A Partial-Step Algorithm for the Nonlinear Estimation Problem

DALE H. BOGGS*
Jet Propulsion Laboratory, Pasadena, Calif.

The "normal equations" solution to the weighted least-squares estimation problem is recast in terms of the singular-value decomposition of the equations-of-condition matrix, A. The superior convergence properties of rank-deficient, pseudoinverse solutions in the presence of nonlinearities and ill-conditionedness of the normal matrix A^TA are examined. An extension of this partial-step technique, which makes use of part of the information contained along "eigen-directions" associated with singular values previously ignored in the standard rank-deficient solution, is presented. Results are given showing the relative convergence powers of the methods in obtaining a solution to the orbit determination problem for a simulated Martian orbiter trajectory. These results indicate that the extended partial-step method will be of value during the orbit phase of the Mariner Mars '71 mission.

Introduction

THE task of obtaining an adequate weighted least-squares solution for the nonlinear estimation problem using an approximating linear model is often a troublesome aspect of orbit determination. If the magnitude of the physical nonlinearities encountered is significant when compared to the noise in the data, the classical Gaussian iterative differential-correction technique may yield an incorrect or divergent result, depending on the initial error in the parameters being estimated. As the nonlinearities become smaller relative to the data noise, the problem becomes less pronounced. Thus, for many trajectory configurations, such as a translunar phase, the linear approximations made in the equations-of-condition and the state transition are quite sufficient, since the neglected second- and higher-order terms are relatively small. However, the nonlinearities associated with planetary and lunar orbiters, for example, are very significant and could well be ruinous to the standard linear batch data filter. The effect of these nonlinearities will typically be most critical at the time of spacecraft reacquisition by tracking stations following injection into orbit about the target body, because of the relatively large uncertainties in the spacecraft state due to orbit insertion maneuver errors.

A closely related, albeit distinct, problem associated with batch data filtering is the inherent loss of precision in the formation of the "normal matrix" A^TA where $A = \partial z/\partial x$ is the matrix of partial derivatives of observables with respect to estimated parameters. It is well known[1] that the conditioning of A^TA is worse than that of A; in fact, the spectral condition number of A^TA is the square of the condition number of A. Thus if the linear system of equations associated with A is ill-conditioned, the normal matrix is even more ill-conditioned. Furthermore, the accuracy of a matrix inverse computation is, in general, dependent (inversely) on the condition number of the original matrix. Therefore, since the commonly used "normal equations" solution to the linear least-squares problem requires the inverse of the normal matrix, the accuracy of this solution is sensitive to any magnification of the original system's ill-conditionedness. This effect becomes more pronounced as the statistical correlations of the estimated parameters increase: in fact, this increase in correlations corresponds qualitatively to a higher degree of linear dependence associated with the columns of A; i.e., a larger condition number.

The loss of precision intrinsic to the formation and subsequent inversion of the normal matrix irretrievably degrades the information content of the original data. That is, the computation of this matrix may lead to numerical difficulties which cannot be removed by any conditioning process applied to A^TA once it has been formed. Furthermore, experience has shown that, at least for orbital estimation applications of least-squares, the deleterious effects of ill-conditionedness are usually magnified by larger nonlinearities. The net result, therefore, of neglected significant nonlinear terms and the formation of the normal matrix, interacting with a relatively large uncertainty in the spacecraft position and velocity, can be the inability to obtain a least-squares solution by iterative refinement with the normal equations differential corrector.

The loss of precision in forming A^TA has led numerical analysts to search for other formulations of the linear least-squares problem. One algorithm for solving the problem using Householder transformations[2] is given by Businger and Golub[3] and extended by Hanson and Lawson.[4] In particular, they apply orthogonal transformations to A to reduce it to a square matrix of dimension equal to the number of solve-for parameters; the solution is then given in terms of the inverse of this reduced matrix. This technique, therefore, transforms an over determined system of equations into an exactly-determined system. The condition number of A is conserved during the reduction process. Thus, assuming that the same precision arithmetic is used for all computations, the orthogonal method would appear to preserve more accuracy than the normal equations method since the latter involves computations with a matrix whose condition number is the square of A's. This, however, is an oversimplification. Golub and Wilkinson[5] give an upper bound for the error in the solution to the linear (single iteration) least-squares problem as a function of the condition number of A. This bound, which is independent of the method of solution, contains a term in $\text{cond}^2(A)$ as well as one in $\text{cond}(A)$. Therefore, even if the orthogonal technique is used, the value of $\text{cond}^2(A)$ is to some extent still relevant. The use of orthogonal transformations nonetheless avoids some of the ill effects inherent in using the normal equations.

A related technique using orthogonal methods is provided by the singular-value decomposition (SVD) algorithm.[4,6,7] The basis of the method is a spectral decomposition of the A matrix from which the pseudoinverse[8] A^+ of A is computed. For the special full-rank case, Rank $(A) = n \equiv$ number of columns of A ($\leq$ number of rows), A^+ has the property

$$A^+ = (A^TA)^{-1}A^T$$

which is the classical batch least-squares filter (disregarding weights) to be applied to the data vector z. The resulting solution is sometimes referred to as the full-rank or full-step solution in nonlinear convergence techniques. If A is not of full rank, A^+ is still well-defined and meaningful; the unique (minimum-length)

Presented as Paper 71-121 at the AIAA 9th Aerospace Sciences Meeting, New York, January 25–27, 1971; submitted February 16, 1971; revision received October 4, 1971. This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract NAS 7-100, sponsored by NASA. The author gratefully acknowledges the contribution made by S. R. McReynolds during the initiation of this work.

Index categories: Computer Technology and Computer Simulation Techniques; Navigation, Control, and Guidance Theory.

* Senior Research Engineer, Precision Orbit Determination Group, Tracking and Orbit Determination Section of the Mission Analysis Division.

solution to the least-squares problem is still given by A^+z (Ref. 9). It is this general property of the pseudoinverse that will be examined and further exploited.

In most applications of linear least-squares the possibility that the rank of A as stored in a computer of, say, 16 decimal places precision, is less than n is quite remote (for an overdetermined system). However, considering the matrix to be of full rank for the purpose of computing the full (classical) step could be hazardous to convergence if, again, the condition number of A is large and the nonlinearities are significant. A fundamental result of Lawson-Hanson[4] is the recognition that a meaningful solution to this least-squares problem can be obtained by computing a less than full-rank solution even when the (numerically computed) rank of A is n. Thus they compute a suitable r-rank ($r \leq n$) approximation to A and A^+ (r is determined from a tolerance parameter) using only the r largest singular values associated with the spectral decomposition of A. The r-rank partial-step solution is then computed as A_r^+z, where A_r^+ is the r-rank approximator for the "true" pseudoinverse. As they show, the rank of A is therefore an essentially irrelevant concept with regard to least-squares filtering. The relevant parameter is the rank of the matrix which is determined to be a permissible replacement for A in the computation of A^+. If the data is not precise due to various error sources, this approximation is justified if the system of linear equations associated with the problem is badly conditioned; in this case it becomes more meaningful to compute the solution of an appropriate neighboring system of lower rank. Also, in the presence of significant nonlinearities, judiciously chosen lower-rank solutions in an iterative process can often yield the solution when divergence might otherwise occur.[10]

In the formulation of the aforementioned r-rank partial-step method, the information associated with the $n - r$ disregarded smaller singular values is not used in computing the solution. This and other considerations led to the development of a related algorithm, which instead of completely ignoring the smaller singular values in the rank-deficient solution, makes some use of this neglected information. The basis of this extended rank-deficient scheme is the inclusion of part of the information contained in A along eigendirections associated with the previously ignored singular values. These eigenvectors are in the (solution) parameter space and are generated as part of the SVD of A. The portion of the information taken along these neglected directions in the n-dimensional parameter space is determined meaningfully from the component of an a priori parameter error covariance ellipsoid lying along the respective directions. The least-squares problems where this extended partial-step algorithm can and should be used are those where (realistic) a priori knowledge concerning the error in the initial state estimate exists. This is nearly always the case in satellite orbit determination.

In this paper the extended partial-step algorithm is developed and its use in solving some nonlinear and ill-conditioned orbital least-squares estimation problems (with numerical experience) is described. Convergence results are given showing favorable comparisons between the extended method, the strict partial-step method, and the classical normal equations (full-rank) method. It should be noted that while only trajectory estimation problems are discussed herein, the results should apply qualitatively to noncelestial applications where similar convergence problems exist.

Linear Least-Squares (LLS) Problem

The unweighted LLS problem can be defined as: given an $m \times n$ matrix A of rank k, and an m-vector of data z, find an n-vector x_s which minimizes $\|Ax - z\| = [(Ax - z)^T(Ax - z)]^{\frac{1}{2}}$, the euclidean length of $Ax - z$. For present purposes it is assumed that all quantities are real-valued, and that $m \geq n$; i.e., the system is overdetermined. Thus, $k \leq n$. These restrictions simplify the development; they are not necessary to it. To recover the weighted LLS problem, z and A are weighted by premultiplication by $W^{\frac{1}{2}}$, where $(W^{\frac{1}{2}})^TW^{\frac{1}{2}} = W$ is a symmetric positive-definite data weighting matrix; i.e.,

$$z \to W^{\frac{1}{2}}z \qquad A \to W^{\frac{1}{2}}A$$

This has the effect of changing the standard Euclidean norm $\|Ax - z\|^2$ which is to be minimized to a more general weighted Euclidean norm

$$\begin{aligned}\|Ax - z\|^2W &\equiv (Ax - z)^TW(Ax - z)\\ &= (W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}z)^T(W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}z)\\ &= \|W^{\frac{1}{2}}Ax - W^{\frac{1}{2}}z\|^2\end{aligned}$$

Thus, the aforementioned LLS definition encompasses the weighted linear least-squares problem if z and A are properly redefined. As applied to linearized orbit estimation, z is the vector of weighted data residuals; i.e., observed-minus-computed values of the observational data, x_s is the incremental correction Δx to be found and added to the initial estimate X_0 of the state to yield an "improved" estimate X_1, and A is the matrix of weighted partial derivatives $\partial f/\partial x$ of the observations f.

To obtain the form of the LLS problem write the differential expression relating x to z

$$z = Ax + \tfrac{1}{2}h + \cdots + \varepsilon \tag{1}$$

where the ith component of h is given by the second-order term

$$h_i = x^T\frac{\partial^2 f_i}{\partial x_j\,\partial x_k}x = \sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 f_i}{\partial x_j\,\partial x_k}x_jx_k \tag{2}$$

ε is a vector random variable representing the (stochastic) error in the data z. Equation (1) results from the Taylor expansion of $f(X_0 + x)$ about the point X_0; retaining only the linear term in x yields

$$\varepsilon = z - Ax \tag{3}$$

(The linear system $z \approx Ax$ is called the equations-of-condition.) The form of the LLS problem defined earlier will now result if we require the solution x_s to be that x which minimizes the quadratic form

$$Q(x) = \varepsilon^T\varepsilon \tag{4}$$

i.e., the sum of squared errors.

The LLS problem always has at least one solution; the solution is unique if $k = n$. If $k < n$, there exists a family of solutions within which can be found a unique vector of minimum euclidean length. Thus, with this constraint, the problem always has a unique solution x_s (Ref. 4). If A is of full rank; i.e., if $k = n$, then the well-known unique solution to the LLS problem is

$$x_s = (A^TA)^{-1}A^Tz \tag{5}$$

The solution (5) to the linearized problem may not adequately solve the actual problem embodied in Eq. (1); in fact, if the magnitude of the neglected nonlinear terms are significant relative to Ax, the resulting solution may be outside the linear solution region of the point-slope method. In any case, since Ax is a linear approximation to z, the solution to the actual problem is usually found by an iterative process whereby x_{i+1} is computed from x_i, $A(x_i)$, $z(x_i)$, for $i = 0, 1, 2, \ldots$. Convergence of this process is usually assumed to have occurred when $\|Ax_i - z\|$ becomes smaller than some prescribed tolerance. If any x_i falls outside the linearity region, the iterative process may yield an erroneous solution; i.e., some other local minimum of $Q(x)$, or it may completely diverge.

Singular-Value Decomposition

An important spectral representation of the matrix A is given by Theorem 1: the $m \times n$ matrix A has the decomposition

$$A = \sum_{i=1}^{n}\lambda_iu_iv_i^T \tag{6}$$

1) the n vectors v_i are eigenvectors associated with the n eigenvalues λ_i^2 of A^TA, and are orthonormal; 2) the n vectors u_i are eigenvectors associated with the n eigenvalues λ_i^2 of AA^T, and are orthonormal; 3) the vectors u_i and v_i satisfy the equations

$$Av_i = \lambda_i u_i \quad (i = 1, 2, \ldots, n) \tag{7}$$

The proof of this theorem can be found in Ref. 6.

The representation (6) can be written in matrix form by defining matrices U, S, V, with respective dimensions $m \times k$, $k \times k$, and $n \times k$ as [k = rank (A)]

$$U \equiv [\, u_1 \,|\, u_2 \,|\, \cdots \,|\, u_k \,], \quad V \equiv [\, v_1 \,|\, v_2 \,|\, \cdots \,|\, v_k \,]$$

$$S \equiv \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k),$$

where the eigenvalues-vectors have been re-indexed so that $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k > 0$. Then

$$A = USV^T \tag{8}$$

To heuristically relate the decomposition (6) to the linear estimation problem, the v_i are n orthogonal eigendirections spanning the n-dimensional parameter vector space X, and the u_i are the eigendirections in an n-dimensional subspace of the m-dimensional data vector space Z. The k u_i's associated with the nonzero eigenvalues span the column (range) space of the matrix A – a k-dimensional subspace of Z. If the rank of A is full ($k = n$), all of the estimated parameters x can be "seen" by the data. If $k < n$, the $n - k$ eigenvectors v_i for which $Av_i = 0$ span a subspace of X which cannot be "seen" by the data: the linear combinations of the x's corresponding to these v_i's cannot be estimated with the available data.

The scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ are called the (nonzero) singular values of A; for a given A they are unique. The matrix spectral norm of A, $\|A\|$, is defined as the scalar

$$\|A\| \equiv \max_{\|x\|=1} \|Ax\| = \max_{\|x\|=1} (x^T A^T A x)^{\frac{1}{2}}$$

i.e., the maximum value attained by the magnitude of the vector Ax over all possible vectors x of length 1. Thus $\|A\| = \lambda_1$. If rank(A) = $k > 0$ the spectral condition number of A is defined as

$$\text{cond}(A) \equiv \lambda_1/\lambda_k$$

Thus

$$\text{cond}(A^T A) = \text{cond}(AA^T) = [\text{cond}(A)]^2$$

Reference 7 gives an efficient numerical procedure for the computation of the singular-value decomposition of A.

Rank-Deficient Solutions to the LLS Problem

The second important result needed for the development of the partial-step method is the generalization of the inverse of a square nonsingular matrix to a nonsquare matrix of arbitrary rank.

Theorem 2: If A is an $m \times n$ matrix, then there exists a unique $n \times m$ matrix A^+ such that 1) $AA^+A = A$, 2) $A^+AA^+ = A^+$, 3) $(AA^+)^T = AA^+$, 4) $(A^+A)^T = A^+A$. A^+ is called the pseudo-inverse of A and has a number of important properties, three of which are

$$A^+ = VS^{-1}U^T \tag{9}$$

$$A^+ = (A^T A)^{-1} A^T \quad \text{and} \quad A^+ A = I_n \tag{10}$$

for the special case rank(A) = n; and finally, the general result is that the unique minimum-length solution to the LLS problem is

$$x_s = A^+ z \tag{11}$$

for any rank(A). Thus, the special (full-rank) case (5) follows from (10) and (11). From (10) we also see that $A^+ = A^{-1}$ if $m = n$ and A is nonsingular. The proof of Theorem 2 and properties (10) and (11) can be found in Ref. 9. Property (9) is easily established by verifying that $VS^{-1}U^T$ satisfies the conditions 1–4 of Theorem 2 for A^+.

Equation (9) can be written in vector form as

$$A^+ = \sum_{i=1}^{k} \lambda_i^{-1} v_i u_i^T \tag{12}$$

Hence Eq. (11) becomes

$$x_s = \left(\sum_{i=1}^{k} \lambda_i^{-1} v_i u_i^T \right) z = \sum_{i=1}^{k} p_i v_i \tag{13}$$

where

$$p_i = u_i^T z/\lambda_i = (u_i, z)/\lambda_i \tag{14}$$

with (.,.) denoting the dot product. Thus the scalar coefficient p_i in the eigenvector expansion of x_s given by Eq. (13) is proportional to the component of the data vector z lying along the u_i eigen-direction of Z. The solution given by Eq. (13) minimizes $\|Ax - z\|$ even when z does not lie entirely in the range (column) space of A, because (13) yields a vector x_s such that Ax_s is the orthogonal projection of z onto that range space. To see this use Eqs. (7), (13), and (14) to compute

$$Ax_s = \sum_{1}^{k} p_i (Av_i) = \sum_{1}^{k} p_i \lambda_i u_i$$

$$= \sum_{1}^{k} (u_i, z) u_i$$

which is the component of z lying in the range of A. Thus, Ax_s is as "close" to z as it can be made, or, in other words, the sum of the squares of the errors is a minimum. The component of z perpendicular to the range of A is

$$z - Ax_s = \sum_{k+1}^{m} (u_i, z) u_i$$

so that

$$\|z - Ax_s\| = \left[\sum_{k+1}^{m} (u_i, z)^2 \right]^{\frac{1}{2}}$$

which is the resulting value of the function being minimized.

As mentioned earlier, the possibility that the rank of A as stored in a computer is less than n is very remote. However, considering the matrix to be of full rank for the purpose of computing x_s according to Eq. (13) could be hazardous if the condition number of A is large. It is clear from Eq. (14) that errors in the observation z lying along eigendirections associated with smaller singular values are magnified more than error components corresponding to larger singular values. This error distortion is obviously intensified as the ratio of the larger to the smaller singular values increases. Hence, it therefore seems likely that a suitable approximation to A of the form (6) could be made by extending the sum over only the first r singular values, where r is the first (ordered) singular-value index such that $\lambda_{r+1}/\lambda_1 < \varepsilon$, with ε being a given tolerance parameter. r is called the pseudorank by Hanson-Lawson[4] and is the rank of the matrices used to approximate A and A^+ in their computation of x_s according to Eq. (13). The resulting approximator to the matrix A is, in essence, an appropriate representative member of a family of matrices, $\mathscr{A}$, all of which are permissible replacements for the "true" A. Among the reasons for the existence of the nonvoid set $\mathscr{A}$ for the linear estimation problem are the following.

1) The components of both A and z are the result of measurements and computer computations of limited precision, and, furthermore, these results are stored in finite-length computer words.

2) The matrix A arises from the linearization of the nonlinear function $z(x)$, and the validity of Ax_s as an approximator for $z(x_s)$ generally decreases as $\|x_s\|$ increases. It is easily shown that $\|x_s\|$ decreases as lower-rank approximators for A^+ are used in computing x_s. Thus, if a small change in A produces a relatively large reduction in $\|x_s\|$, such a change may be desirable.[4]

3) Columns of A may be nearly dependent indicating a correlation between parameters that is too high to allow them to be distinguishable using the available data z. Thus, the rank of A is an essentially irrelevant concept with regard to the LLS problem. The significant parameter is the rank of the matrix which replaces A.

Extension of the Rank-Deficient Technique

In general, the use of the strict rank-deficient method for solving an LLS problem beset with nonlinearities or ill-conditionedness, or both, is not straightforward. The user may wish to compute many of the different-rank solutions during each iteration by varying the pseudorank r of A^+; he must then choose the

pseudorank which is "best" (for each iteration) based upon criteria such as the predicted or computed decrease in the function $\|z - Ax\|$ resulting from the different solutions. This decision must be tempered by the condition number of A^+ which results from the different pseudoranks

$$\text{cond}(A_r^+) = \lambda_1/\lambda_r$$

which must be maintained below a certain level depending on the precision of the arithmetic and the magnitude of the data error. Or the user may simply choose an ε at the outset of the problem, and hope that the resulting ε pseudoranks chosen automatically during each iteration are suitable. Usually ε is picked as a function of the data noise, with larger ε's (smaller pseudoranks) associated with larger data uncertainties. In either case, the user must make somewhat subjective decisions for controlling the success of the convergence.

An investigation by Dyer, McReynolds, and Hanson[10] of the convergence difficulties caused by nonlinearities associated with a (simulated) Mars-orbiter was conducted using the rank-deficient solution technique discussed in the previous section. The nonlinearities they confronted approximate those that will be present during the orbital phase of Mariner '71. Their study consisted essentially of computing varied-rank solutions by successively setting $k = n, n-1, \ldots$, etc., in Eq. (13), and then selecting the highest-rank solution x_s that resulted in $\|z(x_s)\|$ being less than the original value of $\|z\|$, the norm of the weighted-residual vector computed before the iterative refinement process began. This procedure was repeated for each iteration and basically amounted to a search for the optimum-rank solution to be used as each particular iteration's correction to the state. Typically, if the initial errors assumed in the values of the estimated parameters were not small, the first few iterations forced low-rank solution steps to be taken. As the accumulated solution got closer to the "true" state, pseudoranks used in their scheme approached n. Unfortunately, their method required many more $z(x)$ (residual function) evaluations than $A(x)$ (partial derivative) evaluations; in fact, $z(x)$ had to be recomputed for each of the varying-rank solutions x_s that were found during an iteration. During an actual mission (or a simulation of a mission) the computer time needed by an orbit determination program such as JPL's DPODP[11] to compute $z(x)$ is significant—usually approaching that needed to find $A(x)$. Thus these extra function evaluations would normally be a luxury too expensive to sustain.

These and other considerations led to the following extension of the preceding rank-deficient method, which instead of completely disregarding the smaller singular values in the rank-deficient solutions, makes some use of this neglected information. The primary motivation for this will be seen by examining the form of the solution given by Eqs. (13) and (14). We have

$$x^{(r)} = \sum_{i=1}^{r} p_i v_i \equiv \sum_{i=1}^{r} \tilde{x}_i \tag{15}$$

Thus the solution is computed as the sum of $r \leq n$ "eigen-solutions" $\tilde{x}_i$, linearly-independent vectors which span an r-dimensional subspace of X. The summation upper limit r in Eq. (15) is the pseudorank chosen by the previous algorithm to yield the desired change in $\|z(x)\|$. The contribution from the eigen solutions

$$p_i v_i = \tilde{x}_i \quad (i = r+1, r+2, \ldots, n) \tag{16}$$

is not included in the solution. It is the intent here to show that, in some cases, the eigensolutions (16) ought not to be discarded in their entirety, but should be appropriately scaled down and included in the solution. The existence of realistic a priori knowledge concerning the initial state estimate X_0 is such a case. The reason for this will become evident below.

From Eq. (14) we recall that errors in the data z are magnified in the computation of p_i, the scaling factor which determines the magnitude of $\tilde{x}_i$, as the associated singular value λ_i gets smaller. For present purposes, the effect of unmodeled nonlinearities can be considered to be a contribution to the noise in the data. Thus the magnitude of the undesired $\tilde{x}_i$'s may be in considerable error due to nonlinearities and other sources. On the other hand, the direction of $\tilde{x}_i$ is given by v_i which is computed from A only. That is, v_i is affected only indirectly by nonlinearities and data errors. Thus higher-rank $\tilde{x}_i$'s (those corresponding to smaller λ_i's) could be considered to be more accurate in direction and less accurate in magnitude as compared to lower-rank $\tilde{x}_i$'s.

The extended partial-step algorithm is therefore the following. Given a user-supplied a priori covariance matrix P reflecting a realistic uncertainty in the estimate of the initial state X_0, and a scalar bound parameter Q, compute $\tilde{x}_i = p_i v_i$ for $i = 1, 2, \ldots, n$. Next compute the weighted norms

$$\|\tilde{x}_i\|_P = (\tilde{x}_i^T P^{-1} \tilde{x}_i)^{\frac{1}{2}}$$

and compare them against Q. For each i, scale the magnitude of each vector $\tilde{x}_i$ so that the resulting vector $\tilde{x}_i'$ has P-norm not greater than Q. Thus set $\tilde{x}_i' = \alpha_i \tilde{x}_i$, where

$$\alpha_i = Q/\|\tilde{x}_i\|_P \quad \text{if } \|\tilde{x}_i\|_P > Q$$

$$\alpha_i = 1 \quad \text{if } \|\tilde{x}_i\|_P \leq Q$$

for $i = 1, 2, \ldots, n$. It is clear that each scaled eigen solution $\tilde{x}_i'$ is thereby constrained to lie on or within a Q-sigma probability ellipsoid in the parameter space X. For example, if P is a diagonal matrix (the usual case), then the applied constraint becomes

$$\|\tilde{x}_i'\|_P^2 = (\tilde{x}_{i_1}'/\sigma_1)^2 + (\tilde{x}_{i_2}'/\sigma_2)^2 + \cdots + (\tilde{x}_{i_n}'/\sigma_n)^2 \leq Q^2$$

for $i = 1, 2, \ldots, n$, where $\tilde{x}_{i_k}'$ is the kth component of $\tilde{x}_i'$ and σ_k is the assumed a priori standard deviation associated with the kth parameter being estimated. Finally, the extended partial-step solution is given by summing the constrained eigensolutions

$$x_s = \sum_{i=1}^{n} \tilde{x}_i' = \sum_{i=1}^{n} \alpha_i \tilde{x}_i \tag{17}$$

Equation (17) yields a constrained solution to the linear least-squares problem; the same scheme is used for each iteration of a nonlinear problem. In general, α_i's corresponding to smaller singular values will be smaller; however, as iteration progresses and the actual solution to the nonlinear problem is approached, the α_i's all approach 1. Note that Eq. (17) is equivalent to the normal equations solution when $\alpha_1 = \alpha_2 = \cdots \alpha_n = 1$. This constraining technique should be compared to the method used by Morrison[12] and Marquandt[13] for solving the same problem. Their scheme, while similar in spirit, does not make use of singular-value decomposition. Instead, they improve the conditioning of the linear system of equations by adding a positive diagonal matrix to the normal matrix before solving the normal equations. This is equivalent to the orbit determination technique of adding a priori data in the form of an a priori covariance matrix on the estimated parameters to the normal equations.

The success of this extended partial-step method obviously depends upon the choice of Q and upon how closely P corresponds to the actual error in the initial state estimate (and, of course, upon the magnitude of the nonlinearities present in the problem). The technique is currently being investigated with Q held constant for all iterations; an obvious extension of the algorithm is to appropriately shrink Q after each iteration, thereby cutting down the allowable step. This Q-reduction technique should be beneficial since smaller steps are needed as the true solution is approached. Such a scheme that works well for all cases has not yet been found. Another possible variation that has not yet been studied is the dynamic alteration of P from iteration to iteration guided by the solution from the last step.

It is felt that this extended partial-step algorithm has a definite application to the Mariner Mars orbiter mission where it should be possible to obtain a suitable value for P. A simulated Mars-orbiter study was conducted in an effort to determine the potential of this partial-step technique in regions of ill-conditionedness and full-rank divergence. This investigation is discussed below.

Some Results of the Extended Algorithm

A simplified orbit determination program using the extended algorithm (17) for obtaining solutions has been developed. The

physical model is identical to that in the program[14] used by Dyer, McReynolds, and Hanson in their study. The physical system is that of a spacecraft orbiting about Mars on a conic path with the movement of the Earth with respect to Mars being approximated by linear motion. All other planets are ignored (e.g., no Earth n-body effect), and light is assumed to have infinite velocity.

The program begins by simulating Doppler data spanning an interval of several hours using a set of nominal or true conditions—position and velocity at an epoch, usually the time of periapsis. This nominal state is then perturbed by a given amount to obtain a "guess" at the correct state. Starting with this guess as its initial conditions, the program attempts to converge to the correct state using the extended partial-step method. The simulated Doppler data is computed without any statistical noise; i.e., the data is computed deterministically from the range rate associated with the orbital trajectory. A separate investigation has shown that this simplification is inconsequential to the nonlinear convergence study: the precision associated with real Doppler observations makes data noise a secondary effect compared to the neglected nonlinear terms.

Initial results of the method have been good. For example, for the state error in $x, y, z, \dot{x}, \dot{y}, \dot{z}$ equal to

$$\Delta x = (10, -10, 10, 0.001, -0.001, -0.001)^T$$
(units are km, km/sec)

convergence (defined to have occurred when $\|z(x)\|$ gets smaller than some tolerance) occurred in three iterations for the $Q = 3$ case. This was the best choice of Q; for example, convergence occurred in five iterations for the $Q = 2$ case. This is to be contrasted with the nine z and nine A evaluations required for convergence using the Dyer rank-deficient scheme. More dramatic is the case

$$\Delta x = (500, -500, 500, -0.1, 0.1, 0.1)^T$$

where convergence occurred in seven iterations ($Q = 2$) as compared to 32 z and 17 A evaluations for convergence with the Dyer scheme. Other errors of the same magnitude (changing signs) show similar results. Some 1000 km position error cases have also converged with appropriate choices for Q (usually $Q = 1$). It is currently unresolved whether or not an optimum value or a range of optimum values exists for Q.

Another study was performed with the program to compare the relative convergence powers of the normal-equations (full-rank) method to the extended partial-step method. The results are tabulated in Table 1. In this investigation errors in the guess of the true state varied from 1 to 250 km (position) and from 0.1 to 100 m/sec (velocity). The important parameter determined in each case was the minimum time interval over which data, starting at the epoch, had to be taken to yield convergence. There were two epochs used; one was periapsis and the other was 50 min after periapsis. The latter epoch was used to approximately simulate the end-of-occultation time for the Mars-orbiter, at which point data will actually begin to be taken.

Table 1 gives the time intervals (minutes) over which convergence took place. For example, the entry under periapsis, full-rank, error = 1, 0.1, is 15–220. This means that the normal equations method converged to the true (known) solution at epoch using a data span which started at epoch (periapsis) and ended anywhere in the interval 15–220 min after epoch. Alternately, it means divergence occurred if less than 15 or greater than 220 min of data were used. An entry with no second number, such as 10–, means convergence occurred when any span of data greater than or equal to 10 min was used. The data intervals investigated extended over one complete orbit (720 min). Finally, the entry "no convergence" means that no amount of data starting from epoch would yield convergence for that particular error. The full-rank divergence cases were characterized by a "runaway" $z(x)$; i.e., the unbounded growth of $\|z(x)\|$ as iteration continued. The partial-step nonconvergence behavior was somewhat different; $\|z(x)\|$ would generally decrease during the first few iterations but would then essentially stop moving—long before convergence was attained.

Table 1 Convergence intervals

Error	Partial-step	Full-rank
Pos, km; Vel, m/sec	Periapsis	
1, 0.1	10–	15–220, min
10, 1	15–	No convergence
100, 25	25–	No convergence
250, 100	40–575	No convergence
	Periapsis + 50 min	
1, 0.1	12–	13–630
10, 1	20–	180–380
25, 10	26–	No convergence
100, 25	55–	No convergence
250, 100	150–640	No convergence

Conclusions

It is evident from Table 1 that, at least for the idealized problem studied, the extended partial-step algorithm increased the region of convergence (largest initial error which would allow convergence) dramatically beyond that attainable with the full-rank normal-equations method. Few absolute conclusions regarding the difficulty of normal equations orbit estimation during an actual mission of this type should be drawn from these results, however. For example, no a priori information was used in any of the full-rank solutions. The inclusion of a priori information on the estimated state plus data types other than Doppler, and the use of data from more than one tracking station, should enhance the convergence capability of any least-squares differential-correction technique. Nonetheless, significant nonlinear effects will be present during the Mariner Mars '71 orbiter phase, and the results summarized above indicate that the extended partial-step algorithm will be valuable during this phase of the mission.

References

[1] Taussky, O., "Note on the Condition of Matrices," *Mathematical Tables and Other Aids to Computation*, Vol. 4, 1950, pp. 111–112.

[2] Householder, A. S., "Unitary Triangularization of a Nonsymmetric Matrix," *Journal of the Association of Computing Machinery*, Vol. 5, 1958, pp. 339–342.

[3] Businger, P. and Golub, G. H., "Linear Least Squares Solutions by Householder Transformations," *Numerische Mathematik*, Vol. 7, 1965, pp. 269–276.

[4] Hanson, R. J. and Lawson, C. L., "Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems," *Mathematics of Computation*, Vol. 23, No. 108, Oct. 1969, pp. 787–812.

[5] Golub, G. H. and Wilkinson, J. H., "Note on the Iterative Refinement of Least Squares Solution," *Numerische Mathematik*, Vol. 9, 1966, pp. 139–148.

[6] Osborne, E. E., "On Least Squares Solutions of Linear Equations," *Journal of the Association of Computing Machinery*, Vol. 8, 1961, pp. 628–636.

[7] Golub, G. H. and Reinsch, C., "Singular Value Decomposition and Least Squares Solution," *Numerische Mathematik*, Vol. 14, 1970, pp. 403–420.

[8] Greville, T. N. E., "The Pseudo-Inverse of a Rectangular Matrix and Its Applications to the Solution of Systems of Linear Equations," *SIAM Review*, Vol. 1, Jan. 1959, pp. 38–43.

[9] Penrose, R., "A Generalized Inverse for Matrices," *Proceedings of the Cambridge Philosophical Society*, Vol. 51, 1955, pp. 406–413.

[10] Dyer, P., McReynolds, S., and Hanson, R., private communication, 1969, Jet Propulsion Lab., Pasadena, Calif.

[11] Warner, M. R. et al., "Double Precision Orbit Determination Program," 900-56, EPD 426, June 1967, Jet Propulsion Lab., Pasadena, Calif.

[12] Morrison, D. D., "Methods for Nonlinear Least Squares Problems and Convergence Proofs," *Proceedings of the Jet Propulsion Laboratory Seminar*, 1960, pp. 1–9.

[13] Marquandt, D. W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM Journal*, Vol. 11, No. 2, June 1963, pp. 431–441.

[14] Johnson, D. E., "Double Precision Approximated Orbit Determination Program," TR 69-1134-213-1, Sept. 1969, Informatics Inc., Sherman Oaks, Calif.