Nonlinear Filtering Based On Sequential Model Error Determination

John L. Crassidis
Assistant Professor
Department of Mechanical Engineering
Catholic University of America
Washington, D.C. 20064
jlc@pluto.ee.cua.edu

F. Landis Markley
Staff Engineer
Guidance and Control Branch. Code 712
Goddard Space Flight Center
Greenbelt, MD 20771
Landis.Markley@gsfc.nasa.gov

Abstract

In this paper, a real-time predictive filter is derived for nonlinear systems. This provides a method of determining optimal state estimates in the presence of significant error in the assumed (nominal) model. The new real-time nonlinear filter determines (i.e., "predicts") the optimal model error trajectory so that the measurement-minus-estimate covariance statistically matches the known measurement-minus-truth covariance. The optimal model error is found by using a one-time step ahead control approach. Also, since the continuous model is used to determine state estimates, the filter avoids discrete state jumps, as opposed to the extended Kalman filter.

Introduction

Many linear control systems require full state knowledge (such as the LQR). But most systems have sensors which cannot measure all states (e.g., spacecraft sensors measure body positions which must be converted to an attitude for control purposes). An essential feature of most control systems involves an algorithm which is used to both estimate unmeasured states and to filter noisy measurements. Since pointing errors are a combination of both control and estimation errors, the robustness of each component must be addressed to insure proper pointing.

Conventional filter methods, such as the Kalman filter [1], have proven been to be extremely useful in a wide range of applications, including: noise reduction of signals, trajectory tracking of moving objects, and in the control of linear or nonlinear systems. The essential feature of the Kalman filter is the utilization of state-space formulations for the system model. Errors in the dynamic system model are treated as "process noise," since system models are not usually improved or updated during the estimation process. The process noise is essentially used to "shift" the emphasis from the model to the measurements.

The Kalman filter satisfies an optimality criterion which minimizes the trace of the covariance of the estimate error between the system model responses and actual measurements. Statistical properties of the process noise and measurement error are used to determine an "optimal" filter design. Therefore, model characteristics are combined with sequential measurements in order to obtain state estimates which meliorate both the measurements and model responses.

In the Kalman filter, the errors in the system model are assumed to be represented by a zero-mean Gaussian noise process with known covariance. However, in actual practice the noise covariance is usually determined by an *ad hoc* and/or heuristic estimation approach which may result in sub-optimal filter designs. Other applications also determine a steady-state gain directly, which may even produce unstable filter designs [2]. Also, in many cases such as nonlinearities in the actual system responses or non-stationary processes, the assumption of a Gaussian model error process can lead to severely degraded state estimates.

In addition to nonlinear model errors, the actual assumed model may be nonlinear (e.g., three-dimensional kinematic and dynamic equations [3]). The filtering problem for nonlinear systems is considerably more difficult and admits a wider variety of solutions than does the linear problem [4]. The extended Kalman filter is a widely used algorithm for nonlinear estimation and filtering [5]. The essential feature of this algorithm is the utilization of a first-order Taylor series expansion of the model and output system equations. The extended Kalman filter retains the linear calculation of the covariance and gain matrices, and it updates the state estimate using a linear function of the measurement residual; however, it uses the original nonlinear equations for state propagation and in the output system equation [5]. But, the model error statistics are still assumed to be represented by a zero-mean Gaussian noise process.

A new approach for performing optimal state estimation in the presence of significant model error has been developed by Mook and Junkins [6]. This algorithm, called the Minimum Model Error (MME) estimator, unlike most filter and smoother algorithms, does not assume that the model error is represented by a Gaussian process. Instead, the model error is determined during the MME estimation process. The algorithm determines the corrections added to the assumed model such that the model and corrections yield an accurate representation of the system behavior. This is accomplished by solving system optimality conditions and an output error covariance constraint. Therefore, accurate state estimates can be determined without the use of precise system representations in the assumed model. Also, the MME estimator can be applied to systems with nonlinear models. The MME estimates are determined from a solution of a two-point-boundary-value-problem ([6-7]). Therefore, the MME estimator is a batch (off-line) estimator which must utilize post-experiment measurements.

The filter algorithm developed in this paper can be implemented in real-time (as can the Kalman filter). However, the algorithm is not limited to Gaussian noise characteristics for the model error. Essentially, this new algorithm combines the good qualities of both the Kalman filter (i.e., a real-time estimator) and the MME estimator (i.e., determines actual model error trajectories). The new algorithm is based on a predictive tracking scheme first introduced by Lu [8]. Although the problem shown in [8] is solved from a control standpoint, the algorithm developed in this paper is reformulated as a filter and estimator with a stochastic measurement process. Therefore, the new algorithm is known as a predictive filter. The advantages of the new algorithm include: (i) the model error is assumed unknown and is estimated as part of the solution, (ii) the model error may take any form (even nonlinear), and (iii) the algorithm can be implemented on-line to both filter noisy measurements and estimate state trajectories.

The organization of this paper proceeds as follows. First, the basic equations and concepts used for the filter development are reviewed. Then, a predictive filter is derived for nonlinear systems. This approach determines optimal state estimates in real-time by minimizing a quadratic cost function consisting of a measurement residual term and a model error term. Then, the concept of the covariance constraint is introduced for determining the optimal model error weighting matrix. Finally, conclusions are summarized for the new algorithm.

Nonlinear Predictive Filter

Preliminaries

In this section, the nonlinear predictive filter algorithm is derived. This development is based upon the duality which exists between the predictive controller for nonlinear systems by Lu [8] and a general estimation problem. In the nonlinear predictive filter it is assumed that the state and output estimates are given by a preliminary model and a to-be-determined model error vector, given by

$$\dot{\hat{\underline{x}}}(t) = \underline{f}\left(\hat{\underline{x}}(t)\right) + G\left(\hat{\underline{x}}(t)\right)\underline{d}(t) \tag{1a}$$

$$\hat{\underline{y}}(t) = \underline{c}\left(\hat{\underline{x}}(t)\right) \tag{1b}$$

where $\underline{f} \in R^n$ is sufficiently differentiable, $\hat{\underline{x}}(t) \in R^n$ is the state estimate vector, $\underline{d}(t) \in R^q$ represents the model error vector, $G\left(\hat{\underline{x}}(t)\right): R^n \to R^{n\times q}$ is the model-error distribution matrix, $\underline{c}\left(\hat{\underline{x}}(t)\right) \in R^m$ is the measurement vector, and $\hat{\underline{y}}(t) \in R^m$ is the estimated output vector. State-observable discrete measurements are assumed for Equation (1b) in the following form

$$\tilde{\underline{y}}_k = \underline{c}\left(\underline{x}(t_k)\right) + \underline{v}_k \tag{2}$$

where $\tilde{\underline{y}}_k \in R^m$ is the measurement vector at time t_k, $\underline{x}(t_k)$ is the true state vector, and $\underline{v}_k \in R^m$ represents the measurement noise vector which is assumed to be a zero-mean, Gaussian white-noise distributed process with

$$E\left\{\underline{v}_k\right\} = \underline{0} \tag{3a}$$

$$E\left\{\underline{v}_k \underline{v}_l^T\right\} = R\,\delta_{kl} \tag{3b}$$

where R is a $m \times m$ positive-definite measurement covariance matrix.

A Taylor series expansion of the output estimate in Equation (1b) is given by

$$\hat{\underline{y}}(t+\Delta t) \approx \hat{\underline{y}}(t) + \underline{z}\left(\hat{\underline{x}}(t), \Delta t\right) + \Lambda(\Delta t)\, S\left(\hat{\underline{x}}(t)\right)\underline{d}(t) \tag{4}$$

where the i^{th} row of $\underline{z}\left(\hat{\underline{x}}(t), \Delta t\right)$ is given by

$$z_i\left(\hat{\underline{x}}(t), \Delta t\right) = \sum_{k=1}^{p_i} \frac{\Delta t^k}{k!} L_f^k\left(c_i\right) \tag{5}$$

where p_i, $i = 1, 2, \ldots, m$, is the lowest order of the derivative of $c_i\left(\hat{\underline{x}}(t)\right)$ in which any component of

$\underline{d}(t)$ first appears due to successive differentiation and substitution for $\hat{\dot{x}}_i(t)$ on the right side. $L_f^k(c_i)$ is a k^{th} order Lie derivative, defined by

$$L_f^k(c_i) = c_i \qquad \text{for } k = 0$$
$$L_f^k(c_i) = \frac{\partial L_f^{k-1}(c_i)}{\partial \underline{\hat{x}}} \underline{f} \qquad \text{for } k \geq 1 \tag{6}$$

$\Lambda(\Delta t)$ is a $m \times m$ diagonal matrix with elements given by

$$\lambda_{yii} = \frac{\Delta t^{p_i}}{p_i!}, \quad i = 1, 2, \ldots, m \tag{7}$$

$S(\underline{\hat{x}}(t))$ is a $m \times q$ matrix with each i^{th} row given by

$$s_i = \left\{ L_{g_1}\left[L_f^{p_i-1}(c_i)\right], \ldots, L_{g_q}\left[L_f^{p_i-1}(c_i)\right] \right\}, \quad i = 1, 2, \ldots, m \tag{8}$$

where the Lie derivative with respect to g_j in Equation (8) is defined by

$$L_{g_j}\left[L_f^{p_i-1}(c_i)\right] \equiv \frac{\partial L_f^{p_i-1}(c_i)}{\partial \underline{\hat{x}}} g_j, \quad j = 1, 2, \ldots, q \tag{9}$$

Nonlinear Filtering

A cost functional consisting of the weighted sum square of the measurement-minus-estimate residuals plus the weighted sum square of the model correction term is minimized, given by

$$J(\underline{d}(t)) = \frac{1}{2}\underline{d}^T(t) W \underline{d}(t) + \frac{1}{2}\left\{\underline{\tilde{y}}(t+\Delta t) - \underline{\hat{y}}(t+\Delta t)\right\}^T R^{-1} \left\{\underline{\tilde{y}}(t+\Delta t) - \underline{\hat{y}}(t+\Delta t)\right\} \tag{10}$$

where W is a $q \times q$ matrix. Also, a constant sampling rate is assumed so that $\underline{\tilde{y}}(t+\Delta t) \equiv \underline{\tilde{y}}_k$. Substituting Equation (4) into (10), and minimizing Equation (10) with respect to $\underline{d}(t)$ leads to the following model error solution

$$\underline{d}(t) = -\left\{\left[\Lambda(\Delta t) S(\underline{\hat{x}})\right]^T R^{-1} \Lambda(\Delta t) S(\underline{\hat{x}}) + W\right\}^{-1} \left[\Lambda(\Delta t) S(\underline{\hat{x}})\right]^T R^{-1} \left[\underline{z}(\underline{\hat{x}}, \Delta t) - \underline{\tilde{y}}(t+\Delta t) + \underline{\hat{y}}(t)\right] \tag{11}$$

where $\underline{\tilde{y}}(t+\Delta t)$ the measurement sampled at a constant interval (Δt). By using the matrix inversion lemma [9], the model error in Equation (11) can be re-written as

$$\underline{d}(t) = -K(t)\left[\Lambda(\Delta t) S(\underline{\hat{x}})\right]^T R^{-1} \times \left[\underline{z}(\underline{\hat{x}}, \Delta t) - \underline{\tilde{y}}(t+\Delta t) + \underline{\hat{y}}(t)\right] \tag{12}$$

where

$$K(t) = W^{-1}\left(I - \left[\Lambda(\Delta t) S(\underline{\hat{x}})\right]^T \left\{\Lambda(\Delta t) S(\underline{\hat{x}}) W^{-1} \left[\Lambda(\Delta t) S(\underline{\hat{x}})\right]^T + R\right\}^{-1} \left[\Lambda(\Delta t) S(\underline{\hat{x}})\right] W^{-1}\right) \tag{13}$$

Equation (11) or (12) is used in Equation (1a) to perform a nonlinear propagation of the state estimates to time t_k, then the measurement is processed at time t_{k+1} to find the new $\underline{d}(t)$ in $\left[t_k, t_{k+1}\right]$, and then the state estimates are propagated to time t_{k+1}. The matrix W serves to weight the amount of model error added to correct the assumed model in Equation (1). As W decreases, more model error is added to correct the model, so that the estimates more closely follow the measurements. As W increases, less model error is added, so that the estimates more closely follow the propagated model.

Covariance Constraint

The $q \times q$ weighting matrix (W) in Equation (11) can be determined on the basis that the measurement-minus-estimate error covariance matrix must match the measurement-minus-truth error covariance matrix (see [6]). This condition is referred to as the "covariance constraint," shown as

$$\frac{1}{N}\sum_{k=0}^{N}\left\{\underline{e}_k - \underline{\bar{e}}\right\}\left\{\underline{e}_k - \underline{\bar{e}}\right\}^T \approx R \tag{14}$$

where $\underline{\bar{e}}$ is the sample mean of $\underline{\tilde{y}} - \underline{\hat{y}}$, and N is a large number. A test for whiteness can be based upon the autocorrelation function matrix of the measurement residual [5]. The maximum likelihood estimate of the $m \times m$ autocorrelation function matrix for N samples is given by

$$C_k = \frac{1}{N}\sum_{i=k}^{N} \underline{e}_i \underline{e}_{i-k}^T \tag{15}$$

A 95% confidence interval for whiteness using a finite sample length is given by [5]

$$\left|\rho_{iik}\right| \leq 1.96 / N^{1/2} \tag{16}$$

where ρ_{ii} corresponds to the diagonal elements resulting by normalizing the autocorrelation matrix by the zero-lag elements, given by

$$\rho_{iik} = \frac{c_{ii_k}}{c_{ii_0}} \quad (17)$$

If the confidence interval in Equation (16) and the covariance constraint in Equation (14) are met, then the weighting matrix is optimal. Therefore, the proper balance between model error and measurement residual has been achieved. If the measurement residual covariance is higher than the known measurement error covariance (R), then W should be decreased to less penalize the model error. Conversely, if the residual covariance is lower than the known covariance, then W should be increased so that less unmodeled dynamics are added to the assumed system model.

The sample measurement covariance can be determined from a recursive relationship given by [10]

$$\hat{R}_{k+1} = \hat{R}_k + \frac{1}{k+1}\left[\frac{k}{k+1}\left(\underline{e}_{k+1} - \underline{\bar{e}}_k\right)\left(\underline{e}_{k+1} - \underline{\bar{e}}_k\right)^T - \hat{R}_k\right] \quad (18a)$$

$$\underline{\bar{e}}_{k+1} = \underline{\bar{e}}_k + \frac{1}{k+1}\left(\underline{e}_{k+1} - \underline{\bar{e}}_k\right) \quad (18b)$$

The covariance constraint is met when $\hat{R}_k \to R$.

Even though the model error is determined by Equation (11) or (12), it still involves stochastic processes. Therefore, a covariance of the model error can be derived. First, the covariance constraint is re-written as

$$E\left\{\left(\underline{\tilde{y}}_k - \underline{\hat{y}}_k\right)\left(\underline{\tilde{y}}_k - \underline{\hat{y}}_k\right)^T\right\} = R \quad (19)$$

Substituting Equation (2) into Equation (19), and using

$$E\left\{\underline{y}_k \underline{v}_k^T\right\} = E\left\{\underline{v}_k \underline{y}_k^T\right\} = E\left\{\underline{\hat{y}}_k \underline{v}_k^T\right\} = E\left\{\underline{v}_k \underline{\hat{y}}_k^T\right\} = 0 \quad (20)$$

leads to

$$E\left\{\underline{\tilde{y}}_k \underline{\tilde{y}}_k^T\right\} - \underline{\hat{y}}_k \underline{\hat{y}}_k^T = R \quad (21)$$

If Equation (14) is met, the process is stationary so that

$$E\left\{\underline{\tilde{y}}_{k+1} \underline{\tilde{y}}_{k+1}^T\right\} - \underline{\hat{y}}_{k+1} \underline{\hat{y}}_{k+1}^T = R \quad (22)$$

For constant a sampling interval, Equation (22) is equivalent to

$$E\left\{\underline{\tilde{y}}(t+\Delta t)\underline{\tilde{y}}^T(t+\Delta t)\right\} = \underline{\hat{y}}(t+\Delta t)\underline{\hat{y}}^T(t+\Delta t) + R \quad (23)$$

As long as the process remains stationary, Equation (23) is valid even if the covariance constraint is not satisfied. Also, since the optimal model error solution in Equation (11) is a function of the stochastic measurement noise process, a test for the whiteness of the "determined" model error can be found by using the correlation function in Equations (15)-(17), replacing $\underline{e}$ with $\underline{d}$. If the model error is sufficiently white, then the covariance of the model error can also be determined using a recursive formula shown in Equation (18), again replacing $\underline{e}$ with $\underline{d}$. Another form for the model error covariance can be determined by using Equation (11), and assuming that

$$E\left\{\underline{\hat{y}}(t)\underline{v}^T(t+\Delta t)\right\} = E\left\{\underline{v}(t+\Delta t)\underline{\hat{y}}^T(t)\right\} = 0 \quad (24a)$$

$$E\left\{\underline{z}(t)\underline{v}^T(t+\Delta t)\right\} = E\left\{\underline{v}(t+\Delta t)\underline{z}^T(t)\right\} = 0 \quad (24b)$$

which leads to

$$E\left\{\underline{d}(t)\underline{d}^T(t)\right\} = M(t)\left\{\left\langle\underline{\hat{y}}(t) - \underline{\hat{y}}(t+\Delta t) + \underline{z}(t)\right\rangle + R\right\}M^T(t) \quad (25)$$

where

$$\langle\underline{a}\rangle \equiv \underline{a}\,\underline{a}^T \qquad \text{for any } \underline{a} \quad (26)$$

Therefore, the relative magnitude of the model error can now be determined. In fact, if the determined model error process is truly white, then the inverse of the weighting matrix (W) can be shown to be the maximum likelihood estimate of the model error.

Stability

Filter Stability

The effect of W on filter stability and bandwidth can be determined by applying a discrete error analysis. The filter residual is given by

$$\underline{e}(t+\Delta t) = \underline{\tilde{y}}(t+\Delta t) - \underline{\hat{y}}(t+\Delta t) \quad (27)$$

Substituting Equation (4) into Equation (27) leads to

$$\underline{e}(t+\Delta t) = \underline{\tilde{y}}(t+\Delta t) - \underline{\hat{y}}(t) - \underline{z}(\underline{\hat{x}},\Delta t) + \underline{\lambda}(t) \quad (28)$$

where

$$\underline{\lambda}(t) = M(t)\left[\underline{z}(\underline{\hat{x}},\Delta t) - \underline{\tilde{y}}(t+\Delta t) + \underline{\hat{y}}(t)\right] \quad (29)$$

A Taylor series expansion of the output measurement is given by

$$\underline{\tilde{y}}(t+\Delta t) \approx \underline{\tilde{y}}(t) + \underline{b}(t,\Delta t) \quad (30)$$

where the i^{th} row of $\underline{b}(t,\Delta t)$ is given by

$$b_i(t,\Delta t) = \Delta t\,\dot{\tilde{y}}_i(t) + \frac{\Delta t^2}{2}\ddot{\tilde{y}}_i(t) + \cdots + \frac{\Delta t^{p_i}}{p_i!}\tilde{y}_i^{p_i}(t), \quad i = 1,2,\ldots,m \quad (31)$$

Equation (27) can now be re-written as

$$\underline{e}(t+\Delta t)=\left[I-M(t)\right]\underline{e}(t) \\ +\left[I-M(t)\right]\left[\underline{b}(t,\Delta t)-\underline{z}(\hat{\underline{x}},\Delta t)\right] \tag{32}$$

If S is square and full rank, then ΛS is also full rank. As $W\to 0$, then $M\to I$, and $(I-M)\to 0$. This approaches a deadbeat response for the filter dynamics. As $W\to\infty$, then $M\to 0$, and $(I-M)\to I$. This yields a filter response with eigenvalues approaching the unit circle. As long as the covariance matrix is positive, the eigenvalues of the filter will lie within the unit circle. Therefore, the filter remains contractive [11]. Also, the matrix $\left[I-M(t)\right]$ defines how the actual error at $(t+\Delta t)$ is reduced from the predicted error by $\underline{d}(t)$, given by

$$\underline{e}(t+\Delta t)=\left[I-M(t)\right]\hat{\underline{e}}(t+\Delta t) \tag{33}$$

where

$$\hat{\underline{e}}(t+\Delta t)\equiv\tilde{\underline{y}}(t+\Delta t)-\hat{\underline{y}}(t)-\underline{z}(\hat{\underline{x}},\Delta t) \tag{34}$$

which is the predicted measurement residual at $t+\Delta t$ assuming $\underline{d}=\underline{0}$.

In order to study the effects of system observability and numerical difficulties, consider the case where $q=3$, $W=wI$, $\Lambda=\lambda I$, and $R=rI$. Then, the matrix inverse in Equation (11) can be written as

$$\left\{\left[\Lambda S\right]^T R^{-1}\Lambda S+W\right\}^{-1}=\left\{(v-\sigma)I+C\right\}^{-1} \tag{35}$$

where

$$C\equiv\frac{\lambda^2}{r}S^T S \tag{36a}$$

$$\sigma=\frac{\lambda^2}{2r}\operatorname{tr}\left(S^T S\right) \tag{36b}$$

$$v=w+\sigma \tag{36c}$$

By the Cayley-Hamilton theorem, any meromorphic function of C can be expressed as a quadratic in C [11], yielding

$$\left\{(v-\sigma)I+C\right\}^{-1}=\frac{1}{\gamma}\left(\alpha I+\beta C+C^2\right) \tag{37}$$

where

$$\alpha=v^2-\sigma^2+k \tag{38a}$$

$$\beta=-(v+\sigma) \tag{38b}$$

$$\gamma=(v-\sigma)\alpha+\Delta \tag{38c}$$

$$k=\operatorname{tr}(\operatorname{adj}C) \tag{38d}$$

$$\Delta=\frac{\lambda^6}{r^3}\det\left(S^T S\right) \tag{38e}$$

If the system is unobservable then, $S^T S$ is not full rank. However, the filter can compensate for this by adding more model correction. It can be shown that the filter remains stable as long as

$$(v-\sigma)\alpha+\Delta>0 \tag{39}$$

If $S^T S$ is not full rank, then $\Delta=0$, which leads to the following condition

$$v>\frac{\lambda^2}{2r}\operatorname{tr}\left(S^T S\right) \tag{40}$$

Therefore, the filter remains contractive as long as $w>0$. This condition is always met, but Equation (40) can be used to help determine any numerical difficulties (i.e., large values of σ may produce numerical difficulties). One possible solution is to make r as large as possible. However, then w will be adjusted to meet the covariance constraint, so that the numerical difficulties remain. Another solution to this problem is to use smaller sampling interval, but this may not be possible. Another solution is to utilize a "$U-D$" factorization of Equation (13), given by

$$K=UDU^T \tag{41}$$

where U is a $q\times q$ unitary upper triangular matrix, and D is a $q\times q$ diagonal matrix. The matrix $K(t)$ can be processed by the i^{th} scalar measurement using [12]

$$K_i(t)=W^{-1}-\frac{\lambda_{ii}^2 W^{-1}s_i^T s_i W^{-1}}{\lambda_{ii}^2 s_i^T W^{-1}s_i+r_{ii}} \tag{42}$$

Using Equation (41) leads to

$$U_i D_i U_i^T=W^{-1}-\frac{\left[\lambda_{ii}W^{-1}s_i^T\right]\left[s_i W^{-1}\lambda_{ii}\right]}{\lambda_{ii}^2 s_i^T W^{-1}s_i+r_{ii}} \tag{43}$$

which is used to propagate the state estimates for each scalar measurement.

Case 1. Let $p_i=1$ for both the state and output systems. Equations (5), (7) and (8) reduce to

$$\underline{z}=\Delta t\,H(\hat{\underline{x}})\underline{f}(\hat{\underline{x}}) \tag{44a}$$

$$H(\hat{\underline{x}})\equiv\frac{\partial\hat{\underline{y}}}{\partial\hat{\underline{x}}} \tag{44b}$$

$$\Lambda=\lambda I \tag{44c}$$

$$S = H(\hat{\underline{x}})G(\hat{\underline{x}}) \quad (44d)$$

Therefore, the model error trajectory in Equation (12) is given by

$$\begin{aligned} \underline{d} &= -\Delta t\left\{I - W^{-1}G^T H^T Z(t) H\, G\right\} \\ &\times PG^T H^T R^{-1}\left\{\hat{\underline{y}}(t) - \tilde{\underline{y}}(t+\Delta t) + \Delta t\, H\, \underline{f}\right\} \end{aligned} \quad (45)$$

where

$$Z(t) = \left[H\, G W^{-1} G^T H^T + \Delta t^{-2} R\right]^{-1} \quad (46)$$

Case 2. Consider the following system

$$\dot{\hat{\underline{x}}}_1 = \underline{f}_1(\hat{\underline{x}}_1, \hat{\underline{x}}_2) \quad (47a)$$

$$\dot{\hat{\underline{x}}}_2 = \underline{f}_2(\hat{\underline{x}}_2) + G_2(\hat{\underline{x}}_2)\underline{d} \quad (47b)$$

$$\hat{\underline{y}} = \underline{c}(\hat{\underline{x}}_1) \quad (47c)$$

with $p_i = 2$. Equation (47a) usually defines the kinematics, and Equation (47b) usually defines the dynamics. Equations (5), (7) and (8) now become

$$\begin{aligned} \underline{z} &= \Delta t\, \underline{L}_f^1 \\ &+ \frac{\Delta t^2}{2}\left[\frac{\partial \underline{L}_f^1}{\partial \hat{\underline{x}}_1}\underline{f}_1(\hat{\underline{x}}_1, \hat{\underline{x}}_2) + \frac{\partial \underline{L}_f^1}{\partial \hat{\underline{x}}_2}\underline{f}_2(\hat{\underline{x}}_2)\right] \end{aligned} \quad (48a)$$

$$\underline{L}_f^1 \equiv \frac{\partial \underline{c}}{\partial \hat{\underline{x}}_1}\underline{f}_1(\hat{\underline{x}}_1, \hat{\underline{x}}_2) + \frac{\partial \underline{c}}{\partial \hat{\underline{x}}_2}\underline{f}_2(\hat{\underline{x}}_2) \quad (48b)$$

$$\Lambda = \lambda I \quad (48c)$$

$$S = \frac{\partial \underline{L}_f^1}{\partial \hat{\underline{x}}_2} G_2(\hat{\underline{x}}_2) \quad (48d)$$

Conclusions

In this paper, a predictive filter was presented for nonlinear systems. The optimal model error is found by using a one-time step ahead control approach. Advantages of the new algorithm over the extended Kalman filter include: (i) the model error is assumed unknown and is estimated as part of the solution, (ii) the model error may take any form (even nonlinear), and (iii) the model error is used to propagate a continuous model which avoids discrete jumps in the state estimate.

Acknowledgments

The first author's work was supported by a National Research Council Postdoctoral Fellowship tenured at NASA-Goddard Space Flight Center. The author greatly appreciates this support.

References

[1] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 82, March 1960, pp. 34-45.

[2] Mason, P.A.C., and Mook, D.J., "Scalar Gain Interpretation of Large Order Filters," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, 1992, pp. 425-439.

[3] Kane, T.R., Likins, P.W., and Levinson, D.A., *Spacecraft Dynamics*, McGraw-Hill, NY, 1983.

[4] Gelb, A., *Applied Optimal Estimation*, MIT Press, MA, 1974.

[5] Stengel, R.F., *Optimal Control and Estimation*, Dover Publications, NY, 1994.

[6] Mook, D.J., and Junkins, J.L., "Minimum Model Error Estimation for Poorly Modeled Dynamic Systems," *Journal of Guidance, Control and Dynamics*, Vol. 3, No. 4, Jan.-Feb. 1988, pp. 367-375.

[7] Crassidis, J.L., Mason, P.A.C., and Mook, D.J., "Riccati Solution for the Minimum Model Error Algorithm," *Journal of Guidance, Control and Dynamics*, Vol. 16, No. 6, Nov.-Dec. 1993, pp. 1181-1183.

[8] Lu, P., "Nonlinear Predictive Controllers for Continuous Systems," *Journal of Guidance, Control and Dynamics*, Vol. 17, No. 3, May-June 1994, pp. 553-560.

[9] Bierman, G.J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, FL, 1977.

[10] Vidyasagar, M., *Nonlinear Systems Analysis*, Prentice Hall, NJ, 1993.

[11] Shuster, M.D., and Oh, S.D., "Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, Jan.-Feb. 1981, pp. 70-77.

[12] Thornton, C.L., and Jacobson, R.A., "Linear Stochastic Control Using the $U\,D\,U^T$ Matrix Factorization," *Journal of Guidance and Control*, Vol. 1, No. 4, July-Aug. 1978, pp. 232-236.