Kalman Filter and Extended Kalman Filter

Namrata Vaswani, namrata@iastate.edu

Kalman Filter Introduction

- Recursive LS (RLS) was for static data: estimate the signal x better and better as more and more data comes in, e.g. estimating the mean intensity of an object from a video sequence
- RLS with forgetting factor assumes slowly time varying x
- Kalman filter: if the signal is time varying, and we know (statistically) the dynamical model followed by the signal: e.g. tracking a moving object

$$\begin{aligned} x_0 &\sim \mathcal{N}(0, \Pi_0) \\ x_i &= F_i x_{i-1} + v_{x,i}, \;\; v_{x,i} \sim \mathcal{N}(0, Q_i) \end{aligned} \tag{1}$$

The observation model is as before:

$$y_i = H_i x_i + v_i, \;\; v_i \sim \mathcal{N}(0, R_i) \tag{2}$$

The signal and observation noises are assumed uncorrelated (with each other and over time). They are also uncorrelated with the initial state x_0.

- Denote $Y_i \triangleq \{y_1, y_2, \ldots y_i\}$.
- **Goal:** get the best (minimum mean square error) estimate of x_i from $Y_i \triangleq \{y_1, y_2, ...y_i\}$ where Mean square error is given by $J(\hat{x}_i) = E[(x_i - \hat{x}_i)^2 | Y_i]$
- Minimizer is the conditional mean $\hat{x}_i = E[x_i | Y_i]$
- Note: This is also the MAP estimate, i.e. $\hat{x}_i$ also maximizes $p(x_i | Y_i)$ ($p(x_i | Y_i)$ is Gaussian (will be shown) and for Gaussian pdfs, mean=MAP).

Kalman filter

At $i = 0$, $\hat{x}_0 = E[x_0] = 0$, $P_0 = \Pi_0$ and $x_0 \sim \mathcal{N}(\hat{x}_0, \Pi_0)$ (given).

For any i, assume that we know $\hat{x}_{i-1} = E[x_{i-1}|Y_{i-1}]$ and $P_{i-1} = Var(x_{i-1}|Y_{i-1})$ and $x_{i-1}|Y_{i-1} \sim \mathcal{N}(\hat{x}_{i-1}, P_{i-1})$. Then using (1), $x_i|Y_{i-1} = F_i x_{i-1}|Y_{i-1} + v_i$ is also Gaussian with

$$\begin{aligned} E[x_i|Y_{i-1}] &= F_i \hat{x}_{i-1} \triangleq \hat{x}_{i|i-1} \\ Var(x_i|Y_{i-1}) &= F_i P_{i-1} F_i^T + Q_i \triangleq P_{i|i-1} \end{aligned} \tag{3}$$

This is the **prediction step**

Filtering or correction step: Let $Z_1 \triangleq x_i|Y_{i-1}$ & let $Z_2 \triangleq y_i|Y_{i-1}$. Then Z_1 is Gaussian (shown above) and Z_2 is a linear function of Z_1 (follows from (2)). Thus Z_1 and Z_2 are jointly Gaussian with

$$\begin{aligned} Z_1 \triangleq x_i|Y_{i-1} &\sim \mathcal{N}(\hat{x}_{i|i-1}, P_{i|i-1}) \quad \text{(follows from (3))} \\ Z_2|Z_1 \triangleq y_i|x_i, Y_{i-1} = y_i|x_i &\sim \mathcal{N}(h_i x_i, R_i) \quad \text{(follows from (2))} \end{aligned} \tag{4}$$

Using Bayes rule, one can compute the conditional distribution of $Z_1|Z_2 = x_i|Y_i$ (which will also be Gaussian).
Applying the formulas from Pg 155 (equation IV.B.49) of Poor's book (An Introduction to Signal Detection and Estimation), we have

$$\begin{aligned} \hat{x}_i \triangleq E[x_i|Y_i] &= \hat{x}_{i|i-1} + K_i(y_i - H_i \hat{x}_{i|i-1}) \\ P_i \triangleq Var(x_i|Y_i) &= (I - K_i H_i) P_{i|i-1}, \\ \text{where} \quad K_i &= P_{i|i-1} H_i^T (R_i + H_i P_{i|i-1} H_i^T)^{-1} \end{aligned} \tag{5}$$

Summarizing the Kalman Filter

$$
\begin{aligned}
\hat{x}_{i|i-1} &= F_i \hat{x}_{i-1} \\
P_{i|i-1} &= F_i P_{i-1} F_i^T + Q_i \\
K_i &= P_{i|i-1} H_i^T (R_i + H_i P_{i|i-1} H_i^T)^{-1} \\
\hat{x}_i &= \hat{x}_{i|i-1} + K_i (y_i - H_i \hat{x}_{i|i-1}) \\
P_i &= (I - K_i H_i) P_{i|i-1}
\end{aligned}
$$

For $F_i = I$, $Q_i = 0$, $h_i = H_i$, get the Recursive LS algorithm.

Example Applications: Kalman Filter v/s Recursive LS

- Kalman filter: Track a moving object (estimate its location and velocity at each time), assuming that velocity at current time is velocity at previous time plus Gaussian noise). Use a sequence of location observations coming in sequentially.
- Recursive LS: Keep updating estimate of location of an object that is static. Use a sequence of location observations coming in sequentially.
- Recursive LS with forgetting factor: object not static but drifts very very slowly.

Extended Kalman Filter

- State space model is nonlinear Gaussian, i.e.

$$
\begin{aligned}
x_0 &\sim \mathcal{N}(0, \Pi_0) \\
x_i &= f_i(x_{i-1}) + v_{x,i}, \;\; v_{x,i} \sim \mathcal{N}(0, Q_i) \qquad (6) \\
z_i &= h_i(x_i) + v_i, \;\; v_i \sim \mathcal{N}(0, R_i) \qquad (7)
\end{aligned}
$$

  where $f_i(x)$, $h_i(x)$ can both be nonlinear.

- Most commonly used form of Extended KF: At each time i,
  1. Linearize (6) about $\hat{x}_{i-1}$ and use the Kalman filter prediction step (3) with $F_i \triangleq \frac{\partial f_i}{\partial x}(\hat{x}_{i-1})$, to compute $\hat{x}_{i|i-1} - f_i(\hat{x}_{i-1})$.
  2. Linearize (7) about $\hat{x}_{i|i-1}$ and use the Kalman filter update step (4) with $H_i \triangleq \frac{\partial h_i}{\partial x}(\hat{x}_{i|i-1})$ and $y_i \triangleq z_i - h_i(\hat{x}_{i|i-1})$ to compute $\hat{x}_i$

Summarizing the Extended KF

$$
\begin{aligned}
F_i &= \frac{\partial f_i}{\partial x}(\hat{x}_{i-1}) \\
\hat{x}_{i|i-1} &= f_i(\hat{x}_{i-1}) \\
P_{i|i-1} &= F_i P_{i-1} F_i^T + Q_i \\
H_i &= \frac{\partial h_i}{\partial x}(\hat{x}_{i|i-1}) \\
K_i &= P_{i|i-1} H_i^T (R_i + H_i P_{i|i-1} H_i^T)^{-1} \\
\hat{x}_i &= \hat{x}_{i|i-1} + K_i(z_i - h_i(\hat{x}_{i|i-1})) \\
P_i &= (I - K_i H_i) P_{i|i-1}
\end{aligned}
$$

Material adapted from

- Chapters 2, 3 & 9 of Linear Estimation, by Kailath, Sayed, Hassibi

- Chapters 4 & 5 of An Introduction to Signal Detection and Estimation, by Vincent Poor