Outline

- Fundamental Theorem of Estimation Theory.
- Bayes formula.
- Derivation of minimum mean-square estimate.

Fundamental Thm. of Estimation Theory

- Minimum mean-square error estimator:
  \[ \hat{x}_k = E\{x_k | z_k^*\} \]
  \[ z_k^* = \text{col}\{z_i, i = 0, 1, \ldots, k\}. \]
- \( z_k^* \) = all measurements up to and including time \( k \)
- State estimate depends on measurements.
- Predictor
  \[ x_{k+1} = \phi_k x_k + w_k \]
  \[ \hat{x}_{k+1} = E\{x_{k+1} | z_k^*\} = \phi_k E\{x_k | z_k^*\} + 0 = \phi_k \hat{x}_k^+ \]

Proof

\[ MSE = E\{[x_k - \hat{x}_k]^T [x_k - \hat{x}_k] | z_k^*\} \]
- Expand and complete squares.
\[ MSE = E\{x_k^T x_k | z_k^*\} + E\{x_k | z_k^*\} - E\{\hat{x}(k) - E\{x_k | z_k^*\}\} \]
\[ -E\{x_k^T z_k^*\} E\{x_k | z_k^*\} \]
- Choose \( \hat{x}_k \) to minimize mean-square error.
- Minimized if \( \hat{x}_k = E\{x_k | z_k^*\} \)
Bayes Formula

\[ f_{x_k|x_k} = \frac{f_{z_k|x_k}f_{x_k}}{f_{z_k}} \]

- All three densities are normal (noise normal).
- Evaluate each density.
- Substitute and simplify to obtain the optimal estimate.

Density of State Vector \( f_{x(k)} \)

- Assume that we already have an **unbiased** optimal prior estimate (based on measurements up to \( k - 1 \))
  \[ E\{x(k)\} = \hat{x}^-(k) \]
- Error Covariance Matrix
  \[ P^-(k) = E\{[x(k) - \hat{x}^-(k)][x(k) - \hat{x}^-(k)]^T\} \]
- Normally distributed state vector
  \[ x(k) \sim N(\hat{x}^-(k), P^-(k)) \]

Density of Measurements \( f_{z(k)} \)

\[ z(k) = H(k)x(k) + v(k) \]
\[ E\{z(k)\} = H(k)\hat{x}^-(k) \]
\[ \tilde{z}(k) = z(k) - H(k)\hat{x}^-(k) \]
\[ = H(k)[x(k) - \hat{x}^-(k)] + v(k) \]

- Covariance
  \[ E\{\tilde{z}(k)\tilde{z}^T(k)\} = H(k)P^-(k)H^T(k) + R(k) \]
\[ z(k) \sim N(H(k)\hat{x}^-(k), H(k)P^-(k)H^T(k) + R(k)) \]

Conditional Density of Measurements \( f_{z(k)|x(k)} \)

\[ z(k) = H(k)x(k) + v(k) \]
\[ E\{z(k)|x(k)\} = H(k)x(k) \]
\[ \tilde{z}(k) = z(k) - H(k)x(k) = v(k) \]

- Covariance
  \[ E\{\tilde{z}(k)\tilde{z}^T(k)|x(k)\} = R(k) \]
- Normally distributed
  \[ z(k)|x(k) \sim N(H(k)x(k), R(k)) \]
Apply Bayes Formula

\[
f_{x(k)\mid z(k)} = \frac{f_{z(k)\mid x(k)} f_x(k)}{f_z(k)}
\]
\[
x(k) \sim \mathcal{N}(\hat{x}^-(k), P^-(k))
\]
\[
z(k) \sim \mathcal{N}(H(k)\hat{x}^-(k), H(k)P^-(k)H^T(k) + R(k))
\]
\[
z(k)\mid x(k) \sim \mathcal{N}(H(k)x(k), R(k))
\]
\[
f_x = \frac{\exp\left\{-\frac{1}{2} [x - m_x]^T C_{xx}^{-1} [x - m_x]\right\}}{\sqrt{(2\pi)^n \det(C_{xx})}}
\]

Results

- Quotient \( f_{x(k)\mid z(k)} \) is in the form of a multivariate normal density.
- Use fundamental theorem of estimation theory.

\[
x(k)\mid z(k) \sim \mathcal{N}\left(m_{x\mid z}, P_{x\mid z}(k)\right)
\]
\[
m_{x\mid z} = \hat{x}^-(k) + K(k)[z(k) - H(k)\hat{x}^-(k)] = \hat{x}^+(k)
\]
\[
K(k) = P^-(k)H^T(k)\left[H(k)P^-(k)H^T(k) + R(k)\right]^{-1}
\]
\[
P_{x\mid z}(k) = \left[(P^-(k))^{-1} + H^T(k)R^{-1}(k)H(k)\right]^{-1} = P^+(k)
\]
- As in the information filter.

Bayesian Estimation

\[
j = E\{C(x - \hat{x})\} = \int_{-\infty}^{\infty} C(\bar{x}) f_{x\mid z}(x\mid z) dx
\]
\[C(\bar{x}) = \text{cost function} \text{ (e.g. square error)}\]

Fact: For a symmetric convex cost and a conditional density \( f_{x\mid z} \) that is symmetric about the mean, all Bayesian estimates are the same

(Sorenson, 1970)

Conclusions

- Obtained linear estimate for Gaussian case without prior assumption.
- For non-Gaussian case the Kalman filter is the best linear filter but there may be better nonlinear filters.
- If knowledge of the process is incomplete (only up to 2\(^{nd}\) order statistics), the Kalman filter is the best possible (based on best Gaussian approximation of the true process).