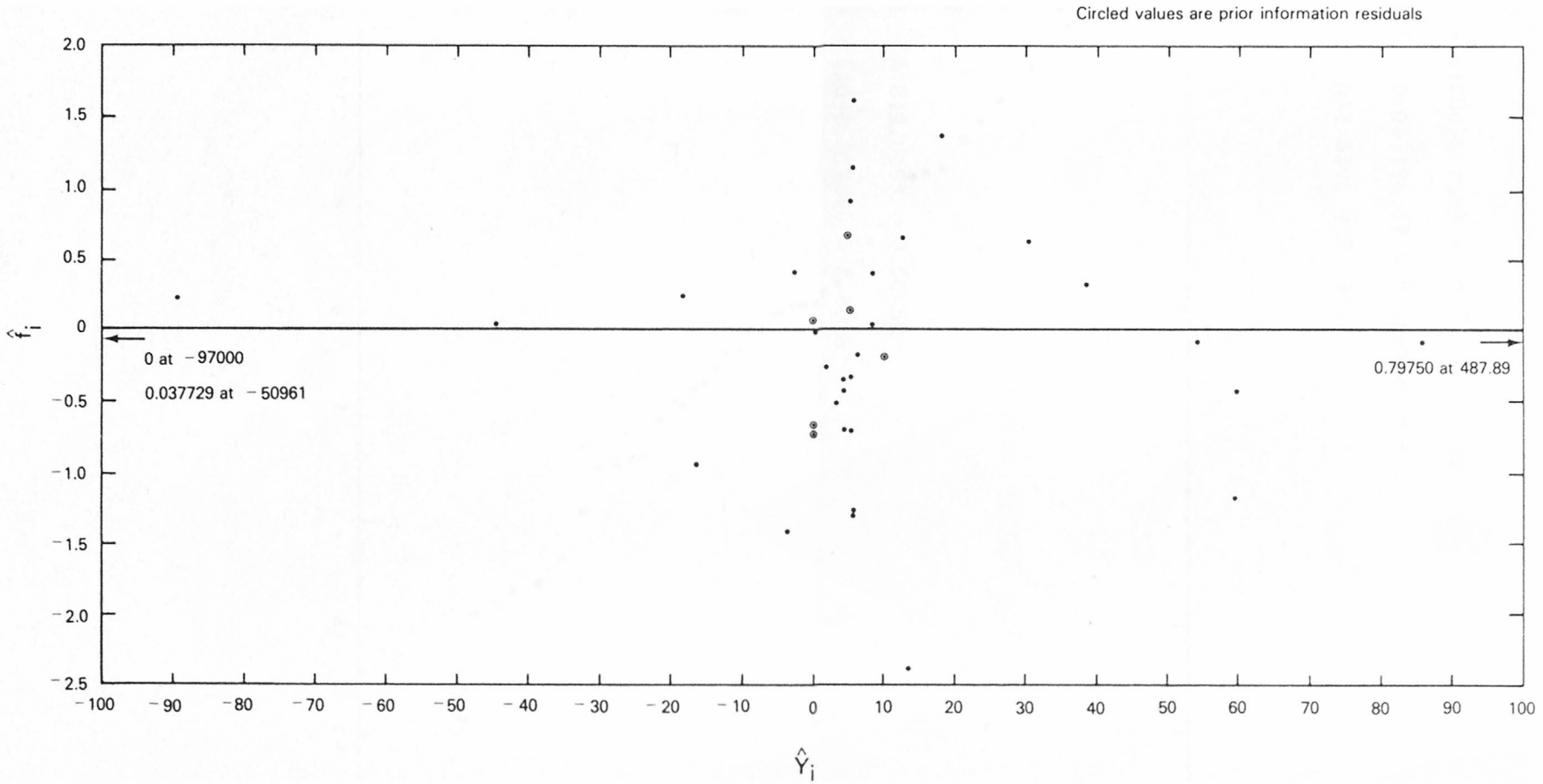

99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
− 2.5
− 2
− 1.5
− 1
− .5
0
.5
1
1.5
2
2.5
g_i

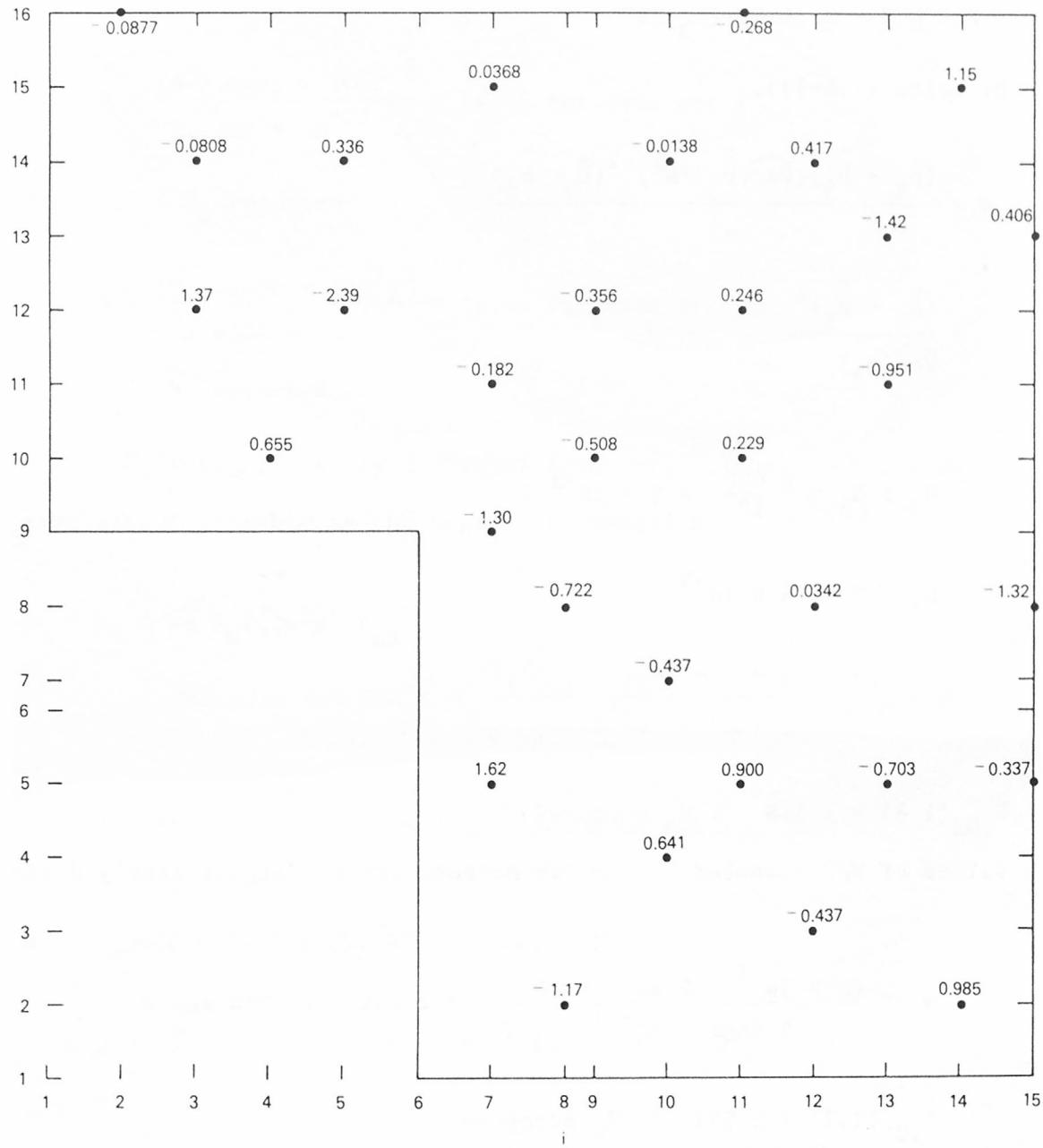
Ordered residual distribution: (The calculated values are prior information residuals $\hat{f}_j = \hat{e}_j \omega_{jj}^{\frac{1}{2}}$, where $\omega_{jj}^{\frac{1}{2}} = s/(Var(\varepsilon_{pj}))^{\frac{1}{2}}$.)

No.	$\hat{f}_j$
1	-2.3900
2	-1.4247
3	-1.3248
4	-1.2986
5	-1.1685
6	-0.95063
7	(0.00031149-0.0004)8278.0 = -0.73269
8	-0.72189
9	-0.70275
10	(0.00013516-0.00017)19478 = -0.67861
11	-0.50751
12	-0.43737
13	-0.43737
14	-0.35588
15	-0.33702
16	-0.26802
17	(10.198-10.4)0.95516 = -0.19294
18	-0.18186
19	-0.087697
20	-0.080833
21	-0.013826
22	$(-97000 + 97000)5.1204 \times 10^{-4} = \sim 0$
23	0.034151
24	0.036783

No.	$\hat{f}_j$
25	$(-50961 + 51000)9.7389 \times 10^{-4} = 0.037982$
26	$(0.080716 - 0.08)124.17 = 0.088906$
27	$(5.4730 - 5.4)1.8396 = 0.13429$
28	0.22868
29	0.24557
30	0.33619
31	0.40574
32	0.41665
33	0.64095
34	0.65517
35	$(5.1211 - 4.8)2.0695 = 0.66452$
36	$(487.89 - 420)0.011826 = 0.80287$
37	0.90021
38	0.98513
39	1.1503
40	1.3703
41	1.6158


99
98
95
90
80
70
60
50
40
30
20
10
5
2
1
F_i
−2.5
− 2
− 1.5
− 1
− .5
0
.5
1
1.5
2
2.5
$\hat{f}_i$


Circled values are prior information residuals
2.0
1.5
1.0
0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5
$\hat{f}_i$
0 at −97000
0.037729 at −50961
0.79750 at 487.89
−100 −90 −80 −70 −60 −50 −40 −30 −20 −10 0 10 20 30 40 50 60 70 80 90 100
$\hat{Y}_i$


16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0.0877
0.268
0.0368
1.15
0.0808
0.336
0.0138
0.417
1.42
0.406
1.37
2.39
0.356
0.246
0.182
0.951
0.655
0.508
0.229
1.30
0.722
0.0342
1.32
0.437
1.62
0.900
0.703
0.337
0.641
0.437
1.17
0.985
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
i

Problem 5.6-1

a) Compute

$$\underline{H}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}\underline{H}^T$$

where $\underline{H} = [0\ 0\ 1]$. Then,

$$\underline{H}(\underline{X}^T \underline{\omega}\ \underline{X})^{-1}\underline{H}^T = \widehat{Var}(\hat{b}_3)/s^2$$

By using (5.6-11),

$$W = \frac{(\tilde{\beta}_3 - \hat{b}_3)(\widehat{Var}(\hat{b}_3)/s^2)^{-1}(\tilde{\beta}_3 - \hat{b}_3)}{s^2}$$

$$= \frac{(\tilde{\beta}_3 - \hat{b}_3)^2}{\widehat{Var}(\hat{b}_3)}$$

$$H_0 : \beta_3 = \frac{0.0003}{10} = 3 \times 10^{-5}$$

$$H_1 : \beta_3 \neq 3 \times 10^{-5}$$

$$W = \frac{(3 \times 10^{-5} - 2.3025 \times 10^{-5})^2}{2.1307 \times 10^{-11}} = 2.283 \text{ for data set 1}$$

$F_{.05}(1,8) = 5.318$ $\therefore$ H_0 accepted

Values of W/T computed by the two methods are not significantly different.

$$W = \frac{(3 \times 10^{-5} - 2.3430 \times 10^{-5})^2}{2.9684 \times 10^{-11}} = 1.454 \text{ for data set 2}$$

$F_{.05}(1,7) = 5.591$ $\therefore$ H_0 accepted

The Maxey-Eakin estimate of W/T could be used as prior information in the regression model, but an estimate of *Var*(W/T) would also be needed.

b) This test is the same as the one in a except that

$$H_0 : \beta_3 = 0 \quad H_1 : \beta_3 \neq 0$$

$$W = \frac{(0-2.3025 \times 10^{-5})^2}{2.1307 \times 10^{-11}} = 24.88 \text{ for data set 1}$$

$\therefore$ H_0 rejected

$$W = \frac{(0-2.3430 \times 10^{-5})^2}{2.9684 \times 10^{-11}} = 18.49 \text{ for data set 2}$$

$\therefore$ H_0 rejected

W/T is significantly different from zero. Thus, recharge is a significant variable in the regression equation.

c) $\tilde{\beta}_3 = \hat{b}_3 \pm \sqrt{q\ F_\alpha(q,n-p)}\ s_{b3}$

$$= 2.3025 \times 10^{-5} \pm \sqrt{5.318} \times 4.6159 \times 10^{-6}$$

$$= 2.3025 \times 10^{-5} \pm 1.0645 \times 10^{-5} \text{ for data set 1}$$

$$\tilde{\beta}_3 = 2.3430 \times 10^{-5} \pm \sqrt{5.591} \times 5.4483 \times 10^{-6}$$

$$= 2.3430 \times 10^{-5} \pm 1.2883 \times 10^{-5} \text{ for data set 2}$$

Problem 5.6-2

For T extreme:

$$T = \hat{T} \pm \frac{\sqrt{2F_{.05}(2,5)}}{s_{bT}} \widehat{Var}(T)$$

$$= 0.11349 \pm \frac{\sqrt{2 \times 5.7861}}{0.0030827} (0.95030 \times 10^{-5}) = 0.11349 \pm 0.010487$$

$$S = \hat{S} \pm \frac{\sqrt{2F_{.05}(2,5)}}{s_{bT}} \widehat{Cov}(T,S)$$

$$= 0.00055221 \pm \frac{\sqrt{2 \times 5.7861}}{0.0030827} (-0.11369 \times 10^{-6}) = 0.00055221 \mp 0.00012546$$

For S extreme:

$$T = \hat{T} \pm \frac{\sqrt{2F_{.05}(2,5)}}{s_{bS}} \widehat{Cov}(T,S)$$

$$= 0.11349 \pm \frac{\sqrt{2 \times 5.7861}}{3.8203 \times 10^{-5}} (-0.11369 \times 10^{-6}) = 0.11349 \mp 0.010124$$

$$S = \hat{S} \pm \frac{\sqrt{2F_{.05}(2,5)}}{s_{bS}} \widehat{Var}(S)$$

$$= 0.00055221 \pm \frac{\sqrt{2 \times 5.7861}}{3.8203 \times 10^{-5}} (0.14595 \times 10^{-8}) = 0.00055221$$

$$\pm 0.00012996$$

Problem 5.6-3

$$\underset{\sim}{b} = \hat{\underset{\sim}{b}} \pm \frac{\sqrt{qF_{\alpha}(q,n-p)}}{s_{bi}} \underset{\sim}{v}_{bi}$$

where $\underset{\sim}{v}_b = (\underset{\sim}{X}^T \underset{\sim}{\omega}\ \underset{\sim}{X})^{-1} s^2$. We have from the output $(\underset{\sim}{Z}^T \underset{\sim}{\omega}\ \underset{\sim}{Z})^{-1} s^2$, but need $(\underset{\sim}{X}^T \underset{\sim}{\omega}\ \underset{\sim}{X})^{-1} s^2$. Thus,

$$(\underset{\sim}{Z}^T \underset{\sim}{\omega}\ \underset{\sim}{Z})^{-1} = ((\underset{\sim}{X}\ \underset{\sim}{B})^T \underset{\sim}{\omega}\ \underset{\sim}{X}\ \underset{\sim}{B})^{-1}$$

$$= \underset{\sim}{B}^{-1} (\underset{\sim}{X}^T \underset{\sim}{\omega}\ \underset{\sim}{X})^{-1} (\underset{\sim}{B}^T)^{-1}$$

or

$$(\underset{\sim}{X}^T \underset{\sim}{\omega}\ \underset{\sim}{X})^{-1} = \underset{\sim}{B} (\underset{\sim}{Z}^T \underset{\sim}{\omega}\ \underset{\sim}{Z})^{-1} \underset{\sim}{B}$$

because $\underset{\sim}{B} = \underset{\sim}{B}^T$. Because $\underset{\sim}{B}$ is also diagonal, the above computation is easy.

Let $(\underset{\sim}{Z}^T \underset{\sim}{\omega}\ \underset{\sim}{Z})^{-1} s^2 = \underset{\sim}{A} = \{A_{ij}\}$. Then

$$\underset{\sim}{B}\ \underset{\sim}{A}\ \underset{\sim}{B} = \{B_{ii} A_{ij} B_{jj}\}$$

Thus,

$$\frac{v_{bij}}{s_{bi}} = \frac{B_{ii} A_{ij} B_{jj}}{\sqrt{B_{ii} A_{ii} B_{ii}}} = \frac{A_{ij} B_{jj}}{\sqrt{A_{ii}}}$$

and

$$\tilde{b}_j = \hat{b}_j \pm \sqrt{qF_{\alpha}(q,n-p)}\ \frac{A_{ij} B_{jj}}{\sqrt{A_{ii}}}$$

$$= \hat{b}_j \pm \sqrt{qF_{\alpha}(q,n-p)}\ \frac{A_{ij}}{\sqrt{A_{ii}}}\ \hat{b}_j$$

where $B_{jj} \cong \hat{b}_j$.

Calculations:

$$\sqrt{qF_{\alpha}(q,n-p)} = \sqrt{2F_{.05}(2,27)} = \sqrt{2 \times 3.359} = 2.592$$

For $T_3 = \beta_{12}$ extreme:

$\tilde{b}_1 = 80.978 \pm 9.2956$

$\tilde{b}_2 = 935.15 \pm 209.10$

$\tilde{b}_3 = -97000 \mp 85.307$

$\tilde{b}_4 = -50961 \mp 29.377$

$\tilde{b}_5 = 10.198 \pm 0.11522$

$\tilde{b}_6 = 5.1211 \pm 1.3759 \times 10^{-2}$

$\tilde{b}_7 = 5.4730 \pm 1.2759 \times 10^{-2}$

$\tilde{b}_8 = 65.754 \pm 7.2417$

$\tilde{b}_9 = 3.1149 \times 10^{-4} \pm 3.1005 \times 10^{-5}$

$\tilde{b}_{10} = 487.89 \pm 2.5669$

$\tilde{b}_{11} = -1.3995 \times 10^{-4} \mp 4.9303 \times 10^{-5}$

$\tilde{b}_{12} = 13.288 \pm 9.4103$

$\tilde{b}_{13} = 1.3516 \times 10^{-4} \pm 9.3157 \times 10^{-5}$

$\tilde{b}_{14} = 8.0716 \times 10^{-2} \pm 3.5369 \times 10^{-5}$

For $q_{B1} = \beta_1$ extreme:

$\tilde{b}_1 = 80.978 \pm 44.058$

$\tilde{b}_2 = 935.15 \pm 490.71$

$\tilde{b}_3 = -97000 \mp 222.36$

$\tilde{b}_4 = -50961 \mp 73.632$

$\tilde{b}_5 = 10.198 \pm 0.27772$

$\tilde{b}_6 = 5.1211 \pm 3.2748 \times 10^{-2}$

$$\tilde{b}_7 = 5.4730 \pm 3.2200 \times 10^{-3}$$

$$\tilde{b}_8 = 65.754 \pm 25.576$$

$$\tilde{b}_9 = 3.1149 \times 10^{-4} \pm 5.4119 \times 10^{-5}$$

$$\tilde{b}_{10} = 487.89 \pm 6.2307$$

$$\tilde{b}_{11} = -1.3995 \times 10^{-4} \pm 1.1721 \times 10^{-4}$$

$$\tilde{b}_{12} = 13.288 \pm 1.9854$$

$$\tilde{b}_{13} = 1.3516 \times 10^{-4} \pm 2.0397 \times 10^{-5}$$

$$\tilde{b}_{14} = 8.0716 \times 10^{-2} \pm 1.5000 \times 10^{-4}$$

Problem 5.7-1

a) Let $\underline{X} = \{X_{ij}\}$, $(\underline{X}^T \underline{\omega}\, \underline{X})^{-1} = \{A_{ij}\}$. Then

$$\underline{X}(\underline{X}^T \underline{\omega}\, \underline{X})^{-1} = \{\sum_{k=1}^{p} X_{ik} A_{k\ell}\} = \{C_{i\ell}\}$$

$$\underline{X}(\underline{X}^T \underline{\omega}\, \underline{X})^{-1}\underline{X}^T = \{\sum_{\ell=1}^{p} C_{i\ell} X_{j\ell}\} = \{\sum_{\ell=1}^{p} \sum_{k=1}^{p} X_{ik} A_{k\ell} X_{j\ell}\}$$

If $i = j$, then the entry is

$$\sum_{\ell=1}^{p} \sum_{k=1}^{p} X_{ik}\, A_{k\ell}\, X_{i\ell} = \sum_{\ell=1}^{p} X_{i\ell} \sum_{k=1}^{p} X_{ik} A_{k\ell}$$

Compute $i = j = 1$ for data set 1

$$\underline{X}_1 = [0.95 \quad 0.05 \quad 23750]$$

$$(\underline{X}^T \underline{\omega}\, \underline{X})^{-1} = \begin{bmatrix} 0.9157290438 & 0.2752072275 & -5.983931169 \times 10^{-6} \\ & 0.7831966299 & -5.318014058 \times 10^{-6} \\ \text{symmetric} & & 6.878625596 \times 10^{-11} \end{bmatrix}$$

$$\underline{X}_1(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\ \underline{X}_1^T = 0.95(0.95 \times 0.9157290438 + 0.05 \times 0.2752072275$$

$$+ 23750 \times (-5.983931169 \times 10^{-6})) + 0.05(0.95 \times 0.2752072275$$

$$+ 0.05 \times 0.7831966299 + 23750 \times (-5.318014058 \times 10^{-6}))$$

$$+ 23750 \times (0.95 \times (-5.983931169 \times 10^{-6}) + 0.05 \times (-5.318014058 \times 10^{-6})$$

$$+ 23750 \times 6.878625596 \times 10^{-11}))$$

$$= 0.6106927103$$

$$\begin{aligned}\underline{X}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}\underline{X}^T &= \underline{X}\ \underline{C}\ \underline{C}^{-1}(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}(\underline{C}\ \underline{C}^{-1})^T\underline{X}^T \\ &= \underline{X}\ \underline{C}(\underline{C}^T\underline{X}^T\underline{\omega}\ \underline{X}\ \underline{C})^{-1}\underline{C}^T\underline{X}^T \\ &= (\underline{X}\ \underline{C})((\underline{X}\ \underline{C})^T\underline{\omega}(\underline{X}\ \underline{C}))^{-1}(\underline{X}\ \underline{C})^T \\ &= \underline{S}\ (\underline{S}^T\underline{\omega}\ \underline{S})^{-1}\underline{S}^T\end{aligned}$$

b) $\widehat{Var}(\hat{Y}_j) = \underline{X}_j\widehat{Var}(\hat{\underline{b}})\underline{X}_j^T$

Let j=1, data set 1.

$$\begin{aligned}\widehat{Var}(Y_1) &= \underline{X}_1(\underline{X}^T\underline{\omega}\ \underline{X})^{-1}s^2\underline{X}_1^T \\ &= 0.6106927103 \times 0.30975625 \\ &= 0.1891658838\end{aligned}$$

c) $Y_{\beta j} = \hat{Y}_j \pm \sqrt{pF_\alpha(p,n-p)}\ s_{yj}$

$F_{.05}(3,8) = 4.066$

$$\begin{aligned}\hat{Y}_1 &= 0.95\hat{b}_1 + 0.05\hat{b}_2 + 23750\hat{b}_3 \\ &= 0.95(50.1204) + 0.05(9.48742) + 23750(2.30248 \times 10^{-5}) \\ &= 48.636\end{aligned}$$

$$\begin{aligned}Y_{\beta 1} &= 48.636 \pm \sqrt{3 \times 4.066} \times 0.434932 \\ &= 48.636 \pm 1.519\end{aligned}$$

for j=1, data set 1.

Problem 6.2-1

Sets of parameters for Beale's measure:

1) (0.12398, 0.00042675)

2) (0.10300, 0.00067767)

3) (0.10337, 0.00068217)

4) (0.12361, 0.00042225)

Only 2 and 4 need be used because the other two are nearly the same. By using 2, drawdown, s, is:

1.5947, 2.2223, 2.5541, 2.8218, 3.0676, 3.3503, 3.5979

By using 4, s is:

1.7821, 2.3239, 2.6058, 2.8320, 3.0390, 3.2763, 3.4839

The resulting value of Beale's measure is

$$\hat{N}_b = 0.016416$$

Because $F_{.05}(2,5) = 5.7861$, the model is effectively linear.

PROGRAMMER _____ DIVISION _____
LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

NUMBER _____
PROJECT _____

SHEET 1 OF 1

IDENTIFICATION: INPUT TØ BEALE'S MEASURE PRØGRAM FØR PRØBLEM 6.2-1
INFORMATION: THEIS DATA SET FRØM PRØBLEM 3.3-1

```
  2       2       7       0       2    .0014328
  .11349  .00055221
  1.6715          2.2521          2.5564          2.8012          3.0256          3.2832          3.5086
          1               1               1               1               1               1               1
 -.91882         -.75264
 -1.4679         -.78421
 -1.7630         -.79343
 -2.0026         -.79866
 -2.2234         -.80223
 -2.4779         -.80527
 -2.7014         -.80724
  .10300  .00067767
  1.5947          2.2223          2.5541          2.8218          3.0676          3.3503          3.5979
  .12361  .00042225
  1.7821          2.3239          2.6058          2.8320          3.0390          3.2763          3.4839
```

444

FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [||] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—88265 p-1

Problem 6.2-2

From the computer output, Beale's measure is 0.043798. Based on $F_{.05}(2,27) = 3.359$, the Beale's measure indicates that the model is almost effectively linear. Hence, linear theory can be applied to use the W statistic based on q = 2.

PROGRAMMER _____ DIVISION _____
LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
COMPUTER CODING FORM

NUMBER _____
PROJECT _____

SHEET 1 OF 2

IDENTIFICATION: INPUT TØ CØMPUTE BEALE'S MEASURE FØR PRØBLEM 6.2-2
INFORMATION: SEVERAL ZØNE, 2-D FLØW DATA SET FRØM PRØBLEM 4.2-2

```
Card set A the same as for problem 4.2-2
 15  16   4   7   0   4   4   2   3  10   1   0   0   0
Card setts C through S the same as for problem 4.2-2
 4
 1  72.996      72.996                  0 3.4249E-4
 2  490.46      490.46                  0-1.8925E-4
 3  22.698      22.698                  0 2.2832E-4
 4  490.46      490.46 8.0751E-2                    0
 1  90.274
 2  1144.3
 3  -97085
 4  -50990
 1  10.313      5.1349
 2  5.1349      5.4858
 1  58.512      58.512                  0 2.8048E-4
 2  485.32      485.32                  0-9.0647E-5
 3  3.8777      3.8777                  0 4.2003E-5
 4  485.32      485.32 8.0681E-2                    0
 1  71.682
 2  726.05
 3  -96915
 4  -50932
 1  10.083      5.1073
 2  5.1073      5.4602
 1  91.330      91.330                  0 3.6561E-4
 2  494.12      494.12                  0-2.5716E-4
 3  15.273      15.273                  0 1.5556E-4
 4  494.12      494.12 8.0866E-2                    0
 1  125.04
 2  1425.9
```

FORM 9-1633 A
JULY 1967

0 = ZERO Ø = ALPHA O 1 = ONE I = ALPHA I 2 = TWO Z = ALPHA Z / = SLASH | = VERT. BAR - = MINUS _ = HORZ. BAR

DATA CODING FORM

GPO 16-88263 p-1

PROGRAMMER _____ DIVISION _____

LOCATION _____ PHONE _____

U. S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

COMPUTER CODING FORM

NUMBER _____

PROJECT _____

SHEET 2 OF 2

IDENTIFICATION: _____

INFORMATION: _____

```
3    -97222
4    -51035
1    10.476      5.1538
2    5.1538      5.4698
1    40.178      40.178                  02.5737E-4
2    481.66      481.66                  0-2.2740E-5
3    11.303      11.303                  01.1476E-4
4    481.66      481.66  8.0566E-2                  0
1    36.920
2    444.44
3    -96778
4    -50887
1    9.9203      5.0884
2    5.0884      5.4762
2
```

447

FORM 9—1633 A
JULY 1967

[0] = ZERO [Ø] = ALPHA O [1] = ONE [I] = ALPHA I [2] = TWO [Z] = ALPHA Z [/] = SLASH [|] = VERT. BAR [-] = MINUS [_] = HORZ. BAR

DATA CODING FORM

GPO 16—86265 p-1

Problem 6.3-1

Compute

$$\hat{\gamma} = (\underline{Y}_p - \underline{X}_p \hat{\underline{b}}^*)^T [s^2 \underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T + \underline{U}]^{-1} (\underline{Y}_p - \underline{X}_p \hat{\underline{b}}^*)$$

For data set 1

$$\hat{\underline{b}}^* = \begin{bmatrix} 50.018 \\ 9.1954 \\ 2.5008 \times 10^{-5} \end{bmatrix}$$

For data set 2

$$\hat{\underline{b}}^* = \begin{bmatrix} 50.055 \\ 9.7914 \\ 2.2766 \times 10^{-5} \end{bmatrix}$$

Also,

$$\underline{X}_p (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} \underline{X}_p^T = \underline{X}_p \underline{B}\ \underline{B}^{-1} (\underline{X}_s^T \underline{V}_s^{-1} \underline{X}_s)^{-1} (\underline{B}\ \underline{B}^{-1})^T \underline{X}_p^T$$

$$= (\underline{X}_p \underline{B}) ((\underline{X}_s \underline{B})^T \underline{V}_s^{-1} (\underline{X}_s \underline{B}))^{-1} (\underline{X}_p \underline{B})^T$$

$$= \underline{Z}_p (\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T$$

For data set 1

$$(\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} s^2 = \begin{bmatrix} 1.0213 \times 10^{-4} & 1.9520 \times 10^{-4} & -1.3870 \times 10^{-3} \\ & 3.0217 \times 10^{-3} & -7.5445 \times 10^{-3} \\ \text{symmetric} & & 3.3124 \times 10^{-2} \end{bmatrix}$$

$$s^2 \underline{Z}_p (\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T = 9.1954 \times 3.0217 \times 10^{-3} \times 9.1954$$

$$= 0.255501$$

$$\underline{U} = (1.1)^2 = 1.21$$

$$s^2 \underline{Z}_p (\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T + \underline{U} = 0.255501 + 1.21$$

$$= 1.46550$$

$$[s^2 \underline{Z}_p (\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T + \underline{U}]^{-1} = 0.68236$$

$$\underline{Y}_p - \underline{X}_p \hat{\underline{b}}^* = 11 - 9.1954 = 1.8046$$

$$\hat{\gamma} = 1.8046 \times 0.68236 \times 1.8046 = 2.222$$

$$\chi^2_{.05}(1) = 3.841 \qquad \therefore H_0 \text{ accepted}$$

Prior and pure regression estimates of $\beta_2 \equiv h_b$ are in agreement.

For data set 2

$$(\underline{Z}_s^T \underline{V}_s^{-1} \underline{Z}_s)^{-1} s^2 = \begin{bmatrix} 2.1601 \times 10^{-4} & 5.3590 \times 10^{-4} & -3.4922 \times 10^{-3} \\ & 5.6452 \times 10^{-3} & -1.7852 \times 10^{-2} \\ \text{symmetric} & & 8.3760 \times 10^{-2} \end{bmatrix}$$

$$s^2 \underline{Z}_p (\underline{Z}_s \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T = 9.7914 \times 5.6452 \times 10^{-3} \times 9.7914$$

$$= 0.541195$$

$$\underline{U} = (0.95)^2 = 0.9025$$

$$s^2 \underline{Z}_p (\underline{Z}_s \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T + \underline{U} = 0.541195 + 0.9025$$

$$= 1.44369$$

$$[s^2 \underline{Z}_p (\underline{Z}_s \underline{V}_s^{-1} \underline{Z}_s)^{-1} \underline{Z}_p^T + \underline{U}]^{-1} = 0.69267$$

$$\underline{Y}_p - \underline{X}_p \hat{\underline{b}}^* = 9.5 - 9.7914 = -0.2914$$

$$\hat{\gamma} = -0.2914 \times 0.69267 \times (-0.2914)$$

$$= 0.05882$$

$\therefore$ H_0 accepted

☆U.S. GOVERNMENT PRINTING OFFICE:1985— 580-310 / 25352