PREDICTIVE FILTERING FOR NONLINEAR SYSTEMS

John L. Crassidis
Assistant Professor, Member AIAA
Department of Mechanical Engineering
Catholic University of America
Washington, D.C. 20064

F. Landis Markley
Staff Engineer, Associate Fellow AIAA
Guidance and Control Branch, Code 712
Goddard Space Flight Center
Greenbelt, MD 20771

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 ("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. The predictive filter is used to estimate the attitude of a spacecraft with the utilization of attitude sensor measurements and rate integrating gyro measurements. Results using this new algorithm indicate that the real-time predictive filter accurately estimates the attitude of the spacecraft by determining actual gyro bias errors.

Introduction

Conventional filter methods, such as the Kalman filter [1], have proven 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, an example involving the estimation of a spacecraft from attitude sensors and gyros is presented.

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 $\underline{f} \in \mathbb{R}^n$ is sufficiently differentiable, $\hat{\underline{x}}(t) \in \mathbb{R}^n$ is the state estimate vector, $\underline{d}(t) \in \mathbb{R}^q$ represents the model error vector, $G\left(\hat{\underline{x}}(t)\right): \mathbb{R}^n \to \mathbb{R}^{n\times q}$ is the model-error distribution matrix, $\underline{c}\left(\hat{\underline{x}}(t)\right) \in \mathbb{R}^m$ is the measurement vector, and $\hat{\underline{y}}(t) \in \mathbb{R}^m$ 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 $\tilde{\underline{y}}_k \in \mathbb{R}^m$ is the measurement vector at time t_k, $\underline{x}(t_k)$ is the true state vector, and $\underline{v}_k \in \mathbb{R}^m$ 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 $R \in \mathbb{R}^{m\times m}$ is a 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

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

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 $\hat{x}_i(t)$ on the right side. $L_f^k(c_i)$ is a k^{th} order Lie derivative, defined by

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

$\Lambda(\Delta t) \in \mathbb{R}^{m\times m}$ is a diagonal matrix with elements given by

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

$S\left(\hat{\underline{x}}(t)\right) \in \mathbb{R}^{m\times q}$ is a 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 L_{g_j} in Equation (8) is defined by

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

Equation (8) is in essence a generalized sensitivity matrix for nonlinear systems.

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

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

where $W \in \mathbb{R}^{q\times q}$ is positive semidefinite. Also, a constant sampling rate is assumed so that $\underline{\tilde{y}}(t+\Delta t) \equiv \underline{\tilde{y}}_{k+1}$. Substituting Equation (4), and minimizing Equation (10) with respect to $\underline{d}(t)$ leads to the following model error

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

By using the matrix inversion lemma [9], the model error in Equation (11) can be re-written as

$$\underline{d}(t) = -M(t)\left[\underline{z}(\underline{\hat{x}},\Delta t) - \underline{\tilde{y}}(t+\Delta t) + \underline{\hat{y}}(t)\right] \tag{12}$$

where

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

The measurement at time t_{k+1} is used in Equation (12) to find the appropriate $\underline{d}$ for the nonlinear propagation of the state estimates from time t_k to time t_{k+1}. By looking ahead to the next measurement to compute $\underline{d}$, the predictive filter gives a continuous state estimate without the jumps exhibited by the Kalman filter. 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 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]

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

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

$$\rho_{iik} = \frac{c_{ii_k}}{c_{ii_0}} \tag{17}$$

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]

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

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

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

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

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

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

leads to

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

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

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

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

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

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

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

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

which leads to

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

where

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

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 covariance. This can be used to determine an adaptive scheme for determining W to satisfy the covariance constraint (which will be reported at a later time).

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

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

Substituting Equations (4) and (12) into Equation (27) leads to

$$\underline{e}(t+\Delta t)=\left[I-\Lambda(\Delta t)S(\hat{\underline{x}})M(t)\right]\hat{\underline{e}}(t+\Delta t) \tag{28}$$

where

$$\hat{\underline{e}}(t+\Delta t)\equiv\tilde{\underline{y}}(t+\Delta t)-\hat{\underline{y}}(t)-\underline{z}(\hat{\underline{x}},\Delta t) \tag{29}$$

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

If S is square and full rank, then $\Lambda S M$ is also full rank. As $W\to 0$, then $\Lambda S M\to I$, and $(I-\Lambda S M)\to 0$. This approaches a deadbeat response for the filter dynamics. As $W\to\infty$, then $M\to 0$, and $(I-\Lambda S M)\to I$. This yields a filter response with discrete eigenvalues approaching the unit circle. As long as the covariance matrix is positive, the discrete eigenvalues of the filter will lie within the unit circle. Therefore, the filter remains contractive [11].

Robustness

Lu [12] has shown that the dual control problem achieves input/output linearization, and asymptotic tracking of any given trajectory if $p_i\le 4$. An analysis of the robustness properties in the face of unmodeled dynamics for $p_i=1$, $W=0$, and square and nonsingular $S(\hat{\underline{x}})$ has also been shown in [12]. In this paper, the case of $p_i=1$, $W\neq 0$, and $S(\hat{\underline{x}})\in \mathrm{R}^{m\times 3}$, where $m\le 3$ is considered. The continuous output estimate for $p_i=1$ is given by [12]

$$\dot{\hat{\underline{y}}}=\underline{L}_f(\underline{c})+S(\hat{\underline{x}})\underline{d} \tag{30}$$

where

$$\underline{L}_f(\underline{c})\equiv\begin{bmatrix}L_f(c_1)\\ \vdots\\ L_f(c_m)\end{bmatrix} \tag{31}$$

Suppose that the unmodeled errors are introduced into the output estimate by

$$\dot{\hat{\underline{y}}}=\underline{L}_f(\underline{c})+\Delta\underline{L}_f(\underline{c})+\left[S(\hat{\underline{x}})+\Delta S(\hat{\underline{x}})\right]\underline{d} \tag{32}$$

and suppose that $\underline{L}_f(\underline{c})$ and $\Delta\underline{L}_f(\underline{c})$ are bounded by

$$\left\|\underline{L}_f(\underline{c})\right\|\le n_1,\quad \left\|\Delta\underline{L}_f(\underline{c})\right\|\le n_2,\quad \text{for all } x\in X \tag{33}$$

Furthermore, assume that $\Delta S(\hat{\underline{x}})$ is represented by

$$\Delta S(\hat{\underline{x}}) = \delta(\hat{\underline{x}}) S(\hat{\underline{x}}) \tag{34}$$

where $\delta(\hat{\underline{x}})$ is a scalar, continuous function with bound given by $-1 < \delta(\hat{\underline{x}}) < n_3$. Assuming that the model errors and measurement errors are isotropic leads to $W = w I$, and $R = r I$. Then, the matrix inverse in Equation (11) can be written as (suppressing arguments)

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

where

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

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

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

By the Cayley-Hamilton theorem, any meromorphic function of C can be expressed as a quadratic in C [13], yielding

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

where

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

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

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

$$k = \operatorname{tr}(\operatorname{adj} C) = \frac{\Delta t^4}{r^2} \operatorname{tr}\left[\operatorname{adj}\left(S^T S\right)\right] \tag{38d}$$

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

Therefore, the error dynamics become

$$\begin{aligned} \underline{\dot{e}} = &-\frac{\Delta t}{r\gamma}(1+\delta) Q \underline{e} + \tilde{\underline{y}} - \underline{L}_f - \Delta \underline{L}_f \\ &+ \frac{\Delta t^2}{r\gamma}(1+\delta) Q \left[\underline{L}_f - \tilde{\underline{y}} \right] \end{aligned} \tag{39}$$

where

$$Q \equiv \left[\alpha\left(S S^T\right) + \frac{\Delta t^2}{r}\beta\left(S S^T\right)^2 + \frac{\Delta t^4}{r^2}\left(S S^T\right)^3 \right] \tag{40}$$

Now, define a Lyapunov function $V = \|\underline{e}\|^2 / 2$. Using the norm inequality [14], and the fact that

$$\underline{e}^T Q \underline{e} \geq \frac{1}{\left\|Q^{-1}\right\|} \underline{e}^T \underline{e} \tag{41}$$

leads to

$$\begin{aligned} \dot{V} \leq &-\frac{2\Delta t(1+\delta)}{r\gamma\left\|Q^{-1}\right\|} V + \left\|\underline{e}^T\right\| \left\| \tilde{\underline{y}} - \underline{L}_f - \Delta \underline{L}_f \right\| \\ &+ \frac{\Delta t^2 (1+\delta)}{r\gamma} \left\|\underline{e}^T\right\| \|Q\| \left\| \underline{L}_f - \tilde{\underline{y}} \right\| \end{aligned} \tag{42}$$

Next, using the well known inequality $a b \leq z a^2 + b^2/(4z)$ for any a, b, and $z > 0$, and defining $\xi \equiv \Delta t(1+\delta)/(r\gamma)$ yields

$$\begin{aligned} \dot{V} \leq &\left(-\frac{2\xi}{\left\|Q^{-1}\right\|} + 4z \right) V + \frac{\left\| \tilde{\underline{y}} - \underline{L}_f - \Delta \underline{L}_f \right\|^2}{4z} \\ &+ \frac{\Delta t^2 \xi^2 \|Q\|^2 \left\| \underline{L}_f - \tilde{\underline{y}} \right\|^2}{4z} \end{aligned} \tag{43}$$

Substituting $4z = \xi / \left\|Q^{-1}\right\|$ leads to

$$\dot{V} \leq -\frac{\xi}{\left\|Q^{-1}\right\|} V + b \tag{44}$$

where

$$b \equiv \left\|Q^{-1}\right\| \left[\frac{1}{\xi} \left\| \tilde{\underline{y}} - \underline{L}_f - \Delta \underline{L}_f \right\|^2 + \Delta t^2 \xi \|Q\|^2 \left\| \underline{L}_f - \tilde{\underline{y}} \right\|^2 \right] \tag{45}$$

Therefore, Equation (44) can be solved to yield

$$V \leq \left(V_0 - \frac{b\left\|Q^{-1}\right\|}{\xi} \right) e^{-\xi t / \left\|Q^{-1}\right\|} + \frac{b\left\|Q^{-1}\right\|}{\xi} \tag{46}$$

where $V_0 \geq b\left\|Q^{-1}\right\| / \xi$ is required to maintain the inequality. Defining the bounds

$$\theta = \inf(1+\delta) > 0 \tag{47a}$$

$$\left\| \tilde{\underline{y}} \right\|_\infty = \max_{t \in [0, t_f]} \sum_{i=1}^{m} \left| \tilde{\underline{y}}_i \right| \tag{47b}$$

and using the matrix norm inequality again leads to

$$\|\underline{e}\| \leq \sqrt{2} \frac{r\gamma\left\|Q^{-1}\right\|}{\Delta t\, \theta} \mu_1^2 + \sqrt{2}\, \Delta t \left\|Q^{-1}\right\| \|Q\| \mu_2^2 \tag{48}$$

where

$$\mu_1 = \left[\left\| \tilde{\underline{y}} \right\|_\infty - n_1 - n_2 \right] \tag{49a}$$

$$\mu_2 = \left[\theta \left(n_1 - \left\| \tilde{\underline{y}} \right\|_\infty \right) \right] \tag{49b}$$

Assuming that $\left\| S\, S^T \right\| \le s_1$ for all $S\, S^T$ leads to the following bound on $\|Q\|$

$$\|Q\| \le \alpha\, s_1 + \frac{\Delta t^2}{r} |\beta| s_1^2 + \frac{\Delta t^4}{r^2} s_1^3 \tag{50}$$

Also, $\text{tr}\left(S^T S\right) \le 3\, s_1$ and $\det\left(S^T S\right) \le s_1^3$, which leads to

$$\sigma \le \frac{3\,\Delta t^2}{2\,r} s_1 \tag{51a}$$

$$k \le \frac{9\,\Delta t^4}{r^2} s_1^2 \tag{51b}$$

$$\Delta \le \frac{\Delta t^6}{r^3} s_1^3 \tag{51c}$$

Substituting Equation (51) into Equations (38) and (50) leads to the following bounds on $\|Q\|$ and γ

$$\|Q\| \le w^2 s_1 + \frac{4\, w\, \Delta t^2 s_1^2}{r} + \frac{13\,\Delta t^4 s_1^3}{r^2} \tag{52a}$$

$$\gamma \le w^3 + \frac{3\, w^2 \Delta t^2 s_1}{r} + \frac{9\,\Delta t^4 s_1^2}{r^2} + \frac{\Delta t^6 s_1^3}{r^3} \tag{52b}$$

A bound on $\left\| Q^{-1} \right\|$ is found by writing it as

$$\begin{aligned} \left\| Q^{-1} \right\| &= \gamma \left\| \left\{ S \left[\frac{\Delta t^2}{r} S^T S + wI \right]^{-1} S^T \right\}^{-1} \right\| \\ &\le \gamma \left\| \left(S\, S^T \right)^{-1} \right\| \left\| \left[\frac{\Delta t^2}{r} S^T S + wI \right] \right\| \\ &\le \gamma \left\| \left(S\, S^T \right)^{-1} \right\| \left\{ \frac{\Delta t^2}{r} \left\| S^T S \right\| + w \right\} \end{aligned} \tag{53}$$

Assuming that $\left\| \left(S\, S^T \right)^{-1} \right\| \le s_2$ for all $S\, S^T$ leads to

$$\left\| Q^{-1} \right\| \le \gamma\, s_2 \left[\frac{\Delta t^2}{r} s_1 + w \right] \tag{54}$$

Therefore Equations (48), (52), and (54) define the bound for the error dynamics under unmodeled uncertainty. Similar results can be obtained for $1 < p_i \le 4$. The case where $q > 3$ can also be determined using a Cayley-Hamilton expansion, but becomes increasingly more complicated.

Numerical Stability

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 for bounded model uncertainties as long as

$$\left(v - \sigma \right) \alpha + \Delta > 0 \tag{55}$$

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

$$v > \frac{\Delta t^2}{2\,r} \text{tr}\left(S^T S \right) \tag{56}$$

Therefore, the filter remains contractive as long as $w > 0$. This condition is always met, but Equation (56) 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. A more practical solution is to utilize a "$U - D$" factorization of Equation (13) [15].

Cases

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

$$\underline{z} = \Delta t\, H\left(\hat{\underline{x}}\right) \underline{f}\left(\hat{\underline{x}}\right) \tag{57a}$$

$$H\left(\hat{\underline{x}}\right) \equiv \frac{\partial \hat{\underline{y}}}{\partial \hat{\underline{x}}} \tag{57b}$$

$$\Lambda = \Delta t\, I \tag{57c}$$

$$S = H\left(\hat{\underline{x}}\right) G\left(\hat{\underline{x}}\right) \tag{57d}$$

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

$$\begin{aligned} \underline{d} = -\Delta t &\left\{ I - W^{-1} G^T H^T \left[H\, G\, W^{-1} G^T H^T + \Delta t^{-2} R \right]^{-1} H\, G \right\} \\ &\times W^{-1} G^T H^T R^{-1} \left\{ \tilde{\underline{y}}(t) - \tilde{\underline{y}}(t + \Delta t) + \Delta t\, H\, \underline{f} \right\} \end{aligned} \tag{58}$$

Case 2. Consider the following system

$$\dot{\hat{\underline{x}}}_1 = \underline{f}_1\left(\hat{\underline{x}}_1, \hat{\underline{x}}_2\right) \tag{59a}$$

$$\dot{\hat{\underline{x}}}_2 = \underline{f}_2\left(\hat{\underline{x}}_2\right) + G_2\left(\hat{\underline{x}}_2\right) \underline{d} \tag{59b}$$

$$\hat{\underline{y}} = \underline{c}\left(\hat{\underline{x}}_1\right) \tag{59c}$$

with $p_i = 2$. Equation (59a) usually defines the kinematics, and Equation (59b) usually defines the dynamics of a system. Equations (5), (7) and (8) now become

$$\underline{z} = \Delta t\, \underline{L}_f^1 + \frac{\Delta t^2}{2}\left[\frac{\partial \underline{L}_f^1}{\partial \underline{\hat{x}}_1} \underline{f}_1\left(\underline{\hat{x}}_1, \underline{\hat{x}}_2\right) + \frac{\partial \underline{L}_f^1}{\partial \underline{\hat{x}}_2} \underline{f}_2\left(\underline{\hat{x}}_2\right)\right] \quad (60a)$$

$$\underline{L}_f^1 \equiv \frac{\partial \underline{c}}{\partial \underline{\hat{x}}_1} \underline{f}_1\left(\underline{\hat{x}}_1, \underline{\hat{x}}_2\right) + \frac{\partial \underline{c}}{\partial \underline{\hat{x}}_2} \underline{f}_2\left(\underline{\hat{x}}_2\right) \quad (60b)$$

$$\Lambda = \frac{\Delta t^2}{2} I \quad (60c)$$

$$S = \frac{\partial \underline{L}_f^1}{\partial \underline{\hat{x}}_2} G_2\left(\underline{\hat{x}}_2\right) \quad (60d)$$

Spacecraft Attitude Estimation

In this section, the predictive filter is used to determine the attitude of a spacecraft using gyros and attitude sensor measurements. First, a brief review of attitude kinematics and gyro dynamics is shown. Next, a brief review of the Kalman filter for attitude estimation is shown. Then, a predictive filter for attitude estimation is developed. Finally, simulation results using the predictive filter to estimate the attitude of the Tropical Rainfall Measurement Mission (TRMM) spacecraft [16] are shown.

Attitude Kinematics and Gyro Dynamics

The attitude is assumed to be represented by the quaternion, defined as

$$\underline{q} = \begin{bmatrix} \underline{q}_{13} \\ q_4 \end{bmatrix} \quad (61)$$

with

$$\underline{q}_{13} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \underline{\hat{n}} \sin\left(\theta/2\right) \quad (62a)$$

$$q_4 = \cos\left(\theta/2\right) \quad (62b)$$

where $\underline{\hat{n}}$ is a unit vector corresponding to the axis of rotation, and θ is the angle of rotation.

The quaternion kinematic equations of motion are derived by using the spacecraft's angular velocity ($\underline{\omega}$),

$$\underline{\dot{q}} = \frac{1}{2}\Omega\left(\underline{\omega}\right)\underline{q} = \frac{1}{2}\Xi\left(\underline{q}\right)\underline{\omega} \quad (63)$$

where $\Omega\left(\underline{\omega}\right)$ and $\Xi\left(\underline{q}\right)$ are defined as

$$\Omega\left(\underline{\omega}\right) \equiv \begin{bmatrix} -\left[\underline{\omega}\times\right] & \vdots & \underline{\omega} \\ \cdots\cdots & \vdots & \cdots\cdots \\ -\underline{\omega}^T & \vdots & 0 \end{bmatrix} \quad (64a)$$

$$\Xi\left(\underline{q}\right) \equiv \begin{bmatrix} q_4 I_{3\times3} + \left[\underline{q}_{13}\times\right] \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ -\underline{q}_{13}^T \end{bmatrix} \quad (64b)$$

where $I_{3\times3}$ is a 3×3 identity matrix. The 3×3 dimensional matrices $\left[\underline{\omega}\times\right]$ and $\left[\underline{q}_{13}\times\right]$ are referred to as cross product matrices since $\underline{a}\times\underline{b} = \left[\underline{a}\times\right]\underline{b}$, with

$$\left[\underline{a}\times\right] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \quad (65)$$

Since a three degree-of-freedom attitude system is represented by a four-dimensional vector, the quaternions cannot be independent. This condition leads to the following normalization constraint

$$\underline{q}^T \underline{q} = \underline{q}_{13}^T \underline{q}_{13} + q_4^2 = 1 \quad (66)$$

The measurement model is assumed to be of the form given by

$$\underline{B}_B = A\left(\underline{q}\right)\underline{B}_I \quad (67)$$

where $\underline{B}_I$ is a 3×1 dimensional vector of some reference object (e.g., a vector to the sun or to a star, or the Earth's magnetic field vector) in a reference coordinate system, $\underline{B}_B$ is a 3×1 dimensional vector defining the components of the corresponding reference vector measured in the spacecraft body frame, and $A\left(\underline{q}\right)$ is given by

$$A\left(\underline{q}\right) = \left(q_4^2 - \underline{q}_{13}^T \underline{q}_{13}\right) I_{3\times3} + 2\,\underline{q}_{13}\underline{q}_{13}^T - 2\,q_4\left[\underline{q}_{13}\times\right] \quad (68)$$

where $A \in SO(3)$, which is a Lie group of orthogonal matrices with determinant 1 (i.e., $A^T A = I$ and $\det\left(A\right) = 1$). A simpler form for the attitude matrix in Equation (68) is given by

$$A\left(\underline{q}\right) = -\Xi^T\left(\underline{q}\right)\Psi\left(\underline{q}\right) \quad (69)$$

where

$$\Psi\left(\underline{q}\right) \equiv \begin{bmatrix} -q_4 I_{3\times3} + \left[\underline{q}_{13}\times\right] \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \underline{q}_{13}^T \end{bmatrix} \quad (70)$$

Also, another useful identity is given by

$$\Psi\left(\underline{q}\right)\underline{\omega}=\Gamma\left(\underline{\omega}\right)\underline{q} \qquad \text{for any } \underline{\omega}_{3\times 1} \tag{71}$$

where

$$\Gamma\left(\underline{\omega}\right)\equiv\begin{bmatrix}-[\underline{\omega}\times] & \vdots & -\underline{\omega}\\ \cdots\cdots & \vdots & \cdots\cdots\\ \underline{\omega}^T & \vdots & 0\end{bmatrix} \tag{72}$$

The true angular velocity is assumed to be modeled by

$$\underline{\omega}=\underline{\tilde{\omega}}_g-\underline{\omega}_b-\underline{\eta}_1 \tag{73}$$

where $\underline{\omega}$ is the true angular velocity, $\underline{\tilde{\omega}}_g$ is the gyro-determined angular velocity, and $\underline{\omega}_b$ is the gyro drift vector, which is modeled by

$$\underline{\dot{\omega}}_b=\underline{\eta}_2 \tag{74}$$

The 3×1 vectors, $\underline{\eta}_1$ and $\underline{\eta}_2$, are assumed to be modeled by Gaussian white-noise processes with

$$E\left\{\underline{\eta}_i\left(t\right)\right\}=\underline{0} \qquad i=1,2 \tag{75a}$$

$$E\left\{\underline{\eta}_i\left(t\right)\underline{\eta}_j^T\left(t'\right)\right\}=Q_i\,\delta_{ij}\,\delta\left(t-t'\right) \qquad i,j=1,2 \tag{75b}$$

Predictive Filtering

The nonlinear predictive filter minimizes the following

$$J=\frac{1}{2}\left\{\underline{\tilde{B}}_B-A\left(\underline{\hat{q}}\right)\underline{B}_I\right\}^T\Big|_{t+\Delta t}R^{-1}\left\{\underline{\tilde{B}}_B-A\left(\underline{\hat{q}}\right)\underline{B}_I\right\}\Big|_{t+\Delta t}+\frac{1}{2}\underline{d}^T\left(t\right)W\,\underline{d}\left(t\right) \tag{76}$$

subject to

$$\underline{\dot{\hat{q}}}=\frac{1}{2}\Omega\left(\underline{\hat{\omega}}\right)\underline{\hat{q}}=\frac{1}{2}\Xi\left(\underline{\hat{q}}\right)\underline{\hat{\omega}}, \qquad \underline{\hat{q}}\left(t_0\right)=\underline{q}_0 \tag{77a}$$

$$\underline{\dot{\hat{\omega}}}_b=\underline{d}, \qquad \underline{\hat{\omega}}_b\left(t_0\right)=\underline{0}_{3\times 1} \tag{77b}$$

$$\underline{\hat{\omega}}=\underline{\tilde{\omega}}_g-\underline{\hat{\omega}}_b \tag{77c}$$

This has the form given by Equation (59), with $\underline{\hat{x}}_1=\underline{\hat{q}}$ and $\underline{\hat{x}}_2=\underline{\hat{\omega}}_b$. Since the body measurements $\left(\underline{\tilde{B}}_B\right)$ are used as the required tracking trajectories, the output vector in Equation (2) is given by

$$\underline{c}\left(\underline{\hat{x}}\right)=A\left(\underline{\hat{q}}\right)\underline{B}_I \tag{78}$$

Since $\underline{c}$ depends on $\underline{\hat{q}}$ and not explicitly on $\underline{\hat{\omega}}_b$, the lowest order time derivative of Equation (78) in which any component of $\underline{d}$ first appears in $\underline{\hat{q}}$ is two, so that $p_i=2$. Therefore, the Λ, $\underline{z}$, and $\underline{e}$ quantities are given by

$$\Lambda=\frac{\Delta t^2}{2}I \tag{79a}$$

$$\underline{z}=\Delta t\underline{L}_f^1+\frac{\Delta t^2}{2}\underline{L}_f^2 \tag{79b}$$

$$\underline{e}=\underline{\tilde{B}}_B-A\left(\underline{\hat{q}}\right)\underline{B}_I \tag{79c}$$

where

$$\left[\underline{L}_f^k\right]_i\equiv L_f^k\left(c_i\right) \tag{80}$$

Using the definitions in Equations (64b), (71), and (72), the derivative of Equation (78) with respect to $\underline{\hat{q}}$ is

$$\frac{\partial\underline{c}}{\partial\underline{\hat{q}}}=-2\,\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right) \tag{81}$$

Therefore, $\underline{L}_f^1$ is given by

$$\underline{L}_f^1=-\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right)\Xi\left(\underline{\hat{q}}\right)\underline{\hat{\omega}} \tag{82}$$

The derivative of Equation (82) with respect to $\underline{\hat{q}}$ is given by

$$\frac{\partial\underline{L}_f^1}{\partial\underline{\hat{q}}}=2\left[\underline{\hat{\omega}}\times\right]\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right) \tag{83}$$

The derivative of Equation (82) with respect to $\underline{\hat{\omega}}_b$ is given by

$$\frac{\partial\underline{L}_f^1}{\partial\underline{\hat{\omega}}_b}=\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right)\Xi\left(\underline{\hat{q}}\right) \tag{84}$$

Therefore, $\underline{L}_f^2$ is given by

$$\underline{L}_f^2=\left[\underline{\hat{\omega}}\times\right]\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right)\Xi\left(\underline{\hat{q}}\right)\underline{\hat{\omega}} \tag{85}$$

The S matrix, which is formed using Equation (8) is given by

$$S=\Xi^T\left(\underline{\hat{q}}\right)\Gamma\left(\underline{B}_I\right)\Xi\left(\underline{\hat{q}}\right)=-\left[A\left(\underline{\hat{q}}\right)\underline{B}_I\times\right] \tag{86}$$

The 3×3 matrix $\left[A\left(\underline{\hat{q}}\right)\underline{B}_I\times\right]$ has at most rank 2, which reflects the fact that there is no information about rotations around the current measurement vector [17].

The extension to multiple measurement sets is achieved by stacking these measurements, e.g.,

$$\underline{e} = \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_{m_{\text{sen}}} \end{bmatrix} \tag{87}$$

where m_{sen} is the total number of vector measurement sets available at time t_k.

An advantage of the nonlinear predictive filter over the conventional Kalman filter for attitude estimation is that quaternion normalization is always preserved. The normalization constraint leads to a singularity in the covariance matrix of the Kalman filter. Several methods are shown in references [17] and [18] which overcome this difficulty. However, since Equation (77) is used to determine the quaternion estimate, normalization is always maintained in the nonlinear predictive filter. Also, the predictive filter determines the model-error trajectory as part of the solution, as opposed to the Kalman filter which assumes that this error is modeled by a Gaussian process with known covariance.

Attitude Estimation of TRMM

In this section, the predictive filter is used to estimate the attitude of the TRMM spacecraft from vector measurement observations and gyro measurements. The TRMM spacecraft (see Figure 1) is due to be launched in 1997 with a nominal mission lifetime of 42 months. The main objectives of this mission include: (i) to obtain multi-year measurements of tropical rainfall, (ii) to understand how interactions between the sea, air, and land masses produce changes in global rainfall and climate, and (iii) to help improve the modeling of rainfall processes and their influence on global circulation.

The spacecraft is three-axis stabilized in a near circular (350 km) orbit with an inclination of 35°. Figure 2 shows the position of the spacecraft for a typical orbit. The nominal Earth-pointing mission mode requires a rotation once per orbit about the spacecraft's y-axis. The attitude determination hardware consists of an Earth Sensor Assembly (ESA), Digital Sun Sensors (DSS), Coarse Sun Sensors (CSS), a Three-Axis Magnetometer (TAM), and gyroscopic rate sensors. The attitude control hardware includes three Magnetic Torquer Bars (MTB) which are used to provide magnetic momentum unloading capability, and a Reaction Wheel Assembly (RWA) which consists of four wheels in a pyramidal arrangement to maximize momentum storage capability along a preferred axis.

Primary attitude knowledge is determined by the ESA and gyros. However, in the event of ESA failure, a contingency mode is used which utilizes the FSS, TAM, and gyros. This mode is used for attitude estimation in the simulations for this paper. The simulated spacecraft completes an orbit in approximately 90 minutes, giving a rotation rate of 236 deg/hr about the spacecraft's y-axis while holding the remaining axis rotations near zero. The magnetic field reference is modeled using a 10th order International Geomagnetic Reference Field (IGRF) model [19]. TAM sensor noise is modeled by a Gaussian white-noise process with a mean of zero and a standard deviation of 0.5 mG.

The two DSS's each have a field of view of about $50° \times 50°$ (see [16] for more details). The two DSS's combine to provide sun measurements for about 2/3 of a complete orbit. The DSS sensor noise is also modeled by a Gaussian white-noise process with a mean of zero and a standard deviation of 0.05°. The gyro "measurements" are simulated using Equations (73) and (74), with a gyro noise standard deviation of 0.062 deg/hr, a ramp noise standard deviation of 0.235 deg/hr/hr, and an initial drift of -0.1 deg/hr.

A plot of the roll, pitch, and yaw attitude errors using the propagated gyro model for a typical simulation run is shown in Figure 3. Clearly, the gyro model used in the simulation produces significant errors in pitch. Also, the roll and yaw errors eventually drift to large errors. A plot of the gyro-bias estimates using the predictive filter is shown in Figure 4. The predictive filter is able to accurately estimate for the gyro-biases. A plot of the predictive filter covariance is shown in Figure 5. This shows that the y-axis gyro-bias covariance is smaller than the x and z-axes. This is most likely due to roll/yaw quarter-orbit coupling for Earth-pointing spacecraft (i.e., a yaw rate error becomes a roll angle error as the spacecraft rotates in the orbital plane). A plot of the attitude errors using the predictive filter is shown in Figure 6.

A plot of the quaternion normalization for the Kalman filter and the predictive filter is shown in Figure 7. For the Kalman filter, the top of Figure 7 shows the post-update normalization error, where the quaternion is explicitly normalized before being propagated to the next measurement. The Kalman filter has larger errors at the start of the simulation run due to large initial updates. This shows that the linearization assumption in the Kalman filter becomes more invalid as the updates in the state estimates become larger. However, from the bottom plot of Figure 7, it is clear that the predictive filter is able accurately estimate the attitude of the spacecraft, while preserving quaternion normalization in the filter (to within numerical accuracy).

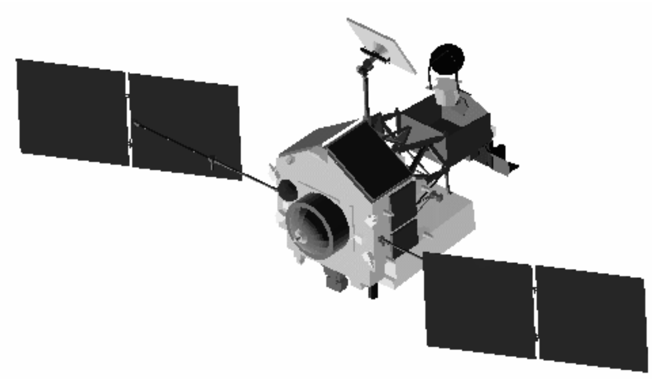

Figure 1 TRMM Spacecraft

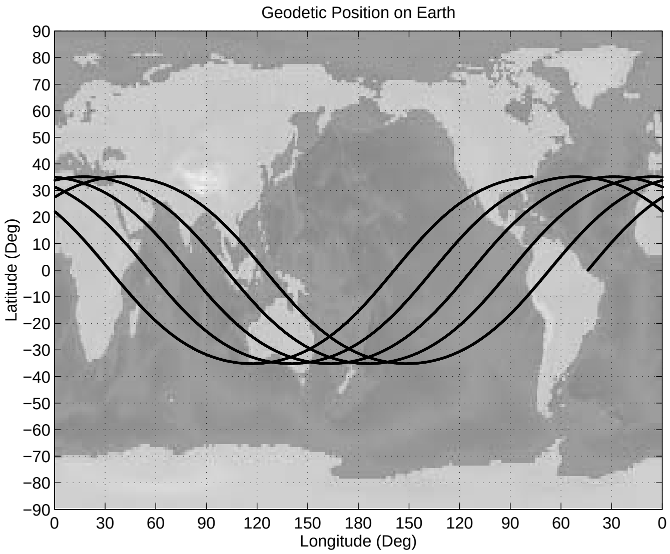

Figure 2 TRMM Spacecraft Position

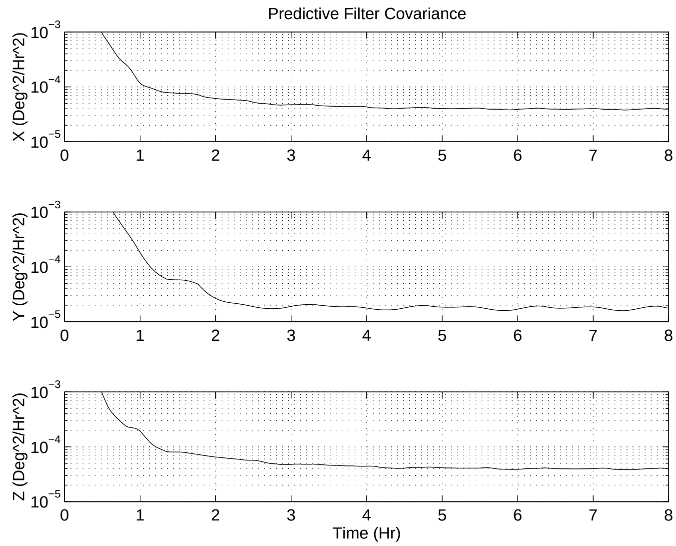

Figure 5 Predictive Filter Covariance

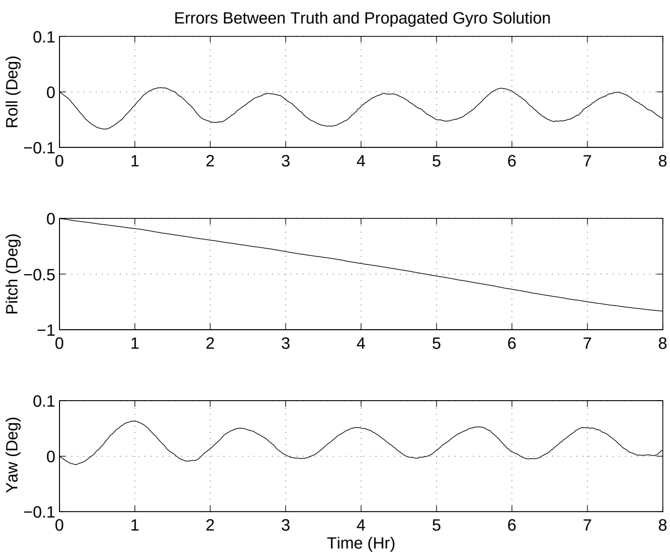

Figure 3 Propagated Gyro Attitude Errors

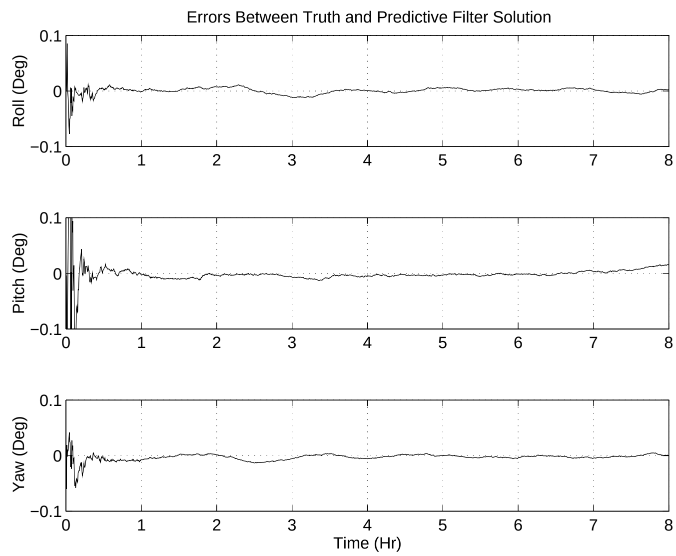

Figure 6 Predictive Filter Attitude Errors

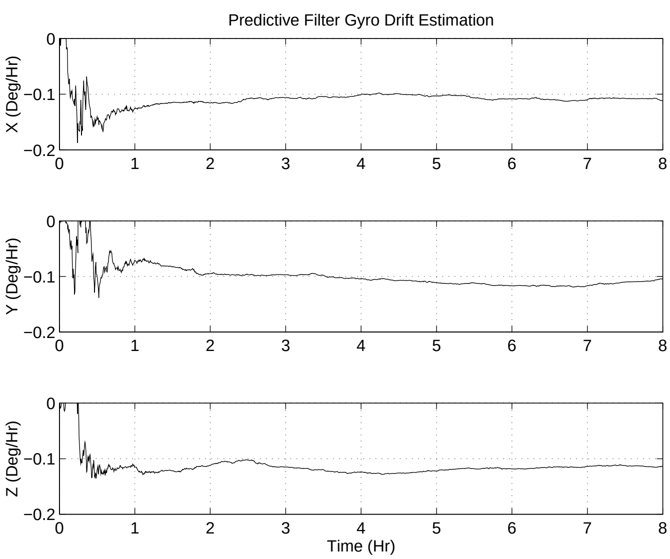

Figure 4 Predictive Filter Gyro-Biases Estimates

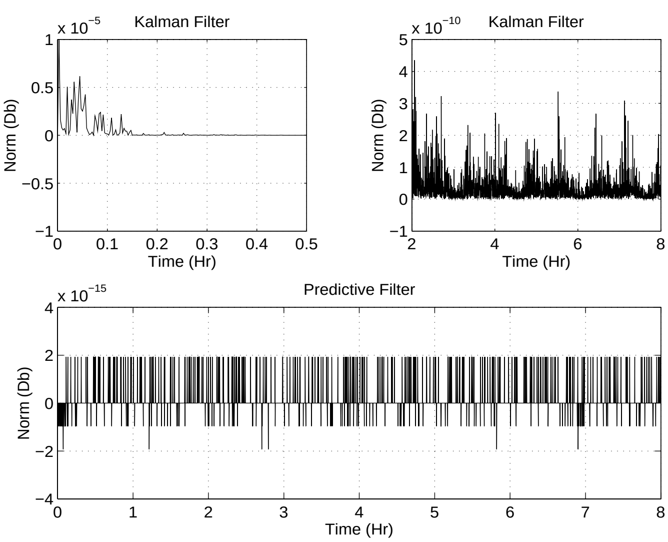

Figure 7 Quaternion Normalization for EKF and PF

Conclusions

In this paper, a predictive filter was presented for nonlinear systems. 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. An example of this algorithm was shown which estimated the attitude of a simulated spacecraft. The predictive filter was shown to preserve quaternion normalization throughout the estimation process, and accurately estimated the spacecraft attitude using attitude sensor observations and gyro measurements.

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. Also, the author wishes to thank Dr. D. Joseph Mook of the State University of New York at Buffalo, and Dr. Ping Lu of Iowa State University for many interesting and helpful discussions.

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] Lewis, F.L., *Optimal Estimation*, John Wiley & Sons, NY, 1986.

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

[12] Lu, P., "Optimal Predictive Control of Continuous Nonlinear Systems," *International Journal of Control*, Vol. 62, No. 3, Sept. 1995, pp. 633-649.

[13] 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.

[14] Horn, R.A., and Johnson, C.R., *Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

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

[16] Crassidis, J.L., Andrews, S.F., Markley, F.L., and Ha, K., "Contingency Designs for Attitude Determination of TRMM," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, 1995, pp. 419-433.

[17] Lefferts, E.J., Markley, F.L., and Shuster, M.D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics*, Vol. 5, No. 5, Sept.-Oct. 1982, pp. 417-429.

[18] Bar-Itzhack, I.Y., and Deutschmann, J.K., "Extended Kalman Filter for Attitude Estimation of the Earth Radiation Budget Satellite," *Proceedings of the AAS Astrodynamics Conference*, Portland, OR, August 1990, AAS Paper #90-2964, pp. 786-796.

[19] Langel, R.A., "International Geomagnetic Reference Field: The Sixth Generation," *Journal of Geomagnetism and Geoelectricity*, Vol. 44, No. 9, 1992, pp. 679-707.