Kalman Filters

Emtiyaz

CS,UBC

A Linear State-Space model

- Consider the following linear (time-invariant) system:

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + B\mathbf{w}_t + F\mathbf{u}_t \quad (1)$$

$$\mathbf{y}_t = C\mathbf{x}_t + D\mathbf{v}_t + G\mathbf{u}_t \quad (2)$$

  with $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0)$, (i.i.d.) $\mathbf{w}_t \sim \mathcal{N}(0, Q)$, $\mathbf{v}_t \sim \mathcal{N}(0, R)$.
- Also $\mathbf{x}_0 \perp \mathbf{w}_t \perp \mathbf{v}_t$ for all t.
- This is also called a linear-dynamical system.
- Problem: Given measurements $\mathbf{y}_t$ and matrices $A - G$, find optimal $\mathbf{x}_t$ and $\mathbf{u}_t$.
- Separability of Control and Estimation: Find optimal state estimate $\hat{\mathbf{x}}_{t|t}$ with $\mathbf{u}_t = 0$, then optimal policy is a linear function of state estimation $\mathbf{u}_t = K_L \hat{\mathbf{x}}_{t|t}$.

A Linear State-Space model

We compute state estimates for the following model:

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + B\mathbf{w}_t \quad (3)$$

$$\mathbf{y}_t = C\mathbf{x}_t + D\mathbf{v}_t \quad (4)$$

Example A 2-D vehicle with speed control driven by noise,

$$\underbrace{\begin{bmatrix} p_t^x \\ s_t^x \\ p_t^y \\ s_t^y \end{bmatrix}}_{\mathbf{x}_t} = \underbrace{\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} p_{t-1}^x \\ s_{t-1}^x \\ p_{t-1}^y \\ s_{t-1}^y \end{bmatrix}}_{\mathbf{x}_{t-1}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} w_t^x \\ w_t^y \end{bmatrix}}_{\mathbf{w}_t} \quad (5)$$

$$\underbrace{\begin{bmatrix} x_t \\ y_t \end{bmatrix}}_{\mathbf{y}_t} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} p_t^x \\ s_t^x \\ p_t^y \\ s_t^y \end{bmatrix}}_{\mathbf{x}_t} + \underbrace{\begin{bmatrix} v_t^x \\ v_t^y \end{bmatrix}}_{\mathbf{v}_t} \quad (6)$$

Demo

Linear Projection

- Denote an estimate of $\mathbf{x}$ given $\mathbf{y}$ as $\hat{\mathbf{x}}_{|y}$ and error covariance $P_{|y} = \mathbb{E}[(\hat{\mathbf{x}}_{|y} - \mathbf{x})(\hat{\mathbf{x}}_{|y} - \mathbf{x})^T]$.
- Minimum mean-square estimate = $\hat{\mathbf{x}}_{|y} = \mathbb{E}(\mathbf{x}|\mathbf{y})$.
- The space of (zero-mean) random variable is an inner-product space with inner-product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbb{E}(\mathbf{x}\mathbf{y}^T)$.
- Hence MMSE estimates can be found by ortho. proj.:

$$\hat{\mathbf{x}}_{|y} = \langle \mathbf{x}, \mathbf{y} \rangle ||\mathbf{y}||^{-2}\mathbf{y} = \mathbb{E}(\mathbf{x}\mathbf{y}^T)\mathbb{E}(\mathbf{y}\mathbf{y}^T)^{-1}\mathbf{y} \quad (8)$$

- If $\mathbf{x}$ is uncorrelated with $\mathbf{y}$, $\hat{\mathbf{x}}_{|y} = 0$ and $P_{|y} = \mathbb{E}(\mathbf{x}\mathbf{x}^T)$.
- $\hat{\mathbf{x}}_{|\mathbf{y}_1,\mathbf{y}_2} = \hat{\mathbf{x}}_{|\mathbf{y}_1} + \hat{\mathbf{x}}_{|\mathbf{y}_2}$ iff $\mathbf{y}_1 \perp \mathbf{y}_2$, and $P_{|y} = P_{|y1} + P_{|y2}$.
- Notation: $\hat{\mathbf{x}}_{t1|t2}, P_{t1|t2}$ and $\mathbf{e}^x_{t1|t2} = \hat{\mathbf{x}}_{t1|t2} - \mathbf{x}_{t1}$

Prediction Estimates

Predict using state equations: $\mathbf{x}_t = A\mathbf{x}_{t-1} + B\mathbf{w}_t$.
Given an estimate at $t-1$, propagate as follows:

$$
\begin{aligned}
\text{Prediction}: \hat{\mathbf{x}}_{t|t-1} &= A\hat{\mathbf{x}}_{t-1|t-1} \\
P_{t|t-1} &= AP_{t-1|t-1}A^T + BQB^T \\
\text{Proof}: \hat{\mathbf{x}}_{t|t-1} &= A\hat{\mathbf{x}}_{t-1|t-1} + B\hat{\mathbf{w}}_t = A\hat{\mathbf{x}}_{t-1|t-1} \\
\mathbf{e}^x_{t|t-1} &= \hat{\mathbf{x}}_{t|t-1} - \mathbf{x}_t \\
&= A\hat{\mathbf{x}}_{t-1|t-1} - \mathbf{x}_t \\
&= A\hat{\mathbf{x}}_{t-1|t-1} - A\mathbf{x}_{t-1} - B\mathbf{w}_t \\
&= A\mathbf{e}^x_{t-1|t-1} - B\mathbf{w}_t \\
\Rightarrow P_{t|t-1} &= AP_{t-1|t-1}A^T + BQB^T
\end{aligned}
$$

Initialize with some $\hat{\mathbf{x}}_{0|0}, P_{0|0}$. Error keeps increasing!!

Innovation

Find "new information" in the current measurement (use measurement equation $\mathbf{y}_t = C\mathbf{x}_t + D\mathbf{v}_t$):

$$\begin{aligned}
\text{Innovations : } \mathbf{e}_t &= \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} = \mathbf{y}_t - C\hat{\mathbf{x}}_{t|t-1} \\
P_{e,t} &= CP_{t|t-1}C^T + DRD^T \\
\text{Proof : } \hat{\mathbf{y}}_{t|t-1} &= C\hat{\mathbf{x}}_{t|t-1} + D\hat{\mathbf{v}}_{t|t-1} = C\hat{\mathbf{x}}_{t|t-1} \\
\mathbf{e}_t &= C\mathbf{x}_t + D\mathbf{v}_t - C\hat{\mathbf{x}}_{t|t-1} \\
&= C\mathbf{e}^x_{t|t-1} + D\mathbf{v}_t \\
\Rightarrow P_{e,t} &= CP_{t|t-1}C^T + DRD^T
\end{aligned}$$

Correction with Kalman Gain

$$\text{Correction : } \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + K_t \mathbf{e}_t \quad (9)$$

$$P_{t|t} = (I - K_t C) P_{t|t-1} + K_t P_{e,t} K_t^T \quad (10)$$

$$\text{Proof : } \mathbf{e}^x_{t|t} = \hat{\mathbf{x}}_{t|t} - \mathbf{x}_t \quad (11)$$

$$= \hat{\mathbf{x}}_{t|t-1} + K_t \mathbf{e}_t - \mathbf{x}_t \quad (12)$$

$$= \mathbf{e}^x_{t|t-1} + K_t \mathbf{e}_t \quad (13)$$

$$\mathbb{E}[\mathbf{e}^x_{t|t-1} \mathbf{e}_t^T K_t^T] = \mathbb{E}[\mathbf{e}^x_{t|t-1} (C \mathbf{e}^x_{t|t-1} + D \mathbf{v}_t)^T K_t^T] \quad (14)$$

$$= P_{t|t-1} C^T K_t^T = K_t C P_{t|t-1} \quad (15)$$

$$P_{t|t} = (I - K_t C) P_{t|t-1} + K_t P_{e,t} K_t^T \quad (16)$$

Find K_t^* (called Kalman gain) such that $P_{t|t}$ is minimized

$$K_t^* = \arg\min K_t P_{e,t} K_t^T - K_t C P_{t|t-1} = P_{t|t-1} C^T P_{e,t}^{-1} \quad (17)$$

Kalman Filters

$$
\begin{aligned}
\text{Predictions : } \hat{\mathbf{x}}_{t|t-1} &= A\hat{\mathbf{x}}_{t-1|t-1} \\
P_{t|t-1} &= AP_{t-1|t-1}A^T + BQB^T \\
\text{Innovations : } \mathbf{e}_t &= \mathbf{y}_t - C\hat{\mathbf{x}}_{t|t-1} \\
P_{e,t} &= CP_{t|t-1}C^T + DRD^T \\
\text{Kalman Gain : } K_t &= P_{t|t-1}C^T P_{e,t}^{-1} \\
\text{Correction : } \hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + K_t\mathbf{e}_t \\
P_{t|t} &= (I - K_tC)P_{t|t-1} + K_tP_{e,t}K_t^T
\end{aligned}
$$

Optimality

- Unbiased if $\hat{\mathbf{x}}_{0|0}$ is unbiased, otherwise asymptotic.
- $P_{t|t}$ is covariance, achieves Cramer-Rao Lower Bound.
- Best linear unbiased estimator.
- Solve the following Riccati equation for P to get the converged covariance.

$$
\begin{aligned}
P_1 &= APA^T + BQB^T \\
P_e &= CP_1C^T + DRD^T \\
K &= P_1C^TP_e^{-1} \\
P &= (I - KC)P_1 + KP_eK^T
\end{aligned}
$$

- then $K = PC^TP_e^{-1}$ can be computed beforehand, reduces the computational complexity.

Implementation

- Square-Root filters.
- Information filters (IF).
- Sparse Kalman Filters.
- EM algorithm to find A, B, C, D.
- Extensions : Extended KF, IF, Unscented KF.

References

- Optimal Filtering, Anderson and Moore (contains almost everything)
- Linear Estimation, T. Kailath (for orthog.-proj. derivations etc.)