Q1. Discuss the evaluation of covariance matrices in relation to Kalman and extended Kalman filtering

Problem Setup

- The KF assumes the following dynamic state-space model:

$$\begin{aligned} \text{Process equation:} \quad x_k &= F_k x_{k-1} + v_{k-1} \\ \text{Measurement equation:} \quad z_k &= H_k x_k + w_k \end{aligned}$$

- Assumptions:
  - Additive uncorrelated Gaussian noise sequences with known statistics. i.e., $v_k \sim \mathcal{N}(0, Q_{k-1})$ and $w_k \sim \mathcal{N}(0, R_k)$
  - Known initial estimate. i.e., $x_0 \sim \mathcal{N}(\hat{x}_{0|0}, P_{0|0})$
- Objective: Estimate the state at time k recursively given $\{z_1, z_2, \ldots z_k\}$.
- Performance criteria: Minimum mean-square error.

Two Basic Operations

- Predict:

$$
\begin{aligned}
\hat{x}_{k|k-1} &= F_k \hat{x}_{k-1|k-1} \\
P_{k|k-1} &= F_k P_{k-1|k-1} F_k^T + Q_{k-1}
\end{aligned}
$$

- Correct:

$$
\begin{aligned}
\hat{z}_{k|k-1} &= H_k \hat{x}_{k|k-1} \\
S_{k|k-1} &= H_k P_{k|k-1} H_k^T + R_k \\
W_k &= P_{k|k-1} H_k^T S_{k|k-1}^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + W_k (z_k - \hat{z}_{k|k-1}) \\
P_{k|k} &= P_{k|k-1} - W_k H_k P_{k|k-1}
\end{aligned}
$$

Error Covariances: Analysis

- As a self-assessment of its own errors, the KF yields the error covariance matrices:
  - state $\Rightarrow$predicted and posterior error covariances
  - measurement $\Rightarrow$ innovation covariance
- Why we evaluate covariances?
  - To verify the credibility of the filter: if the actual error is consistent with the filter-computed error?
  - To compare various filter performances
  - To probe into modeling errors

Error Covariances (Cont'd)

- Tools for evaluation:
  - Mean-Squared Error (MSE)
  - Posterior Cramer-Rao Lower Bound (PCRLB)
  - Normalized Innovation-Squared (NIS)
  - Normalized Estimation Error-squared(NEES)
- The PCRLB and the NIS can be used in real-time applications.

Mean-Squared Error

- Given the true state x_k, the MSE (matrix) of the filter estimate is defined by

$$\mathrm{MSE}(k) \quad = \quad \mathbb{E}((x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T)$$

- The gain-posterior covariance relationship: $W_k = P_{k|k} H^T R^{-1}$.
- 3 practical cases:
  - $P_{k|k} = \mathrm{MSE}(k) \Rightarrow$ optimal
  - $P_{k|k} > \mathrm{MSE}(k) \Rightarrow$ pessimistic
  - $P_{k|k} < \mathrm{MSE}(k) \Rightarrow$ optimistic
- MSE = variance + bias-squared (in a scalar case).

PCRLB

- The covariance matrix $P_{k|k}$ of an unbiased state estimator $\widehat{\mathbf{x}}_{k|k}$ has a lower bound

$$P_{k|k} \quad \succeq \quad J_k^{-1}$$

where the Fisher information matrix

$$J_k = D_{k-1}^{22} - D_{k-1}^{21} \left(J_{k-1} + D_{k-1}^{11} \right)^{-1} D_{k-1}^{12} \qquad (k > 0)$$

where

$$\begin{aligned}
D_{k-1}^{11} &= -E\left\{ \nabla_{\mathbf{x}_{k-1}} \left[\nabla_{\mathbf{x}_{k-1}} \ln p(\mathbf{x}_k|\mathbf{x}_{k-1}) \right]^T \right\} \\
D_{k-1}^{21} &= -E\left\{ \nabla_{\mathbf{x}_{k-1}} \left[\nabla_{\mathbf{x}_k} \ln p(\mathbf{x}_k|\mathbf{x}_{k-1}) \right]^T \right\} \\
D_{k-1}^{12} &= -E\left\{ \nabla_{\mathbf{x}_k} \left[\nabla_{\mathbf{x}_{k-1}} \ln p(\mathbf{x}_k|\mathbf{x}_{k-1}) \right]^T \right\} = \left[D_{k-1}^{21} \right]^T \\
D_{k-1}^{22} &= -E\left\{ \nabla_{\mathbf{x}_k} \left[\nabla_{\mathbf{x}_k} \ln p(\mathbf{x}_k|\mathbf{x}_{k-1}) \right]^T \right\} - E\left\{ \nabla_{\mathbf{x}_k} \left[\nabla_{\mathbf{x}_k} \ln p(\mathbf{z}_k|\mathbf{x}_k) \right]^T \right\}
\end{aligned}$$

PCRLB for the KF

- For the LG case, information matrices

$$\begin{aligned} D_{k-1}^{11} &= F_{k-1}^T Q_{k-1}^{-1} F_{k-1} \\ D_{k-1}^{12} &= [D_{k-1}^{21}]^T = -F_{k-1}^T Q_{k-1}^{-1} \\ D_{k-1}^{22} &= Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k \end{aligned}$$

- Hence we get the Fisher information matrix at time k as

$$\begin{aligned} J_k &= Q_{k-1}^{-1} - Q_{k-1}^{-1} F_{k-1} \left(J_{k-1} + F_{k-1}^T Q_{k-1}^{-1} F_{k-1} \right)^{-1} F_{k-1}^T Q_{k-1}^{-1} + J_k^z \\ &= \left(Q_{k-1} + F_{k-1} J_{k-1}^{-1} F_{k-1}^T \right)^{-1} + H_k^T R_k^{-1} H_k \end{aligned}$$

- Obtain the same when replacing J_k with $P_{k|k}^{-1} \Rightarrow$ KF is an efficient estimator for a linear-Gaussian system.

Credibility Check on Innovation Covariance

- Properties of innovation sequences
  - zero-mean Gaussian
  - uncorrelated.
- Define normalized innovation-squared (NIS):

$$\epsilon(k) = \nu_k^T S_{k|k}^{-1} \nu_k$$

where the innovation and its covariance

$$\nu_k = (z_k - \hat{z}_{k|k-1})$$
$$S_{k|k-1} = \text{cov}(\nu_k)$$

- Distribution of $\epsilon(k)$:

$$\epsilon(k) \sim \chi_m^2$$

where m is the measurement-vector dimension or the dof.

Credibility Check (Cont'd)

- Postulate the null hypothesis

$$H_0: \quad \mathbb{E}[\epsilon(k)] \quad = \quad m.$$

- Accept both the innovation and its covariance commensurate with theoretical results if

$$\epsilon(k) \quad \in \quad [r_1, \quad r_2]$$

- The acceptance interval $[r_1 \quad r_2]$ is determined such that

$$P(\epsilon(k) \in [r_1 \quad r_2] | H_0) \quad = \quad 1 - \alpha$$

  where α is the level of significance.

- For $\alpha = 0.05$, the limits

$$r_{1,2} \quad \approx \quad \frac{1}{2m}\big(\pm 1.96 + \sqrt{2m-1}\big)^2$$

Remarks

- For an unbiased estimator, the NIS check is directly comparable to a credibility check on the innovation covariance.
- In an N Monte Carlo runs, we use the averaged NIS statistics.

Credibility Check on Posterior Error Covariance

- Define normalized estimation error-squared (NEES):

$$\epsilon(k) = \tilde{x}_k^T P_{k|k}^{-1} \tilde{x}_k$$

where the estimation error

$$\tilde{x}_k = (x_k - \hat{x}_{k|k})$$

- Distribution of $\epsilon(k)$:

$$\epsilon(k) \sim \chi_n^2$$

where n is the state-vector dimension.

- Following a similar procedure as in the NIS case, we may check the credibility of filter-estimated covariance at an α level.

Extended Kalman Filters

- The EKF assumes the following dynamic state-space model:

$$\text{Process equation:} \quad x_k = f(x_{k-1}) + v_{k-1} \tag{1}$$

$$\text{Measurement equation:} \quad z_k = h(x_k) + w_k \tag{2}$$

- Idea: Linearize nonlinear functions using the first-order Taylor series expansion.

- Let $x \sim \mathcal{N}(\bar{x}, \Sigma_x)$. Then we write y, where $y = f(x)$ as

$$y = f(x) \approx f(\bar{x}) + \underbrace{F(x - \bar{x})}_{\text{linear}},$$

where the Jacobian

$$F = [\nabla_x f(x)^T]^T_{x=\bar{x}}.$$

EKF (Cont'd)

- Approximates y to be Gaussian with the following mean and covariance:

$$\begin{aligned} \bar{y} &= \mathbb{E}(f(x)) \\ &\approx f(\bar{x}) \\ \Sigma_y &\approx F\Sigma_x F^T. \end{aligned}$$

Final Remarks

- For a nonlinear system, all conditional densities are non-Gaussian. However all the above methods assume Gaussianity suggesting that they wont fully characterize a nonlinear filter accuracy
- EKF estimate is biased $\Rightarrow P_{k|k} <$ MSE
- To meet the condition of zero-mean error (unbiased mean error), we subtract off mean errors before applying the NIS or NEES.
- In the EKF case, the information matrices of the CRLB are given by:

$$\begin{aligned} D_{k-1}^{11} &= \mathbb{E}(F_{k-1}^T Q_{k-1}^{-1} F_{k-1}) \\ D_{k-1}^{12} &= [D_{k-1}^{21}]^T = -\mathbb{E}(F_{k-1}^T) Q_{k-1}^{-1} \\ D_{k-1}^{22} &= Q_{k-1}^{-1} + \mathbb{E}(H_k^T R_k^{-1} H_k) \end{aligned}$$

  where F and H are Jacobians of the state and measurement functions.

- The expectation operators are replaced by Monte Carlo averages

(useful in simulations!)

References

- Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with applications to target tracking* , Wiley, 2001.
- D. Simon *Optimal state estimation*, Wiley 2006.

Thank you!