AD-A049 704 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO F/G 9/3
SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM: SQUARE ROOT --ETC(U)
SEP 77 P S MAYBECK

UNCLASSIFIED AFIT-TR-77-6 NL

| OF |
AD
A049704

END
DATE
FILMED
3-78
DDC

(1)

SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM: SQUARE ROOT AND U-D COVARIANCE FACTORIZATIONS

AFIT-TR-77-6

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson Air Force Base, Ohio

(See 1473)

AFIT-TR-77-6

SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM: SQUARE ROOT AND U-D COVARIANCE FACTORIZATIONS

Peter S. Maybeck
Assoc. Prof. of Electrical Engineering

September 1977

AIR FORCE INSTITUTE OF TECHNOLOGY
Wright-Patterson AFB OH 45433

SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM:
SQUARE ROOT AND U-D COVARIANCE FACTORIZATIONS

ABSTRACT

This report presents the concept of square root filters and the closely related U-D covariance factorization filter as viable alternatives to conventional Kalman filters. For a modest increase in computational loading, one obtains optimal filter algorithms equivalent to the Kalman filter if infinite wordlength is assumed, but with vastly superior numerical characteristics with finite wordlength. These algorithms are at least as good a solution to troublesome measurement update computations as implementing a Kalman filter in double precision, since the Kalman filter inherently involves unstable numerics.

The filter algorithms are developed and presented in a form convenient for implementation. Of the covariance square root forms, the Carlson filter is more efficient than the Potter form computationally, and it also maintains triangularity of the square root matrices. The U-D covariance factorization filter is even more efficient, not requiring square root computations. In comparison, the inverse covariance square root filter is often considerably more burdensome, although it too becomes competitive if the measurement dimension is very large.

SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM: SQUARE ROOT AND U-D COVARIANCE FACTORIZATIONS

1. Introduction: The Kalman Filter

The Kalman filter is the optimal state estimator for a system described by a linear model driven by white, Gaussian noise. On the basis of such a model and incomplete, noise-corrupted data from available sensors, it provides estimates of the system state variables that are optimal with respect to many different criteria for evaluating filter performance.

Assume that modelling techniques have produced an adequate system description in the form of a linear stochastic differential equation to describe the state propagation, with discrete-time noise-corrupted linear measurements available as system outputs. To describe this model, let a stochastic process be defined as a real vector-valued function $\underset{\sim}{a}(\cdot,\cdot)$ of two arguments, the first of which is an element of some time set T of interest (continuous or discrete), and the second an element ω of a probability sample space Ω, such that for any fixed t out of the set T, $\underset{\sim}{x}(t,\cdot)$ is a vector-valued random variable. If we fix the second argument instead of the first, we can say that, to each point ω_k out of Ω there can be associated a time function $\underset{\sim}{x}(\cdot,\omega_k) = \underline{x}(\cdot)$, each of which is a sample from the stochastic process. Sometimes the second argument is not included explicitly in the notation, so the underscored tilde is used to distinguish between a process $\underset{\sim}{a}(\cdot)$ and one of its samples $\underline{a}(\cdot)$. For instance, one can develop a stochastic process model $\underset{\sim}{z}(\cdot,\cdot)$ to

describe the measurements becoming available over time from sensors, while a single sample $\underset{\sim}{z}(\cdot,\omega_k) = \underset{\sim}{z}(\cdot)$ would correspond to a single measurement time history, and $\underset{\sim}{z}(t_i,\omega_k) = \underset{\sim}{z}(t_i) = \underset{\sim}{z}_i$ would be the vector of sensor outputs on a given time history at the particular time t_i.

Let the *state* process $\underset{\sim}{x}(\cdot,\cdot)$ of the system model satisfy the linear equation

$$\dot{\underset{\sim}{x}}(t) = \underline{F}(t)\underset{\sim}{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{G}(t)\,\underset{\sim}{w}(t) \tag{1}$$

where

$\underset{\sim}{x}(\cdot,\cdot) \triangleq$ n-vector state process

$\underline{u}(\cdot) \triangleq$ r-vector of piecewise continuous deterministic control input functions (more general input functions are possible, but piecewise continuous is adequate for practical applications)

$\underset{\sim}{w}(\cdot,\cdot) \triangleq$ s-vector dynamic noise process, modelled as white and Gaussian, with mean and autocorrelation for all t and t' out of T given by

$$E\{\underset{\sim}{w}(t)\} = \underline{0} \tag{2a}$$

$$E\{\underset{\sim}{w}(t)\underset{\sim}{w}^T(t')\} = \underline{Q}(t)\,\delta\,(t-t') \tag{2b}$$

where $\underline{Q}(\cdot)$ is an s-by-s matrix of piecewise continuous functions (most generally) such that $\underline{Q}(t)$ is symmetric and positive semidefinite for all t (i. e. all eigenvalues positive or zero), and $\delta(\cdot)$ is the Dirac delta function

$\underline{F}(\cdot) \triangleq$ n-by-n system dynamics matrix (of piecewise continuous functions in its general form

$\underline{B}(\cdot) \triangleq$ n-by-r deterministic input matrix

$\underline{G}(\cdot) \triangleq$ n-by-s dynamic noise input matrix

Note that if $\underline{F}$, $\underline{B}$, and $\underline{G}$ are in fact constant matrices, then the system dynamics model is termed time-invariant, rather than time-varying.

The state differential equation (1) is propagated forward from the initial condition $\underset{\sim}{x}(t_o)$. For any particular operation of the real system, the initial state assumes a specific value $\underline{x}(t_o)$. However, because this value may not be known precisely apriori, it is modelled as a Gaussian random vector. Thus, the probabilistic description of $\underline{x}(t_o)$ is completely specified by its mean $\hat{\underline{x}}_o$ and covariance $\underline{P}_o$:

$$E\{\underset{\sim}{x}(t_o)\} = \hat{\underline{x}}_o \tag{3a}$$

$$E\{[\underset{\sim}{x}(t_o) - \hat{\underline{x}}_o][\underset{\sim}{x}(t_o) - \hat{\underline{x}}_o]^T\} = \underline{P}_o \tag{3b}$$

where $\underline{P}_o$ is an n-by-n matrix that is symmetric and positive semidefinite. Allowing $\underline{P}_o$ to be positive semidefinite, instead of requiring positive definiteness (i.e. all eigenvalues strictly positive), admits the case of singular $\underline{P}_o$: the case in which some initial states or some linear combination of initial states are known precisely.

Measurements are available at discrete time points t_1, t_2,..., t_i,... (often, but not necessarily, equally spaced in time), and are modelled by the relation (for all t_i out of T):

$$\underset{\sim}{z}(t_i) = \underline{H}(t_i)\ \underset{\sim}{x}(t_i) + \underset{\sim}{v}(t_i) \tag{4}$$

where

$\underset{\sim}{z}(\cdot,\cdot) \triangleq$ m-vector discrete-time measurement process

$\underset{\sim}{x}(\cdot,\cdot) \triangleq$ state process; $\underset{\sim}{x}(t_i,\cdot)$ is a random vector corresponding to the state stochastic process at the particular sample time t_i

$\underset{\sim}{v}(\cdot,\cdot) \triangleq$ m-vector discrete-time measurement corruption noise process, modelled as white and Gaussian, with statistics (for all t_i, t_j taken from T):

$$E\ \{\underset{\sim}{v}(t_i)\ \} = \underline{0} \tag{5a}$$

$$E\ \{\underset{\sim}{v}(t_i)\underset{\sim}{v}^T(t_j)\} = \begin{cases} \underline{R}(t_i) & t_i = t_j \\ \underline{0} & t_i \neq t_j \end{cases} \tag{5b}$$

where $\underline{R}(t_i)$ is an m-by-m, symmetric, positive definite matrix for all t_i

$\underline{H}(\cdot) \triangleq$ m-by-n measurement matrix

In this description, positive definiteness of $\underline{R}(t_i)$ implies that all components of the measurement vector are corrupted by noise,

and there is no linear combination of these components that would be noise-free. The measurements modelled as in equation (4), and knowledge of our own inputs $\underline{u}$, are all that we have available from the real system under consideration.

It is further assumed that $\underset{\sim}{x}(t_o)$, $\underset{\sim}{w}(\cdot,\cdot)$ and $\underset{\sim}{v}(\cdot,\cdot)$ are independent of each other. Since all are assumed Gaussian, this is equivalent to assuming that they are uncorrelated with each other.

It is desired to combine the measurement data from the actual system with the information provided by the system model and statistical description of uncertainties, in order to obtain an "optimal" estimate of the system state. Under the modelling assumptions made, the Kalman filter provides state estimates that are optimal with respect to essentially all criteria used to evaluate estimators: Bayesian, maximum likelihood, weighted least squares, minimum variance unbiased, and maximum a posteriori, to name the most common. Some assumptions can be relaxed to yield a more general problem formulation, yielding restrictions upon the optimality of the filter or modifications in the algorithm itself (as in the case of allowing $\underset{\sim}{w}(\cdot,\cdot)$ and $\underset{\sim}{v}(\cdot,\cdot)$ to be correlated). However, the stated assumptions are the most common, and will serve to define the estimation problem for this report.

Before the algorithm is delineated, it will be convenient to introduce some additional notation. Define a composite vector which comprises the entire measurement history to the current sample time t_i, and denote it as $\underset{\sim}{Z}(t_i)$:

$$\underset{\sim}{Z}(t_i) \triangleq \begin{bmatrix} \underset{\sim}{z}(t_1) \\ \underset{\sim}{z}(t_2) \\ \vdots \\ \underset{\sim}{z}(t_i) \end{bmatrix} \tag{6}$$

This is a vector of growing dimension: at time t_i, it is a vector random variable of dimension $(i \cdot m)$. Its realized value, analogous to $\underset{\sim}{z}(t_i, \omega_k) = \underset{\sim}{z}_i$, is $\underset{\sim}{Z}_i$:

$$\underset{\sim}{Z}_i \triangleq \begin{bmatrix} \underset{\sim}{z}_1 \\ \underset{\sim}{z}_2 \\ \vdots \\ \underset{\sim}{z}_i \end{bmatrix} \tag{7}$$

The state estimate provided by the Kalman filter at time t_i, <u>before</u> the measurement $\underset{\sim}{z}(t_i, \omega_k) = \underset{\sim}{z}_i$ is available, is denoted as $\hat{\underset{\sim}{x}}(t_i^-)$. In fact, this estimate is the conditional mean of $\underset{\sim}{x}(t_i)$, conditioned on the previous measurement history $\underset{\sim}{Z}(t_{i-1}, \omega_k) = \underset{\sim}{Z}_{i-1}$, denoted as $E\{\underset{\sim}{x}(t_i) \mid \underset{\sim}{Z}(t_{i-1}) = \underset{\sim}{Z}_{i-1}\}$:

$$\hat{\underset{\sim}{x}}(t_i^-) \triangleq E\{\underset{\sim}{x}(t_i) \mid \underset{\sim}{Z}(t_{i-1}) = \underset{\sim}{Z}_{i-1}\} \tag{8}$$

Also provided by the algorithm is the conditional covariance of the state $\underset{\sim}{x}(t_i)$, conditioned on the previous measurement history:

$$\underline{P}(t_i^-) \triangleq E\{[\underset{\sim}{x}(t_i) - \hat{\underset{\sim}{x}}(t_i^-)]\ [\underset{\sim}{x}(t_i) - \hat{\underset{\sim}{x}}(t_i^-)]^T \mid \underset{\sim}{Z}(t_{i-1}) = \underline{Z}_{i-1}\} \tag{9}$$

Since $\hat{\underset{\sim}{x}}(t_i^-)$ is used as the state estimate, this is also the covariance of the error committed by the filter in its attempt to estimate the state. Moreover, it can be shown that $\underline{P}(t_i^-)$ is also the unconditional covariance of this error, so that its evaluation is completely independent of the particular measurement history one sees on each individual use of the filter.

The corresponding state estimate at the same time t_i, but <u>after</u> incorporating the measurement $\underline{z}_i$, is denoted as $\hat{\underline{x}}(t_i^+)$:

$$\hat{\underline{x}}(t_i+) \triangleq E\{\underset{\sim}{x}(t_i) \mid \underset{\sim}{Z}(t_i) = \underline{Z}_i\} \tag{10}$$

The associated error covariance matrix is then

$$\underline{P}(t_i^+) \triangleq E\{[\underset{\sim}{x}(t_i) - \hat{\underline{x}}(t_i^+)]\ [\underset{\sim}{x}(t_i) - \hat{\underline{x}}(t_i^+)]^T \mid \underset{\sim}{Z}(t_i) = \underline{Z}_i\} \tag{11}$$

The algorithm is composed of two parts, one for updating the state estimate and error covariance when a measurement becomes available, and the other for propagating those values to the next sample time. First, assuming $\hat{\underline{x}}(t_i^-)$ and $\underline{P}(t_i^-)$ to be already computed, the measurement $\underline{z}_i$ will be processed to yield $\hat{\underline{x}}(t_i^+)$ and $\underline{P}(t_i^+)$. Then the relations for determining $\hat{\underline{x}}(t_{i+1}^-)$ and $\underline{P}(t_{i+1}^-)$ will be evaluated, to complete one cycle of the recursion.

The <u>update</u> equations are

$$\underline{K}(t_i) = \underline{P}(t_i^-)\underline{H}^T(t_i)\ [\underline{H}(t_i)\underline{P}(t_i^-)\underline{H}^T(t_i) + \underline{R}(t_i)]^{-1} \tag{12}$$

$$\hat{\underline{x}}(t_i^+) = \hat{\underline{x}}(t_i^-) + \underline{K}(t_i)\,[\underline{z}_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^-)] \tag{13}$$

$$\underline{P}(t_i^+) = \underline{P}(t_i^-) - \underline{K}(t_i)\underline{H}(t_i)\underline{P}(t_i^-) \tag{14}$$

The $\underline{K}(t_i)$ evaluated in equation (12) is the Kalman filter gain. "High gains" would be caused by "large" $\underline{P}(t_i^-)$ - large uncertainty in the output of the filter's propagation model (to be discussed; due to "large" $\underline{Q}$), or by "small" $\underline{R}(t_i)$ - "small" corruption on the measurements. In either case, the state estimate in equation (13) is strongly influenced by the measurement data $\underline{z}_i$. In this equation, the bracketted term (the "innovation" or "residual") is the difference between the data $\underline{z}_i$ obtained from the actual measuring devices, and the best prediction of its value before it becomes available: $[\underline{H}(t_i)\hat{\underline{x}}(t_i^-)]$.

To propagate to the next sample time, t_{i+1}, the following differential equation can be integrated numerically:

$$\dot{\hat{\underline{x}}}(t) = \underline{F}(t)\hat{\underline{x}}(t) + \underline{B}(t)\underline{u}(t) \tag{15}$$

$$\dot{\underline{P}}(t) = \underline{F}(t)\underline{P}(t) + \underline{P}(t)\underline{F}^T(t) + \underline{G}(t)\underline{Q}(t)\underline{G}^T(t) \tag{16}$$

starting from the initial conditions $\hat{\underline{x}}(t_i^+)$ and $\underline{P}(t_i^+)$ respectively. The result of this integration would be $\hat{\underline{x}}(t_{i+1}^-)$ and $\underline{P}(t_{i+1}^-)$. A direct solution can be written in the form of

$$\hat{\underline{x}}(t_{i+1}^-) = \underline{\Phi}(t_{i+1}, t_i)\hat{\underline{x}}(t_i^+) + \int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1}, \tau)\underline{B}(\tau)\underline{u}(\tau)\,d\tau \tag{17}$$

$$\underline{P}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\underline{P}(t_i^+)\,\underline{\Phi}^T(t_{i+1},t_i) + \tag{17}$$

$$+ \int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1},\tau)\underline{G}(\tau)\underline{Q}(\tau)\underline{G}^T(\tau)\underline{\Phi}^T(t_{i+1},\tau)d\tau \tag{18}$$

where $\underline{\Phi}(t_{i+1},t)$ is the n-by-n state transition matrix that satisfies

$$\frac{d}{dt}\underline{\Phi}(t, t_i) = \underline{F}(t)\,\underline{\Phi}(t, t_i) \tag{19a}$$

for all t in the interval from t_i to t_{i+1}, starting from the initial condition

$$\underline{\Phi}(t_i,t_i) = \underline{I} \tag{19b}$$

Note that, if $\underline{u}(t)$ is held constant over each sample period (as would be the case for $\underline{u}$ being generated by a digital controller operating with the same sample period as the filter), then equation (17) simplifies to

$$\hat{\underline{x}}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\hat{\underline{x}}(t_i^+) + \left[\int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1},\tau)\underline{B}(\tau)d\tau\right]\underline{u}(t_i) \tag{20}$$

Equations (18) and (20) are propagation difference equations. They are the proper relations not only for a state model described by the stochastic differential equation (1), but also a discrete-time

model described by the stochastic difference equation

$$\underset{\sim}{x}(t_{i+1}) = \underline{\Phi}(t_{i+1},t_i)\underset{\sim}{x}(t_i) + \underline{B}_d(t_i)\underline{u}(t_i) + \underline{G}_d(t_i)\underset{\sim}{w}_d(t_i) \tag{21}$$

where the discrete-time control input matrix $\underline{B}_d(t_i)$ is defined by

$$\underline{B}_d(t_i) = \int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1},\tau)\underline{B}(\tau)d\tau \tag{22}$$

Furthermore, $\underset{\sim}{w}_d(\cdot,\cdot)$ is zero-mean white Gaussian discrete-time noise of strength $\underline{Q}_d(t_i)$ for all t_i,

$$E\{\underset{\sim}{w}_d(t_i)\underset{\sim}{w}^T(t_j)\} = \begin{cases} \underline{Q}_d(t_i) & t_i = t_j \\ \underline{0} & t_i \neq t_j \end{cases} \tag{23}$$

chosen such that the covariance contribution of $[\underline{G}_d(t_i)\underset{\sim}{w}_d(t_i)]$ to the difference equation (21) is equal to the covariance contribution of $[\underline{G}(t)\underset{\sim}{w}(t)]$ to the continuous-time model over the interval from t_i to t_{i+1} :

$$\underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i) = \int_{t_i}^{t_{i+1}} \underline{\Phi}(t_{i+1},\tau)\underline{G}(\tau)\underline{Q}(\tau)\underline{G}^T(\tau)\,\underline{\Phi}^T(t_{i+1},\tau)d\tau \tag{24}$$

In many instances, $\underline{G}_d(t_i)$ is chosen to be the identity matrix, and $\underset{\sim}{w}_d(\cdot,\cdot)$ is then an n-dimensional process with the properties just described.

Equation (21) is often termed an "equivalent discrete-time model" to the stochastic differential equation (1). It is "equivalent" in the sense that the statistical characterization of the process defined by (21) to (24) and initial condition (3), is identical to properties of the process described by (1) to (3) when viewed only at the sample times $t_1, t_2, \ldots$ etc. In the following sections, the system model and filtering algorithms will be posed in terms of this "equivalent" discrete-time system model.

2. The Wordlength Problem

Although the Kalman filter is the theoretically optimal solution to the filtering problem described in the previous section, the algorithm itself is prone to serious numerical difficulties. Measurement updating of the covariance matrix by equation (14) can involve the small difference of large numbers, especially in the case of very precise measuring devices. With finite wordlength, this can cause severe truncation and precision problems. Moreover, the Kalman filter algorithm can be shown to be numerically unstable, so that once such problems arise, their effects propagate rather than die out.

Thus, rather long wordlength is often required to maintain acceptable numerical accuracy. For small on-line computers, double precision computations are usually required. In fact, without double precision arithmetic, these numerical characteristics can easily become the dominant error source corrupting estimation precision, and unfortunately an error source usually not included in designers' error budgets.

The difficulties encountered in converting a tuned Kalman filter on a long wordlength, large computer system used for engineering design to an effective algorithm on a smaller wordlength online computer are well documented.[28,31] For instance, although it is theoretically impossible for the covariance matrix to have negative eigenvalues, such a condition can, and often does, result due to numerical computation using finite wordlength, especially when (1) the measurements are very accurate (eigenvalues of

$\underline{R}(t_i)$ are small relative to those of $\underline{P}(t_i^-)$, this being accentuated by large eigenvalues in $\underline{P}_o$) or (2) a linear combination of state vector components is known with great precision while other combinations are nearly unobservable (i.e., there is a large range of magnitudes of state covariance eigenvalues). Such a condition can lead to subsequent divergence or total failure of the recursion. On close inspection, even Kalman filters that maintain adequate estimation accuracy exhibit instances of negative covariance diagonal terms.[33]

To circumvent these problems in numerics inherent to the Kalman filter algorithm, alternate recursion relationships have been developed to propagate and update a state estimate and error covariance square root or inverse covariance square root instead of the covariance or its inverse themselves. Although equivalent algebraically to the conventional Kalman filter recursion, these "square root filters" exhibit improved numerical precision and stability, particularly in ill-conditioned problems (i.e., the cases described that yield erroneous results due to finite wordlength). The square root approach can yield twice the effective precision of the conventional filter in ill-conditioned problems. In other words, the same precision can be achieved with approximately half the wordlength. Moreover, this method is completely successful in maintaining the positive semidefiniteness of the error covariance.

These excellent numerical characteristics, combined with modest additional computation cycle time and memory storage requirements, make the square root filter approach a viable alternative to the conventional filter in many applications, especially when computer wordlength is limited or the estimation problem is ill-conditioned. The formulation of the square root filter for the case of no dynamic noise is especially attractive because of its

computational simplicity, and its outstanding numerical characteristics led to its implementation in the Apollo spacecraft navigation filters.

A number of practitioners have argued, with considerable logic, that square root filters should always be adopted in preference to the standard Kalman filter recursion, rather than switching to these more accurate and stable algorithms when and if numerical problems occur.[33] Even though Kalman algorithms can be made to work in most applications, by using double precision mathematics or scaling variables to reduce dynamic range or employing ad hoc modifications, numerics degrade performance from that achievable by numerically better conditioned recursions. Recent investigations tend to support an approach of designing and tuning an optimal filter by the methods of the two previous chapters, ignoring the errors caused by numerics, but then implementing the square root equivalent for online operation. Nevertheless, one can expect conventional Kalman algorithms to be applied rather extensively as well.

Section 3 introduces the concept of matrix square roots, and then Section 4 develops the initially designed and simplest covariance square root filter, applicable to the case of no dynamic driving noise and scalar measurements. The succeeding two sections generalize these results, first incorporating vector-valued measurements and then allowing dynamic driving noise. In Section 7, the square root counterpart to the inverse covariance formulation of the optimal filter is considered. Although it is not actually a square root filter, the U-D covariance factorization filter is very closely related to square root filtering, and it is depicted in Section 8. Finally, Section 9 presents the tradeoff of numerical advantages and increased computational burden of the square root filters.

3. Matrix Square Roots

Let $\underline{A}$ be an n-by-n, symmetric, positive semidefinite matrix. Then there exists at least one n-by-n "square root" matrix, denoted as $\sqrt{\underline{A}}$, such that

$$\sqrt{\underline{A}}\ \sqrt{\underline{A}}^{T} = \underline{A} \tag{25}$$

In fact, there are many matrices $\sqrt{\underline{A}}$ which satisfy (25) in general. The essential idea of square root filters is to replace the recursion for the error covariance $\underline{P}$ with a recursion for its square root, $\sqrt{\underline{P}}$, and to compute the state estimate using an optimal gain calculated in terms of $\sqrt{\underline{P}}$ instead of $\underline{P}$ itself. To motivate this, consider the scalar case: if dynamic range numerical precision problems are encountered in a filter that propagates the variance $P = \sigma^2$, the problem can be reduced by expressing the result in terms of the standard deviation σ since the dynamic range will be effectively reduced to half the original range. This basic idea can be generalized to the vector case by defining the state error covariance square roots, before and after measurement incorporation at time t_i, as $\underline{S}(t_i^-)$ and $\underline{S}(t_i^+)$ respectively, via:

$$\underline{S}(t_i^-)\ \underline{S}^{T}(t_i^-) \triangleq \underline{P}(t_i^-) \tag{26}$$

$$\underline{S}(t_i^+)\ \underline{S}^{T}(t_i^+) \triangleq \underline{P}(t_i^+) \tag{27}$$

Similarly, define the square root of the covariances depicting the strengths of discrete-time white Gaussian noises $\underset{\sim}{w}_d(\cdot,\cdot)$ and $\underset{\sim}{v}(\cdot,\cdot)$ as:

$$\underline{W}_d(t_i)\underline{W}_d^{T}(t_i) \triangleq \underline{Q}_d(t_i) \triangleq E\{\underset{\sim}{w}_d(t_i)\underset{\sim}{w}_d^{T}(t_i)\} \tag{28}$$

$$\underline{V}(t_i)\underline{V}^{T}(t_i) \triangleq \underline{R}(t_i) \triangleq E\{\underset{\sim}{v}(t_i)\underset{\sim}{v}^{T}(t_i)\} \tag{29}$$

The covariance square roots are <u>not uniquely</u> defined by (26) through (29), and square root filters can be formulated in terms of general matrix square roots. One means of exploiting this fact is to develop algorithms which

maintain a particularly attractive square root form, namely an upper or lower triangular matrix (with all zeroes below or above the main diagonal respectively), thereby requiring computation and storage of only $n(n+1)/2$ instead of n^2 scalar variables.

This lack of uniqueness does not cause difficulties in converting from a problem description in terms of initial $\underline{P}_o$ and time histories of $\underline{Q}_d(t_i)$ and $\underline{R}(t_i)$ to corresponding $\underline{S}_o$, $\underline{W}_d(t_i)$, and $\underline{V}(t_i)$ values, as might first appear to be the case. The reason is that any positive semidefinite matrix can be factored into the product of a lower triangular matrix and its transpose by the Cholesky decomposition algorithm.[12] Although (25) does not uniquely define $\sqrt{\underline{A}}$, a unique Cholesky lower triangular square root $\sqrt[c]{\underline{A}}$ can be defined such that $\sqrt[c]{\underline{A}}\ \sqrt[c]{\underline{A}}^T = \underline{A}$:

$$\begin{bmatrix} \sqrt[c]{A}_{11} & 0 & \cdots & 0 \\ \sqrt[c]{A}_{12} & \sqrt[c]{A}_{22} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt[c]{A}_{1n} & \sqrt[c]{A}_{2n} & \cdots & \sqrt[c]{A}_{nn} \end{bmatrix} \begin{bmatrix} \sqrt[c]{A}_{11} & \sqrt[c]{A}_{12} & \cdots & \sqrt[c]{A}_{1n} \\ 0 & \sqrt[c]{A}_{22} & \cdots & \sqrt[c]{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt[c]{A}_{nn} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{12} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

The elements of the Cholesky square root matrix can be generated sequentially, row by row, from the recursion: for i=1,2,...,n, compute:

$$\sqrt[c]{A}_{ij} = \begin{cases} \dfrac{1}{\sqrt[c]{A}_{jj}} \left[A_{ij} - \sum_{k=1}^{j-1} \sqrt[c]{A}_{ik}\ \sqrt[c]{A}_{jk} \right] & j=1,2,\ldots,i-1 \\ \sqrt{A_{ii} - \sum_{k=1}^{i-1} \sqrt[c]{A}_{ik}^2} & j=i \\ 0 & j > i \end{cases} \tag{30}$$

Thus, $\underline{A}$ is scanned and $\sqrt[c]{\underline{A}}$ is generated in the order depicted in Fig 1.

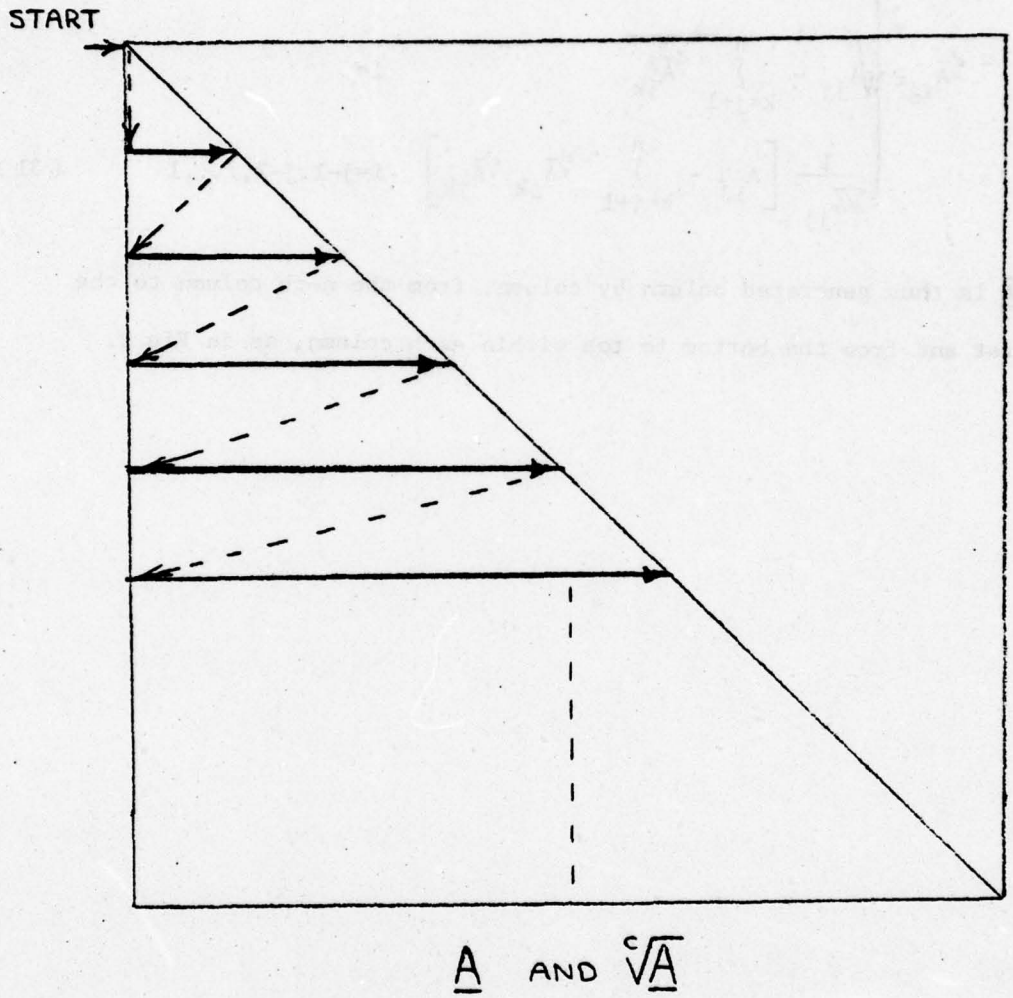

Fig 1. Scanning of A and Generation of $\sqrt[c]{A}$

Later in the Carlson filter[10] of Section 6, we will have occasion to seek an upper triangular Cholesky square root ${}^{c'}\sqrt{\underline{A}}$ such that ${}^{c'}\sqrt{\underline{A}}\ {}^{c'}\sqrt{\underline{A}}^{T} = \underline{A}$. Such a matrix can be found by operating (30) backwards, or specifically, for j=n,n-1,...,1, perform the following computations:

$$ {}^{c'}\sqrt{A}_{ij} = \begin{cases} 0 & i > j \\ \sqrt{A_{jj} - \sum_{k=j+1}^{n} {}^{c'}\sqrt{A}_{jk}^{2}} & i=j \\ \frac{1}{{}^{c'}\sqrt{A}_{jj}} \left[A_{ij} - \sum_{k=j+1}^{n} {}^{c'}\sqrt{A}_{ik}\ {}^{c'}\sqrt{A}_{jk} \right] & i=j-1,j-2,\ldots,1 \end{cases} \qquad (31) $$

${}^{c'}\sqrt{\underline{A}}$ is thus generated column by column, from the n-th column to the first and from the bottom to top within each column, as in Fig 2.

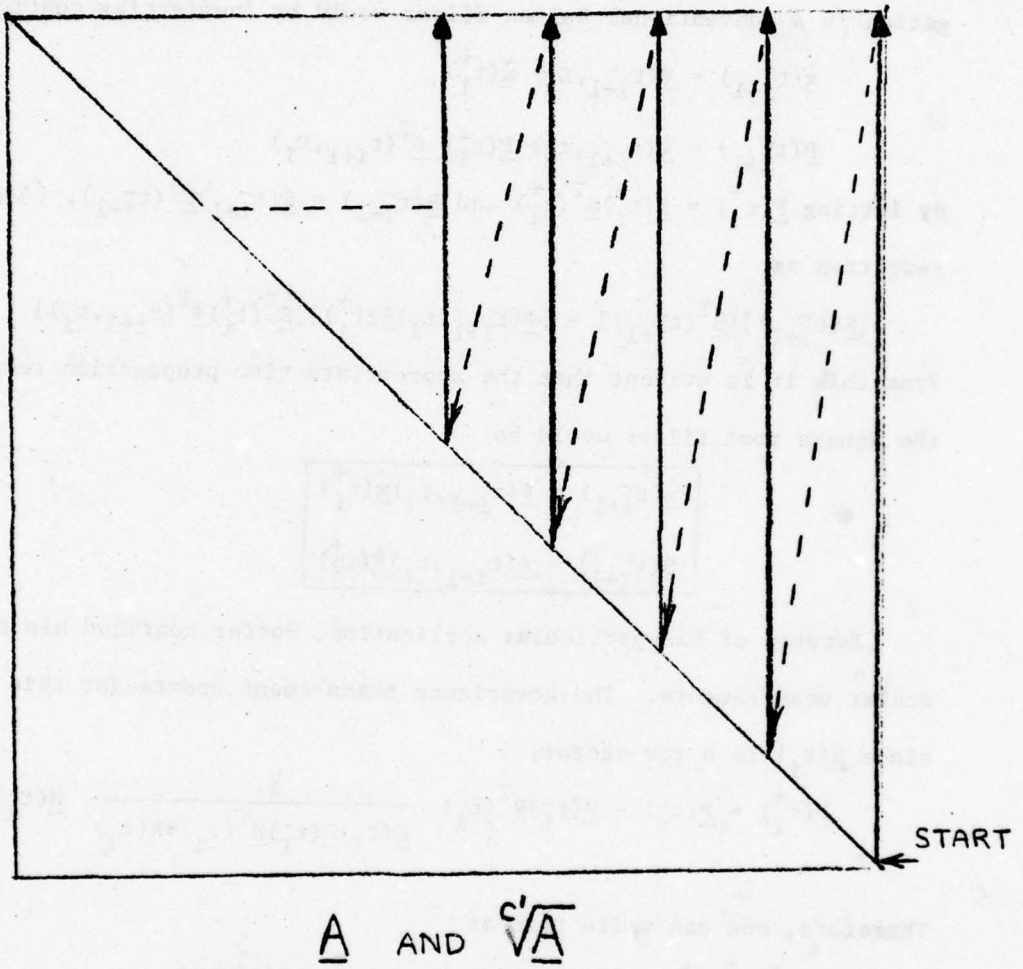

Fig 2. Scanning of $\underline{A}$ and Generation of $\sqrt[c']{\underline{A}}$

4. Covariance Square Root Filter for $\underline{Q}_d \equiv \underline{0}$

In 1964, Potter developed a square root filter implementation for space navigation applications in which there was no dynamic driving noise in the system model, i.e. $\underline{Q}_d(t_i) \equiv \underline{0}$ for all time, motivated by restricted wordlength in the Apollo Guidance Computer. For this case, the time propagation in a conventional Kalman filter would be (neglecting control inputs):

$$\hat{\underline{x}}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\,\hat{\underline{x}}(t_i^+) \tag{32a}$$

$$\underline{P}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\,\underline{P}(t_i^+)\,\underline{\Phi}^T(t_{i+1},t_i) \tag{32b}$$

By letting $\underline{P}(t_i^+) = \underline{S}(t_i^+)\underline{S}^T(t_i^+)$ and $\underline{P}(t_{i+1}^-) = \underline{S}(t_{i+1}^-)\underline{S}^T(t_{i+1}^-)$, (32b) can be rewritten as

$$[\underline{S}(t_{i+1}^-)][\underline{S}^T(t_{i+1}^-)] = [\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+)][\underline{S}^T(t_i^+)\underline{\Phi}^T(t_{i+1},t_i)]$$

From this it is evident that the appropriate time propagation relations for the square root filter would be

$$\boxed{\hat{\underline{x}}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\hat{\underline{x}}(t_i^+)} \tag{33a}$$

$$\boxed{\underline{S}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+)} \tag{33b}$$

Because of his particular application, Potter confined his attention to scalar measurements. The covariance measurement update for this case is, since $\underline{H}(t_i)$ is a row vector,

$$\underline{P}(t_i^+) = \underline{P}(t_i^-) - \underline{P}(t_i^-)\underline{H}^T(t_i)\,\frac{1}{\underline{H}(t_i)\underline{P}(t_i^-)\underline{H}^T(t_i)+R(t_i)}\,\underline{H}(t_i)\underline{P}(t_i^-) \tag{34}$$

Therefore, one can write this as

$$\underline{S}(t_i^+)\underline{S}^T(t_i^+) = \underline{S}(t_i^-)[\underline{I}-b(t_i)\underline{a}(t_i)\underline{a}^T(t_i)]\underline{S}^T(t_i^-) \tag{35}$$

where by n-by-1 $\underline{a}(t_i)$ and scalar $b(t_i)$ are defined by

$$\underline{a}(t_i) = \underline{S}^T(t_i^-)\underline{H}^T(t_i) \tag{36a}$$

$$\frac{1}{b(t_i)} = \underline{a}^T(t_i)\underline{a}(t_i) + R(t_i) \tag{36b}$$

Potter showed that the bracketted term in (35) can be factored into

$$[\underline{I}-b\underline{a}\underline{a}^T] = [\underline{I} - b\gamma\underline{a}\underline{a}^T][I - b\gamma\underline{a}\underline{a}^T]^T \tag{37}$$

where γ is a scalar defined by

$$\gamma = \frac{1}{1 + \sqrt{bR}} \tag{38}$$

Substituting this into (35) yields the covariance update as

$$\begin{aligned}\underline{S}(t_i^+) &= \underline{S}(t_i^-)\ [\underline{I} - b(t_i)\gamma(t_i)\underline{a}(t_i)\underline{a}^T(t_i)] \\ &= \underline{S}(t_i^-) - b(t_i)\gamma(t_i)\underline{S}(t_i^-)\underline{a}(t_i)\underline{a}^T(t_i)\end{aligned} \tag{39}$$

The state estimate measurement update is of the conventional form, but with the Kalman gain evaluated as $[b(t_i)\underline{S}(t_i^-)\underline{a}(t_i)]$. Thus, the measurement update becomes:

$$\boxed{\begin{aligned}\underline{a}(t_i) &= \underline{S}^T(t_i^-)\underline{H}^T(t_i) \\ b(t_i) &= 1/[\underline{a}^T(t_i)\underline{a}(t_i)+R(t_i)] \\ \gamma(t_i) &= 1/[1 + \sqrt{b(t_i)R(t_i)}] \\ \underline{K}(t_i) &= b(t_i)\ \underline{S}(t_i^-)\ \underline{a}(t_i) \\ \hat{\underline{x}}(t_i^+) &= \hat{\underline{x}}(t_i^-) + \underline{K}(t_i)\ [z_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^-)] \\ \underline{S}(t_i^+) &= \underline{S}(t_i^-) - \gamma(t_i)\underline{K}(t_i)\underline{a}^T(t_i)\end{aligned}} \tag{40}$$

An equivalent form that is often employed is:

$$\boxed{\begin{aligned}a(t_i) &= \underline{S}^T(t_i^-)\underline{H}^T(t_i) \\ \sigma(t_i) &= \sqrt{\underline{a}^T(t_i)\underline{a}(t_i)+R(t_i)} \\ \alpha(t_i) &= \sigma(t_i) + V(t_i) \\ \beta(t_i) &= 1/[\sigma(t_i)\alpha(t_i)] \\ \underline{g}(t_i) &= \beta(t_i)[\underline{S}(t_i^-)\underline{a}(t_i)] \\ \hat{\underline{x}}(t_i^+) &= \hat{\underline{x}}(t_i^-) + \underline{g}(t_i)\{[\alpha(t_i)/\sigma(t_i)][z_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^-)]\} \\ \underline{S}(t_i^+) &= \underline{S}(t_i^-) - \underline{g}(t_i)\underline{a}^T(t_i)\end{aligned}} \tag{41}$$

Note that, even if $\underline{S}(t_i^-)$ is lower triangular, $\underline{S}(t_i^+)$ will generally not be lower triangular when this update form is used. A method that preserves the lower triangular nature of the covariance square root will be discussed later.

5. Vector-Valued Measurements

The preceding section considered scalar measurement updates. Bellantoni and Dodge[3] extended these results to the vector measurement case by using eigenvalue decompositions, but their algorithm is inefficient for the typical case in which the measurement vector dimenstion m is significantly less than the state dimension n. Andrews[2] also developed an update that processed an m-dimensional measurement vector in a single update, without requiring diagonalization:

$$\left.\begin{aligned}
\underline{A}(t_i) &= \underline{S}^T(t_i^-)\underline{H}^T(t_i) \\
\underline{\Sigma}(t_i) &= \sqrt[c]{\underline{A}^T(t_i)\underline{A}(t_i) + \underline{R}(t_i)} \\
\hat{\underline{x}}(t_i^+) &= \hat{\underline{x}}(t_i^-) + \underline{S}(t_i^-)\underline{A}(t_i)[\underline{\Sigma}^{-1}(t_i)]^T\underline{\Sigma}^{-1}(t_i)[\underline{z}_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^-)] \\
\underline{S}(t_i^+) &= \underline{S}(t_i^-) - \underline{S}(t_i^-)\underline{A}(t_i)[\underline{\Sigma}^{-1}(t_i)]^T[\underline{\Sigma}(t_i)+\underline{V}(t_i)]^{-1}\underline{A}^T(t_i)
\end{aligned}\right\} \quad (42)$$

This can be seen to be a direct extension of (41), and it is more efficient computationally than the Bellantoni and Dodge algorithm. Processing a measurement entails a Cholesky decomposition of an m-by-m matrix to generate $\underline{\Sigma}(t_i)$, (the extension of $\sigma(t_i)$) and inversion of two triangular m-by-m matrices, $\underline{\Sigma}(t_i)$ and $[\underline{\Sigma}(t_i) + \underline{V}(t_i)]$.

For the covariance square root filter, the most efficient means of performing a vector measurement update is to employ the Potter scalar update, (40) or (41), , repeatedly m times. An m-dimensional measurement vector $\underline{z}_i$ can <u>always</u> be processed equivalently as m scalar measurements. If $\underline{R}(t_i)$ is diagonal, the m components can be treated as independent measurements and processed sequentially. If $\underline{R}(t_i)$ is not diagonal, the procedure is somewhat more complicated. First the Cholesky decomposition of $\underline{R}(t_i)$ is computed, yielding $\sqrt[c]{\underline{R}(t_i)}$ as a lower triangular matrix. Then a transformation of variables is used to convert

$$\underset{\sim}{z}(t_i) = \underline{H}(t_i)\underset{\sim}{x}(t_i) + \underset{\sim}{v}(t_i) \tag{43}$$

into

$$\underset{\sim}{z}^*(t_i) = \underline{H}^*(t_i)\underset{\sim}{x}(t_i) + \underset{\sim}{v}^*(t_i) \tag{44}$$

where

$$\sqrt[c]{\underline{R}(t_i)}\ \underset{\sim}{z}^*(t_i) = \underset{\sim}{z}(t_i) \tag{45a}$$

$$\sqrt[c]{\underline{R}(t_i)}\ \underline{H}^*(t_i) = \underline{H}(t_i) \tag{45b}$$

$$\sqrt[c]{\underline{R}(t_i)}\ \underset{\sim}{v}^*(t_i) = \underset{\sim}{v}(t_i) \tag{45c}$$

Note that (45c) implies that $\underset{\sim}{v}^*(\cdot,\cdot)$ is a unit power white Gaussian noise, i.e., $E\{\underset{\sim}{v}^*(t_i)\underset{\sim}{v}^{*T}(t_i)\} = \underline{I}$, since

$$E\{\underset{\sim}{v}(t_i)\underset{\sim}{v}^T(t_i)\} = \underline{R}(t_i) = E\{\sqrt[c]{\underline{R}(t_i)}\ \underset{\sim}{v}^*(t_i)\underset{\sim}{v}^{*T}(t_i)\ \sqrt[c]{\underline{R}(t_i)}^T\}$$

$$= \sqrt[c]{\underline{R}(t_i)}\ E\{\underset{\sim}{v}^*(t_i)\underset{\sim}{v}^{*T}(t_i)\}\ \sqrt[c]{\underline{R}(t_i)}^T$$

Thus, the components of $\underset{\sim}{z}^*(t_i,\omega_j) = \underline{z}_i^*$ can be processed one at a time sequentially. Moreover, (45a) and (45b) can be solved to yield $\underline{z}_i^*$ and $\underline{H}^*(t_i)$ by simple back substitution rather than matrix inversion, because $\sqrt[c]{\underline{R}(t_i)}$ is lower triangular and thus the j-th component of $\underline{z}_i^*$ is a linear combination of the first j components of $\underline{z}_i$.

As noted in the previous section, the $\underline{S}(t_i^+)$ matrix generated by these update forms is generally not lower triangular, even if $\underline{S}(t_i^-)$ is.

6. Covariance Square Root Filter for $\underline{Q}_d \neq 0$

If dynamic driving noise enters the system model, the conventional Kalman filter propagates the covariance matrix from one measurement time to the next by means of:

$$\underline{P}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\,\underline{P}(t_i^+)\,\underline{\Phi}^T(t_{i+1},t_i) + \underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i) \qquad (46)$$

where this might be an equivalent discrete-time representation of a continuous-time system with sampled output (in which case $\underline{G}_d(t_i) \equiv \underline{I}$). Now we wish to develop an analogous recursion to yield $\underline{S}(t_{i+1}^-)$ in terms of $\underline{S}(t_i^+)$. It would be desirable to generate a lower triangular $\underline{S}(t_{i+1}^-)$ since then only 1/2 n(n+1) elements would require computation rather than n^2.

One means of achieving the desired result is called the Matrix RSS (Root-Sum-Square)[10] method:

$$\left.\begin{aligned} \underline{X}(t_{i+1}) &= \underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+) \\ \underline{P}(t_{i+1}^-) &= \underline{X}(t_{i+1})\,\underline{X}^T(t_{i+1}) + [\underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i)] \\ \underline{S}(t_{i+1}^-) &= {}^c\sqrt{\underline{P}(t_{i+1}^-)} \end{aligned}\right\} (47)$$

This method actually computes $\underline{P}(t_{i+1}^-)$ and then generates $\underline{S}(t_{i+1}^-)$ as its lower triangular Cholesky square root. Although this is a rapid method, it does suffer in having only the same numerical precision as the conventional filter time propagation. Nevertheless, since it is the measurement update and not the time propagation that causes the critical numerical problems in the filter, (47) may well be acceptable for many applications.

The other means of establishing the time propagation relations is called the Triangularization Method.[22] In Section 4, the desired result (33b) was established by writing both sides of the covariance propagation (32b) in terms of a factor times its own transpose, and then equating the indi-

vidual factors. Let us attempt to apply the same logic to (46). Assume that the square roots of $\underline{P}(t_i^+)$ and $\underline{Q}_d(t_i)$ are available: for $\underline{P}_o$ and $\underline{Q}_d(t_i)$ for all t_i, a Cholesky decomposition could be used, and for $\underline{P}(t_i^+)$ in general, assume the square root has been propagated and updated by the filter algorithm. Thus, we have:

$$\underline{P}(t_i^+) = \underline{S}(t_i^+)\ \underline{S}^T(t_i^+) \tag{48a}$$

$$\underline{Q}_d(t_i) = \underline{W}_d(t_i)\ \underline{W}_d^T(t_i) \tag{48b}$$

Note that $\underline{S}(t_i^+)$ need not be lower triangular (important in view of the preceding section). Equation (46) can, therefore be written as

$$\underline{P}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+)\underline{S}^T(t_i^+)\underline{\Phi}^T(t_{i+1},t_i) + \underline{G}_d(t_i)\underline{W}_d(t_i)\underline{W}_d^T(t_i)\underline{G}_d^T(t_i) \tag{49}$$

Now it is desired to find the propagation equation for the square root of $\underline{P}(t_{i+1}^-)$: to find the relation to yield $\underline{S}(t_{i+1}^-)$ such that $\underline{S}(t_{i+1}^-)\underline{S}^T(t_{i+1}^-)$ is equal to the right hand side of (49).

One such matrix would be $\tilde{\underline{S}}(t_{i+1}^-)$ defined by

$$\tilde{\underline{S}}(t_{i+1}^-) = [\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+) \;\vdots\; \underline{G}_d(t_i)\underline{W}_d(t_i)] \tag{50}$$

However, if $\underline{S}(t_i^+)$ is n-by-n, then $[\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+)]$ is n-by-n and $[\underline{G}_d(t_i)\underline{W}_d(t_i)]$ is n-by-s, so $\tilde{\underline{S}}(t_{i+1}^-)$ would be an n-by-(n+s) square root of $\underline{P}(t_{i+1}^-)$. Since this type of square root increases the dimension of the covariance square root matrix for each propagation interval, it must be rejected as computationally impractical.

However, this does in fact provide a fruitful insight. If $\tilde{\underline{S}}(t_{i+1}^-)$ is a square root of $\underline{P}(t_{i+1}^-)$, then so is $[\tilde{\underline{S}}(t_{i+1}^-)\underline{T}]$ if $\underline{T}$ is an orthogonal (n+s)-by-(n+s) matrix, i.e. $\underline{T}\ \underline{T}^T = \underline{I}$, since

$$\tilde{\underline{S}}(t_{i+1}^-)\underline{T}\,\underline{T}^T\tilde{\underline{S}}^T(t_{i+1}^-) = \tilde{\underline{S}}(t_{i+1}^-)\tilde{\underline{S}}^T(t_{i+1}^-) \tag{51}$$

Therefore, if an orthogonal matrix $\underline{T}$ can be found such that

$$\tilde{\underline{S}}(t_{i+1}^-)\,\underline{T} = [\underbrace{\underline{S}(t_{i+1}^-)}_{n \text{ columns}} \;\vdots\; \underbrace{\underline{0}}_{s \text{ columns}}] \;\}n \text{ rows} \tag{52}$$

then in fact an n-by-n square root matrix $\underline{S}(t_{i+1}^-)$ will have been found which satisfies the desired relationship. If, in addition, this $\underline{S}(t_{i+1}^-)$ were lower triangular, the result would be especially advantageous. Two methods[22] of generating such a $\underline{S}(t_{i+1}^-)$, known as triangularization algorithms, are the Gram-Schmidt orthogonalization[14,3] procedure and the Householder transformation[12,13] technique. Note that the same procedure could also be applied to

$$\underline{P}(t_i^+) = [\underline{I}-\underline{K}(t_i)\underline{H}(t_i)]\underline{P}(t_i^-)[\underline{I}-\underline{K}(t_i)\underline{H}(t_i)]^T + \underline{K}(t_i)\underline{R}(t_i)\underline{K}^T(t_i)$$

or

$$\underline{P}^{-1}(t_i^+) = \underline{P}^{-1}(t_i^-) + \underline{H}^T(t_i)\underline{R}^{-1}(t_i)\underline{H}(t_i)$$

for vector measurement updates; the latter of these will be discussed subsequently.

First let us demonstrate that the Gram-Schmidt procedure yields the desired result. Let $\underline{e}^k$ denote the n-dimensional vector composed of all zeroes except for a one as the k-th component, so that $\underline{e}^1, \underline{e}^2, \ldots, \underline{e}^n$ form the standard basis for n-dimensional space. Then $[\tilde{\underline{S}}^T(t_{i+1}^-)\underline{e}^k]$ is just the k-th column of $\tilde{\underline{S}}^T(t_{i+1}^-)$, of dimension (n+s):

$$\tilde{\underline{S}}^T(t_{i+1}^-) = \underbrace{\left[\begin{array}{cccc} | & | & & | \\ \tilde{\underline{s}}^1 & \tilde{\underline{s}}^2 & \cdots & \tilde{\underline{s}}^n \\ | & | & & | \end{array}\right]}_{n \text{ columns}} \Bigg\}(n+s) \text{ rows} \tag{53}$$

where

$$\left.\begin{aligned} \tilde{\underline{s}}^1 &= \tilde{\underline{S}}^T(t_{i+1}^-)\ \underline{e}^1 \\ \tilde{\underline{s}}^2 &= \tilde{\underline{S}}^T(t_{i+1}^-)\ \underline{e}^2 \\ &\vdots \\ \tilde{\underline{s}}^n &= \tilde{\underline{S}}^T(t_{i+1}^-)\ \underline{e}^n \end{aligned}\right\} \quad (54)$$

Construct the orthonormal basis vectors $\underline{b}^1, \underline{b}^2, \ldots, \underline{b}^n$ (each of dimension (n+s)) by the Gram-Schmidt procedure as:

$$\left.\begin{aligned} \underline{b}^1 &= \text{unit}\ (\tilde{\underline{s}}^1) \\ \underline{b}^2 &= \text{unit}\ (\tilde{\underline{s}}^2 - [\tilde{\underline{s}}^{2T}\underline{b}^1]\underline{b}^1) \\ \underline{b}^3 &= \text{unit}\ (\tilde{\underline{s}}^3 - [\tilde{\underline{s}}^{3T}\underline{b}^1]\underline{b}^1 - [\tilde{\underline{s}}^{3T}\underline{b}^2]\underline{b}^2) \\ &\vdots \end{aligned}\right\} \quad (55)$$

If $\underline{P}(t_i^+)$ is positive definite, then $\tilde{\underline{S}}^T(t_{i+1}^-)$ is of rank n, so n such orthogonal unit basis vectors can be obtained. Now the desired orthogonal transformation matrix $\underline{T}$ can be defined as the (n+s)-by-(n+s) matrix

$$\underline{T} = [\underline{b}^1, \underline{b}^2, \ldots, \underline{b}^n, \underline{b}^{n+1}, \ldots, \underline{b}^{n+s}] \quad (56)$$

where $\underline{b}^1, \underline{b}^2, \ldots, \underline{b}^n$ have been computed as in (55) and the remaining s columns, $\underline{b}^{n+1}, \ldots, \underline{b}^{n+s}$, are additional orthogonal unit basis vectors of (n+s)-dimensional space. However, it will be shown that they do not have to be computed to obtain $\underline{S}(t_{i+1}^-)$, so their generation will not be specified explicitly.

At this point, $\underline{T}^T\ \tilde{\underline{S}}^T(t_{i+1}^-)$ can be written as

$$\underline{T}^T \tilde{\underline{S}}^T(t_{i+1}^-) = \begin{bmatrix} — \ \underline{b}^{1T} \ — \\ — \ \underline{b}^{2T} \ — \\ \vdots \\ — \ \underline{b}^{(n+s)T} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \tilde{\underline{s}}^1 & \tilde{\underline{s}}^2 & \cdots & \tilde{\underline{s}}^n \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \underline{b}^{1T}\tilde{\underline{s}}^1 & \underline{b}^{1T}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{1T}\tilde{\underline{s}}^n \\ \underline{b}^{2T}\tilde{\underline{s}}^1 & \underline{b}^{2T}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{2T}\tilde{\underline{s}}^n \\ \vdots & \vdots & & \vdots \\ \underline{b}^{nT}\tilde{\underline{s}}^1 & \underline{b}^{nT}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{nT}\tilde{\underline{s}}^n \\ \vdots & \vdots & & \vdots \\ \underline{b}^{(n+s)T}\tilde{\underline{s}}^1 & \underline{b}^{(n+s)T}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{(n+s)T}\tilde{\underline{s}}^n \end{bmatrix} \qquad (57)$$

However, since the rank of $\tilde{\underline{S}}^T(t_{i+1}^-)$ is n and $\{\underline{b}^1, \underline{b}^2, \ldots, \underline{b}^n\}$ span its range space while $\{\underline{b}^{n+1}, \ldots, \underline{b}^{n+s}\}$ are orthogonal to this spanning set, it follows that the last s rows in (57) are all zeroes. By the manner in which the basis vectors were chosen, it is also true that

$$\underline{b}^{kT}\tilde{\underline{s}}^j = 0 \qquad k > j$$

by the same reasoning. Thus, (57) becomes

$$\underline{T}^T\tilde{\underline{S}}^T(t_{i+1}^-) = \left[\begin{array}{cccc} \underline{b}^{1T}\tilde{\underline{s}}^1 & \underline{b}^{1T}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{1T}\tilde{\underline{s}}^n \\ & \underline{b}^{2T}\tilde{\underline{s}}^2 & \cdots & \underline{b}^{2T}\tilde{\underline{s}}^n \\ \underline{0} & & & \vdots \\ & & & \underline{b}^{nT}\underline{s}^n \\ \hline \underline{0} & & & \end{array}\right] \begin{array}{l} \} \text{n rows} \\ \\ \} \text{s rows} \end{array}$$

n columns (58)

The upper n-by-n partition of this matrix is just the $\underline{S}^T(t_{i+1}^-)$ we have been seeking, so that $\underline{S}(t_{i+1}^-)$ is in fact an n-by-n lower triangular matrix:

$$\underline{S}(t_{i+1}^-) = \begin{bmatrix} \underline{b}^{1T}\underline{s}^1 & & & \\ \underline{b}^{1T}\underline{s}^2 & \underline{b}^{2T}\underline{s}^2 & & \underline{0} \\ \vdots & \vdots & \ddots & \\ \underline{b}^{1T}\underline{s}^n & \underline{b}^{2T}\underline{s}^n & \cdots & \underline{b}^{nT}\underline{s}^n \end{bmatrix} \tag{59}$$

An efficient computational form of the Gram-Schmidt orthogonalization called the Modified Gram-Schmidt (MGS)[22,30] algorithm has been shown to be numerically superior to the straightforward classical procedure, i.e. less susceptible to roundoff errors.[8,22] Moreover, it requires no more arithmetic operations than the conventional Gram-Schmidt procedure, uses less storage, and has been shown[25] to have numerical accuracy comparable to Householder[18,23] and Givens[14] transformations. Generation of $\underline{S}(t_{i+1}^-)$ through this recursion proceeds as follows. Define the initial condition on $\underline{A}^k$, an (n+s)-by-n matrix, as

$$\boxed{\underline{A}^1 = \underline{\tilde{S}}^T(t_{i+1}^-) = [\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+) \;\vdots\; \underline{G}_d(t_i)\underline{W}_d(t_i)]^T} \tag{60}$$

Notationally let $\underline{A}_j^k$ denote the j-th column of $\underline{A}^k$. Perform the n-step recursion, for k=1,2,...,n:

$$\boxed{\begin{aligned} a^k &= \sqrt{\underline{A}_k^{kT}\,\underline{A}_k^k} \\ C_{kj} &= \begin{cases} 0 & j = 1,\ldots,k-1 \\ a^k & j = k \\ \left[(1/a^k)\,\underline{A}_k^{kT}\right]\underline{A}_j^k & j = k+1,\ldots,n \end{cases} \\ \underline{A}_j^{k+1} &= \underline{A}_j^k - C_{kj}[(1/a^k)\underline{A}_k^k] \qquad j = k+1,\ldots,n \end{aligned}} \tag{61}$$

Note that on successive iterations, the new $\underline{A}_j^{k+1}$ vectors can be "written over" the $\underline{A}_j^k$ vectors to conserve memory. At the end of this recursion,

$$\boxed{\underline{S}(t_{i+1}^-) = \underline{C}^T} \tag{62}$$

Notice that the computational algorithm never calculates or stores $\underline{T}$ explicitly in generating $\underline{S}(t_{i+1}^-)$.

A Householder transformation[19] can also be used to solve (52) for the square root matrix $\underline{S}(t_{i+1}^-)$. Conceptually, it generates $\underline{T}$ as

$$\underline{T} = \underline{T}^n \, \underline{T}^{(n-1)} \, \ldots \, \underline{T}^1$$

where $\underline{T}^k$ is generated recursively as

$$\underline{T}^k = \underline{I} - d^k \, \underline{u}^k \, \underline{u}^{kT}$$

with the scalar d^k and the (n+s)-vector $\underline{u}^k$ defined below. However, the computational algorithm never calculates these $\underline{T}^k$'s or $\underline{T}$ explicitly. The initial condition on the (n+s)-by-n $\underline{A}^k$ is

$$\boxed{\underline{A}^1 = \tilde{\underline{S}}^T(t_{i+1}^-) = [\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+) \;\vdots\; \underline{G}_d(t_i)\underline{W}_d(t_i)]^T} \tag{63}$$

Again, letting $\underline{A}_j^k$ represent the j-th column of $\underline{A}^k$, perform the n-step recursion, for k=1,2,...,n:

$$\boxed{\begin{aligned}
a^k &= \sqrt{\sum_{j=k}^{n+s} [A_{jk}^k]^2} \cdot \operatorname{sgn}\{A_{kk}^k\} \\
d^k &= \frac{1}{a^k(a^k + A_{kk}^k)} \\
u_j^k &= \begin{cases} 0 & , \; j < k \\ a^k + A_{kk}^k & , \; j = k \\ A_{jk}^k & , \; j = (k+1),\ldots,(n+s) \end{cases} \\
y_j^k &= \begin{cases} 0 & j < k \\ 1 & j = k \\ d^k \, \underline{u}^{kT} \, \underline{A}_j^k & , \; j = (k+1),\ldots,n \end{cases} \\
\underline{A}^{k+1} &= \underline{A}^k - \underline{u}^k \underline{y}^{kT}
\end{aligned}} \tag{64}$$

At stage k, the first (k-1) columns of $\underline{A}^k$ are zero below the diagonal of the upper square partition, and $\underline{u}^k$ has been chosen so that the subdiagonal elements of $\underline{A}_k^{k+1}$ will be zero. After the n iterations of (64).

$$\underline{A}^{n+1} = \begin{bmatrix} \underline{C} \\ \hline \underline{0} \end{bmatrix} \begin{matrix} \} n \\ \} s \end{matrix} \qquad \underbrace{\phantom{\underline{C}}}_{n} \tag{65}$$

and then $\underline{S}(t_{i+1}^-)$ is generated as

$$\underline{S}(t_{i+1}^-) = \underline{C}^T \tag{66}$$

The Householder triangularization requires $1/6[4n^3 + 6n^2(s+1) + 2n]$ multiplies, $1/6[4n^3 + 6sn^2 + 8n]$ adds, n divides, and n square roots. This is $1/6[2n^3 - 8n]$ fewer multiplies and $1/6[2n^3 + 3n^2 - 11n]$ fewer adds than required by the Modified Gram-Schmidt algorithm. However, the MGS algorithm becomes slightly more precise numerically as the residual size increases, and thus is a viable alternative.

A Householder transformation method has also been proposed for performing measurement updates.[11,22] However, this has been shown to be equivalent to the Potter method described previously, but not as efficient computationally.

Thus, the covariance square root filter (Potter filter) algorithm can be specified as follows. The propagation of the state estimate from one sample time to the next is given by (33a). Covariance square root time propagations are calculated by means of the matrix RSS method (47), the MGS algorithm given by (60) to (62), or the Householder transformation as in (63) to (66). Of these, the latter two are preferable since they are more accurate numerically than the computationally efficient first

method, and numerics are the basic motivation for square root forms. Measurement updates would be processed through m iterations of the Potter algorithm (40) or (41). If the $\underline{R}(t_i)$ matrix is not diagonal, the transformation of variables given by (43) to (45) must first be performed.

One significant drawback of the covariance square root filter just described is that the triangularity of the square root matrix is generally destroyed during the measurement updating. Consequently, all n^2 elements must be computed and stored. A more recent method developed by Carlson[12] provides substantial improvement in both computational speed and required storage by maintaining the covariance square root matrix in triangular form. By doing so, only $n(n+1)/2$ memory locations need be allocated for $\underline{S}(t_i^+)$, and the product $[\underline{\Phi}(t_{i+1},t_i)\underline{S}(t_i^+)]$ for the subsequent time propagation requires only half the usual number of computations.

Like the Potter measurement update, the Carlson algorithm processes vector measurements iteratively as scalars. Therefore, consider the general square root solution to (35):

$$\begin{aligned}\underline{S}(t_i^+) &= \underline{S}(t_i^-)[I-b(t_i)\underline{a}(t_i)\underline{a}^T(t_i)]^{1/2} \\ &= S(t_i^-)[I-\underline{a}(t_i)\underline{a}^T(t_i)/d(t_i)]^{1/2}\end{aligned} \tag{67}$$

where, for convenience, $d(t_i)$ has been defined as $1/b(t_i)$. Assuming $\underline{S}(t_i^-)$ to be upper triangular, we seek a matrix $[\underline{I}-\underline{a}(t_i)\underline{a}^T(t_i)/d(t_i)]^{1/2}$ such that the $\underline{S}(t_i^+)$ computed in (67) is also upper triangular. The choice between upper and lower triangular form is arbitrary, governed by selecting either forward or backward recursion algorithms for the Cholesky, Householder, and Gram-Schmidt procedures. Upper triangular forms are motivated to some extent by state vector partitioning, as discussed by Carlson.[12]

The desired square root matrix is in fact derived by means of an analytic Cholesky decomposition, and can be expressed as

$$\begin{bmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{bmatrix} - \begin{bmatrix} 0 & a_1 & a_1 & \cdots & a_1 \\ & 0 & a_2 & \cdots & a_2 \\ & & 0 & \ddots & \vdots \\ & & & \ddots & a_{n-1} \\ & & & & 0 \end{bmatrix} \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{bmatrix}$$

where a_k is the k-th component of $\underline{a}(t_i)$, and b_k and c_k for k=1,2,...,n will be described presently. However, the computational algorithm neither computes this square root explicitly nor requires a matrix multiplication as in (67) to generate $\underline{S}(t_i^+)$. The algorithm is initialized by setting the scalar d_o and n-vectors $\underline{e}_o$ and $\underline{a}$ as:

$$\boxed{\begin{aligned} d_o &= R(t_i) \\ \underline{e}_o &= \underline{0} \\ \underline{a} &= \underline{S}^T(t_i^-)\underline{H}^T(t_i) \end{aligned}} \tag{68}$$

and iterating for k=1,2,...,n on:

$$\boxed{\begin{aligned} d_k &= d_{k-1} + a_k^2 \\ b_k &= (d_{k-1}/d_k)^{1/2} \\ c_k &= a_k/(d_{k-1}d_k)^{1/2} \\ \underline{e}_k &= \underline{e}_{k-1} + \underline{S}_k^- a_k \\ \underline{S}_k^+ &= \underline{S}_k\, b_k - \underline{e}_{k-1}\, c_k \end{aligned}} \tag{69}$$

In the recursion, $\underline{S}_k^-$ denotes the k-th column of $\underline{S}(t_i^-)$, and both it and $\underline{e}_k$ consist of zeroes below the k-th element. After the n iterations, $\underline{S}(t_i^+)$ is produced as

$$\boxed{\underline{S}(t_i^+) = [\underline{S}_1^+ \;\; \underline{S}_2^+ \cdots \underline{S}_n^+]} \tag{70}$$

which is an upper triangular matrix. The state vector update is then given by

$$\boxed{\hat{\underline{x}}(t_i^+) = \hat{\underline{x}}(t_i^-) + \underline{e}_n \{[z_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^-)]/d_n\}} \tag{71}$$

For time propagations, Carlson suggested the matrix RSS method, (47), but with an upper triangular Cholesky square root as generated in (31) replacing the lower triangular form in (47). However, the triangularization methods could also be employed, thereby sacrificing some computational speed for increased numerical precision. A Modified Gram-Schmidt algorithm[32] can be written by initializing $\underline{A}$ according to (60) and then iterating for k=n,n-1,...,1 on:

$$\boxed{\begin{aligned} &S_{kk}(t_{i+1}^-) = \sqrt{\underline{A}_k^T \underline{A}_k} \\ &\underline{v}_k = \underline{A}_k / S_{kk}(t_{i+1}^-) \\ &\left.\begin{aligned} &S_{jk}(t_{i+1}^-) = \underline{A}_j^T \underline{v}_k \\ &\underline{A}_j \leftarrow \underline{A}_j - S_{jk}(t_{i+1}^-)\underline{v}_k \end{aligned}\right\} \; j=1,2,\ldots,(k-1) \end{aligned}} \tag{72}$$

where "$\leftarrow$" denotes replacement by means of "writing over" old variables.

7. Inverse Covariance Square Root Filter

The inverse covariance formulation[22,25] of the optimal filter is an algorithm which is algebraically equivalent to the Kalman filter but has substantially different characteristics, such as being able to incorporate unknown initial conditions and being more efficient if the measurement vector is very large in dimension (m>n). Now we consider the square root filter analog of such a formulation.

If we define the covariance square root matrix $\underline{S}(t_i^+)$ through

$$\underline{P}(t_i^+) \triangleq \underline{S}(t_i^+)\,\underline{S}^T(t_i^+) \tag{73}$$

then it is consistent that an inverse covariance square root $\underline{S}^{-1}(t_i^+)$ be defined through

$$\underline{P}^{-1}(t_i^+) \triangleq \underline{S}^{-T}(t_i^+)\,\underline{S}^{-1}(t_i^+) \tag{74a}$$

where S^{-T} denotes $[\underline{S}^{-1}]^T = [\underline{S}^T]^{-1}$. Similarly, $\underline{S}^{-1}(t_i^-)$ would be defined through

$$\underline{P}^{-1}(t_i^-) \triangleq \underline{S}^{-T}(t_i^-)\,\underline{S}^{-1}(t_i^-) \tag{74b}$$

To develop the inverse covariance square root filter,[5,22] first consider the measurement update equation in the inverse covariance filter:

$$\underline{P}^{-1}(t_i^+) = \underline{P}^{-1}(t_i^-) + \underline{H}^T(t_i)\underline{R}^{-1}(t_i)\underline{H}(t_i) \tag{75}$$

Using (29) and (74), this can be written as

$$\underline{P}^{-1}(t_i^+) = \underline{S}^{-T}(t_i^-)\underline{S}^{-1}(t_i^-) + \underline{H}^T(t_i)\underline{V}^{-T}(t_i)\underline{V}^{-1}(t_i)\underline{H}(t_i) \tag{76}$$

We now seek an update relation for $\underline{S}^{-1}(t_i^+)$ such that $\{[\underline{S}^{-1}(t_i^+)]^T[\underline{S}^{-1}(t_i^+)]\}$ is equivalent to the right hand side of (76). One such matrix would be the (n+m)-by-n matrix

$$\tilde{\underline{S}}^{-1}(t_i^+) = \begin{bmatrix} \underline{S}^{-1}(t_i^-) \\ \hline \underline{V}^{-1}(t_i)\,\underline{H}(t_i) \end{bmatrix} \tag{77}$$

As in the previous section, such an $\underset{\sim}{\tilde{S}}(t_i^+)$ would be unacceptable due to the increasing matrix dimensions it would cause. However, if an orthogonal matrix $\underset{\sim}{T}$ can be constructed such that

$$\underset{\sim}{T}^T \underset{\sim}{\tilde{S}}^{-1}(t_i^+) = \begin{bmatrix} \underset{\sim}{S}^{-1}(t_i^+) \\ \hline \underset{\sim}{0} \end{bmatrix} \begin{matrix} \} \text{ n rows} \\ \\ \} \text{ m rows} \end{matrix} \tag{78}$$

or

$$\underset{\sim}{\tilde{S}}^{-T}(t_i^+)\underset{\sim}{T} = [\underset{\sim}{S}^{-T}(t_i^+) \vdots \underset{\sim}{0}] \tag{78'}$$

then the resulting n-by-n $\underset{\sim}{S}^{-1}(t_i^+)$ is the desired square root matrix. In analogy to the previous development, it could be especially beneficial if the $S^{-1}(t_i^+)$ so generated were upper triangular. Either the Modified Gram-Schmidt orthogonalization procedure or the Householder transformation algorithm can be employed to solve for the desired $\underset{\sim}{S}^{-1}(t_i^+)$, and this will be developed in detail after the state estimate is discussed.

In its operation the inverse covariance filter does not compute a state estimate directly, but rather

$$\hat{\underset{\sim}{y}}(t_i^-) \triangleq \underset{\sim}{P}^{-1}(t_i^-)\, \hat{\underset{\sim}{x}}(t_i^-) \tag{79a}$$

$$\hat{\underset{\sim}{y}}(t_i^+) \triangleq \underset{\sim}{P}^{-1}(t_i^+)\, \hat{\underset{\sim}{x}}(t_i^+) \tag{79b}$$

which are related by

$$\hat{\underset{\sim}{y}}(t_i^+) = \hat{\underset{\sim}{y}}(t_i^-) + \underset{\sim}{H}^T(t_i)\underset{\sim}{R}^{-1}(t_i)\underset{\sim}{z}_i \tag{80}$$

Analogously, the inverse covariance square root filter does not generate an estimate of the state explicitly, but instead calculates

$$\hat{\underset{\sim}{a}}(t_i^-) \triangleq \underset{\sim}{S}^{-1}(t_i^-)\, \hat{\underset{\sim}{x}}(t_i^-) \tag{81a}$$

and

$$\hat{\underset{\sim}{a}}(t_i^+) \triangleq \underset{\sim}{S}^{-1}(t_i^+)\, \hat{\underset{\sim}{x}}(t_i^+) \tag{81b}$$

The update relationship between these estimates can be shown to be

$$\begin{matrix} n \text{ rows } \{ \\ m \text{ rows } \{ \end{matrix} \left[\begin{array}{c} \underline{\hat{\alpha}}(t_i^+) \\ \hline \underline{\beta}(t_i) \end{array}\right] = \underline{T}^T \left[\begin{array}{c} \underline{\hat{\alpha}}(t_i^-) \\ \hline \underline{V}^{-1}(t_i)\underline{z}_i \end{array}\right] \begin{matrix} \} \text{ n rows} \\ \} \text{ m rows} \end{matrix} \tag{82}$$

where $\underline{T}$ is the same orthogonal matrix as in (78), and $\underline{\beta}(t_i)$ is an m-dimensional vector (the residual after processing the measurement, $[\underline{z}_i - \underline{H}(t_i)\hat{\underline{x}}(t_i^+)]$) that need not be calculated. Since the first n rows of $\underline{T}^T$ are the result of an n-step Gram-Schmidt or Householder process, $\underline{\hat{\alpha}}(t_i^+)$ can be computed without knowledge of any additional portion of $\underline{T}^T$ than that generated by either of the triangularization algorithms discussed previously.

The Modified Gram-Schmidt (MGS) measurement update initializes an (n+m)-by-n matrix $\underline{A}^k$ and an n-vector $\underline{b}^k$ as

$$\underline{A}^1 = \tilde{\underline{S}}^{-1}(t_i^+) = \left[\begin{array}{c} \underline{S}^{-1}(t_i^-) \\ \hline \underline{V}^{-1}(t_i)\underline{H}(t_i) \end{array}\right] \tag{83a}$$

$$\underline{b}^1 = \underline{0} \tag{83b}$$

Then an n-step recursion is performed that is identical to (61) except for two additional equations for eventual generation of $\underline{\hat{\alpha}}(t_i^+)$; for k=1,2,...,n:

$$\begin{aligned} a^k &= \sqrt{\underline{A}_k^{k\,T}\,\underline{A}_k^k} \\ C_{kj} &= \begin{cases} 0 & j = 1,\ldots,k-1 \\ a^k & j = k \\ [(1/a^k)\,\underline{A}_k^{k\,T}]\underline{A}_j^k & j = k+1,\ldots,n \end{cases} \\ e^k &= [(1/a^k)\,\underline{A}_k^{k\,T}]\underline{b}^k \\ \underline{A}_j^{k+1} &= \underline{A}_j^k - C_{kj}[(1/a^k)\underline{A}_k^k] \qquad j = k+1,\ldots,n \\ \underline{b}^{k+1} &= \underline{b}^k - e^k\,[(1/a^k)\underline{A}_k^k] \end{aligned} \tag{84}$$

At the end of this recursion,

$$\boxed{\begin{aligned} \underline{S}^{-1}(t_i^+) &= \underline{C} \\ \underline{\hat{\alpha}}(t_i^+) &= \underline{b}^{n+1} \end{aligned}} \tag{85}$$

A Householder measurement update[17] can also be employed. The $(n+m)$-by-n matrix $\underline{A}^k$ and $(n+m)$-vector $\underline{b}^k$ (note the different dimension on $\underline{b}^k$) are initialized as in (83). Subsequently an n-step recursion identical in form to (64) except for auxiliary steps to calculate $\underline{\hat{\alpha}}(t_i^+)$ is performed[22]: for $k = 1,2,\ldots,n$,

$$\boxed{\begin{aligned} a^k &= \sqrt{\sum_{j=k}^{n+m} [A_{jk}^k]^2} \cdot \operatorname{sgn}\{A_{kk}^k\} \\ d^k &= \frac{1}{a^k(a^k + A_{kk}^k)} \\ u_j^k &= \begin{cases} 0 & j < k \\ a^k + A_{kk}^k & j = k \\ A_{jk}^k & j = k+1,\ldots,(n+m) \end{cases} \\ y_j^k &= \begin{cases} 0 & j < k \\ 1 & j = k \\ d^k \underline{u}^{kT} \underline{A}_j^k & j = k+1,\ldots,n \end{cases} \\ e^k &= a^k \underline{u}^{kT} \underline{b}^k \\ \underline{A}^{k+1} &= \underline{A}^k - \underline{u}^k \underline{y}^{kT} \\ \underline{b}^{k+1} &= \underline{b}^k - \underline{u}^k e^k \end{aligned}} \tag{86}$$

After the n iterations of (86), $\underline{S}^{-1}(t_i^+)$ and $\underline{\hat{\alpha}}(t_i^+)$ are obtained from:

$$\underline{A}^{n+1} = \begin{bmatrix} \underline{S}^{-1}(t_i^+) \\ \hline \underline{0} \end{bmatrix}$$

$$\underline{b}^{n+1} \quad \begin{bmatrix} \underline{\alpha}(t_i^+) \\ \hline \underline{\beta}(t_i) \end{bmatrix} \tag{87}$$

For time propagations in which the dynamic driving noise is s-dimensional, s scalar recursions analogous to the Potter measurement update in the covariance square root filter are performed. Thus $\underline{Q}_d(t_i)$ is assumed diagonal, perhaps after a change of variables, and the effects of $\underline{w}_d(\cdot,\cdot)$ are incorporated component by component. Letting $\underline{G}_{dk}$ be the k-th column of $\underline{G}_d(t_i)$ and Q_{dk} be the k-th diagonal element of $\underline{Q}_d(t_i)$, the algorithm becomes, for k = 1,2,...,s:

$$\begin{aligned}
\underline{a}(t_i) &= \underline{S}^{-1}(t_i^+)\ \underline{\Phi}(t_i,t_{i+1})\underline{G}_{dk} \\
b(t_i) &= 1/[\underline{a}^T(t_i)\underline{a}(t_i) + \{1/Q_{dk}\}] \\
\gamma(t_i) &= 1/[1 + \sqrt{b(t_i)\{1/Q_{dk}\}}] \\
\underline{\ell}^T(t_i) &= b(t_i)\underline{a}^T(t_i)\underline{S}^{-1}(t_i^+)\ \underline{\Phi}(t_i,t_{i+1}) \\
\underline{\alpha}(t_{i+1}^-) &= \underline{\alpha}(t_i^+) - b(t_i)\gamma(t_i)\underline{a}(t_i)\underline{a}^T(t_i)\underline{\alpha}(t_i^+) \\
\underline{S}^{-1}(t_{i+1}^-) &= \underline{S}^{-1}(t_i^+)\ \underline{\Phi}(t_i,t_{i+1}) - \gamma(t_i)\underline{a}(t_i)\underline{\ell}^T(t_i)
\end{aligned} \tag{88}$$

Note the order of the time indices on the state transition matrix and that $\underline{\Phi}(t_i,t_{i+1}) = \underline{\Phi}^{-1}(t_{i+1},t_i)$. After the first of the s iterations of (88), $\underline{\Phi}(t_i,t_{i+1})$ is replaced by the identity matrix. In analogy to the covariance square root filter, a Householder transformation has also been proposed for performing the time propagation, but it has been shown to be equivalent to, but less efficient than, the Potter type algorithm given in (88).

Thus, in the inverse covariance square root filter, measurement updates are conducted in vector form through a triangularization procedure, and time propagations involve iterative applications of a Potter-type scalar incorporation algorithm. This is in direct opposition to the covariance square root filter. As a result, its time propagations are more efficient than those of the covariance square root filter for the typical case in which the state dimension is much greater than the dynamic noise dimension s. On the other hand, its measurement update is more efficient only when the measurement dimension m is considerably greater than n. Although most applications have shown the covariance square root filter to be more efficient computationally, there are circumstances ($m>>n>>s$) under which the inverse covariance square root formulation is preferable. Section 9 will compare the various forms explicitly.

8. U-D Covariance Factorization Filter [1,6,7,12-15,32-37]

Another approach to enhancing the numerical characteristics of the optimal filter algorithm is known as "U-D covariance factorization." Rather than decomposing the covariance into its square root factors as in (26) and (27) this method expresses the covariances before and after measurement incorporation as

$$\underline{P}(t_i^-) = \underline{U}(t_i^-)\underline{D}(t_i^-)\underline{U}^T(t_i^-) \tag{89}$$

$$\underline{P}(t_i^+) = \underline{U}(t_i^+)\underline{D}(t_i^+)\underline{U}^T(t_i^+) \tag{90}$$

where the $\underline{U}$ matrices are upper triangular and unitary (i.e., ones along the diagonal) and the $\underline{D}$ matrices are diagonal. Although covariance square roots are never explicitly evaluated in this method, this filter algorithm is included in this chapter because: (1) $\underline{U}\underline{D}^{1/2}$ corresponds directly to the covariance square root of the Carlson filter in Section 6, and the Carlson filter in fact partially motivated this filter development, and (2) the U-D covariance factorization filter shares the advantages of the square root filters discussed previously: guaranteeing nonnegativity of the computed covariance and being numerically accurate and stable. (Merely being a square root filter is not a sufficient condition for numerical accuracy and stability, but the algorithms discussed previously do have these attributes.) Like the Carlson filter, triangular forms are maintained so that this algorithm is considerably more efficient in terms of computations and storage than the Potter filter. Though similar in concept and computation to the Carlson filter, this algorithm does not require any of the (nm+s) computationally expensive scalar square roots as processed in the former.

Before considering the filter algorithm itself, let us demonstrate that, given some $\underline{P}$ as an n-by-n symmetric, positive semidefinite matrix, a unit upper triangular factor $\underline{U}$ and diagonal factor $\underline{D}$ such that $\underline{P} = \underline{U}\underline{D}\underline{U}^T$ can always be generated. Although such $\underline{U}$ and $\underline{D}$ matrices are not unique, a uniquely defined pair can in fact be generated through an algorithm closely related to the backward running Cholesky decomposition algorithm, (31). This will be shown by explicitly displaying the result. First, for the n-th column

$$\left.\begin{aligned} D_{nn} &= P_{nn} \\ U_{in} &= \begin{cases} 1 & i = n \\ P_{in}/D_{nn} & i = n-1, n-2, \ldots 1 \end{cases} \end{aligned}\right\} \quad (91a)$$

Then for the remaining columns, for j = n-1,n-2,...,1, compute:

$$\left.\begin{aligned} D_{jj} &= P_{jj} - \sum_{k=j+1}^{n} D_{kk}\, U_{jk}^2 \\ U_{ij} &= \begin{cases} 0 & i > j \\ 1 & i = j \\ [P_{ij} - \sum_{k=j+1}^{n} D_{kk}\, U_{ik}\, U_{jk}]/D_{jj} & i = j-1, j-2, \ldots, 1 \end{cases} \end{aligned}\right\} \quad (91b)$$

This is useful for defining the required U-D factors of $\underline{P}_o$ and the $\underline{Q}_d$ time history for a given application.

To develop the filter algorithm itself, first consider a scalar measurement update, for which $H(t_i)$ is 1-by-n. For convenience, we drop the time index and let $\underline{P}(t_i^-) = \underline{P}^-$, $\underline{P}(t_i^+) = \underline{P}^+$, and so forth. The Kalman update

$$\begin{aligned} \underline{P}^+ &= \underline{P}^- - (\underline{P}^-\underline{H}^T)\,\frac{1}{a}\,(\underline{H}\underline{P}^-) \\ a &= \underline{H}\underline{P}^-\underline{H}^T + R \end{aligned} \quad (92)$$

can be factored as

$$\begin{aligned} \underline{U}^+\underline{D}^+\underline{U}^{+T} &= \underline{U}^-\underline{D}^-\underline{U}^{-T} - \frac{1}{a}\,(\underline{U}^-\underline{D}^-\underline{U}^{-T}\underline{H}^T)\,\underline{H}\,\underline{U}^-\underline{D}^-\underline{U}^{-T} \\ &= \underline{U}^-\,[\underline{D}^- - \frac{1}{a}\,(\underline{D}^-\underline{U}^{-T}\underline{H}^T)(\underline{D}^-\underline{U}^{-T}\underline{H}^T)^T]\underline{U}^{-T} \end{aligned} \quad (93)$$

Defining the n-vectors $\underline{f}$ and $\underline{v}$ as

$$\underline{f} = \underline{U}^{-T}\underline{H}^{T} \tag{94a}$$

$$\underline{v} = \underline{D}^{-}\underline{f} \quad \text{i.e., } v_j = D^{-}_{jj} f_j \text{ , } j = 1,2,\ldots,n \tag{94b}$$

and substituting into (93) yields

$$\underline{U}^{+}\underline{D}^{+}\underline{U}^{+T} = \underline{U}^{-}[\underline{D}^{-} - \frac{1}{a}\underline{v}\,\underline{v}^{T}]\underline{U}^{-T} \tag{95}$$

Now let $\overline{\underline{U}}$ and $\overline{\underline{D}}$ be the U-D factors of $[\underline{D}^{-} - \frac{1}{a}\underline{v}\,\underline{v}^{T}]$:

$$\overline{\underline{U}}\,\overline{\underline{D}}\,\overline{\underline{U}}^{T} = [\underline{D}^{-} - \frac{1}{a}\underline{v}\,\underline{v}^{T}] \tag{96}$$

so that (95) can be written as

$$\underline{U}^{+}\underline{D}^{+}\underline{U}^{+T} = [\underline{U}^{-}\overline{\underline{U}}\,]\overline{\underline{D}}\,[\underline{U}^{-}\overline{\underline{U}}\,]^{T} \tag{97}$$

Since $\underline{U}^{-}$ and $\overline{\underline{U}}$ are unit upper triangular, this then yields

$$\underline{U}^{+} = \underline{U}^{-}\overline{\underline{U}} \tag{98a}$$

$$\underline{D}^{+} = \overline{\underline{D}} \tag{98b}$$

In this manner, the problem of factoring the Kalman filter measurement update has been reduced to the problem of factoring a symmetric matrix, $[\underline{D} - \frac{1}{a}\underline{v}\,\underline{v}^{T}]$ into $\overline{\underline{U}}$ and $\overline{\underline{D}} = \underline{D}^{+}$. These factors can be generated recursively by letting $a_o = R$ and computing, for j=1,2,...,n

$$\left.\begin{aligned} a_j &= \sum_{k=1}^{j} D_{kk} f_j^{\,2} + R \\ \overline{D}_{jj} &= D_{jj}\, a_{j-1}/a_j \\ \overline{U}_{ij} &= \begin{cases} -D_{ii} f_i f_j/a_{j-1} & i = 1,2,\ldots,j-1 \\ 1 & i = j \\ 0 & i = j+1,\ j+2,\ldots,n \end{cases} \end{aligned}\right\} \tag{99}$$

Thus, $[\underline{D} - \frac{1}{a}\underline{v}\,\underline{v}^T]$ is scanned and $\underline{U}$ is generated column by column, as depicted in Figure 3. The validity of the terms generated in (99) can be demonstrated by substituting them into (91) and showing that the resulting $[P_{ij}]$ matrix is in fact $[\underline{D} - \frac{1}{a}\underline{v}\,\underline{v}^T]$.

The scalar measurement update for the U-D covariance factorization filter can now be specified. At time t_i, $\underline{U}(t_i^-)$ and $\underline{D}(t_i^-)$ are available from a previous time propagation (to be discussed). Using the measurement value z_i and the known 1-by-n $\underline{H}(t_i)$ and scalar $R(t_i)$, one computes

$$\boxed{\begin{aligned} &\underline{f} = \underline{U}^T(t_i^-)\underline{H}^T(t_i) \\ &v_j = D_{jj}(t_i^-)f_j \qquad\qquad j = 1,2,\ldots,n \\ &a_o = R \end{aligned}} \tag{100}$$

Then, for k = 1,2,...,n, calculate the results of (99), but in a more efficient manner as

$$\boxed{\begin{aligned} &a_k = a_{k-1} + f_k v_k \\ &D_{kk}(t_i^+) = D_{kk}(t_i^-)\, a_{k-1}/a_k \\ &b_k \leftarrow v_k \\ &p_k = -f_k/a_{k-1} \\ &\left.\begin{aligned} &U_{jk}(t_i^+) = U_{jk}(t_i^-) + b_j\, p_k \\ &b_j \leftarrow b_j + U_{jk}(t_i^-)v_k \end{aligned}\right\} \; j = 1,2,\ldots,(k-1) \end{aligned}} \tag{101}$$

In (101), "$\leftarrow$" denotes replacement, exploiting the technique of "writing over" old variables for efficiency. For k=1, only the first three equations need be processed. After the n iterations of (101), $\underline{U}(t_i^+)$ and $\underline{D}(t_i^+)$ have been computed, and the filter gain $\underline{K}(t_i)$ can be calculated in terms of the n-vector $\underline{b}$ made up of components $b_1, b_2, \ldots, b_n$ computed in the last

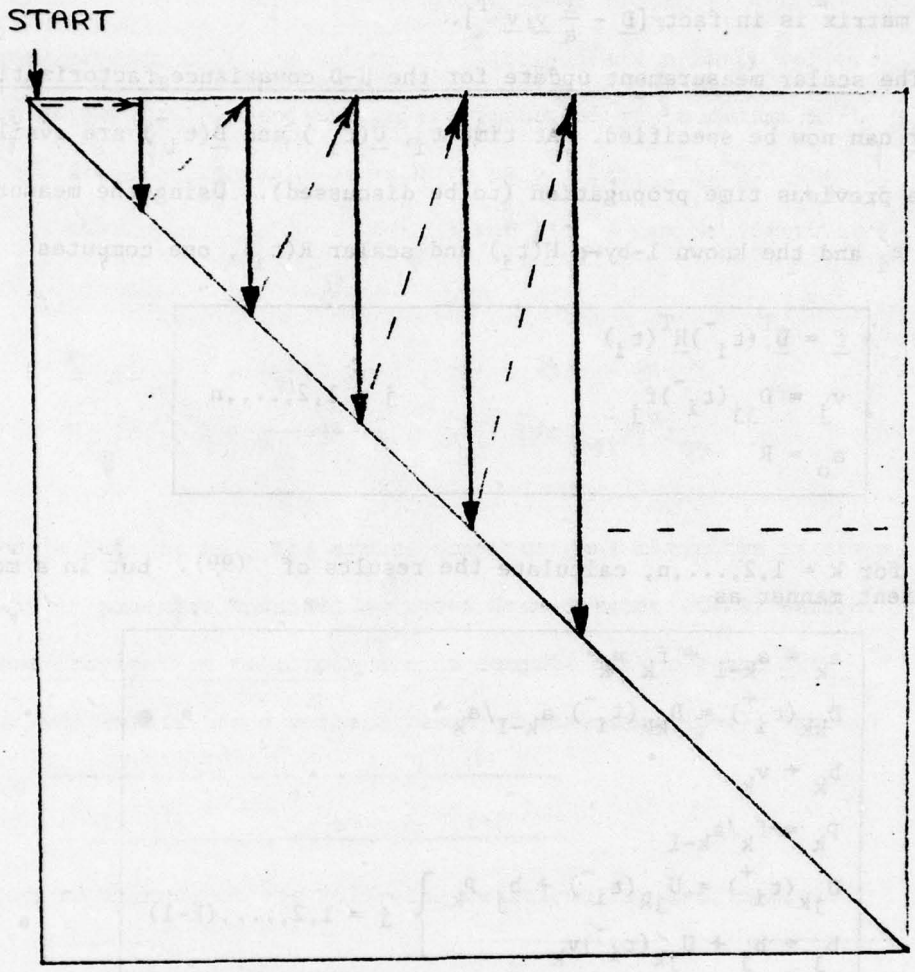

Fig 3 Scanning of $[\underline{D} - \frac{1}{a}\underline{v}\underline{v}^T]$ and Generation of $\bar{\underline{U}}$

iteration of (101), and the state updated as

$$\boxed{\begin{aligned} \underline{K}(t_i) &= \underline{b}/a_n \\ \underline{\hat{x}}(t_i^+) &= \underline{\hat{x}}(t_i^-) + \underline{K}(t_i)\,[z_i - \underline{H}(t_i)\underline{\hat{x}}(t_i^-)] \end{aligned}} \tag{102}$$

Vector measurement updates would be performed component by component, requiring a transformation of variables as in Section 5 if $\underline{R}(t_i)$ is not originally diagonal.

The time propagation of the U-D factors employs a generalized Gram-Schmidt orthogonalization to preserve numerical accuracy while attaining computational efficiency.[52] Given the covariance time propagation relation

$$\underline{P}(t_{i+1}^-) = \underline{\Phi}(t_{i+1},t_i)\underline{P}(t_i^+)\underline{\Phi}^T(t_{i+1},t_i) + \underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i) \tag{103}$$

and the U-D factors of $\underline{P}(t_i^+)$, we desire the factors $\underline{U}(t_{i+1}^-)$ and $\underline{D}(t_{i+1}^-)$ such that $[\underline{U}(t_{i+1}^-)\underline{D}(t_{i+1}^-)\underline{U}^T(t_{i+1}^-)]$ equals the right hand side of (103). Without loss of generality, $\underline{Q}_d(t_i)$ is assumed diagonal, since, given the n-by-n matrix $[\underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i)]$, (91) can be used to generate $\underline{G}_d(t_i)$ as its U-factor and $\underline{Q}_d(t_i)$ as its D-factor.

If an n-by-(n+s) matrix $\underline{Y}(t_{i+1}^-)$ and an (n+s)-by-(n+s) diagonal matrix $\tilde{\underline{D}}(t_{i+1}^-)$ are defined as

$$\underline{Y}(t_{i+1}^-) = [\underline{\Phi}(t_{i+1},t_i)\underline{U}(t_i^+) \vdots \underline{G}_d(t_i)] \tag{104a}$$

$$\tilde{\underline{D}}(t_{i+1}^-) = \left[\begin{array}{c|c} \underline{D}(t_i^+) & \underline{0} \\ \hline \underline{0} & \underline{Q}_d(t_i) \end{array}\right] \tag{104b}$$

then it can be seen that $[\underline{Y}(t_{i+1}^-)\tilde{\underline{D}}(t_{i+1}^-)\underline{Y}^T(t_{i+1}^-)]$ satisfies (103).

Similar to the development of (53) to (79) of Section 6, the desired result can be generated through a Gram-Schmidt procedure applied to $\underline{Y}(t_{i+1}^-)$. The only significant modification is that the inner products used in the procedure are weighted inner products: whereas in (55) the inner product of $\tilde{\underline{s}}^j$ (a column of $\tilde{\underline{S}}^T(t_{i+1}^-)$) and a basis vector $\underline{b}^k$ was written as $[\underline{s}^{j\,T}\,\underline{b}^k]$, here the inner product of $\underline{y}^j$ (a column of $\underline{Y}^T(t_{i+1}^-)$) and a basis vector $\underline{b}^k$ would be written as $[\underline{y}^{jT}\,\underline{\tilde{D}}(t_{i+1}^-)\underline{b}^k]$. When an analogous development is made, $\underline{D}(t_i^+)$ and $\underline{U}(t_i^+)$ can be identified as, for $j = 1,2,\ldots n$ and $k = j,j+1,\ldots,n$:

$$D_{jj}(t_{i+1}^-) = [\underline{b}^j]^T\,\underline{\tilde{D}}(t_{i+1}^-)\underline{b}^j \tag{105a}$$

$$U_{jk}(t_{i+1}^-) = \frac{1}{D_{kk}(t_{i+1}^-)}\,\{[\underline{y}^j]^T\,\underline{\tilde{D}}(t_{i+1}^-)\underline{b}^k\} \tag{105b}$$

As in Section 6, the actual computational algorithm is the efficient, numerically superior Modified Weighted Gram-Schmidt (MWGS) method. Thus, the <u>time propagation relations</u> are to compute $\underline{Y}(t_{i+1}^-)$ and $\underline{\tilde{D}}(t_{i+1}^-)$ as in (104), and initialize n vectors, each of dimension (n+s), through

$$\boxed{[\underline{a}_1 \,\vdots\, \underline{a}_2 \,\vdots\, \text{---} \,\vdots\, \underline{a}_n] = \underline{Y}^T(t_{i+1}^-)} \tag{106}$$

and then to iterate on the following relations for k=n, n-1, ..., 1:

$$\boxed{\begin{aligned}
&\underline{c}_k = \underline{\tilde{D}}(t_{i+1}^-)\underline{a}_k \qquad (c_{kj} = \tilde{D}_{jj}(t_{i+1}^-)a_{kj};\ j=1,2,\ldots,n)\\
&D_{kk}(t_{i+1}^-) = \underline{a}_k^{\,T}\,\underline{c}_k\\
&\underline{d}_k = \underline{c}_k / D_{kk}(t_{i+1}^-)\\
&\left.\begin{aligned}
&U_{jk}(t_{i+1}^-) = \underline{a}_j^{\,T}\,\underline{d}_k\\
&\underline{a}_j \leftarrow \underline{a}_j - U_{jk}(t_{i+1}^-)\underline{a}_k
\end{aligned}\right\}\ j = 1,2,\ldots,k-1
\end{aligned}} \tag{107}$$

As before, "←" denotes replacement, or "writing over" old variables to reduce storage requirements. On the last iteration, for k=1, only the first two relations need be computed. The state estimate is given by

$$\boxed{\underline{\hat{x}}(t_{i+1}^{-}) = \underline{\Phi}(t_{i+1},t_i)\underline{\hat{x}}(t_i^{+})} \tag{108}$$

9. Filter Performance and Requirements

The algorithms of this chapter have been investigated in order to implement the optimal filtering solution given by the Kalman filter, but without the numerical instability and inaccuracy of that algorithm when processed with finite wordlength. In this section, both the numerical advantages and the increased computational burden of these filters will be delineated.

An algorithm can be said to be numerically stable if the computed result of the algorithm corresponds to an exactly computed solution to a problem that is only slightly perturbed from the original one.[35] By this criterion, the Kalman filter is numerically unstable,[6] in both the conventional and Joseph formulations. In contrast, all of the filters described in this chapter can be shown to be numerically stable.

The numerical conditioning of a set of computations can be described in part by what is called a "condition number", a concept which is often used to analyze the effects of perturbations in linear equations. If $\underline{A}$ is a matrix, not necessarily square, then the condition number $k(\underline{A})$ associated with $\underline{A}$ is defined by[22]

$$k(\underline{A}) = \sigma_{max}/\sigma_{min} \tag{109}$$

where σ_{max}^2 and σ_{min}^2 are the maximum and minimum eigenvalues of $\underline{A}^T\underline{A}$ respectively. When computing in base 10 (or base 2) arithmetic with N significant digits (or bits), numerical difficulties may be expected as $k(\underline{A})$ approaches 10^N (or 2^N). For instance, if the maximum and minimum numbers of interest, σ_{max} and σ_{min}, were 100000 and 000001 (in base 10 or 2), then to add these values together and obtain 100001 without numerical difficulties would require at least 6 significant figures (digits

or bits). But,

$$k(\underline{P}) = k(\underline{S}\underline{S}^T) = [k(\underline{S})]^2 \tag{110}$$

Therefore, while numerical operations on the covariance $\underline{P}$ may encounter difficulties when $k(\underline{P}) = 10^N$ (or 2^N), those same numerical problems would arise when $k(\underline{S}) = 10^{N/2}$ (or $2^{N/2}$) according to (110): the same numerical precision is achieved with half the wordlength.

EXAMPLE

This example and the next illustrate the improved numerical conditioning of the square root filters. To simulate roundoff, let $e << 1$ be such that

$$1 + e \overset{r}{\neq} 1$$

$$1 + e^2 \overset{r}{=} 1$$

where "$\overset{r}{=}$" means equal due to rounding. Consider a scalar measurement update of a two-state problem, with

$$\underline{P}(t_i^-) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \underline{H}(t_i) = [1 \; 0], \quad R(t_i) = e^2$$

and compare the computed results of conventional filters and those of this report. Note that

$$\underline{P}(t_i^-) = \underline{P}^{-1}(t_i^-) = \underline{S}(t_i^-) = \underline{S}^{-1}(t_i^-) = \underline{U}(t_i^-) = \underline{D}(t_i^-) = \underline{I}$$

and that the exact covariance $\underline{P}(t_i^+)$ for this example is:

$$\underline{P}(t_i^+) = \begin{bmatrix} e^2/(1+e^2) & 0 \\ 0 & 1 \end{bmatrix}$$

The computed results are:

(a) Conventional Kalman $\quad \underline{P}(t_i^+) \overset{r}{=} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

(b) Joseph Form Kalman $\quad \underline{P}(t_i^+) \overset{r}{=} \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$

(c) Potter Covariance Square Root $\quad \underline{S}(t_i^+) \overset{r}{=} \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$

(d) Carlson Covariance Square Root $\quad \underline{S}(t_i^+) \overset{r}{=} \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$

(e) Inverse Covariance $\quad \underline{P}^{-1}(t_i^+) \overset{r}{=} \begin{bmatrix} 1/e^2 & 0 \\ 0 & 1 \end{bmatrix}$

(f) Inverse Covariance Square Root $\quad \underline{S}^{-1}(t_i^+) \overset{r}{=} \begin{bmatrix} 1/e & 0 \\ 0 & 1 \end{bmatrix}$

(g) U-D Factor $\quad \underline{U}(t_i^+) \overset{r}{=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$\underline{D}(t_i^+) \overset{r}{=} \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$$

For this example, all but the conventional Kalman filter yield non-singular and nearly exact answers. Although the difference between 0 and e^2 in the upper left element of $\underline{P}(t_i^+)$ may seem insignificant, it can have grave consequences. For instance, assume no dynamics and let a second measurement of the same form be taken. The gain $\underline{K}$ computed by the conventional Kalman filter would be

$$\underline{K}(t_i^{++}) = \underline{P}(t_i^+)\,\underline{H}(t_i)/\,[\underline{H}(t_i)\underline{P}(t_i^+)\underline{H}^T(t_i) + R(t_i)]$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} / [0 + 1] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

whereas the correct value is

$$\underline{K}(t_i^{++}) = \frac{e^2}{1 + e^2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Bigg/ \left\{ \frac{e^2}{1 + e^2} + e^2 \right\} \cong \tfrac{1}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

as would be calculated correctly by the Joseph form in this case. ■

EXAMPLE

Consider the same problem as in the previous example, but let $\underline{H}(t_i)$ be [1 1] instead of [1 0]. In this case, the exact answer is

$$\underline{P}(t_i^+) = \frac{1}{2 + e^2} \begin{bmatrix} 1 + e^2 & -1 \\ -1 & 1 + e^2 \end{bmatrix}$$

The computed results are:

(a) Conventional Kalman $\quad \underline{P}(t_i^+) \overset{r}{=} \tfrac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

(b) Joseph Form Kalman $\quad \underline{P}(t_i^+) \overset{r}{=} \tfrac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

(c) Potter Covariance Square Root

$$\underline{S}(t_i^+) \triangleq \begin{bmatrix} 1 + e/\sqrt{2} & 0 \\ -1 + e/\sqrt{2} & 1 + e/\sqrt{2} \end{bmatrix}$$

(d) Carlson Covariance Square Root

$$\underline{S}(t_i^+) \triangleq \begin{bmatrix} e & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

(e) Inverse Covariance

$$\underline{P}^{-1}(t_i^+) \triangleq \frac{1}{e^2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

(f) Inverse Covariance Square Root

$$\underline{S}^{-1}(t_i^+) \triangleq \frac{1}{e} \begin{bmatrix} -1 & -1 \\ 0 & e\sqrt{2} \end{bmatrix}$$

(g) U-D Factor

$$\underline{U}(t_i^+) \triangleq \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$\underline{D}(t_i^+) \triangleq \begin{bmatrix} e^2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

In this case, only the square root and U-D implementations yield non-singular results. Such singular $\underline{P}(t_i^+)$ or $\underline{P}^{-1}(t_i^+)$ matrices would again yield a zero gain $\underline{K}$ if a second measurement of the same form were processed, while the square root and U-D factor filters compute a gain which is nearly exact. Moreover, even though $\underline{S}, \underline{S}^{-1}$, $\underline{U}$, and $\underline{D}$ are non-singular, the associated value of $\underline{P}(t_i^+)$ or $\underline{P}^{-1}(t_i^+)$ found by multiplication would be rounded to a singular matrix; thus it is better not to perform such computations explicitly, and the time propagations based on triangularization are to be preferred over the RSS method which performs such multiplication. ■

The improved numerical characteristics of the square root and U-D factorization filters are achieved at the expense of increased computational burden. Letting n be the dimension of the state vector $\underset{\sim}{x}$, s be the dimension of the dynamic driving noise $\underset{\sim}{w}$, and m be the dimension of the measurement $\underset{\sim}{z}$ and its corruptive noise $\underset{\sim}{v}$, we now determine the number of mathematical operations required by the various filters, assuming that:

1) all implementations take advantage of symmetry and zeros as they appear in general forms

2) $\underline{R}(t_i)$ and $\underline{Q}_d(t_i)$ are diagonal

3) the inverse covariance and inverse covariance square root filters generate explicit state estimates, $\hat{\underline{x}}(t_i^-)$ and $\hat{\underline{x}}(t_i^+)$.

Table 1 presents the number of operations for one time propagation and one measurement update required by

1. Kalman Filter - with conventional and Joseph form measurement update
2. Potter Covariance Square Root Filter - with MGS and Householder time propagations
3. Carlson Covariance Square Root Filter - with Matrix RSS and MGS time propagations
4. Inverse Covariance Filter
5. Inverse Covariance Square Root filter (MGS update)
6. U-D Covariance Factorization Filter.

This table can be used to project the computation time required by each filter formulation for a given application. Note that if, instead of assuming $\underline{Q}_d(t_i)$ to be diagonal, we were to assume that the n-by-n $[\underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i)]$ or the n-by-s $[\underline{G}_d(t_i)\underline{W}_d(t_i)]$ were known, there would be $\frac{1}{2}$ ns (n + 3) fewer multiplies and $\frac{1}{2}$ n (n + 1)(s - 1) fewer adds in filter forms 1, 3 with RSS time propagation, and 4, or ns fewer multiplies in forms 2 and 3 with MGS time propagation.

EXAMPLE

To put the algebraic expressions of Table 1 into perspective, Table 2 presents the number of operations required for one time propagation and one measurement update for the case of n = 10, s = 10, and m = 2. The noise dimension s was intentionally set equal to n to correspond to the n-by-n $[\underline{G}_d(t_i)\underline{Q}_d(t_i)\underline{G}_d^T(t_i)]$ being of full rank, typical of an equivalent discrete-time model. The last column in Table 2 portrays computer time required for one total filter recursion, neglecting the computations associated with the various subscripting and storage operations for each filter (roughly the same for each), and using single precision instruction times typical of the IBM 360 and some smaller state-of-the-art computers:

time for addition = 2.7 μsec
time for multiplication = 4.1 μsec

Table 1 Operations Required for One Time Propagation and One Measurement Update

FILTER	ADDS (all times $1/6$)	MULTIPLIES (all times $1/6$)	DIVIDES	SQUARE ROOTS
Conventional Kalman	$9n^3+3n^2(3m+s-1)+n(15m+3s-6)$	$9n^3+3n^2(3m+s+3)+n(27m+9s)$	m	0
Joseph Form Kalman	$18n^3+3n^2(5m+s-10)+n(9m^2+6m+3s)+3m^3-6m^2+3m$	$18n^3+3n^2(5m+s+4)+n(9m^2+24m+9s)+3m^3+9m^2-6m$	2m-1	0
Potter Cov $\sqrt{}$ (MGS)	$12n^3+3n^2(6m+2s)+n(6m-6)+6m$	$12n^3+3n^2(6m+2s+2)+n(24m+6s)+12m$	n+2m	n+m
Potter Cov $\sqrt{}$ (Householder)	$10n^3+3n^2(6m+2s-1)+n(6m+5)+6m$	$10n^3+3n^2(6m+2s+2)+n(24m+6s+8)+12m$	n+2m	n+m
Carlson Cov $\sqrt{}$ (RSS)	$5n^3+3n^2(3m+s+1)+n(9m+3s-14)+2s^3+4s$	$5n^3+3n^2(4m+s+3)+n1(30m+9s-2)+2s^3+6s^2-2s$	2mn+s	mn+s
Carlson Cov $\sqrt{}$ (MGS)	$9n^3+3n^2(3m+s-1)+3n(3m+3s-8)+2s^3+6s^2+4s$	$9n^3+3n^2(4m+s+2)+3n(10m+5s-7)+2s^3+12s^2+4s$	2mn+s	mn+s
Inverse Covariance	$10n^3+3n^2(m+3s+2)+n(9m+9s-16)$	$10n^3+3n^2(m+3s+6)+n(15m+21s-10)$	2s-1	0
Inverse Cov $\sqrt{}$	$9n^3+3n^2(2m+6s+5)+n(12m+6s-6)$	$9n^3+3n^2(2m+6s+6)+n(12m+24s+3)+6s$	2n+2s	n+s
U-D Factor	$9n^3+3n^2(3m+2s+2)+3n(3m+1)$	$9n^3+3n^2(3m-2s+7)+3n(m+4s-4)-6s$	n(m+1)-1	0

Table 2 Operations for One Total Filter Recursion; n = s = 10 and m = 2

FILTER	ADDS	MULTIPLIES	DIVIDES	SQUARE ROOTS	TIME (Millisec)
Conventional Kalman	2340	2690	2	0	17.36
Joseph Form Kalman	3631	4498	3	0	28.27
Potter Cov $\sqrt{}$ (MGS)	3612	3884	14	12	26.49
Potter Cov $\sqrt{}$ (Householder)	3247	3564	14	12	24.19
Carlson Cov $\sqrt{}$ (RSS)	2080	2560	50	30	18.24
Carlson Cov $\sqrt{}$ (MGS)	2830	3355	50	30	23.53
Inverse Covariance	3520	3950	19	0	25.82
Inverse Cov $\sqrt{}$	5080	5455	40	20	37.55
U-D Factor	2935	3330	29	0	21.77

time for division = 6.6 μsec
time for square root = 60.0 μsec

As can be seen from Table 2, the covariance square root filters and the U-D covariance factorization filter involve a computational load greater than the conventional Kalman filter, but not so great as to be prohibitive. In fact, the increase is less than that caused by employing the Joseph form of the update equation, which is inferior to these filters in performance. Moreover, since the Kalman filter would probably require double precision operations instead of the single precision assumed to establish Table 2, these filters are even more competitive with the Kalman filter than indicated in the table.

Of the square root type filters, the Carlson covariance square root and the U-D covariance factorization filters are the most efficient computationally. The Carlson filter with Matrix RSS time propagations requires the least computer time, but this is offset by the degraded numerical accuracy of the Matrix RSS method. Thus, the U-D covariance factorization filter would appear to be an exceptionally efficient and numerically advantageous alternative to the conventional Kalman filter for this particular application. ■

10. Conclusion

This report presented the concept of square root filters and the closely related U-D covariance factorization filter as viable alternatives to conventional Kalman filters. For a modest increase in computational loading, one obtains optimal filter algorithms equivalent to the Kalman filter if infinite wordlength is assumed, but with vastly superior numerical characteristics with finite wordlength. From a numerical analysis standpoint, this is at least as good a solution to troublesome measurement update computations as implementing a Kalman filter in double precision, since the Kalman filter inherently involves unstable numerics.

Of the covariance square root forms, the Carlson filter is more efficient than the Potter form computationally, and it also maintains triangularity of the square root matrices. The U-D covariance factorization filter is comparable to the Carlson filter and does not require square root computations. In comparison, the inverse covariance square root filter is often considerably more burdensome computationally, although it too becomes competitive if the measurement dimension m is very large.

Chandrasekhar type square root algorithms have also been reported in the literature. However, these have been omitted because they do not appear to be computationally competitive with algorithms presented herein for the nonstationary linear discrete-time estimation problem.

BIBLIOGRAPHY

1. Agee, W. S., and R. H. Turner, "Triangular Decomposition of a Positive Definite Matrix Plus a Symmetric Dyad with Applications to Kalman Filtering," Mathematical Services Branch, Analysis and Computation Division, Tech. Report 38, White Sands Missile Range, 1972.

2. Andrews, A., "A Square Root Formulation of the Kalman Covariance Equations," AIAA Journal, Vol. 6, No. 6, June 1968, pp. 1165-1166.

3. Bellantoni, J. F., and K. W. Dodge, "A Square Root Formulation of the Kalman - Schmidt Filter," AIAA Journal, Vol. 5, No. 7, July 1967, pp. 1309-1314.

4. Bierman, G. J., "A Comparison of Discrete Linear Filtering Algorithms," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-9, No. 1, Jan 1973, pp 28-37.

5. Bierman, G. J., "Sequential Square Root Filtering and Smoothing of Discrete Linear Systems," Automatica, Vol. 10, 1974, pp. 147-158.

6. Bierman, G. J., "Measurement Updating Using the U-D Factorization," Proceedings of the 1975 IEEE Control and Decision Conference, Houston, Texas, Dec 1975, pp. 337-346.

7. Bierman, G. J., Factorization Methods for Discrete Sequential Estimation, Academic Press, New York, N. Y., 1976.

8. Björck, A., "Solving Linear Least Squares Problems by Gram-Schmidt Orthogonalization," BIT, Vol. 7, 1967, pp 1-21.

9. Businger, P., and G. H. Golub, "Linear Least Squares Solution by Householder Transformations," Numerische Mathematik, Vol. 7, 1965, pp. 269-276.

10. Carlson, N. A., "Fast Triangular Formulation of the Square Root Filter," AIAA Journal, Vol. 11, No. 9, Sept 1973, pp. 1259-1265.

11. Dyer, P., and S. McReynolds, "Extension of Square-Root Filtering to Include Process Noise," Journal of Optimization Theory and Applications, Vol. 3, No. 6, 1969, pp. 444-459.

12. Fadeeva, V. N., Computational Methods of Linear Algebra, Dover Publications, Inc., New York, N. Y., 1959.

13. Fletcher, R., and M. J. D. Powell, "On the Modification of LDL^T Factorizations," Mathematics of Computation, Vol. 28, No. 128, Oct 1974, pp. 1067-1087.

14. Gentleman, W. M., "Least Squares Computations by Givens Transformations without Square Roots," J. Inst. Math. Applic., Vol. 12, 1973, pp. 329-336.

15. Gill, P. E., G. H. Golub, W. Murray, and M. A. Saunders, "Methods for Modifying Matrix Factorizations," Mathematics of Computation, Vol. 28, No. 126, 1974, pp. 505-535.

16. Gill, P. E., W. Murray, and M. A. Saunders, "Methods for Computing and Modifying the LDV Factors of a Matrix," Mathematics of Computation, Vol. 29, No. 132, Oct 1975, pp. 1051-1077.

17. Golub, G. H., "Numerical Methods for Solving Linear Least Squares Problems," Numerische Mathematik, Vol. 7, 1965, pp. 206-216.

18. Hanson, R. J., and C. L. Lawson, "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.

19. Householder, A. S., The Theory of Matrices in Numerical Analysis, Blaisdell Publishing Co., Waltham, Mass., 1964.

20. Jordan, T., "Experiments on Error Growth Associated with Some Linear Least Squares Procedures," Mathematics of Computation, Vol. 22, 1968, pp. 579-588.

21. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," Transactions of the ASME, Vol. 82D, Mar 1960, pp. 35-50.

22. Kaminski, P. G., A. E. Bryson Jr., and S. F. Schmidt, "Discrete Square Root Filtering: A Survey of Current Techniques," IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, Dec 1971, pp. 727-735.

23. Lawson, C. L., and R. J. Hanson, Solving Linear Least Squares Problems, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974.

24. Maybeck, P. S., "The Kalman Filter: An Introduction for Potential Users," Air Force Flight Dynamics Laboratory Report TM-72-3, Wright Patterson AFB, OH, June 1972.

25. Maybeck, P. S., Stochastic Estimation and Control Systems, Air Force Institute of Technology, Wright-Patterson AFB, OH, Feb 1975.

26. Morf, M., and T. Kailath, "Square-Root Algorithms for Least-Squares Estimation," IEEE Transactions on Automatic Control, Vol. AC-20, No. 4, Aug 1975, pp. 487-497.

27. Potter, J. E., "W Matrix Augmentation," MIT Instrumentation Laboratory Memo SGA 5-64, Jan 1964.

28. Practical Aspects of Kalman Filtering Implementation, NATO Advisory Group for Aerospace Research and Development, AGARD-LS-82, London, Mar 1976.

29. Rice, J. R., "Experiments on Gram-Schmidt Orthogonalization," Mathematics of Computation, Vol. 20, 1966, pp. 325-328.

30. Schmidt, S. F., "Computational Techniques in Kalman Filtering," Chapter 3 in Theory and Applications of Kalman Filtering, AGARDograph 139, London, Feb 1970.

31. Theory and Applications of Kalman Filtering, NATO Advisory Group for Aerospace Research and Development, AGARDograph 139, London, Feb 1970.

32. Thornton, C. L., and G. J. Bierman, "Gram-Schmidt Algorithms for Covariance Propagation," Proceedings of the 1975 IEEE Control and Decision Conference, Houston, Texas, Dec 1975, pp. 489-498.

33. Thornton, C. L., and G. J. Bierman, "A Numerical Comparison of Discrete Kalman Filtering Algorithms: An Orbit Determination Case Study," preprint of paper submitted for publication.

34. Wampler, R. H., "A Report on the Accuracy of Some Widely Used Least Squares Computer Programs," Journal of the American Statistical Association, Vol. 65, No. 330, 1970, pp. 549-565.

35. Wilkinson, J. H., Rounding Errors in Algebraic Processes, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) AFIT-TR-77-6	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) SOLUTIONS TO THE KALMAN FILTER WORDLENGTH PROBLEM: SQUARE ROOT AND U-D COVARIANCE FACTORIZATIONS	5. TYPE OF REPORT & PERIOD COVERED (9) Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Peter S. Maybeck, Associate Professor of Electrical Engineering	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT-ENG) Wright-Patterson AFB, OH 45433	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Avionics Laboratory (11) Wright-Patterson AFB, OH 45433	12. REPORT DATE Sep[redacted]77
	13. NUMBER OF PAGES ~~61~~ (12) 67 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES Approved for public release; IAW AFR 190-17

For James E. Sharrah MSgt USAF
JERRAL F. GUESS, Capt, USAF Deputy Director of Information
Director of Information

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Kalman Filter	Algorithms
Square-Root Filter	Filter
U-D Covariance Factorization Filter	Estimation
Wordlength	Optimal Filter
Numerics	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report presents the concept of square root filters and the closely related U-D covariance factorization filter as viable alternatives to conventional Kalman filters. For a modest increase in computational loading, one obtains optimal filter algorithms equivalent to the Kalman filter if infinite wordlength is assumed, but with vastly superior numerical characteristics with finite wordlength. These algorithms are at least as good a solution to troublesome measurement update computations as implementing a Kalman filter in double precision, since the Kalman filter inherently involves unstable numerics.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

012200

next page

20. continued

The filter algorithms are developed and presented in a form convenient for implementation. Of the covariance square root forms, the Carlson filter is more efficient than the Potter form computationally, and it also maintains triangularity of the square root matrices. The U-D covariance factorization filter is even more efficient, not requiring square root computations. In comparison, the inverse covariance square root filter is often considerably more burdensome, although it too becomes competitive if the measurement dimension is very large.