Nonlinear State Estimation

Particle, Sigma-Points Filters

Robert Stengel

Optimal Control and Estimation, MAE 546
Princeton University, 2017

- Particle filter
- Sigma-Points (“Unscented Kalman”) filter
  - Transformation of uncertainty
  - Propagation of mean and variance
- Helicopter, HIV state estimation examples
- Additional nonlinear filters

Criticisms of the Basic Extended Kalman Filter*

*Julier and Uhlmann, 1997; Wan and van der Merwe, 2001

- State estimate prediction is deterministic, i.e., not based on an expectation
  - (Not true; expectation is the same as the deterministic prediction)
- 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
Transformation of Uncertainty

Nonlinear transformation of a random variable

\[ x : \text{Random variable with mean, } \bar{x}, \text{ and covariance, } P_{xx} \]

\[ y = f[x] \]

Estimate the mean and covariance of the transformation’s output

\[ \bar{y}(\bar{x}, P_{xx}) \text{ and } P_{yy}(\bar{x}, P_{xx}) \]

The transformation is said to be “unscented”* if its probability distribution is

Consistent, Efficient, and Unbiased

---

Consistent Estimate of a Dynamic State

Let \( x_k \triangleq \bar{x}, \ x_{k+1} \triangleq y \)

\[ \bar{x}_{k+1}(\bar{x}_k, P_{x_k x_k}) = \bar{y}(\bar{x}, P_{xx}) \text{ and } P_{x_{k+1} x_{k+1}}(\bar{x}_k, P_{x_k x_k}) = P_{yy}(\bar{x}, P_{xx}) \]

Consistent state estimate converges in the limit

\[ \{P_{x_{k+1} x_{k+1}} - E[(x_{k+1} - \bar{x}_{k+1})(x_{k+1} - \bar{x}_{k+1})^T]\} \geq 0 \]

\[ \{\text{Estimated Covariance} - \text{Actual Covariance}\} \geq 0 \]

**Lesson**: In filtering, add sufficient “process noise” to the filter gain computation to prevent filter divergence
Adding Process Noise Improves Consistency

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

\[ \text{Efficient and Unbiased Estimate of a Dynamic State} \]

Efficient state estimator converges more quickly than an inefficient estimator

\[ \min_{\text{Added Process Noise}} \{ \text{Estimated Covariance} - \text{Actual Covariance} \} \]

Add “just enough” process noise

Unbiased estimate

\[ \bar{x}_{k+1} = E(x_{k+1}) \]  \[ \text{[Estimated Mean = Actual Mean]} \]
Empirical Determination of a Probability Distribution

- Monte Carlo evaluation
  - Propagate random noise through nonlinearity for given prior distribution (e.g., Gaussian)
  - Generate histogram of output
    - Repeated evaluation (N trials) of nonlinear system propagation
    - Use histogram as numerical representation of distribution, or
    - Identify associated theoretical distribution using functional minimization

- Particle Filter
  - Propagate mean, $E(x)$, using the empirical probability distribution
  - As $N \to \infty$, estimate error $\to 0$

Empirical Determination of Mean and Variance

- Sample mean for N data points, $x_1, x_2, \ldots, x_N$
  \[
  \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i
  \]

- Sample variance for same data set
  \[
  \sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2
  \]
  - Divisor is $(N - 1)$ rather than $N$ to produce an unbiased estimate

- Sigma Points Filter:
  - Propagate mean, $E(x)$, using the sample mean and variance
Sigma Points of $\mathbf{P}_{xx}$ and Mean

2-D Example

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

3-D Example
**Sigma Points of Estimate Uncertainty, \( P_{xx} \)**

State covariance matrix

\[ P_{xx} : \text{Symmetric, positive-definite covariance matrix} \]

**Eigenvalues are real and positive**

\[ (sI_n - P_{xx}) = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_n) \]

**Eigenvectors**

\[ (\lambda_i I_n - P_{xx}) \alpha e_i = 0, \quad i = 1, n \]

**Modal matrix**

\[ E = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix} \]

---

**Sigma Points of \( P_{xx} \)**

Diagonalized covariance matrix

**Eigenvalues are the Variances**

\[ \Lambda = E^T P_{xx} E = E^T P_{xx} E = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \]

1) Principal axes of the covariance matrix are defined by modal matrix, \( E \)

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

\[ \begin{bmatrix} \pm \Delta x(\sigma_1) & \pm \Delta x(\sigma_2) & \cdots & \pm \Delta x(\sigma_n) \end{bmatrix} = E \begin{bmatrix} \pm \sigma_1 & 0 & \cdots & 0 \\ 0 & \pm \sigma_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \pm \sigma_n \end{bmatrix} \]
Propogation of the Mean Value and Covariance Matrix

Propagation of the Mean Value and the Sigma Points

- Mean value at $t_k$
  $$\bar{x}(t_k) = \bar{x}_k$$

- Sigma points (relative to mean value)
  $$\sigma_{i\ell} \triangleq \begin{cases} 
  \bar{x}_k - \Delta x_k(\sigma_i), & i = 1, n \\
  \bar{x}_k + \Delta x_k(\sigma_i), & i = (n+1), 2n
  \end{cases}$$

- Projection from the prior mean
  $$\bar{x}_{k+1} = \bar{x}_k + \int_{t_k}^{t_{k+1}} f[\bar{x}(t), u(t), \bar{w}(t), t] dt$$

Nonlinear projection from each prior sigma point
  $$\sigma_{i\ell+1} = \sigma_{i\ell} + \int_{t_k}^{t_{k+1}} f[\sigma_i(t), u(t), \bar{w}(t), t] dt, \quad i = 1, 2n$$
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

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

Weighted Ensemble Average for the Mean Value

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

Projected Covariance Matrix

Unbiased ensemble estimate of the covariance matrix

\[ P_{x_{k+1}x_{k+1}} = \frac{1}{(2n+1)-1} \left\{ (\bar{x}_{k+1} - \hat{x}_{k+1})(\bar{x}_{k+1} - \hat{x}_{k+1})^T + \sum_{i=1}^{2n} (\sigma_{i_{k+1}} - \hat{x}_{k+1})(\sigma_{i_{k+1}} - \hat{x}_{k+1})^T \right\} \]

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

Does not require calculation of Jacobian matrices
Sigma Points of Disturbance Uncertainty, $Q$

$Q: \quad (s \times s)$ Symmetric, positive-definite covariance matrix

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

$(\lambda_i I_s - Q)\alpha e_i = 0, \quad i = 1, s$

$E_Q = \begin{bmatrix} e_1 & e_2 & \cdots & e_s \end{bmatrix}

\begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\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 \\
0 & 0 & \cdots & \sigma_s^2
\end{bmatrix}_Q$

$A_Q = E_Q^T Q E_Q =

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

\begin{bmatrix}
\pm \sigma_1 & 0 & \cdots & 0 \\
0 & \pm \sigma_2 & \cdots & 0 \\
\cdots & \cdots & \cdots \\
0 & 0 & \cdots & \pm \sigma_s
\end{bmatrix}_Q

\text{Propagation of the Disturbed Mean Value}

Sigma points of disturbance (relative to mean value)

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

Incorporation of effects of disturbance uncertainty on state propagation

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

\[ \hat{x}_{k+1} = \frac{\bar{x}_{k+1} + \sum_{i=1}^{2n} \sigma_{i_{k+1}} + \sum_{i=1}^{2s} (\bar{x}_{\omega_i})_{k+1}}{2(n+s)+1} \]

Weighted Ensemble Average for the Mean Value

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

Covariance Propagation with Disturbance Uncertainty

Unbiased sampled estimate of the covariance matrix

\[
P_{x_{k+1}x_{k+1}} = \frac{1}{[2(n+s)+1]-1} \left( P_{\text{mean}} + P_{\text{sigma}} + P_{\text{disturbance}} \right)
\]

\[
P_{\text{mean}} = (\bar{x}_{k+1} - \hat{x}_{k+1})(\bar{x}_{k+1} - \hat{x}_{k+1})^T
\]

\[
P_{\text{sigma}} = \sum_{i=1}^{2n} (\sigma_{i_{k+1}} - \hat{x}_{k+1})(\sigma_{i_{k+1}} - \hat{x}_{k+1})^T
\]

\[
P_{\text{disturbance}} = \sum_{i=1}^{2s} \left[ (\bar{x}_{\omega_i})_{k+1} - \hat{x}_{k+1} \right] \left[ (\bar{x}_{\omega_i})_{k+1} - \hat{x}_{k+1} \right]^T
\]
Sigma-Points
("Unscented Kalman") Filter

System Vector Notation

System vector
\[ \mathbf{v} = \begin{bmatrix} \mathbf{x} \\ \mathbf{w} \\ \mathbf{n} \end{bmatrix} \]
\[ \dim(\mathbf{v}) = (n + r + s) \times 1 \]

Expected value of system vector
\[ \hat{\mathbf{v}}_0 = \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{w}}_0 \\ \hat{\mathbf{n}}_0 \end{bmatrix} = E \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w}_0 \\ \mathbf{n}_0 \end{bmatrix} \overset{\Delta}{=} \mathbf{\chi}_0 = \begin{bmatrix} \mathbf{\chi}_0^x \\ \mathbf{\chi}_0^w \\ \mathbf{\chi}_0^n \end{bmatrix} \]

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

Measurement vector, corrupted by noise
\[ \mathbf{\psi} = \mathbf{h}(\mathbf{\chi}_k^x, \mathbf{\chi}_k^n) \]
\[ \text{Analogous to } \mathbf{z}(t) = \mathbf{Hx}(t) + \mathbf{n}(t) \]
Matrix Array of System and Sigma-Point Vectors

Expected value of system vector

\[ \chi_0 = \hat{\chi}_0 = \begin{bmatrix} \hat{x}_0 \\ \hat{\omega}_0 \\ \hat{n}_0 \end{bmatrix} ; \quad \text{dim}(\chi_0) = (n + r + s) \times 1 = L \times 1 \]

Weighted sigma points for system vector

\[ \chi_i = \begin{cases} \hat{\chi}_i + \xi(S), & i = 1, L \\ \hat{\chi}_i - \xi(S), & i = L + 1, 2L \end{cases} ; \quad \text{dim}(\chi_i) = 2L \times 1 \]

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

Matrix of mean and sigma-point vectors

\[ X \triangleq \begin{bmatrix} \chi_0 & \chi_1 & \cdots & \chi_{2L} \end{bmatrix} ; \quad \text{dim}(X) = L \times (2L + 1) \]

Initialize Filter

State and covariance estimates

\[ \hat{x}_0 = E[x(0)] = \chi^x(0) \]

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

Covariance matrix of system vector

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

\[ = \begin{bmatrix} P^x(0) & 0 & 0 \\ 0 & Q^w(0) & 0 \\ 0 & 0 & R^n(0) \end{bmatrix} \]
Propagate State Mean and Covariance

Incorporate disturbance sigma points

\[
\chi^x_{i+1} = \chi^x_i + \int_{t_k}^{t_{k+1}} f(\chi^x_i(t), u(t), \chi^w_i(t), t) \, dt
\]

Ensemble average estimates of mean and covariance

\[
\hat{x}_{k+1}(-) = \sum_{i=0}^{2L} \eta_i \chi^x_i|_{k+1}
\]

\[
\hat{P}^x_{k+1}(-) = \sum_{i=0}^{2L} \eta_i \left\{ \left[ \chi^x_i(-) - \hat{x}(-) \right]_{k+1} \left[ \chi^x_i(-) - \hat{x}(-) \right]^T \right\}
\]

Incorporate Measurement Error in Output

Mean/sigma-point projections of measurement

\[
\psi_{i+1} = h(\chi^x_{k+1}, \chi^n_{k+1}) , \quad i = 0, 2L
\]

Weighted estimate of measurement projection

\[
\hat{y}_{k+1}(-) = \sum_{i=0}^{2L} \eta_i \psi_i|_{k+1}
\]
Incorporate Measurement Error in Covariance

Prior estimate of measurement covariance

\[
P_{y,k+1}^{x}(-) = \sum_{i=0}^{2L} \eta_i \left\{ \left[ \psi_i (-) - \hat{y} (-) \right]_{k+1} \left[ \psi_i (-) - \hat{y} (-) \right]_{k+1}^T \right\}
\]

Prior estimate of state/measurement cross-covariance

\[
P_{xy,k+1}^{x}(-) = \sum_{i=0}^{2L} \eta_i \left\{ \left[ \chi_i^{x} (-) - \hat{x} (-) \right]_{k+1} \left[ \psi_i (-) - \hat{y} (-) \right]_{k+1}^T \right\}
\]

Compute Kalman Filter Gain

Original formula (eq. 3, Lecture 18)

\[
K_k = P_{k}^{x}(-)H_k^T \left[ H_k P_{k}^{x}(-) H_k^T + R_k \right]^{-1}
\]

Sigma points version w/index change

\[
K_{k+1}^{xy} = P_{k+1}^{xy}(-) \left[ P_{k+1}^{y}(-) \right]^{-1}
\]
Post-Measurement State and Covariance Estimate

State estimate update

\[
\hat{x}_{k+1}(+) = \hat{x}_{k+1}(-) + K_{k+1} \left[ z_{k+1} - \hat{y}_{k+1}(-) \right]
\]

Covariance estimate “update”

\[
P_{k+1}^x(+) = P_{k+1}^x(-) - K_{k+1} P_{k+1}^y(-) K_{k+1}^T
\]

Example: Simulated Helicopter UAV Flight

*van der Werwe and Wan, 2004*
Comparison of Pitch, Roll, and Yaw Errors

Comparison of RMS Errors for EKF, UKF, and Gauss-Hermite Filter (GHF)

Banks et al, A COMPARISON OF NONLINEAR FILTERING APPROACHES IN THE CONTEXT OF AN HIV MODEL, 2010
Comparison of Various Filters for “Blind Tricyclist Problem” (Psiaki, 2013)

- Extended Kalman filter
- Unscented Kalman filter
- Particle filter
- Batch least-squares filter
- Backward-smoothing EKF

Observations

- Jacobians vs. no Jacobians
- Number of nonlinear propagation steps
- Gaussian vs. approximate non-Gaussian distributions
- Best choice of averaging weights is problem-dependent
- Comparison of filters is problem-dependent

- Are these filters better than a quasi-linear filter? (TBD)
More on Nonlinear Estimators


Next Time:
Adaptive State Estimation