AAS 06-035

Nonlinear Consider Covariance Analysis Using a Sigma-Point Filter Formulation

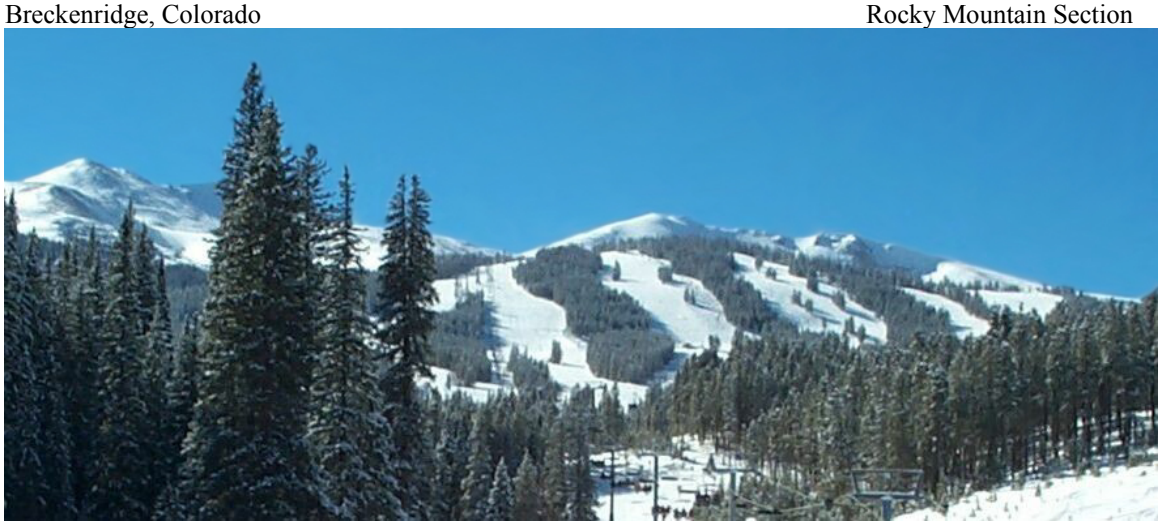
Dr. Michael E. Lisano

Jet Propulsion Laboratory, California Institute of Technology

29th ANNUAL AAS GUIDANCE AND CONTROL CONFERENCE

February 4-8, 2006
Breckenridge, Colorado

Sponsored by
Rocky Mountain Section



AAS Publications Office, P.O. Box 28130 - San Diego, California 92198

AAS 06-035

NONLINEAR CONSIDER COVARIANCE ANALYSIS USING A SIGMA-POINT FILTER FORMULATION

Dr. Michael E. Lisano*

The research reported here extends the mathematical formulation of nonlinear, sigma-point estimators to enable consider covariance analysis for dynamical systems. This paper presents a novel sigma-point consider filter algorithm, for consider-parameterized nonlinear estimation, following the unscented Kalman filter (UKF) variation on the sigma-point filter formulation, which requires no partial derivatives of dynamics models or measurement models with respect to the parameter list. It is shown that, consistent with the attributes of sigma-point estimators, a consider-parameterized sigma-point estimator can be developed entirely without requiring the derivation of any partial-derivative matrices related to the dynamical system, the measurements, or the considered parameters, which appears to be an advantage over the formulation of a linear-theory sequential consider estimator. It is also demonstrated that a consider covariance analysis performed with this "partial-derivative-free" formulation yields equivalent results to the linear-theory consider filter, for purely linear problems.

INTRODUCTION

Consider covariance analysis entails conservatively calculating the formal uncertainty associated with estimating a list of states and parameters, while omitting to estimate some constant parameters whose values are in error.[1] Considering parameters in a dynamical system while estimating its state provides a upper bound on the formal estimated state error covariance, which is a conservative approach to designing estimators for problems of general guidance, navigation and control. This is true because error in the value of a non-estimated parameter increases the estimation error of the system beyond the level formally indicated by the covariance of an estimator that neglects errors or uncertainty in that parameter.

Because consider covariance analysis represents a conservative engineering approach of estimating and predicting spacecraft trajectories, consider covariance capabilities are featured in some established orbit determination filter codes that are used in ground-based space flight operations. Hence, the ability to perform consider covariance analysis within the mathematical framework of a branch of estimation theory (e.g. with a batch filter formulation, or in classical linearized or extended sequential estimator formulations) enables that theoretical branch to be employed in the context of operational orbit determination, particularly for deep-space missions with stringent targeting requirements.

Recent literature in applied estimation theory reflects broad and growing interest in applying sigma-point (also called "unscented") formulations [c.f. 2, 5] to problems in nonlinear optimal sequential state estimation. These references often provide performance comparisons with the extended Kalman filter (EKF) algorithm, as applied to specific dynamical problems [c.f. 3, 4]. Favorable attributes of sigma-point filters over linear-theory sequential estimators such as the EKF are described as including a lower expected estimate error for nonlinear – even non-differentiable – dynamical systems, and a straightforward

* Technical Group Supervisor, Mission Analysis Software Engineering. Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109.

formulation not requiring derivation or implementation of any partial derivative Jacobian matrices. In light of these advantages of the sigma-point filter, it was sought in the present research to extend sigma-point estimation formulation (i.e. a Jacobian-free formulation, in which the state formal uncertainty is mapped in time by using nonlinear equations of motion to propagate a set of points from the surface of the covariance hyperellipsoid) to consider covariance analysis for dynamical systems.

The remainder of the paper will discuss a novel sigma-point consider filter algorithm developed as part of this research, with an example demonstration of its use. Aspects of implementing a consider covariance tool following this sigma-point approach will be compared with implementing consider covariance in standard sequential estimators. In particular it is shown that, consistent with the attributes of sigma-point estimators, a consider-parameterized sigma-point estimator can be developed entirely without requiring the derivation of any partial-derivative matrices related to the dynamical system, the measurements, or the considered parameters. This simplicity of formulation, with no loss in estimation accuracy, appears to be a distinct advantage over developing a consider covariance capability using a standard, linear-theory sequential estimation algorithm such as the EKF. As will be demonstrated, consider covariance analysis techniques such as sensitivity analyses can be addressed with this "partial-derivative-free" formulation.

First, an overview of the Unscented Kalman Filter (UKF) sigma-point algorithm in the next section, followed by a review of the linear-theory consider-covariance filter algorithm, provides context and a basis of comparison for the new sigma-point consider covariance filter algorithm. Then, the new sigma-point consider covariance filter algorithm is described and discussed. Finally, a numerical example comparing the use of the linear-theory consider filter with the sigma-point consider filter will illustrate that the two significantly different filter formulations result in identical results for a linear system, as expected.

OVERVIEW OF THE UNSCENTED KALMAN FILTER ALGORITHM

The unscented Kalman filter (UKF) is one member of the family of sigma-point filter algorithms. The UKF was developed as an alternative algorithm to the prevalently-used Extended Kalman Filter (EKF) for sequential nonlinear estimation. Like all sigma-point filters, the UKF features a *weighted statistical linear regression* technique for calculating terms in the Kalman update rule. Weighted statistical linear regression techniques linearize nonlinear functions of random variables while including information on the random variable's actual uncertainty. The random-variable function, of dimension L, is linearized by regressing between (2L + 1) points sampled from the a-priori uncertainty, that are updated using the full nonlinear models of the function. This differs from the standard linear-theory approach to estimation, in that the latter approach operates on truncated Taylor-series expansions about the expected value of the random-variable function, i.e. only using information from a single point. For further discussion of the theory of sigma-point filters and its distinctions from linear-theory filtering, the reader is referred to an article by van der Merwe and Wan[5].

Comprehensive discussions of the UKF origins and derivation, and demonstrations of its accuracy in solving nonlinear estimation problems, its runtime performance, and implementation aspects, in comparison with the EKF can be found in numerous references in the recent literature of applied estimation theory [c.f. 2, 5, 6, 8].. However, a summary of the algorithm and brief discussion of the UKF algorithm facilitates and frames the later discussion of consider-covariance extensions to the UKF. Summarizing the steps of the UKF implementation as described by Crassidis and Junkins[6], given the estimated parameter list $\mathbf{X}_{est} \in \Re^n$, its associated (n x n) error covariance matrix P_{k-1}^{xc+} at epoch t_{k-1}, plus (n x n) process noise matrix Q_{k-1}, we compute a set of (2n + 1) sigma points:

$$\sigma_{k-1} = 2n \text{ columns from} \pm \sqrt{(n+\lambda_x)P_{k-1}^{xx+} + Q_{k-1}} \tag{1}$$

$$\chi_{k-1,0} = X_{est,k-1}^{+} \tag{2}$$

$$\chi_{k-1,i} = X^{+}_{est,k-1} + \sigma_{k-1,i} \qquad i = 0,1,\cdots 2n \tag{3}$$

The (2n + 1) sigma points are mapped in time from epoch t_{k-1} to epoch t_k using the (n x 1) vector integral, **f**, of the nonlinear equations of motion for the dynamical system:

$$\chi_{k,i} = f\left(t_{k-1}, t_k, \chi_{k-1,i}, u_{k-1}\right), \qquad i = 0, 1, \cdots 2n \tag{4}$$

The predicted expected value of the estimated state is calculated as:

$$X^{-}_{est,k} = \frac{1}{n+\lambda_x}\left\{\lambda_x \chi_{k,o} + \frac{1}{2}\sum_{i=1}^{2n}\chi_{k,i}\right\} \tag{5}$$

The predicted state error covariance matrix (n x n) at epoch t_k is calculated as the weighted sum of auto-correlations of the vector differences between (2n + 1) predicted χ points and the predicted expected value of the state, i.e.:

$$P^{xx-}_{k} = \frac{1}{n+\lambda_x}\left\{\lambda_x\left(\chi_{k,o} - X^{-}_{est,k}\right)\left(\chi_{k,o} - X^{-}_{est,k}\right)^T + \frac{1}{2}\sum_{i=1}^{2n}\left(\chi_{k,i} - X^{-}_{est,k}\right)\left(\chi_{k,i} - X^{-}_{est,k}\right)^T\right\} \tag{6}$$

The mean modeled measurement is given by (m x 1) matrix, which is calculated as:

$$\hat{y}^{-}_{k} = \frac{1}{n+\lambda_x}\left\{\lambda_x \gamma_{k,o} + \frac{1}{2}\sum_{i=1}^{2n}\gamma_{k,i}\right\} \tag{7}$$

in which the (2n + 1) sigma points are mapped to measurement space by the (m x 1) vector nonlinear measurement model **h**:

$$\gamma_{k,i} = h\left(t_k, \chi_{k,i}, u_k\right), \qquad i = 0, 1, \cdots 2n \tag{8}$$

The cross-correlation matrix between the state errors and the associated measurement model errors is (n x m) matrix:

$$P^{xy}_{k} = \frac{1}{n+\lambda_x}\left\{\lambda_x\left(\chi_{k,o} - X^{-}_{est,k}\right)\left(\gamma_{k,o} - \hat{y}^{-}_{k}\right)^T + \frac{1}{2}\sum_{i=1}^{2n}\left(\chi_{k,i} - X^{-}_{est,k}\right)\left(\gamma_{k,i} - \hat{y}^{-}_{k}\right)^T\right\} \tag{9}$$

Now, (m x 1) measurement residual ν_k is computed by differencing the vector Y_k of measurements at epoch t_k with the mean modeled measurement:

$$\nu_k = Y_k - \hat{y}^{-}_{k} \tag{10}$$

The covariance of the modeled measurement errors (due to predicted state errors) is (m x m) matrix:

$$P^{yy}_{k} = \frac{1}{n+\lambda_x}\left\{\lambda_x\left(\gamma_{k,o} - \hat{y}^{-}_{k}\right)\left(\gamma_{k,o} - \hat{y}^{-}_{k}\right)^T + \frac{1}{2}\sum_{i=1}^{2n}\left(\gamma_{k,i} - \hat{y}^{-}_{k}\right)\left(\gamma_{k,i} - \hat{y}^{-}_{k}\right)^T\right\} \tag{11}$$

The covariance of the observation residual, which is a function of both intrinsic measurement noise covariance R (m x m) and also measurement model error covariance P^{yy}, (m x m), is simply:

$$P_k^{\nu\nu} = P_k^{yy} + R \tag{12}$$

The (n x m) optimal gain matrix mapping the measurement residuals into estimated state space (i.e. the Kalman gain) is:

$$K_k = P_k^{xy}\left(P_k^{\nu\nu}\right)^{-1} \tag{13}$$

Finally, the estimated parameter list $\mathbf{X}_{est}$ at epoch t_k and its associated error covariance P_k^{xc+} are updated with observation information by the equations:

$$X_{est}^{+}(t_k) = X_{est}^{-}(t_k) + K_k \nu_k \text{, and} \tag{14}$$

$$P_k^{xx+} = P_k^{xx-} - K_k P_k^{\nu\nu} K_k^T \tag{15}$$

Note that the UKF algorithm above requires no partial derivatives with respect to the parameter list of either the dynamical equations of motion (i.e. a state transition matrix) or the measurement model (i.e. the measurement partial derivative Jacobian). Also, note that Eqs. (12) through (15) constitute *generic* expressions for a Kalman filter algorithm. For instance, linear-theory versions of the Kalman filter (such as the Extended Kalman Filter algorithm, as presented, for example, in Tapley, Schutz and Born[1]) contains identical expressions for measurement model error covariance, Kalman gain, and update of parameter list and its error covariance, if one defines linear-theory versions of the two covariance terms below:

$$P_{linear}^{yy} \equiv H_x P^{xx-} H_x^T \text{, and} \tag{16}$$

$$P_{linear}^{xy} \equiv P^{xx-} H_x^T \tag{17}$$

H_x is the partial derivative Jacobian of the measurement model **h** with respect to the parameter list $\mathbf{X}_{est}$. With these definitions, for example, the linear-theory expression for the Kalman gain is seen to be identical to Equation (13) above:

$$K_{linear} = P^{xx-} H_x^T \left(H_x P^{xx-} H_x^T + R\right)^{-1} = P_{linear}^{xy}\left(P_{linear}^{\nu\nu}\right)^{-1} \tag{18}$$

LINEAR-THEORY CONSIDER KALMAN FILTER ALGORITHM

In order to provide a basis for discussing the extension of the sigma-point filter formulation to consider-covariance analysis, this section recapitulates (in brief summary, with minor changes in symbology) the algorithm for a linear-theory consider covariance filter algorithm that is fully derived in the text by Tapley, Schutz and Born[1].

In the problem of consider filter analysis, a vector $\mathbf{C} \in \Re^p$ of constant non-estimated parameters (which appear in the filter models of the system dynamics or of the measurements) has errors with constant (p x p) covariance matrix P^{cc}. The effect of the error in **C** on the (n x 1) estimated parameter list $\mathbf{X}_{est}$ is to increase the uncertainty P^{xx} of $\mathbf{X}_{est}$. The increased uncertainty is due to the additional effect of error covariance P^{cc} for consider parameter list **C**.

The list of constant consider parameters **C** augments the (n x 1) estimated parameter list $\mathbf{X}_{est}$, to form partitioned consider state vector, $\mathbf{X}_{cons} \in \Re^{n+p}$:

$$X_{cons} \equiv \begin{bmatrix} X_{est} \\ \hdashline C \end{bmatrix} \begin{matrix} (n \text{ x } 1) \\ (p \text{ x } 1) \end{matrix} \tag{19}$$

The associated consider covariance P_{cons} is partitioned according to $\mathbf{X}_{cons}$, i.e.

$$P_{cons} \equiv \left[\begin{array}{c:c} P^{xx}_{cons(n \text{ x } n)} & P^{xc}_{(n \text{ x } p)} \\ \hdashline P^{cx}_{(n \text{ x } p)} = \left(P^{xc}\right)^T & P^{cc}_{(p \text{ x } p)} \end{array}\right] \tag{20}$$

Note that the sub-matrix P^{xx}_{cons} associated with $\mathbf{X}_{est}$ in the context of considering errors in **C** will not in general be equal to the estimation-only covariance P^{xx} associated with estimating $\mathbf{X}_{est}$ (and neglecting errors in **C**). The relationship between these two matrices will be defined further below.

At epoch t_k, the inputs $X^+_{est,k-1}$, P^{xx+}_{k-1}, P^{xc+}_{k-1}, Y_k, R_k, Q_{k-1}, **C**, and P^{cc} are given, in which the symbols "-" and "+" reflect that a quantity is a best estimate, respectively before or after measurement information is incorporated, and in which:

Y_k (m x 1) is a vector of measurements made of the system at epoch t_k
R_k (m x m) is the covariance of intrinsic (vs. modeled) measurement noise, and
Q_{k-1} (n x n) is the process noise covariance for time interval $[t_{k-1}, t_k]$

With these input data, the linear-theory consider filter starts by propagating the state vector and covariance information through time from epoch t_{k-1} to epoch t_k by the following procedure.

$$X^-_{est,k} = f\left(t_{k-1}, t_k, X^+_{est,k-1}, C, u_{k-1}\right) \tag{21}$$

$$\Phi\left(t_{k-1}, t_k\right) = \int_{t_{k-1}}^{t_k} \frac{\partial f}{\partial X_{est}} \Phi\left(t_{k-1}, \tau\right) d\tau; \quad \Phi\left(t_{k-1}, t_{k-1}\right) = I \tag{22}$$

$$\theta\left(t_{k-1}, t_k\right) = \int_{t_{k-1}}^{t_k} \left[\frac{\partial f}{\partial X_{est}} \theta\left(t_{k-1}, \tau\right) + \frac{\partial f}{\partial C}\right] d\tau; \quad \theta\left(t_{k-1}, t_{k-1}\right) = 0 \tag{23}$$

$$P^{xx-}_k = \Phi\left(t_{k-1}, t_k\right) P^{xx+}_{k-1} \Phi^T\left(t_{k-1}, t_k\right) \tag{24}$$

$$P^{xc-}_k = \Phi\left(t_{k-1}, t_k\right) P^{xc+}_{k-1} + \theta\left(t_{k-1}, t_k\right) P^{cc}, \quad P^{xc+}_0 = 0 \tag{25}$$

$$P^{xx-}_{cons,k} = P^{xx-}_k + P^{xc-}_k \left(P^{cc}\right)^{-1} \left(P^{xc-}_k\right)^T \tag{26}$$

in which:

f is the (n x 1) vector integral of the equations of motion
$\Phi(t_{k-1}, t_k)$ is the state transition matrix (n x n) for the system

$\theta(t_{k-1}, t_k)$ is the state transition matrix (n x p) for the system with respect to **C**
H_x (m x n)is the partial derivative of the measurement model **h** with respect to $\mathbf{X}_{est}$
H_c (m x p) is the partial derivative of the measurement model **h** with respect to **C**

Next, the a-priori values of the augmented parameter list and the consider covariance are innovated with measurement data at epoch tk:

$$K_k = P_k^{xx-} H_{x,k}^T (H_{x,k} P_k^{xx-} H_{x,k}^T + R_k)^{-1} \tag{27}$$

$$\nu_k = Y_k - h\left(t_k, X_{est,k-1}^-, u_k\right) \tag{28}$$

$$X_{est,k}^+ = X_{est,k}^- + K_k \nu_k \tag{29}$$

$$P_k^{xx+} = \left(I - K_k H_{x,k}\right) P_{k-1}^{xx-} = P_{k-1}^{xx-} - K_k (H_{x,k} P_k^{xx-} H_{x,k}^T + R_k) K_k^T \tag{30}$$

$$P_k^{xc+} = \left(I - K_k H_{x,k}\right) P_k^{xc-} - K_k H_c P^{cc} \tag{31}$$

$$P_{cons,k}^{xx+} = P_k^{xx+} + P_k^{xc+} \left(P^{cc}\right)^{-1} \left(P_k^{xc+}\right)^T \tag{32}$$

The above algorithm is applied recursively in time, until all measurements have been applied or a desired final epoch has been reached.

SIGMA-POINT CONSIDER KALMAN FILTER ALGORITHM

The sigma-point consider filter algorithm discussed here was developed to enable sigma-point filters to be used in conditions when a conservative, consider-analysis approach to estimation is needed. Thus, a consider filtering algorithm was sought following the formulation of sigma-point filters, that would not require the derivation of partial-derivative terms, but would provide the same results as the linear-theory filters for linear problems.

The sigma-point consider filter algorithm is as follows. As with the linear-theory consider filter, a (p x 1) list **C** of constant, non-estimated parameters whose errors are to be considered augments the (n x 1) estimated parameter list $\mathbf{X}_{est}$, to form a partitioned consider state vector, $\mathbf{X}_{cons}$:

$$X_{cons} \equiv \begin{bmatrix} X_{est} \\ \hline C \end{bmatrix} \begin{matrix} (n \times 1) \\ (p \times 1) \end{matrix} \tag{33}$$

Also as with the linear-theory consider filter, the associated consider covariance P_{cons} is partitioned according to $\mathbf{X}_{cons}$, i.e.

$$P_{cons} \equiv \left[\begin{array}{c|c} P_{cons(n \times n)}^{xx} & P_{(n \times p)}^{xc} \\ \hline P_{(n \times p)}^{cx} = \left(P^{xc}\right)^T & P_{(p \times p)}^{cc} \end{array}\right] \tag{34}$$

Now, adapting the UKF formulation approach to computing a-priori sigma points in an unscented filter, first we factorize P_{cons} (best estimate of covariance from previous epoch t_{k-1}) into $S_{cons,k-1}(S_{cons,k-1})^T$. In order to preserve the constancy of P^{cc}, the lower-triangular square matrix $S_{cons,k-1}$ is computed using a block-Cholesky decomposition[7], with the initial pivot starting in the P^{cc} partition.

$$S_{cons,k-1} = \left[\begin{array}{c|c} S^{xx}_{cons,k-1} & 0_{(n\ x\ p)} \\ \hline S^{cx}_{k-1} & S^{cc}_{k-1} \end{array}\right] \tag{35}$$

where

$$S^{cc}_{k-1(p\ x\ p)} \equiv \sqrt{(\lambda_c + p)P^{cc}} = \text{const.} \tag{36}$$

$$S^{cx}_{k-1(p\ x\ n)} \equiv \left(S^{cc}_{k-1}\right)^{-1} P^{cx}_{k-1}\text{, and} \tag{37}$$

$$S^{xx}_{k-1(n\ x\ n)} \equiv \sqrt{(\lambda_x + n)P^{xx}_{cons,k-1} - \left(S^{cx}_{k-1}\right)^T S^{cx}_{k-1}}\,. \tag{38}$$

and in which $\lambda_c = 3 - p$, and $\lambda_x = 3 - n$. Equation (5) for S^{cx} will always be analytic, because sub-matrix S^{cc} is a square root of constant positive definite matrix P^{cc}.

As with the linear-theory consider algorithm, the sigma-point consider filter requires that $\mathbf{X}_{est}$ and P^{xx} be estimated with the standard sequential estimator, in this case, the UKF algorithm described in the section above.

Now, in addition to estimating this nth order system, to compute the additional uncertainty from consider the non-estimating parameters, at epoch t_{k-1} we generate (2p + 1) consider sigma-points, which are a set[1] of samples of system state realizations at its a-priori expected value and also at 2p select points on the uncertainty ellipsoid:

$$\chi_{cons,k-1,i} = X^{+}_{k-1} + \sigma_{cons,k-1,i}\,, \qquad i = 0, 1, \cdots 2p \tag{39}$$

where $\sigma_{cons,k-1,0} = 0$, and $\sigma_{cons,k-1,i}$ = ±columns (n + 1) through (n + p) of $(S_{cons,k-1})^T$, containing the square-root of the dispersed values of consider parameter list **C** plus expected variations in a-priori estimated parameter list $\mathbf{X}_{est,k-1}$ due to cross-correlations with **C**.

These (2p + 1) consider sigma-points are then propagated from epoch t_{k-1} to epoch t_k, using the ((n + p) x 1) vector of nonlinear expressions for the time-evolution of the augmented state vector, e.g. the equations of motion for the dynamical system, augmented with a (p x 1) list of zeros to maintain the constancy of the consider parameters. In this manner, we obtain the (2p + 1) predicted sigma-points at epoch t_k:

$$\chi_{cons,k,i} = f\left(t_{k-1}, t_k, \chi_{cons,k-1,i}, u_{k-1}\right), \qquad i = 0, 1, \ldots 2p \tag{40}$$

Per the unscented Kalman filter formulation, the predicted mean consider state is computed using the predicted sigma-points as:

$$X^{-}_{cons,k} = \frac{1}{p + \lambda_c}\left(\lambda_c \chi_{cons,k,o} + \frac{1}{2}\sum_{i=1}^{2p} \chi_{cons,k,i}\right) \tag{41}$$

Then, the predicted consider covariance is computed as:

[1] For parameter list $X \in R^L$ belonging to a multivariate probability distribution, 2L + 1 samples is the minimum needed to capture the first two moments (expected value and covariance) of the distribution.

$$P^{-}_{cons,k} = \left[\begin{array}{c|c} P^{xx-}_{cons,k} & P^{xc-}_{k} \\ \hline P^{cx-}_{k} & P^{cc} \end{array}\right] = \frac{1}{p+\lambda_c}\left\{\lambda_c\left(\chi_{cons,k,o} - X^{-}_{cons,k}\right)\left(\chi_{cons,k,o} - X^{-}_{cons,k}\right)^T + \frac{1}{2}\sum_{i=1}^{2p}\left(\chi_{cons,k,i} - X^{-}_{cons,k}\right)\left(\chi_{cons,k,i} - X^{-}_{cons,k}\right)^T\right\} \tag{42}$$

The primary utility of matrix $P^{-}_{cons,k}$ is that it yields the (n x p) sub-matrix P^{xc-}_{k} of predicted cross-correlations between consider and estimated parameters. At epochs for which measurements are available, the change due to the measurement information on the cross-correlation sub-matrix is calculated the following way:

$$P^{xc+}_{k} = P^{xc-}_{k} - K_k P^{\nu\nu}_{k} K_k{}^{T}\left(P^{xx-}_{k}\right)^{-1} P^{xc-}_{k} \tag{43}$$

The optimal gain K_k, residual covariance $P^{\nu\nu}_{k}$ and a-priori state covariance P^{xx-}_{k} are all obtained from a standard sigma-point filter (that estimates $\mathbf{X}_{est}$ and neglects errors in **C**) operating simultaneously in time with the consider filter algorithm. The above cross-correlation update expression serves the same purpose as, and is similar to, the cross-correlation update step (Eq. (31)) in the linear-theory consider algorithm. Finally, the additive uncertainty from considering non-estimated parameters is calculated as:

$$dP_{cons,k} = P^{xc+}_{k}\left(P^{cc}\right)^{-1}\left(P^{xc+}_{k}\right)^{T} \tag{44}$$

and the updated consider covariance matrix, post-measurement, for epoch t_k is given by:

$$P_{cons,k} \equiv \left[\begin{array}{c|c} P^{xx+}_{cons} & P^{xc+} \\ \hline P^{cx+} & P^{cc} \end{array}\right] \tag{45}$$

in which the state covariance augmented for consider-parameter uncertainty ($P^{xx+}_{cons,k}$) is related to the standard state covariance P^{xx+}_{k} by:

$$P^{xx+}_{cons,k} = P^{xx+}_{k} + dP_{cons,k} \tag{46}$$

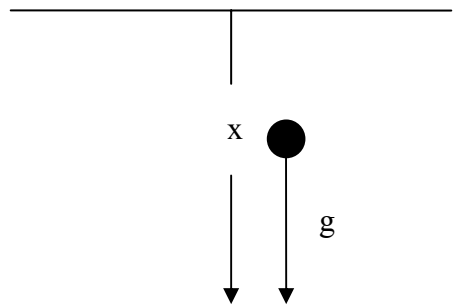
If no observation is available to be processed at time t_k, Equations (44) and (45) are evaluated using P^{xx-}_{k} and P^{xc-}_{k} instead of P^{xx+}_{k} and P^{xc+}_{k}. This algorithm is recursive, so that after calculating $P_{cons,k}$, if filter processing is to continue, the time index k is set to (k-1), and the steps above are repeated at the next processing epoch, until all measurements have been applied or the desired end time has been reached.

EXAMPLE CONSIDER COVARIANCE ANALYSES FOR A LINEAR SYSTEM

A simple linear example problem appearing in the text by Tapley et al,[1] is extended to the sigma-point consider filter, to illustrate the differences in the computational steps of the linear-theory and sigma-point versions of the consider filter algorithm, and to compare the resulting solutions.

In the example problem, a point mass accelerates in one dimension, x, under the influence of a constant acceleration g. The system is observed at two epochs, t = 0 and t = 1, with a one-dimensional range from the origin to the point mass location. The effect of error in the value of the constant

acceleration on a sequential estimate of the location and velocity of the point mass will be considered in this example.



For both example filters, the system model is as follows:

$$\text{State vector } X(t) = \begin{bmatrix} x \\ v \end{bmatrix} \tag{47}$$

The system dynamics model is given by:

$$\dot{X}(t) = \begin{bmatrix} v \\ g \end{bmatrix}, \text{and } X(t_2) = f(t_1, t_2, X_1, u_1) = \begin{bmatrix} x_1 + v_1(t_2 - t_1) + \dfrac{g(t_2 - t_1)^2}{2} \\ v_1 + g(t_2 - t_1) \end{bmatrix} \tag{48}$$

$$\text{Scalar measurement model } G(t, X(t), u(t)) = x. \tag{49}$$

and let g = 10, arbitrarily. Also, the following data are used in both example filters:

$$\text{Initial (true) state vector } X(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\text{State vector initial covariance } P^{xx}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Intrinsic measurement noise R(t) = const = 1

$$\text{Process noise } Q(t) = \text{const} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Finally, there is an error in the model of constant "gravity" acceleration g = 1; and so the error in g is considered in both the linear-theory and sigma-point consider filters. The cross-correlation P^{xc} between the initial estimated state uncertainty and the error in g is assumed to be zero.

$$P^{cc} = \text{const} = 1$$

$$P^{xc}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Next, the consider covariance is computed at times t = 0 and t = 1 using both a sequential linear-theory formulation and the sigma-point formulation.

(1) Linear-Theory Sequential Consider Filter Solution:

This linear-theory filter solution is replicated from Tapley et al[1], pages 416-419, and presented here to compare with the sigma-point filter solution below. First, the Jacobian matrices required for the sequential linear-theory consider filter are calculated. The partial derivative matrices of the measurement with respect to the estimated state list (H_x), and also with respect to the consider parameter list (H_c), are:

$$H_x = [dx/dx \;\; dx/dv] = [1 \;\; 0]$$

$$H_c = [dx/dg] = 0$$

And the partial derivatives of the state vector at epoch t_2 with respect to the state vector at epoch t_1 (Φ, i.e. the state transition matrix), and also with respect to the consider parameter list (θ), are:

$$\Phi(t_1, t_2) = \begin{bmatrix} 1 & (t_2 - t_1) \\ 0 & 1 \end{bmatrix}$$

$$\theta(t_1, t_2) = \begin{bmatrix} (t_2 - t_1)^2 / 2 \\ (t_2 - t_1) \end{bmatrix}$$

So, for this example, with a fixed interval of 1 second between t_1 and t_2,

$$\Phi(t_1, t_1 + 1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\theta(t_1, t_1 + 1) = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

At epoch t = 0, following the linear-theory consider filter algorithm, the a-priori covariance is updated with the range measurement of t = 0. First, the Kalman gain at time t = 0 is calculated:

$$K(t = 0) = P^{xx}(t = 0) H_x{}^T \left(H_x P^{xx}(t = 0) H_x{}^T + R \right)^{-1} = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}$$

and using the Kalman gain, the covariance is updated with measurement information:

$$P^{xx+}(t = 0) = [I - K(t = 0) H_x] P^{xx}(t = 0) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

The updated value of the cross-correlation sub-matrix at t = 0 is:

$$P_k^{xc+}(t = 0) = (I - K(t = 0) H_x) P^{xc}(t = 0) - K(t = 0) H_c P^{cc} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

At epoch t = 1, following the linear-theory consider filter algorithm, the best-estimated covariance from time t = 0 is propagated to time t = 1:

$$P^{xx-}(t = 1) = \Phi(0,1) P^{xx+}(t = 0) \Phi^T(0,1) + Q = \begin{bmatrix} 3/2 & 1 \\ 1 & 1 \end{bmatrix}$$

The Kalman gain at time t = 1 is:

$$K(t=1) = P^{xx-}(t=1)H_x{}^T\left[H_x P^{xx-}(t=1)H_x{}^T + R\right]^{-1} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$

and we update the covariance with measurement information:

$$P^{xx+}(t=1) = \left[I - K(t=1)H_x\right]P^{xx-}(t=1) = \begin{bmatrix} 3/5 & 2/5 \\ 2/5 & 3/5 \end{bmatrix}$$

Then, cross-correlation matrix P^{xc} is propagated by:

$$P_k^{xc-}(t=1) = \Phi(0,1)P^{xc+}(t=0) + \theta(0,1)P^{cc} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

and then is also updated, by:

$$P^{xc+}(t=1) = \left(I - K_k(t=1)H_x\right)P_k^{xc-}(t=1) - K(t=1)H_c P^{cc} = \begin{bmatrix} 1/5 \\ 4/5 \end{bmatrix}$$

Finally, the consider covariance is

$$P_{cons}^{xx+}(t=1) = P^{xx+} + P^{xc+}\left(P^{cc}\right)^{-1}\left(P^{xc+}\right)^T = \begin{bmatrix} 0.64 & 0.56 \\ 0.56 & 1.24 \end{bmatrix}$$

(2) Sigma-Point Consider Filter Solution:

In the sigma-point consider formulation, as noted above, there are no state transition matrices, or partial derivatives of the measurement models, and we start by defining the augmented state vector and consider covariance matrix:

$$X_{cons}(\text{initial}) = \begin{bmatrix} X_{est} \\ \hline C \end{bmatrix}_{t=0} = \begin{bmatrix} x \\ v \\ \hline g \end{bmatrix}_{t=0} = \begin{bmatrix} 1 \\ 0 \\ \hline 10 \end{bmatrix}, \text{ and}$$

$$P_{cons}(\text{initial}) = \left[\begin{array}{c|c} P_{cons}^{xx} & P^{xc} \\ \hline P^{cx} & P^{cc} \end{array}\right]_{t=0} = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$

Now for this system, n = 2 and p = 1, so $\lambda_x = 1$ and $\lambda_c = 2$. Next, per the definitions in Eqs. (36) through (38), we take the lower-triangular block-Cholesky root of P_{cons}, prior to incorporating the first measurement at t = 0:

$$S_{cons}(\text{initial}) = \left[\begin{array}{c|c} S_{cons}^{xx} & 0 \\ \hline S^{cx} & S^{cc} \end{array}\right] = \left[\begin{array}{cc|c} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ \hline 0 & 0 & \sqrt{3} \end{array}\right]$$

The initial sigma points at time t = 0, which are positive-and-negative vector pairs for each column of the S_{cons} matrix, added to the a-priori best-estimated state vector, are calculated as:

$$\chi(t=0) = X_{cons}(\text{initial}) + \sigma_i, \quad i = 0, 1, \cdots 2n$$

$$= \left\{ \begin{bmatrix} 1 \\ 0 \\ \hline 10 \end{bmatrix}, \begin{bmatrix} 1+\sqrt{3} \\ 0 \\ \hline 10 \end{bmatrix}, \begin{bmatrix} 1-\sqrt{3} \\ 0 \\ \hline 10 \end{bmatrix}, \begin{bmatrix} 1 \\ 0+\sqrt{3} \\ \hline 10 \end{bmatrix}, \begin{bmatrix} 1 \\ 0-\sqrt{3} \\ \hline 10 \end{bmatrix} \right\}$$

and the initial consider sigma points, incorporating just the consider partition of S_{cons}, are:

$$\chi_{cons}(t=0) = X_{cons}(\text{initial}) + \sigma_{cons,i}, \quad i = 0, 1, \cdots 2p = \left\{ \begin{bmatrix} 1 \\ 0 \\ \hline 10 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \hline 10+\sqrt{3} \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ \hline 10-\sqrt{3} \end{bmatrix} \right\}$$

The (2n + 1) measurement points at time t = 0, corresponding to dispersions of the nonlinear measurement model G(t, X(t), u(t)) from Eq. (49), are:

$$\begin{aligned} \gamma(t=0) = &\quad G\left(t=0,\ \chi_i(t=0),\ u(t=0)=0\right), \quad i = 0, 1, \cdots 2n \\ = &\quad \left\{1,\quad 1+\sqrt{3},\quad 1-\sqrt{3},\quad 1,\quad 1\right\} \end{aligned}$$

The mean predicted measurement is given by Eq. (7):

$$\hat{y}^{-}(t=0) = \frac{1}{n+\lambda_x}\left\{\lambda_x \gamma_{i=o}(t=0) + \frac{1}{2}\sum_{i=1}^{2n}\gamma_i(t=0)\right\} = 1$$

The cross-correlation between the a-priori state vector dispersion and the measurement dispersion is calculated using Eq. (9), and results in value:

$$P^{xy}\left(t=0\right) = \begin{bmatrix} 1 \\ 0 \\ \hline 0 \end{bmatrix}$$

The measurement model error covariance is given by Eq. (11):

$$P^{yy}\left(t=0\right) = \left[1\right]$$

The residual covariance is given by Eq. (12):

$$P^{\nu\nu}\left(t=0\right) = P^{yy}(t=0) + R = \left[2\right]$$

and the Kalman gain is calculated using Eq. (13) as:

$$K\left(t=0\right) = P^{xy}(t=0)\left(P^{\nu\nu}(t=0)\right)^{-1} = \begin{bmatrix} 1/2 \\ 0 \\ \hline 0 \end{bmatrix}$$

The covariance matrix, augmented with the cross-correlation and consider parameter covariance sub-matrices at t = 0, and updated with the first measurement, is given by Eq. (15):

$$P^{xx+}\left(t=0\right) = P^{xx+}\left(t=0\right) = \left[\begin{array}{cc|c} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$

from which we see that the state-consider-parameter cross-correlation submatrix is:

$$P^{xc+}(t=0)=\begin{bmatrix}0\\0\end{bmatrix}$$

Thus, per Eq. (44), the increase in estimated covariance due to the considered parameters is zero for the initial time step. Next, a new, a-posteriori block-Cholesky root is taken of $P^{xx+}(t = 0)$, resulting in:

$$S_{cons}(t=0)=\left[\begin{array}{cc|c}\sqrt{3/2} & 0 & 0\\ 0 & \sqrt{3} & 0\\ \hline 0 & 0 & \sqrt{3}\end{array}\right]$$

Then, $S_{cons}(t = 0)$ is used to calculate new sigma points, which will be used to propagate the state and covariance forward in time from t = 0 to t = 1:

$$\begin{aligned}\chi(t=0) &= X_{cons}(t=0)+\sigma_i, \quad i=0,1,\cdots 2n\\ &=\left\{\begin{bmatrix}1\\0\\\hline 10\end{bmatrix},\begin{bmatrix}1+\sqrt{3/2}\\0\\\hline 10\end{bmatrix},\begin{bmatrix}1-\sqrt{3/2}\\0\\\hline 10\end{bmatrix},\begin{bmatrix}1\\0+\sqrt{3}\\\hline 10\end{bmatrix},\begin{bmatrix}1\\0-\sqrt{3}\\\hline 10\end{bmatrix}\right\}\end{aligned}$$

and the initial consider sigma points, incorporating just the consider partition of S_{cons}, are:

$$\chi_{cons}(t=0)=X_{cons}(t=0)+\sigma_{cons,i}, \quad i=0,1,\cdots 2p=\left\{\begin{bmatrix}1\\0\\\hline 10\end{bmatrix},\begin{bmatrix}1\\0\\\hline 10+\sqrt{3}\end{bmatrix},\begin{bmatrix}1\\0\\\hline 10-\sqrt{3}\end{bmatrix}\right\}$$

The sigma points χ and χ_{cons} are propagated in time from t = 0 to t = 1, using the system dynamics model, Eq. (48):

$$\begin{aligned}\chi(t=1) &= f\big(t=0,\ t=1,\ \chi(t=0),\ u(t=0)=0\big)\\ &=\left\{\begin{bmatrix}6\\10\\\hline 10\end{bmatrix},\begin{bmatrix}6+\sqrt{3/2}\\10\\\hline 10\end{bmatrix},\begin{bmatrix}6-\sqrt{3/2}\\10\\\hline 10\end{bmatrix},\begin{bmatrix}6+\sqrt{3}\\10+\sqrt{3}\\\hline 10\end{bmatrix},\begin{bmatrix}6-\sqrt{3}\\10-\sqrt{3}\\\hline 10\end{bmatrix}\right\}\end{aligned}$$

and

$$\chi_{cons}(t=1)=f\big(t=0,t=1,\chi_{cons}(t=1),u(t=0)=0\big)=\left\{\begin{bmatrix}6\\10\\\hline 10\end{bmatrix},\begin{bmatrix}6+\sqrt{3}/2\\10+\sqrt{3}\\\hline 10+\sqrt{3}\end{bmatrix},\begin{bmatrix}6-\sqrt{3}/2\\10-\sqrt{3}\\\hline 10-\sqrt{3}\end{bmatrix}\right\}$$

The quantities $P^{xx-}(t = 1)$, $P^{xy}(t = 1)$, $P^{yy}(t = 1)$, and $P^{vv}(t = 1)$ are computed in the same manner as for t = 0, using the sigma points $\chi(t = 1)$. $P^{yy}(t = 1)$, $P^{vv}(t = 1)$, the Kalman gain K(t = 1), and $P^{xx-}(t = 1)$ are found to be:

$$P^{yy}(t=1)=[3/2]$$

$$P^{vv}(t=1)=P^{yy}(t=1)+R=[5/2]$$

$$K(t=1)=\begin{bmatrix}3/5\\2/5\\0\end{bmatrix}$$

$$P^{xx-}(t=1)=\left[\begin{array}{cc|c}3/2 & 1 & 0\\1 & 1 & 0\\\hline 0 & 0 & 0\end{array}\right]$$

in the same manner as was done for step t = 0. However, now the consider version of the expected value of the state vector is found using $\chi_{cons}(t = 1)$, per Eq. (41):

$$X^{-}_{cons}(t=1)=\begin{bmatrix}6\\10\\10\end{bmatrix}$$

and the predicted consider covariance matrix, per Eq. (42), is:

$$P^{-}_{cons}(t=1)=\left[\begin{array}{c|c}P^{xx-}_{cons} & P^{xc-}\\\hline P^{cx-} & P^{cc}\end{array}\right]=\left[\begin{array}{cc|c}1/4 & 1/2 & 1/2\\1/2 & 1 & 1\\\hline 1/2 & 1 & 1\end{array}\right]$$

From the predicted consider covariance matrix, which the predicted cross-correlation sub-matrix is seen to be:

$$P^{xc-}(t=1)=\begin{bmatrix}1/2\\1\end{bmatrix}$$

Then, from Eq. (43), the updated value of the cross-correlation sub-matrix is calculated to be:

$$P^{xc+}(t=1)=P^{xc-}-KP^{vv}K^{T}\left(P^{xx-}\right)^{-1}P^{xc-}_{k}=\begin{bmatrix}1/5\\4/5\end{bmatrix}$$

Finally, applying Eqs. (15), (44) and (46), the estimate-only and consider covariance values for epoch t = 1 are found to be, respectively:

$$P^{xx+}(t=1)=\left[\begin{array}{cc|c}3/5 & 2/5 & 0\\2/5 & 3/5 & 0\\\hline 0 & 0 & 0\end{array}\right],$$

$$dP_{cons}(t=1)=P^{xc+}\left(P^{cc}\right)^{-1}\left(P^{xc+}\right)^{T}=\left[\begin{array}{cc|c}1/25 & 4/25 & 0\\4/25 & 16/25 & 0\\\hline 0 & 0 & 0\end{array}\right],\text{ and}$$

$$P^{xx+}_{cons}(t=1)=\left[\begin{array}{cc|c}0.64 & 0.56 & 1/5\\0.56 & 1.24 & 4/5\\\hline 1/5 & 4/5 & 1\end{array}\right]$$

Note that, despite the very different formulations, the sigma-point consider filter and the linear-theory consider filter yield identical answers for this example involving a linear system. The advantages of the sigma-point formulation are that the user-supplied inputs are fewer and simpler (i.e. only the nonlinear system dynamics model and the nonlinear measurement model are needed) than for the linear-theory filter (which requires the same nonlinear models, plus the partial derivatives of the nonlinear models with respect to the state vector and considered parameter list). Hence, the derivation and implementation of the sigma-point consider filter requires fewer formulation steps than the linear-theory consider filter. No partial-derivative code needs be debugged, documented or maintained. Moreover, the computational run-time throughput is similar for the two consider covariance analysis techniques.

RESULTS AND CONCLUSIONS

A novel algorithm for a sigma-point consider filter has been described, and demonstrated with a comparison to a simple, standard-form consider filter example from the published estimation theory literature. The algorithm enables consider covariance using a sigma-point filter framework, in which no partial derivatives of dynamical models or measurement models are necessary to achieve the same results as linear-theory filters on linear problems. An example problem was used to demonstrate that equivalent answers are provided by linear-theory and sigma-point versions of a sequential consider filter when applied to a linear problem. Sigma-point filters have been demonstrated [2,3,4,5] to provide more accurate state estimates and covariance representations than linear-theory filters such as extended Kalman filters (EKF's), when applied to estimation of systems with nonlinear dynamics and/or measurement models.

Assessment of the sigma-point consider filter when applied to nonlinear problems, and the overall utility of this filtering technique, is a part of planned future work. This future work will include using the sigma-point consider filter in processing inertial measurement unit (IMU) data telemetered to Earth by the Mars Reconnaisance Orbiter during its aerobraking operations starting in March, 2006. Consider covariance analysis capability is a significant new addition to sigma-point estimation theory and algorithms, as consider covariance is essential for ground-based navigation of deep-space vehicles, which is an important prospective future application for sigma-point filters. Also included in future work is to explore re-casting the sigma-point consider algorithm into factorized form, the feasibility of which is indicated by the square-root UKF algorithm reported by van der Merwe and Wan[8].

ACKNOWLEDGEMENTS

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. Tapley, B., B. Schutz, G. Born. Statistical Orbit Determination. Elsevier Academic Press, 2004. pp. 387-437.
2. Julier, S. and J. Uhlmann. *Unscented Filtering and Nonlinear Estimation*, Proceedings of the IEEE, vol. 92, no. 3, March 2004.
3. Jah, M. and M. Lisano. Mars Aerobraking Spacecraft State Estimation By Processing Inertial Measurement Unit Data. NASA Tech Publication NASA/TP 2005-213901, in revision, Dec. 2005.
4. Crassidis, J.L., and F.L. Markley. *Unscented Filtering for Spacecraft Attitude Estimation*, AIAA Journal of Guidance, Control, and Dynamics, Vol. 26, No. 4, July-Aug. 2003, pp. 536-542.
5. van der Merwe, R. and E. Wan. *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*, in Proceedings of the Workshop on Advances in Machine Learning,Montreal, Canada, June 2003.
6. Crassidis, J. and J. Junkins. Optimal Estimation of Dynamic Systems. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science, Vol. 2, 2004.

7. Golub. G. and C. van Loan. Matrix Computations, 2nd Edition. Baltimore, MD, Johns Hopkins University Press, 1989, p. 144.
8. van der Merwe, R. and E. Wan. *The Square-Root Unscented Kalman Filter for State and Parameter Estimation*, in International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, May 2001.