APPENDIX I-A

A UNIFIED APPROACH TO MINIMUM VARIANCE ESTIMATION WITH APPLICATIONS TO KALMAN FILTERING

BY

Mack W. Alford

I. INTRODUCTION

The purpose of this report is to show the relationships of least squares estimation, minimum variance estimation, Kalman filtering, and general sequential minimum variance estimation via the concept of the pseudo-inverse of a matrix, and inversion of matrices via partitioning and/or bordering. Although some familiarity with the above concepts is desirable, this report is meant to be self-contained.

Section II introduces the concept of a pseudo-inverse via the technique of preferred coordinate systems. Section III shows the relationship between minimum variance estimators and least squares estimators; i.e., they are the same in the preferred coordinate system.

Section IV shows that a priori estimates can be treated as more data points. Section V discusses sequential estimation, and shows that a set of data points may be reduced to an a priori estimate with an appropriate covariance matrix. Two computational algorithms are compared and contrasted. Section VI introduces the general concept of noise in the state for discrete measurements. Section VII derives the Kalman filter equations and shows these are just a special case of the foregoing. The problem of divergence of the Kalman filter is placed in its proper context as a model problem or a numeric problem of propagation of roundoff errors.

In Section VIII, a different error model is proposed which leads to a computational algorithm not prone to such roundoff problems. In Section IX, some applications to nonlinear equations are discussed and some approaches for the derivation of computational algorithms are given.

Section X gives a brief look at the subject of stochastic processes via the foregoing error-model pseudo-inverse approach.

It is probably fruitful to start out by discussing the philosophy of the error model approach used in this paper. The following comments may at first seem trivially true, but unless they are explicitly understood, great difficulties can ensue.

Whenever one is discussing the simulation of a process with incomplete or corrupted (noisy) measurements, one must talk in terms of an error model. To fix ideas, take the simple linear case. Suppose one has some variables which describe the state or condition of a system, denoted by a vector X; suppose further that a measurement vector Y is obtained from X by a linear transformation A, but that the measurements also contain additive noise ε. In this case, one models the process by writing the following equations: if there were no noise,

$$AX = Y \tag{1-1}$$

What is measured is $b = Y + \varepsilon$, and so this equation becomes

$$AX = b - \varepsilon \tag{1-2}$$

Notice that numerically A and b are provided. The crucial point is that there are more unknown variables than equations in (1-2). Thus, there is no solution in the strict sense of the word.

The usual next statement is the basis for the subject of statistics: "Although I am not able to tell you what the values of the ε_i are, I am willing to characterize them collectively; e.g., in some sense, their sum "should" equal zero, in some sense, any ε_i "should" not be too far away from zero." The point is that the ε_i are not <u>specified</u>, but they are <u>characterized</u>; i.e., even though the ε_i are not known exactly, you are still able to say <u>something</u> about them. Because you are willing to characterize the ε_i, then the X_i are not totally unknown; they are not able to be found exactly, but something can be said about them; e.g., the most probable value of X, how much the value of any X_i could vary about it, etc; i.e., the X_i may be characterized.

The philosophy of the error model approach is to characterize the errors, and then on the basis of this characterization, to deduce what characterizations of the X_i are consistent with those of the ε_i.

This is not the end of the process. To be complete, one should reverse the process and evaluate the model for consistency. Thus, if the ε_i are assumed to be sequentially independent zero mean random normal numbers, and the residuals (our estimates of the ε_i) are grossly serially correlated, the result of that model (our estimate of the X_i) is suspect because the assumptions by which they were derived are not satisfied.

To place this discussion on "earthy" terms, remember that whenever one works with error models, GIGO, or: garbage in, garbage out.

With this information firmly in mind, we next take up the subject of what is meant by linear transformation, and the inverse of a linear transformation.

II. PREFERRED COORDINATE SYSTEMS

The problem of linearly extracting a signal from a measurement, sometimes called "filtering", can be placed in the following abstract framework. Consider a linear transformation $\overline{A}$ which maps a finite dimensional vector space $\overline{X}$ (the parameter space) into the finite dimensional vector space $\overline{Y}$ (the measurement space). Of course, we assume that the measurements are corrupted with noise of some sort.

Assume for the moment that a basis $\{X_i\}$ is given for $\overline{X}$, and a basis $\{Y_i\}$ is given for $\overline{Y}$. Now relative to these two sets of basis vectors, the linear transformation $\overline{A}$ may be represented in the traditional manner by a matrix A, called the measurement matrix.

Further assume that A is diagonal. Then the measurement equations may be written:

$$AX = Y \tag{2-1}$$

or

$$\begin{pmatrix} \lambda_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \lambda_k & \\ & 0 & & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_k \\ X_{k+1} \\ \vdots \\ X_m \end{pmatrix} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \\ Y_{k+1} \\ \vdots \\ Y_n \end{pmatrix} \tag{2-2}$$

Notice that A is partitioned into three submatrices: (1) a k x k diagonal matrix; (2) a (m-k) x k matrix of zeros; and (3) a (n-k) x m matrix of zeros.

Everything of interest to the problem may be immediately read off of the matrix equations by sight. The following items are of interest:

(1) The vectors $X_1, \ldots, X_k$ may be varied to match $Y_1, \ldots, Y_k$. The matrix A is said to be of rank k. The subspace of $\underline{Y}$ spanned by $Y_1, \ldots, Y_k$ is called the Range of A.

(2) The vectors $X_{k+1}, \ldots, X_m$ are mapped into 0 by A. Thus, the subspace of $\underline{X}$ spanned by those vectors is called the Null-Space of A. In this context $X_{k+1}, \ldots, X_m$ are called unobservable.

(3) The error component of the measurement is $Y_{k+1}, \ldots, Y_n$. Note that this is the null space of A^T, where T denotes transpose, and that the range of A^T is $X_1, \ldots, X_k$.

(4) It is obvious that the matrix A has no inverse matrix. However, the matrix A^+, defined by

$$A^+ = \left(\begin{array}{ccc|c} 1/\lambda_1 & & 0 & \\ & \ddots & & \\ 0 & & 1/\lambda_k & 0 \\ \hline & 0 & & \end{array}\right) \qquad (2\text{-}3)$$

has the properties that AA^+ is the identity in the range of A, and A^+A is the identity in the range of A^T. Moreover,

$$X = A^+Y \qquad (2\text{-}4)$$

has the property that

$$Y - AX = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ Y_{k+1} \\ \vdots \\ Y_n \end{pmatrix} \qquad (2\text{-}5)$$

and

$$X = \begin{pmatrix} Y_1/\lambda_1 \\ \vdots \\ Y_k/\lambda_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad (2\text{-}6)$$

In words, it solves the equations where it is possible to do so, and sets the component of X which lies in the nullspace of A equal to zero. That is as good as can be done. For this reason, A^+ is called the Pseudo-inverse of A.

Another characterization of A^+ is as follows: Define the inner product in both $\underline{\bar{X}}$ and $\underline{\bar{Y}}$ as Euclidian; i.e.,

$$(X_1, X_2)_{\underline{\bar{X}}} = X_1^T X_2 = \sum_j X_1(j)\, X_2(j) \tag{2-7}$$

$$(Y_1, Y_2)_{\underline{\bar{Y}}} = Y_1^T Y_2 \tag{2-8}$$

Then A^+ has the property that

$$\hat{X} = A^+ Y \tag{2-9}$$

is the smallest least squares solution of (1), i.e., of all X such that

$$(AX - Y, AX - Y)_{\underline{\bar{Y}}} = \min \tag{2-10}$$

it is also true that

$$(\hat{X}, \hat{X})_{\underline{\bar{X}}} \leq (X, X)_{\underline{\bar{X}}} \tag{2-11}$$

The point of the above discussion is that any linear transformation may be represented by a diagonal matrix as above by choosing the correct coordinate system. This is a direct consequence of the Generalized Spectral Decomposition Theorem for finite dimensional vector spaces. This theorem states: Given any linear transformation $\bar{A}$ which maps $\underline{\bar{X}}$ into $\underline{\bar{Y}}$, there are orthonormal bases $\{X_i\}$ for $\underline{\bar{X}}$ and $\{Y_i\}$ for $\underline{\bar{Y}}$ such that $\bar{A}$ may be represented by:

$$A = \sum \lambda_i Y_i X_i^T \tag{2-12}$$

We call $\{X_i\}$ and $\{Y_i\}$ the preferred coordinate system for the linear transformation $\underline{A}$, because in this coordinate system, the solution is trivial.

At any rate, the pseudo-inverse of A^+ of A can be defined as

$$A^+ = \Sigma\, 1/\lambda_i\, X_i\, Y_i^T \tag{2-13}$$

This definition is consistent with the previous definition, is valid in any coordinate system, and corresponds to our intuitive notion of what a generalized inverse should do.

If the columns of A are independent, i.e., if A has an empty nullspace, then the pseudo-inverse may be given by:

$$A^+ = (A^T A)^{-1} A^T \tag{2-14}$$

when the inner product is Euclidian.

The concept of a generalized (non-Euclidian) inner product arises naturally in the following manner. Suppose the vector space $\underline{X}$ has two sets of basis vectors, $\{p_i\}$ and $\{q_i\}$. Any vector X can be expressed in terms of either basis; i.e.,

$$X = \Sigma\alpha_i p_i \tag{2-15}$$

or

$$X = \Sigma\beta_i q_i \tag{2-16}$$

From linear algebra, we know that there is a matrix, say P, such that

$$\beta = P\,\alpha \tag{2-17}$$

The concept of inner product has associated with it the concepts of angle and length. In those terms we have

$$(X_1, X_1) = ||X_1||^2 \tag{2-18}$$

where $||X_1||$ is the length of X_1, and

$$(X_1, X_2) = ||X_1|| \; ||X_2|| \cos\theta \tag{2-19}$$

where θ is the angle between X_1 and X_2. Suppose that the inner product is Euclidian in the $\{q_i\}$ coordinate system. One may define a generalized inner product in the $\{p_i\}$ coordinate system so that lengths and angles are preserved by noticing that

$$(X_1, X_2) = \beta_1^T \beta_2 = \alpha_1^T P^T P \alpha_2 \tag{2-20}$$

If the inner product is not Euclidian, in $\underline{\bar{X}}$ and $\underline{\bar{Y}}$, it still may be represented by a quadratic form. Set

$$(X_1, X_2)_{\underline{\bar{X}}} = X_1^T P X_2 \tag{2-21}$$

where P is a symmetric positive definite matrix. Also set

$$(Y_1, Y_2)_{\underline{\bar{Y}}} = Y_1^T W^{-1} Y_2 \tag{2-22}$$

where W is another (non-singular) symmetric positive definite matrix. Then,

$$A^+ = (A^*A)^{-1} A^* \tag{2-23}$$

where A* is defined by

$$(AX, Y)_{\underline{\bar{Y}}} = (X, A^*Y)_{\underline{\bar{X}}} \text{ for all } X \in \underline{\bar{X}} \text{ and } Y \in \underline{\bar{Y}}. \tag{2-24}$$

Equating the two definitions, we obtain

$$(X, A^*Y)_{\underline{\bar{X}}} = X^T P A^* Y \tag{2-25}$$

$$(AX, Y)_{\underline{\bar{Y}}} = X^T A^T W^{-1} Y \tag{2-26}$$

$$= X^T P P^{-1} A^T W^{-1} Y \tag{2-27}$$

Hence,

$$A^* = P^{-1} A^T W^{-1} \tag{2-28}$$

Now,

$$A^+ = (A^*A)^{-1}A^* \tag{2-29}$$

$$= (P^{-1} A^T W^{-1} A)^{-1} P^{-1} A^T W^{-1} \tag{2-30}$$

$$= (A^T W^{-1} A)^{-1} PP^{-1} A^T W^{-1} \tag{2-31}$$

$$= (A^T W^{-1} A)^{-1} A^T W^{-1} \tag{2-32}$$

This is the familiar minimum variance estimator. To say what that means, we need some notions from statistics.

III. SOME STATISTICS

Suppose $Y_{k+1}, \ldots, Y_n$ are identically independently normally distributed random numbers with zero mean. Then an estimate of the variance σ^2 of the population from which they were sampled is:

$$S = \frac{\sum_{i=k+1}^{n} Y_i^2}{n - k} \tag{3-1}$$

The covariance matrix P of a random vector Z is defined by

$$P = E(Z\, Z^T) \tag{3-2}$$

or

$$P_{ij} = E(Z_i\, Z_j) \tag{3-3}$$

where $E(\cdot)$ denotes the expected value of $(\cdot)$. Suppose

$$Z = QY \tag{3-4}$$

Then, by definition,

$$\text{Cov } Z = E(QY\ (QY)^T) \tag{3-5}$$

$$= E(QY\ Y^T\ Q^T) \tag{3-6}$$

If

$$E(Y\ Y^T) = \sigma^2\ I,$$

then

$$\text{Cov } Z = \sigma^2\ QQ^T = \sigma^2 W. \tag{3-7}$$

Note that the components of the vector Z are (depending on the matrix Q) dependently distributed random variables.

Note also that

$$\begin{aligned} Y^T Y &= (Q^{-1} Z)^T (Q^{-1} Z) \\ &= Z^T (Q^{-1})^T Q^{-1} Z \\ &= Z^T W^{-1} Z \end{aligned} \tag{3-8}$$

Turning these equations around, suppose we have

$$AX - b = Z \tag{3-9}$$

where Z is postulated to have covariance matrix W. The minimum variance estimate of X is that which minimizes

$$Z^T W^{-1} Z. \tag{3-10}$$

Hence, any minimum variance estimate is a least squares estimate in the preferred coordinate system. The matrix Q^{-1} exists because W^{-1} is a positive definite symmetric matrix by hypothesis.

Assume the columns of A are independent. The minimum variance solution $\hat{X}$ given by

$$\hat{X} = A^+(b + Z) \tag{3-11}$$

where

$$A^+ = (A^T W^{-1} A)^{-1} A^T W^{-1} \tag{3-12}$$

and Z is a zero mean random vector. The mean $\overline{X}$ of $\hat{X}$ is given by

$$\overline{X} = A^+ b \tag{3-13}$$

and, hence, the error vector of $\hat{X}$ is given by

$$\hat{X} - \overline{X} = A^+ Z \tag{3-14}$$

Now, by definition,

$$\text{cov } \hat{X} = E[(\hat{X} - \bar{X})(\hat{X} - \bar{X})^T] \quad (3\text{-}15)$$

$$= E[A^+ ZZ^T A^{+T}] \quad (3\text{-}16)$$

$$= A^+ E[ZZ^T] A^{+T} \quad (3\text{-}17)$$

$$= A^+ W A^{+T} \quad (3\text{-}18)$$

$$= [(A^T W^{-1} A)^{-1} A^T W^{-1}] W [W^{-1} A (A^T W^{-1} A)^{-1}] \quad (3\text{-}19)$$

$$\text{cov } \hat{X} = (A^T W^{-1} A)^{-1} (A^T W^{-1} A)(A^T W^{-1} A)^{-1} \quad (3\text{-}20)$$

$$= (A^T W^{-1} A)^{-1} \quad (3\text{-}21)$$

A second computational method for finding the solution X and its covariance matrix is to perform a Gram-Schmidt orthogonalization. This results in decomposing A into the product of two matrices,

$$A = UM \quad (3\text{-}22)$$

where the columns of U are orthonormal and M is triangular.

$$U^T W^{-1} U = I \quad (3\text{-}23)$$

$$\hat{X} = A^+ Z = M^{-1} U^+ Z \quad (3\text{-}24)$$

$$U^+ = (U^T W^{-1} U)^{-1} U^T W^{-1} = (I)^{-1} U^T W^{-1} = U^T W^{-1} \quad (3\text{-}25)$$

and

$$A^+ = M^{-1} U^T W^{-1} \quad (3\text{-}26)$$

Hence,

$$E(A^+ ZZ^T A^{+T}) = A^+ W A^{+T} \quad (3\text{-}27)$$

$$= M^{-1} U^T W^{-1} W W^{-1} U M^{-1T} \quad (3\text{-}28)$$

$$= M^{-1} M^{-1T} \quad (3\text{-}29)$$

Finding M^{-1} is easy since M is triangular.

IV. A PRIORI ESTIMATES

The problem of minimum variance estimation with a priori estimates is given as follows: If

$$AX - b = Z \tag{4-1}$$

Z is a zero-mean random vector with covariance matrix W, X_a is the a priori estimate vector with covariance matrix J, and it is desired to minimize the quantity

$$(X - X_a)^T J^{-1}(X - X_a) + Z^T W^{-1} Z \tag{4-2}$$

Define the new set of equations written in matrix form

$$\overline{A}X = \begin{pmatrix} I \\ A \end{pmatrix} X = \begin{pmatrix} X_a \\ b \end{pmatrix} = \overline{b} \tag{4-3}$$

Let the augmented covariance matrix $\overline{W}$ be defined by:

$$\overline{W} = \begin{pmatrix} J & 0 \\ 0 & W \end{pmatrix} \tag{4-4}$$

Then the X which minimizes

$$(\overline{A}X - \overline{b})^T \overline{W}^{-1} (\overline{A}X - \overline{b}) \tag{4-5}$$

solves the stated problem.

Hence: a priori estimates can simply be included as more equations in the general formulation.

Note: The form of the covariance matrix assumes the a priori estimates and the measurement errors are uncorrelated.

V. SEQUENTIAL ESTIMATION

The idea behind sequential estimation is to process data while still collecting it. For example, suppose measurements are made in time with the state changing linearly in time. We assume that

$$X(t_i) = \phi(t_i)X \tag{5-1}$$

where $\phi(t_i)$ is called a state transition matrix. Then the problem is to find X such that

$$\underset{\text{all } X}{\text{Minimum}} = (X - X_a)^T J^{-1}(X - X_a) + (AX - b)^T W^{-1}(AX - b) \tag{5-2}$$

where

$$A = \begin{pmatrix} M_1\phi(t_1) \\ M_2\phi(t_2) \\ \vdots \\ M_p\phi(t_p) \end{pmatrix} \tag{5-3}$$

Given the complete set of P measurements, we can find X which is the minimum variance estimate for the total problem. If we perform the same type of calculations after every set of measurements has been taken, we will obtain a sequence of estimates X_a, X_1, X_2, ..., X_n, and X_n will solve the total problem, and each X_k for $k \leq n$ will be the best estimate given all data up to this point. Note that a total fit is actually performed by the time the n^{th} measurement set is reached.

It is interesting to note however, that if residuals of the k^{th} data set are computed using the sequential estimate X_k for k = 1, ..., n, then the only set of residuals which are the true residuals from the total fit are the last set of residuals computed. Because of their nature, the residuals from a sequential fit must be interpreted differently from those of a total fit.

A simple example of the above is to fit a quadratic with a straight line. The residuals from a total fit will clearly show the quadratic term. The residuals from a sequential fit (unless recomputed in a total fit form) will not show this clear quadratic, but instead will show a gradual (quadratic) divergence. Interpretation of the residuals becomes much more difficult when a priori estimates of the coefficients are included as part of the data.

For the no-noise case, we may do the sequential estimation in the following manner. In either case, of course, the solution is obtained by setting

$$\bar{A} = \begin{pmatrix} I \\ M_1 \quad \phi(t_1) \\ M_2 \quad \phi(t_2) \\ \vdots \\ M_p \quad \phi(t_p) \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} X_a \\ Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix} \tag{5-4}$$

Then $\hat{X}$ is given by

$$\hat{X} = \bar{A}^{+}\bar{b} \tag{5-5}$$

where

$$\bar{A}^{+}\bar{b} = (\bar{A}^T W^{-1} \bar{A})^{-1} \bar{A}^T W^{-1} \bar{b} \tag{5-6}$$

It is clear that the columns are independent because of the I on top of $\bar{A}$.

If W^{-1} is a block diagonal with blocks J_o^{-1} and W_i^{-1}, these equations reduce to

$$\hat{X} = \bar{A}^{+}\bar{b} = [J_o^{-1} + \sum_{i=1}^{P} \phi^T(t_i)\, M_i^T\, W_i^{-1}\, M_i\, \phi(t_i)]^{-1}[J_o^{-1}\, X_a + \sum_{i=1}^{P} \phi^T(t_i)\, M_i^T\, W_i^{-1}\, Y_i] \tag{5-7}$$

The matrix corresponding to $(\bar{A}^T W^{-1} \bar{A})$ and the vector corresponding to $\bar{A}^T W^{-1} \bar{b}$ may be accumulated and at each time point a matrix inversion and a matrix multiplication will yield the desired sequential estimate.

We may get a different computational method by use of Ho's Lemma (see Appendix I).

Suppose $\hat{X}_k$ solves

$$A_k X_k = b_k \quad \text{with covariance } \bar{W}_k = \text{diag}(W_i),\ i \leq k \tag{5-8}$$

where

$$A_k = \begin{pmatrix} I & \\ M_1 & \phi(t_i) \\ \vdots & \\ M_k & \phi(t_k) \end{pmatrix}, \quad b_k = \begin{pmatrix} X_a \\ Y_1 \\ \vdots \\ Y_k \end{pmatrix} \tag{5-9}$$

Let

$$A_{k+1} = \begin{pmatrix} A_k & \\ M_{k+1} & \phi(t_{k+1}) \end{pmatrix}, \quad b_{k+1} = \begin{pmatrix} b_k \\ Y_{k+1} \end{pmatrix} \tag{5-10}$$

Now

$$A_{k+1}^T \bar{W}_{k+1}^{-1} A_{k+1} = A_k^T \bar{W}_k^{-1} A_k + \phi^T(t_{k+1}) M_{k+1}^T W_{k+1}^{-1} M_{k+1} \phi(t_{k+1}) \tag{5-11}$$

Set

$$J_k = (A_k^T \bar{W}_k^{-1} A_k)^{-1} \tag{5-12}$$

$$J_{k+1} = (A_{k+1}^T \bar{W}_{k+1}^{-1} A_{k+1})^{-1} \tag{5-13}$$

$$D = W_{k+1} + M_{k+1} \phi(t_{k+1}) J_k \phi^T(t_{k+1}) M_{k+1}^T$$

Thus, by use of Ho's Lemma,

$$J_{k+1} = J_k - J_k \phi^T(t_{k+1}) M_{k+1}^T D^{-1} M_{k+1} \phi(t_{k+1}) J_k \tag{5-14}$$

$$= [I - J_k \phi^T(t_{k+1}) M_{k+1}^T D^{-1} M_{k+1} \phi(t_{k+1})] J_k \tag{5-15}$$

Now

$$X_{k+1} = J_{k+1}(A_{k+1}^T \bar{W}_{k+1}^{-1} b_{k+1}) \quad (5\text{-}16)$$

$$= J_{k+1}(J_o^{-1} X_a + \sum_{i=1}^{k} M_i W_i^{-1} Y_i + M_{k+1} W_{k+1}^{-1} Y_{k+1}) \quad (5\text{-}17)$$

$$= (I - \ldots) J_k(J_o^{-1} X_a + \sum_{i=1}^{k} M_i W_i^{-1} Y_i) + J_{k+1} M_{k+1} W_{k+1}^{-1} Y_{k+1} \quad (5\text{-}18)$$

$$= (I - \ldots) \hat{X}_k + J_{k+1} M_{k+1} W_{k+1}^{-1} Y_{k+1} \quad (5\text{-}19)$$

If the size of M is smaller than the size of J, i.e., if there are fewer measured variables than state variables, then this algorithm necessitates only the inversion of a matrix the size of M rather than the size of J, and is hence faster.

Still, there is a price paid. The price is that any error made in either inversion or matrix multiplication in the first step is propagated through the rest of the sequential scheme. This is not the case if $A^T W^{-1} A$ is accumulated. If one wanted both a sequential estimate and a reliable total estimate of $\hat{X}$, it would be desirable to compute the sequential estimate by Method II and when all data was processed, use the accumulated $A^T W^{-1} A$ and Method I to find the total estimate. The difference is the error due to roundoff.

Incidentally, in the derivation of the above results, we also proved a rather interesting fact: <u>equations may be reduced to a priori estimates</u>. This is merely a consequence of the fact that the pseudo-inverse of a matrix can be computed by partitioning. This becomes rather an important fact when there is noise in the state.

The fact that equations may be reduced to a priori estimates may be proven more succinctly in the following way. By matrix manipulations, we obtain:

$$AX - b = AX - b + AA^+ b - AA^+ b \quad (5\text{-}20)$$

$$= (AA^+ b - b) + (AX - AA^+ b) \quad (5\text{-}21)$$

$$(AX - b)^T(AX - b) = (AA^+b - b)^T(AA^+b - b)$$
$$+ 2(AA^+b - b)^T(AX - AA^+b) \qquad (5\text{-}22)$$
$$+ (AX - AA^+b)^T(AX - AA^+b)$$

Set $\hat{X} = A^+b$. Then, these equations become

$$(AX - b)^T(AX - b) = (A\hat{X} - b)^T(A\hat{X} - b)$$
$$+ 2(A\hat{X} - b)^T(AX - A\hat{X}) \qquad (5\text{-}24)$$
$$+ (X - \hat{X})^T(A^TA)(X - \hat{X})$$

In this form, it is directly perceivable that the X which minimized $(AX - b)^T(AX - b)$ is simply $\hat{X}$; the proof is that it zero's every term which contains X at all. The last term in the last equation may be interpreted as follows: $\hat{X}$ is an estimate of X with zero mean and associated covariance matrix $(A^TA)^{-1}$.

Thus, if more observations were taken, and the errors were independent of the preceding errors, all of the information necessary for finding an updated estimate of X could be condensed into the value $\hat{X}$, taken as an a priori estimate with covariance matrix $(A^TA)^{-1}$.

VI. NOISE IN THE STATE - DISCRETE CASE

Frequently, a process in time cannot be modeled as fully deterministic, so a model involving "noise in the state" is resorted to. This says that instead of $X(t_i)$ being a deterministic function of $X(t_o)$, i.e.,

$$X(t_i) = \phi(t_i)\, X(t_o) \tag{6-1}$$

rather the state at the next time point is assumed to be almost linearly predictable from all past states; i.e., the quantity

$$X(t_i) - \sum_j k_{ij}\, \phi(t_i, t_j)\, X(t_j) \tag{6-2}$$

is taken to be a random vector with zero mean with an associated total covariance matrix R (it is not required that the errors at the $i^{\underline{th}}$ time point be independent of any of the rest). Quite simply, this has the effect of increasing the total number of variables being sought from, say m, to nm, where n is the number of time points at which measurements are taken. For the reason that the solution will be correspondingly less well-determined, it is advisable to avoid this if at all possible.

Note: The larger the covariance matrix R, the less any sort of predictability is demanded. An infinite R for any variable corresponds to a problem with no predictability assumed for that variable.

At any rate, any noise in the state problem may now be posed in matrix form in the following manner: find the minimum variance estimate of X which satisfies

$$AX = b$$

where W is an arbitrary covariance matrix, where A, b, and X may be written as

$$\begin{pmatrix} I & & & O \\ M_1 & & & \\ (-k_{21}\phi_{21}) & I & & \\ 0 & M_2 & & \\ (-k_{31}\phi_{31}) & (-k_{32}\phi_{32}) & M_3 & \\ 0 & 0 & \ddots & \\ \vdots & \vdots & \ddots & \\ (-k_{n1}\phi_{n1}) & (-k_{n2}\phi_{n2}) & \cdots\ (-k_{n,n-1}\phi_n) & I \\ 0 & 0 & 0 \quad 0 & M_n \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} X_a \\ Y_1 \\ 0 \\ Y_2 \\ 0 \\ Y_3 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \tag{6-3}$$

If preferred, these equations may be rewritten in terms of the noise variables ε explicitly. Set

$$\varepsilon_i = X(t_i) - \sum_j K_{ij}\phi(t_i, t_j)\, X(t_j) \tag{6-4}$$

or

$$X(t_i) = \sum_j K_{ij}\phi(t_i, t_j)\, X(t_j) + \varepsilon_i \tag{6-5}$$

$$= K_{io}\phi(t_i, t_o)\, X(t_o) + \sum_{j=1}^{i} \bar{K}_{ij}\phi(t_i, t_j)\varepsilon_j \tag{6-6}$$

with the $\bar{K}_{ij}$ found as functions of the original K_{ij} by matrix manipulations. In this coordinate system Equations (6-3) now become

$$\begin{pmatrix} I & & & & \\ M_1 & & & & \\ 0 & & & & \\ M_2\bar{K}_{21}\phi_{21} & M_2 & & & \\ 0 & 0 & I & \cdots & 0 \\ M_3\bar{K}_{31}\phi_{31} & M_3\bar{K}_{32}\phi_{32} & M_3 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & I \\ M_n\bar{K}_{n1}\phi_{n1} & M_n\bar{K}_{n2}\phi_{n2} & M_n\bar{K}_{n3}\phi_{n3} & \cdots & M_n \end{pmatrix} \begin{pmatrix} X_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} X_a \\ Y_1 \\ 0 \\ Y_2 \\ 0 \\ Y_3 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \tag{6-7}$$

Great computational simplifications can now be made by demanding that the ε_i be independent random vectors and the errors added into the Y_i be independent random vectors; i.e., the covariance matrix W be block diagonal.

VII. KALMAN FILTERING

An examination of the error model and equations for Kalman filtering will now reveal that it is a special case of the above techniques. First we assume that the X_i are observable; i.e., the null space of $\bar{A}$ is empty, or the columns of $\bar{A}$ are independent. This is assumed by requiring a priori estimates of all state variables; i.e., $X = X_a$ with covariance matrix J_o. The error model for noise in the state is given by stipulating that the quantities

$$X_i - \phi(t_i, t_{i-1})\, X_{i-1} = X_i - \phi_{i-1}\, X_{i-1} \tag{7-1}$$

are independently distributed random vectors for each i with covariance matrix R_{i-1}.

At time t_i, we assume that

$$\varepsilon_i = MX_i - Y_i \tag{7-2}$$

is a random vector with mean 0 and covariance matrix W_i, and the measurement errors, a priori errors, and state noise errors are all independent of each other. Since the covariance matrix is block diagonal, substitution and simplification of Equation 3-10 shows that the minimum variance estimate of X is now that which minimizes

$$(X_1 - X_a)^T J_o^{-1} (X_1 - X_a) + \sum_i (MX_i - Y_i)^T W_i^{-1} (MX_i - Y_i) \tag{7-3}$$

$$+ \sum_i (X_{i+1} - \phi(t_{i+1}, t_i)X_i)^T R_i^{-1} (X_{i+1} - \phi(t_{i+1}, t_i)X_i)$$

The entire set of equations may now be written in the previous notation as

$$\bar{A}X = \bar{b} \tag{7-4}$$

with covariance matrix $\bar{W}$, assumed to be block diagonal. Written out in all its gory detail, we have :

$$\begin{pmatrix} I & & & & & \\ M_1 & & & & & \\ -\phi_1 & I & & & & \\ 0 & M_2 & & & & \\ 0 & -\phi_2 & I & & & \\ 0 & 0 & M_3 & \ddots & & \\ \vdots & & & \ddots & -\phi_{n-1} & I \\ 0 & & & & 0 & M_n \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} X_a \\ Y_1 \\ 0 \\ Y_2 \\ 0 \\ Y_3 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \tag{7-5}$$

with associated covariance matrix

$$\bar{W} = \begin{pmatrix} J_o & & & & 0 \\ & W_1 & & & \\ & & R_1 & & \\ & & & \ddots & \\ & & & & R_{n-1} \\ 0 & & & & W_n \end{pmatrix} \tag{7-6}$$

Assume further that we would like to do a sequential estimation; i.e., after every set of measurements are taken, an updated estimate of the state is required. We wish to derive a computational algorithm which will yield the desired results.

A glance at the fundamental matrix $\bar{A}$ shows that the inner product of the columns which are multiplied by X_i and those which are multiplied by X_j are zero if $|i - j| \geq 2$. This says that $\bar{A}^T \bar{W}^{-1} \bar{A}$ will be block-tri-diagonal. In fact, a little messy algebra shows that

$$\bar{A}^T\bar{W}^{-1}\bar{A} = \begin{pmatrix} (J_o^{-1} + M_1^T W_1^{-1} M_1 + \phi_1^T R_1^{-1}\phi_1) & (-\phi_1^T R_1^{-1}) & \cdots & 0 \\ (-R_1^{-1}\phi_1) & (R_1^{-1} + M_2^T W_2^{-1} M_2 + \phi_2^T R_2^{-1}\phi_2) & \cdots & 0 \\ 0 & (-R_2^{-1}\phi_2) & \cdots & 0 \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & R_{n-1}^{-1} + M_n^T W_n^{-1} M_n \end{pmatrix} \tag{7-7}$$

and

$$\bar{A}^T\bar{W}^{-1}\bar{b} = \begin{pmatrix} J_o^{-1} X_a \\ M_1^T W_1^{-1} Y_1 \\ M_2^T W_2^{-1} Y_2 \\ \vdots \\ M_n^T W_n^{-1} Y_n \end{pmatrix} \tag{7-8}$$

If it were required that the best estimate of the state be found at each time point, given <u>all</u> data, and its associated covariance matrix, then the entire matrix $\bar{A}^T\bar{W}^{-1}\bar{A}$ would have to be inverted. If the number of measurements were large, this could pose problems in computer storage. If we wish only an estimate of X_n, the final state, then much less is demanded. This may be seen by consideration of the fact that since $\bar{A}^T\bar{W}^{-1}\bar{A}$ is block-tri-diagonal, it lends itself easily to a sequential inversion by repeated use of the first matrix identity and Ho's Lemma, appearing in Appendix I.

The algebra gets messy, so this is left as an exercise for the beginning student.

A somewhat different derivation, more in keeping with Kalman's approach, is to use the method of sequential orthogonalization of the blocks of columns of $\bar{A}$. This method is derived in Appendix II.

It may be applied as follows: By the Section V discussion, we may reduce all previous data to our estimate of X_n, and its associated covariance matrix J_n. Let $X_{i/j}$ mean the estimate of X_i given all measurements up to the jth time point, since our best estimate of X_{n+1} is given by

$$X_{n+1/n} = \phi_n X_{n/n} \tag{7-9}$$

By Equation (3-6), the associated covariance matrix is

$$J_{n+1/n} = \phi_n J_{n/n} \phi_n^T \tag{7-10}$$

The equations reduce to

$$\begin{pmatrix} I & 0 \\ -I & I \\ 0 & M_{n+1} \end{pmatrix} \begin{pmatrix} \phi_n X_n \\ X_{n+1} \end{pmatrix} = \begin{pmatrix} X_a \\ 0 \\ Y_{n+1} \end{pmatrix} \quad \begin{pmatrix} \text{Cov} \\ J_{n+1/n} \\ R_n \\ W_{n+1} \end{pmatrix} \tag{7-11}$$

By Appendix II, set

$$A_1 = \begin{pmatrix} I \\ -I \\ 0 \end{pmatrix} \tag{7-12}$$

$$A_1^+ = (A_1^T W_1^{-1} A_1)^{-1} A_1^T W_1^{-1} \tag{7-13}$$

$$= (J_{n+1/n}^{-1} + R_n^{-1})^{-1} (J_{n+1/n}^{-1}, -R_n^1, 0)$$

$$X_{n+1/n+1} = [(I - A_1 A_1^+) A_2]^+ Y_{n+1} = ((J_{n+1/n} + R_n)^{-1} + M_{n+1}^T W_{n+1}^{-1} M_{n+1})^{-1}$$

$$(J_{n+1/n}^{-1} X_a + M_{n+1}^T W_{n+1}^{-1} Y_{n+1})$$

It is well known that a process of sequential orthogonalization is prone to a build-up of round-off errors. The reason is quite simple: any error made at any step propagates into all future steps. Thus, one would expect that successive sequential estimates would get more and more inaccurate. Note that this assumes a perfect model, and is due purely to accumulation of round-off errors.

This type of divergence should not be confused with the divergence due to model errors (e.g., fitting a quadratic with a straight line will lead to endpoint divergence), although the residuals may be indistinguishable.

Experiments with Kalman filtering have shown divergence problems do exist, and, in fact, of the type indicated above. Note that although a "better fit" of the measurements will result from an increase of the state noise covariance matrix R, this solves a different problem than that originally given. A physical interpretation of this larger R is that less predictability is demanded of the state variables. If R is very very large (infinite), then the state variables are adjusted <u>independently</u> at each time point to fit that set of measurement data.

A final difficulty inherent in the Kalman filter is that: (a) a final estimate of the state variables at any time point is difficult to find, and must be specifically provided for by subsequent manipulations. These equations may be derived from repeated iteration of the techniques of Appendix II; (b) because these final estimates are so difficult to find, the residuals at each time point of the measurement variables are also difficult to compute. These residuals are of course necessary for the complete <u>evaluation</u> of the model. Because of the inherent difficulties, such evaluation is almost never done in practice.

VIII. A DIFFERENT ERROR MODEL

As was stated in the last section, Kalman filtering is a special case of the general model of Section VI. The Kalman equations are simple because of the special structure of the Kalman error model. This leads one quite naturally to ask: is there some other special structure which also leads to a set of equations which may also be solved simply?

The answer to this question is yes. The task of this section will be to find such a model, derive a method of solution, and to explore its ramifications.

The basic general error model Equation (6-2) states that

$$\varepsilon_j = X(t_i) - \sum_j k_{ij}\, \phi(t_i, t_j)\, X(t_j) \qquad (8\text{-}1)$$

is a random vector. The Kalman filter problem is that where

$$k_{ij} = \begin{cases} 0 & j \neq i\text{-}1 \\ 1 & j = i\text{-}1 \end{cases} \qquad (8\text{-}2)$$

This is the so-called Markov Hypothesis, i.e., that even if one knows all the past history of $X(t)$, all of the information about $X(t_i)$ is contained in the estimate of $X(t_{i-1})$.

A different error model, which is explicitly non-Markovian, is to say that the variable

$$\varepsilon_i = X(t_i) - \phi(t_i, t_o)\, X(t_o) \qquad (8\text{-}3)$$

is a random vector, and that ε_i and ε_j are independent of one another if $i \neq j$. One way to interpret this model is as follows: In the Kalman model we have

$$\varepsilon_1 = X(t_1) - \phi(t_1, t_o)\, X(t_o) \qquad (8\text{-}4)$$

or

$$X(t_1) = \phi(t_1, t_o)\, X(t_o) + \varepsilon_1 \qquad (8\text{-}5)$$

Now, by iteration, we have

$$X(t_2) = \phi(t_2, t_1)\, X(t_1) + \varepsilon_2$$

$$= \phi(t_2, t_1)\, [\phi(t_1, t_o)\, X(t_o) + \varepsilon_1] + \varepsilon_2$$

$$= \phi(t_2, t_o)\, X(t_o) + \phi(t_2, t_1)\varepsilon_1 + \varepsilon_2 \qquad (8\text{-}6)$$

and

$$X(t_3) = \phi(t_3, t_o)\, X(t_o) + \phi(t_3, t_1)\, \varepsilon_1 + \phi(t_3, t_2)\varepsilon_2 + \varepsilon_3 \qquad (8\text{-}7)$$

Notice that the error ε, propogates through to all successive time points. This may be thought of as Noise in the Derivative.

The new model, however, has the following form:

$$X(t_i) = \phi(t_i, t_o)\, X(t_o) + \varepsilon_i = \phi_i\, X_o + \varepsilon_i \qquad (8\text{-}8)$$

with ε_i and ε_j independent for $i \neq j$. Thus, given $X(t_o)$, $X(t_i)$ may be predicted except for additional noise. This may be called Noise in the State. With this error model, Equations 6-3 may be rewritten

$$\begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ -\phi(t_1,t_o) & I & 0 & \cdots & 0 \\ 0 & M_1 & 0 & \cdots & 0 \\ -\phi(t_2,t_o) & 0 & I & \cdots & 0 \\ 0 & 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\phi(t_n,t_o) & 0 & 0 & \cdots & I \\ 0 & 0 & 0 & \cdots & M_n \end{pmatrix} \begin{pmatrix} X_o \\ X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} X_a \\ 0 \\ Y_1 \\ 0 \\ Y_2 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \qquad (8\text{-}9)$$

with the covariance matrix

$$W = \begin{bmatrix} J_o & & & & & 0 \\ & R_1 & & & & \\ & & W_1 & & & \\ & & & R_2 & & \\ & & & & W_2 & \\ & & & & & \ddots \\ & & & & & & R_n \\ 0 & & & & & & & W_n \end{bmatrix} \tag{8-10}$$

Because of computational considerations, it is desirable to perform a change of variables and write Equations 8-9 in terms of X_o and ε_i instead of just the X_i. Then we have $AX = b$ rewritten as:

$$\begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ M_1\emptyset_1 & M_1 & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ M_2\emptyset_2 & 0 & M_2 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & I \\ M_n\emptyset_n & 0 & 0 & \cdots & M_n \end{bmatrix} \begin{pmatrix} X_o \\ \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} X_a \\ 0 \\ Y_1 \\ 0 \\ Y_2 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \tag{8-11}$$

At this point, we may generalize things a bit further by allowing the ε_i to have a smaller dimension than X_o, i.e., not every state variable must have noise. In this case, the dimensions of ε_i must be adjusted to the correct size by a premultiplication by a matrix Ψ_i. Thus, Equations 8-11 become:

$$\begin{pmatrix} I & 0 & 0 & & \\ 0 & I & 0 & & \\ M_1\emptyset_1 & M_1\Psi_1 & 0 & & \\ 0 & 0 & I & & \\ M_2\emptyset_2 & 0 & M_2\Psi_2 & & \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \cdots & I \\ M_n\emptyset_n & 0 & 0 & \cdots & M_n\Psi_n \end{pmatrix} \begin{pmatrix} X_o \\ \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} X_a \\ 0 \\ Y_1 \\ 0 \\ Y_2 \\ \vdots \\ 0 \\ Y_n \end{pmatrix} \tag{8-12}$$

Now a few computations are in order.

$$A^T W^{-1} A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{8-13}$$

with

$$a_{11} = J_o^{-1} + \sum_i \emptyset_i^T M_i^T W_i^{-1} M_i \emptyset_i \tag{8-14}$$

$$a_{12} = (\emptyset_1^T M_1^T W_1^{-1} M_1 \Psi_1, \ \ldots, \ \emptyset_n^T M_n^T W_n^{-1} M_n \Psi_n) \tag{8-15}$$

$$= (k_1, k_2, \ldots, k_n)$$

$$a_{21} = a_{12}^T \tag{8-16}$$

$$a_{22} = \text{diag} \ (\Psi_i^T M_i^T W_i^{-1} M_i \Psi_i + R_i^{-1}) \tag{8-17}$$

$$= \text{diag} \ (d_i)$$

$$A^T W^{-1} b = \begin{matrix} J_o^{-1} X_a + \sum_i \phi_i^T M_i^T W_i^{-1} Y_i \\ \psi_1^T M_1^T W_1^{-1} Y_1 \\ \vdots \\ \psi_n^T M_n^T W_n^{-1} Y_n \end{matrix} \qquad (8\text{-}18)$$

Now we can solve the fundamental equation by

$$X = (A^T W^{-1} A)^{-1} A^T W^{-1} b \qquad (8\text{-}19)$$

By use of Appendix I, we invert $A^T W^{-1} A$ by partitioning. The inverse of a_{22} is diagonal. Depending on taste,

$$d_i^{-1} = (\psi_i^T M_i^T W_i^{-1} M_i \psi_i + R_i^{-1})^{-1} \qquad (8\text{-}20)$$

may be found directly or by use of Ho's Lemma. At any rate, from Appendix I

$$(A^T W^{-1} A)^{-1} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \qquad (8\text{-}21)$$

with $b_{11} = (a_{11} - a_{12} a_{22}^{-1} a_{21})^{-1}$ (8-22)

or

$$b_{11} = [J_o^{-1} + \sum_i \phi_i^T M_i^T W_i^{-1} M_i \phi_i - \sum_i \phi_i^T M_i^T W_i^{-1} M_i \psi_i \; d_i^{-1} \; \psi_i^T M_i^T W_i^{-1} M_i \phi_i]^{-1} \qquad (8\text{-}23)$$

or

$$b_{11} = [J_o^{-1} + \sum_i \phi_i^T M_i^T W_i^{-1} M_i \phi_i - \sum_i k_i \; d_i^{-1} \; k_i^T]^{-1} \qquad (8\text{-}24)$$

$$b_{12} = - b_{11}\, a_{12}\, a_{22}^{-1}$$

$$= (-b_{11}\, k_1\, d_1^{-1}, \ldots, -b_{11}\, k_n\, d_n^{-1}) \qquad (8\text{-}25)$$

$$b_{21} = - a_{22}^{-1}\, a_{21}\, b_{11}$$

$$= \begin{matrix} - d_1^{-1}\ k_1^T \\ \vdots \\ - d_n^{-1}\ k_n^T \end{matrix} b_{11} \qquad (8\text{-}26)$$

$$b_{22} = a_{22}^{-1} + a_{22}^{-1}\, a_{21}\, b_{11}\, a_{12}\, a_{22}^{-1}$$

$$= \begin{pmatrix} d_1^{-1} & & 0 \\ & \ddots & \\ 0 & & d_n^{-1} \end{pmatrix} + \begin{pmatrix} d_1^{-1} k_1^T b_{11} k_1 d_1^{-1}, \ldots, d_1^{-1} k_1^T b_{11} k_n d_n^{-1} \\ \vdots \\ d_n^{-1} k_n^T b_{11} k_1 d_1^{-1}, \ldots, d_n^{-1} k_n^T b_{11} k_n d_n^{-1} \end{pmatrix}$$

(8-27)

Now, by substitution into Equation 8-19, we have

$$X_o = b_{11}(J_o^{-1} X_a + \sum_i \phi_i^T M_i^T W_i^{-1} Y_i)$$

$$- b_{11}[\sum_j k_j\, d_j^{-1}\ \psi_j^T M_j^T W_j^{-1} Y_j] \qquad (8\text{-}28)$$

and

$$\varepsilon_i = - d_i^{-1} k_i^T b_{11}(J_o^{-1} X_a + \sum_i \phi_i^T M_i^T W_i^{-1} Y_i)$$

$$+ d_i^{-1} \psi_i^T M_i^T W_i^{-1} Y_i \qquad (8\text{-}29)$$

$$+ d_i^{-1} k_i^T b_{11}[\sum_j k_j\, d_j^{-1}\ \psi_j^T M_j^T W_j^{-1} Y_j]$$

or

$$\varepsilon_i = - d_i^{-1} k_i^T X_o + d_i^{-1} \psi_i^T M_i^T W_i^{-1} Y_i \qquad (8\text{-}30)$$

Having derived the relevant equations, it will now pay to try to digest the results. In the first place, the form of the equations allows a complete solution of the entire system of equations in two passes. The first pass is to accumulate the matrix sums appearing in Equations 8-23, 8-28, etc. At the end of the first pass, the estimate of X_o may be found immediately from Equations 8-23. If sequential estimates of X_o are desired, they may also be obtained.

Now the data may be reprocessed in a second pass, and each $\hat{\varepsilon}_i$ may be found along with its estimated variance. Further, the residuals

$$r_i = Y_i - M_i' \Phi_i X_o - M_i' \Phi_i \hat{\varepsilon}_i$$

may be computed at each time point. The $\hat{\varepsilon}_i$ may be plotted (approximately scaled by $R^{1/2}$) and also the residuals r_i may be plotted, and hence the assumption of independence and proper scaling may be evaluated. This latter must done in order to properly vindicate the model and its applicability to the problem actually solved by the method.

An alternative method of viewing this model is from the perspective of separation of errors. An examination of the equations of Section V and those of this section shows that the answers for X_o are the same except for two additional terms, one which corrects the b_{11} term and another which corrects for the ε_i terms. Thus, the total error

$$Y_i - M_i' \Phi_i X_o$$

is separated into two parts; one set of independent random variables whose covariance is a function of the measurement matrix M', another set of independent random variables whose covariance is independent of M'.

Notice that this algorithm, unlike the Kalman filter, is computationally stable; i.e., the more data, the better the results.

IX. APPLICATIONS TO NONLINEAR EQUATIONS

Nonlinear Minimum Variance Estimation (NMVE) is derived in the same manner as is MVE, and from the same motivation.

No Noise In State

The model here is

$$Y = f(X) \tag{9-1}$$

$$Y^* = Y + \varepsilon \tag{9-2}$$

$$\text{cov}(\varepsilon) = W \tag{9-3}$$

where X is the "state," and Y* are the measurements. In this case, the quantity

$$(Y^* - f(X))^T W^{-1}(Y^* - f(X)) \tag{9-4}$$

is to be minimized by varying X.

Because the equations are nonlinear, the equations cannot be solved analytically in one fell swoop. The technique of linearization and iteration is used for the solution of the problem. Equations (9-1) and (9-4) are replaced by:

$$Y = f(\bar{X}) + \left(\frac{\partial f}{\partial X}\right)_{\bar{X}} \delta X \tag{9-5}$$

and

$$\underset{\delta X}{\text{Minimum}} = (Y^* - f(\bar{X}) - \left(\frac{\partial f}{\partial X}\right)_{\bar{X}} \delta X)^T W^{-1}(Y^* - f(\bar{X}) - \left(\frac{\delta f}{\delta X}\right)_{\bar{X}} \delta X) \tag{9-6}$$

If $\delta X = 0$, then $\bar{X}$ is a local solution of (9-4); if $\delta X \neq 0$, we set $\bar{X} = \bar{X} + \delta X$, and iterate until $\delta X = 0$.

If a priori estimates of the state are included as equations, appropriate adjustments must be made.

$$X = X_a \qquad \text{Covariance } J_o \tag{9-7}$$

is transposed into

$$\delta X = X_a - \bar{X} \qquad \text{Covariance } J_o \tag{9-8}$$

This is true whether $\bar{X} = X_a$ as it will for the first iteration, or not, as is the case after the first iteration.

Noise in the State

The noise in the state case is approximated in one of two ways, called Linear Models I and II. Linear Model 0, the no noise case, is also an approximation assuming the noise to be essentially zero.

Linear Model I assumes the noise in the state is sufficiently small that the equation

$$g(X^*) = g(X + \varepsilon) \tilde{\sim} g(\bar{X}) + \left(\frac{\partial g}{\partial X}\right) \delta X + \left(\frac{\partial g}{\partial X}\right)\varepsilon \tag{9-9}$$

is approximately valid. This means that the ε_i are re-solved for at each iteration.

Linear Model II assumes the noise in the state is a little bit larger, but not too much larger, so that the equation

$$\begin{aligned} g(X^*) &= g(X + \varepsilon) \\ &= g(\bar{X} + \bar{\varepsilon}) + \left(\frac{\partial g}{\partial X}\right)\delta X + \left(\frac{\partial g}{\partial X}\right)\delta\varepsilon \end{aligned} \tag{9-10}$$

is valid. Notice that in Linear Model I, only the state $\bar{X}(t_i)$ need be saved, while in Linear Model II both $\bar{X}(t_i)$ and $\bar{\varepsilon}(t_i)$ must be saved at each time point.

The solution of the linear equations which result from these linearizations is covered in the preceding pages.

In particular, if the uncorrelated noise in the state model is considered, the following is derived. The model equations are:

$$\left.\begin{aligned} X &= f(X_o) \\ X^* &= X + \varepsilon \\ Y &= g(X^*) \\ Y^* &= Y + \eta \end{aligned}\right\} \qquad (9\text{-}11)$$

where X_o is the initial state, ε is the state noise, X is the predicted state, X* is the "noisy state," Y is the computed measurement, η is the measurement noise, and Y* is the measurement data; the state and measurement noise are characterized by

$$E(\varepsilon_i \varepsilon_j^T) = \left\{\begin{array}{ll} 0 & i \neq j \\ R_i & i = j \end{array}\right\}$$

$$E(\varepsilon_i \eta_j) = 0 \qquad (9\text{-}12)$$

$$E(\eta_i \eta_j) = \left\{\begin{array}{ll} 0 & i \neq j \\ W_i & i = j \end{array}\right\}$$

The quantity

$$(X_a - X_o)^T J_o^{-1}(X_a - X_o) + [Y^* - g(X^*)]^T W^{-1}[Y^* - g(X^*)] + \varepsilon^T R \varepsilon \qquad (9\text{-}13)$$

is to be minimized with respect to X_o and ε.

Linear Model I makes the approximation

$$\begin{aligned} g(X^*) &= g(X + \varepsilon) \\ &= g(f(X_o) + \varepsilon) \\ &= g(f(X_o^n)) + \left(\frac{\partial g}{\partial X^*}\right)\varepsilon + \left(\frac{\partial g}{\partial X^*}\right)\left(\frac{\partial f}{\partial X_o}\right)\delta X \end{aligned} \qquad (9\text{-}14)$$