JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 21, No. 5, September–October 1998

Ill-Conditioned Covariance Matrices in the First-Order Two-Step Estimator

James L. Garrison*
NASA Langley Research Center, Hampton, Virginia 23681
Penina Axelrad†
University of Colorado, Boulder, Colorado 80309-0429
and
N. Jeremy Kasdin‡
Stanford University, Stanford, California 94305-4085

The first-order two-step estimator is found to occasionally produce first-step covariance matrices with very low, sometimes negative, eigenvalues. These low eigenvalues can cause large errors or meaningless estimates. A single matrix is found, which is shown to have a rank equal to the difference between the number of first- and second-step states. Furthermore, it is demonstrated that the basis of the column space of this matrix remains fixed once the large initial state error has decreased. A test matrix containing the (constant) basis of this column space and the partial derivative matrix relating first and second step states is derived. This matrix numerically drops rank at the same locations that the first-step covariance does. A simple example problem involving dynamics described by two states and a range measurement illustrates the cause of this anomaly and application of the aforementioned numerical test. Suggested modifications to the filter that can mitigate the numerical problems caused by this anomaly are given.

I. Introduction

THE two-step optimal estimator derived in Haupt et al.[1] and Kasdin and Haupt[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 that permit an exact linear measurement model and a nonlinear relationship between these first-step states and the second-step states that describe the dynamics. In attempting to apply the first-order version of this filter to the problem of navigating two satellites relative to each other in an elliptical orbit,[3] 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 estimates. This paper presents an analysis of the cause of ill-conditioned first-step state covariance matrices and of this problem and derives a numerical test for the location of this problem.

A covariance matrix with very low or nearly 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. In the two-step filter, the first-step state vector is often selected to have a larger dimension than the second-step state vector. This does open the possibility that there may exist linear combinations of the first-step states that can at times have very low estimation error without the corresponding second-step states exhibiting any such characteristics.

II. Two-Step Estimator Overview

Given a continuous nonlinear system described by a state vector $\boldsymbol{x}$ of dimension m and the differential equation $\dot{\boldsymbol{x}} = \boldsymbol{g}(\boldsymbol{x}, t)$, the general estimation problem is to process noisy discrete vector measurements, $\boldsymbol{z}_i + \boldsymbol{\nu}_i$, of dimension l related to the state through the nonlinear function $\boldsymbol{z}_i = \boldsymbol{h}_i(\boldsymbol{x}_i)$ and generate the best estimate of the state at those points, $\hat{x}_i$. The measurement noise ($\boldsymbol{\nu}_i$) is described by its covariance matrix $R_i \equiv E\{\boldsymbol{\nu}_i \boldsymbol{\nu}_i^T\}$. Haupt et al.[1] define a first-step state vector $\boldsymbol{y}$ of dimension n related to $\boldsymbol{x}$, the second-step state, through a nonlinear function $\boldsymbol{y}_i = \boldsymbol{f}_i(\boldsymbol{x}_i)$. The first-step state must be defined such that the second-step state is observable from it at each time step. This, as a minimum, requires that the first-step state vector have a dimension equal to or larger than that of the second-step state vector ($n \geq m$). The first step-state vector $\boldsymbol{y}_i$ is specifically chosen so that the measurement vector can be expressed as a linear combination of the first-step states $\boldsymbol{z}_i = H\boldsymbol{y}_i$. (In all cases considered in this study H is a constant $l \times n$ matrix.) 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 Refs. 1 and 2. At the ith time step, the a priori estimates of the first- and second-step states and their covariance matrices are available. A caret over the state vector indicates an estimate of that vector, the notation of $(+)$ signifies the a posteriori conditions and $(-)$ signifies the a priori conditions.

One complete step of the two-step filter, incorporating the ith observation ($\boldsymbol{z}_i$), proceeds as follows.

1) Perform a standard linear measurement update of the first-step state estimate and covariance based on the observations

$$\hat{y}_i(+) = \hat{y}_i(-) + K_i(\boldsymbol{z}_i - H\hat{y}_i(-)) \tag{1}$$

$$K_i = P_{y_i}(-)H^T\left[HP_{y_i}(-)H^T + R_i\right]^{-1} \tag{2}$$

$$P_{y_i}^{-1}(+) = P_{y_i}^{-1}(-) + H^T R_i^{-1} H \tag{3}$$

2) Compute the second-step state estimates by a nonlinear estimation process such as the Gauss–Newton or Lavenberg–Marquardt[4] methods, which numerically minimize the cost function

$$J_x = (\hat{y}_i(+) - \boldsymbol{f}_i(\boldsymbol{x}_i))^T P_{y_i}^{-1}(+)(\hat{y}_i(+) - \boldsymbol{f}_i(\boldsymbol{x}_i)) \tag{4}$$

and produce the a posteriori second-step state estimate $\hat{x}_i(+)$. The second-step state covariance is obtained as

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

Presented as Paper AAS 97-195 at the AAS/AIAA Space Flight Mechanics Meeting, Huntsville, AL, Feb. 10–12, 1997; received May 16, 1997; revision received Jan. 13, 1998; accepted for publication Jan. 13, 1998.

*Aerospace Technologist, Spacecraft and Sensors Branch, MS 328. E-mail: j.l.garrison@larc.nasa.gov. Senior Member AIAA.

†Assistant Professor, Department of Aerospace Engineering Sciences. Senior Member AIAA.

‡Chief Systems Engineer, Gravity Probe B Relativity Mission, W. W. Hansen Experimental Physics Laboratory. Senior Member AIAA.

in which $\partial f/\partial x$ is an $n \times m$ matrix of partial derivatives of the first-step states with respect to the second-step states.

3) Propagate the second-step state and covariance matrix forward to the time of the next measurement using numerical integration.

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 (from step 3). Reference 1 approximates the first-step covariance propagation from the ith time step to the $i+1$ step to first order,

$$P_{y_{i+1}}(-) = P_{y_i}(+) + \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{i+1}(-)} P_{x_{i+1}}(-) \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_{i+1}(-)} - \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(+)} P_{x_i}(+) \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_i(+)} \tag{6}$$

This last step is the focus of the research presented here. Equation (6) generates the a priori covariance matrix at the $i+1$ time step by adjusting the a posteriori covariance from the ith step by the difference between two positive semidefinite matrices. This is not guaranteed to produce a positive-definite covariance matrix, owing to the negative sign on the third term. In fact, when applying the two-step filter to the aforementioned orbital mechanics problem, this equation occasionally results in covariance matrices with very low eigenvalues. Sometimes these small eigenvalues become negative, which represents a physically meaningless condition and can result in failure of the iterative second-step solution. The following analysis seeks to understand the cause of these ill-conditioned first-step covariance matrices by determining the conditions under which they occur and deriving a test to predict their location. As the smallest eigenvalues would have to go through zero before becoming negative, this test also identifies nonpositive-definite matrices in addition to rank deficient matrices.

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, a UD covariance factorization of the two-step filter is used.[2] The original two-step filter from Ref. 1 will be used in the analysis because it is mathematically identical but simpler to understand than the factored form. The UD covariance factored algorithm used in all of the numerical simulations is given in Ref. 2.

III. 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 that identifies points in which P_y may become ill-conditioned is derived using this set of basis vectors and the partial derivative matrix.

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 preceding section are expressed in terms of another matrix C, defined at each time step for the a priori and a posteriori conditions,

$$C_i(\pm) \equiv P_{y_i}(\pm) - \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(\pm)} P_{x_i}(\pm) \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_i(\pm)} \tag{7}$$

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.

Assumption 1: $P^{-1}_{x_i}(-)$ is approximated as

$$P^{-1}_{x_i}(-) \approx \left.\frac{\partial f}{\partial x}\right|^T_{\hat{x}_i(-)} P^{-1}_{y_i}(-) \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(-)} \tag{8}$$

This equation is found by taking Eq. (5), defined for the a posteriori conditions and applying it to the a priori conditions.

Assumption 2: The partial derivative matrix is approximately constant across measurement updates

$$\left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(+)} \approx \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_i(-)} \approx \left.\frac{\partial f}{\partial x}\right|_i \tag{9}$$

It is also assumed that computing this partial derivative using the estimated state gives approximately the same result as computing it using the true state. Henceforth, no reference will be made to which side of the measurement update the partial derivative matrix is computed on.

These two assumptions, taken together, amount to approximating the second-step covariance matrix propagation along a reference trajectory as equal to that of a linearized Kalman filter, which processes measurements and directly updates the second-step states. This is shown by premultiplying the first-step covariance update (an $n \times n$ matrix equation) by $(\partial f/\partial x)|^T_i$ and postmultiplying it by $(\partial f/\partial x)|_i$. The resulting $m \times m$ matrix expression is reduced to the form of a second-step state covariance update,

$$P^{-1}_{x_i}(+) = P^{-1}_{x_i}(-) + H^T_{x_i} R^{-1}_i H_{x_i} \tag{10}$$

in which the linearized measurement matrix is $H_{x_i} = H(\partial f/\partial x)|_i$. The propagation of the second-step states between measurements would be the same for the two-step filter as well as the linearized filter.

Properties of the C Matrix

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^{-1}_y 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.

Property 1: The rank of C is $n-m$. This property is obtained by arranging Eq. (7) as an expression for $P_{y_i}(+)$ given $C_i(+)$ and $P_{x_i}(+)$

$$P_{y_i}(+) = C_i(+) + \left.\frac{\partial f}{\partial x}\right|_i P_{x_i}(+) \left.\frac{\partial f}{\partial x}\right|^T_i \tag{11}$$

and noting that the term

$$\left.\frac{\partial f}{\partial x}\right|_i P_{x_i}(+) \left.\frac{\partial f}{\partial x}\right|^T_i \tag{12}$$

is of rank m. This, of course, assumes that 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. For $P_{y_i}(+)$ to be full rank, $C_i(+)$ must contain at least $n-m$ column vectors that are linearly independent of the column vectors of Eq. (12). Hence, there is the requirement

$$\text{rank}(C_i(+)) \geq n - m \tag{13}$$

Postmultiplying Eq. (7) by the matrix $P^{-1}_{y_i}(+)[\partial f/\partial x]|_i$ and using Eq. (5) result in

$$C_i(+) P^{-1}_{y_i}(+) \left.\frac{\partial f}{\partial x}\right|_i = 0 \tag{14}$$

Equation (14) shows that the m columns of $P^{-1}_{y_i}(+)\partial f/\partial x|_i$ are in the nullspace of $C_i(+)$.

The rank of the matrix product $P^{-1}_{y_i}(+)\partial f/\partial x|_i$ is, therefore, less than or equal to m [$P^{-1}_{y_i}(+)$ must be full rank and the rank of $\partial f/\partial x|_i$ is m]. If the rank of this product was less than m, however, then there would exist some vector $\boldsymbol{a} \neq 0$ such that $P^{-1}_{y_i}(+)\partial f/\partial x|_i \boldsymbol{a} = 0$. Postmultiplying the partial derivative matrix, which must have m linearly independent columns, by a nonzero vector $\boldsymbol{a}$ will give another

nonzero vector $\boldsymbol{b}$; $(\partial f/\partial x)|_i \boldsymbol{a} = \boldsymbol{b} \neq 0$. This results in the contradiction $P_{y_i}^{-1}(+)\boldsymbol{b} = 0$. Therefore, the matrix product $P_{y_i}^{-1}(+)\partial f/\partial x|_i$ 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.

This finding, combined with the fundamental relationship between the dimension of the nullspace and the rank for any $n \times n$ matrix and Eq. (13), results in the conclusion that the rank of $C(+)$ is exactly $n - m$.

Equation (6), in terms of C, reduces to

$$C_{i+1}(-) = C_i(+) \tag{15}$$

Therefore, this property of the rank of C applies to both the a priori as well as the a posteriori conditions

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

Property 2: The column space of C is fixed in R^n for a filter with $l \leq m$ after the large initial state error has been removed. Consider that the nullspace and row space of a matrix are orthogonal complements. It will be shown that the nullspace of C remains fixed; therefore, the row space of C (and, hence, the columns space because C is symmetric) must also be fixed.

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

$$N_i(\pm) \equiv P_{y_i}^{-1}(\pm)\frac{\partial f}{\partial x}\bigg|_i \tag{17}$$

The nullspace of C remains the same between measurement updates, as shown in Eq. (15). This can be written as

$$N_{i+1}(-) = N_i(+)A_{i+1} \tag{18}$$

in which A_{i+1} is a matrix 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 $i+1$ time step is given by Eq. (3). Postmultiplying Eq. (3) by $(\partial f/\partial x)|_{i+1}$, expressing this in terms of the $N_{i+1}(\pm)$ matrix defined in Eq. (17), and substituting in Eq. (18) result in

$$N_{i+1}(+) = N_i(+)A_{i+1} + H^T R_{i+1}^{-1} H \frac{\partial f}{\partial x}\bigg|_{i+1} \tag{19}$$

Consider the general condition where H^T does not initially lie in the column space of $N_i(+)$. As each new measurement is processed, a new term of the form $H^T R^{-1} H(\partial f/\partial x)$ is added to $N_i(+)A_{i+1}$. The only direction in which the column space of N can be changed is in the direction of the span of the columns 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. These arguments, of course, do not rigorously prove this property. Numerical simulations run on models with $l \leq m$ all have produced a C matrix with a column space that stays fixed after the initial state transients are reduced and a nullspace that contains the columns of H^T.

Once the column space of $N_i(+)$ contains the column space of H^T as a subspace then the columns of H^T can 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 $n \times m - l$ matrix $\bar{N}$. The matrix $N_i(+)$ can, therefore, be expressed as the linear combination

$$N_i(+) = \bar{N}B_i + H^T D_i \tag{20}$$

Substituting this into Eq. (19) gives

$$N_{i+1}(+) = \bar{N}B_i A_{i+1} + H^T \left[D_i A_{i+1} + R_{i+1}^{-1} H \frac{\partial f}{\partial x}\bigg|_{i+1} \right] \tag{21}$$

This shows that 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 $C_i(+)$. Therefore, if $l \leq m$ and H^T is in the nullspace of C, then the column space of C remains fixed in R^n.

For the case of $l > m$, the same arguments just used can be used to show that the columns of H^T are the only modifications that are possible to the column space of C. However, 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

$$H^T R_i^{-1} H \frac{\partial f}{\partial x}\bigg|_i \tag{22}$$

does not stay fixed because of the partial derivative factor. It is a changing m-dimensional subspace of a fixed l-dimensional space. The basis of the nullspace of C would be determined by the accumulation of terms like Eq. (22) 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 that is of interest, not the actual elements of the C matrix. For the remainder of this paper the notation $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{n-m}\}$ will be used to indicate the (fixed for $l \leq m$) basis of 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 is assumed that P_y^{-1} still exists and that the filter will never actually generate a singular P_y. The test is one that sets a numerical tolerance on how close to singular P_y is allowed to become. This is derived starting with Eq. (11).

As mentioned before, Eq. (11) expresses a matrix that must be of rank n as the sum of a matrix of rank $n - m$ and one of rank m. This is possible as long as the $n - m$ basis vectors $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{n-m}\}$ are linearly independent of the column space of $(\partial f/\partial x)P_x(\partial f/\partial x)^T$. Now, consider the case where at least one basis vector of C is close to being linearly dependent on the column space of $(\partial f/\partial x)P_x(\partial f/\partial x)^T$. This can be simplified by removing the first-step state covariance P_x because 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 $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{n-m}\}$ become close to being linearly dependent upon the set of m column vectors $\{\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_m\}$.

All of this can be expressed as a test of the numerical rank of an $n \times n$ matrix

$$\text{rank}\left(\left[\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{n-m}, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\right], \epsilon\right) < n \tag{23}$$

for some tolerance ϵ. Points for which the condition described in Eq. (23) is true are the locations where an ill-conditioned first-step covariance matrix would occur. The numerical rank of an $n \times n$ matrix is defined as the number of singular values greater than ϵ (Ref. 5). The span of $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_{n-m}\}$ is fixed (property 2) and the partial derivative matrix can be computed along a reference trajectory (assumption 1). Therefore, the test defined in Eq. (23) 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. The importance of Eq. (23) in this analysis, however, is that it leads to a geometric interpretation of the cause of ill-conditioned P_y.

Geometric Interpretation

The condition in which Eq. (23) would be true is visualized geometrically using a system of two second-step states and three first-step states ($n = 3$ and $m = 2$), as shown in Fig. 1. The covariance matrix P_x is represented by the two-dimensional error ellipse in the

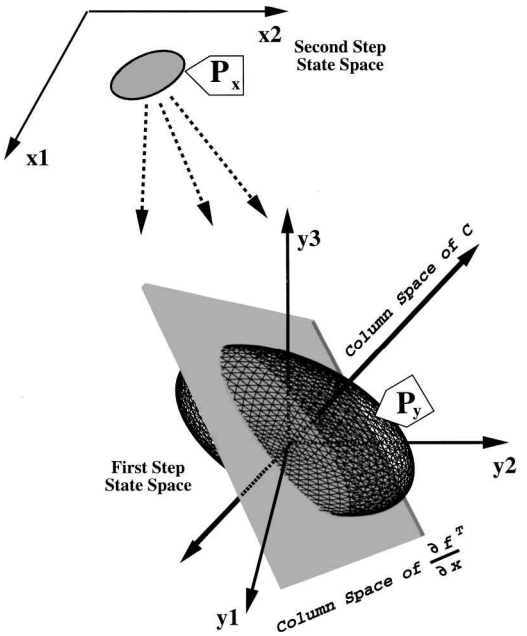


Fig. 1 Geometric interpretation of the column space of P_y.

x_1–x_2 space. The effect of Eq. (12) 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 Fig. 1. The effect of Eq. (11) is to add the additional basis vector from C to the column space of P_y.

The second property of C states that the orientation of this vector remains fixed in y_1–y_2–y_3 space. The plane defined by the column space of $(\partial f/\partial x)$, however, changes orientation with the state vector evolution. This line and plane together span R^3 under nondegenerate conditions. The orientation of the plane with respect to the line controls the linear independence of the columns of the P_y matrix as described by Eq. (23).

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 one direction in the y_1–y_2–y_3 space, namely, perpendicular to the column space of $(\partial f/\partial x)$, in which the first-step state estimate is predicted to have very little uncertainty.

IV. Numerical Example

A simple two-state example problem is used to illustrate this anomaly in the two-step filter. This example problem did not exhibit a catastrophic numerical failure as a result of ill-conditioned P_y. It did, however, produce an ill-conditioned covariance matrix in a similar manner to the more complex orbital mechanics problem. Although the existence of low eig(P_y) is still very evident, these eigenvalues never get small enough to cause a failure of the second-step iteration. This example is used to demonstrate how the filter structure gives rise to the low-rank P_y and to demonstrate the use of Eq. (23) to predict the location of these points. The geometry of this problem is shown in Fig. 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 set of two position coordinates x_1 and x_2 is sufficient to determine the particle motion and, therefore, is used as the two-dimensional second-step state vector. The nonlinear differential equations in Cartesian coordinates are

$$\frac{dx_1}{dt} = \frac{x_1 v_0}{\sqrt{x_1^2 + x_2^2}} - x_2\omega_0 \tag{24}$$

$$\frac{dx_2}{dt} = \frac{x_2 v_0}{\sqrt{x_1^2 + x_2^2}} + x_1\omega_0 \tag{25}$$

Table 1 Numerical data for example problem

Independent variable step, Δt	0.0002
Filter initial state, $\hat{x}_0$	{2.0, 0.0}
True initial state, $\boldsymbol{x}(0)$	{2.583, 0.313}
Measurement covariance, R	10^{-4}
Filter initial second-step state covariance, P_{x0}	diag{0.25, 0.25}
Filter initial first-step state covariance, P_{y0}	$\begin{bmatrix} 0.226 & 0.209 & 0.000 \\ 0.209 & 0.247 & 0.000 \\ 0.000 & 0.000 & 0.248 \end{bmatrix}$
Second-step state discrete time process noise, Q_d	diag{10^{-12} 10^{-12}}

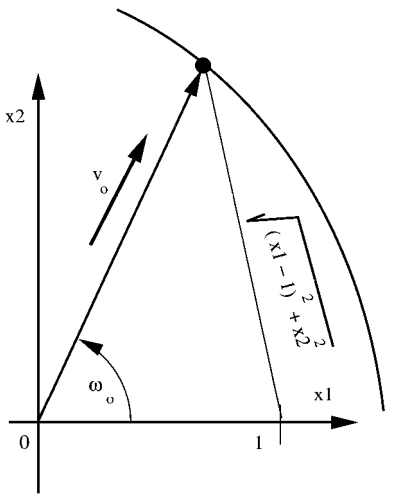


Fig. 2 Example problem geometry.

The second-step covariance matrix is propagated between measurements using the state transition matrix Φ and a discrete process noise term Q_{d_i},

$$P_{x_{i+1}}(-) = \Phi_{i+1,i} P_{x_i}(+) \Phi_{i+1,i}^T + Q_{d_i} \tag{26}$$

The state transition matrix is obtained by numerically integrating $\dot{\Phi} = A(\hat{x}, t)\Phi$ from t_i to t_{i+1} along with the predicted state estimate using the partial derivative matrix of the dynamics

$$A(\hat{x}, t) \equiv \left.\frac{\partial g(\boldsymbol{x}(t), t)}{\partial x}\right|_{\boldsymbol{x}(t)=\hat{x}(t)} \tag{27}$$

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

$$z = \left[\sqrt{(x_1 - 1)^2 + x_2^2}\right] \tag{28}$$

The first-step state vector is defined as the second-step state vector augmented by the measurement equation, following the suggestion in Ref. 1,

$$\boldsymbol{y} = \begin{bmatrix} \sqrt{(x_1 - 1)^2 + x_2^2} \\ x_1 \\ x_2 \end{bmatrix} \tag{29}$$

Therefore, $H = [1\ 0\ 0]$.

Specific numbers used in the example and the filter implementations are 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 prevent the filter from losing sensitivity to new measurements as $t \to \infty$ and $P_{y,x} \to 0$. No dynamic noise, however, is simulated in the truth model. A plot of the particle motion and one set of noisy measurements is shown in Fig. 3. A large number (30,000) of data points are generated to illustrate the sharp decreases

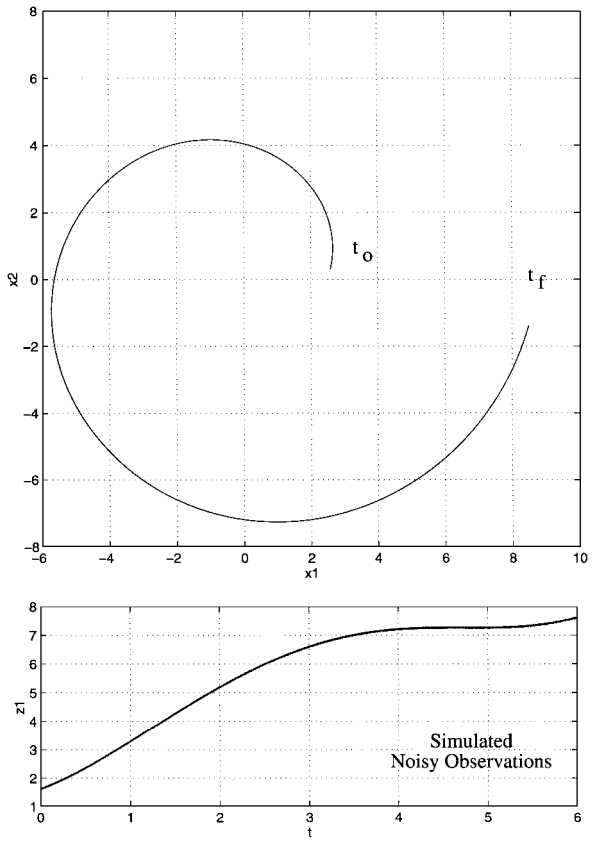


Fig. 3 Example problem state and observations.

in eigenvalues of P_y that can occur over a very small time period. It was found that spacing points farther apart lessens the decrease in the eigenvalues.

Eigenvalues of the $P_y(+)$ matrix on the order of 10^{-15} occur near $t \approx 1.3$ and 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 is used as the basis of the column space of C. Figure 4 plots the numerical values demonstrating the properties of C. The three singular values of C [svd(C)] are plotted in Fig. 4a. Note that the first singular values are 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 $\boldsymbol{c}(\infty)$, computed as the average of the last 5000 and each previous $\boldsymbol{c}_i$, is plotted in Fig. 4b. This dot product is very nearly unity once the large initial state error has been reduced by the filter. Hence, it has been numerically demonstrated that the column space of C is fixed (property 2).

The explanation given in this study for the constant basis of the column space of C is shown in Fig. 4c. In this figure the dot product between the H matrix and the basis vector $\boldsymbol{c}$ is plotted. After $t \approx 0.5$, this product becomes very small in comparison to the magnitude of the H^T and $\boldsymbol{c}$ vectors (both unity) indicating that the two vectors are orthogonal and, therefore, the column vector H^T is in the nullspace of C.

Figure 5 shows the rank test defined in Eq. (23) using a numerical tolerance of 10^{-4}, and Fig. 5a shows a plot of the eigenvalues of $P_y(+)$. This test matrix is computed from the steady-state eigenvector of C and the partial derivative matrix computed along the true trajectory. Therefore, other than the determination of the steady-state basis vector $\boldsymbol{c}$, no output from the filter is used in the calculation of the numerical rank. The test correctly predicts the location of low eigenvalues in the P_y time history, as shown in Fig. 5. The use of a numerical definition of matrix rank to resolve the apparent contradiction in using P_y^{-1} to derive properties of the C matrix, which are then used to look for conditions under which P_y^{-1} does not exist, is demonstrated in these results. The smallest singular values of C are 10^{-18}, the size of the second largest singular value, as shown on Fig. 4a. This numerical tolerance is used to determine the rank of

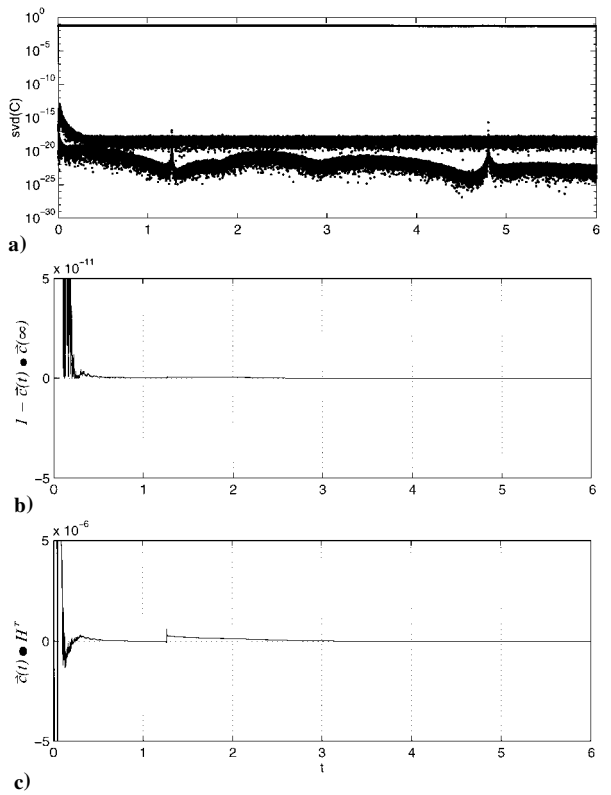
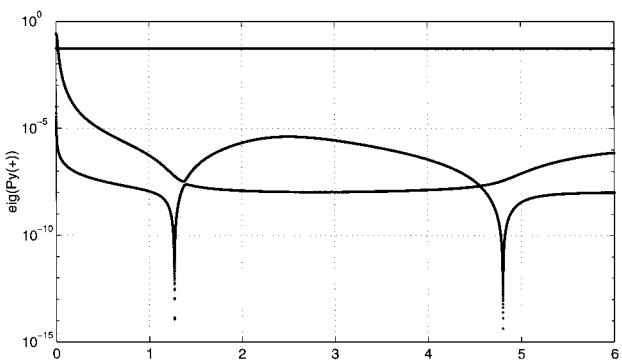
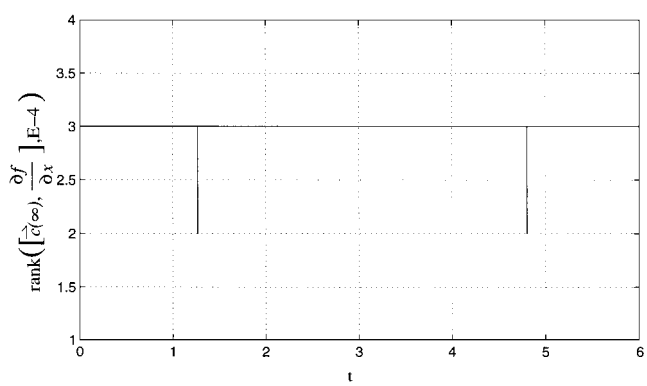


Fig. 4 Properties of the column space of C.



a) Eigenvalues of first-step state covariance matrix



b) Demonstration of Eq. (23) in predicting where P_y drops rank

Fig. 5 Numerical rank test for ill-conditioned P_y.

C and, hence, its properties. It is much smaller than the numerical tolerances used to identify a low-rank P_y (in this example 10^{-4}) using Eq. (23), as plotted in Fig. 5. Therefore, the properties of C are expected to be numerically valid even under situations in which P_y is determined to be ill conditioned. The C matrix may, however, change if the filter generates large state errors as the result of numerical problems caused by the ill-conditioned P_y. This is not a serious concern in these examples because the partial derivatives for the test

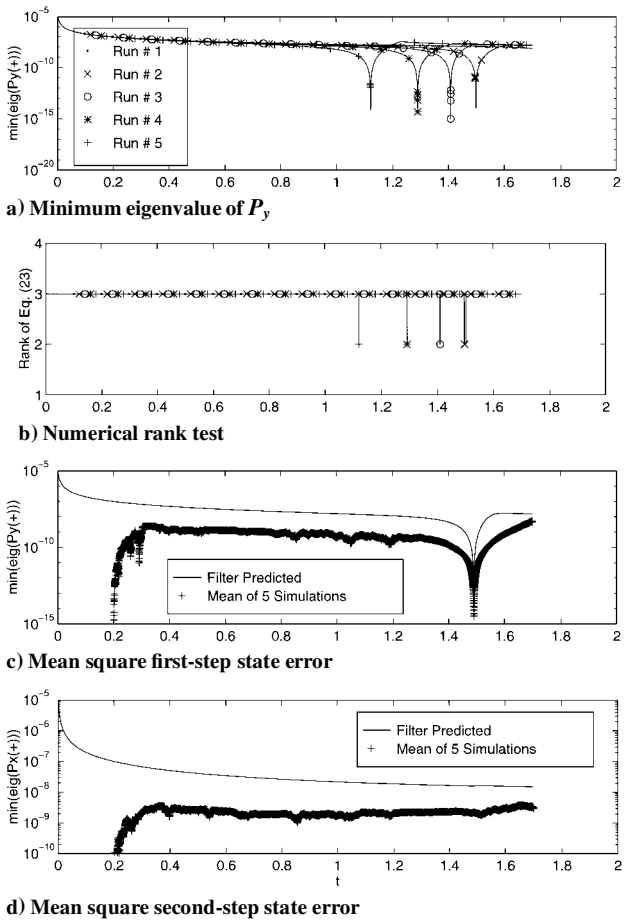


Fig. 6 Comparison of minimum eig(P_y) location with mean squared first- and second-step state error for five simulations with measurement noise.

matrix are computed using the true state time history. For reference, the floating point relative accuracy on the computer used in these simulations is 2.2×10^{-16}. This is defined as the difference between 1.0 and the next largest floating point number.

It is now demonstrated through numerical simulation that this anomaly represents an actual instance of first-step states with very low uncertainty in one direction and, furthermore, that this does not have a similar effect on the estimation error in the corresponding second-step state estimates. A series of five simulations of the example problem was done for 8500 points, each starting from exactly the same initial conditions as given in Table 1. The random measurement errors produce a slightly different state estimate time history for each simulation. The minimum eigenvalue of $P_y(+)$ for each of these five simulations is shown in Fig. 6a. A different C matrix column vector $\boldsymbol{c}$ was computed from the average of the last 5000 points of each run and used in the test given in Eq. (23), with the partial derivatives computed along the true trajectory and a tolerance of $\epsilon = 10^{-4}$. Again, the only numerical datum obtained from the filter itself is the C matrix. The results of this test are plotted in Fig. 6b. What this shows is that the wide difference in location of the points with low eigenvalues is the result of the different orientation of the column space of C depending on the state estimate time history. Recall that Eqs. (8) and (9), which are necessary for C matrix to remain fixed, are only approximations. These are not valid at the start of each filter simulation but rapidly become so, as the plots in Fig. 4 show. The different measurement and state estimate time histories for the five simulations in Fig. 6, as a result of the random measurement noise, would cause the C matrix to take a different path to this converged orientation for each run. A different column space basis vector $\boldsymbol{c}$ would, therefore, be generated for each simulation.

This point is further emphasized by the plot in Fig. 6c. The filter was run five times with exactly the same initial conditions but with perfect measurements for the first 1001 points. This forces the C matrix to converge to the same orientation in each of the simulations before noisy data are encountered. On this plot two curves are shown: the minimum eigenvalue of the mean of predicted $P_y(+)$ matrices from the five runs and the minimum eigenvalue of the mean squared error $(\frac{1}{5})\sum^5(\hat{y}-\boldsymbol{y})(\hat{y}-\boldsymbol{y})^T$. Both of these plots go through a minimum at the same point. This indicates that the anomaly discovered for the two-step estimator actually does predict a location where some linear combination of the first-step state mean squared error reaches very low values. This, however, does not result in any corresponding linear combination of the second-step state estimates, which have very low error, as Fig. 6d shows. In Fig 6d, the two eigenvalues of the mean predicted $P_x(+)$ are plotted along with the eigenvalues of the mean squared second step state error $(\frac{1}{5})\sum^5(\hat{x}-\boldsymbol{x})(\hat{x}-\boldsymbol{x})^T$. The eigenvalues of $P_x(+)$ do not exhibit a similar minimization at that location, and so it can be concluded that this anomaly only represents a feature of the set of (underdetermined) first-step states. It does not imply that a linear combination of second-step states exists that has very small error.

V. Filter Modifications

Two straightforward modifications are suggested for the two-step filter that can mitigate the effect of the ill-conditioned P_y. The first of these is the addition of a small positive diagonal matrix (ϵI) similar to process noise onto the first-step time update in Eq. (6),

$$P_{y_{i+1}}(-) = P_{y_i}(+) + \frac{\partial f}{\partial x}\bigg|_{\hat{x}_{i+1}(-)} P_{x_{i+1}}(-)\frac{\partial f}{\partial x}\bigg|^T_{\hat{x}_{i+1}(-)} - \frac{\partial f}{\partial x}\bigg|_{\hat{x}_i(+)} P_{x_i}(+)\frac{\partial f}{\partial x}\bigg|^T_{\hat{x}_i(+)} + \epsilon I \tag{30}$$

The small positive scalar ϵ can be chosen large enough to prevent P_y from becoming so ill conditioned that it causes the second-step iteration to fail, but small enough so that it does not significantly effect the filter covariance propagation. As already mentioned, addition of process noise in the conventional manner to the second-step state propagation would not effect the possibility that ill-conditioned covariance matrices would be generated.

The other suggested modification is to monitor some indication of the numerical rank of P_y, such as the rank itself, the value of the test condition given in Eq. (23) computed using a predicted state, or the ratio of smallest to largest diagonal in the UD factorization of the first-step covariance. When this indicates that the matrix P_y has dropped rank, then some number of new measurements are rejected until the predicted state has advanced past the location where low eigenvalues will occur.

This is related to the earlier statement that spacing of the data points farther apart reduces the chances of a failure. This is because the rank of P_y drops rapidly over a very short time period. The closer measurements are together, the greater the chance of encountering a P_y with low enough eigenvalues to cause the second step iteration to fail.

Use of the UD covariance factorization does reduce the numerical problems resulting from the small eigenvalues of P_y but it does not eliminate their existence nor even the possibility of negative eigenvalues. This is because the factored form of Eq. (6) contains a negative block diagonal from the third term in that equation. The UD covariance factorization of Eq. (6) is $P_{y_i}(-) = Y_y \tilde{D}_y Y_y^T$ in which

$$Y_y = \left[U_{y_{i-1}}(+) \;\vdots\; \frac{\partial f}{\partial x}\bigg|_{\hat{x}_i(-)} U_{x_i}(-) \;\vdots\; \frac{\partial f}{\partial x}\bigg|_{\hat{x}_{i-1}(+)} U_{x_{i-1}}(+) \right] \tag{31}$$

$$\tilde{D}_y = \begin{bmatrix} D_{y_{i-1}}(+) & \vdots & 0 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & D_{x_i}(-) & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & 0 & \vdots & -D_{x_{i-1}}(+) \end{bmatrix} \tag{32}$$

and the U and D matrices in the block diagonals are the $U\,D$ factors of the three covariance matrices on the right-hand side of Eq. (5).

Another possibility is to allow the existence of ill-conditioned P_y and to develop a better second-step parameter estimation that is robust enough to still produce good estimates of the first-step state. This option must also allow for the potential of small negative eigenvalues of P_y, possibly by setting them equal to zero.

These modifications to the filter are suggested to overcome the numerical difficulties resulting from the occurrence of an ill-conditioned first-step covariance matrix. They do not change the structure of the filter so as to avoid this occurrence altogether. It is not clear that there is any general guidance for selecting first-step states so as to avoid ill-conditioned first-step state covariances, or even if this would ever be possible. Reformulation of the two-step filter from the start with this objective in mind is beyond the scope of this study.

VI. Conclusion

The occurrence of ill-conditioned first-step covariance matrices in the two-step filter has been explained. This anomaly results when the two positive semidefinite matrices in the linear time update of the second-step covariance matrix do not combine to produce a full set of linearly independent column vectors. The presence of very low first-step eigenvalues is a result of the two-step filter formulation. Consequences of this for numerical implementation of the filter must be considered. The time update is an approximation, but the numerical simulations indicate that the locations of low eigenvalues are representative of a real condition of the first-step state estimation error.

The test derived correctly predicted the locations of ill-conditioned first-step covariance matrices given the true state history. Use of the test matrix in this manner to identify this problem in a real application of the filter is impossible (because the truth is unknown). This demonstration is done, however, to show that the difference in location of the ill-conditioned matrices between different samples of an ensemble of runs was determined only by the orientation of the column space of the C matrix. As mentioned, this test could conceivably be applied as an operational test as well, using the estimated state. The real use of the test matrix lies in the geometric interpretation that it gives for why the first-step covariance time update does not always provide a set of column vectors that are sufficiently independent to span the first-step state space.

References

[1] Haupt, G. T., Kasdin, N. J., Keiser, G. M., and Parkinson, B. W., "Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 3, 1996, pp. 643–649.

[2] Kasdin, N. J., and Haupt, G. T., "Second-Order Correction and Numerical Considerations for the Two-Step Optimal Estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 2, 1997, pp. 362–369.

[3] Garrison, J. L., "Recursive Nonlinear Estimation for Relative Navigation in Elliptical Orbits," Ph.D. Dissertation, Dept. of Aerospace Engineering Sciences, Univ. of Colorado, Boulder, CO, Dec. 1997.

[4] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes in C*, Cambridge Univ. Press, New York, 1994, pp. 683–685.

[5] "MATLAB Reference Guide," Mathworks, Inc., Natick, MA, Aug. 1992, p. 405.