

Nonlinear State Estimation

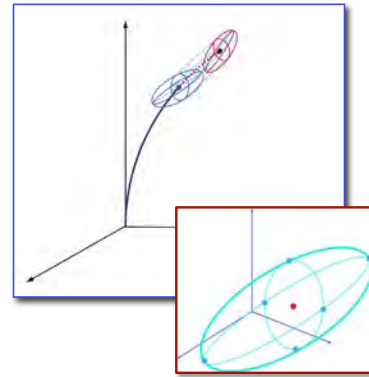
Sigma Points Filter

Robert Stengel

Optimal Control and Estimation, MAE 546

Princeton University, 2015

- Sigma Points (“Unscented Kalman”) nonlinear filter
 - Transformation of uncertainty
 - Propagation of mean and variance
- Helicopter state estimation example
- Introduction to advanced nonlinear filters



Copyright 2015 by Robert Stengel. All rights reserved. For educational use only.
<http://www.princeton.edu/~stengel/MAE546.html>
<http://www.princeton.edu/~stengel/OptConEst.html>

1

Criticisms of the Basic Extended Kalman Filter*

- State estimate **prediction** is deterministic, i.e., not based on an expectation (actually not true)
- State estimate **update** is linear
- **Jacobians** must be evaluated to calculate covariance prediction and update
- Not all comments apply to iterated, quasilinear or adaptive extended Kalman filters

* Julier and Uhlmann, 1997; van der Werpe and Wan, 2004

2

Transformation of Uncertainty

Nonlinear transformation of a random variable

$$\mathbf{x} : \text{Random variable with mean, } \bar{\mathbf{x}}, \text{ and covariance, } \mathbf{P}_{\mathbf{xx}} \\ \mathbf{y} = \mathbf{f}[\mathbf{x}]$$

Estimate the mean and covariance of the transformation's output

$$\bar{\mathbf{y}}(\bar{\mathbf{x}}, \mathbf{P}_{\mathbf{xx}}) \text{ and } \mathbf{P}_{\mathbf{yy}}(\bar{\mathbf{x}}, \mathbf{P}_{\mathbf{xx}})$$

The transformation is said to be “**unscented**”^{*}
if its probability distribution is
Consistent, Efficient, and Unbiased

^{*} Julier and Uhlmann, 1997

3

Consistent Estimate of a Dynamic State

$$\text{Let } \mathbf{x}_k \triangleq \mathbf{x}, \quad \mathbf{x}_{k+1} \triangleq \mathbf{y}$$

$$\bar{\mathbf{x}}_{k+1}(\bar{\mathbf{x}}_k, \mathbf{P}_{\mathbf{x}_k \mathbf{x}_k}) = \bar{\mathbf{y}}(\bar{\mathbf{x}}, \mathbf{P}_{\mathbf{xx}}) \quad \text{and} \quad \mathbf{P}_{\mathbf{x}_{k+1} \mathbf{x}_{k+1}}(\bar{\mathbf{x}}_k, \mathbf{P}_{\mathbf{x}_k \mathbf{x}_k}) = \mathbf{P}_{\mathbf{yy}}(\bar{\mathbf{x}}, \mathbf{P}_{\mathbf{xx}})$$

Consistent state estimate converges in the limit

$$\left\{ \mathbf{P}_{\mathbf{x}_{k+1} \mathbf{x}_{k+1}} - E \left[(\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1})^T \right] \right\} \geq \mathbf{0} \\ \{ \text{Estimated Covariance} - \text{Actual Covariance} \} \geq \mathbf{0}$$

Lesson: In filtering, add sufficient “process noise” to the filter gain computation to prevent filter divergence

4

Adding Process Noise Improves Consistency

$$\left\{ \mathbf{P}_{\mathbf{x}_{k+1}|\mathbf{x}_{k+1}} - E \left[(\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1})(\mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1})^T \right] \right\} \geq \mathbf{0}$$

{Estimated Covariance – Actual Covariance} ≥ 0

- **Satellite orbit determination**
 - Aerodynamic drag produced unmodeled bias
 - Optimal filter did not estimate bias
- **Process noise increased for filter design**
 - Divergence was contained

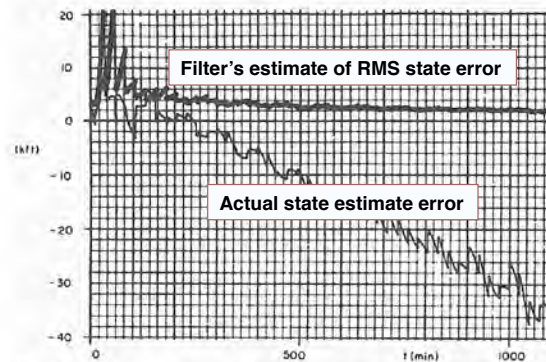


Fig. 6. Divergence due to ignoring drag in satellite navigation (forward component of position error).

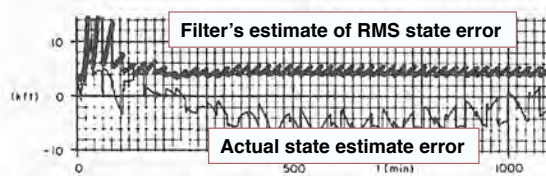


Fig. 7. Elimination of divergence by increasing assumed process noise.

Fitzgerald, 1971

5

Efficient and Unbiased Estimate of a Dynamic State

Efficient state estimator converges more quickly than an inefficient estimator

$$\min_{\text{Added Process Noise}} \left\{ \text{Estimated Covariance} - \text{Actual Covariance} \right\}$$

Add “just enough” process noise
Unbiased estimate

$$\bar{\mathbf{x}}_{k+1} = E(\mathbf{x}_{k+1}) \quad [\text{Estimated Mean} = \text{Actual Mean}]$$

6

Recall: Experimental Determination of Mean and Variance

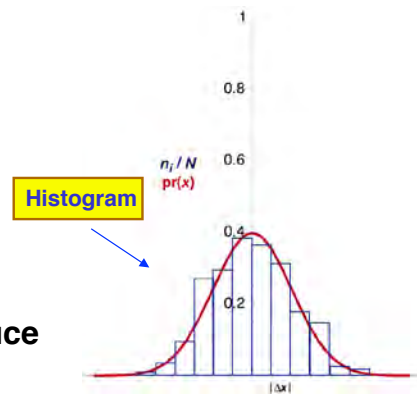
- Sample mean for N data points, x_1, x_2, \dots, x_N

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$

- Sample variance for same data set

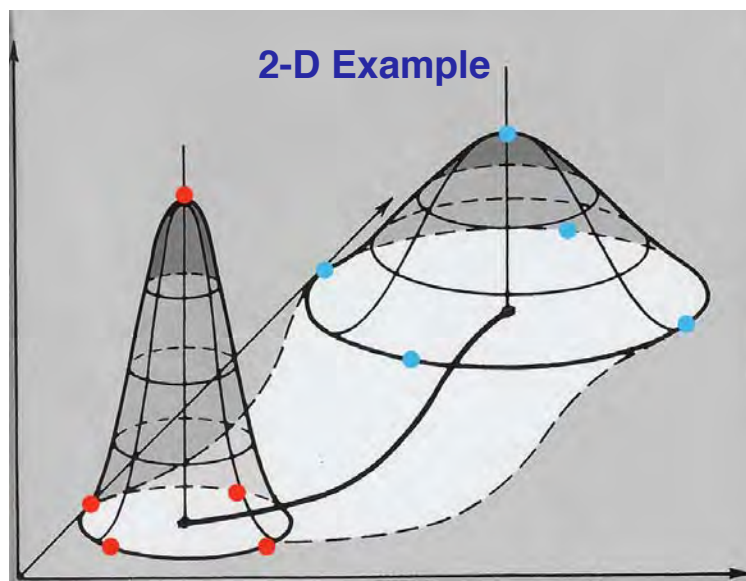
$$\sigma_x^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{(N-1)}$$

- Divisor is $(N-1)$ rather than N to produce an unbiased estimate
 - $(N-1)$ terms are independent
 - Inconsequential for large N
- Distribution is not necessarily Gaussian



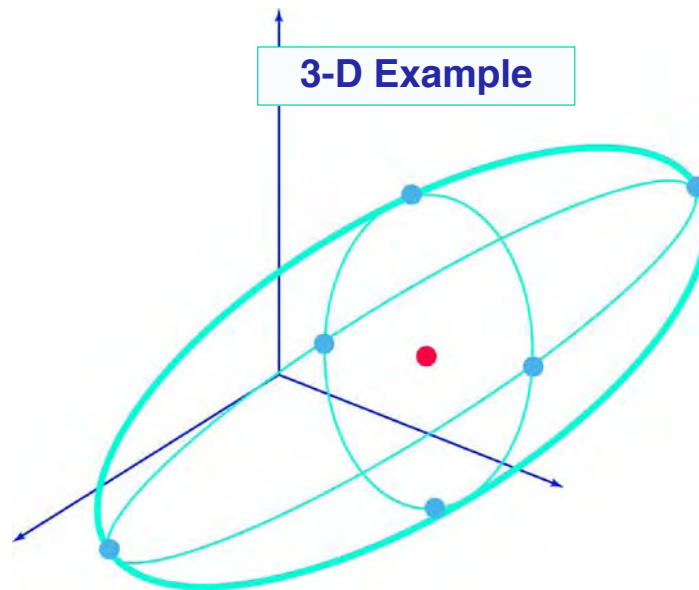
7

Sigma Points of P_{xx} and Mean



8

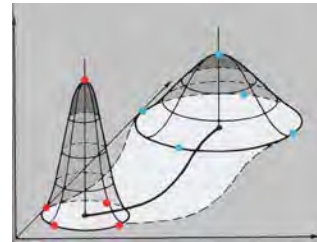
Sigma Points of \mathbf{P}_{xx} and Mean



9

Sigma Points of \mathbf{P}_{xx}

State covariance matrix



\mathbf{P}_{xx} : Symmetric, positive-definite covariance matrix

Eigenvalues are real and positive

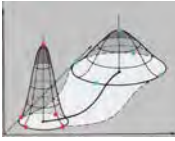
$$|s\mathbf{I}_n - \mathbf{P}_{xx}| = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_n)$$

Eigenvectors and the modal matrix

$$(\lambda_i \mathbf{I}_n - \mathbf{P}_{xx}) \alpha \mathbf{e}_i = 0, \quad i = 1, n$$

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}$$

10



Sigma Points of \mathbf{P}_{xx}

Diagonalized covariance matrix

Eigenvalues are the Variances

$$\mathbf{\Lambda} = \mathbf{E}^{-1} \mathbf{P}_{xx} \mathbf{E} = \mathbf{E}^T \mathbf{P}_{xx} \mathbf{E} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}$$

1) Principal axes of the covariance matrix are defined by modal matrix, \mathbf{E}

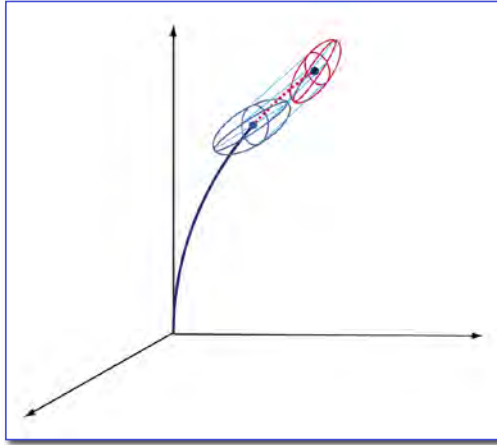
2) Location of $2n$ one-sigma points in state space given by

$$\begin{bmatrix} \pm \Delta \mathbf{x}(\sigma_1) & \pm \Delta \mathbf{x}(\sigma_2) & \dots & \pm \Delta \mathbf{x}(\sigma_n) \end{bmatrix} = \mathbf{E} \begin{bmatrix} \pm \sigma_1 & 0 & \dots & 0 \\ 0 & \pm \sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \pm \sigma_n \end{bmatrix}$$

11

*Propagation of the
Mean Value and
Covariance Matrix*

Propagation of the Mean Value and the Sigma Points



- Mean value at $t_k =$

$$\bar{\mathbf{x}}(t_k) = \bar{\mathbf{x}}_k$$

- Sigma points** (relative to mean value)

$$\sigma_{i_k} \triangleq \begin{cases} \bar{\mathbf{x}}_k - \Delta \mathbf{x}_k(\sigma_i), & i = 1, n \\ \bar{\mathbf{x}}_k + \Delta \mathbf{x}_k(\sigma_i), & i = (n+1), 2n \end{cases}$$

- Projection from the mean**

$$\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{x}}_k + \int_{t_k}^{t_{k+1}} \mathbf{f}[\bar{\mathbf{x}}(t), \mathbf{u}(t), \bar{\mathbf{w}}(t), t] dt$$

- Projection from each sigma point**

$$\sigma_{i_{k+1}} = \sigma_{i_k} + \int_{t_k}^{t_{k+1}} \mathbf{f}[\sigma_i(t), \mathbf{u}(t), \bar{\mathbf{w}}(t), t] dt, \quad i = 1, 2n$$

13

Estimation of the Propagated Mean Value

Assumptions:

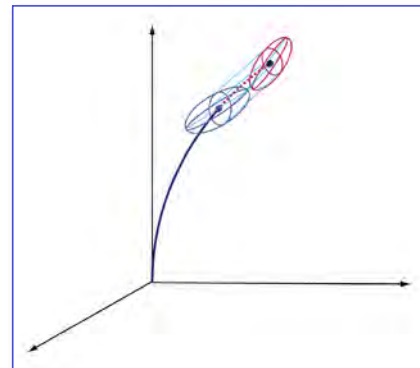
- To 2nd order, the propagated probability distribution is symmetric about its mean
- New mean is estimated as average or weighted average of projected points (arbitrary choice by user)

Ensemble Average for the Mean Value

$$\hat{\mathbf{x}}_{k+1} = \frac{\bar{\mathbf{x}}_{k+1} + \sum_{i=1}^{2n} \sigma_{i_{k+1}}}{2n+1}$$

Weighted Ensemble Average for the Mean Value

$$\hat{\mathbf{x}}_{k+1} = \frac{\bar{\mathbf{x}}_{k+1} + \xi \sum_{i=1}^{2n} \sigma_{i_{k+1}}}{2\xi n + 1}$$



14



Projected Covariance Matrix

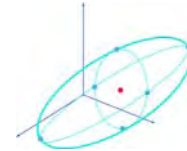
Unbiased ensemble estimate of the covariance matrix

$$\mathbf{P}_{\mathbf{x}_{k+1}|\mathbf{x}_{k+1}} = \frac{1}{(2n+1)-1} \left\{ (\bar{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_{k+1})(\bar{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_{k+1})^T + \sum_{i=1}^{2n} (\boldsymbol{\sigma}_{i_{k+1}} - \hat{\mathbf{x}}_{k+1})(\boldsymbol{\sigma}_{i_{k+1}} - \hat{\mathbf{x}}_{k+1})^T \right\}$$

This estimate neglects effects of disturbance uncertainty during the state propagation from t_k to t_{k+1}

15

Sigma Points of Disturbance Uncertainty, Q



\mathbf{Q} : ($s \times s$) Symmetric, positive-definite covariance matrix

$$|s\mathbf{I}_s - \mathbf{Q}| = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_s) \quad [s = \text{Laplace operator}]$$

$$(\lambda_i \mathbf{I}_s - \mathbf{Q})\alpha \mathbf{e}_i = 0, \quad i = 1, s$$

$$\mathbf{E}_Q = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_s \end{bmatrix}_Q$$

- Eigenvalues
- Modal Matrix
- Transformation

$$\Lambda_Q = \mathbf{E}_Q^T \mathbf{Q} \mathbf{E}_Q = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_s \end{bmatrix}_Q = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_s^2 \end{bmatrix}_Q$$

$$\begin{bmatrix} \pm \Delta \mathbf{w}(\sigma_1) & \pm \Delta \mathbf{w}(\sigma_2) & \cdots & \pm \Delta \mathbf{w}(\sigma_s) \end{bmatrix} = \mathbf{E}_Q \begin{bmatrix} \pm \sigma_1 & 0 & \cdots & 0 \\ 0 & \pm \sigma_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \pm \sigma_s \end{bmatrix}_Q$$

16

Propagation of the Disturbed Mean Value

Sigma points of disturbance (relative to mean value)

$$\boldsymbol{\omega}_{i_k} \triangleq \begin{cases} \bar{\mathbf{w}}_k + \Delta \mathbf{w}_k(\sigma_i), & i = 1, s \\ \bar{\mathbf{w}}_k - \Delta \mathbf{w}_k(\sigma_i), & i = (s+1), 2s \end{cases}$$

Incorporation of effects of disturbance uncertainty on state propagation

$$\left(\bar{\mathbf{x}}_{\boldsymbol{\omega}_i} \right)_{k+1} = \bar{\mathbf{x}}_k + \int_{t_k}^{t_{k+1}} \mathbf{f} \left[\bar{\mathbf{x}}(t), \mathbf{u}(t), \boldsymbol{\omega}_i(t), t \right] dt, \quad i = 1, 2s$$

17

Estimation of the Propagated Mean Value with Disturbance Uncertainty

Estimate now includes effect of disturbance uncertainty
Estimate of the mean is the average or weighted average of projected points

Ensemble Average for the Mean Value

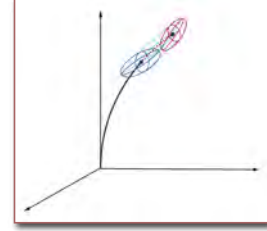
$$\hat{\mathbf{x}}_{k+1} = \frac{\bar{\mathbf{x}}_{k+1} + \sum_{i=1}^{2n} \sigma_{i_{k+1}} + \sum_{i=1}^{2s} \left(\bar{\mathbf{x}}_{\boldsymbol{\omega}_i} \right)_{k+1}}{2(n+s)+1}$$

Weighted Ensemble Average for the Mean Value

$$\hat{\mathbf{x}}_{k+1} = \frac{\bar{\mathbf{x}}_{k+1} + \xi \left[\sum_{i=1}^{2n} \sigma_{i_{k+1}} + \sum_{i=1}^{2s} \left(\bar{\mathbf{x}}_{\boldsymbol{\omega}_i} \right)_{k+1} \right]}{2\xi(n+s)+1}$$

18

Covariance Propagation with Disturbance Uncertainty



Unbiased sampled estimate of the covariance matrix

$$\mathbf{P}_{\mathbf{x}_{k+1}\mathbf{x}_{k+1}} = \frac{1}{[2(n+s)+1]-1} (\mathbf{P}_{mean} + \mathbf{P}_{sigma} + \mathbf{P}_{disturbance})$$

$$\mathbf{P}_{mean} = (\bar{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_{k+1})(\bar{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_{k+1})^T$$

$$\mathbf{P}_{sigma} = \sum_{i=1}^{2n} (\boldsymbol{\sigma}_{i_{k+1}} - \hat{\mathbf{x}}_{k+1})(\boldsymbol{\sigma}_{i_{k+1}} - \hat{\mathbf{x}}_{k+1})^T$$

$$\mathbf{P}_{disturbance} = \sum_{i=1}^{2s} \left[\left(\bar{\mathbf{x}}_{\omega_i} \right)_{k+1} - \hat{\mathbf{x}}_{k+1} \right] \left[\left(\bar{\mathbf{x}}_{\omega_i} \right)_{k+1} - \hat{\mathbf{x}}_{k+1} \right]^T$$

19

Sigma Points Filter

System Vector Notation

System vector

$$\mathbf{v} \triangleq \begin{bmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{n} \end{bmatrix}$$

$\dim(\mathbf{v}) = (n + r + s) \times 1$

Expected value of system vector

$$\hat{\mathbf{v}}_0 = \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{w}}_0 \\ \hat{\mathbf{n}}_0 \end{bmatrix} = E \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w}_0 \\ \mathbf{n}_0 \end{bmatrix} \triangleq \boldsymbol{\chi}_0 = \begin{bmatrix} \boldsymbol{\chi}_0^x \\ \boldsymbol{\chi}_0^w \\ \boldsymbol{\chi}_0^n \end{bmatrix}$$

Propagation of the mean

$$\boldsymbol{\chi}_{k+1}^x = \boldsymbol{\chi}_k^x + \int_{t_k}^{t_{k+1}} \mathbf{f}[\boldsymbol{\chi}^x(t), \mathbf{u}(t), \boldsymbol{\chi}^w(t), t] dt$$

Measurement vector, corrupted by noise

$$\boldsymbol{\psi} = \mathbf{h}(\boldsymbol{\chi}^x, \boldsymbol{\chi}^n)$$

$$\text{Analogous to } \mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)$$

21

Matrix Array of System and Sigma-Point Vectors

Expected value of system vector

$$\boldsymbol{\chi}_0 = \hat{\mathbf{v}}_0 = \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{w}}_0 \\ \hat{\mathbf{n}}_0 \end{bmatrix}; \quad \dim(\boldsymbol{\chi}_0) = (n + r + s) \times 1 \triangleq L \times 1$$

Weighted sigma points for system vector

$$\boldsymbol{\chi}_i = \begin{cases} \hat{\mathbf{v}}_i + \xi(\mathbf{S})_i, & i = 1, L \\ \hat{\mathbf{v}}_i - \xi(\mathbf{S})_i, & i = L + 1, 2L \end{cases}; \quad \dim(\boldsymbol{\chi}_i) = 2L \times 1$$

\mathbf{S} : Square root of \mathbf{P} ; $(\mathbf{S})_i \triangleq i^{\text{th}}$ column of \mathbf{S}

Matrix of mean and sigma-point vectors

$$\mathbf{X} \triangleq \begin{bmatrix} \boldsymbol{\chi}_0 & \boldsymbol{\chi}_1 & \cdots & \boldsymbol{\chi}_{2L} \end{bmatrix}; \quad \dim(\mathbf{X}) = L \times (2L + 1)$$

22

Initialize Filter

State and covariance estimates

$$\hat{\mathbf{x}}_o = E[\mathbf{x}(0)] = \boldsymbol{\chi}^x(0)$$

$$\mathbf{P}^x(0) = E\left\{[\mathbf{x}(0) - \hat{\mathbf{x}}(0)][\mathbf{x}(0) - \hat{\mathbf{x}}(0)]^T\right\}$$

Covariance matrix of system vector

$$\begin{aligned} \mathbf{P}^v(0) &= E\left\{[\mathbf{v}(0) - \hat{\mathbf{v}}(0)][\mathbf{v}(0) - \hat{\mathbf{v}}(0)]^T\right\} \\ &= \begin{bmatrix} \mathbf{P}^x(0) & 0 & 0 \\ 0 & \mathbf{Q}^w(0) & 0 \\ 0 & 0 & \mathbf{R}^n(0) \end{bmatrix} \end{aligned}$$

23

Propagate State Mean and Covariance

Incorporate disturbance sigma points

$$\left(\boldsymbol{\chi}_i^x\right)_{k+1} = \left(\boldsymbol{\chi}_i^x\right)_k + \int_{t_k}^{t_{k+1}} \mathbf{f}\left[\left(\boldsymbol{\chi}_i^x\right)(t), \mathbf{u}(t), \left(\boldsymbol{\chi}_i^w\right)(t), t\right] dt$$

Ensemble average estimates of mean and covariance

$$\hat{\mathbf{x}}_{k+1}(-) = \sum_{i=0}^{2L} \eta_i \left(\boldsymbol{\chi}_i^x\right)_{k+1}$$

$$\begin{aligned} \eta_i &: \text{Weighting factor} \\ &\text{Typically} \\ \eta_i &= \begin{cases} 1/(L+1), & i=0 \\ 1/2(L+1), & i=1, 2L \end{cases} \end{aligned}$$

$$\mathbf{P}_{k+1}^x(-) = \sum_{i=0}^{2L} \sum_{j=0}^{2L} \frac{\eta_i \eta_j}{1 - \eta_i \eta_j} \left(\boldsymbol{\chi}_i^x\right)_{k+1} \left(\boldsymbol{\chi}_j^x\right)_{k+1}^T$$

24

Incorporate Measurement Error in Output

Mean/sigma-point projections of measurement

$$(\psi_i)_{k+1} = \mathbf{h} \left[(\chi_i^x)_{k+1}, (\chi_i^n)_{k+1} \right], \quad i = 0, 2L$$

Weighted estimate of measurement projection

$$\hat{\mathbf{y}}_{k+1}(-) = \sum_{i=0}^{2L} \eta_i (\psi_i)_{k+1}$$

25

Incorporate Measurement Error in Covariance

Prior estimate of measurement covariance

$$\mathbf{P}_{k+1}^y(-) = \sum_{i=0}^{2L} \sum_{j=0}^{2L} \frac{\eta_i \eta_j}{1 - \eta_i \eta_j} (\psi_i)_{k+1} (\psi_j)_{k+1}^T$$

Prior estimate of state/measurement cross-covariance

$$\mathbf{P}_{k+1}^{xy}(-) = \sum_{i=0}^{2L} \sum_{j=0}^{2L} \frac{\eta_i \eta_j}{1 - \eta_i \eta_j} (\chi_i^x)_{k+1} (\psi_j)_{k+1}^T$$

26

Compute Kalman Filter Gain

Original formula (eq. 3, Lecture 18)

$$\mathbf{K}_k = \mathbf{P}_k(-) \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k(-) \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}$$

Sigma points version w/index change

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^{\text{xy}}(-) \left[\mathbf{P}_{k+1}^{\text{y}}(-) \right]^{-1}$$

27

Post-Measurement State and Covariance Estimate

State estimate update

$$\hat{\mathbf{x}}_{k+1}(+) = \hat{\mathbf{x}}_{k+1}(-) + \mathbf{K}_{k+1} \left[\mathbf{z}_{k+1} - \hat{\mathbf{y}}_{k+1}(-) \right]$$

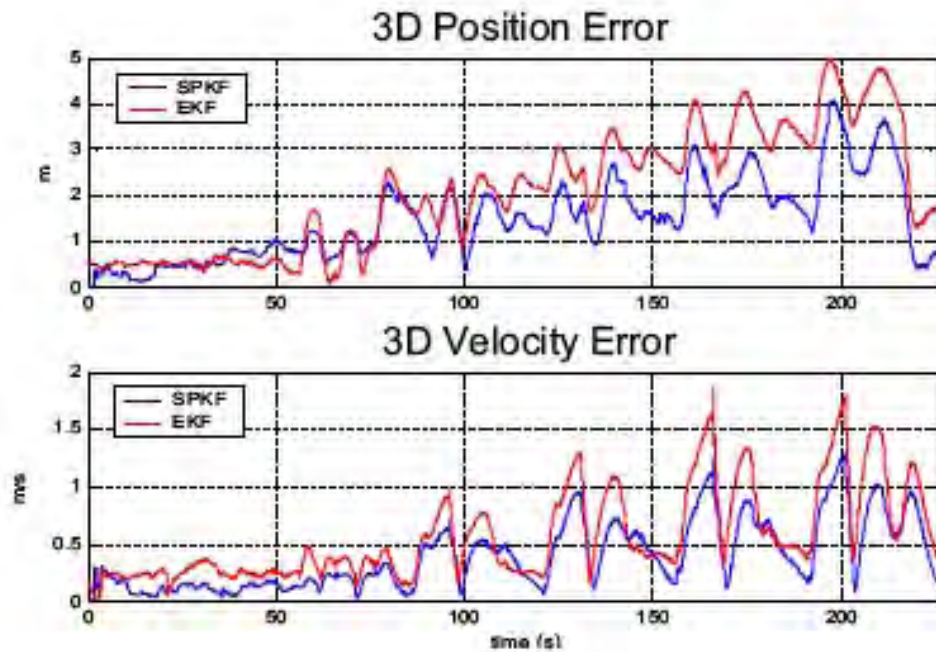
Covariance estimate “update”

$$\mathbf{P}_{k+1}^{\text{x}}(+) = \mathbf{P}_{k+1}^{\text{x}}(-) - \mathbf{K}_{k+1} \mathbf{P}_{k+1}^{\text{y}}(-) \mathbf{K}_{k+1}^T$$

28

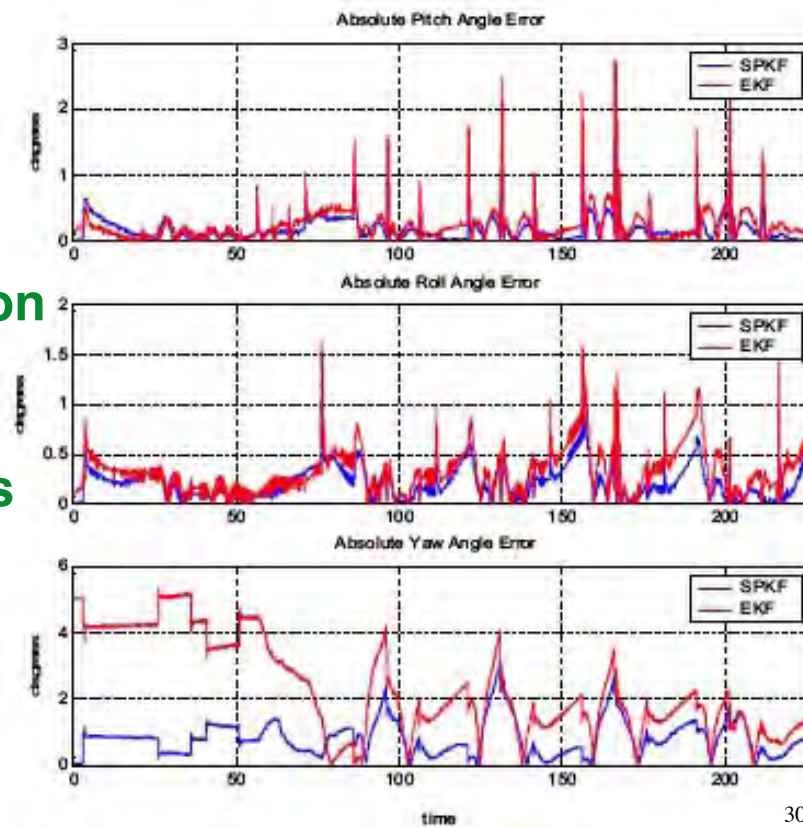
Example: Simulated Helicopter UAV Flight

van der Werwe and Wan, 2004



29

Comparison of Pitch, Roll, and Yaw Errors



30

Observations

- No Jacobians calculated
- Large number of propagation steps
- Approximation for Gaussian distributions
- Estimate is equivalent to that from a second-order EKF filter
- Best choice of averaging weights is problem-dependent
- Alternative sigma-point filter formulations
- Is it better than a quasi-linear estimate?
(TBD)

31

Advanced Nonlinear Estimators

- Discussion of
 - Psiaki, M. L., “The Blind Tricyclist Problem and a Comparison of Nonlinear Filters,” *IEEE Control Systems Magazine*, June, 2013, pp. 48-54
 - Psiaki, M. L., Schoenberg, J. R., and Miller, I. T., “Gaussian Sum Reapproximation for Use in a Nonlinear Filter,” *Journal of Guidance, Control, and Dynamics*, 38 (2), Feb 2015, pp. 292-303
 - Psiaki, M. L., “The ‘Blob’ Filter: Gaussian Mixture Nonlinear Filtering with Re-Sampling for Mixand Narrowing,” *IEEE/ION PLANS 2014*, May 2014, pp. 1-14

32

Next Time:
Adaptive State Estimation