ON THE POSSIBILITY OF ILL-CONDITIONED COVARIANCE MATRICES IN THE FIRST-ORDER TWO-STEP ESTIMATOR

James L. Garrison [1]

Penina Axelrad [2]

N. Jeremy Kasdin [3]

The first-order two-step nonlinear estimator, when applied to a problem of orbital navigation, is found to occasionally produce first step covariance matrices with very low eigenvalues at certain trajectory points. This anomaly is the result of the linear approximation to the first step covariance propagation. The study of this anomaly begins with expressing the propagation of the first and second step covariance matrices in terms of a single matrix. This matrix is shown to have a rank equal to the difference between the number of first step states and the number of second step states. Furthermore, under some simplifying assumptions, it is found that the basis of the column space of this matrix remains fixed once the filter has removed the large initial state error. A test matrix containing the basis of this column space and the partial derivative matrix relating first and second step states is derived. This square test matrix, which has dimensions equal to the number of first step states, numerically drops rank at the same locations that the first step covariance does. It is formulated in terms of a set of constant vectors (the basis) and a matrix which can be computed from a reference trajectory (the partial derivative matrix). A simple example problem involving dynamics which are described by two states and a range measurement illustrate the cause of this anomaly and the application of the aforementioned numerical test in more detail.

INTRODUCTION

The two-step optimal estimator derived in Haupt, et al. [1] and Kasdin, et al. [2] provides an improved recursive solution to the state estimation problem involving nonlinear measurements. This method breaks the problem into two parts by defining a set of first step states which permit an exact linear measurement model and a nonlinear relationship between the desired states and the first step states. In attempting to apply this first order version of this method to the problem of navigating two satellites relative to each other in an elliptical orbit it is found that occasionally a covariance matrix for the first step states with one very small eigenvalue results. This makes the second step state update ill-conditioned and potentially causes subsequent filter divergence or meaningless state

---

[1] Aerospace Technologist, MS 328 Spacecraft and Sensors Branch, NASA Langley Research Center, Hampton, VA 23681, email: j.l.garrison@larc.nasa.gov

[2] Assistant Professor, Department of Aerospace Engineering Sciences, University of Colorado, Boulder CO 80309-0429

[3] Chief Systems Engineer, Gravity Probe B Relativity Mission, W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA, 94305-4085

estimates. This paper presents an analysis of the cause of this problem and derives a numerical test for the location of the ill-conditioned first step state covariances.

A covariance matrix with very low or zero eigenvalues indicates that a strong correlation has developed between the states and that some linear combination of the states can be known almost exactly. The problems for which the two step filter was developed, however, are non-linear. For a nonlinear function relating the first and second step states there would not exist, in general, a linear combination of first step states which could be known exactly. The analysis presented in this paper is interpreted as identifying a problem with the linear approximations used for the first step covariance propagation in the filter, not to imply that it is possible to find some linear combination of first step states which truly have zero error.

TWO-STEP ESTIMATOR OVERVIEW

Given a non-linear system described by a state vector, $\vec{x}$, of dimension m and the differential equation $\dot{\vec{x}} = g(\vec{x}, t)$, the general estimation problem is to process discrete vector measurements, $\vec{z}_i$, of dimension l related to the state through the nonlinear function $\vec{z}_i = h(\vec{x}_i, t_i)$ and generate the best estimate of the state at those points, $\hat{\vec{x}}_i$. Haupt, et al. [1] define a "first step" state vector, $\vec{y}$, of dimension n (for the cases considered here $n > m$) related to $\vec{x}$, the "second step" state, through a nonlinear function $\vec{y}_i = f(\vec{x}_i, t_i)$. $\vec{y}_i$ is chosen such that the measurement vector can be expressed as a linear combination of the first step states $\vec{z}_i = H\vec{y}_i$. (In all cases considered in this study H is constant.) Thus each new measurement is incorporated through a linear system where traditional methods of linear estimation are applicable.

Mechanization of the two step filter is given in the references [1] and [2] as follows: 1) Perform a standard linear measurement update of the first step state estimate and covariance based on the observations. 2) Compute the second step state estimates by a nonlinear estimation process such as the Gauss-Newton or Lavenberg-Marquardt [4] methods and update the second step covariance matrix by a first order approximation. 3) Propagate the second step states and covariance matrix forward to the time of the next measurement using the state transition matrix or numerical integration as required. 4) Propagate the first step states and covariance matrix to the time of the next measurement using the *a posteriori* first and second step covariance matrices from the last measurement update and the propagated (*a priori*) second step covariance matrix for the upcoming measurement epoch.

This last step is the focus of the research presented here. Equation (33) from reference [1] approximates the first step covariance propagation from the i^{TH} time step to the $i+1^{ST}$ step to first order.

$$P_{y_{i+1}}(-) = P_{y_i}(+) + \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{i+1}(-)} P_{x_{i+1}}(-) \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_{i+1}(-)} - \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(+)} P_{x_i}(+) \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_i(+)} \tag{1}$$

In which a hat $\hat{()}$ over the state vector indicates an estimate of that vector and the notation of $(+)$ signifies the *a posteriori* conditions and $(-)$ signifies the *a priori* conditions.

This generates the *a priori* covariance matrix at the $i+1^{ST}$ time step by adjusting the *a posteriori* covariance from the i^{TH} step by the difference between two positive semidefinite matrices. It is not clear that Eq. (1) is guaranteed to produce a non-singular and positive-definite covariance matrix, owing to the negative sign on the third term. In fact, when applying the two-step filter to an orbital mechanics problem, this equation occasionally resulted in covariance matrices with very low or negative eigenvalues. The following analysis seeks to understand the conditions under which the two step filter would produce such an ill-conditioned first step covariance matrix and also to derive a test to predict when this condition will occur. As the smallest eigenvalues would have to go through zero before becoming negative, this is also a test for non-positive-definite matrices in addition to rank deficient matrices.

The combination of Eqs. (25) and (19) from reference [1] gives.

$$P_{x_i}^{-1}(+) = \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_i(+)} P_{y_i}^{-1}(+) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(+)} \tag{2}$$

2

This expression represents the update of the second step covariance from the (post-measurement updated) first step covariance following the iterative update of the of the second step state (step number 2 above).

It should be noted that in actual practice, both for the orbital mechanics problem which was the motivation for this work and in the simple numerical example presented later in this paper, a UDU^T factored form of the two step filter is used [2]. The original two-step filter from [1] will be used in the analysis in this paper because it is mathematically identical but simpler to understand than the factored form. The specific UDU^T factored algorithm used in all of the numerical simulations in given in reference [2] and the Appendix.

ANALYSIS OF THE PROBLEM

An analysis into the cause of ill-conditioned first step covariances can be broken into three steps. First, the filter equations are expressed in terms of a new matrix, C, which is a function of the two state covariance matrices, P_x and P_y. Second, two properties are found for the column space of C which allow this space to be represented by a reduced number of constant basis vectors. Finally, a numerical test is derived using the set of basis vectors and the reference or predicted state trajectory which identifies points in which P_y may become ill-conditioned.

Formulation in Terms of the Matrix C

The covariance propagation equations for P_x and P_y in the two step estimator described in the previous section are expressed in terms of another matrix, C, defined at each time step for the *a priori* and *a posteriori* conditions.

$$C_i(\pm) \equiv P_{y_i}(\pm) - \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(\pm)} P_{x_i}(\pm) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(\pm)}^T \tag{3}$$

Equation (1) reduces to

$$C_{i+1}(-) = C_i(+) \tag{4}$$

Next, two assumptions are made about the evolution of the first and second step covariance matrices once any large initial state error has been removed.

<u>Assumption 1:</u> $P_{x_i}^{-1}(-)$ is approximated as:

$$P_{x_i}^{-1}(-) \approx \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(-)}^T P_{y_i}^{-1}(-) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(-)} \tag{5}$$

This equation is found by taking Eq. (2), defined for the *a posteriori* conditions, and applying it the the *aposteriori* conditions.

<u>Assumption 2: The partial derivative matrix is approximately constant across measurement updates.</u>

$$\left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(+)} \approx \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(-)} \approx \left.\frac{\partial f}{\partial x}\right|_i \tag{6}$$

Computing this partial matrix from the true, reference, or predicted state are all assumed to be equivalent. The state estimate is of interest in this analysis only as it enters the two-step estimator covariance prediction through the partial derivative matrix, $\partial f/\partial x$. Henceforth, no reference will be made to which side of the measurement update the partial matrix is computed on.

These two assumptions, taken together, amount to approximating the second step covariance matrix propagation along a nominal trajectory as equal to that of a linearized Kalman filter which processes measurements and directly updates the second step states. This is shown as follows; The first step covariance update in the two step filter is given by

$$P_{y_i}^{-1}(+) = P_{y_i}^{-1}(-) + H^T R_i^{-1} H \tag{7}$$

3

Premultiply this n by n matrix equation by $(\partial f/\partial x)|_i^T$ and postmultiplied it by $(\partial f/\partial x)|_i$. Then use (5) and (6) to reduce the resulting m by m matrix expression to the form of a second step state covariance update

$$P_{x_i}^{-1}(+) = P_{x_i}^{-1}(-) + H_{x_i}^T R_i^{-1} H_{x_i} \tag{8}$$

The measurement matrix is $H_{x_i} = H(\partial f/\partial x)|_i$. Equation (8) is recognized as the covariance update in an linearized Kalman filter processing measurements $\vec{z}_i$ and updating the second step state directly from them. The propagation of the second step states between measurements would be the same for the two step filter as well as the linearized filter. Therefore, these two assumptions can be interpreted as the assumption that the evolution of the covariance matrices in two step filter is the same as that for a linearized Kalman filter applied to the same problem. This was found to be true once the large initial state errors were reduced.

Properties of the C Matrix

Here, two important properties of the C matrix which greatly simplify its propagation and ultimate use in the test for ill-conditioned P_y, are derived. Although conditions under which P_y becomes ill-conditioned are of interest here, it is necessary to assume that the matrix P_y^{-1} exists. These two points are reconciled by interpreting this analysis as an inquiry into where P_y becomes *numerically* lower rank, assuming that it will always have an inverse *analytically*.

First Property: The Rank of C is $n - m$. This property is obtained by considering Eq. (3) not as a definition for $C_i(+)$ but as an expression for $P_{y_i}(+)$ given a known $C_i(+)$ and $P_{x_i}(+)$

$$P_{y_i}(+) = C_i(+) + \left. \frac{\partial f}{\partial x} \right|_i P_{x_i}(+) \left. \frac{\partial f}{\partial x} \right|_i^T \tag{9}$$

and noting that the term

$$\left. \frac{\partial f}{\partial x} \right|_i P_{x_i}(+) \left. \frac{\partial f}{\partial x} \right|_i^T \tag{10}$$

is of rank m (assuming $P_{x_i}(+)$ is full rank). In order for $P_{y_i}(+)$ to be full rank, $C_i(+)$ must contain at least $n - m$ column vectors which are linearly independent of the column vectors of (10). Hence, the requirement

$$rank(C_i(+)) \geq n - m \tag{11}$$

Post-multiplying Eq. (3) by the matrix $P_{y_i}^{-1}(+) \left. \frac{\partial f}{\partial x} \right|_i$

$$C_i(+) P_{y_i}^{-1}(+) \left. \frac{\partial f}{\partial x} \right|_i = \left. \frac{\partial f}{\partial x} \right|_i - \left. \frac{\partial f}{\partial x} \right|_i P_{x_i}(+) \left. \frac{\partial f}{\partial x} \right|_i^T P_{y_i}^{-1}(+) \left. \frac{\partial f}{\partial x} \right|_i \tag{12}$$

and using Eq. (2) results in

$$C_i(+) P_{y_i}^{-1}(+) \left. \frac{\partial f}{\partial x} \right|_i = 0 \tag{13}$$

Equation (13) indicates that the m columns of $P_{y_i}^{-1}(+)\partial f/\partial x|_i$ are in the nullspace of $C_i(+)$. If the first step states are independent of each other at each time step, then the matrix $(\partial f/\partial x|_i)$ is of rank m. This would always be true in the suggested practice [1] of defining the first step states as the set of second step states augmented by the nonlinear measurement equations. In this case m rows of the partial derivative matrix form an m by m identity.

The rank of the matrix product $P_{y_i}^{-1}(+)\partial f/\partial x|_i$ is therefore less than or equal to m ($P_{y_i}^{-1}(+)$ must be full rank). If the rank of this product was less than m, however, then there would exist some vector $\vec{a} \neq 0$ such that $P_{y_i}^{-1}(+)\partial f/\partial x|_i \vec{a} = 0$. Post-multiplying the partial derivative matrix, which must have m linearly independent columns, by a nonzero vector, $\vec{a}$, will give another non-zero vector, $\vec{b}$; $(\partial f/\partial x)|_i \vec{a} = \vec{b} \neq 0$. This results in the contradiction; $P_{y_i}^{-1}(+)\vec{b} = 0$. Therefore, the matrix product

4

$P_{y_i}^{-1}(+)\partial f/\partial x|_i$ must have a rank equal to m. It is therefore concluded that $C_i(+)$ also has at least m linearly independent null vectors and hence the dimension of its nullspace is at least m.

$$null(C_i(+)) \geq m \tag{14}$$

Combining Eqs. (11) and (14) and using the fundamental relationship for any n by n matrix

$$rank(C_i(+)) + null(C_i(+)) = n \tag{15}$$

results in the conclusion

$$rank(C_i(+)) = n - m \tag{16}$$

and, consequently

$$null(C_i(+)) = m \tag{17}$$

From Eq. (4), the nullspace and column space of $C_{i+1}(-)$ are identical to those same spaces defined for $C_i(+)$. This rank property would consequently apply to the *a priori* condition as well.

$$rank(C_{i+1}(-)) = n - m \tag{18}$$

Second Property: The Column Space of C is Fixed in R^n for a steady state filter with $l \leq m$. Consider that the nullspace and row space of a matrix are orthogonal complements, ie., every vector in one must be orthogonal to every vector in the other. Therefore, if it can be shown that the nullspace of C remains fixed, then the row space of C must also be fixed. The row space and column space of C are equivalent because it is a symmetric matrix (defined as the sum of two symmetric matrices). This analysis will therefore focus on showing that the nullspace of C remains constant.

To simplify the notation, the matrices $N_i(+)$ and $N_i(-)$ will be defined for the products

$$N_i(\pm) \equiv P_{y_i}^{-1}(\pm) \left. \frac{\partial f}{\partial x} \right|_i \tag{19}$$

The statement that the nullspace of C remains the same between measurement updates (Eq. (4)) is equivalent to the statement that the columns of $N_{i+1}(-)$ span the same space as the columns of $N_i(+)$. This can be written as

$$N_{i+1}(-) = N_i(+)A_{i+1} \tag{20}$$

in which A_{i+1} is a set of coefficients expressing each column vector of $N_{i+1}(-)$ as a linear combination of the column vectors of $N_i(+)$.

The update of the first step covariance at the $i + 1^{ST}$ time step using the standard Kalman filter is

$$P_{y_{i+1}}^{-1}(+) = P_{y_{i+1}}^{-1}(-) + H^T R_{i+1}^{-1} H \tag{21}$$

Post-multiplying by $(\partial f/\partial x)|_{i+1}$, expressing this in terms of the $N_{i+1}(\pm)$ matrix defined in (19) and substituting in (20) results in

$$N_{i+1}(+) = N_i(+)A_{i+1} + H^T R_{i+1}^{-1} H \left. \frac{\partial f}{\partial x} \right|_{i+1} \tag{22}$$

First, consider the special case in which the columns of H^T lie in the column space of $N_i(+)$. The columns of H^T can then be used as l of the basis vectors for the column space of $N_i(+)$. The other basis vectors are defined as the columns of an n by $m - l$ matrix $\bar{N}$. The matrix $N_i(+)$ can therefore be expressed as the linear combination

$$N_i(+) = \bar{N}B_i + H^T D_i \tag{23}$$

Substituting this into (22) gives

$$N_{i+1}(+) = \bar{N}B_i A_{i+1} + H^T \left[D_i A_{i+1} + R_{i+1}^{-1} H \left. \frac{\partial f}{\partial x} \right|_{i+1} \right] \tag{24}$$

5

Therefore the same basis $\{\bar{N}, H^T\}$ spans the column space of both $N_{i+1}(+)$ and $N_i(+)$ and consequently spans the nullspace of both $C_{i+1}(+)$ and $C_i(+)$. This shows that if $l \leq m$ and H^T is in the nullspace of C then the column space of C remains fixed in R^n.

The arguments given above can be extended heuristically to the situations in which H^T does not initially lie in the column space of $N_i(+)$. As each new measurement that is processed, a new term of the form $H^T R^{-1} H (\partial f/\partial x)$ is added to $N_i(+)$. The column vectors of this term, which are in the column space of H^T, are the only modification that the filter can make to the nullspace of C. The only direction in which the columns of N can be changed is within the span of H^T. One would expect that as i increases, some l dimensional subspace of the column space of $N_i(\pm)$ would approach the column space of H^T. This, of course, does not rigorously prove this property. The numerical simulations, however, run on models with $l \leq m$ all have produced a C with a column space that stays fixed after the initial state transients are reduced.

In the special case in which $l = m$ and the filter is started with little *a priori* state knowledge ($P_{y_o}^{-1}$ very small), the matrix $N_o(+)$ could be approximated as

$$N_o(+) \approx H^T R_o^{-1} H \left. \frac{\partial f}{\partial x} \right|_o \tag{25}$$

for the initial time steps because the magnitudes of the column vectors of $H^T R_o^{-1} H$ are much larger than those of the columns of $P_{y_o}^{-1}(+)$. For this special case, the columns of H^T would span the nullspace of C.

In the case of $l > n$, the same arguments used above can be used to show that the columns of H^T are the only modifications which are possible to the column space of C, but there are too many column vectors in H^T to form a basis. For this reason, it will not generally be true that the column space of C stays fixed for $l > m$. The span of the columns of the matrix product

$$H^T R_i^{-1} H \left. \frac{\partial f}{\partial x} \right|_i \tag{26}$$

does not stay constant because of the partial derivative factor. The basis of the nullspace of C would be determined by the accumulation of terms like (26) from all previous time steps. This case would be rare in actual practice. Most estimation problems do not involve an observation vector which has a larger dimension than the state vector.

It is only the space spanned by the columns of C which is of interest, not the actual elements of the C matrix. This space has been demonstrated to remain fixed in R^n when $l \leq m$. For the remainder of this paper we will use the notation $\{\vec{c}_1, \vec{c}_2, ... \vec{c}_{n-m}\}$ to indicate the basis for the column space of C without any reference to a specific time step or *a priori* or *a posteriori* state.

It should be emphasized that these results are independent of how the second step state covariance is propagated between measurements. Including process noise would not effect the possibility of generating the ill-conditioned covariance matrices. It may, however, change the specific location of these anomalies by generating a different C matrix.

Test for Ill-conditioned First Step Covariance

The C matrix is now used to derive a test to predict the location of points in which the P_y matrix may become ill-conditioned. It must be assumed that P_y^{-1} still exists and that the filter will never actually generate a singular P_y. The test which is desired will be one which sets a *numerical* tolerance on how close to singular we allow P_y to become. Also, it is desired to have a test which is independent of the filter operations so that it could be operated on a reference or predicted trajectory to determine beforehand which sections of the trajectory are susceptible to the ill-conditioned covariances. Such a rank test can be derived by starting with (9).

As mentioned before, Eq. (9) expresses a matrix which must be of rank n as the sum of a matrix of rank $n - m$ and one of rank m. This would be possible as long as the $n - m$ basis vectors $\{\vec{c}_1, \vec{c}_2, ... \vec{c}_{n-m}\}$ are linearly independent of the column space of $(\partial f/\partial x) P_x (\partial f/\partial x)^T$. This has already been shown when deriving the first property of C. We now consider cases where

at least one basis vector of C becomes *close* to being linearly dependent on the column space of $(\partial f/\partial x)P_x(\partial f/\partial x)^T$. To simplify this further and remove the time-dependent and filter-state-dependent covariance P_x, consider that the column space of $(\partial f/\partial x)P_x(\partial f/\partial x)^T$ is spanned by the m columns vectors $\{\partial f/\partial x_1, \partial f/\partial x_2, \ldots \partial f/\partial x_m\}$. This leads to a test to look for conditions under which the set of $n-m$ vectors $\{\vec{c}_1, \vec{c}_2, \ldots \vec{c}_{n-m}\}$ become close to being linearly dependent upon the set of m column vectors $\{\partial f/\partial x_1, \partial f/\partial x_2, \ldots \partial f/\partial x_m\}$. (From the assumption in Eq. (6) these partial derivatives can be computed from a reference trajectory or from a state trajectory predicted ahead of the current filter time).

All of this can be summarized in a test of the numerical rank test of an n by n matrix.

$$rank\left(\left[\vec{c}_1, \vec{c}_2, \ldots \vec{c}_{n-m}, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_m}\right], \epsilon\right) < n \tag{27}$$

for some tolerance ϵ. Points for which the condition described in Eq. (27) is true are the locations where an ill-conditioned first step covariance matrix would occur. The numerical rank of an n by n matrix is defined as the number of singular values greater than ϵ [3]. The span of $\{\vec{c}_1, \vec{c}_2, \ldots \vec{c}_{n-m}\}$ is fixed (property 2) and the partial derivative matrix is computed along a reference trajectory (assumption 1). Therefore, the test defined in (27) could conceivably be applied to points ahead of the present filter state to identify future trajectory points in which generating an ill-conditioned covariance matrix is possible.

Geometric Interpretation

These concepts are visualized geometrically using a system of two second step states and three first step states ($n = 3$ and $m = 2$) as shown in Figure 1. The covariance matrix P_x is represented by the two-dimensional error ellipse in the $x_1 - x_2$ plane. The effect of (10) is to rotate and scale that ellipse to a new orientation in the $y_1 - y_2 - y_3$ space. The column space of $\partial f/\partial x$ defines the plane of that ellipse. The first property of the C matrix is that its rank is 1, hence its column space is a single line in the $y_1 - y_2 - y_3$ space as shown in Figure 1.

The second property of C states that the orientation of this line would remain fixed in $y_1 - y_2 - y_3$ space. The plane defined by the column space of $(\partial f/\partial x)$, however, would change orientation with the state vector evolution. This line and plane together span R^3 under non-degenerate conditions. The orientation of the plane with respect to the line control the linear independence of the columns of the P_y matrix as described by Eq. (27).

As long as this line is not coplanar with the column space of $(\partial f/\partial x)$ P_y remains rank 3. If this line does fall in the plane defined by the column space then the union of these two subspaces would not span R^3 and consequently P_y would be singular. If the line lies *nearly* in the plane, then the P_y matrix is ill-conditioned. In this case, there will exist a coordinate transformation such that in one direction in the $y_1 - y_2 - y_3$ space the linear covariance approximation given in (1) would predict that the first step state estimate will have very little error. This transformation will define a linear combination of the over-determined first step states which the filter predicts can be known almost exactly. In reality, because the first and second step states are *nonlinearly* related, there cannot be a linear combination of first step states which are known exactly.

NUMERICAL EXAMPLE

A simple two state example problem is used to illustrate this anomaly in the two step filter. The geometry of this problem is shown in Figure 2. The kinematics consist of a particle following a spiral path defined by a constant angular velocity ω_0 and a constant radial velocity v_0 as if the particle was attached to a string of increasing length. The nonlinear differential equations in Cartesian coordinates are

$$\frac{dx_1}{dt} = \frac{x_1 v_0}{\sqrt{x_1^2 + x_2^2}} - x_2 \omega_0 \tag{28}$$
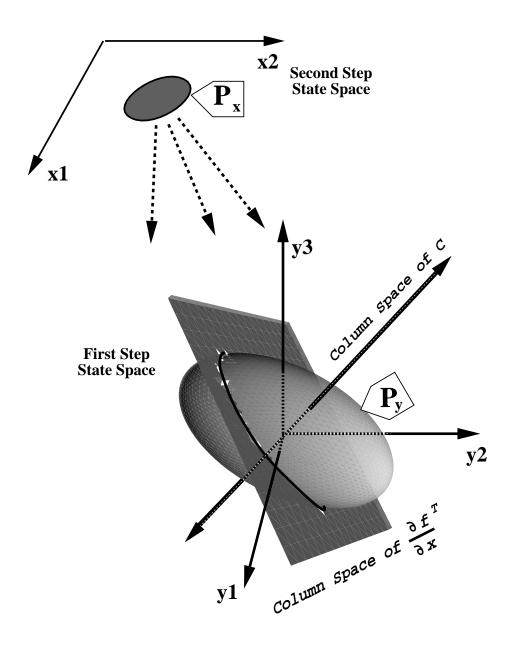
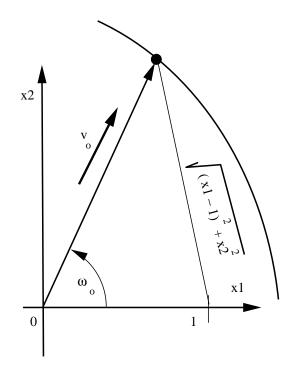Figure 1: Geometric Interpretation of the Column Space of P_y

Figure 2: Example Problem Geometry

$$\frac{dx_2}{dt} = \frac{x_2 v_0}{\sqrt{x_1^2 + x_2^2}} + x_1 \omega_0 \tag{29}$$

The linearized state dynamics matrix, $A(\vec{x})$, used to propagate the state transition matrix by numerically integrating $\dot{\Phi} = A(\vec{x})\Phi$ is

$$A = \begin{bmatrix} \frac{x_2^2 v_0}{(x_1^2+x_2^2)^{\frac{3}{2}}} & -\frac{v_0 x_1 x_2}{(x_1^2+x_2^2)^{\frac{3}{2}}} - \omega_0 \\ -\frac{v_0 x_1 x_2}{(x_1^2+x_2^2)^{\frac{3}{2}}} + \omega_0 & \frac{x_1^2 v_0}{(x_1^2+x_2^2)^{\frac{3}{2}}} \end{bmatrix} \tag{30}$$

The measurement is the range from a fixed point located at the coordinates $(1,0)$ to the particle. This gives the measurement equation.

$$\vec{z} = \left[\sqrt{(x_1 - 1)^2 + x_2^2} \right] \tag{31}$$

The first step states are defined as the second step states augmented by the measurement, following the suggestion in reference [1].

$$\vec{y} = \begin{bmatrix} \sqrt{(x_1 - 1)^2 + x_2^2} \\ x_1 \\ x_2 \end{bmatrix} \tag{32}$$

Hence $H = [\begin{array}{ccc} 1 & 0 & 0 \end{array}]$ and the partial matrix, $\partial f / \partial x$, is given by;

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{x_1 - 1}{\sqrt{(x_1-1)^2+x_2^2}} & \frac{x_2}{\sqrt{(x_1-1)^2+x_2^2}} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{33}$$

Specific numbers used in the example and the filter implementations are all listed in table 1. The initial conditions are in error from the reference starting conditions $(2,0)$ and a normally distributed error is added to the measurements. A process noise term is included in the filter to

9

Table 1:

NUMERICAL DATA FOR EXAMPLE PROBLEM

Number of Filtered Points:	30001
Independent Variable (t) range:	$0:6$
Filter Initial State ($\hat{x}_o(+)$):	$\{2.0, 0.0\}$
True Initial State ($\vec{x}(0)$):	$\{2.583, 0.313\}$
Measurement Noise (1σ, normal distribution):	0.01
Measurement Covariance (R):	10^{-4}
Filter *a priori* Second Step State Covariance ($P_{x0}(-)$):	$diag\{0.25, 0.25\}$
Filter *a priori* First Step State Covariance ($P_{y0}(-)$):	$\begin{bmatrix} 0.226 & 0.209 & 0.0 \\ 0.209 & 0.247 & 1.85\times 10^{-4} \\ 0.0 & 1.85\times 10^{-4} & 0.248 \end{bmatrix}$
Second Step State Discrete Time Process Noise (Q_d):	$diag\{\ 10^{-12} \quad 10^{-12}\}$

prevent the filter from losing sensitivity to new measurements as $t \to \infty$ and $P_{y,x} \to 0$. No dynamic noise was simulated in the truth model, however. A plot of the particle motion and one set of noisy measurements is shown in Figure 3. A large number (30000) of data points were generated, this was to illustrate the sharp decreases in eigenvalues of P_y which can occur only over a very small time period.

Eigenvalues of the $P_y(+)$ matrix on the order of 10^{-15} occurred near $t \approx 1.3$ and $t \approx 4.8$. The matrix C is computed from each point in the filter time history from the *a posteriori* state covariances and estimated state. The largest eigenvector of C was used as the basis of the column space of C. Figure 4 plots the numerical values demonstrating the aforementioned properties of C. The three singular values of C are plotted in Figure 4(a). Note that the first singular values on the order of 10^{18} larger than the other two, and consequently the rank of C is 1 (property 1). The dot product between the "steady state" $\vec{c}$, computed as the average of the last 5000 points, and each previous $\vec{c}_i$ is plotted in Figure 4(b). This dot product is very nearly unity once the large initial state error has been reduced by the filter. Hence, it has been demonstrated that the column space of C is fixed (property 2).

The explanation given in this study for the constant basis of C's column space is illustrated in Figure 4(c). In this figure the dot product between the H matrix and the basis vector $\vec{c}$ is plotted. After $t \approx 0.5$ this product becomes very small in comparison to the magnitude of the H^T and $\vec{c}$ vectors (both unity) indicating that the two vectors are orthogonal and therefore the column of H^T is in the nullspace of C. The small jump near $t \approx 1.3$ in Figure 4(c) is possibly caused by numerical problems resulting from the near singularity at that point. The magnitude of this deviation is still very small as compared to unity.

Figure 5 illustrates the rank test defined in Eq. (27) using a numerical tolerance of 10^{-4} on top of a plot of the eigenvalues of $P_y(+)$. This test condition is computed from the steady state eigenvector of C and the partial derivative matrix computed along the true trajectory. The rank test correctly predicts the two locations of low eigenvalues in the P_y time history as illustrated on Figure 5.

Figure 6 plots the square of the error between the filter *a posteriori* prediction for the first step states and the true values of those states, rotated into the direction of the minimum eigenvalue of P_y. On top of that plot is the minimum eigenvalue of P_y from the filter predicted first step covariance. All of the numerical results presented herein were from a single filter run. Monte Carlo simulations of a statistically distributed ensemble of runs would produce similar results. The specific locations of the points in which P_y becomes ill-conditioned, however, would be slightly different for each member of the ensemble because of differences in the orientation of the space $\{\vec{c}_1, \vec{c}_2, ...\vec{c}_{n-m}\}$ depending upon the starting conditions and the (noisy) measurement time history.
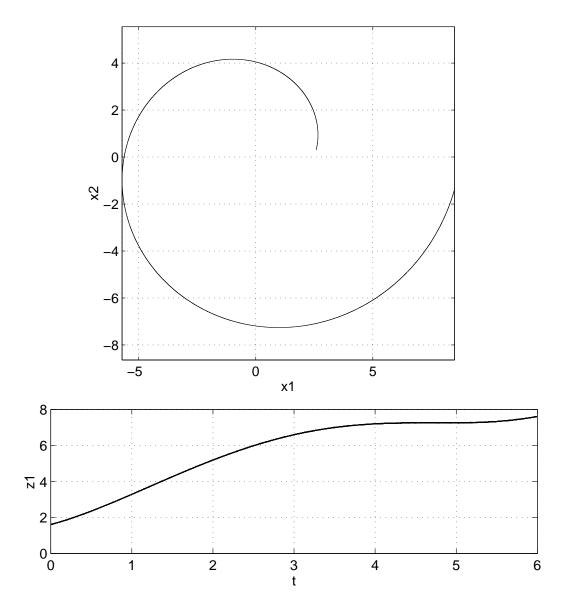
Figure 3: Example Problem State and Observations
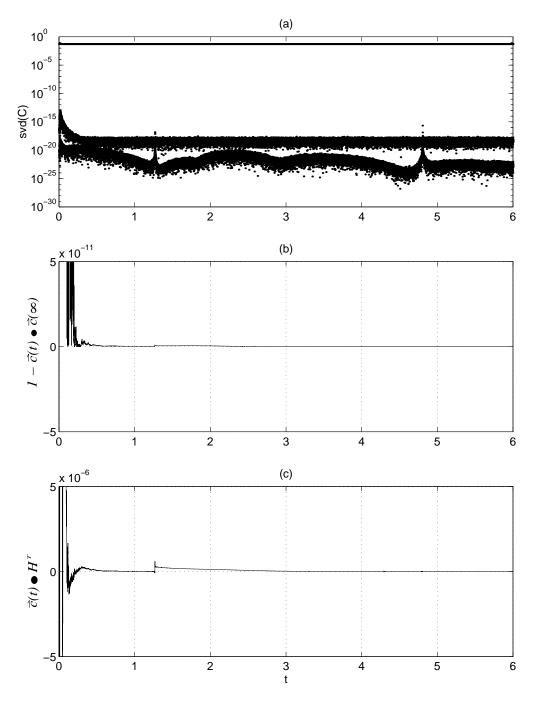
Figure 4: Properties of the Column Space of C

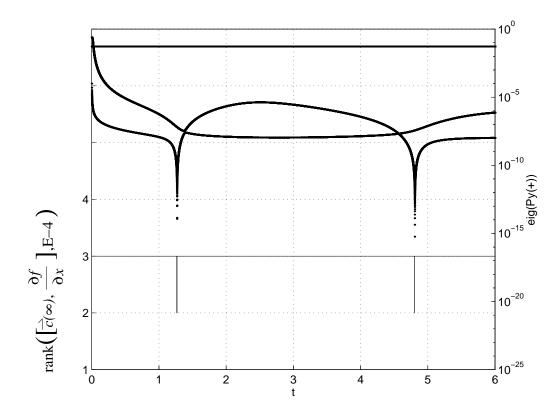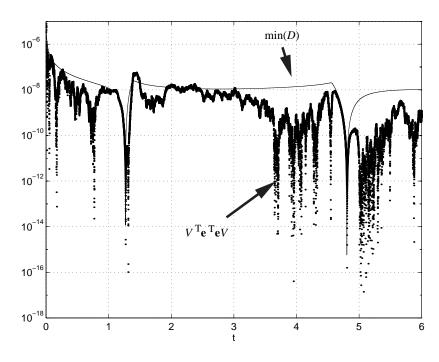Figure 5: Numerical Rank Test for Ill-Conditioned P_y

Eigenvalue decompositon of Py is $Py = V\,D\,V^{\mathrm{T}}$
Error vector is computed as: $e = y - y_{\mathrm{TRUE}}$

Figure 6: First Step State Error in Minimum $eig(P_y)$ Direction

13

CONCLUSION

The occurrence of ill-conditioned time updates for the first step covariance matrix in the two step filter has been explained. This anomaly is the result of the linearized approximation to the first step covariance propagation. In this approximation, there were conditions under which the positive definite second step covariance matrices did not combine to produce a full set of linearly independent column vectors for the first step covariance matrix. This analysis was simplified by the ability to formulate the filter covariance propagation in terms of a reduced set of basis vectors which, under certain conditions and assumptions, spanned a fixed space. One result of this study was a numerical test of the linear dependence of those basis vectors and the partial derivative matrix which was shown to correspond with the first step covariance matrix becoming ill conditioned.

APPENDIX

$\underline{UDU^T \text{ Factored Algorithm for the Two Step Filter}}$ [2], [5]

- Initialization of the following:

$$P_{x_o}(+), \quad P_{y_o}(+), \quad \vec{x_o}(+), \quad \vec{y_o}(+) = f(\vec{x_o}(+)) \tag{34}$$

- † Propagation of the second step states and the state transition matrix $\Phi_{(i,i-1)}$ by numerical integration from the $i-1^{ST}$ to the i^{TH} time step. The first step state is estimated from

$$\hat{y}_i(-) = \hat{y}_{i-1}(+) + f(\hat{x}_i(-)) - f(\hat{x}_{i-1}(+)) \tag{35}$$

The second step covariance is propagated by computing the block matrices Y_x and $\tilde{D}_x$

$$Y_x = \left[\begin{array}{ccc} \Phi_{i,i-1} U_{x_{i-1}}(+) & \vdots & G_d \end{array} \right] \tag{36}$$

$$\tilde{D}_x = \left[\begin{array}{ccc} D_{x_{i-1}}(+) & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & Q_d \end{array} \right] \tag{37}$$

and then reducing them to a m by m upper triangular and diagonal matrices, $U_{x_i}(-)$ and $D_{x_i}(-)$, respectively, such that

$$U_{x_i}(-)D_{x_i}(-)D_{x_i}^T(-) = Y_x \tilde{D}_x Y_x^T = P_{x_i}(-) \tag{38}$$

This is done using the Modified Weighted Gram-Schmidt (MWGS) [5] method.

Similarly, the subsequent propagation of the first step states as defined in Eq. (1) is done by computing the matrices Y_y and $\tilde{D}_y$.

$$Y_y = \left[\begin{array}{ccccc} U_{y_{i-1}}(+) & \vdots & \left. \frac{\partial f}{\partial x} \right|_{\hat{x}_i(-)} U_{x_i}(-) & \vdots & \left. \frac{\partial f}{\partial x} \right|_{\hat{x}_{i-1}(+)} U_{x_{i-1}}(+) \end{array} \right] \tag{39}$$

$$\tilde{D}_y = \left[\begin{array}{ccccccccc} D_{y_{i-1}}(+) & \vdots & 0 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & D_{x_i}(-) & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & 0 & \vdots & -D_{x_{i-1}}(+) \end{array} \right] \tag{40}$$

The MWGS process is then used to compute $U_{y_i}(-)$ and $D_{y_i}(-)$ such that

$$U_{y_i}(-)D_{y_i}(-)U_{y_i}^T(-) = Y_y \tilde{D}_y Y_y^T \tag{41}$$

- The measurement update is performed by converting the observation vector $\vec{z}_i$ into l uncorrelated scalar observations $(\vec{z}^*)$ with unity variance through a coordinate transformation.

$$WW^T = R_i \tag{42}$$

$$\vec{z}^* = W^{-1}\vec{z} \tag{43}$$

$$H^* = W^{-1}H \tag{44}$$

Details of the scalar measurement update are given in [5]. This step will result in the *a posteriori* first step states $(\hat{y}_i(+))$ and $U_{y_i}(+)$ and $D_{y_i}(+)$ factors. The inverse of $P_{y_i}^{-1}(+)$ can then be computed by the more numerically stable method

$$P_{y_i}^{-1}(+) = U_{y_i}^{-T}(+) \begin{bmatrix} 1/D_{y_i}(+)(1,1) & 0 & \cdots & 0 \\ 0 & 1/D_{y_i}(+)(2,2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/D_{y_i}(+)(n,n) \end{bmatrix} U_{y_i}^{-1}(+) \tag{45}$$

- The *a posteriori* second step states are found from the $\hat{y}_i(+)$, $U_{y_i}(+)$ and $D_{y_i}(+)$ by numerically minimizing the cost function

$$J(\hat{x}) = \frac{1}{2}(\hat{y}_i(+) - f(\hat{x}))^T P_{y_i}^{-1}(+)(\hat{y}_i(+) - f(\hat{x})) \tag{46}$$

The Lavenberg-Marquardt method [4] is used for this numerical minimization.
‡ The iterations are initialized with $\hat{x}^o = \hat{x}_i(-)$ and $\lambda = 0.001$. At each k^{TH} iteration, the Hessian is computed from the approximation

$$H_G \approx \left.\frac{\partial f}{\partial x}\right|_{\hat{x}^k}^T P_{y_i}^{-1}(+) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}^k} \tag{47}$$

The linear system

$$[\alpha]\delta\hat{x} = -\left.\frac{\partial J(\hat{x})}{\partial \hat{x}}\right|_{\hat{x}^k} \tag{48}$$

is solved to update the state estimate

$$\hat{x}^{k+1} = \hat{x}^k + \delta\hat{x} \tag{49}$$

in which the gradient of J is

$$\left.\frac{\partial J}{\partial \hat{x}}\right|_{\hat{x}_k} = -(\hat{y}_i - f(\hat{x}_k))P_{y_i}^{-1}(+) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_k} \tag{50}$$

and α is computed from

$$\alpha_{l,l} = H_{G_{l,l}}(1 + \lambda) \tag{51}$$

$$\alpha_{p,q} = H_{G_{p,q}} \quad \text{for } p \neq q \tag{52}$$

The change in the cost function $J(\hat{x}^{k+1})$ determines how the next iteration is performed.

- If $J(\hat{x}^{k+1}) \geq J(\hat{x}^k)$ then $\lambda = 10\lambda$;
  $k = k + 1$; $\hat{x}^{k+1} = \hat{x}^k$ and execution is returned to ‡.
- If $J(\hat{x}^{k+1}) < J(\hat{x}^k)$ then $\lambda = \lambda/10$; $k = k + 1$; and execution is returned to ‡.

The iterations are continued until the cost function decreases by less than 0.01. After this indicates convergence, the *a posteriori* second step states are updated by

$$\hat{x}_i(+) = \hat{x}^k \tag{53}$$

and the *a posteriori* second step covariance is computed from Eq. (2). $U_{x_i}(+)$ and $D_{x_i}(+)$ are then obtained by Cholesky factorization. The time step is incremented $i = i + 1$ and execution is returned to †

References

[1] Haupt, Gordon T., Kasdin, Jeremy N., Keiser, George M., and Parkinson, Bradford W., "Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation", *AIAA Journal of Guidance, Control and Dynamics*, Vol. 19, No. 3, May-June 1996, pp 643-649.

[2] Kasdin, Jeremy N., and Haupt, Gordon T., "Second-Order Correction and Numerical Considerations for the Two-Step Optimal Estimator", *AIAA Journal of Guidance, Control and Dynamics*, in press.

[3] *MATLAB Reference Guide* , the Mathworks, Inc., Natick, MA, 1992, pp 405.

[4] Press, William H., Teukolsky, Saul A., Vetterling, William T., and Flannery, Brian P. *Numerical Recipes in C* , Cambridge University Press, New York, 1994, pp 683-685.

[5] Maybeck, Peter S., *Stochastic Models, Estimation and Control*, Volume 1., Academic Press, New York, 1979.

[6] Strang, Gilbert, *Linear Algebra and Its Applications*, Second Edition, Academic Press, 1980.

[7] Noble, Ben, and Daniel, James W., *Applied Linear Algebra* , Second Edition, Prentice-Hall, Inc., 1977.