BAYESIAN KALMAN FILTER

A short study note

Author
Brian W.

Contents

1 Some propositions

1.1 $Var(x+y)$

$$\begin{aligned}
Var(x+y)_{ij} &= E[(x_i + y_i - E[x_i] - E[y_i])(x_j + y_j - E[x_j] - E[y_j])] \\
&= E[(x_i - E[x_i])(x_j - E[x_j])] + E[(y_i - E[y_i])(y_j - E[y_j])] \\
&\quad + E[(x_i - E[x_i])(y_j - E[y_j])] + E[(x_j - E[x_j])(y_i - E[y_i])] \\
&= Var(x)_{ij} + Var(y)_{ij} + Cov(x,y)_{ij} + Cov(y,x)_{ij} \\
\Rightarrow Var(x+y) &= Var(x) + Var(y) + Cov(x,y) + Cov(x,y)^T
\end{aligned}$$

1.2 $Cov(Ax,y)$

$$\begin{aligned}
Cov(Ax,y)_{ij} &= Cov((Ax)_i, y_j) \\
&= Cov(\sum_k A_{ik} x_k, y_j) \\
&= \sum_k A_{ik} Cov(x_k, y_j) \\
\Rightarrow Cov(Ax,y) &= ACov(x,y)
\end{aligned}$$

1.3 $Cov(y,Ax)$

$$\begin{aligned}
Cov(y,Ax)_{ij} &= Cov(y_i, (Ax)_j) \\
&= Cov(y_i, \sum_k A_{jk} x_k) \\
&= \sum_k A_{jk} Cov(y_i, x_k) \\
\Rightarrow Cov(y,Ax) &= Cov(y,x)A^T
\end{aligned}$$

1.4 $Var(Ax)$

$$\begin{aligned}
Var(Ax)_{ij} &= Cov(\sum_{k_1} A_{ik_1} x_{k_1}, \sum_{k_2} A_{jk_2} x_{k_2}) \\
&= \sum_{k_1} \sum_{k_2} A_{ik_1} A_{jk_2} Cov(x_{k_1}, x_{k_2}) \\
&= \sum_{k_1} A_{ik_1} \sum_{k_2} Cov(x_{k_1}, x_{k_2}) A_{jk_2} \\
&= \sum_{k_1} A_{ik_1} (Var(x)A^T)_{k_1 j} \\
&= (AVar(x)A^T)_{ij} \\
\Rightarrow Var(Ax) &= AVar(x)A^T
\end{aligned}$$

2 Linear Dynamic Systems

$$x_k = F_k x_{k-1} + v_k$$
$$z_k = H_k x_k + w_k$$

- F_k: the state-transition matrix
- H_k: the observation matrix
- x_k: true state
- z_k: measurement
- v_k: process noise, $v_k \sim \mathcal{N}(0, Q_k)$
- w_k: the covariance of the observation noise, $w_k \sim \mathcal{N}(0, R_k)$

3 Kalman Filter: Predict

$P(x_{k-1}|z_{k-1}) = \mathcal{N}(m_{k-1}, P_{k-1})$,
$P(x_k|x_{k-1}) = \mathcal{N}(F_k x_{k-1}, Q_k)$,
$P(z_k|x_k) = \mathcal{N}(H_k x_k, R_k)$

$$\begin{bmatrix} x_{k-1}|z_{k-1} \\ x_k|z_{k-1} \end{bmatrix} \sim N(\begin{bmatrix} m_{k-1} \\ F_k m_{k-1} \end{bmatrix}, \begin{bmatrix} P_{k-1} & P_{k-1}F_k^T \\ F_k P_{k-1} & Q_k + F_k P_{k-1} F_k^T \end{bmatrix})$$

$$P(x_k|z_{k-1}) = \mathcal{N}(F_k m_{k-1}, Q_k + F_k P_{k-1} F_k^T) = \mathcal{N}(x_{k|k-1}, P_{k|k-1})$$

3.1 Derivation

$$\begin{aligned} E[x_k|z_{k-1}] &= \int x_k P(x_k|z_{k-1})\, dx_k \\ &= \iint x_k P(x_k, x_{k-1}|z_{k-1})\, dx_k\, dx_{k-1} \\ &= \iint x_k P(x_k|x_{k-1}) P(x_{k-1}|z_{k-1})\, dx_k\, dx_{k-1} \\ &= \int P(x_{k-1}|z_{k-1}) \int x_k P(x_k|x_{k-1})\, dx_k\, dx_{k-1} \\ &= \int F_k x_{k-1} \mathcal{N}(m_{k-1}, P_{k-1})\, dx_{k-1} \\ &= F_k m_{k-1} \end{aligned}$$

$$
\begin{aligned}
Cov(x_{k-1}, x_k|z_k)_{ij} &= Cov(x_{k-1}^{(i)}, x_k^{(j)}|z_k) \\
&= E[x_{k-1}^{(i)}, x_k^{(j)}|z_k] - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \iint x_{k-1}^{(i)} x_k^{(j)} P(x_{k-1}^{(i)}, x_k^{(j)}|z_k)\, dx_{k-1}^{(i)}\, dx_k^{(j)} - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \iint x_{k-1}^{(i)} x_k^{(j)} \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1}) \mathcal{N}(x_k; F_k x_{k-1}, Q_k)\, dx_{k-1}^{(i)}\, dx_k^{(j)} - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \int x_{k-1}^{(i)} \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1}) \int x_k^{(j)} \mathcal{N}(x_k; F_k x_{k-1}, Q_k)\, dx_k^{(j)}\, dx_{k-1}^{(i)} - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \int x_{k-1}^{(i)} \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1}) (F_k x_{k-1})_j\, dx_{k-1} - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \sum_\alpha (F_k)_{j\alpha} \int x_{k-1}^{(\alpha)} x_{k-1}^{(i)} \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1})\, dx_{k-1} - (m_{k-1})_i(F_k m_{k-1})_j \\
&= \sum_\alpha (F_k)_{j\alpha} [(P_{k-1})_{\alpha i} + (m_{k-1})_\alpha (m_{k-1})_i] - (m_{k-1})_i(F_k m_{k-1})_j \\
&= (F_k P_{k-1})_{ij}^T \\
\Rightarrow Cov(x_{k-1}, x_k|z_{k-1}) &= (F_k P_{k-1})^T = P_{k-1}^T F_k^T = P_{k-1} F_k^T
\end{aligned}
$$

$$
\begin{aligned}
Var(x_k|z_{k-1}) &= E[x_k x_k^T|z_{k-1}] - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= \int x_k x_k^T P(x_k|z_{k-1})\, dx_k - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= \iint x_k x_k^T \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1}) \mathcal{N}(x_k; F_k x_{k-1}, Q_k)\, dx_{k-1}\, dx_k - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= \int \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1}) \int x_k x_k^T \mathcal{N}(x_k; F_k x_{k-1}, Q_k)\, dx_k\, dx_{k-1} - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= \int (Q_k + F_k x_{k-1} x_{k-1}^T F_k^T) \mathcal{N}(x_{k-1}; m_{k-1}, P_{k-1})\, dx_{k-1} - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= Q_k + F_k P_{k-1} F_k^T + F_k m_{k-1} m_{k-1}^T F_k^T - F_k m_{k-1} m_{k-1}^T F_k^T \\
&= Q_k + F_k P_{k-1} F_k^T
\end{aligned}
$$

4 Kalman Filter: Update

$P(x_k, z_k|z_{k-1}) = P(z_k|x_k)P(x_k|z_{k-1})$, where $P(z_k|x_k) = \mathcal{N}(H_k x_k, R_k)$, $P(x_k|z_{k-1}) = \mathcal{N}(x_{k|k-1}, P_{k|k-1})$

$$
\begin{bmatrix} x_k|z_{k-1} \\ z_k|z_{k-1} \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} x_{k|k-1} \\ H_k x_{k|k-1} \end{bmatrix}, \begin{bmatrix} P_{k|k-1} & P_{k|k-1} H_k^T \\ H_k P_{k|k-1} & R_k + H_k P_{k|k-1} H_k^T \end{bmatrix})
$$

$$
P(x_k|z_k) = \mathcal{N}(E[x_k|z_k],\, Var(x_k|z_k)) = \mathcal{N}(x_{k|k}, P_{k|k})
$$

4.1 Derivation

$$\begin{aligned}
E[z_k|z_{k-1}] &= \int z_k P(z_k|z_{k-1})\,dz_k \\
&= \iint z_k P(x_k, z_k|z_{k-1})\,dx_k\,dz_k \\
&= \iint z_k \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})\mathcal{N}(z_k; H_k x_k, R_k)\,dx_k\,dz_k \\
&= \int \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1}) \int z_k \mathcal{N}(z_k; H_k x_k, R_k)\,dz_k\,dx_k \\
&= \int H_k x_k \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})\,dx_k \\
&= H_k x_{k|k-1}
\end{aligned}$$

$$\begin{aligned}
Cov(x_k, z_k|z_{k-1}) &= E[x_k z_k^T|z_{k-1}] - x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= \int \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1}) \int x_k z_k^T \mathcal{N}(z_k; H_k x_k, R_k)\,dz_k\,dx_k - x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= \int x_k x_k^T H_k^T \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})\,dx_k - x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= P_{k|k-1}H_k^T + x_{k|k-1}x_{k|k-1}^T H_k^T - x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= P_{k|k-1}H_k^T
\end{aligned}$$

$$\begin{aligned}
Var(z_k|z_{k-1}) &= E[z_k z_k^T|z_{k-1}] - H_k x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= \int \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1}) \int z_k z_k^T \mathcal{N}(z_k; H_k x_k, R_k)\,dz_k\,dx_k - H_k x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= \int (R_k + H_k x_k x_k^T H_k^T)\mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})\,dx_k - H_k x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= R_k + H_k P_{k|k-1}H_k^T + H_k x_{k|k-1}x_{k|k-1}^T H_k^T - H_k x_{k|k-1}x_{k|k-1}^T H_k^T \\
&= R_k + H_k P_{k|k-1}H_k^T
\end{aligned}$$

$P(x_k|z_k) =?$

Let $a = Ax_k + Bz_k \perp z_k \Rightarrow AP_{k|k-1}H_k^T + BR_k + H_k P_{k|k-1}H_k^T = 0 \Rightarrow a = x_k - P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1}z_k = x_k + Mz_k$ $M = -P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1}$

$$\begin{aligned}
E[x_k|z_k] &= E[a - Mz_k|z_k] \\
&= E[a] - Mz_k \\
&= E[x_k] + M(E[z_k] - z_k) \\
&= x_{k|k-1} + P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1}(z_k - H_k x_{k|k-1})
\end{aligned}$$

$$\begin{aligned}
Var(x_k|z_k) &= Var(a - Mz_k|z_k) \\
&= Var(a) \\
&= Var(x_k) + MVar(z_k)M^T \\
&= P_{k|k-1} - P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1}(R_k + H_k P_{k|k-1}H_k^T)(R_k + H_k P_{k|k-1}H_k^T)^{-1}H_k P_{k|k-1} \\
&= P_{k|k-1} - P_{k|k-1}H_k^T(R_k + H_k P_{k|k-1}H_k^T)^{-1}H_k P_{k|k-1}
\end{aligned}$$

5 Summary

5.1 Algorithm

Predict:

$$x_{k|k-1} = F_k x_{k-1|k-1}$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

Update:

$$x_{k|k} = x_{k|k-1} + P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} (z_k - H_k x_{k|k-1})$$
$$P_{k|k} = P_{k|k-1} - P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} H_k P_{k|k-1}$$

5.2 Introducing Kalman Gain and Some Variables

We introduce the following variables:

- $z_{k|k-1} = H_k x_{k|k-1}$: This is the expected measurement.
- $S_k = R_k + H_k P_{k|k-1}$: This represents the expected measurement covariance.
- $v_k = z_k - z_{k|k-1}$: This is the measurement residual.
- $K_k = P_{k|k-1} H_k^T S_k^{-1}$: This is the Kalman Gain.

5.3 Rewrite the algorithm with the introduced variables

Predict:

$$x_{k|k-1} = F_k x_{k-1|k-1}$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

Update:

$$x_{k|k} = x_{k|k-1} + K_k v_k$$
$$P_{k|k} = P_{k|k-1} - K_k S_k^{-1} K_k^T$$

6 Reference

References

[1] Särkkä, S., *Bayesian Filtering and Smoothing*, Cambridge University Press, 2013.