LYON, CHARLES CROSBY

THE SOLUTION OF ILL-POSED SYSTEMS OF LINEAR EQUATIONS IN THE PRESENCE OF NOISE, WITH APPLICATIONS IN GEOTOMOGRAPHY

THE UNIVERSITY OF ARIZONA M.S. 1984

THE SOLUTION OF ILL-POSED SYSTEMS OF LINEAR EQUATIONS IN THE PRESENCE OF NOISE, WITH APPLICATIONS IN GEOTOMOGRAPHY

by

Charles Crosby Lyon

A Thesis Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE
WITH A MAJOR IN ELECTRICAL ENGINEERING

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 4

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Charles Lyon

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

D. G. Dudley
Professor of Electrical Engineering

4/24/84
Date

ACKNOWLEDGMENTS

I wish to express my gratitude to the following people:

Dr. Donald Dudley of the Department of Electrical and Computer Engineering and the Department of Applied Mathematics at the University of Arizona, for serving as my thesis advisor.

Dr. Carl DeVito of the Department of Mathematics at the University of Arizona, for verifying the theorem proofs and suggesting numerous valuable improvements.

Dr. Carl Glass of the Department of Mining and Geological Engineering at the University of Arizona, and the United States Nuclear Regulatory Commission; for funding the initial portion of this research under contract #NRC-04-78-269.

Sarah Oordt, for patiently typing the manuscript under difficult conditions.

Allan Honda, for being a good friend and coming through in a pinch.

The Miller family - Chris; Norman, Sr.; and Norman, Jr. - for taking me into their home during the difficult

period while I simultaneously finished this thesis and started a new job.

My parents - Mary K. Lyon and A. C. Lyon, Jr. - whose unfailing support has made my career possible.

These contributions are deeply appreciated.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--Continued

ABSTRACT

Analyses are presented for ill-posed systems of linear equations and various means of solving them. Emphasis is placed on methods related to the projection method (also referred to as the Kaczmarz method or an algebraic reconstruction technique (ART)). It is shown that stopping these algorithms prior to convergence in effect imposes a stability constraint on the solutions. Criteria are proposed by which to estimate the optimal stopping point. The various projection algorithms have advantages over other methods for calculating stabilized solutions in that

(1) they can be made efficient for large, sparse systems;

(2) they allow easy application of inequality constraints; and

(3) if the proper stopping criterion is employed, they can automatically impose stronger stability constraints when more noise is present.

As an application, the capabilities of geotomography (a means of subsurface geophysical exploration) are evaluated via computer simulations. It is verified that geotomography is effective for several applications.

CHAPTER 1

INTRODUCTION

A problem of current interest in geophysical exploration is that of determining with high resolution the geologic features of a region at a relatively great depth beneath the earth's surface. This problem arises in coal gasification, where it is necessary to monitor the gasification of beds too deep or inaccessible to be mined by conventional methods [33, 43]. It also arises in nuclear waste disposal site evaluation, where it is important to detect even relatively small imperfections in the host rock around the repository site [16]. This type of exploration could also be used to map mineral deposits and could be used prior to tunneling to detect hazards such as pockets of gas or water. Because the currently well-established geophysical exploration techniques can detect only large scale features at great depth, they are not applicable to problems of this nature. Furthermore, many of these techniques can detect only certain types of features, such as layering. To obtain more detailed information about a region at great depth, it is necessary to position instruments nearby, in tunnels or boreholes.

Ray-Optic Geotomography (ROG) and the Volume Eddy Current Techinque of Reconstruction (VECTOR) are two

approaches to the high resolution geophysical exploration problem in which the region between two boreholes is probed by transmitting electromagnetic radiation between the holes. The ROG method was introduced by Lager and Lytle [32] and VECTOR has been proposed by Howard [27]. Although this paper will address only the ROG method directly, both methods must overcome a similar, difficult mathematical problem. ROG and VECTOR both ultimately require the solution in the presence of noise of systems of linear equations which may be very large, rank-deficient, and highly ill-posed. (Rank-deficient systems are those having fewer independent equations than unknowns. Ill-posed systems are those for which a small perturbation of the equations can cause a large perturbation of the solution.) It is also desirable to apply inequality constraints to the solutions, such as requiring the unknowns to be non-negative. These special circumstances make the solution of these equations by conventional methods impractical. Chapters 2 and 3 of this paper are devoted to the study of systems of equations of this type, and to the investigation and development of appropriate algorithms for their solution. Particular attention is given to the error due to noise and means of reducing it. The projection method and related methods are studied in depth and shown to be able to apply inequality constraints and suppress noise error

efficiently for large systems. The results presented in Chapters 2 and 3 are general enough to be applicable to a wide range of problems, including VECTOR as well as ROG. In Chapter 4, ROG is discussed and its performance is evaluated by inverting computer-generated data in the presence of noise with various algorithms. ROG is shown to be an effective means of short-range, high-resolution probing in low-loss media such as salt or dry rock.

Both mathematically rigorous and empirically derived results are presented here. An effort is made to distinguish between the two, with major rigorous results usually being presented in theorem form. In an effort to make this paper readily accessible to a wide audience, relevent background material is briefly reviewed as it is needed, and the more involved derivations are included in Appendix 2. Parts of this appendix demand more mathematical background of the reader than the main text. Some key concepts of linear algebra are reviewed in Appendix 1.

CHAPTER 2

THEORY PERTINENT TO THE SOLUTION OF ILL-POSED SYSTEMS OF LINEAR EQUATIONS IN THE PRESENCE OF NOISE

A system of real linear equations (subsequently referred to here simply as a system) may be expressed in the form

$$Ax = b, \tag{2.1}$$

where $A \in E^{m \times n}$ (i.e., A is an $m \times n$ real matrix), $x \in E^{n}$ is the vector of unknowns, and $b \in E^{m}$ is known. (E^{k} denotes the Euclidean space of k dimensions i.e., the space of all real k-dimensional vectors.) The search for a solution of (2.1) may be complicated by inconsistency (the failure of a solution to exist), non-uniqueness (the existence of more than one solution) and instability (the sensitivity of a solution to errors in A or b). Because of these conditions it can be far from simple in practice even to define the word "solution". This chapter develops adequate background material to permit definition of solutions which always exist and have practical value. Sections 2.1 through 2.4 present various means of dealing with

inconsistency and non-uniqueness, emphasizing the use of the least-squares approach and the minimum-norm constraint. The effects of errors in A and b on the solution are examined in Section 2.5, and it is concluded that in the presence of noise ill-posed systems are best solved by first perturbing them in such a way as to render them well-posed. Various algorithms capable of doing this will be discussed in Chapter 3.

Each equation of (2.1) may be viewed geometrically as a hyperplane in E^n, and a solution of (2.1) is a point of intersection of all the hyperplanes. When $n = 2$ the hyperplanes are lines. Figure 1 is a graph of the hyperplanes of the two-dimensional system

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x(1) \\ x(2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \tag{2.2}$$

which has the unique solution $x(1) = x(2) = 1$. Simple graphical examples of this nature will provide valuable intuitive insight into many of the theoretical results presented in this paper.

It is convenient at this point to establish some notation which will be used throughout the remainder of this paper. The elements of A, x, and b are denoted by

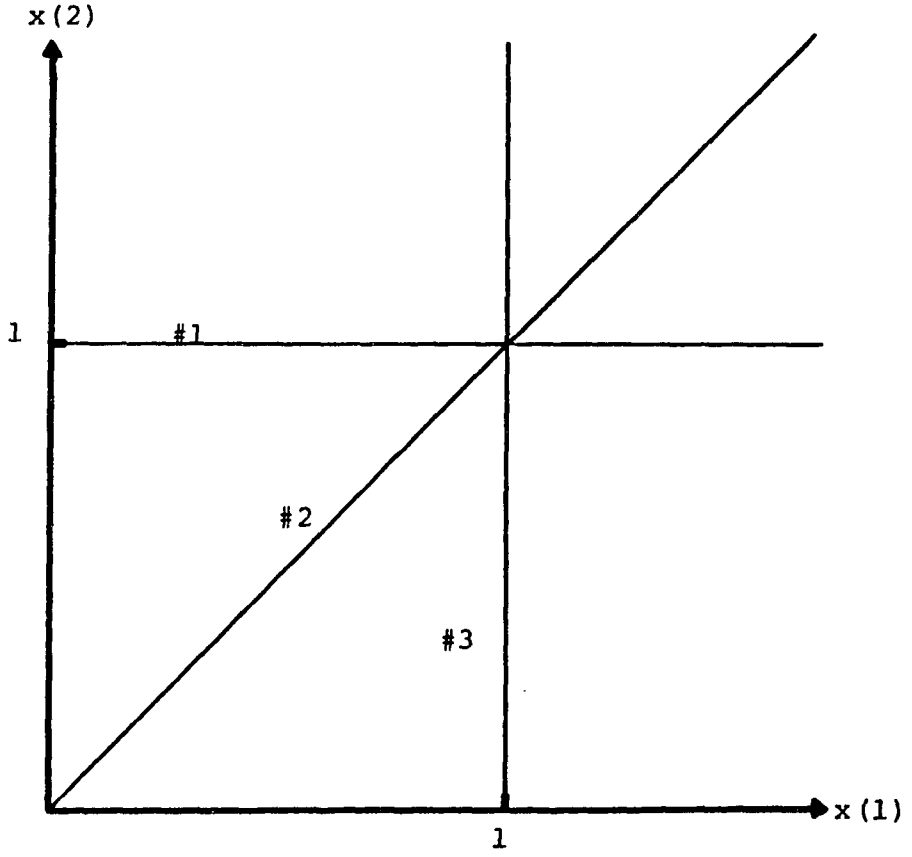


Figure 1. A consistent system with a unique solution.
(The hyperplanes are numbered in the order of the equations which define them.)

A(i,j), x(j) and b(i) respectively, and the row vectors of A are denoted by a_i, where $1 \leq i \leq m$ and $1 \leq j \leq n$. The number of equations is m, and the number of unknowns is n. In general, m may be less than, equal to, or greater than n. Although all matrices and vectors are assumed to be real and finite-dimensional in this treatment, many of the results to be presented can readily be extended to systms of complex linear equations. (Note that any complex linear equation can be reformulated as two real linear equations simply by equating the real and imaginary parts of both sides.)

The inner product of two equally-dimensioned vectors x and y is defined by

$$\langle x,y\rangle = x^T y, \tag{2.3}$$

where x^T denotes the transpose of x. Vector norm is defined by

$$\|x\| = \langle x,x\rangle^{1/2}, \tag{2.4}$$

and matrix norm is defined by

$$\|A\| = \sup\{\frac{\|Ax\|}{\|x\|}: \|x\| \neq 0\}, \tag{2.5}$$

where sup denotes the least upper bound. Occasional reference will be made to the Frobenius norm, defined by

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} A(i,j)^2 \right)^{1/2}. \tag{2.6}$$

The rank of A (i.e., the number of independent equations) is denoted $r(A)$. I and $\underline{0}$ denote the identity matrix and the zero matrix (or vector) with the dimensions appropriate to the context. $N(A)$ and $R(A)$ denote the null space and range of A, respectively, which are defined by

$$N(A) = \{x : Ax = \underline{0}\} \tag{2.7}$$

and

$$R(A) = \{b : Ax = b \text{ for some } x\}. \tag{2.8}$$

(Equation (2.7) is read "The null space of A is the set of all vectors x such that Ax equals the zero vector".) Upper case script letters will be used to denote spaces, and the dimension of a space S will be denoted $\dim(S)$. The orthogonal complement of a subspace S will be denoted $S^{\perp}$, and the projection of a vector x onto S, denoted x_S, will sometimes be referred to here as the component of x in S. Orthogonal complements and projections are reviewed in Appendix 1.

2.1. Inconsistency and Least Squares

In the presence of noise, A and b are assumed to have been perturbed. Thus

$$A = A_u + \delta A \tag{2.9}$$

and

$$b = b_u + \delta b, \tag{2.10}$$

where A_u and b_u are the unperturbed values of A and b, and δA and δb are the perturbations or noise terms. A and b are known, but A_u and b_u are assumed to be unavailable. In practice, a small amount of noise is virtually always present due to round-off errors by digital computers. Many problems of interest have high levels of noise.

It is easily seen that (2.1) is consistent, (i.e, has a solution) if and only if $b \in R(A)$. Noise can render a system inconsistent by moving b outside $R(A)$, or by changing $R(A)$ so that it no longer contains b. In geometric terms, this corresponds to a perturbation of the hyperplanes in such a way that they no longer intersect at a single point. Even if the unperturbed system

$$A_u x_u = b_u \tag{2.11}$$

is known to be consistent, in general it must be assumed that (2.1) may be inconsistent due to noise. (The only exception to this is when the equations are independent (i.e., $r(A) = m$), since they cannot contradict each other unless some of them are redundant. In this case $R(A) = E^m$.)

Figure 2 shows an inconsistent system which could be obtained by perturbing the system depicted in Figure 1. Perturbations of A alter the angles of the hyperplanes from the horizontal, while perturbations of b move the hyperplanes laterally. Although inconsistent systems such as this have no solutions, it is possible to find vectors which come close to satisfying the equations in some sense. The classical least-squares method is most often used for this purpose.

A least-squares solution x_{ls} is a vector which minimizes the residual norm $\|b - Ax\|$. It is well-known that the least-squares solutions of (2.1) are the solutions of the consistent system

$$Ax_{ls} = b_{R(A)}. \tag{2.12}$$

(Recall that $b_{R(A)}$ denotes the component of b in the range of A.) Thus, in efect, the least-squares approach renders (2.1) consistent by subtracting from b its component outside the range of A. This constitutes the

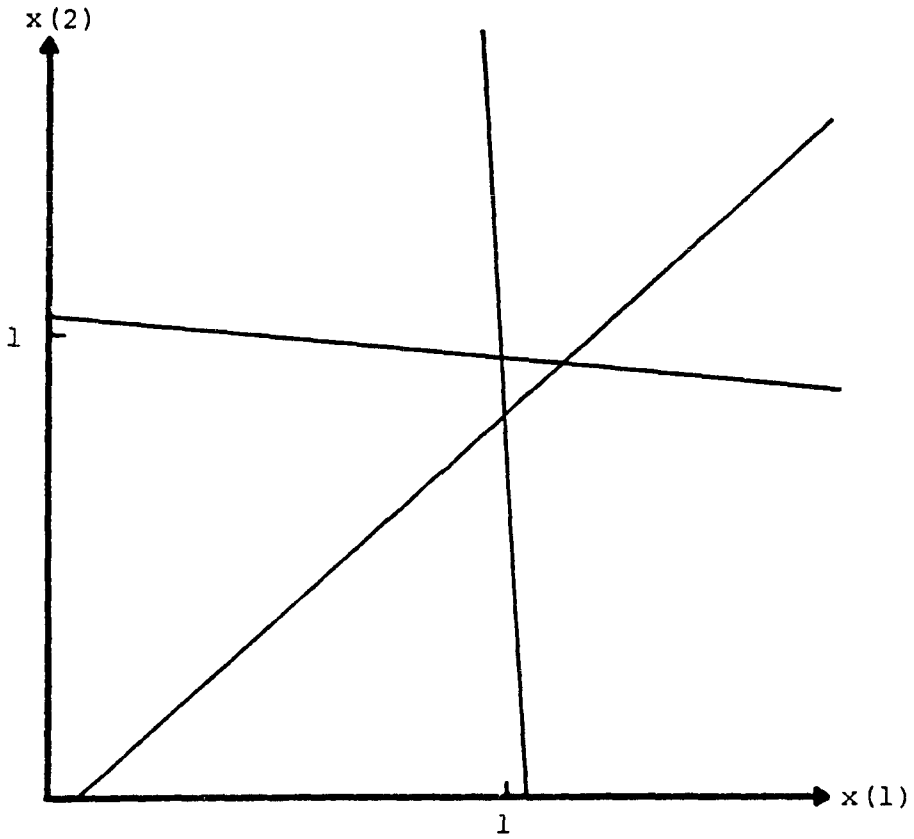


Figure 2. An inconsistent system.

smallest possible perturbation of b which renders (2.1) consistent [18].

Although (2.12) lends insight to the least-squares method, other means are used to find x_{ls} in practice. Some of these will be considered in subsequent sections, notable among them being use of the Gauss normal equations

$$A^T A x = A^T b. \tag{2.13}$$

Equation (2.13) is consistent and its solutions are the least-squares solutions of (2.1).

Treatments of the least-squares method may be found in [1, 2, 35, 41, 48] and in most introductory texts on linear algebra.

2.2. Total Least Squares

In the previous section it was observed that the least-squares method in effect renders (2.1) consistent by making the smallest possible perturbation of b and leaving A unchanged. This approach is valid if one assumes that the inconsistency of (2.1) is primarily due to error in b. It is possible, however, that b could be accurately known and that any inconsistency could be largely due to error in A. In such a case, it might be preferable to render the equations consistent by slightly altering A instead of b, or

by altering both A and b. This is known as the total least-squares problem, and it has been the subject of recent research [20]. Although the total least-squares method will not be applied in its usual form in this paper, the general approach of perturbing A prior to solution will prove useful. The total least-squares approach also serves to illustrate that while the traditional least-squares solution is valuable and universally accepted, it is not the only valid "solution" of an inconsistent system of linear equations.

2.3. Weighted Least Squares

The residual norm may be expressed as

$$\|b - Ax\| = \left[\sum_{i=1}^{m} (b(i) - \langle a_i, x\rangle)^2\right]^{1/2} \tag{2.14}$$

(see Equation (A1.27) in Appendix 1). Since a_i contains the coefficients of the i'th equation, it is clear that if one equation has large coefficients it may be dominant in determining the residual norm, and may therefore play an unduly large role in determining the least-squares solutions. If some of the equations are considered to be more important or more reliable than the others, then their influence on the least-squares solutions may be increased by scaling them so that the norms of their row vectors are larger than the others. This may be referred to as weighting the equations,

and it is done by simply multiplying both sides of each equation by an appropriate factor. A graphical example of weighting appears in Figure 3. The systems

$$\begin{bmatrix} -2 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \tag{2.15}$$

and

$$\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1/\sqrt{2} \\ 1 \\ 1 \end{bmatrix} \tag{2.16}$$

both describe the set of hyperplanes graphed in Figure 3. However, the first row vector of (2.15) has a larger norm than the others, while the row vectors of (2.16) all have unit norm. Therefore, the least-squares solution of (2.15) is closer to the first hyperplane than the least-squares solution of (2.16). (Methods for calculating these least-squares solutions are discussed later in this paper.)

Least-squares solutions of systems with unit-norm row vectors will be referred to as equally weighted least-squares solutions. They are defined formally as follows:

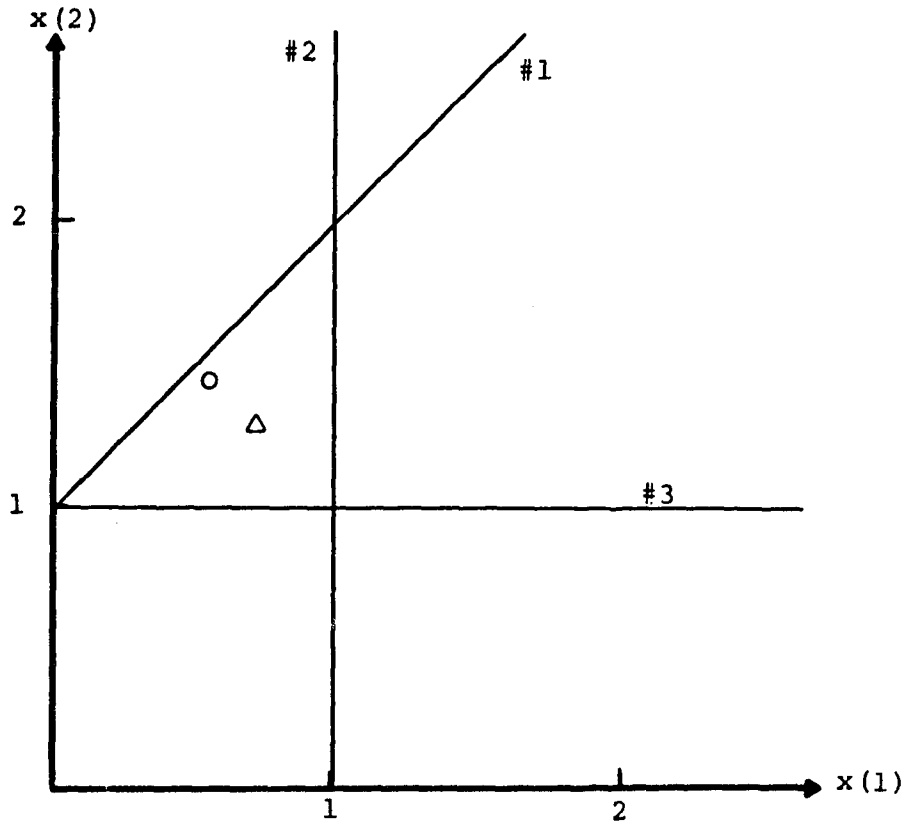


Figure 3. Examples of weighted least squares solutions.
o - The least squares solution of (2.15).
Δ - The least-squares solution of (2.16).

DEFINITION 2.1. Let $Ax = b$ be any real system of linear equations for which $A \in E^{m\times n}$ and the rows of A are non-zero. Let a_i denote the i'th row-vector of A, let $A_o \in E^{m\times n}$ denote the matrix having $\frac{a_i}{\|a_i\|}$ as its i'th row vector, and let $b_o \in E^m$ denote the vector having $\frac{b(i)}{\|a_i\|}$ as its i'th element, where $1 \leq i \leq m$. Then the least-squares solutions of $A_o x = b_o$ are said to be <u>equally weighted least-squares solutions</u> of $Ax = b$.

Systems for which the $\|a_i\|$ vary widely are known to be ill-posed [56], so

$$A_o x = b_o \tag{2.17}$$

is often less sensitive to round-off errors made during its solution than the original system (2.1). Equation (2.17) can still be highly ill-posed, however. Equally weighted least-squares solutions have the geometric significance of minimizing $\sum_{i=1}^{m} \|p_i(x)\|^2$, where $p_i(x)$ is the vector of minimum norm from the point x to the i'th hyperplane of (2.1) and $\|p_i(x)\|$ is the Euclidean distance of x from the i'th hyperplane. This is readily seen from the following theorem if it is remembered that equally weighted least-squares solutions of (2.1) minimize $\|b_o - A_o x\|$ and hence minimize $\|b_o - A_o x\|^2$.

THEOREM 2.1. Let $Ax = b$ be any real system of linear equations for which the rows of A are non-zero, let b_o and A_o be as defined in Definition 2.1, and let $p_i(x)$ denote the vector of minimum norm from the point x to the i'th hyperplane of $Ax = b$ (i.e., $p_i(x)$ is the vector of minimum norm such that $x + p_i(x)$ satisfies the i'th equation of $Ax = b$). Then

$$\sum_{i=1}^{m} \|p_i(x)\|^2 = \|b_o - A_o x\|^2.$$

(A proof is included in Appendix 2.)

2.4. Non-Uniqueness and the Minumum-Norm Constraint

The material presented here on non-uniqueness and the use of the minimum-norm constraint to resolve it is well-known, and treatments of it may be found in [1, 2, 35, 41, 48].

The least-squares solutions of (2.1) are given by

$$x_{ls} = A^{\dagger}b + x_1, \tag{2.18}$$

where $A^{\dagger}$ is the pseudo-inverse of A, and x_1 may be any arbitrary vector in $N(A)$. ($A^{\dagger}$ is defined and discussed in Appendix 1.) For any x_{ls}, $A^{\dagger}b$ and x_1 are the components of x_{ls} in $N^{\perp}(A)$ and $N(A)$ respectively. If $r(A) = n$

then N_A is trivial (i.e., $N(A) = \{\underline{0}\}$) and (2.1) has the unique least-squares solution

$$x_{ls} = A^{\dagger}b. \tag{2.19}$$

If $r(A) < n$, then x_1 is arbitrary within $N(A)$ and x_{ls} is not unique. This is because

$$\|b - Ax_{ls}\| = \|b - A(A^{\dagger}b + x_1)\| = \|b - AA^{\dagger}b\|, \tag{2.20}$$

and therefore x_1 has no effect on how well x_{ls} satisfies the equation in the least-squares sense. (I.e., the equations contain no information about the component of x in $N(A)$.)

From a geometric viewpoint, consistent systems have non-unique solutions when the hyperplanes intersect at an infinite number of points. For $n = 3$ the hyperplanes are planes, and a system of several planes intersecting in a line (like a revolving door) would have non-unique solutions. A two-dimensional example of non-uniqueness is shown in Figure 4a. It represents a system for which all the equations are multiples of each other, and hence define the same hyperplane. Each point on that hyperplane is a solution, and the minimum-norm solution is circled. (The norm of a solution is simply its distance from the origin.) Figure 4b represents

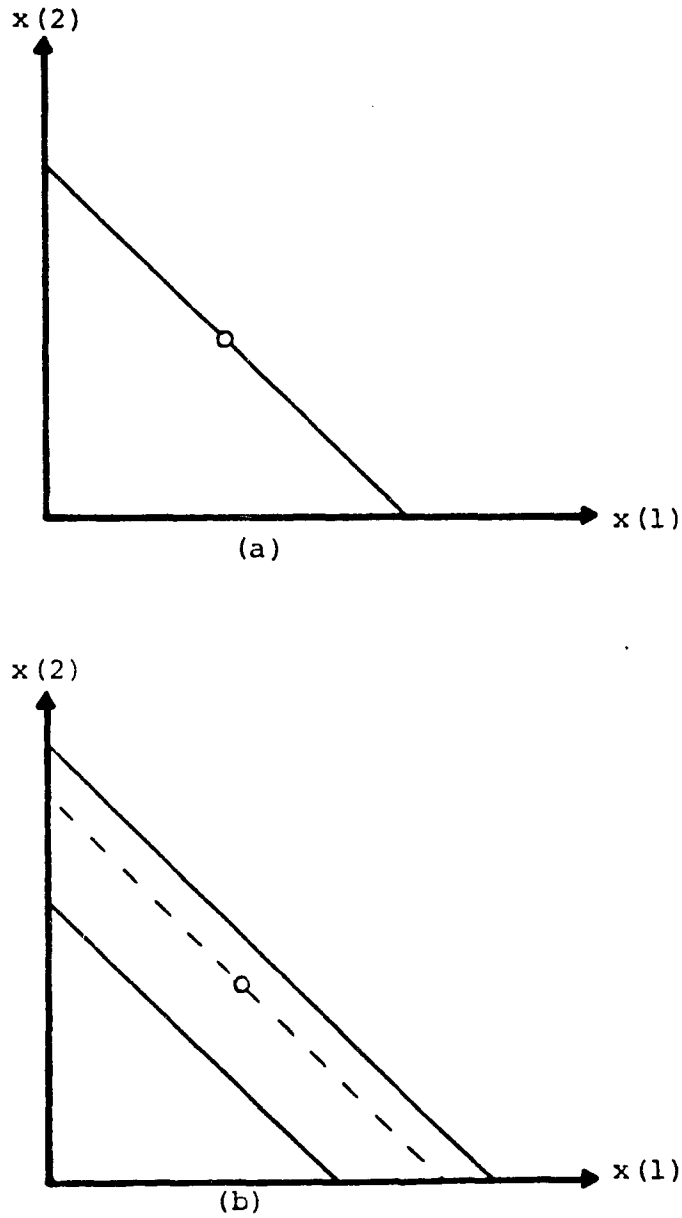


Figure 4. Examples of non-uniqueness.

(a) A consistent system with non-unique solutions .
(b) An inconsistent system with non-unique least-squares solutions.

an inconsistent system with non-unique least-squares solutions. The least-squares solutions are shown as a dashed line between the two hyperplanes and parallel to them. The minimum-norm least-squares solution is circled.

If x_{ls} is non-unique, most applications require that a single solution be selected. This may be viewed as the two separate problems of finding $A^{\dagger}b$ and selecting x_1. If all that is required is a solution which satisfies the equations in the least-squares sense, then it makes no difference which solution is selected, and x_1 may be set arbitrarily to any vector in $N(A)$. Often, however, there is only one acceptable solution, and the mathematical non-uniqueness of the solution merely reflects the failure of the equations to describe the actual problem completely. In such cases it is not adequate for the solution to satisfy the equations, and it is necessary to determine x_1 by the application of an appropriate additional constraint on x_{ls}. This is often done by selecting the unique least-squares solution having minimum norm. For the remainder of this paper, the minimum-norm least-squares solution of (2.1) will be denoted by x and will be referred to simply as "the solution" of (2.1). Because

$$\|x_{ls}\| = (\|A^{\dagger}b\|^2 + \|x_1\|^2)^{1/2}, \tag{2.21}$$

x is obtained by constraining x_{ls} to $N^{\perp}(A)$ (i.e., setting x_1 to $\underline{0}$) and

$$x = A^{\dagger}b. \tag{2.22}$$

The minimum-norm constraint is used very widely, due to its simplicity and its physical significance for many problems. (It often serves as a minimum-energy constraint.) Other constraints can be applied to obtain unique solutions to non-unique problems, however. Two examples are the maximum-entropy constraint [6] and the minimum-variance constraint [25]. The entropy of x is defined by

$$s = \frac{-1}{\ell n\, n} \sum_{j=1}^{n} \frac{x(j)}{x_{av}} \cdot \ell n\left(\frac{x(j)}{x_{av}}\right), \tag{2.23}$$

and the variance of x is defined by

$$V = \sum_{j=1}^{n} (x(j) - x_{av})^2, \tag{2.24}$$

where

$$x_{av} = \frac{1}{n} \sum_{j=1}^{n} x(j). \tag{2.25}$$

When the elements of x are non-negative, the minimum-variance constraint has been shown to be equivalent to the

minimum-norm constraint [25]. There are also other schemes for resolving non-uniqueness [29, 53].

2.5 Perturbation Theory and the Suppression of Error Due to Noise

When a system of linear equations must be solved in the presence of noise, the actual goal is not to find x, the solution of perturbed system $Ax = b$. It is, rather, to approximate x_u, the solution of the unperturbed system

$$A_u x_u = b_u, \tag{2.26}$$

given the perturbed values of A and b. (Recall that A_u and b_u were defined in Section 2.1 by equations (2.9) and (2.10).) Even when noise levels are very low, it is possible for x to be very different from x_u. Systems for which small perturbations of A or b can cause large perturbations of x are said to be ill-posed and their solutions are said to be unstable. Here, A is termed ill-conditioned when x is unstable with respect to perturbation of b. (A more precise definition of ill-conditioning will be given shortly.) Square, non-singular matrices which are ill-conditioned are often called near singular. Whatever terminology is used to describe them, ill-posed systems perturbed by noise present a special problem. Even when accurately solved, their solutions can be useless due to

the effects of noise unless special precautions are taken. For reasons to be discussed later in this chapter, extremely ill-posed systems arise routinely in many applications. By examining the effects of perturbation of b and A on x for such systems, means of suppressing these effects can be studied.

It is shown in Section 2.5.1 that if A is ill-conditioned then it can be approximated by any of a number of well-conditioned matrices A', for which

$$r(A') < r(A) \tag{2.27}$$

and

$$N^{\perp}(A') \perp N^{\perp}(\delta A'), \tag{2.28}$$

where

$$\delta A' = A' - A. \tag{2.29}$$

A contribution to perturbation theory is made in Section 2.5 by showing that most of the potentially massive error resulting from perturbation of an ill-posed system is concentrated in $N(A')$. (This is demonstrated by mathematically

rigorous theorems for perturbation of b and by non-rigorous arguments supported by empirical observations for perturbation of A.) The ability to estimate not only the size of the noise error but its location as well can aid in the development of new means of suppressing noise error and shed new light on existing algorithms. Specifically, it is shown that replacement of A by A' prior to solving the system reduces error by applying a stability constraint to the solution. The use of a more well-conditioned A' suppresses noise error more effectively, at a cost of extracting less information from the equations. The properties of the A' matrices are discussed in a general manner in Chapter 2. Various means of finding and employing different A' matrices are presented in Chapter 3.

2.5.1. Perturbation of b

The stability of x with respect to perturbation of b will now be examined. In this section it will be assumed that A is unperturbed, and δx will denote the perturbation of x resulting from perturbation of b by the addition of δb. Because

$$x = x_u + \delta x = A^{\dagger}(b_u + \delta b) = x_u + A^{\dagger}\delta b, \tag{2.30}$$

$$\delta x = A^{\dagger}\delta b. \tag{2.31}$$

Clearly a least upper bound on $\|\delta x\|$ is given by

$$\|\delta x\| \leq \|A^{\dagger}\|\|\delta b\|. \tag{2.32}$$

An upper bound on relative error is given by [48]

$$\frac{\|\delta x\|}{\|x_u\|} \leq \kappa(A)\frac{\|\delta b\|}{\|b_u\|}, \tag{2.33}$$

where

$$\kappa(A) = \|A^{\dagger}\|\|A\|. \tag{2.34}$$

$\kappa(A)$ serves as a condition number of A, and in this paper A will be termed ill-conditioned when $\kappa(A)$ is large compared to 1.

It will be shown that A is ill-conditioned if and only if it can be approximated by some matrix A' of lower rank. When this is the case A is said to have weak rank, and if certain conditions are met $r(A')$ may be referred to as the pseudo-rank of A. These points are presented formally by the following definitions and theorems:

DEFINITION 2.2. If A is a non-zero matrix and $\epsilon > 0$ then A is said to have <u>pseudo-rank with tolerance ϵ</u>, denoted $r_p(A,\epsilon)$, if $r_p(A,\epsilon)$ is the smallest positive

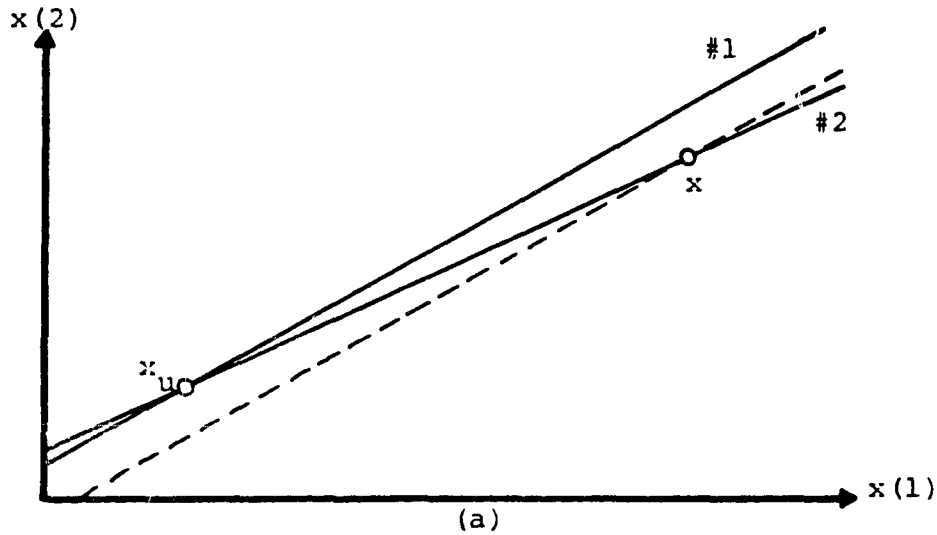
integer such that there is a matrix A' with $\|A' - A\| \leq \epsilon$ and $r(A') = r_p(A,\epsilon)$.

DEFINITION 2.3. If A is a non-zero matrix and $\epsilon > 0$ then A is said to have <u>weak rank with tolerance ϵ</u> if $r_p(A,\epsilon) < r(A)$. A is said to have <u>strong rank with tolerance ϵ</u> if $r_p(A,\epsilon) = r(A)$.

The definitions of strong and weak rank are original, but the term "pseudo-rank" has appeared in the literature previously in a more general sense, with no specification of tolerance [35].

THEOREM 2.2. For any non-zero matrix A and any constant $\kappa_0 \geq 1$, $\kappa(A) \geq \kappa_0$ if and only if A has weak rank with tolerance $\frac{\|A\|}{\kappa_0}$. (A proof is included in Appendix 2.)

Theorem 2.2 is illustrated by the systems shown in Figure 5. The dashed lines represent the positions of the hyperplanes defined by the first equations after small perturbations of b. The system depicted in Figure 5a has an ill-conditioned A matrix, and the perturbed solution x is far from the unperturbed solution x_u. For the system shown in Figure 5b, A is well-conditioned and x is near x_u. The hyperplanes of Figure 5a are at nearly the same



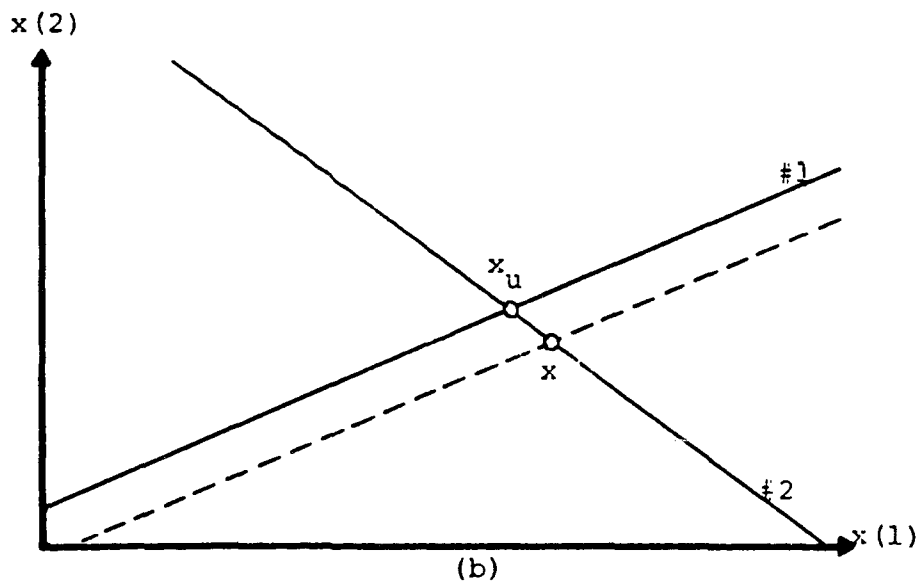


Figure 5. An illustration of Theorem 2.2.

(a) A system for which A is ill-conditioned.
(b) A system for which A is well-conditioned.

angle from the horizontal, so the row vectors of A (which determine these angles) are nearly multiples of each other. Therefore, slight perturbations of either or both of the row vectors of A could make them exact multiples of each other. Such perturbations of A would reduce r(A) from 2 to 1, and give parallel hyperplanes. It is clear, therefore, that the ill-conditioned matrix of Figure 5a can be approximated by any of an infinite number of matrices of reduced rank. These matrices correspond to pairs of parallel hyperplanes with angles near the angles of the original hyperplanes. The well-conditioned matrix of Figure 5b can not be approximated by a matrix of reduced rank, because it would obviously require a large perturbation of A to make its hyperplanes parallel.

At this point it is convenient to make use of the fact that every matrix A has a singular value decomposition (SVD) of the form

$$A = U\Sigma V^T. \tag{2.35}$$

U and V are orthogonal matrices, m x n matrix Σ satisfies

$$\Sigma(i,j) = \begin{cases} \sigma_i & \text{if } i = j \text{ and } i \leq r(A) \\ 0 & \text{otherwise} \end{cases} \tag{2.36}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$, and

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{r(A)} > 0. \tag{2.37}$$

(The SVD is briefly reviewed in Appendix 1.) It is easily shown that

$$\kappa(A) = \frac{\sigma_1}{\sigma_{r(A)}} \tag{2.38}$$

(Equation (A1.24) in Appendix 1). Therefore, if A is ill-conditioned then it has some singular values which are small relative to σ_1. It follows that A can be approximated by a well-conditioned matrix whose SVD is obtained from the SVD of A by replacing the small singular values with zeros. This is demonstrated by Theorem 2.3.

THEOREM 2.3. Let matrix A have the singular value decomposition $U\Sigma V^T$ and let A_k denote $U\Sigma_k V^T$ for $1 \leq k \leq r(A)$, where Σ_k is the matrix obtained from Σ by replacing all but the k largest singular values with zero. Then, if δA_k denotes $A_k - A$,

1) $\kappa(A_k) = \frac{\sigma_1}{\sigma_k}$,

2) $\|\delta A_k\| = \sigma_{k+1}$,

3) $r(A_k) = k$, and

4) $N^{\perp}(\delta A_k) \perp N^{\perp}(A_k)$.

(A proof is included in Appendix 2.)

Clearly, an ill-conditioned matrix A can be approximated by some relatively well-conditioned matrix A_k of rank k. Smaller k gives a more well-conditioned A_k matrix which less accurately approximates A. There are other matrices A' in addition to the A_k which are well-conditioned, are near A, have reduced rank, and satisfy (2.28). (It is easily shown that matrices of the form

$$A_k' = U\Sigma_k'V^T \tag{2.39}$$

satisfy these conditions, where Σ_k' is a matrix formed from Σ_k by slightly perturbing its singular values. If the largest and smallest singular values are not excessively perturbed then

$$\kappa(A_k') \cong \kappa(A_k) \tag{2.40}$$

by (2.38). The other conditions follow easily, as in the proof of Theorem 2.3.) Matrices fitting this description may be broadly referred to here as A' matrices, and they will prove to be very useful. It should be noted that the A_k matrices defined in Theorem 2.3 are as close to A as any matrices of their rank, in the sense that [18]

$$\|A - A_k\|_F = \min_{r(B)\leq k} \|A - B\|_F, \tag{2.41}$$

where $B \in E^{m \times n}$. It is common for systems to have extremely ill-conditioned A matrices which are extremely close to well-conditioned A' matrices. (This point will be discussed in detail in Section 2.5.2.)

Consider a linear system for which A is highly ill-conditioned and b is perturbed, so that

$$Ax = b_u + \delta b. \tag{2.42}$$

The fact that A is near some relatively well-conditioned A' matrix will make it possible to partition E^n into two subspaces S_e and S_c, so that

$$E^n = S_e \oplus S_c, \tag{2.43}$$

where

$$S_e = N^{\perp}(A') \tag{2.44}$$

and

$$S_c = N(A'). \tag{2.45}$$

(The reason for this choice of symbols will be made clear shortly.) The subscripts e and c will be used to denote the components of a vector in S_e and S_c, respectively. The solution x_u of the unperturbed system

(2.46) $$Ax_u = b_u$$

may be expressed as

(2.47) $$x_u = x_{ue} + x_{uc},$$

so that the problem of approximating x_u may be viewed as the two separate problems of approximating x_{ue} and x_{uc}. The solution x of the perturbed system (2.42) may similarly be expressed as

(2.48) $$x = x_e + x_c.$$

Therefore

(2.49) $$\|\delta x\| = \|x - x_u\| = (\|x_e - x_{ue}\|^2 + \|x_c - x_{uc}\|^2)^{1/2}.$$

In the following theorem it is demonstrated that $\|x_e - x_{ue}\|$ is relatively small, so most of the massive error associated with solving a system with an ill-conditioned matrix and noisy data is concentrated in S_c. The instability of x_c is also demonstrated.

THEOREM 2.4. Let $Ax = b$ be a system of real, linear equations such that $A \in E^{m\times n}$, and let $A' \in E^{m\times n}$

be any matrix such that $r(A') < r(A)$ and $N^{\perp}(\delta A') \perp N^{\perp}(A')$, where $\delta A'$ denotes $A' - A$. Further, let S_e and S_c denote $N^{\perp}(A')$ and $N(A')$, respectively. Then,

1) for any vector $\delta b \in E^m$ if δx denotes $A^{\dagger}\delta b$ (i.e., δx is the perturbation of x which would result from perturbation of b by δb), and δx_e and δx_c denote the projections of δx onto S_e and S_c, respectively, then

$$\|\delta x\| \leq \frac{\kappa(A')}{\|A'\|}(\|\delta b\| + \|\delta A'\|\|\delta x_c\|).$$

Furthermore,

2) for any vector $\delta x \in S_c \cap N^{\perp}(A)$, there is a vector δb (specifically, $A\delta x$) such that $A^{\dagger}\delta b = \delta x$ (i.e., perturbation of b by δb would perturb x by δx) and

$$\|\delta x\| \geq \frac{\|\delta b\|}{\|\delta A'\|}$$

Part 1 of Theorem 2.4 demonstrates that x_e is relatively stable with respect to perturbation of b, with more stability when A' is more well-conditioned and nearer to A. (If $\|\delta x_c\|$ is very large, $\|\delta x_e\|$ is still small compared to $\|\delta x_c\|$ if $\frac{\kappa(A')\|\delta A'\|}{\|A'\|}$ is sufficiently small.) Therefore,

(2.50) $$x_e \cong x_{ue},$$

and x_{ue} can be accurately determined from the equations; hence the subscript e for "equations".

Part 2 of Theorem 2.4 clearly demonstrates that x_c is unstable with respect to perturbations of b, with less stability for A' nearer to A. (δx is restricted to $N^{\perp}(A)$ because, as mentioned in Section 2.4, x is restricted to $N^{\perp}(A)$ by the minimum-norm constraint.) If A' is sufficiently close to A, then x_{uc} can be approximated more accurately by applying the minimum-norm constraint than by using the equations (i.e.,

(2.51) $$\|x_{uc} - \underline{0}\| < \|x_{uc} - x_c\|);$$

hence the subscript c for "constraint".

The minimum-norm constraint is applied to S_c by constraining the solution to S_e. This is accomplished by finding the vector $x' \in S_e$ of minimum norm for which the residual $\|b - Ax'\|$ is minimal. Because

(2.52) $$\|b - Ax'\| = \|b - (A - A')x' + A'x'\|$$

$$= \|b - A'x'\|,$$

x' is the minimum-norm, least-squares solution of

(2.53) $$A'x' = b.$$

Therefore,

(2.54) $$x' = A'^{\dagger}b.$$

Noise error may thus be reduced by replacing A with A' or, equivalently, by perturbing A by the addition of $\delta A'$, where

(2.55) $$\delta A' = A' - A.$$

Various means of calculationg A' and x' from given values of A and b will be presented in Chapter 3.

While standard minimum-norm, least-squares approaches constrain the solution to $N^{\perp}(A)$ (which is the subspace in which the component of the solution is uniquely determined), the use of (2.54) constrains the solution to $N^{\perp}(A')$ (which is a subspace in which the solution is both uniquely determined and stable). Thus, the perturbation of A by $\delta A'$ in effect constitutes the application of a stability constraint, at the cost of violating the least-squares constraint. Experience strongly indicates that even though $\|x - x'\|$ may be very large, the residual $\|b - Ax'\|$ is near the minimal value $\|b - Ax\|$ if $\|\delta A'\|$ is sufficiently small.

Figure 6 gives an example of the suppression of noise error by the perturbation of A. Figure 6a represents

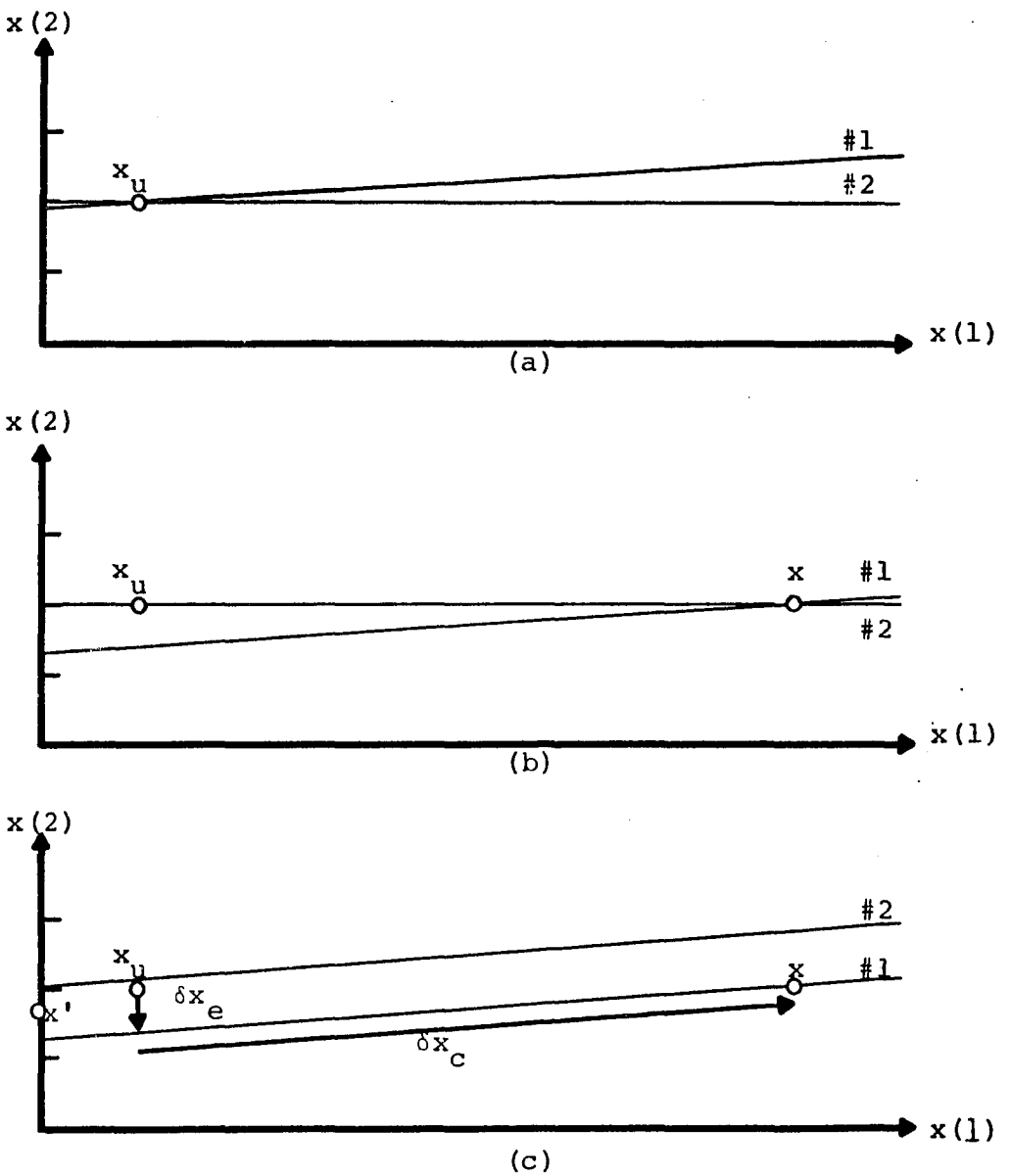


Figure 6. The suppression of error due to perturbation of b.

(a) The unperturbed system $Ax_u = b_u$

(b) The perturbed system $Ax = b$.

(c) The stabilized system $A'x' = b$.

the unperturbed, ill-posed system

$$Ax_u = b_u. \tag{2.56}$$

In Figure 6b, hyperplane number 1 has been moved laterally. This corresponds to the perturbed system

$$Ax = b = b_u + \delta b. \tag{2.57}$$

Although the hyperplane was moved only slightly, the perturbed solution, x, is quite far from x_u. In Figure 6c the matrix A has been replaced by an A' matrix having the properties discussed earlier. It is easily verified that in two dimensions any matrix which approximates A and gives parallel hyperplanes is an A' matrix. For this example, hyperplane 1 was left unchanged and A was altered to make hyperplane 2 parallel to hyperplane 1 (i.e., the second row vector of A was replaced with its projection onto the first row vector). The resultant system clearly has greatly improved stability with respect to perturbation of b, and the resultant minimum-norm least-squares solution, x', is much closer to x_u than x is. The solutions x_u and x from Figures 6a and 6b are shown in Figure 6c, and the perturbation error

(2.58) $$\delta x = x - x_u$$

is shown broken into its components in S_e and S_c. (Note that the hyperplanes are perpendicular to their corresponding row vectors and S_e is the row space of A'.) Obviously,

(2.59) $$\|\delta x_e\| << \|\delta x_c\|,$$

as predicted by Theorem 2.4.

Although different $\delta A'$ matrices are selected by the different algorithms to be presented in Chapter 3, the selection of $\|\delta A'\|$ can be discussed here in a general manner. To improve the stability of x' with respect to perturbations of b (i.e., to reduce $\kappa(A')$) generally requires a decrease in $r(A')$ and an accompanying increase in $\|\delta A'\|$, as discussed in Section 2.5.1. While the perturbation of A by $\delta A'$ may be viewed as a source of error in the solution, it serves to reduce the sensitivity of the solution to perturbation of b, thus reducing the overall error. From the definitions of S_e and S_c it is easily seen that

(2.60) $$\dim(S_e) = r(A')$$

and

(2.61) $$\dim(S_c) = n - r(A').$$

Clearly, therefore, as $r(A')$ is reduced to improve stability and $\|\delta A'\|$ increases, the equations play a smaller role and the minimum-norm constraint plays a larger role in determining the solution x'. As $\|\delta A'\|$ continues to increase, there comes a point at which the benefit of improved stability of the solution with respect to perturbation of b begins to be outweighed by the cost of increased perturbation of the solution by $\delta A'$. Thus it is clear that each linear system has an optimal value of $\|\delta A'\|$ for which the net error in the solution is minimal. It is also clear that this optimal value is strongly dependent on the level of noise in b, with larger $\|\delta b\|$ requiring a smaller $\kappa(A')$ and a larger $\|\delta A'\|$. If $\|\delta A'\|$ is chosen too small, then the resultant solution may have excessive error due to noise. If $\|\delta A'\|$ is chosen too large, however, all the useful information may not be extracted from the equations.

2.5.2. Perturbation of A

The stability of x with respect to perturbation of A will now be considered. In this section b is assumed to be unperturbed, so that (2.1) may be re-expressed as

$$(A_u + \delta A)(x_u + \delta x) = b, \tag{2.62}$$

where here δx is the perturbation of x resulting from the perturbation of A by δA. For systems with full rank (i.e., $r(A) = n$), an upper bound on relative error for small perturbations of A is presented in Theorem 2.5 [48]:

THEOREM 2.5. Let δA_1 and δA_2 denote the matrices having as column vectors the projections of the corresponding column vectors of δA onto $R(A)$ and $R^{\perp}(A)$, respectively; and let b_1 and b_2 denote the projections of b onto $R(A)$ and $R^{\perp}(A)$, respectively. If $r(A) = n$ and $\|\delta A_1\| \leq \frac{1}{2\|A^{\dagger}\|}$, then

$$\frac{\|\delta x\|}{\|x_u\|} \leq \kappa(A) \frac{\|\delta A_1\|}{\|A\|} + 4\kappa(A)^2 \frac{\|\delta A_2\|}{\|A\|} + 8\kappa(A)^3 \frac{\|\delta A_2\|}{\|A\|}.$$

$\kappa(A)$ is as defined previously by Equation (2.34). For square, non-singular matrices,

$$\delta A_1 = \delta A \tag{2.63}$$

and

$$\delta A_2 = \underline{0}, \tag{2.64}$$

so only the first term remains.

There are many applications (including most tomography problems) for which Theorem 2.5 does not apply, because $r(A) < n$. To investigate the effect of perturbation of A for such systems, it is helpful to observe that

$$\delta x = (x_u + \delta x) - x_u \tag{2.65}$$

$$= (A_u + \delta A)^{\dagger} b - A_u^{\dagger} b$$

$$= ((A_u + \delta A)^{\dagger} - A_u^{\dagger}) b.$$

Clearly, therefore, a system is unstable with respect to perturbations of A when a small perturbation of A causes a large perturbation of its pseudo-inverse. The vector b also plays a key role in determining this stability. It will be shown that for systems with $r(A) < \min(m,n)$ an infinitesimally small perturbation δA can increase the rank of A and cause an unbounded increase in the norm of the pseudo-inverse. ($\min(m,n)$ denotes the smaller of m and n.) Therefore, such systems are infinitely unstable with respect to perturbation of A. This same δA also causes infinite instability with respect to perturbation of b, as demonstrated by Theorem 2.6.

THEOREM 2.6. If $r(A) < \min(m,n)$, then for any $\epsilon > 0$ there is a matrix δA such that:

1) $\|\delta A\| = \epsilon$,

2) $N^{\perp}(\delta A) \perp N^{\perp}(A)$

3) $r(A + \delta A) > r(A)$,

4) $\|(A + \delta A)^{\dagger}\| \geq \frac{1}{\epsilon}$,

5) $\kappa(A + \delta A) \geq \frac{1}{\epsilon}(\|A\| - \epsilon)$.

(A proof is included in Appendix 2.)

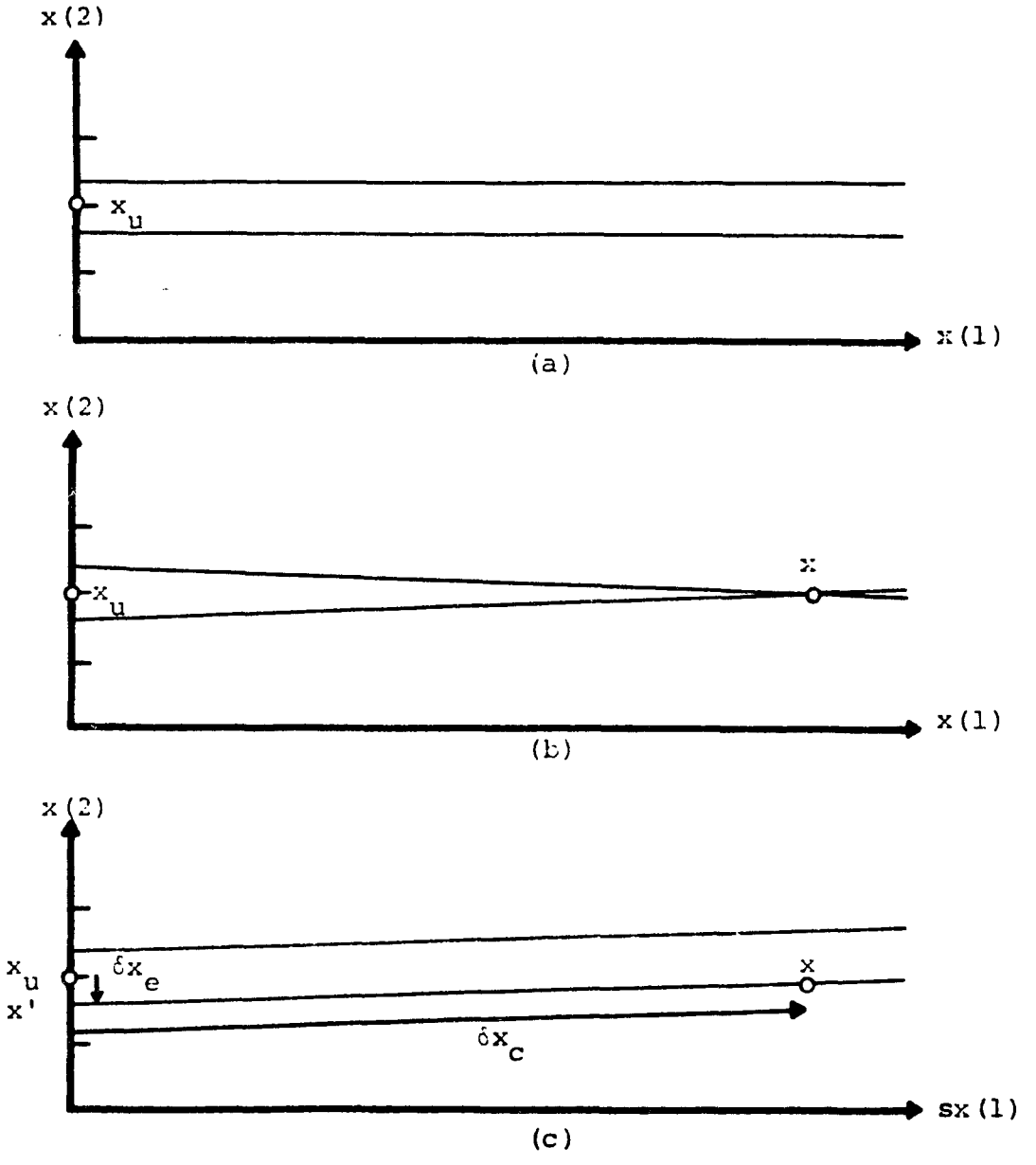
An unusual tiwst of this phenomenon is that smaller perturbations of A tend to cause larger perturbations of x and larger increases in $\kappa(A)$. This explains why in practice matrices with $r(A) < \min(m,n)$ are very often extremely ill-conditioned. A well-conditioned matrix A_u with $r(A_u) < \min(m,n)$ can easily be rendered extremely ill-conditioned by extremely small perturbations, such as the ever-present computer round-off errors. The use of double precision merely tends to produce an even more highly ill-conditioned A matrix. Such A matrices are extremely close to A' matrices satisfying equations (2.27) and (2.28). (The distance $\|A' - A\|$ from A to the nearest A' matrix is $\sigma_{r(A)}$ (see Theorem 2.3 and Equation (2.41)). However, it is evident from Equation (2.38) that $\sigma_{r(A)}$ is extremely small. (Recall that $\sigma_1 = \|A\|$.)) Experience indicates that if

A_u is well-conditioned and A is close to A_u, then A is close to an A' matrix which is well-conditioned in addition to satisfying equations (2.27) and (2.28). These points are illustrated by an example in Figure 7.

Figure 7a represents a system with parallel hyperplanes for which $m = n = 2$ and $r(A) = 1$. The minimum-norm, least-squares solution x_u is shown. Because the hyperplanes are parallel, any non-zero perturbation of A will change the angles of the hyperplanes from the horizontal and cause them to intersect at some point x. (The only exception to this is the unlikely case when $r(A)$ is not increased by the perturbation and the hyperplanes remain parallel.) Figure 7b shows the system of Figure 7a with A perturbed by the addition of some small matrix δA. It is clear that a smaller $\|\delta A\|$ tends to leave the hyperplanes more nearly parallel, giving an x that is farther from x_u. This corresponds to a larger $\|\delta x\|$, where, as before,

$$\delta x = x - x_u. \tag{2.66}$$

It is also clear that a smaller $\|\delta A\|$ tends to produce a system with less stability with respect to perturbation of b. (Recall that perturbation of b moves the hyperplanes laterally.) It is easily seen that computer round-off



Firqure 7. A system with infinite instability with respect to perturbation of A.

(a) The unperturbed system $A_u x_u = b$.

(b) The perturbed system $Ax = b$.

(c) The stabilized system $A'x' = b$. (x_u and x' are very close together and are shown here as a single point.)

errors, which often leave the first 10 significant figures unchanged, would cause an immense perturbation of x and produce an extremely ill-conditioned A matrix.

In Figure 7c, the matrix of Figure 7b has been replaced by an A' matrix, as discussed in the previous section. The resultant minimum-norm, least-squares solution, x', is not far from x_u. As was done for Figure 6, the perturbation error δx is shown broken into its components in S_e and S_c. For this example, most of the error due to perturbation of A was contained in S_c.

An upper bound on error due to perturbation of A is given by Theorem 2.7.

THEOREM 2.7. Let $A, B \in E^{m\times n}$; let $b \in E^m$; and let δx denote $B^\dagger b - A^\dagger b$. (Note that δx is the perturbation of the minimum-norm least-squares solution of $Ax = b$ which results when A is replaced by B.) Then

$$\|\delta x\| \leq (\|A^\dagger\| + \|B^\dagger\|)\|b\|.$$

PROOF.
$$\begin{aligned}\|\delta x\| &= \|(B^\dagger - A^\dagger)b\| \\ &\leq \|B^\dagger - A^\dagger\|\|b\| \\ &\leq (\|B^\dagger\| + \|A^\dagger\|)\|b\|.\end{aligned}$$

□

This bound is not low enough in general to establish that x' approximates x_u. (Recall that x' and x_u are the solutions of $A'x = b$ and $A_u x = b$ respectively.) However, it does show that if A_u and A' are well-conditioned then the error $\|x' - x_u\|$ can not attain the potentially massive size predicted by Theorem 2.6. (Recall that the errors described by Theorem 2.6 are large only if the perturbed matrix is ill-conditioned.) It is tempting to speculate, therefore, that most of the massive error $\|x - x_u\|$ predicted by Theorem 2.6 is conveniently confined to S_c, along with most of the error due to perturbation of b. (Such speculation is encouraged by the example in Figure 7.) If this is the case, then this error is easily suppressed by replacing A with an appropriate A' matrix, as in the previous section.

Experience verifies that if A' satisfies the conditions outlined in Section 2.5.1, then x' (defined by Equation (2.54)) is in fact stable with respect to perturbations of both b and A. (The systems referred to are those of ROG. Solutions of these systems are presented in Chapter 4.) It should be noted, however, that in general $r(A')$ may be less than $\min(m,n)$. As a result, x' may be infinitely unstable with respect to any further perturbation of A' which would increase its rank (Theorem 2.6).

Therefore, once A has been replaced by A' it is necessary to prevent any subsequent perturbation from increasing its rank and rendering it ill-conditioned.

Although, as noted above, smaller perturbations of A tend to cause larger errors, these errors can be more effectively suppressed because A' is nearer to A. Therefore, in accordance with intuition, precision in A is, in fact, desirable.

Not as much is known about stability with respect to perturbation of A as has been discovered about stability with respect to perturbation of b. However, those problems caused by perturbation of A which have been identified here are solved by replacing A with an appropriate A' matrix. Extensive experience indicates that if A' is properly selected as discussed previously, then

$$x' = A'^{\dagger}b \tag{2.67}$$

is a useful solution of $Ax = b$ for a very wide range of problems. Various specific means of calculating A' and x' are presented in Chapter 3.

CHAPTER 3

ALGORITHMS FOR THE SOLUTION OF ILL-POSED SYSTEMS OF LINEAR EQUATIONS IN THE PRESENCE OF NOISE

In this chapter, several algorithms for calculating stabilized, minimum-norm, least-squares solutions of ill-posed or rank-deficient systems are considered. Each applies a stability constraint to the solution by approximating the system matrix A with some well-conditioned matrix A', as discussed in Chapter 2. Two of these algorithms are the singular value decomposition method and the Lagrangian multiplier method. While these are both widely used and effective methods, they each require large amounts of computer memory when applied to very large problems. Furthermore, neither method can easily apply inequality constraints to the solution. These factors make them less than ideal for use on the ROG problems to be solved in Chapter 4. Therefore, emphasis here is placed on a relatively new class of iterative methods related to the Projection Method (PM).

These projection algorithms are known to be efficient for very large systems and easily capable of applying inequality constraints [10]. Stabilized solutions are obtained by stopping these algorithms before they completely converge. This means of stabilization has not been well understood in

the past, and various ad hoc stopping criteria have been employed to determine the optimal point at which to stop the algorithms. A contribution to the understanding of the projection algorithms is made here by demonstrating that stopping prior to convergence is approximately equivalent to solving an altered system

$$A'x = b, \tag{3.1}$$

where A' is a well-conditioned matrix near A as discussed in Chapter 2.

Based on this new understanding of the effect of stopping, new stopping criteria are proposed which are superior to those in current use. It is also demonstrated here that the Simultaneous Projection Method (SPM), an algorithm closely related to PM, suppresses noise error more effectively than PM, although it is much slower. An additional advantage of SPM over PM is that it can be proven to converge to an equally weighted least-squares solution. (A proof of this will be presented.) The convergence of PM is not well-defined for inconsistent systems, although PM does converge (in a sense to be discussed shortly) to an approximation of an equally weighted least-squares solution.

A modification of SPM is introduced here which has increased speed, although it is still slower than PM. It

will be referred to as the Fast Simultaneous Projection Method (FSPM).

It is concluded that both PM and FSPM are valuable algorithms for certain applications. FSPM suppresses noise more effectively and gives more accurate solutions, but PM is faster and simpler to implement. The performance of PM, FSPM, and the Lagrangian multiplier method will be compared in Chapter 4 for an ROG problem.

3.1. The Singular Value Decomposition Method

A review of the singular value decomposition is included in Appendix 1. More comprehensive treatments of this well-known method may be found in [18, 19, 35, 41, 48].

$A^{\dagger}$ is readily calculated from the SVD of A as follows:

$$A^{\dagger} = V\Sigma^{\dagger}U^{T}, \tag{3.2}$$

where $\Sigma^{\dagger}$ is calculated from Σ simply by replacing each singular value with its reciprocal. From $A^{\dagger}$, x is easily obtained by Equation (2.22). It is clear from prior discussion that the stability of x may be improved by replacing the small singular values of A with zero before calculating $A^{\dagger}$. (This was proposed by Golub and Kahan [18] and is well-known.) Neglecting more singular values more effectively

suppresses noise error but allows less information to be extracted from the equations. The optimal level at which to truncate singular values depends on the individual problem and the noise level. It may often be estimated adequately by trial and error.

Although the singular value decomposition is a powerful theoretical and computational tool, it does have drawbacks for some systems. If A is very large and sparse, then A^TA will not be sparse in general, and the required eigenvalue-eigenvector problem may become quite cumbersome. Furthermore, inequality constraints are not easily applied by this method.

3.2. The Lagrangian Multiplier Method

Instead of seeking x, the vector of minimum norm for which $\|b - Ax\|$ is minimal, this widely known method seeks the vector x' of minimum norm for which

$$\|b - Ax'\| \leq \epsilon, \tag{3.3}$$

where $\epsilon > 0$ is an estimate of the degree of inconsistency of the equations. (Usually ϵ is simply an estimate of $\|\delta b\|$, the noise level of the data.) For an ill-posed system, x' can be vastly more stable than x while still satisfying the equations within the uncertainties in the

data, so noise error can be dramatically reduced.

To find x' [53], the normal equations (2.13) are solved by any convenient means after first being perturbed as follows:

$$(A^TA + \gamma I)x' = A^Tb. \tag{3.4}$$

Here, I is the identity matrix and γ is the Lagrangian multiplier, a small positive number which must be selected for each application. (γ is also referred to in the literature as a regularizing parameter or a ridge parameter.) Equation (3.4) is non-singular for $\gamma > 0$ [1], so it always has a unique solution x'. Of all vectors x such that $\|b - Ax\| \leq \|b - Ax'\|$, x' is the vector of minimum norm [53], so clearly it is necessary to select the value of γ which gives

$$\|b - Ax'\| \cong \epsilon. \tag{3.5}$$

Larger values of γ give x' vectors for which $\|b - Ax'\|$ is larger, so γ may be selected by trying various values and selecting post facto the value for which (3.5) is satisfied [53].

The perturbation of A^TA by γI gives (3.4) full rank (i.e., $r(A^TA + \gamma I) = n$), which is in contrast to previous discussions in Chapter 2 and Section 3.1 in which

perturbation of A by a matrix $\delta A'$ yielded a rank-deficient matrix A'. For all practical purposes, however, the perturbation of A^TA by γI has very much the same effect as perturbation of A by $\delta A'$, with larger perturbations giving more stability as responsibility for determining the solution is shifted from the equations to the minimum-norm constraint. An excessively large value of γ produces an overly smoothed solution, while setting γ too small may yield a solution with excessive noise error.

The Lagrangian multiplier method has proven to be very effective, particularly when noise levels can be adequately estimated. A disadvantage of using the normal equations is that they are more ill-conditioned than the original system [47] and therefore have an increased sensitivity to round-off errors made during their solution. It is of little concern in this application, however, because experience shows that (3.4) is well-conditioned if γ is sufficiently large. Furthermore, if there is an appreciable amount of error in A or b, round-off errors are likely to be negligible in comparison.

Like the singular value decomposition method, the Lagrangian multiplier method can become somewhat cumbersome for very large, sparse systems because it uses A^TA, which is not generally sparse. Another drawback for some applications (including ROG) is that inequality constraints are not

easily applied with this method. (It is possible to solve (3.4) by a method capable of applying inequality constraints, such as the Projection Method to be discussed in the next section. Such an approach has not been established in a rigorous manner, however, and would destroy the sparseness of large, sparse systems.) Treatments of the Lagrangian multiplier method may be found in [26, 52, 53].

3.3. The Projection Method

The Projection Method (PM) is an iterative method for solving systems of linear equations. Although it apparently evolved from a similar method reported by Morris [39,40] in 1935, it originated with Kaczmarz [30] in 1937. It is sometimes referred to as the Kaczmarz method. In the decades following its discovery, PM received little attention. In the early 1970's Tanabe [51] resurrected the then virtually unknown algorithm. Simultaneously, Gordon, Bender and Herman [22] rederived PM independently while working on x-ray tomography and electron microscopy. (In x-ray tomography a reconstruction of an object is created from a series of x-rays. This requires the solution of a large system of linear equations.) In this context, they named PM and some related techniques the Algebraic Reconstruction Techniques (ART). This terminology is widely used. However, because this treatment is not restricted to reconstruction problems, it is not employed here.

Following its use by Gordon, Bender and Herman, PM became an accepted method for solving tomography-related problems, particularly in the field of medical imaging [45]. Lager and Lytle [32] subsequently adopted it for use in geophysical exploration. As more is learned about PM and related algorithms, they promise to find wider application.

PM is demonstrated by an example in Figure 8. It starts with some initial point $x^1 \in E^n$ (usually $\underline{0}$) and projects that point onto the first hyperplane (i.e., it adds the smallest possible correction vector to make the point satisfy the first equation). This point is then projected in a similar manner onto the second hyperplane and onto each of the other hyperplanes in turn. After projection onto the last hyperplane, one iteration is said to have been completed. This process is then repeated.

The equation which defines PM is [51]

$$x^{q+1} = x^q + p_i(x^q), \tag{3.6}$$

where

$$p_i(x^q) = (b(i) - \langle a_i, x^q \rangle)\frac{a_i}{\|a_i\|^2}. \tag{3.7}$$

x^q is the previous estimate of the solution vector x, i

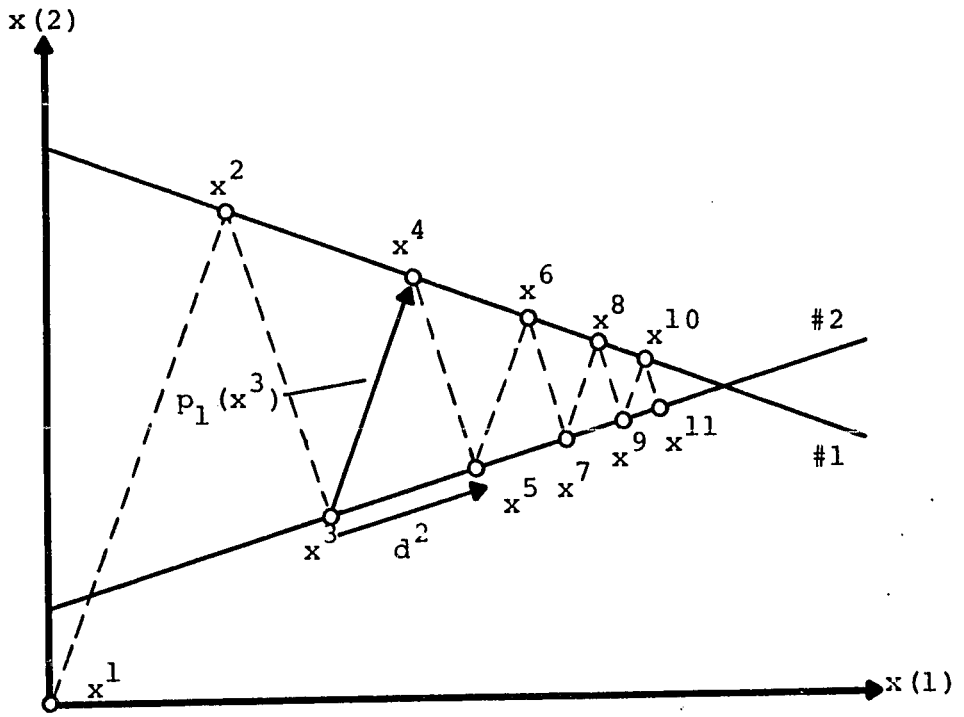


Figure 8. Application of PM to a consistent system.

is the row number of the current equation, and $p_i(x^q)$ is the current correction vector which projects x^q onto the i'th hyperplane. The total correction made during the k'th iteration may be denoted d^k, and

$$d^k = \sum_{i=1}^{m} p_i(x^{(k-1)m+i}). \tag{3.8}$$

(Examples of $p_i(x^q)$ and d^k are shown in Figure 8.)

Since PM uses A directly, without first forming A^TA, it can sometimes be made more efficient by taking advantage of the sparseness of large, sparse systems. Also, since only one row of A is needed at a time, computer memory requirements can be greatly reduced for some very large systems by recalculating each row of A as it is needed, thus eliminating the need to store A.

If (2.1) is a consistent system, real or complex, Tanabe [51] proved that

$$\lim_{q\to\infty} x^q = A^{\dagger}B + x^1_{N(A)}. \tag{3.9}$$

Thus, if $x^1 \in N^{\perp}(A)$, PM converges to the minimum-norm solution of (2.1).

For any inconsistent system, real or complex, Tanabe [51] proved that x^q converges in a cyclical manner as shown in Figure 9. Although x^q does not converge in the

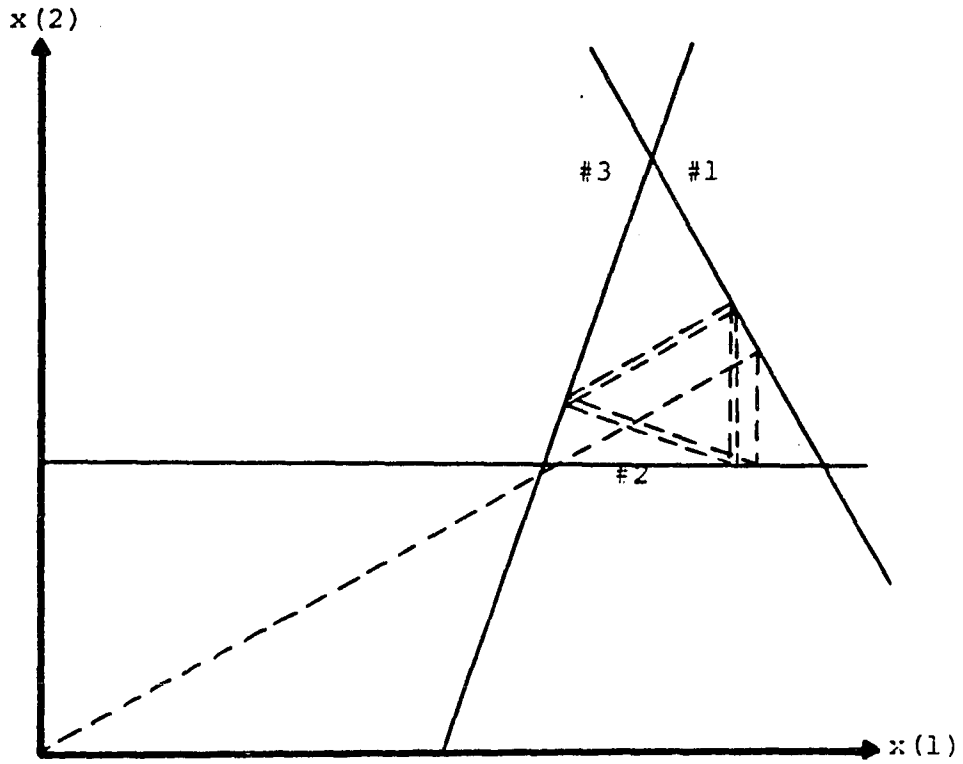


Figure 9. Application of PM to an inconsistent system.

usual sense of the word, the average of x^q over a cycle does. Experience has shown it to be a useful solution for tomographic problems. (This will be demonstrated by examples in Chapter 4.) This average may be denoted by

$$\bar{x}^q = \frac{1}{m} \sum_{p=q-m}^{q} x^p, \tag{3.10}$$

where $q > m$. $\lim_{q\to\infty} \bar{x}^q$ may be denoted by $\bar{x}^\infty$. In practice, $\bar{x}^q$ need only be calculated for the final iteration. In many applications the values of x^q are little different from $\bar{x}^q$ during the final iteration. Because of this, acceptable results have often been obtained by simply using the final value of x^q as a solution.

By Theorem 2.1,

$$\|b_o - A_o\|^2 = \sum_{i=1}^{m} \|p_i(x)\|^2, \tag{3.11}$$

where $\|p_i(x)\|$ is the distance from x to the i'th hyperplane of (2.1), and b_o and A_o are as defined in Definition 2.1. The projection of x^q onto the i'th hyperplane by PM reduces $\|p_i(x^q)\|$ to zero. It would appear, therefore, that PM attempts to make $\|b_o - A_o x^q\|$ small. Experience verifies that, in fact, $\bar{x}^\infty$ gives a near-minimal residual norm and that, if A_o is well conditioned, then [51]

$$\bar{x}^{\infty} \simeq A_o^{\dagger} b_o + x^1_{N(A)}. \tag{3.12}$$

Thus, if A_o is well-conditioned, PM produces an approximation of an equally weighted, least-squares solution. If $x^1 \in N^i(A)$, then PM produces an approximation of the equally weighted least-squares solution of minimum norm.

Since, in practice, ill-conditioned matrices are replaced by nearby well-conditioned matrices, as discussed in Chapter 2, (3.12) holds for most systems of practical interest. It should be noted, however, that for inconsistent systems the solution obtained by PM may be dependent on the order in which the equations are considered. This effect is usually slight in tomographic applications. It is reduced, but not eliminated, by use of the final value of $\bar{x}^q$ as a solution, as opposed to the final value of x^q.

3.3.1. The Weighted Projection Method

It is immediately apparent from (3.6) and (3.7) that the PM corrections are unaffected by multiplication of the equations by non-zero constants. As a result, weighting can not be achieved by scaling the equations as discussed in Section 2.3. (Recall that PM produces an approximation of an equally weighted least-squares solution, which is not necessarily a least-squares solution.) Weighting can be achieved [32], however, by replacing (3.6) with

$$x^{q+1} = x^q + w(i)p_i(x^q), \tag{3.13}$$

where $0 \leq w(i) \leq 1$ for $i = 1, \ldots, m$. Equations which are to be given more weight are assigned larger values of $w(i)$. This will be referred to here as the Weighted Projection Method (WPM). Although it has not been rigorously proven to converge to a weighted least-squares solution, intuition and extensive experience strongly suggest that it does, in an approximate sense [32].

WPM is illustrated by a simple example in Figure 10. Here, $w(1)$ and $w(2)$ are 1 and .5 respectively. Because the hyperplanes are parallel, cyclic convergence is achieved after one iteration. Clearly, the solution obtained, $\bar{x}^{\infty}$, is closer to the first hyperplane than the second.

3.3.2. Suppression of Noise Error by Application of Inequality Constraints

For some problems it is known in advance that the elements of x must lie within certain bounds. For example, they may be known to be non-negative, or for physical reasons they may be unlikely to exceed some upper limit. If at any point during the application of PM an element of x^q exceeds such a limit, this may be assumed to be the result of error in A or b, and the value of that element may justifiably be truncated to that limit. This may be referred to as the

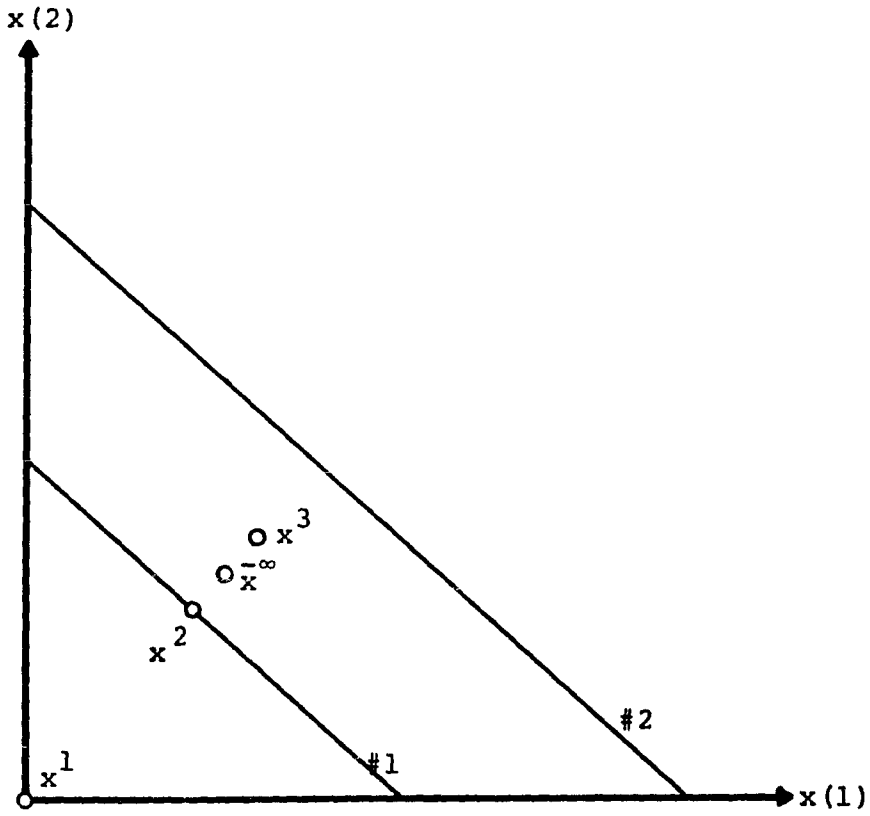


Figure 10. Application of WPM to an inconsistent system.

Constrained Projection Method (CPM). It has been shown to converge for consistent systems [25]. For inconsistent systems, CPM has not been proven to converge, although experience indicates that, like PM, it converges cyclically. In practice, CPM has proven very effective on tomographic problems, giving significant improvement over PM results when the noise level is high. (Examples will be presented in Chapter 4.)

3.3.3. Suppression of Noise Error by Algorithm Termination

When PM has been applied to tomography problems in the presence of noise, x^q has been observed to attain a good reconstruction of the object and then begin a long, slow process of deterioration. This deterioration is so slow and massive that for all practical purposes the algorithm diverges [25], in apparent violation of Twomey's proof that PM converges for any system. It has become common practice, therefore, to stop the PM algorithm prior to convergence, at a point in time when the reconstruction has formed but has not yet significantly deteriorated. A number of stopping criteria have been proposed, in a somewhat heuristic manner, to estimate the optimal stopping point [21]. An explanation of this phenomenon is offered in this section. Based on this new understanding of the problem, improved stopping criteria will be proposed in Section 3.3.4.

In most tomographic applications, $r(A) < \min(m,n)$. (This will be demonstrated for ROG in Chapter 4.) As a result, these problems have the extreme instability with respect to perturbations of A predicted by Theorem 2.6. Consequently, even if the unperturbed matrix A_u is well-conditioned, A is virtually always rendered highly ill-conditioned by round-off or other errors. Therefore, there are relatively well-conditioned A' matrices near A. As discussed earlier, if any one of these A' matrices is selected then E^n may be partitioned into S_e and S_c as in Equation (2.43).

Because of the instability associated with S_c and because of the inevitable presence of noise, $\bar{x}_c^{\infty}$ (the component of $\bar{x}^{\infty}$ in S_c) is virtually always very large. The rows of A, and hence the corrections $p_i(x^q)$, lie almost wholly within S_e (the row space of A'). Therefore, PM progresses very slowly in S_c. It will be shown that noise error can be dramatically reduced by allowing the PM algorithm to proceed long enough to extract most of the useful information from the equations and then stopping before convergence nears completion in S_c, where most of the noise error is concentrated.

Because $\bar{x}_c^{\infty}$ tends to be large and because PM progresses slowly in S_c, PM has a very strong tendency to

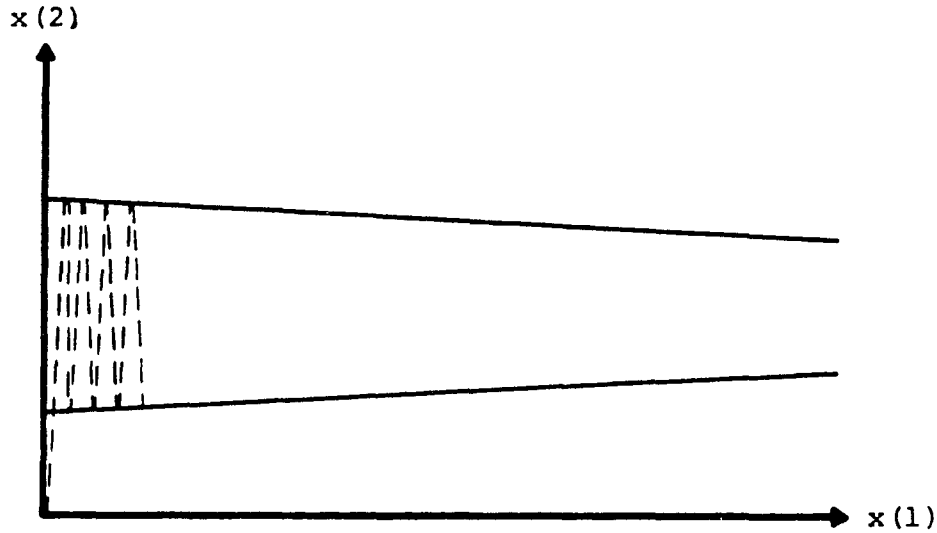
converge very slowly when A is highly ill-conditioned. This is illustrated by an example in Figure 11a. Conversely, intuition and extensive experience strongly suggest that PM converges rapidly when A is well-conditioned, as illustrated in Figure 11b. On occasion it will be fruitful to make use of this evidence and assume that PM does in fact converge quickly in some broad sense when A is well-conditioned. When the assumption is made, however, it will be identified as such.

Since the rows of A are approximated by the rows of A', it is obvious from (3.7) and (3.8) that when PM is applied to Ax = b,

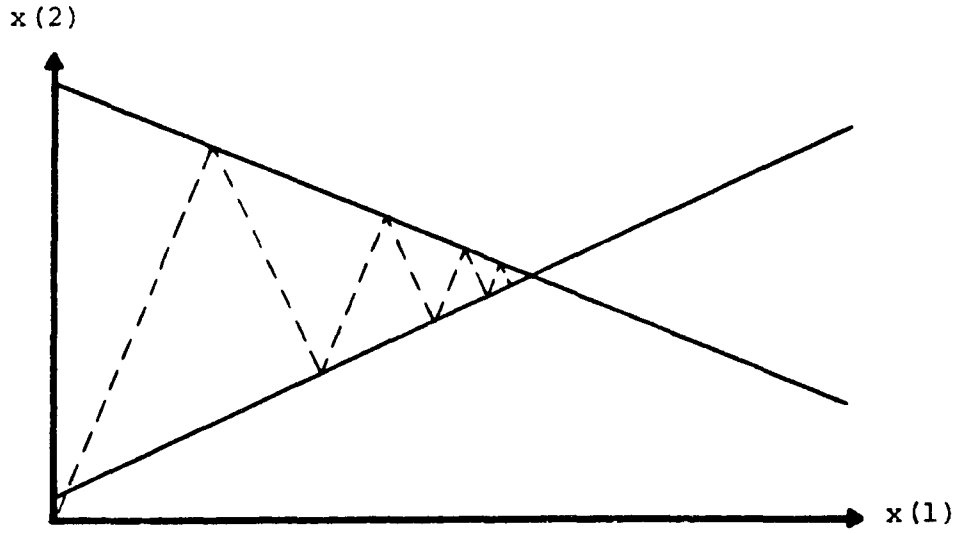
$$d^k = d'^k + e^k, \tag{3.14}$$

where d'^k denotes the k'th correction vector which would be obtained if PM were applied to A'x = b, and e^k is some error vector which is small for small k. The d'^k are in S_e, and the e^k contribute to noise error in S_c. Let x'_o denote the minimum-norm, equally weighted, least-squares solution of A'x = b; let $\bar{x}'^{\infty}$ denote the vector to which $\bar{x}^q$ converges when PM is applied to A'x = b; and recall from Section 3.3 that

$$x'^{\infty} \simeq x'_o. \tag{3.15}$$



(a)



(b)

Figure 11. Application of PM to systems with ill-conditioned and well-conditioned matrices.

(a) A system with an ill-conditioned matrix.
(b) A system with a well-conditioned matrix.

Since A' is well-conditioned, it is assumed that PM converges relatively quickly when applied to $A'x = b$. Therefore, the d'^k of Equation (3.14) decrease in size quickly and, after a relatively small number of iterations, become as small as the e^k. At this point, there have been enough iterations for PM to be near convergence if applied to $A'x = b$, but not enough for the e^k to have made a substantial contribution to $\bar{x}^q$. Thus, employing (3.15) yields

$$\bar{x}^q \cong \bar{x}'^{\infty} \cong x'_o. \tag{3.16}$$

If additional iterations are performed, the d'^k continue their rapid decrease and the error terms e^k start to dominate the corrections. As a result, $\bar{x}^q$ stops getting closer to x'_o and begins the long, slow movement away from x'_o observed by tomographers as a deterioration of their reconstruction. If the minimum-norm, equally weighted, least-squares solution is desired, then the point in time when $\bar{x}^q$ attains its closest proximity to x'_o is the optimal stopping point. (Convergence is largely completed in S_e but has barely begun in S_c.) By stopping near this point, PM approximates the minimum-norm, equally weighted least-squares solution to $A'x = b$, and noise error is thus suppressed, as discussed in Chapter 2.

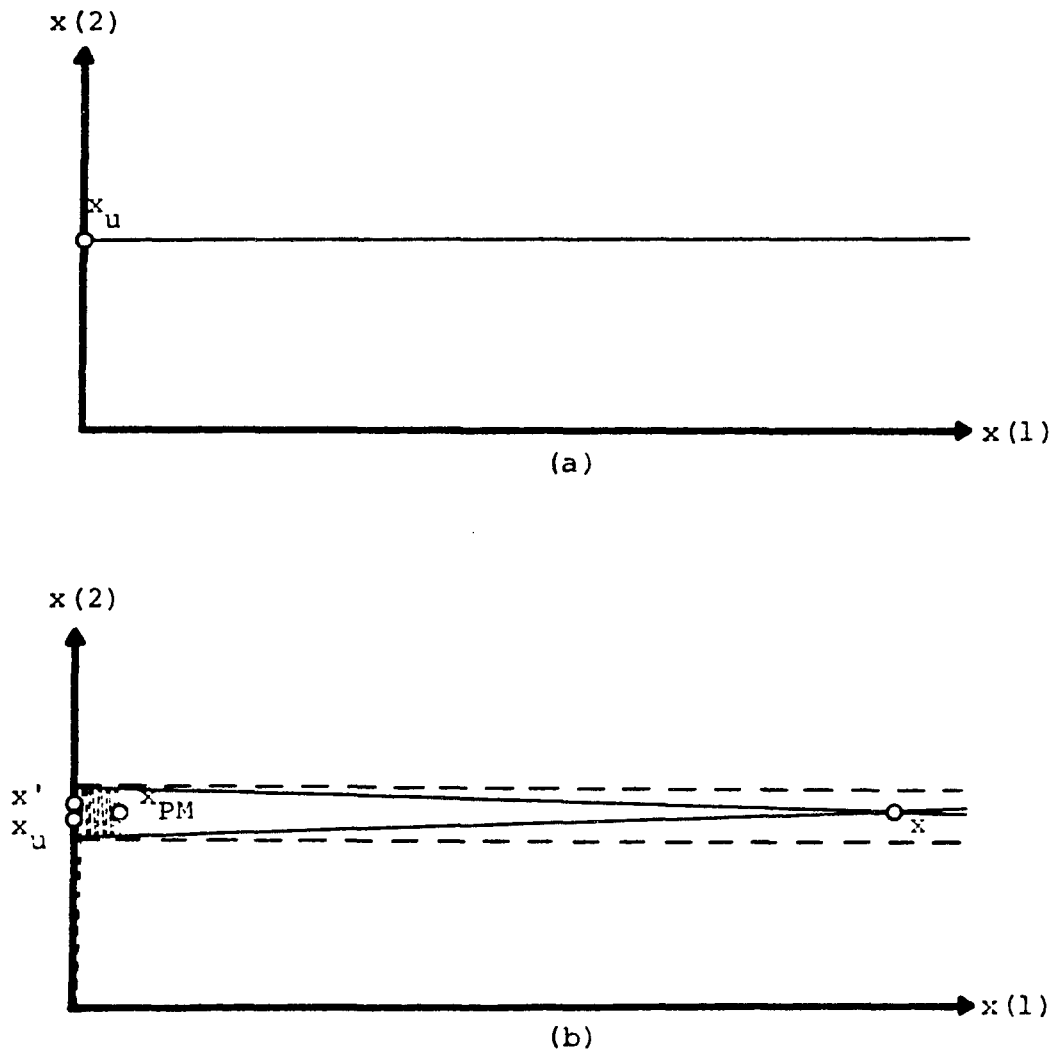
The size of $\|\delta A'\|$ (recall that $\delta A' = A' - A$) may be selected by varying the number of iterations performed. Because PM converges more slowly for systems with more ill-conditioned matrices, performing more iterations allows the algorithm to near convergence for less well-conditioned A' matrices nearer to A. As a result, $\|\delta A'\|$ is reduced by stopping later. From discussion in Chapter 2 it follows that later stopping allows extraction of more information from the equations while less effectively suppressing noise error. It is interesting to note that at no point in this discussion has it been necessary to identify a specific A' matrix. The conclusions drawn here hold for any A' and the corresponding subspaces S_e and S_c, as long as A' is well-conditioned, near A, and satisfies (2.28).

An example of noise suppression by algorithm termination is shown in Figure 12. Figure 12a shows the system

$$A_u x_u = b_u \tag{3.17}$$

with two coincident hyperplanes. Here, $m = n = 2$ and $r(A_u) = 1$. Thus, (3.17) has the extreme instability predicted by Theorem 2.6. Figure 12b shows the system

$$Ax = b \tag{3.18}$$



Frigure 12. Noise error suppression by termination of the PM algorithm.

(a) The unperturbed system $A_u x_u = b_u$.

(b) PM applied to the perturbed system $Ax = b$.

obtained from (3.17) by slightly perturbing A_u and b_u. The resultant solution, x, is far from the unperturbed solution x_u. When PM is applied to (3.18), it progresses very slowly after the first iteration. If it is stopped after some moderate number of iterations, then the final value of $\bar{x}^q$, denoted here by x_{PM}, is much nearer to x_u than x is. It is easily seen that for a moderate number of iterations PM corrections are approximately the same as they would have been if A had been replaced by any A' matrix. (Recall from earlier discussion that for an example of this type an A' matrix would give parallel hyperplanes at some angle near the angles of the hyperplanes of (3.18).) The hyperplanes of the system

$$A'x = b \tag{3.19}$$

for some A' appear in Figure 12b as dashed lines, and the corresponding minimum-norm, least-squares solution, x', is shown. Clearly, for this example,

$$x_{PM} \cong x' \cong x_u \tag{3.20}$$

as predicted.

3.3.4. Stopping Criteria

When PM (or any other iterative method) is used to solve the system $Ax = b$, it must be decided when to cease iterations. This decision can be based on the known quantities A and b and the present and past values of x^k. Sometimes additional information such as the degree of precision desired of the solution or the estimated noise level may be employed. If A is well-conditioned, PM converges quickly to an approximation of x. The algorithm may be stopped when the corrections become small; for instance, when

$$SC1^k < SC1_o, \tag{3.21}$$

where

$$SC1^k = \frac{\|d^k\|}{\|x^{km+1}\|} \tag{3.22}$$

and $SC1_o$ is some small constant selected to reflect the degree of precision desired. The only cost associated with performing additional iterations is the computing time required. If A is ill-conditioned, however, stopping must be used to suppress noise error as discussed previously. Deciding when to stop is therefore more important and more

difficult. The remainder of this section is concerned with means of making this decision.

It must be emphasized that, as discussed in Chapter 2, the error measure which should be minimized is $\|x^k - x_u\|$. (Recall that x_u is the solution of the unperturbed system (2.26).) Because x_u is unknown, this must be done in an indirect manner. The residual norm $\|b - Ax^k\|$ is also an error measure in the sense that it measures how well x^k satisfies the equations. It is very convenient to use because it can be calculated from known quantities. However, it is important to remember from Chapter 2 that if A is ill-conditioned, x^k can be very far from the true solution x_u and still produce a small residual norm. Thus, for x^k to be near the true solution x_u, it is necessary but not sufficient that x^k has a small residual norm.

The optimal stopping point varies with A, b, and the noise levels (i.e., $\|\delta b\|$ and $\|\delta A\|$). It is also affected by the application of inequality constraints. Larger $\|\delta b\|$ requires more noise suppression and hence earlier stopping. Smaller $\|\delta A\|$ tends to increase noise error but slows convergence in S_c, with the net effect of requiring later stopping. Inequality constraints reduce noise error and slow convergence, thus delaying the optimal stopping point.

At low noise levels, the solution deteriorates slowly after reducing its optimal value. Therefore, stopping

a moderate number of iterations past the optimal point causes little harm. At higher noise levels, however, the solution deteriorates more quickly. This makes it more important to accurately estimate the optimal stopping point. Some stopping criteria currently used for this purpose are discussed here, and some new stopping criteria are proposed.

Of the various stopping criteria in current use, the simplest is to stop the algorithm when

$$SC2^k = SC2_o, \tag{3.23}$$

where

$$SC2^k = k \tag{3.24}$$

and $SC2_o$ is an integer selected in advance. The optimal value of $SC2_o$ may be selected by running a test case.

To run a test case for a given type of problem, a typical matrix A_u and a typical solution x_u are selected and $b_u = A_u x_u$ is calculated. Then, b_u is perturbed by a typical type and amount of noise. (A_u may also be perturbed if desired.) If PM is then applied to the resultant perturbed system, the optimal value of $SC2_o$ is the value of $SC2^k$ when $\|x_u - x^k\|$ reaches its minimum. For best results,

$SC2_o$ should be reselected for significantly different noise levels or different types of problems.

$SC2^k$ is a rather crude stopping criterion. It is unable to adjust to any of the varying conditions mentioned earlier unless $SC2_o$ is reselected for each different problem. Although $SC2^k$ can provide adequate stopping for some purposes, its use is not recommended when accuracy and flexibility are required.

Two stopping criteria which have been applied to tomography problems in the past are based on the variance and entropy of x^q [21]. (The variance and entropy of a vector are defined by equations (2.24) and (2.23).) If V^k and S^k denote the variance and entropy of x^q after the k'th iteration, then the inequalities which indicate the stopping point when satisfied are

$$SC3^k \leq SC3_o \tag{3.25}$$

and

$$SC4^k \leq SC4_o, \tag{3.26}$$

where

$$SC3^k = \frac{|V^{k+1}-V^k|}{V^k} \tag{3.27}$$

and

$$SC4^k = \frac{|s^{k+1} - s^k|}{s^k}. \tag{3.28}$$

$SC3_o$ and $SC4_o$ are small constants selected by running test cases.

Although these stopping criteria apparently evolved out of an interest in selecting a unique solution by the application of minimum variance or maximum entropy constraints [21, 22], use of $SC3^k$ or $SC4^k$ does not constitute an application of these constraints. Earlier discussion showed that a unique solution is determined by the selection of x^1, and that the stopping criterion only controls the degree of noise suppression. As PM proceeds toward convergence, each successive iteration tends to cause less change in x^q. As a result, functions of x^q also tend to change more slowly. Therefore, because $SC3^k$ and $SC4^k$ tend to decrease with each iteration, they can be used to estimate the optimal stopping point.

Values of $SC3_o = .01$ [25] and $SC4_o = .005$ [43] have been used successfully for specific tomography problems. Extensive experience with ROG problems, however, shows that the optimal values of $SC3_o$ and $SC4_o$ vary widely as the optimal stopping point varies with A, b, $\|\delta b\|$, $\|\delta A\|$ and the effects of inequality constraints.

Although experience verifies that $SC3^k$ and $SC4^k$ tend to decrease as the PM algorithm progresses for many

ROG problems, it also shows that they sometimes behave erratically. The application of inequality constraints is particularly disruptive to $SC3^k$ and $SC4^k$. When these constraints are applied to systems with appreciable noise levels, $SC3^k$ and $SC4^k$ usually fail to decrease beyond a certain level, and thus have virtually no value as stopping criteria. It should also be noted that $SC4^k$ can not be applied unless the elements of x^q are constrained non-negative.

Another stopping criterion [21] which has been applied in tomography problems is based on the residual norm of x^q, $\|b - Ax^q\|$. (This is sometimes referred to as the discrepancy.) If $SC5^k$ denotes the residual norm of x^q after completion of the k'th interation, then the stopping point is indicated when

$$SC5^k \leq SC5_o, \tag{3.29}$$

where $SC5_o$ is a constant whose selection will be discussed shortly. ($\bar{x}^q$ could be used instead of x^q to calculate the residual norm, but this is not necessary.) Because later stopping less effectively truncates the massive noise error in S_c, $\|x^q\|$ has a strong tendency to increase with q, particularly if $x^1 = \underline{0}$. This tendency becomes stronger as the algorithm progresses. Each successive iteration forces

x^q to satisfy the equations more exactly, so $SC5^k$ tends to decrease toward a minimum. By stopping when $SC5^k$ has decreased to the level of uncertainty in the equations, one selects (approximately) the x^q of minimum norm which satisfies the equations within the level of uncertainty in the data. This level of uncertainty ($SC5_o$) is usually simply set equal to the estimated noise level in b, $\|\delta b\|$. However, noise in A may also be taken into account by increasing $SC5_o$. This approach is similar in principle to the Lagrangian multiplier method of Section 3.2.

$SC5^k$ is an effective stopping criterion and adjusts well to different problems and different noise levels as long as a rough estimate of the noise level can be made. Care must be taken, however, not to underestimate the noise level, because $SC5^k$ may never decrease far enough to stop the algorithm. On the other hand, grossly overestimating $\|\delta b\|$ is also obviously an undesirable source of error.

For problems with unpredictable and potentially high noise levels and for which it may be important to obtain optimal solutions (such as ROG), selecting $SC5_o$ can pose a difficult problem. In such cases, there are ways to estimate the noise level [6], such as computing $\|b_{R^\perp(A)}\|$ from

$$b_{R^\perp(A)} = (I - AA^\dagger)b \tag{3.30}$$

or by applying PM without stopping and observing the level toward which $\|Ax^q - b\|$ descends. The latter type of noise-level estimation has been done in the past for ROG problems [43], although with somewhat different motivation. This potentially difficult and imprecise process may be avoided by the use of two new stopping criteria which will be introduced here shortly. These new stopping criteria will be able to compensate, to a large extent, for unpredictable variations in the noise level.

The optimal stopping point occurs approximately when d^k shifts from being primarily in S_e to being primarily in S_c. If a stopping criterion could detect this transition, it could provide near-optimal stopping and would automatically compensate for any variations in A, b, $\|\delta A\|$, or $\|\delta b\|$.

Speaking in broad terms, the fundamental difference between S_e and S_c is that for any unit vector $\delta x_c \in S_c \cap N_A^{\perp}$, $\|A\delta x_c\|$ is small, while for any unit vector $\delta x_e \in S_e \cap N_A^{\perp}$, $\|A\delta x_e\|$ is not small. (For δx_c this follows immediately from Theorem 2.4, Part 2. For δx_e, consider Theorem 2.4, Part 1, and let $\delta b = A^{\dagger}x_e$. Then

$$\delta x = A^{\dagger}A\delta x_e = \delta x_e$$

by Equation (A1.12) in Appendix 1. It follows that

$$\|\delta x_e\| \leq \frac{\kappa(A')}{\|A'\|}\|A\delta x_e\|. \tag{3.31}$$

Recall that $\|\delta A'\|$ is small and $\kappa(A')$ is not large.) Because the d^k are linear combinations of the row vectors of A and $N^{\perp}(A)$ is the row space of A,

$$d^k \in N^{\perp}(A). \tag{3.32}$$

Therefore, if d^k is primarily in S_c, Ax^q, and hence the residual norm $\|b - Ax^q\|$, change slowly from iteration to iteration. If d^k is primarily in S_e, then Ax^q changes more rapidly and it is assumed that $\|b - Ax^q\|$ tends to change more rapidly. (Experience supports this assumption.) Thus, it is possible to estimate which of the subspaces S_e and S_c contains more of d^k by observing the rate of change of the residual norm. The optimal stopping point is indicated when

$$SC6^k \leq SC6_o, \tag{3.33}$$

where

$$SC6^k = \frac{\|b-Ax^{(k-1)m}\| - \|b-Ax^{km}\|}{\|b-Ax^{(k-1)m}\|} \tag{3.34}$$

and $SC6_o$ is a small positive constant whose selection will be discussed shortly.

For very similar reasons, d^k changes more slowly when it is primarily in S_c than when it is primarily in S_e, suggesting the stopping criterion

$$SC7^k \leq SC7_o, \tag{3.35}$$

where

$$SC7^k = \frac{\|d^k - d^{k-1}\|}{\|d^{k-1}\|}$$

and $SC7_o$ is a small positive constant to be determined.

Application of PM to ROG problems to be discussed in Chapter 4 has shown that $SC6^k$ and $SC7^k$ both adjust well, but not perfectly, to variations in A, b, $\|\delta A\|$, and $\|\delta b\|$. Widely differing problems do require different values of $SC6_o$ and $SC7_o$ for optimal stopping, with higher noise levels requiring higher values. These variations are much less pronounced, however, than they are for the other stopping criteria, making it much less crucial to estimate noise levels accurately. $SC6_o = .1$ has proven effective for a wide variety of ROG problems. Application of an inequality constraint can seriously disrupt the fall of $SC7^k$, but $SC6^k$ seems considerably less sensitive to this problem. Since $SC6^k$ also appears to adjust more effectively to varying conditions, it appears to be preferable to $SC7^k$.

None of the stopping criteria discussed here are able to compensate fully for the later stopping required for CPM. Therefore, when applying inequality constraints the practice is introduced here of stopping when

$$SC6^k \leq SC6_o \cdot \max\left(\left(\frac{n-n_t^k}{n}\right)^4, .02\right), \tag{3.37}$$

where n_t^k denotes the number of elements of x which had to be truncated after the k'th iteration. As before, n denotes the dimension of x. This rule of thumb gives significantly improved results for ROG problems.

To summarize, the preferred stopping criteria are $SC1^k$, $SC5^k$, and $SC6^k$. $SC1^k$ is recommended only for well-conditioned systems. $SC5^k$ is very effective if an estimate of noise level is available. Use of $SC5^k$ is also convenient, because it does not require running a test case. Because underestimating the noise level can cause a failure to stop, however, approximate knowledge of the noise level is necessary for optimal results with $SC5^k$. Use of $SC6^k$ may require running an occasional test case for optimal results on widely differing problems, and in rare cases $SC6^k$ can be severely disrupted by the application of inequality constraints. However, $SC6^q$ is able to adapt to unpredictable variations in noise level. (This is important in ROG.)

The proper stopping criterion should be selected for each application, and sometimes, by using more than one simultaneously, one criterion can compensate for the weaknesses of another. $SC6^k$ is a good general purpose stopping criterion which works well for well-posed as well as ill-posed problems. Its performance will be demonstrated in Chapter 4.

3.3.5. The Relaxed Projection Method

The Relaxed Projection Method (RPM) is defined by

$$x^{q+1} = x^q + \Delta p_i(x^q), \tag{3.38}$$

where $p_i(x^q)$ is as defined previously (Equation 3.7) and Δ is a relaxation parameter, usually between 0 and 1. (Δ is sometimes referred to as a damping factor.) Δ is assumed here to remain fixed throughout any given application of PM. If $0 < \Delta < 2$, it has been proven [13] that RPM converges for any system, with inconsistent systems exhibiting the cyclic convergence described earlier.

A relaxation parameter was first applied to PM by Sweeney and West while solving optical interferometry problems [50]. They employed it during the final iterations to reduce the excessive influence of the last few equations on the solution. As mentioned earlier, this could also be

achieved by using $\bar{x}^q$ instead of x^q as a solution. However, there are other benefits of relaxation beyond those initially foreseen by Sweeny and West. Even if the effect of the ordering of the equations is all but eliminated by using $\bar{x}^q$ as the solution, use of a small relaxation parameter allows noise error to more effectively be suppressed by algorithm termination, at a cost of slow convergence. This is demonstrated by Figure 13.*

Figure 13 displays the optimal error, ϵ_o, as a function of Δ for RPM applied to an ROG problem in the presence of noise. For a fixed value of Δ, ϵ_o is defined by

$$\epsilon_o = \frac{\|\bar{x}_o^q - x_u\|}{\|x_u\|}, \tag{3.39}$$

where, as before, x_u is the solution of the unperturbed problem and $\bar{x}_o^q$ is the value of $\bar{x}^q$ for which $\|\bar{x}^q - x_u\|$ is minimal. (Note that x_u is not known in most applications.

---

*Radcliff and Balanis [43] reported that for each problem there was an optimal value of Δ which was dependent on the noise level. They were among the first to recognize the importance of the noise level. However, they failed to realize that it was $SC3_o$, and not Δ, which depended on the noise level. By adjusting Δ they apparently varied the speed of the algorithm until their stopping criterion aligned with a good stopping point. This made it appear that there was an optimal value of Δ. In fact, smaller Δ is always better, but slower, as demonstrated by Figure 13.

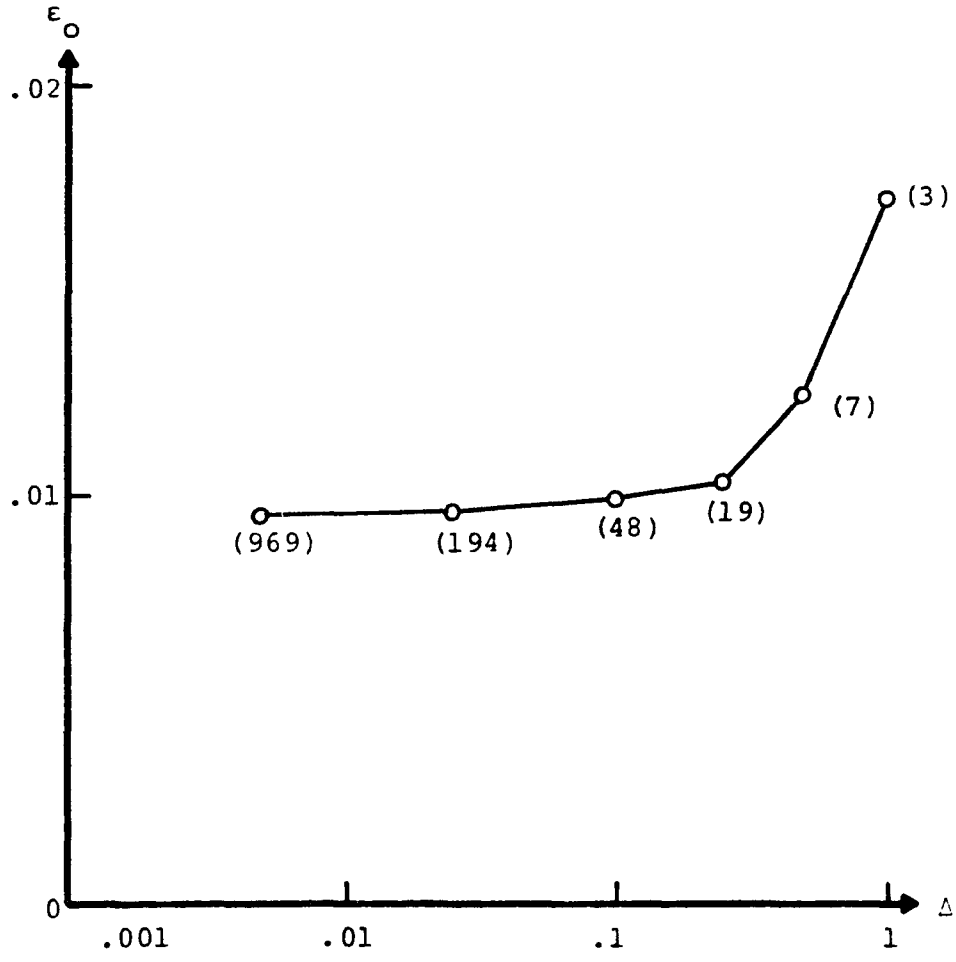


Figure 13. Minimum error vs. relaxation parameter for RPM.

Advance knowledge of the solution is employed for this example to identify the optimal stopping point and thus determine the upper limit of RPM performance as a function of Δ. In practice, errors would be greater than those shown here due to the use of an imperfect stopping criterion.) Each circled point in Figure 13 corresponds to an application of RPM with the indicated value of Δ. Near each of these points, the optimal number of iterations appears in parentheses. It is clear that, with diminishing returns after a point, decreasing Δ improves the solution at a cost of slower convergence.

A likely explanation for RPM's superior noise suppression is that relaxation slows the progress of the PM algorithm in S_c (where most of the noise error is) more than it slows the progress in S_e. (This is admittedly speculation.) An examination of the effects of relaxation will lead to the development of a new algorithm in the next section.

A final important point is that when the stopping criterion is based on a rate of change, it is necessary to compensate for the altered rate of convergence. Thus, for example, $SC6^k$ should be applied by stopping when

$$SC6^k \leq SC6_o \cdot \Delta. \tag{3.40}$$

(This appears not to have been recognized in the past [43].)

3.4. The Simultaneous Projection Method

RPM with a small relaxation parameter has been observed to suppress noise effects more effectively than PM. It will be fruitful to examine exactly how an iteration of PM with a small relaxation parameter changes x^q.

The first iteration starts with vector X^1 and calculates

$$x^2 = x^1 + \Delta P_1(x^1), \tag{3.41}$$

and then

$$x^3 = x^2 + \Delta P_2(x^2) \tag{3.42}$$

$$= x^1 + \Delta P_1(x^1) + \Delta(\langle a_2, x^1 + \Delta P_1(x^1)\rangle - b(2))\frac{a_2}{\|a_2\|^2}.$$

(Recall that $P_i(x^q)$ is defined by (3.7).) Since Δ is small, second order terms may be neglected, yielding

$$x^3 \cong x^1 + \Delta P_1(x^1) + \Delta(\langle a_2, x^1\rangle - b(2))\frac{a_2}{\|a_2\|^2} \tag{3.43}$$

$$= x^1 + \Delta(P_1(x^1) + P_2(x^2)).$$

If this process is continued, neglecting higher order terms of Δ, the change made by the first iteration is found to be

$$x^{l+m} \cong x^{l} + \Delta \sum_{i=1}^{m} p_i(x^{l}). \tag{3.44}$$

Apparently, (3.44) represents a very desirable correction, since the neglected higher order Δ terms appear to contribute more heavily to noise error than the first order terms do.

At this point, an obvious question is whether the higher order terms could be eliminated by simply calculating $\sum_{i=1}^{m} p_i(x^{l})$ directly, making the time-consuming use of a small relaxation parameter unnecessary. Although experience indicates that the algorithm

$$x^{k+1} = x^{k} + \sum_{i=1}^{m} p_i(x^{k}) \tag{3.45}$$

does not always converge, the algorithm

$$x^{k+1} = x^{k} + d^{k} \tag{3.46}$$

with

$$d^{k} = \frac{1}{m} \sum_{i=1}^{m} p_i(x^{k}) \tag{3.47}$$

does converge and gives the same superior noise suppression as RPM with an infinitesimally small relaxation parameter. This algorithm considers all the equations simultaneously by averaging the corrections necessary to satisfy the individual equations. It is essentially the same as the Simultaneous Iterative Reconstruction Technique (SIRT) proposed by Gilbert [14], and in the context of this treatment it will be termed the Simultaneous Projection Method (SPM).

SPM has other advantages over PM in addition to its superior noise suppression. It is demonstrated by Theorem 3.1 that SPM converges to an equally weighted least-squares solution, and that if $x^1 \in N^{\perp}(A)$, then SPM converges to the minimum-norm equally weighted least-squares solution.

THEOREM 3.1. When SPM (as defined by equations (3.46), (3.47), and (3.7)) is applied to any real system of linear equations $Ax = b$ for which the rows of A are non-zero,

$$\lim_{k\to\infty} x^k = x^1_{N(A)} + A_o^{\dagger} b_o,$$

where A_o and b_o are as defined in Definition 2.1. (A proof is included in Appendix 2.)

Subsequent to proving Theorem 3.1, it came to the attention of the author that the corrections d^k for the

SPM algorithm are identical to the corrections produced by the application of a special case of Richardson's method to the normal equations of the scaled system

$$A_o x = b_o. \tag{3.48}$$

Here, A_o and b_o are as defined in Definition 2.1. Richardson's method [57] is an iterative method defined by

$$x^{k+1} = x^k + \mu(b - Ax^k), \tag{3.49}$$

where μ is a scalar. If (3.49) is applied to the normal equations

$$A^T Ax = A^T b \tag{3.50}$$

with

$$0 < \mu < \frac{2}{\|A\|^2}, \tag{3.51}$$

then

$$\lim_{k\to\infty} x^k = x^1_{N(A)} + A^\dagger b. \tag{3.52}$$

(A proof is given in [55].) Because the SPM correction d^k can be expressed in the alternate form

$$d^k = \frac{1}{m} A_o^T (b_o - A_o x^k) \tag{3.53}$$

(see Equation (A1.26) in Appendix 1), it follows that application of Richardson's method to

$$A_o^T A_o x = A_o^T b_o \tag{3.54}$$

with $\mu = \frac{1}{m}$ is equivalent to SPM. The similarity of Richardson's method to SIRT, which in turn is similar to SPM, was reported in [34].

Noise error is suppressed by stopping SPM much the same way it was for PM. (The same stopping criteria are used. Optimal stopping levels, such as $SC6_o$, may be different, however.) In fact, some of the discussions in the PM section which had to be treated approximately or supported by empirical evidence become more precise when applied to SPM, with its well-defined convergence properties. When SPM is applied to an ill-posed system and stopped by a stopping criterion, the resultant solution approximates an equally weighted least-squares solution of $A'x = b$, where A' is a well-conditioned matrix near A, as before. The norm of d^k serves as a measure of how well x^k

satisfies the equations, because it is proportional to the residual of the normal equations of the scaled problem $A_o x = b_o$ (see Equation (3.53)). Geometrically, d^k is the average of the shortest vectors from x^k to each of the hyperplanes (Equation (3.47)). Lemma A2 in Appendix 2 establishes that the sequence $\{\|d^k\|\}_{k=1}^{\infty}$ is monotonically decreasing.

Inequality Constraints can be applied with SPM, just as they were for PM, by simply truncating the elements of x^k as necessary after each iteration. This constrained algorithm may be termed the Constrained Simultaneous Projection Method (CSPM). It is possible to assign weights to the equations by replacing (3.47) with

$$d^k = \frac{1}{m}\sum_{i=1}^{m} w(i)p_i(x^k), \tag{3.55}$$

where, as before, the elements of $w \in E^m$ are weighting coefficients between 0 and 1. This will be referred to here as the Weighted Simultaneous Projection Method (WSPM). If diagonal matrix $Q \in E^{m\times m}$ is defined by

$$Q = \mathrm{diag}(w(1)^{1/2}, w(2)^{1/2}, \ldots, w(m)^{1/2}), \tag{3.56}$$

then, by (3.55), (3.7) and (A1.26),

$$d^k = \frac{1}{m}Q^2A_o^T(b_o - A_ox^k) \tag{3.57}$$

$$= \frac{1}{m}(QA_o)^T(Qb_o - QA_ox^k).$$

But this is equivalent to Richardson's Method (3.49) applied to the normal equations of

$$QA_ox = Qb_o; \tag{3.58}$$

so WSPM converges to a weighted least squares solution of (2.1) (i.e., the least-squares solution of the scaled system (3.58)). If $x^1 \in N^\perp(A)$, then WSPM converges to the minimum-norm, weighted least-squares solution.

3.5. The Fast Simultaneous Projection Method (FSPM)

Although SPM converges faster than RPM with a very small damping factor, it converges much slower than PM. For ROG problems SPM is typically slower than PM by a factor of from 5 to 50 or more. This motivated the development of the Fast Simultaneous Projection Method (FSPM), which is introduced here. FSPM speeds convergence by replacing (3.46) with

$$x^{k+1} = x^k + \Delta_k d^k, \tag{3.59}$$

where d^k is defined by (3.45), as before, and

$$\Delta_k \geq 1. \tag{3.60}$$

In this context, Δ_k may be referred to as a pushing factor.

Δ_k should be chosen carefully to speed the algorithm as much as possible without disrupting its convergence or reducing the effectiveness of the stopping criterion. Most of the stopping criteria discussed earlier are severely disturbed by excessive pushing. $SC6^k$, however, is relatively insensitive to this, and $SC5^k$ is even more tolerant of pushing. The required value of Δ_k often varies widely as the algorithm progresses. Therefore, it is best to adjust it after each iteration.

In order to adjust Δ_k, it is helpful to calculate

$$\theta_k = \frac{\langle d^k, d^{k-1} \rangle}{\|d^k\| \|d^{k-1}\|}. \tag{3.61}$$

When θ_k is near 1, the corrections are changing direction slowly. Thus Δ_k can be increased. If θ_k falls significantly below 1, Δ_k should be decreased (but never below 1). It is also important to monitor the progress of the stopping criterion. A decrease in Δ_k is indicated if the stopping criterion rises or if its rate of decrease

suddenly drops. Particular care should be taken when the algorithm is near termination. It is often advisable to keep θ_k closer to 1 as the algorithm nears its stopping point. If

$$\Delta^k < \frac{2}{\|A\|^2} \tag{3.62}$$

for $k > 0$, then FSPM converges to the same solution as SPM. (This follows from its similarity to Richardson's algorithm with a variable value of μ [57].)

When a stopping criterion is used which is based on a rate of change, then it is necessary to compensate for variations in Δ_k as in Equation (3.40). Inequality constraints may be applied with FSPM as they were with SPM. The resulting algorithm may be referred to as Constrained FSPM (CFSPM). For ROG problems, FSPM and CFSPM have been found to be as effective as SPM and CSPM, respectively, while converging typically 10 to 20 times as fast.

CHAPTER 4

GEOTOMOGRAPHY

Tomography is an area of medical science concerned with producing images of internal areas of the human body. Geotomography is a related field in which tomographic techniques are applied to geophysical exploration. After a brief discussion of tomography and the various methods of geotomography, this chapter focuses on Ray-Optic Geotomography (ROG). ROG is a method of geotomography in which the region between two boreholes is probed by the transmission of ray-optic electromagnetic radiation. The performance limitations of ROG are determined by applying it to computer-simulated data corrupted by noise. It is concluded that ROG is an effective means of short-range, high-resolution probing for low-loss media such as salt or dry rock. The approximate nature of the straight-line model of propagation employed in this implementation of ROG is found to be balanced by the insensitivity of ROG to errors in the equations. It is shown that the linear equations which arise in ROG must be solved by an algorithm capable of suppressing noise error. The CPM and CFSPM algorithms discussed in Chapter 3 are found to be very effective for this purpose. VECTOR is discussed briefly as a potential alternative to ROG.

4.1. Tomography

In medical x-ray tomography a display of the x-ray attenuation (i.e., the degree of opaqueness to x-rays) of a selected plane of the body is derived from several x-rays taken at different angles. Since the individual x-rays view not only the area of interest, but also obstacles in front of or behind that area, the tomograph can reveal details not visible in conventional x-rays. The intensity of a given point of an x-ray is determined by the amount of radiation which strikes that point. This, in turn, is determined by the line integral of the attenuation along the path taken by that radiation. Thus, in tomography a function is reconstructed from a number of line integrals of that function. These line integrals are often referred to as projections.

Problems of this nature occur in a number of different fields. Two medical imaging techniques which are closely related to x-ray tomography are ultrasound scanning, in which a velocity profile is determined from time of flight measurements; and nuclear emission scanning, in which a patient is scanned after ingesting radioactive material [23, 24]. Other areas in which unknown functions are reconstructed from their line integrals include electron-microscopy, radio astronomy, optical inteferometry and geophysical exploration [25, 5, 50, 10].

Lorentz is credited with being the first to solve a problem of this type [46]. The mathematical foundations of this process were subsequently established by Radon in 1917 [44]. (More recent treatments taking a mathematical point of view include [8, 9, 23, 31, 46].) The first reconstruction of an image from its line integrals was reported by Bracewell [4] in 1956. He independently developed a technique of reconstruction employing Fourier transforms for application in radio astronomy. Bracewell and Riddle [5] subsequently modified this technique, eliminating the need to actually calculate the Fourier transforms. In the Fourier approach, noise error is reduced by filtering the high frequency content of the Fourier transforms.

A different approach was taken by Gordon, Bender and Herman [22], who approximated the integral equations governing tomography and electron-microscopy with large systems of linear equations and solved them directly with their Algebraic Reconstruction Technique (ART) (i.e., PM), without any use of Fourier techniques. As observed earlier, they suppressed noise error by stopping their algorithm.

When computer-aided tomography (CAT) machines for medical imaging were introduced in the early 1970's, they employed ART. Most of the more recent machines employ Fourier methods similar to Bracewell's. Two likely reasons for this trend are (1) the Fourier approach is faster (which can be

important for the extremely large systems which arise in medical imaging) and (2) the process of stopping ART prior to convergence is not well understood [31, 45]. One of the primary contributions reported here is the analysis of the stopping process in Chapters 2 and 3, along with the improved stopping criteria proposed as a result of that analysis. Although Fourier methods work well for many applications, it will be shown shortly that, for several reasons, a direct approach similar to ART is preferable for the geophysical applications to be considered here.

4.2. Ray-Optic Geotomography

Lager and Lytle [32] adapted tomographic techniques to geophysical exploration. The resulting branch of geophysical exploration has been termed "geophysical tomography" or "geotomography". It has been the subject of recent research [10, 33, 37, 43]. In its usual form, the goal of geotomography is to characterize the plane segment between two boreholes by dividing it into n cells and assigning some characteristic value $x(j)$ to each cell.

This can be done by transmitting radiation along m different paths between the holes. Seismic radiation can be employed [3, 17]. This treatment, however, is restricted to ray-optic geotomography (ROG), in which electromagnetic radiation is assumed to propagate in a ray-optic manner.

For this implementation of ROG, a straight-line, ray-optic model of propagation is assumed. The approximate nature of this model will be discussed shortly. The configuration of paths and cells for ROG is shown in Figure 14 for an example with 30 cells and 36 paths.

For electromagnetic radiation propagating in a ray-optic manner along the i'th path, it is well-known that the power at the receiver and the power at the transmitter are related by

$$P_{rec_i} = P_{tran_i} \exp(-2\int_i \alpha ds), \tag{4.1}$$

where $\int_i ds$ denotes line integration along the i'th path and α is the spatially varying attenuation of the medium in nepers/m. It is also well-known [49] that

$$\alpha = \omega(\frac{\mu\varepsilon}{2})^{1/2}[(1 + [\frac{\sigma}{\omega\varepsilon}]^2)^{1/2} - 1]^{1/2}, \tag{4.2}$$

$$\alpha \cong \frac{\sigma}{2}(\frac{\mu}{\varepsilon})^{1/2} \quad (\text{for} \quad \omega >> \frac{\sigma}{\varepsilon}), \tag{4.3}$$

and

$$\alpha \cong (\frac{1}{2}\omega\mu\sigma)^{1/2} \quad (\text{for} \quad \omega << \frac{\sigma}{\varepsilon}). \tag{4.4}$$

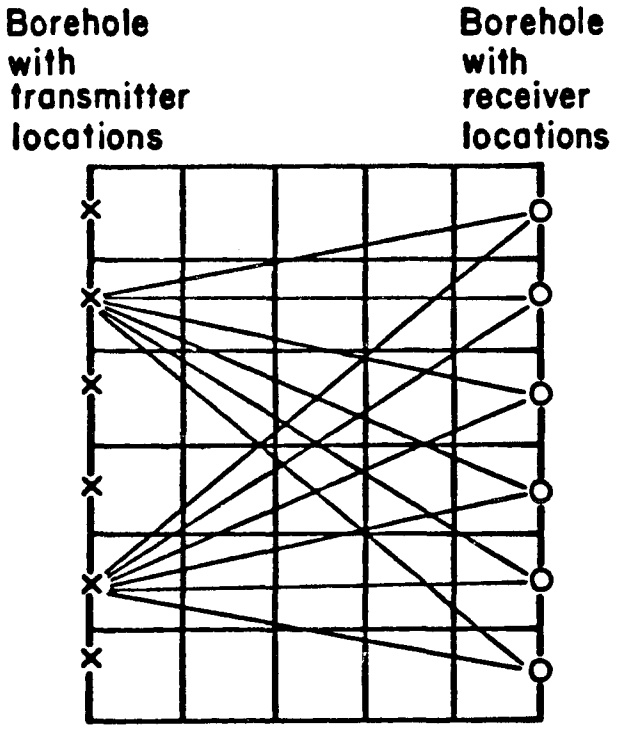


Figure 14. The geometry for ROG.
(Not all paths are shown.)

Here, ω denotes the angular frequency (radians/second) and σ, ε, and μ denote the conductivity (mhos/meter), permittivity (farads/meter) and permeability (henries/meter) of the medium. Values of σ, ε, and μ for various geologic materials may be found in [17]. The frequency dependence of α is shown in Figure 15.

The object of ROG is to reconstruct α as a function of position given P_{tran_i} and P_{rec_i} for the various paths. If $y(i)$ denotes the net attenuation in the nepers of the radiation transmitted along the i'th path, then

$$y(i) = \int_i \alpha ds \tag{4.5}$$

for $i = 1, 2, \ldots, m$. It follows from (4.1) that

$$y(i) = \frac{1}{2}\ln\left(\frac{kP_{tran_i}}{P_{rec_i}}\right), \tag{4.6}$$

where k is a constant determined by such practical considerations as antenna gains and cable losses [32]. If α is modeled as being uniform within each cell and $x(j)$ denotes the value of α in the j'th cell, then (4.5) may be approximated by the system of linear equations

$$Dx = y, \tag{4.7}$$

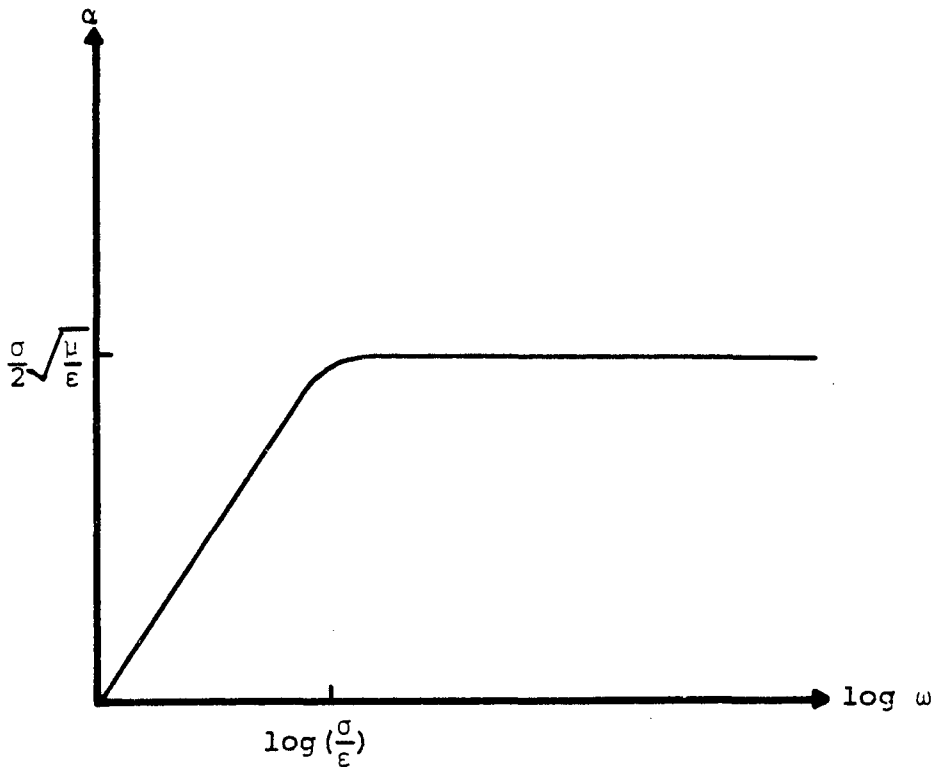


Figure 15. Attenuation as a function of frequency.

where D is the $m \times n$ matrix such that $D(i,j)$ is the length in meters of the segment of the i'th path which is inside the j'th cell. The desired vector of unknown attenuations is x, and y is determined from measured data by (4.6). The approximate nature of (4.7) may be viewed as a source of error in D and y.

4.2.1. Limitations of ROG

The straight-line model of propagation upon which (4.7) is based is quite simple. It requires the following assumptions [32]:

1) displacement currents dominate conduction currents (i.e., $\omega >> \frac{\sigma}{\varepsilon}$);
2) borehole separation is greater than $\lambda/2\pi$, where λ is the wave length in the medium;
3) any spatial variation of the permittivity of the medium is smooth.

The first two conditions are easily met by using a sufficiently high frequency. It is clear from Figure 15 that this can result in a large attenuation of the signal, which limits the effective probing range. Because of this, ROG is best suited for short-range probing (perhaps a few hundred meters) in low-loss media such as salt or dry rock. (Low-loss media are those for which (4.2) gives a small value of α.) In media with appreciable moisture content,

the range of ROG is severely limited. (For such media, seismic radiation might be appropriate.)

It is probable that the assumption of smoothly varying permittivity will be violated to some extent in practice. When this is the case, the straight-line model of propagation fails to account for partial reflection and ray bending which occur at interfaces between dissimilar media [49]. Although partial reflection may be viewed as a source of error because it causes an interface to appear as a region of high attenuation, this actually proves useful in the detection of anomalies such as fractures or air-filled voids. Ray bending, on the other hand, is a potentially significant source of error in D.

Error due to ray bending can be greatly reduced if the time of flight is measured for each path [38]. From these measurements, it is possible to estimate the velocity of the radiation in each cell. Because the velocity of electromagnetic radiation does not vary widely for different geologic materials, it is not as good a diagnostic as attenuation. However, velocity calculations can be used to estimate the true ray paths. Thus, prior to solving (4.7), D can be adjusted to compensate for ray bending. Although this scheme for correcting D is desirable, its implementation is beyond the scope of this work. An in depth treatment appears in [38].

There are numerous other potential sources of error in ROG. Radiation can propagate by extraneous modes, such as up one hole, along the surface, and down the other hole. Interfaces between different layers of rock can allow unwanted modes of propagation due to reflection, and a layered earth can act as a waveguide [54]. Instrument errors and the assumption of uniform attenuation within each cell also contribute to the overall error. It is clear that, in practice, (4.7) may be a highly perturbed system. Because many of the errors are determined by unknown geologic features, it is also clearly very difficult to estimate the noise level in advance.

Yet another potential source of error is associated not with inaccuracy in (4.7), but with the null space of D. Since the equations of (4.7) have no information about the component of x in $N(D)$, an anomaly which has a significant component in $N(D)$ can not be reconstructed by ROG. An example of such an anomaly is the slab shown in Figure 16. In the forward problem of calculating y from x, the slab increases the net attenuation along each path by the same ratio, so it generates the same data as would a uniform region with an attenuation of 2 nepers/m. Such difficulties are easily avoided by probing a region of sufficient vertical extent so that no anomaly is likely to

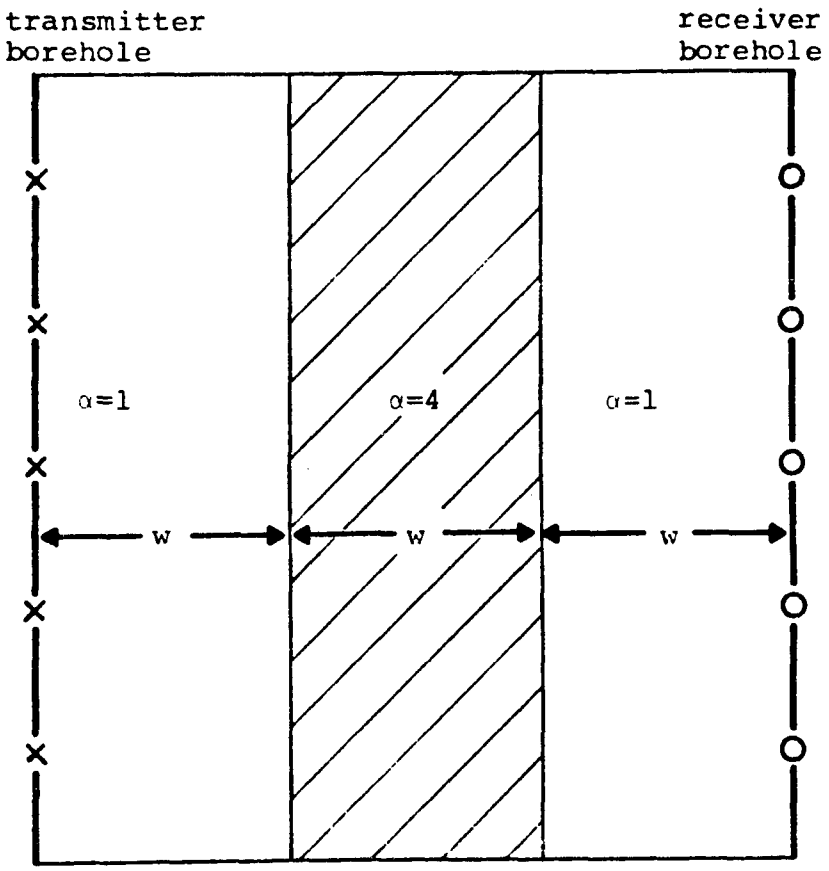


Figure 16. An anomaly which can not be detected by ROG.

intersect all the paths, as illustrated in Figure 17. If this is done, then experience indicates that ROG can reconstruct anomalies of interest in geophysical exploration. Excessively long paths should be omitted, to keep the received signal well above the noise level. Null-space error for VECTOR is discussed in [11], and a discussion of null-space error in medical imaging appears in [31].

4.2.2. Algorithm Selection

Equation (4.7) is not simple to solve for several reasons, one of which is the potentially large size of D. If, for example, a square region 200 meters wide is partitioned into square cells 1 meter wide, then there are 40,000 cells and D has 40,000 columns. D is sparse, however. (For convenience, the systems solved here are considerably smaller than this, but such a problem could easily arise in practice.) The problem is further complicated by the fact that D is rank-deficient. (This is demonstrated in Appendix 3). Thus $r(D)$ is virtually always less than $\min(m,n)$, and (4.7) therefore has the extreme instability discussed in Chapter 2. This, coupled with the high and unpredictable noise level discussed in the previous section, makes efficient suppression of noise error very important. All these factors must be weighed when selecting the method by which to solve (4.7).

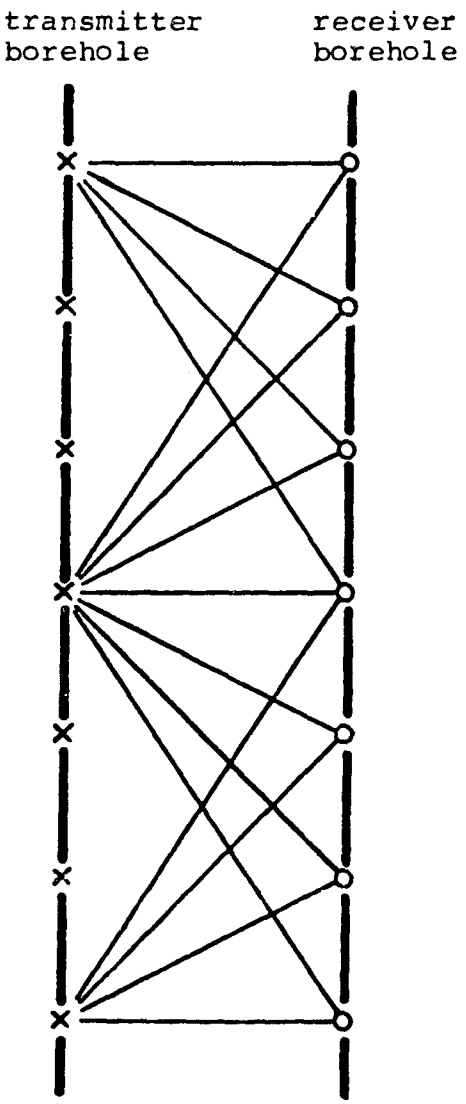


Figure 17. A configuration to reduce null-space error.
(Paths are shown only for selected transmitters.)

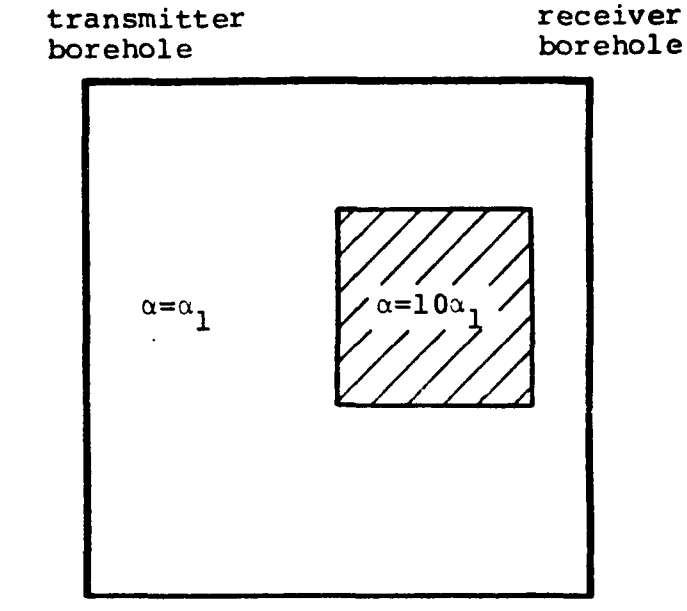
The Fourier methods mentioned earlier, which work well for some applications, are not well-suited for ROG because they do not readily adapt to the varying geometric configurations which arise in geotomography. They are also believed to be less readily adaptable to correction for ray-bending than direct methods such as the projection algorithms [10]. (Here, "projection algorithms" refers collectively to PM and SPM and all their modified forms.) Because high noise levels often result in geotomography solutions with physically impossible negative values of attenuation, the ability of the projection algorithms to impose inequality constraints makes them well-suited for solving geotomography problems. Inequality constraints are not readily applied with the singular value decomposition or Lagrangian multiplier methods. An additional advantage of using projection algorithms for this application is their ability to suppress noise error efficiently when the noise level is unknown by using the stopping criterion $SC6^k$.

For the reasons given here and in previous sections, CPM and CFSPM have been selected here as the best means of data inversion for ROG. The best stopping criterion was found to be $SC6^k$, employed as in Equation (3.37), with $SC6_o = .1$ for CPM and $SC6_o = .005$ for CFSPM. CPM is considerably easier to implement and is faster than CFSPM, but CFSPM usually gives noticeably superior results. (This

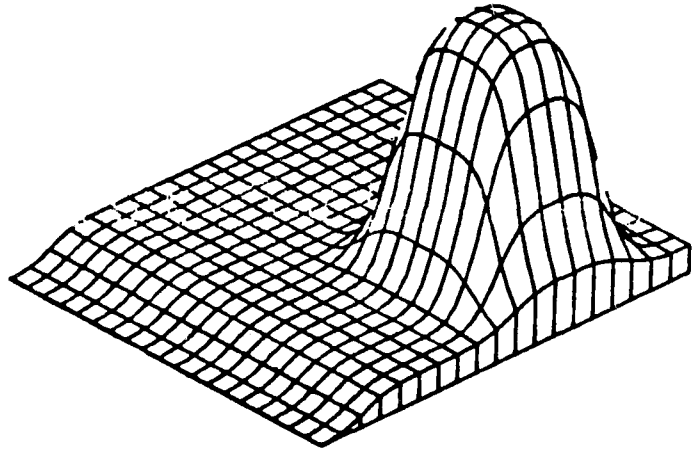
will be illustrated by examples in Figures 19 and 26.) For systems of moderate size with low noise levels, efficiency and the ability to impose inequality constraints are not essential. Under these conditions the Lagrangian multiplier method can be used effectively, particularly if the noise level can be estimated.

In order to compare the solutions obtained by different algorithms visually, it is helpful to plot these solutions in three dimensions. Figure 18a is a two-dimensional representation of a void, and Figure 18b shows the same void represented by a three-dimensional perspective block plot. Some smoothing was done prior to plotting, resulting in rounded edges. To provide a reference level, the front left edge of the perspective block plot has been set to zero.

Figure 19 shows perspective block plots of the solutions obtained by applying various algorithms to a simulated ROG problem. This problem was simulated by calculating the y vector for the void of Figure 18, assuming straight-line propagation. This vector was then perturbed by the addition of uniformly distributed noise. The noise level for this case was 10% (i.e., the norm of the added noise vector was 10% of $\|y\|$). Eighty-one paths and 81 cells were used , and the attenuation contrast was 10:1 (i.e., the

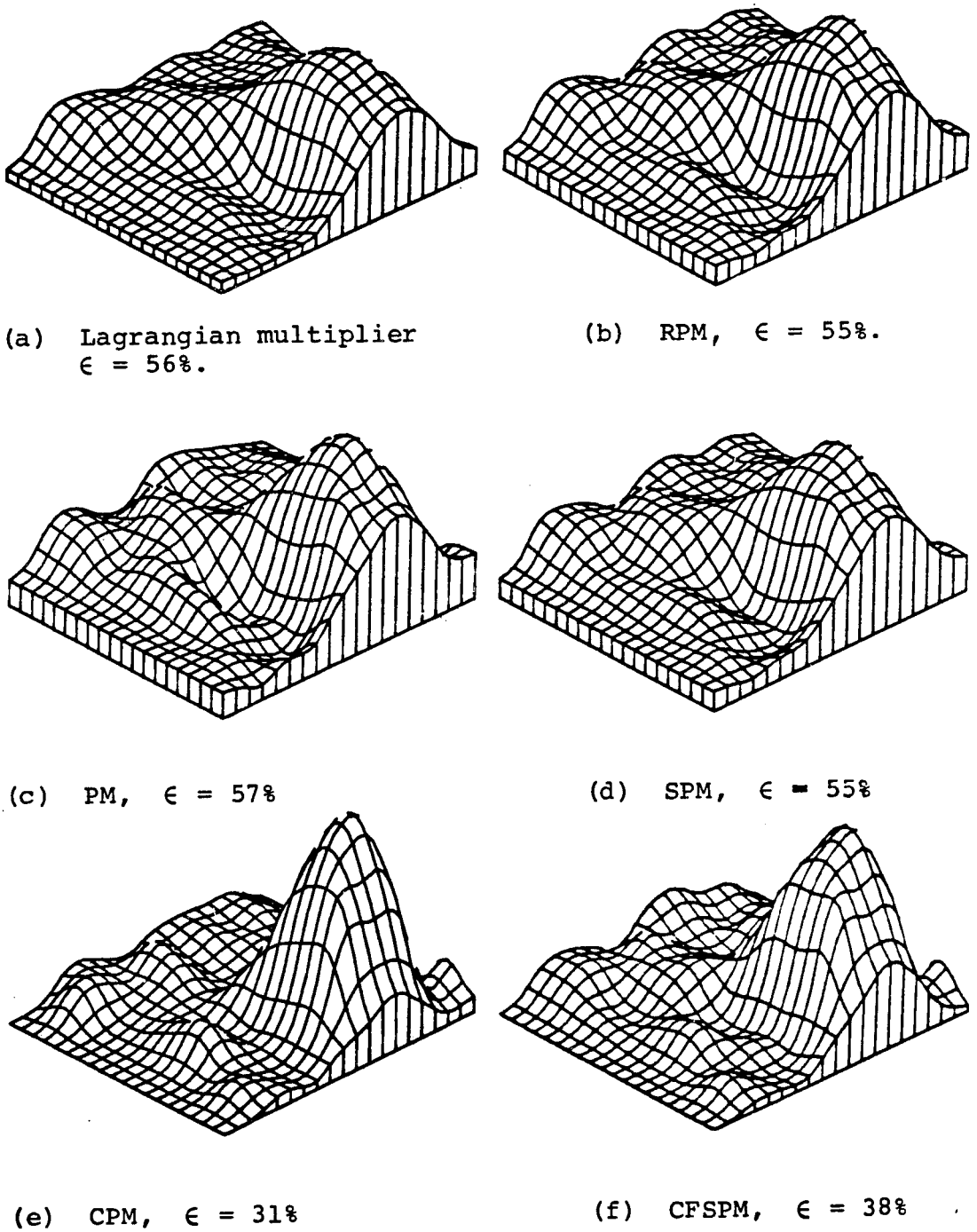


(a) Two-dimensional diagram



(b) Three-dimensional perspective-block plot

Figure 18. Two and three dimensional representations of a void.



(a) Lagrangian multiplier ϵ = 56%.

(b) RPM, ϵ = 55%.

(c) PM, ϵ = 57%

(d) SPM, ϵ = 55%

(e) CPM, ϵ = 31%

(f) CFSPM, ϵ = 38%

Figure 19. Comparison of various algorithms.

attenuation of the void was 10 times that of the background medium).

In order to compare the limitations of the various methods equitably, advance knowledge of the solution was used to select optimal stopping points and the optimal Lagrangian multiplier for this example. In practice, slightly degraded results would be obtained due to the use of imperfect stopping criteria. (Test cases utilizing stopping criteria will be presented shortly.)

The per cent error

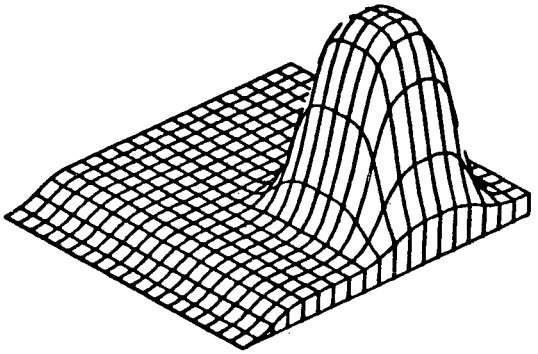
$$\epsilon = \frac{\|x - x_{tm}\|}{\|x_{tm}\|} \cdot 100 \tag{4.8}$$

is shown for each algorithm, where x is the solution obtained with the algorithm and x_{tm} is the test model solution, corresponding to the original void. The solutions of Figures 19e and f were constrained non-negative. Comparison of these solutions with the others in Figure 19 clearly demonstrates the value of applying inequality constraints and the superiority of the CPM and CFSPM algorithms for this application. (The PM, RPM, and CPM solutions were averaged over the final iteration.) The similarity of Figures 19b and d reflects the similarity of SPM to RPM with a small relaxation parameter. (Here, the relaxation parameter was 0.1.) For this example, PM and FSPM produced solutions of

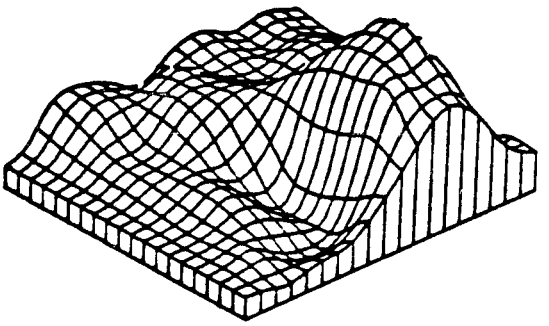
similar quality, both with and without inequality constraints. In most cases, however, FSPM (CFSPM) solutions are noticably superior to PM (CPM) solutions. This will be illustrated later by another example, in Figure 26.

4.2.3. Inequality Constraints for ROG

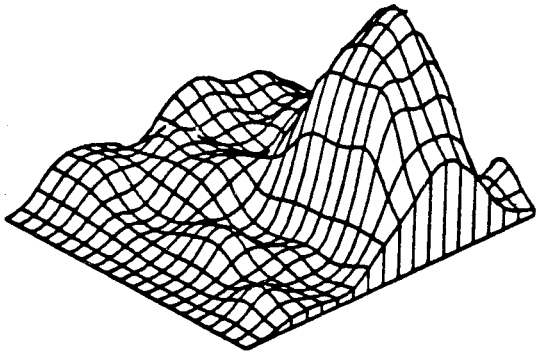
Because the elements of x are known to be non-negative, noise error can be reduced by constraining them to be non-negative. This has been done in previous implementations of ROG [10, 43]. However, when probing in a low-loss medium such as salt or dry rock, any geologic anomalies are highly likely to have attenuations greater than that of the background medium. Even an air-filled void, which has very low attenuation, would appear as an area of high attenuation due to partial reflection, as discussed previously. Therefore, it is proposed here that for low-loss media noise error may be further suppressed by constraining the $x(j)$ above the attenuation level of the background medium. For similar reasons it may also be beneficial to impose an appropriate upper bound on the $x(j)$, although no upper bound is used here. Use of a background level constraint can significantly improve the solution, as demonstrated by the example shown in Figure 20. Here, the CFSPM algorithm has been used to impose various inequality constraints on the



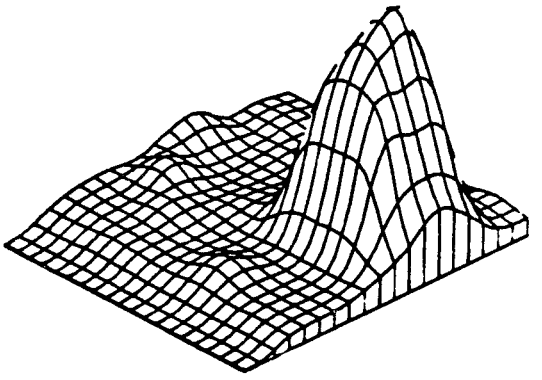
(a) test model



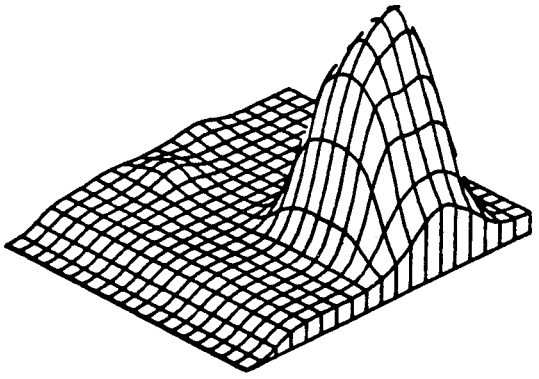
(b) unconstrained, ϵ = 55%



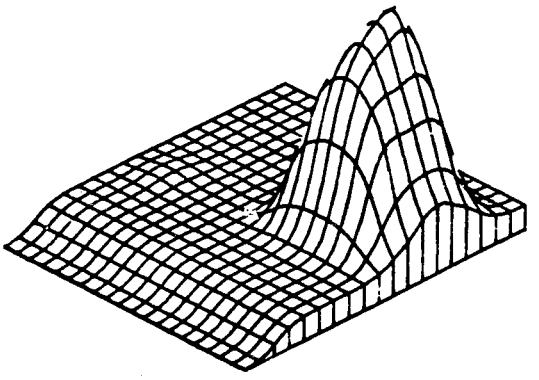
(c) CL = 0, ϵ = 38%



(d) CL = 80%, ϵ = 21%



(e) CL = 100%, ϵ = 19%



(f) CL = 120%, ϵ = 20%

Figure 20. CFSPM solutions for various constraint levels.

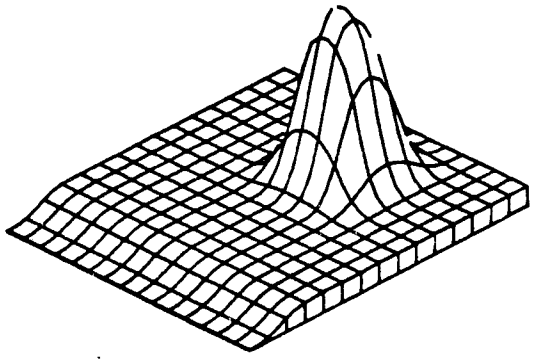
solution of the ROG problem from Figures 18 and 19. The constraint level (CL) as percentage of background attenuation and the percent error ϵ are given for each case.

When the solution was constrained above the background level (Figure 20e), it more closely resembled the correct solution in Figure 20a than it did when it was just constrained non-negative (Figure 20c). Both the constrained solutions in Figures 20c and e approximate the correct solution much more accurately than does the unconstrained solution in Figure 20b. Figures 20d and f show that good results are still obtained if the inequality constraint is applied 20% above or below the background attenuation level. Thus, the background attenuation level need not be precisely known.

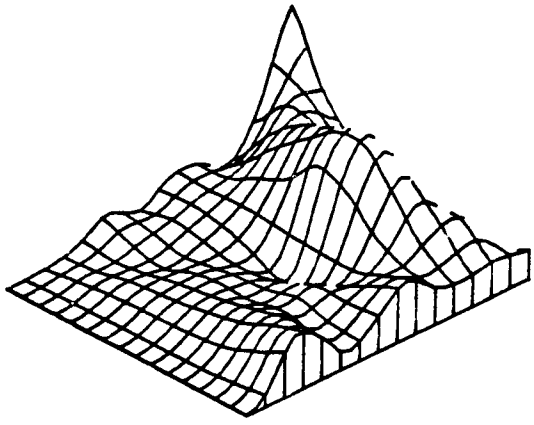
4.2.4. Additional Test Cases

Figures 21 and 22 show the solutions obtained by the PM algorithm after various numbers of iterations. For this example, a model of a void with a 10:1 attenuation contrast was used to generate the data (Figure 21a). Eighty-one cells and 81 paths were used, and 15% uniform noise was added to the data. The number of iterations, k; the percent error, ϵ; and the normalized residual,

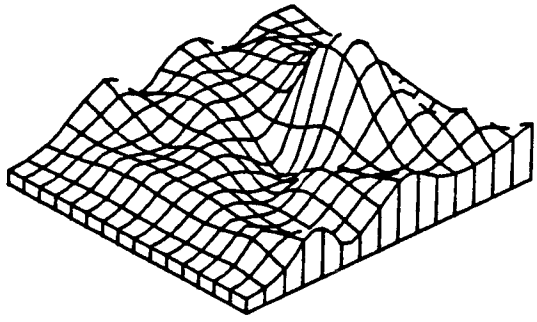
$$r = \frac{\|Dx-y\|}{\|y\|}, \tag{4.9}$$



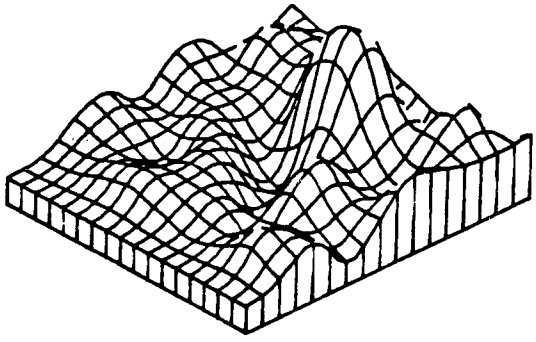
(a) test model



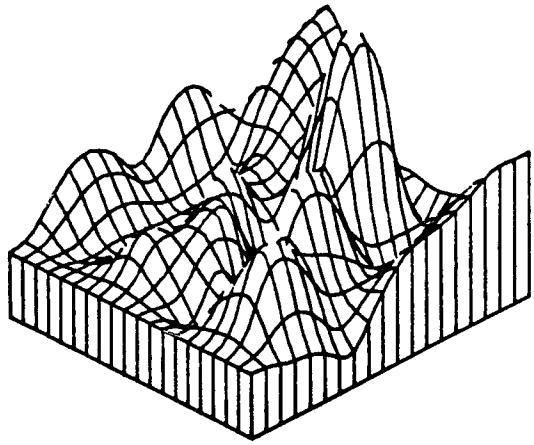
(b) k = 1, ϵ = 77%, r = 45%



(c) k = 3, ϵ = 62%, r = 14%

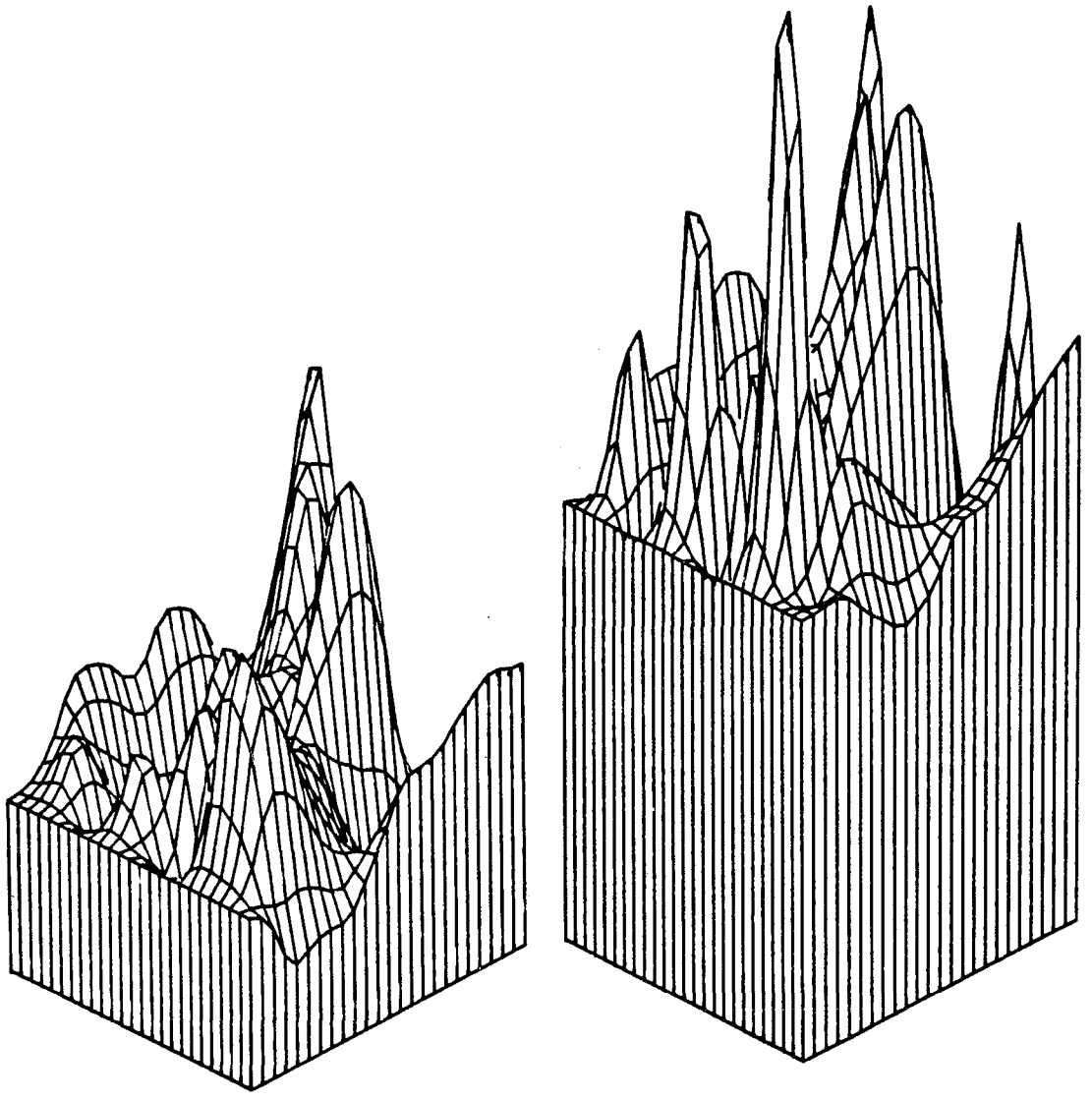


(d) k = 10, ϵ = 70%, r = 11%



(e) k = 100, ϵ = 135%, r = 9.8%

Figure 21. PM solutions for various numbers of iterations.



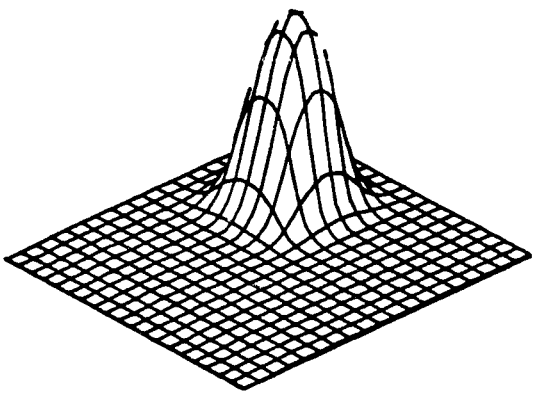
(a) k = 1,000, ϵ = 308%, r = 8.5%

(b) k = 10,000, ϵ = 656%, r = 8.7%

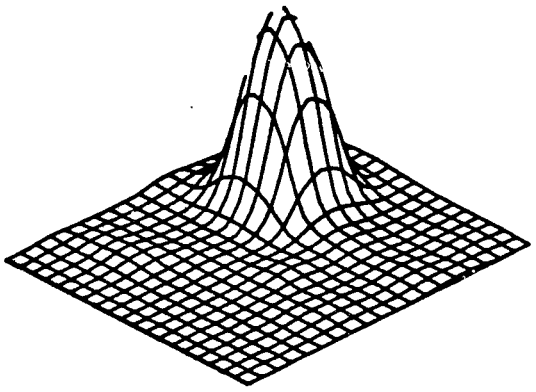
Figure 22. PM solutions for various numbers of iterations-continued.

are given for each case. Here, D, y and ϵ are as previously defined and x is the PM solution. It is easily seen that PM reaches its best reconstruction quickly (Figure 21c) and then begins the slow, massive deterioration discussed in Chapter 3. The best reconstruction is not very good because no inequality constraint was imposed for this example. The importance of stopping the algorithm may be appreciated by observing the massive deterioration of the solution in Figures 22a and b. While PM progresses slowly after 10,000 iterations, it has by no means fully converged. Note that r continues to decrease even while the solution is deteriorating. This supports the assertion, made in Chapter 2, that a small residual norm is no guarantee of an acceptable solution. (After about 3,000 iterations, r leveled off and drifted back up slightly.)

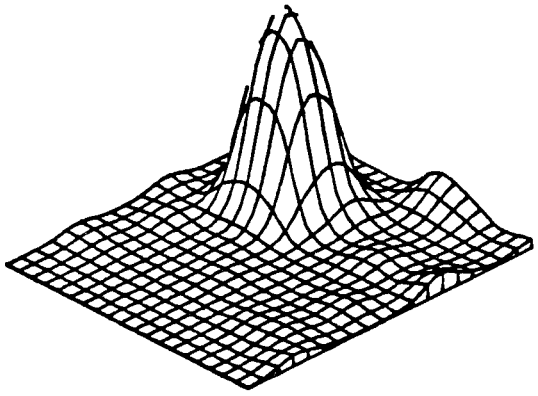
Figure 23 shows the performance of CFSPM for different levels of uniform noise. A void with a 100:1 attenuation contrast (Figure 23a) was used to generate data. (This model resembles a fluid-filled cavity in dry rock, such as a brine pocket in salt.) Eighty-one paths and 81 cells were used. The solution was constrained non-negative, and the algorithm was stopped by $SC6^k$. For Figures 23b-f, the per cent noise level, N; the per cent error, ϵ; and the required number of iterations, k, are given. (Here, k is



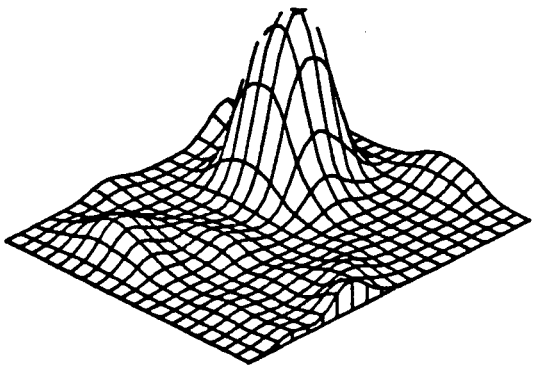
(a) test model



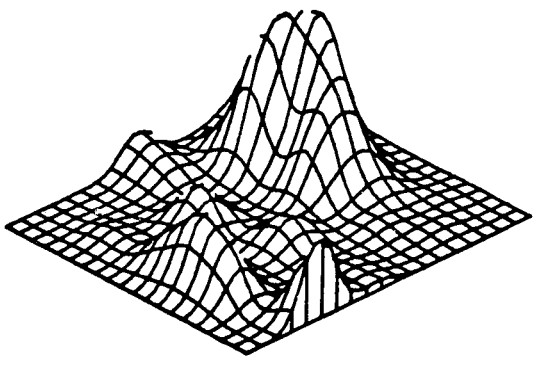
(b) N = 10%, ϵ = 14%, k = 1,083



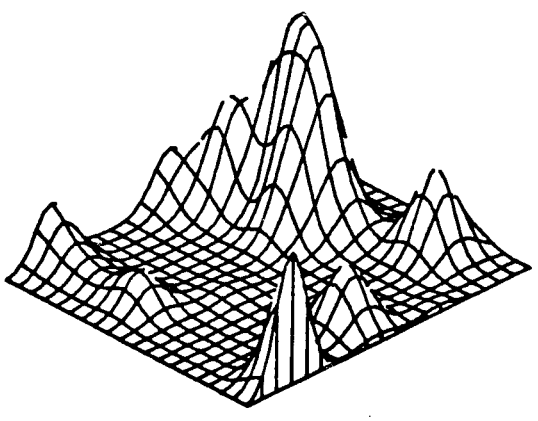
(c) N = 20%, ϵ = 20%, k = 947



(d) N = 50%, ϵ = 39%, k = 721



(e) N = 100%, ϵ = 83%, k = 348



(f) N = 200%, ϵ = 110%, k = 360

Figure 23. CFSPM solutions for various noise levels.

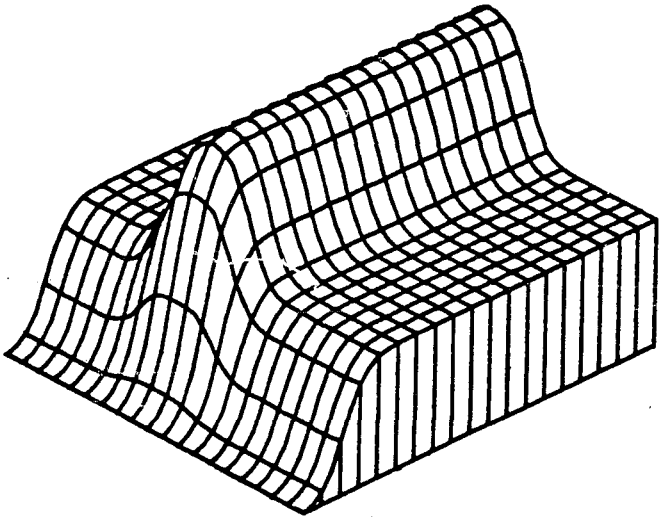
actually the sum of the pushing factors, and serves as a rough estimate of how many iterations CSPM would have required.) These examples clearly demonstrate the ability of the stopping criterion $SC6^k$ to compensate for variations in noise level. For 10% noise (Figure 23b) $SC6^k$ indicated the stopping point at $k = 1,083$, while for 200% noise (Figure 23f) it stopped the alogorithm at $k = 360$. (Recall from Chapter 3 that earlier stopping more effectively suppresses noise error and thus is desirable at higher noise levels.) The efficiency of CFSPM is illustrated by observing that for N = 10% CFSPM actually required only 51 interations, while CSPM would have required on the order of 1,000 iterations.

Although the least-squares solutions of ROG problems are extremely sensitive to noise (Chapter 2) and the projection algorithms produce extremely unstable solutions if a very large number of iterations are performed (Figure 22), Figure 23 demonstrates that ROG problems can be remarkably insensitive to noise if these algorithms are stopped after a reasonable number of iterations. This tolerance of noise is particularly high when the geologic features have highly contrasting attenuations, as is the case for Figure 23. This compensates to a large extent for the approximate nature of the ROG equations.

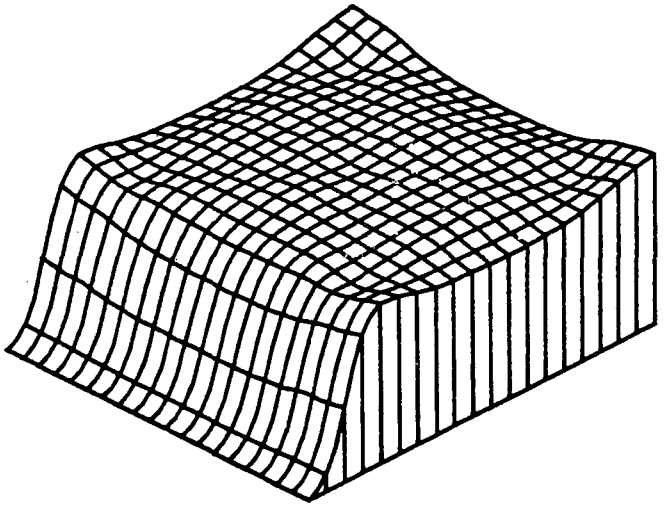
An attempt to detect a vertical slab via ROG is shown in Figure 24. Figure 24a is the model used to generate the data and Figure 24b is the solution obtained by the CFSPM algorithm. The algorithm was stopped by $SC6^k$. There were 81 cells and 81 paths, and the attenuation contrast was 2:1. No noise was added to the data. As expected, the data does not contain sufficient information to reconstruct the slab. The uniformly high attenuation of the solution in Figure 24b is a symptom of null-space error and indicates a need to take more measurements.

In Figure 25, models of horizontal layers with various attenuation contrasts (Figures 25a, c and e) are shown beside their corresponding CFSPM solutions (Figures 25b, d and f). For each figure the attenuation contrast is given. There were 81 cells and 81 paths, and the noise level was 10%. The solutions were bound non-negative, and the algorithm was stopped after the optimal number of iterations. It is evident from these figures that ROG has no problem detecting horizontal layers. As one would expect, the layers with higher attenuation contrasts are more easily reconstructed.

Figure 26a shows a model of a fault block with a 5:1 attenuation contrast, and Figures 26b and c show the solutions obtained for this problem by the CFSPM and CPM algorithms. There were 81 cells and 81 paths, and the noise

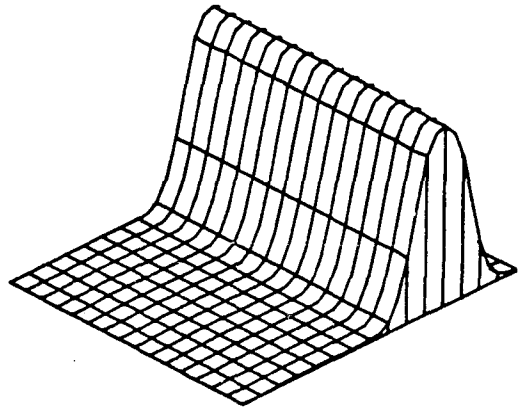


(a) test model

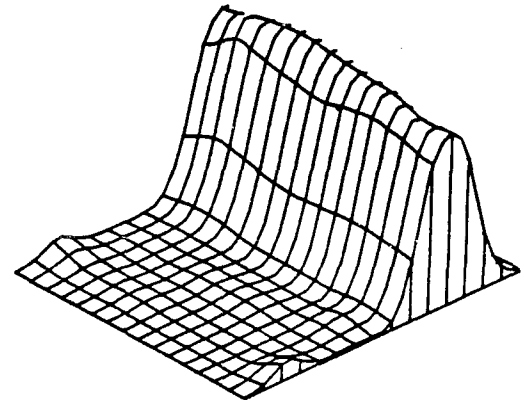


(b) CFSPM solution

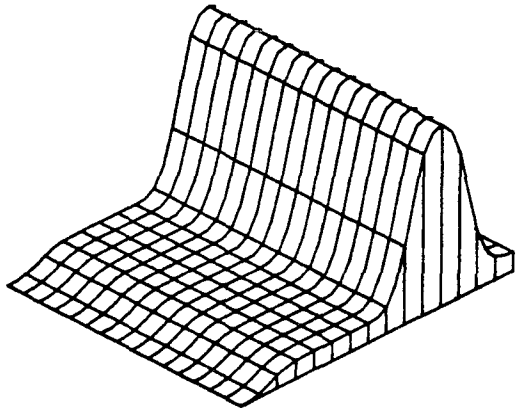
Figure 24. A vertical slab - an example of null-space error.



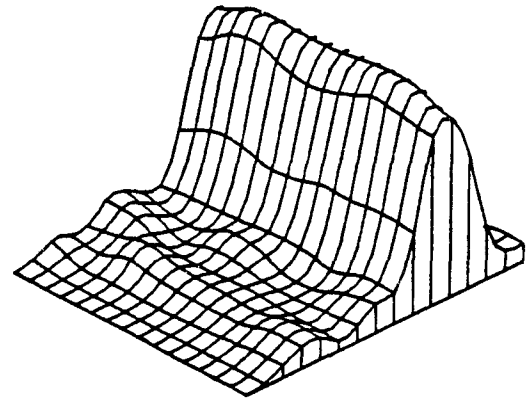
(a) test model, 100:1



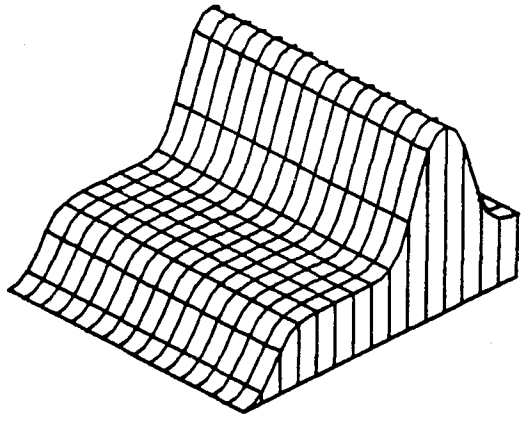
(b) CFSPM solution, 100:1



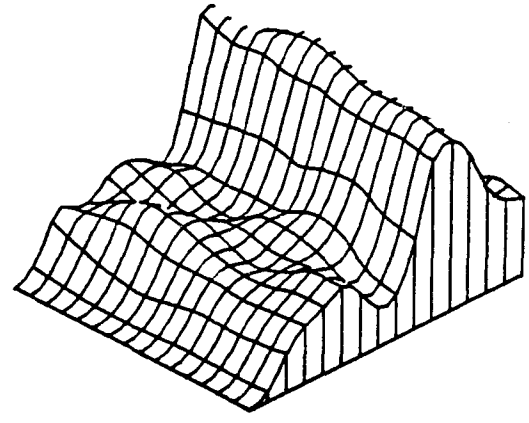
(c) test model, 10:1



(d) CFSPM solution, 10:1

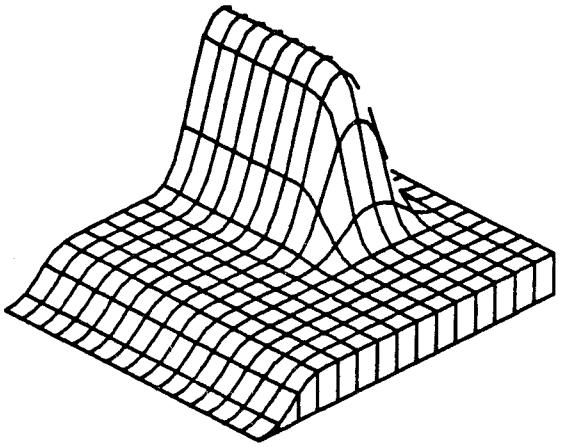


(e) test model, 3:1

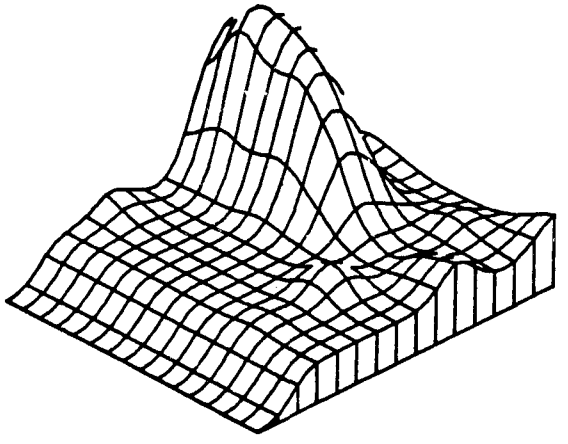


(f) CFSPM solution, 3:1

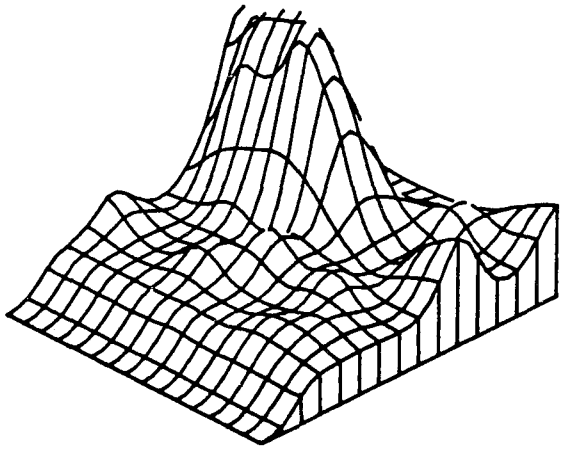
Figure 25. CFSPM solutions for horizontal layers.



(a) test model



(b) CFSPM solution
ϵ = 22%



(c) CPM solution
ϵ = 38%

Figure 26. CFSPM and CPM solutions for a fault block.

level was 20%. The inequality constraint was applied at the background level, and the algorithms were stopped after the optimal number of iterations. In addition to giving an example of a different type of geologic anomaly, Figure 26 illustrates that for most ROG problems CFSPM is noticeably superior to CPM. (Percent error was 22% for CFSPM, as opposed to 38% for CPM.)

4.3. Vector

The Volume Eddy Current Technique of Reconstruction (VECTOR) is a method of geotomography which, like ROG, probes the region between two boreholes by transmitting time harmonic electromagnetic radiation between them. VECTOR was proposed by Howard [27, 28] as an alternative to ROG, and it is sometimes referred to as wave-diffusion geotomography. While no pretense is made of treating VECTOR in depth here, a brief description is given and some results reported elsewhere are discussed. References [11, 27, 28] are detailed treatments of VECTOR.

The configuration for applying VECTOR is shown in Figure 27. VECTOR seeks to determine the currents induced in the earth by the sources. It does this by modeling each cell as an electrical monopole antenna (i.e., a cylindrical source, whose circular cross-section is shown for some of the cells in Figure 27). VECTOR then assigns excitation

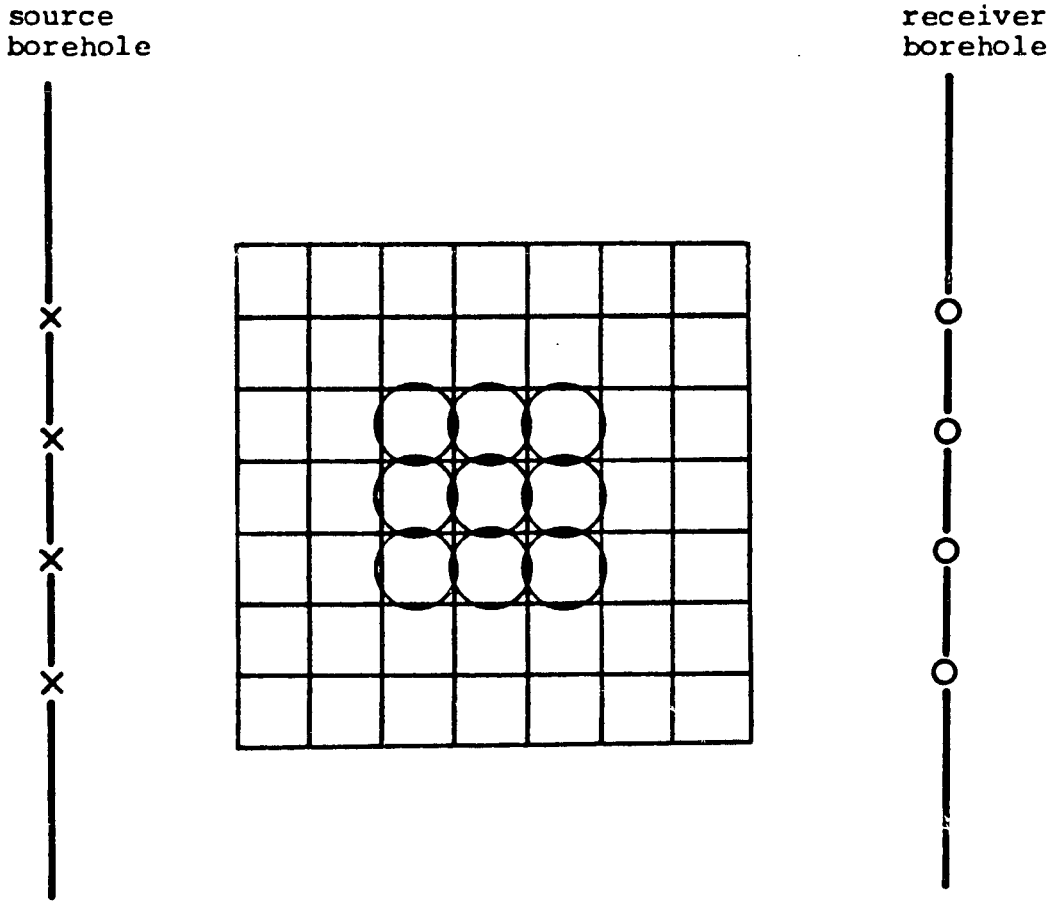


Figure 27. The geometry for VECTOR.

currents to each of these hypothetical antennas in such a way that the resultant field matches the field measured at the receiver borehole. This requires the solution of a linear system of the form

$$Ax = b. \tag{4.10}$$

The resulting profile of current values is an indicator of the actual eddy currents in the observation area. Under certain frequency restrictions, resitivity and permittivity profiles can be determined from the current profile [28].

VECTOR does not neglect conduction currents, so the electromagnetic theory it employs is more sophisticated than that employed in ROG. This is very satisfying from a theoretical point of view. It also avoids some of the limitations of ROG, such as error due to ray bending and the restriction to high frequencies. By lowering the frequency, signal attenuation can be reduced and probing range increased (Figure 15).

VECTOR, however, has its own unique limitations. For example, some error is introduced by modeling a square cell with a round one. Although VECTOR is theoretically valid at low frequencies, a trade-off inherent to electromagnetic probing is that decreasing the frequency to improve penetration reduces the ability to resolve fine detail. This is

easily seen by solving the forward problem (i.e., calculating the electric fields at the receivers which would result from a given geologic formation). If, for example, the boreholes are ten ambient wavelengths apart, then even a perfectly conducting anomaly centered in the interrogation area causes virtually no change in the received field unless it is larger than one wavelength [15]. Thus, to attain the high resolution required in many geotomography applications, the frequency must be high enough so that conduction currents are negligible and ray-optic approximations apply. A detailed discussion of frequency selection for VECTOR is given in [11].

In order to model the eddy currents adequately, it is necessary to have cells no more than one-tenth of an ambient wavelength wide. It is also advisable to take twice as many received field measurements as there are cells [28]. At frequencies high enough for high-resolution probing, this makes the A matrix so large that only trivially short-range problems can be solved, due to computer limitations. (Unfortunately, A is not sparse.) Various schemes for overcoming this problem are currently under investigation [28].

In computer simulations with a perfectly conducting cylindrical anomally, VECTOR has successfully calculated eddy currents which reproduce the measured data (i.e.,

solutions of (4.10) have been found for which the residual norm is very small) [28]. However, since A is known to be rank-deficient and ill-conditioned [28] (i.e., the equations do not completely describe the actual problem) there may be many widely varying solutions for which the residual is very small. (This is discussed in [11] and Chapter 2 of this paper.) Thus, it is not clear as yet whether VECTOR could in practice detect realistic geologic features. (Note that a perfectly conducting anomaly plotted in the manner of Figures 19-26 would have infinite height.)

While VECTOR is a promising new technique worthy of more research, further refinement and testing are necessary to determine whether it will be a practical means of short-range, high-resolution probing.

Chapter 5

CONCLUDING REMARKS

To date, the use of projection algorithms has been limited to a relatively small number of applications (primarily those related to tomography, where the immense size of the system dictates the use of efficient methods). This is largely due to a lack of understanding of the properties of these algorithms. Of particular concern to potential users of projection algorithms have been the following questions [31, 34]:

(1) Do these algorithms converge?

(2) If they converge, what do they converge to?

(3) If the algorithms are to be stopped prior to convergence, how is this justified and what is the optimal stopping point?

Much of this paper has been directed toward answering these questions.

Confusion about the convergence of projection algorithms can arise for several reasons. The cyclical manner in which PM converges for an inconsistent system can be interpreted as a failure to converge. In practice this is not a problem, however, since there is usually little variation in the solutions obtained in a given cycle. Any variation

that does occur is easily eliminated by averaging over the final cycle. Cyclic convergence can be avoided altogether by using SPM.

The fact that the limit to which PM converges is not well-defined for inconsistent systems has also been a source of concern. Analysis of PM (Section 3.3) indicates that it approximates the minimum-norm equally weighted least-squares solution if $x^1 = \underline{0}$. Experience verifies this, and many inconsistent tomography problems have been solved successfully by PM. SPM was rigorously proven to converge exactly to the minimum-norm equally weighted least-squares solution if $x^1 = \underline{0}$. The similarity of PM and SPM solutions (Figure 19) helps to confirm the validity of PM solutions.

The property of the projection algorithms which is probably the most disturbing when first encountered is that for noisy rank-deficient and ill-posed systems, these algorithms appear to diverge after initially attaining good approximate solutions. However, far from being a problem, this behavior represents one of the most valuable assets of the projection algorithms. This was demonstrated by showing that the minimum-norm, least-squares solution of a system of linear equations may be expressed as the sum of two components: x_e, which is stable with respect to perturbations of the equations; and x_c, which is unstable. Because the

projection algorithms reconstruct x_c more slowly then x_e, stopping these algorithms prior to convergence has the effect of imposing a stability constraint on the solution. Such a constraint is required when solving rank-deficient equations because they are extremely ill-posed. New criteria were proposed for estimating the optimal stopping point. These criteria, particularly $SC6^k$, were found to be superior to criteria previously developed in an ad hoc manner.

SPM was found to suppress noise error more effectively than PM, but it is much slower. FSPM, a modified version of SPM, was introduced here. While significantly faster than SPM, FSPM is slower and more complicated than PM. PM is simple, and very easy to implement.

The projection algorithms have several advantages over other methods currently used to calculate stabilized (or regularized) solutions of linear systems. They can easily impose inequality constraints, and can be made efficient for large, sparse systems. Furthermore, if stopping criterion $SC6^k$ is employed, the projection algorithms can compensate for unpredictable variations in the noise level by automatically imposing stronger stability constraints (i.e., by stopping earlier) when more noise is present. For these reasons, the projection algorithms are well suited for

many types of problems, including ROG. As more is learned about them, they promise to find wider application.

The test cases presented here verify that ROG is an effective method for short-range, high resolution probing in low-loss media such as salt or dry rock. This makes it highly suitable for nuclear waste disposal site evaluation, because salt is one of the preferred host rocks [36]. The problem of null-space error appears to be fairly easy to detect, and in most cases it is easily avoided by taking a sufficient number of measurements. When a stability constraint is imposed, ROG's high tolerance to noise offsets the approximate nature of its equations. ROG has an advantage over VECTOR in that it can employ the high frequencies required for high-resolution probing without the need to solve unmanagably large systems of equations. Furthermore, the ROG matrix is sparse, so its equations can be solved efficiently by projection algorithms. (The VECTOR matrix is not sparse.) ROG has demonstrated its ability to detect realistic geologic features both in computer simulations and in the field [12, 32, 33, 43].

APPENDIX 1

REVIEW OF LINEAR ALGEBRA

A1.1. The Pseudo-Inverse

There are several equivalent definitions of $A^{\dagger}$ in the literature. Penrose [42] defined $A^{\dagger}$ as the unique matrix satisfying the conditions

$$AA^{\dagger} = (AA^{\dagger})^{T}, \tag{A1.1}$$

$$A^{\dagger}A = (A^{\dagger}A)^{T}, \tag{A1.2}$$

$$AA^{\dagger}A = A, \tag{A1.3}$$

and

$$A^{\dagger}AA^{\dagger} = A^{\dagger}. \tag{A1.4}$$

Another, equivalent definition, is

$$A^{\dagger} = \lim_{\epsilon \to 0} (A^{T}A + \epsilon^{2}I)^{-1}A^{T} \qquad [1]. \tag{A1.5}$$

$A^{\dagger}$ may also be expressed in terms of the singular value decomposition (see Section A1.3), and this expression has been used to define $A^{\dagger}$ [48]. When $r(A) = n$ it is

well-known that

$$A^{\dagger} = (A^{T}A)^{-1}A^{T}, \tag{A1.6}$$

and in some limited treatments this expression also is used to define $A^{\dagger}$. It is important to note, however, that $A^{\dagger}$ exists for any A, even if $A^{T}A$ is singular.

Perhaps the most useful property of $A^{\dagger}$ is that $A^{\dagger}b$ is the minimum-norm, least-squares solution of the linear system $Ax = b$. $A^{\dagger}$ is also useful for calculating projections (Section A1.2). The following properties of $A^{\dagger}$ are well-known:

$$(A^{T})^{\dagger} = (A^{\dagger})^{T}, \tag{A1.7}$$

$$N(A^{\dagger}) = N(A^{T}) = R^{\perp}(A), \tag{A1.8}$$

and

$$R(A^{\dagger}) = R(A^{T}) = N^{\perp}(A). \tag{A1.9}$$

Finally, it follows easily from (A1.3) and (A1.4) that

$$A^{\dagger} = \underline{0} \text{ if and only if } A = \underline{0}. \tag{A1.10}$$

Treatments of the pseudo-inverse (also referred to as a generalized inverse or the Moore-Penrose inverse) may be found in [1, 2, 35, 42, 48].

A1.2. Orthogonality and Projections

The material reviewed here is well-known, and treatments of it may be found in [1, 48].

Two equally dimensioned vectors x and y are said to be orthogonal, denoted $x \perp y$, if $\langle x,y \rangle = 0$. Two subspaces S_1 and S_2 of E^n are said to be orthogonal, denoted $S_1 \perp S_2$, if $x \perp y$ for all $x \in S_1$ and all $y \in S_2$. If S is a subspace of E^k then the orthogonal complement of S in E^k, denoted $S^\perp$, is the unique subspace of E^k such that $S \cup S^\perp = E^k$, $S \cap S^\perp = \{\underline{0}\}$, and $S \perp S^\perp$.

The notation $E^k = S_1 \oplus S_2$ or $S_1 = E^k \ominus S_2$ is often used to indicate that S_1 and S_2 are orthogonal complements in E^k. It is well-known that for any $A \in E^{m \times n}$ the row space of A (i.e., the space generated by the row vectors of A) and $N(A)$ are orthogonal complements in E^n.

The concept of the projection of a vector onto a subspace may be derived from the projection theorem. The projection theorem states that if $E^k = S_1 \oplus S_2$ then for any vector x in E^k there are unique vectors $x_{S_1} \in S_1$ and

$x_{S_2} \in S_2$ such that $x = x_{S_1} + x_{S_2}$. The vectors x_{S_1} and x_{S_2} are the projections of x onto S_1 and S_2 respectively, and in this paper may be referred to simply as the components of x in S_1 and S_2. It is easily shown that

$$\|x\|^2 = \|x_{S_1}\|^2 + \|x_{S_2}\|^2. \tag{A1.11}$$

If $A \in E^{m\times n}$ then for any vector $x \in E^n$ the projection of x onto $N^{\perp}(A)$ is

$$x_{N^{\perp}(A)} = A^{\dagger}Ax; \tag{A1.12}$$

for any vector $b \in E^m$ the projection of b onto $R(A)$ is

$$b_{R(A)} = AA^{\dagger}b; \tag{A1.13}$$

and for any vectors x and $y \in E^n$ the projection of y onto x (i.e., the projection of y onto the space generated by x) is

$$y_x = \langle x,y\rangle \frac{x}{\|x\|^2}. \tag{A1.14}$$

Finally,

(A1.15) $$N^{\perp}(A) = R(A^T)$$

and

(A1.16) $$N(A) = R^{\perp}(A^T).$$

A1.3. The Singular Value Decomposition

It is well-known that for every matrix $A \in E^{m\times n}$ there is a singular value decomposition (SVD) having the form

(A1.17) $$A = V\Sigma U^T,$$

where V and U are orthogonal matrices (i.e., they have orthonormal column vectors), $m \times n$ matrix Σ satisfies

(A1.18) $$\Sigma(i,j) = \begin{cases} \sigma_i & \text{if } i = j \text{ and } i \leq r(A) \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$, and

(A1.19) $$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{r(A)} > 0.$$

Σ is uniquely determined for each A. The numbers $\sigma_1, \ldots, \sigma_{r(A)}$ are called the singular values of A, and they are the positive square roots of the non-zero eigenvalues of A^TA. U and V can be calculated from the eigenvectors of A^TA

[41]. Because U and V are orthogonal,

(A1.20) $$\|A\| = \|\Sigma\|$$

and

(A1.21) $$r(A) = r(\Sigma),$$

where $r(\Sigma)$ is simply the number of singular values [48]. $A^{\dagger}$ is readily obtained from the SVD of A by

(A1.22) $$A^{\dagger} = V\Sigma^{\dagger}U^{T},$$

where $\Sigma^{\dagger}$ is the matrix formed by replacing each non-zero element of Σ with its reciprocal [18]. The SVD of $A^{\dagger}$ can be obtained from (A1.22) simply by reversing the order of the columns of V and U and reversing the order of the non-zero elements of $\Sigma^{\dagger}$ to put them in order of decreasing size. If σ_{max} and σ_{min} denote the largest and smallest singular values of A then [48]

(A1.23) $$\|A\| = \sigma_{max}.$$

Because $\frac{1}{\sigma_{min}}$ is the largest singular value of $A^{\dagger}$, it follows that

(A1.24) $$\kappa(A) = \|A^{\dagger}\|\|A\| = \frac{\sigma_{max}}{\sigma_{min}}.$$

If $r(A) = 1$ then A has only one singular value, so $\sigma_{max} = \sigma_{min}$ and

$$\kappa(A) = 1. \tag{A1.25}$$

More comprehensive treatments of the SVD may be found in [18, 35, 41, 48].

A1.4. Miscellaneous Linear Algebra

Let $A \in E^{m\times n}$ have column vectors $C_1, \ldots, C_n$ and row vectors $a_1, \ldots, a_m$, and let $x \in E^n$. Then, by the rules of matrix-vector multiplication it is easily shown that

$$Ax = \sum_{j=1}^{n} x(j)C_j = \begin{bmatrix} \langle a_1, x\rangle \\ \vdots \\ \langle a_m, x\rangle \end{bmatrix}, \tag{A1.26}$$

and therefore

$$\|Ax\|^2 = \sum_{i=1}^{m} \langle a_i, x\rangle^2. \tag{A1.27}$$

APPENDIX 2

THEOREM PROOFS

Proof of Theorem 2.1

(The theorem is stated on p. 17.)

$x + p_i(x)$ satisfies the i'th equation of $A_o x = b_o$, so

$$\langle a_{oi}, x + p_i(x) \rangle = b_o(i)$$

and hence

$$\langle a_{oi}, p_i(x) \rangle = b_o(i) - \langle a_{oi}, x \rangle \tag{A2.1}$$

for $i = 1, \ldots, m$, where a_{oi} denotes the i'th row vector of A_o. Recall that $\|a_{oi}\| = 1$. By the Cauchy-Schwarz-Bunyakovski inequality,

$$\|p_i(x)\| \geq |\langle a_{oi}, p_i(x) \rangle| = |b_o(i) - \langle a_{oi}, x \rangle| \tag{A2.2}$$

with equality if and only if $p_i(x) = \alpha a_{oi}$ for some α. Because $p_i(x)$ is the vector of minimum norm of those satisfying (A2.1), the equality holds in (A2.2) and $\|p_i(x)\|^2 = (b_o(i) - \langle a_{oi}, x \rangle)^2$. Therefore, by (A1.26),

$$\sum_{i=1}^{m} \|p_i(x)\|^2 = \sum_{i=1}^{m} (b_o(i) - \langle a_{oi}, x\rangle)^2$$

$$= \|b_o - \begin{bmatrix} \langle a_{o1}, x\rangle \\ \vdots \\ \langle a_{om}, x\rangle \end{bmatrix}\|^2 = \|b_o - A_o x\|^2. \qquad \square$$

Proof of Theorem 2.2

(The theorem is stated on p. 26.)

If $\kappa(A) \geq \kappa_o$ then by Theorem 2.3 and equations (A1.24) and (A1.23) there is a matrix A' such that $r(A') = r(A) - 1$ and

$$\|A - A'\| = \sigma_{r(A)} = \frac{\sigma_1}{\kappa(A)} = \frac{\|A\|}{\kappa(A)} \leq \frac{\|A\|}{\kappa_o},$$

where σ_1 and $\sigma_{r(A)}$ are the largest and smallest singular values of A respectively. (Theorem 2.3 is stated on page 29. Its proof immediately follows this proof and does not employ Theorem 2.2.) It follows that A has weak rank with tolerance $\frac{\|A\|}{\kappa_o}$.

To show the converse, assume A has weak rank with tolerance $\frac{\|A\|}{\kappa_o}$. Then there is a matrix A' such that $r(A') < r(A)$ and $\|A - A'\| \leq \frac{\|A\|}{\kappa_o}$. Because

$$\dim(\mathcal{N}^{\perp}(A')) = r(A') < r(A) = \dim(\mathcal{N}^{\perp}(A)),$$

there is a vector x_0 such that $\|x_0\| = 1$ and $x_0 \in \mathcal{N}^{\perp}(A) \ominus \mathcal{N}^{\perp}(A')$ [48, p. 313], and

$$1 = \|x_0\| = \|A^\dagger A x_0\| \leq \|A^\dagger\|\|Ax_0\| = \|A^\dagger\|\|(A - A')x_0 + A'x_0\|$$

$$= \|A^\dagger\|\|(A - A')x_0\| \leq \|A^\dagger\|\|A - A'\|$$

$$\leq \frac{\|A^\dagger\|\|A\|}{\kappa_0} = \frac{\kappa(A)}{\kappa_0}.$$

Theorem 2.2 follows. □

Proof of Theorem 2.3

(The theorem is stated on p. 29.)

Part 1 follows immediately from (A1.24). To establish Part 2, note that

$$\text{(A2.3)} \qquad \delta A_k = A - A_k = U\Sigma V^T - U\Sigma_k V^T = U(\Sigma - \Sigma_k)V^T.$$

The SVD of δA_k may be obtained from $U(\Sigma - \Sigma_k)V^T$ by simply reordering the columns of U, $\Sigma - \Sigma_k$, and V in such a way that the elements of $\Sigma - \Sigma_k$ are arranged in order of decreasing size. Therefore, the largest singular value of δA_k is σ_{k+1} and by (A1.23), $\|\delta A_k\| = \sigma_{k+1}$. Part 3 is immediate, by (A1.21). It remains only to show that for $y \in N^\perp(\delta A_k)$ and $z \in N^\perp(A_k)$, $y^T z = 0$. By (A1.2) and (A1.12),

$$y^T z = (\delta A_k^T \delta A_k^{\dagger T} y)^T (A_k^T A_k^{\dagger T} z) = y^T \delta A_k^\dagger \delta A_k A_k^T A_k^{\dagger T} z$$

so it will suffice to show $\delta A_k A_k^T$ is zero. By (A2.3),

$$\delta A_k A_k^T = U(\Sigma - \Sigma_k)V^T(U\Sigma_k V^T)^T = U(\Sigma - \Sigma_k)\Sigma_k U^T.$$

Clearly $(\Sigma - \Sigma_k)\Sigma_k = \underline{0}$ so Theorem 2.3 follows. □

Proof of Theorem 2.4

(The theorem is stated on page 32.)

To establish Part 1 let $\delta b \in E^m$ and let δx denote $A^\dagger \delta b$. Note that because $N^\perp(\delta A') \perp N^\perp(A')$, $S_e \subset N(\delta A')$.

$$\|\delta b\| \geq \|\delta b_{R(A)}\| = \|AA^\dagger \delta b\| = \|A\delta x\| = \|A'\delta x - \delta A'\delta x\|$$

$$= \|A'\delta x_e - \delta A'\delta x_c\| \geq \|A'\delta x_e\| - \|\delta A'\delta x_c\|$$

$$\geq \|A'\delta x_e\| - \|\delta A'\|\|\delta x_c\|$$

so

$$\|A'\delta x_e\| \leq \|\delta b\| + \|\delta A'\|\|\delta x_c\|.$$

Therefore,

$$\|\delta x_e\| = \|A'^\dagger A'\delta x_e\| \leq \|A'^\dagger\|\|A'\delta x_e\|$$

$$\leq \|A'^\dagger\|(\|\delta b\| + \|\delta A'\|\|\delta x_c\|)$$

$$= \frac{\kappa(A')}{\|A'\|}(\|\delta b\| + \|\delta A'\|\|\delta x_c\|).$$

For Part 2, let $\delta x \in S_c \cap N^{\perp}(A)$ and let δb denote $A\delta x$. By (A1.12),

$$A^{\dagger}\delta b = A^{\dagger}A\delta x = \delta x.$$

Because

$$\|\delta b\| = \|A\delta x\| = \|A'\delta x - \delta A'\delta x\| = \|\delta A'\delta x\|$$

$$\leq \|\delta A'\|\|\delta x\|,$$

Theorem 2.4 follows. □

Proof of Theorem 2.6

(The theorem is stated on page 41.)

Let $\epsilon > 0$ be given and assume $r(A) < \min(m,n)$. Because

$$\dim(N(A)) = n - r(A) > 0$$

there is some vector $x_0 \in N(A)$ with unit norm. Because $r(A) < m$ there is a row vector of A, a_{i_o}, which is a linear combination of the other row vectors of A. Let δA denote the $m \times n$ matrix having $\epsilon \cdot x_0$ as its i_0'th row vector with zeroes elsewhere. By Lemma A1, which follows

this proof, δA satisfies Part 1. $N^{\perp}(\delta A)$ is the row space of δA, which is a subspace of $N(A)$, so δA clearly satisfies Part 2. The addition of δA to A renders its i_0'th row vector independent of the other row vectors, so Part 3 follows. $A + \delta A$ clearly has weak rank with tolerance ϵ, so Part 4 follows by Theorem 2.2 and (2.34). Part 5 is established by employing (2.34) and Parts 1 and 4 as follows:

$$\kappa(A + \delta A) = \|(A + \delta A)^{\dagger}\|\|A + \delta A\| \geq \frac{1}{\epsilon} \cdot \|A + \delta A\|$$

$$\geq \frac{\|A\| - \|\delta A\|}{\epsilon} = \frac{\|A\| - \epsilon}{\epsilon}. \qquad \square$$

Lemma A1. Suppose that $A \in E^{m\times n}$ has only one non-zero row vector, a_k. Then, $\|A\| = \|a_k\|$.

Proof. Let $x \in E^n$ and note that the vector Ax has only a single non-zero element, $\langle a_k, x\rangle$. Therefore,

$$\|Ax\| = |\langle a_k, x\rangle| \leq \|a_k\|\|x\|$$

by the Cauchy-Schwarz-Bunyakovski inequality. It follows that $\|A\| \leq \|a_k\|$, and it will suffice to show that $\|A\| \geq \|a_k\|$. Let $x_0 = \frac{a_k}{\|a_k\|}$, so $\|x_0\| = 1$. Then,

$$\|A\| = \|A\|\|x_0\| \geq \|Ax_0\| = |\langle a_k, x_0\rangle| = |\langle a_k, \frac{a_k}{\|a_k\|}\rangle| = \|a_k\|$$

and the lemma follows. □

Proof of Theorem 3.1

(The theorem is stated on page 88.)

Let $a_{o1}, \ldots, a_{om}$ denote the row vectors of A_o and let $\sum_{i=1}^{m}$ be abbreviated by $\sum_i$. Theorem 3.1 will be proven by first establishing the following series of lemmas:

Lemma A2. The sequence $\{\|d^k\|\}_{k=1}^{\infty}$ is monotonically decreasing, and hence convergent.

Proof. By (3.53) and (3.46), for $k \geq 1$

$$d^{k+1} = \frac{1}{m}A_o^T(b_o - A_o x^{k+1}) = \frac{1}{m}A_o^T(b_o - A_o(x^k + d^k))$$

$$= d^k - \frac{1}{m}A_o^T A_o d^k.$$

If d_i^k denotes the projection of d^k onto a_{oi} then by (A1.26) and (A1.14)

$$A_o^T A_o d^k = \sum_i a_{oi}\langle a_{oi}, d^k\rangle = \sum_i d_i^k$$

so

$$d^{k+1} = d^k - \frac{1}{m}\sum_i d_i^k = \frac{1}{m}\sum_i (d^k - d_i^k). \tag{A2.4}$$

However,

$$\|d^k - d_i^k\| = (\|d^k\|^2 - \|d_i^k\|^2)^{1/2} \leq \|d^k\|,$$

so

$$\|d^{k+1}\| = \|\frac{1}{m}\sum_i (d^k - d_i^k)\| \leq \frac{1}{m}\sum_i \|d^k - d_i^k\|$$

$$\leq \frac{1}{m}\sum_i \|d^k\| = \|d^k\|. \qquad \square$$

Lemma A3. $\lim_{k\to\infty} \|d^k\| = 0.$

Proof. By Equation (A2.4) of the Lemma A2 proof,

$$\|d^{k+1}\|^2 = \langle d^{k+1}, d^{k+1}\rangle = \langle d^k - \frac{1}{m}\sum_i d_i^k,\ d^k - \frac{1}{m}\sum_i d_i^k\rangle$$

$$= \|d^k\|^2 + \frac{1}{m^2}\|\sum_i d_i^k\|^2 - \frac{2}{m}\sum_i \langle d^k, d_i^k\rangle$$

and thus

$$m^2(\|d^k\|^2 - \|d^{k+1}\|^2) = 2m\sum_i \langle d^k, d_i^k\rangle - \|\sum_i d_i^k\|^2.$$

However,

$$\langle d^k, d_i^k\rangle = \langle d^k, a_{oi}\langle d^k, a_{oi}\rangle\rangle = \langle d^k, a_{oi}\rangle^2 = \|d_i^k\|^2$$

and

$$\|\sum_i d_i^k\|^2 \leq (\sum_i \|d_i^k\|)^2,$$

so

$$(A2.5) \qquad m^2(\|d^k\|^2 - \|d^{k+1}\|^2) = 2m\sum_i \|d_i^k\|^2 - \|\sum_i d_i^k\|^2$$

$$\geq 2m\sum_i \|d^k\|^2 - (\sum_i \|d_i^k\|)^2.$$

However, if α and β denote $\frac{1}{m}\sum_i \|d_i^k\|$ and $\frac{1}{m}\sum_i \|d_i^k\|^2$, respectively, then

$$\beta - \alpha^2 = \beta - 2\alpha^2 + \alpha^2 = \frac{1}{m}\sum_i (\|d_i^k\|^2 - 2\alpha\|d_i^k\| + \alpha^2)$$

$$= \frac{1}{m}\sum_i (\|d_i^k\| - \alpha)^2 \geq 0.$$

Therefore,

$$(\sum_i \|d_i^k\|)^2 = m^2\alpha^2 \leq m^2\beta = m\sum_i \|d_i^k\|^2$$

and it follows from (A2.5) that

$$m^2(\|d^k\|^2 - \|d^{k+1}\|^2) \geq m\sum_i \|d_i^k\|^2 \geq 0.$$

From Lemma A2 it follows that

$$\lim_{k\to\infty} \sum_i \|d_i^k\|^2 = 0.$$

But, by (A1.14) and (A1.26),

$$\sum_i \|d_i^k\|^2 = \sum_i \langle a_{oi}, d^k \rangle^2 = \|A_o d^k\|^2,$$

so

$$\lim_{k\to\infty} \|A_o d^k\|^2 = 0$$

and hence

$$\lim_{k\to\infty} \|A_o d^k\| = 0. \tag{A2.6}$$

Because $d^k \in N^\perp(A_o)$ for all k, (A1.12) may be employed to give

$$0 \leq \|d^k\| = \|A_o^\dagger A_o d^k\| \leq \|A_o^\dagger\| \|A_o d^k\|.$$

It follows from (A2.6) that

$$\lim_{k\to\infty} \|d^k\| = 0. \qquad \square$$

Lemma A4. If $x^1 \in N^\perp(A)$, then $\lim_{k\to\infty} x^k = A_o^\dagger b_o$.

Proof. Let b_o' denote the projection of b_o onto $R(A_o)$, and note that

$$R(A_o) = N^\perp(A_o^T) = N^\perp(A_o^\dagger),$$

so

$$A_o^T b_o = A_o^T b_o'$$

and

$$A_o^\dagger b_o' = A_o^\dagger b_o.$$

It is clear from (3.53) that $d^k \in R(A_o^T) = N^\perp(A_o)$, so $x^k \in N^\perp(A_o)$ for all k. Furthermore,

$$(b_o' - A_o x^k) \in R(A_o) = N^\perp(A_o^T),$$

so (3.53) and (A1.12) yield

$$m\|d^k\| = \|A_o^T(b_o - A_o x^k)\| = \|A_o^T(b_o' - A_o x^k)\|$$

$$\geq \frac{\|A_o^{T\dagger} A_o^T (b_o' - A_o x^k)\|}{\|A_o^{T\dagger}\|} = \frac{\|b_o' - A_o x^k\|}{\|A_o^\dagger\|} \geq \frac{\|A_o^\dagger b_o' - A_o^\dagger A_o x^k\|}{\|A_o^\dagger\|^2}$$

$$= \frac{\|A_o^\dagger b_o - x^k\|}{\|A_o^\dagger\|^2} \geq 0.$$

(Recall that A is non-zero so by (A1.10) $\|A_o^{T\dagger}\| = \|A_o^\dagger\| \neq 0$.) Lemma A4 follows, by Lemma A3. □

It will be proven by induction that if $\{x^k\}_{k=1}^\infty$ and $\{x_1^k\}_{k=1}^\infty$ are the sequences produced by the application of

SPM to $Ax = b$ with initial vectors x^1 and $x^1_{N^\perp(A)}$, respectively, then

$$x^k - x_1^k = x^1_{N(A)} \tag{A2.7}$$

for $k \geq 1$. Since

$$\lim_{k\to\infty} x_1^k = A_o^\dagger b_o$$

by Lemma A4, this will suffice to prove Theorem 3.1. For $k = 1$, (A2.7) is clearly true. Assume that

$$x^k - x_1^k = x^1_{N(A)}$$

for some integer $k \geq 1$ and let d^k and d_1^k denote the SPM corrections to x^k and x_1^k given by (3.53). Then

$$x^{k+1} - x_1^{k+1} = x^{k+1} - x_1^{k+1} - (x^k - x_1^k) + x^1_{N(A)}$$

$$= d^k - d_1^k + x^1_{N(A)} = \tfrac{1}{m}A_o^T(b_o - A_o x^k)$$

$$- \tfrac{1}{m}A_o^T(b_o - A_o x_1^k) + x^1_{N(A)} = \tfrac{1}{m} A_o^T A_o(x_1^k - x^k) + x^1_{N(A)}$$

$$= - \tfrac{1}{m}A_o^T A_o x^1_{N(A)} + x^1_{N(A)} = x^1_{N(A)}.$$

(Recall that, because A_o and A have the same row space, $N(A_o) = N(A)$ and hence $A_o x^1_{N(A)} = \underline{0}$.) □

APPENDIX 3

PROOF THAT THE ROG PROBLEM IS RANK-DEFICIENT

Theorem A3.1. Let D be the matrix arising from a straight-line ROG problem between two boreholes, as described in Chapter 4. (See Figure 14 for the geometry of the problem.) Further, let all cells be uniform rectangles. Then

$$r(D) \leq n - \frac{1}{2}(N_{vz} - 1),$$

where N_{vz} is the number of vertical zones into which the region is partitioned.

Proof. Because $r(D) = n - \dim(N(D))$, it will suffice to show that $N(D)$ contains a set of orthogonal vectors whose number is greater than or equal to $\frac{1}{2}(N_{vz} - 1)$.

Let the vertical zones be labeled by the integer k, where $1 \leq k \leq N_{vz}$, and let $x_k \in E^n$ denote the vector such that $x_k(j) = 1$ if the j'th cell is in the k'th vertical zone and $x_k(j) = 0$ otherwise. If w is the width of each cell and θ_i is the angle of the i'th path from the horizontal, then the i'th element of Dx_k is $w/\cos\theta_i$. Because Dx_k is independent of k, for any integers e and

f between 1 and N_{vz},

$$Dx_e = Dx_f.$$

Therefore,

$$D(x_e - x_f) = \underline{0}$$

and

$$(x_e - x_f) \in N(D).$$

If e, f, g, and h are distinct positive integers less than or equal to N_{vz}, then clearly

$$(x_e - x_f) \perp (x_g - x_h)$$

and

$$\{(x_e - x_f), (x_g - x_h)\} \subset N(D).$$

Because the number of pairs of x_k vectors which can be formed using each vector only once is greater than or equal to $\frac{1}{2}(N_{vz} - 1)$, Theorem A3.1 follows. □

REFERENCES

[1] Albert, A., "Regression and the Moore-Penrose Pseudoinverse", Academic Press, 1972.

[2] Ben-Israel, A. and T. N. E. Greville, "Generalized Inverses: Theory and Applications", John Wiley and Sons, New York, 1974.

[3] Bois, P., M. LaPorte, M. Lavergne and G. Thomas, "Well-to-Well Seismic Measurements", Geophysics, Vol. 37, No. 3, pp. 471-480, 1972.

[4] Bracewell, R., "Strip Integration in Radio Astronomy", Aust. J. Phys., Vol. 9, pp. 198-217, 1956.

[5] Bracewell, R. and A. Riddle, "Inversion of Fan-beam Scans in Radio Astronomy", The Astrophys. J., Vol. 150, pp. 427-434, 1967.

[6] Censor, Y., "Finite Series-Expansion Reconstruction Methods", Proc. IEEE, Vol. 71, No. 3, 1983.

[7] Censor, Y., P. B. Eggermont and D. Gordon, "Strong Underrelaxation in Kaczmarz's Method for Inconsistent Systems", Numer. Math., Vol. 41, pp. 83-92, 1983.

[8] Cormack, A., "Representation of a Function by its Line Integrals, with some Radiological Applications", J. Appl. Phys., Vol. 34, No. 9, pp. 2722-2727, 1963.

[9] Cormack, A., "Representation of a Function by its Line Integrals, with some Radiological Applications-II", J. Appl. Phys., Vol. 35, No. 10, pp. 2908-2913, 1964.

[10] Dines, K. and R. J. Lytle, "Computerized Geophysical Tomography", Proc. IEEE, Vol. 67, No. 7, pp. 1065-1073, 1979.

[11] Doerr, T., "An Analysis of Errors in the Algebraic Reconstruction Technique with an Application to Geotomography", Ph. D. Dissertation, University of Arizona, Tucson, Az., 1983. (A reprint appears in [12].)

[12] Doerr, T. A., C. E. Glass, D. G. Dudley and C. C. Lyon, "Ray-Optic Geotomography". In United States Nuclear Regulatory Commission Report No. NUREG/CR-3143, Vol. 3, 1983.

[13] Eggermont, P. P. B., G. T. Herman and A. Lent, "Iterative Algorithms for Large Partitioned Linear Systems, with Applications to Image Reconstruction", Linear Algebra and its Applications, No. 40, pp. 37-67, 1981.

[14] Gilbert, P., "Iterative Methods for the Three-dimensional Reconstruction of an Object from Projections", J. Theor. Biol., Vol. 36, pp. 105-117, 1972.

[15] Gill, Daniel, University of Arizona, Dept. of Electrical and Computer Engineering, Tucson, Az., personal communication.

[16] Glass, C., "Indirect Rock Mass Investigations for Optimized Borehole Drilling Programs", U. S. Nuclear Regulatory Commission Report No. NUREG/CR-3143, Vol. 1.

[17] Glass, C. E., D. G. Dudley and A. Q. Howard, Jr., "Summary of Wave Propagation Geophysical Techniques". In United States Nuclear Regulatory Commission Report No. NUREG/CR-3143, Vol. 2, 1983.

[18] Golub, G. and W. Kahan, "Calculating the Singular Values and Psuedo-Inverse of a Matrix", J. SIAM Numer. Anal., Ser. B, Vol. 2, No. 2, 1965.

[19] Golub, G. H. and C. Reinsch, "Singular value decomposition and least squares solutions", Numer. Math. Vol. 14, pp. 403-420, 1970.

[20] Golub, G. H. and C. Van Loan, "An Analysis of the Total Least Squares Problem", SIAM J. Numer. Anal., Vol. 17, No. 6, December 1980.

[21] Gordon, R., "A Tutorial on ART (Algebraic Reconstruction Techniques)", IEEE Transactions on Nuclear Science, Vol. NS-21, June 1974.

[22] Gordon, R., R. Bender and G. Herman, "Algebraic Reconstruction Techniques (ART) for Three-dimensional Electron Microscopy and X-ray Photography", J. Theor. Biol., Vol. 29, pp. 471-481, 1970.

[23] Herman, G., Editor, "Image Reconstruction from Projections-Implementation and Applications", Springer-Verlag, Berlin, Heidelberg, New York, 1979.

[24] Herman, G., Editor, "Scanning the Issue - the Special Issue on Computerized Tomography", Proc. IEEE, Vol. 71, No. 3, 1983.

[25] Herman, G. T., A. Lent and S. W. Rowland, "ART: Mathematics and Applications. A Report on the Mathematical Foundations and on the Applicability to Real Data of the Algebraic Reconstruction Techniques", Journal of Theoretical Biology, Vol. 42, pp. 1-32, 1973.

[26] Hilgers, J. W., "A Theory for Optimal Regularization in the Finite Dimensional Case", Linear Algebra and its Applications, No. 48, 1982.

[27] Howard, A. Q., Jr., M. R. Barclay and D. B. Henry, "A Low Frequency Electromagnetic Inversion Technique", University of Arizona Topical Report to the U. S. Nuclear Regulatory Commission, Contract No. NRC-04-78-269.

[28] Howard, A. Q., Jr., C. E. Glass, D. B. Henry, D. M. N'Geuessan and D. M. Siemers, "Wave Diffusion Geotomography". In United States Nuclear Regulatory Commission Report No. NUREG/CR-3143, Vol. 4, 1983.

[29] Jackson, D. D., "The use of *a priori* data to resolve non-uniqueness in linear inversion", Geophys. J. R. Astr. Soc., Vol. 57, pp. 137-157, 1979.

[30] Kaczmarz, S., "Angenäherte Auflösung von Systemen Linearer Gleichungen", Bull. Acad. Polon. Sci. Lett., A, pp. 355-357 , 1937.

[31] Katz, M. B., "Questions of Uniqueness and Resolution in Reconstruction from Projections", Lecture Notes in Biomathematics, Vol. 26, Springer-Verlag, Berlin, 1978.

[32] Lager, D. L. and R. J. Lytle, "Computer Algorithms Useful for Determining a Subsurface Electrical Profile via High Frequency Probing", Lawrence Livermore Lab. Rep. UCRL-51748, 1975.

[33] Lager, D. L. and R. J. Lytle, "Determining a subsurface electromagnetic profile from high-frequency measurements by applying reconstruction-technique algorithms", Radio Science, Vol. 12, No. 2, pp. 249-260, March-April 1977.

[34] Lakshminarayanan, A. and A. Lent, "Methods of Least Squares and SIRT in Reconstruction", J. Theor. Biol., Vol. 76, pp. 267-295, 1979.

[35] Lawson, C. and R. Hanson, "Solving Least Squares Problems", Prentice-Hall Inc., Englewood Cliffs, N. J., 1974.

[36] Lomenick, T. F., S. Gonzales, K. S. Johnson and D. Byerly, "Regional Geological Assessment of the Devonian-Mississippian Shale Sequence of the Appalachian, Illinois and Michigan Basins Relative to Potential Storage/Disposal of Radioactive Wastes", ORNL Report ORNL-5703, 1983.

[37] Lytle, R. J., K. Dines, E. Laine and D. Lager, "Electromagnetic Cross-borehole Survey of a Site Proposed for a Future Urban Transit Station", Lawrence Livermore Lab Rep. UCRL-52484, 1978. (Available from National Technical Information Services, Springfield, VA.)

[38] Lytle, R. J. and K. Dines, "Iterative Ray Tracing Between Boreholes for Underground Image Reconstruction", IEEE Trans. Geos. and Rem. Sens., Vol. GE 18, No. 3, 1980.

[39] Moris, J., "On a simple method for solving simultaneous linear equations by a successive approximation process", Journal of the Roy. Aero. Soc., April, 1935.

[40] Morris, J., "A Successive Approximation Process for Solving Simultaneous Linear Equations", Aeronaut. Res. Com. Rep. #1711 (1936), pp. 1-12.

[41] Noble, B. and J. W. Daniel, "Applied Linear Algebra", Prentice-Hall Inc., Englewood Cliffs, N. J., 1977.

[42] Penrose, R., "A Generalized Inverse for Matrices", Cambridge Philosophical Society Proceedings, Vol. 51, pp. 406-413, 1955.

[43] Radcliff, R. D. and C. A. Balanis, "Reconstruction Algorithms for Geophysical Applications in Noisy Environments", Proceedings of the IEEE, Vol. 67, No. 7, July 1979.

[44] Radon, J., "Uber die Bestimmung von Funktionen durch Ihre Interralwerte Laengs Gewisser Mannigfaltigkeiten "(On the determination of functions from their integrals along certain manifolds)", Berichte Saechsische Akad. Wissenschaft Leipzig, Math. Physics Klass, Vol. 69, pp. 262-271, 1917. (A reprint appears in [46].)

[45] Scudder, H., "Introduction to Computer Aided Tomography", Proc. IEEE, Vol. 66, No. 6, pp. 628-637, 1978.

[46] Shep, L. A., Editor, "Computed Tomography", Proeedings of Symposia in Applied Mathematics, Vol. 27, American Mathematical Society, Providence, Rhode Island, 1983.

[47] Sluis, A., "Stability of the Solutions of Linear Least Squares Problems", Numer. Math., Vol. 23, pp. 241-254, 1975.

[48] Stewart, G., "Introduction to Matrix Computations", Academic Press, 1973.

[49] Stratton, J., "Electromagnetic Theory", McGraw-Hill, 1941.

[50] Sweeney, D. and C. West, "Reconstruction of Three-dimensional Refractive Index Fields from Multidirectional Interferometric Data", Applied Optics, Vol. 12, No. 11, pp. 2649-2664, November 1973.

[51] Tanabe, K., "Projection Method for Solving a Singular System of Linear Equations and its Applications", Numer. Math., Vol. 17, pp. 203-214, 1971.

[52] Tikhonov, A. and V. Arsenin, "Solutions of Ill-Posed Problems", John Wiley and Sons, 1977.

[53] Twomey, S., "Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements", Elsevier Scientific Publishing Co., 1977.

[54] Wait, J. R., "Note on the Theory of Transmission of Electromagnetic Waves in a Coal Seam", Radio Science, Vol. 11, No. 4, pp. 263-265, 1976.

[55] Whitney, T. M. and R. K. Meany, "Two Algorithms Related to the Method of Steepest Descent", SIAM J. Numer. Anal., Vol. 4, No. 1, 1967.

[56] Wilkinson, J. H., "The Algebraic Eigenvalue Problem", Oxford University Press, 1965.

[57] Young, D., "On Richardson's Method for Solving Linear Systems with Positive Definite Matrices", J. Math. Phys., Vol. 32, pp. 243-255, 1954.