Exercise Sheet 8

Robin Schmidt
Probabilistic Inference & Learning

December 7, 2018

Gauss-Markov Models

1. Kalman Filters

For the Kalman predicition we get:

$$
\begin{aligned}
p(x_t|Y_{0:t-1}) &= \int p\left(x_t|x_{t-1}\right) p\left(x_{t-1}|Y_{0:t-1}\right) dx_{t-1} \\
&= \int \mathcal{N}\left(x_t; Ax_{t-1}, Q\right) \cdot \left(\int \mathcal{N}\left(x_{t-1}; Ax_{t-2}, Q\right) \cdot p(x_{t-2}|Y_{0:t-2}) \right)
\end{aligned}
$$

Through this recursive property going back to state x_0 with $p\left(x_0\right) = \mathcal{N}\left(x_0; m_0^-, P_0^-\right)$ one can see that: $p\left(x_t|Y_{0:t-1}\right) = \mathcal{N}\left(x_t; Am_{t-1}, AP_{t-1}A^\top + Q\right)$.

For the Kalman estimation we get:

$$
\begin{aligned}
p\left(x_t|Y_{0:t}\right) &= \frac{p\left(y_t|x_t\right) p\left(x_t|Y_{0:t-1}\right)}{p\left(y_t\right)} \\
&= \frac{\mathcal{N}\left(y_t; Hx_t, R\right) \int p\left(x_t|x_{t-1}\right) p\left(x_{t-1}|Y_{0:t-1}\right) dx_{t-1}}{\sum_{j=1}^{n} \mathcal{N}\left(y_t; Hx_t, R\right) \mathcal{N}\left(x_t; m_t^-, P_t^-\right)} \\
&= \frac{\mathcal{N}\left(y_t; Hx_t, R\right) \mathcal{N}\left(x_t; m_t^-, P_t^-\right)}{\sum_{j=1}^{n} \mathcal{N}\left(y_t; Hx_t, R\right) \mathcal{N}\left(x_t; m_t^-, P_t^-\right)} \\
&= \mathcal{N}\left(x_t; m_t^- + P_t^- H^\top \left(HP_t^- H^\top + R\right)^{-1} \left(y_t - Hm_t^-\right), (I - P_t^- H^\top \left(HP_t^- H^\top + R\right)^{-1} H)P_t^-\right)
\end{aligned}
$$

Which is essentially $\mathcal{N}\left(x_t; m_t^- + K_t\left(y_t - Hm_t^-\right), (I - KH)P_t^-\right)$ for $K_t := P_t^- H^\top \left(HP_t^- H^\top + R\right)^{-1}$.

For the Rauch-Tung-Striebel smoothed estimation we get:

$$
\begin{aligned}
p\left(x_t|Y\right) &= p\left(x_t|Y_{0:t}\right)\int p\left(x_{t+1}|x_t\right)\frac{p\left(x_{t+1}|Y\right)}{p\left(x_{t+1}|Y_{0:t}\right)}dx_{t+1} \\
&= \mathcal{N}\left(x_t; m_t, P_t\right)\int \mathcal{N}\left(x_{t+1}; Ax_t, Q\right)\frac{p\left(x_{t+1}|Y\right)}{\mathcal{N}\left(x_{t+1}; m_{t+1}, P_{t+1}\right)}dx_{t+1}
\end{aligned}
$$

2. Stochastic Differential Equations

$$
\begin{aligned}
m(t) &= e^{F(t-t_0)}x_0 \\
k\left(t_a, t_b\right) &= \int_{t_0}^{\min(t_a,t_b)} e^{F(t_a-\tau)}LL^\top e^{F^\top(t_b-\tau)}d\tau \qquad = \left[-\frac{LL^\top e^{F^\top(t_b-\tau)+F(t_a-\tau)}}{F+F^\top}\right]_{t_0}^{\min(t_a,t_b)}
\end{aligned}
$$

(a)

Starting off with $F=0$ and $L=\theta$ and substituting in the equations from above gives:

$$
\begin{aligned}
m(t) &= e^{0(t-t_0)}x_0 \\
&= x_0 \\
k\left(t_a, t_b\right) &= \int_{t_0}^{\min(t_a,t_b)} e^{F(t_a-\tau)}LL^\top e^{F^\top(t_b-\tau)}d\tau \\
&= \int_{t_0}^{\min(t_a,t_b)} e^{0(t_a-\tau)}\theta\theta^\top e^{0^\top(t_b-\tau)}d\tau \\
&= \int_{t_0}^{\min(t_a,t_b)} \theta\theta^\top d\tau \\
&= \int_{t_0}^{\min(t_a,t_b)} \theta\theta^\top d\tau \\
&= \left[\theta\theta^\top\tau\right]_{t_0}^{\min(t_a,t_b)} \\
&= \theta\theta^\top\cdot\min\left(t_a, t_b\right) - \theta\theta^\top t_0 \\
&= \theta^2(\min\left(t_a, t_b\right) - t_0)
\end{aligned}
$$

(b)

Starting off with $F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $L = \begin{bmatrix} 0 \\ \theta \end{bmatrix}$ and substituting in the equations from above gives:

$$
\begin{aligned}
m(t) &= e^{F(t-t_0)}x_0 \\
&= e^{Ft-Ft_0}x_0
\end{aligned}
$$

Computing $Ft - Ft_0$ gives the matrix $A = \begin{bmatrix} 0 & t - t_0 \\ 0 & 0 \end{bmatrix}$, also rewriting $e^x := \sum_{n=0}^{\infty} \frac{\chi^n}{n!}$ gives:

$$\begin{aligned} m(t) &= e^A x_0 \\ &= x_0 \sum_{n=0}^{\infty} \frac{A^n}{n!} \\ &= x_0 \sum_{n=0}^{\infty} \frac{\begin{bmatrix} 0 & t - t_0 \\ 0 & 0 \end{bmatrix}^n}{n!} \end{aligned}$$

Now looking at the matrix A shows, that it is nilpotent for all $n \geq 2$ this yields the following:

$$\begin{aligned} m(t) &= x_0 \left(\frac{\begin{bmatrix} 0 & t - t_0 \\ 0 & 0 \end{bmatrix}^1}{1!} + \frac{\begin{bmatrix} 0 & t - t_0 \\ 0 & 0 \end{bmatrix}^0}{0!} \right) \\ &= x_0 \left(\begin{bmatrix} 0 & t - t_0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} 1 & t - t_0 \\ 0 & 1 \end{bmatrix} x_0 \end{aligned}$$

Using the general solution for the integral of the kernel and rewriting $e^x := \sum_{n=0}^{\infty} \frac{\chi^n}{n!}$ again gives the following:

$$\begin{aligned} k(t_a, t_b) &= \left[-\frac{LL^\top e^{F^\top (t_b - \tau) + F(t_a - \tau)}}{F + F^\top} \right]_{t_0}^{\min(t_a, t_b)} \\ &= \left[-\frac{[0 \ \theta^2] \sum_{n=0}^{\infty} \frac{(F^\top (t_b - \tau) + F(t_a - \tau))^n}{n!}}{F + F^\top} \right]_{t_0}^{\min(t_a, t_b)} \end{aligned}$$

Because F is nilpotent for all $n \geq 2$ this yields the following:

$$k(t_a, t_b) = \left[-\frac{\begin{bmatrix} 0 \\ \theta^2 \end{bmatrix} \cdot \left(\frac{(F^\top(t_b-\tau)+F(t_a-\tau))^1}{1!} + \frac{(F^\top(t_b-\tau)+F(t_a-\tau))^0}{0!} \right)}{F+F^\top} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[-\frac{\begin{bmatrix} 0 \\ \theta^2 \end{bmatrix} \cdot \left(F^\top(t_b-\tau)+F(t_a-\tau)+I \right)}{F+F^\top} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[-\frac{\begin{bmatrix} 0 \\ \theta^2 \end{bmatrix} \cdot \left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}(t_b-\tau) + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}(t_a-\tau)+I \right)}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[-\frac{\begin{bmatrix} 0 \\ \theta^2 \end{bmatrix} \cdot \begin{bmatrix} 1 & t_a-\tau \\ t_b-\tau & 1 \end{bmatrix}}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[-\frac{\begin{bmatrix} \theta^2(t_a-\tau) \\ \theta^2 \end{bmatrix}}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[\begin{bmatrix} \theta^2(t_a-\tau) \\ \theta^2 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}^{-1} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[\begin{bmatrix} \theta^2(t_a-\tau) \\ \theta^2 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \left[\begin{bmatrix} -\theta^2 \\ -\theta^2(t_a-\tau) \end{bmatrix} \right]_{t_0}^{\min(t_a,t_b)}$$

$$= \begin{bmatrix} -\theta^2 \\ -\theta^2(t_a - \min(t_a,t_b)) \end{bmatrix} - \begin{bmatrix} -\theta^2 \\ -\theta^2(t_a-t_0) \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ \theta^2 \min(t_a,t_b) - \theta^2 t_0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ \theta^2(\min(t_a,t_b) - t_0) \end{bmatrix}$$

(c)

Starting off with $F = -\xi$ and $L = \theta$ and substituting in the equations from above gives:

$$m(t) = e^{-\xi(t-t_0)} x_0$$

Using the general solution for the integral of the kernel gives the following:

$$\begin{aligned}
k\left(t_a, t_b\right) &= \left[-\frac{\theta\theta^\top e^{(-\xi)^\top (t_b-\tau)+(-\xi)(t_a-\tau)}}{(-\xi)+(-\xi)^\top} \right]_{t_0}^{\min(t_a,t_b)} \\
&= \left[-\frac{\theta^2 e^{-\xi(t_a+t_b-2\tau)}}{-2\xi} \right]_{t_0}^{\min(t_a,t_b)} \\
&= \left(-\frac{\theta^2 e^{-\xi(t_a+t_b-2\min(t_a,t_b))}}{-2\xi} \right) - \left(-\frac{\theta^2 e^{-\xi(t_a+t_b-2t_0)}}{-2\xi} \right) \\
&= \frac{-(\theta^2 e^{-\xi(t_a+t_b-2\min(t_a,t_b))}) + \theta^2 e^{-\xi(t_a+t_b-2t_0)}}{-2\xi} \\
&= \frac{\theta^2 e^{-\xi(t_a+t_b)} (e^{2\xi min(t_a,t_b)} + e^{2\xi t_0})}{-2\xi}
\end{aligned}$$

(d)

Starting off with $F = \begin{bmatrix} 0 & 1 \\ -\xi^2 & -2\xi \end{bmatrix}$ and $L = \begin{bmatrix} 0 \\ \theta \end{bmatrix}$ and using the hint we first compute the eigenvalues λ from F:

$$\begin{aligned}
det(\lambda I - F) &= det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -\xi^2 & -2\xi \end{bmatrix} \right) \\
&= det\left(\begin{bmatrix} \lambda & -1 \\ \xi^2 & \lambda + 2\xi \end{bmatrix} \right) \\
&= \lambda(\lambda + 2\xi) + \xi^2 \\
&= \lambda^2 + 2\xi\lambda + \xi^2
\end{aligned}$$

Solving the characteristic polynomial gives the eigenvalues λ:

$$\begin{aligned}
\lambda_{1,2} &= \frac{-2\xi \pm \sqrt{(2\xi)^2 - 4 \cdot 1 \cdot \xi^2}}{2} \\
\lambda_1 &= -\xi \\
\lambda_2 &= -\xi
\end{aligned}$$

The eigenvectors are aquired by solving $Fx = \lambda x$ like:

$$\begin{aligned}
\begin{bmatrix} 0 & 1 \\ -\xi^2 & -2\xi \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
\begin{bmatrix} 0 & 1 \\ -\xi^2 & -2\xi \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} -\xi x_1 \\ -\xi x_2 \end{bmatrix} \\
\text{I} \quad 0 \cdot x_1 + 1 \cdot x_2 &= -\xi x_1 \\
\text{II} \quad -\xi^2 x_1 - 2\xi x_2 &= -\xi x_2
\end{aligned}$$

This yields the eigenvector for $\lambda = -\xi$ with $x_\lambda = \lambda \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. If we try to compute the mean we get:

$$\begin{aligned}
m(t) &= e^{F(t-t_0)} x_0 \\
&= e^{Ft - Ft_0} x_0
\end{aligned}$$

Computing $Ft - Ft_0$ gives the matrix $A = \begin{bmatrix} 0 & t - t_0 \\ \xi^2(t_0 - t) & 2\xi(t_0 - t) \end{bmatrix}$, also rewriting $e^x := \sum_{n=0}^{\infty} \frac{\chi^n}{n!}$ gives:

$$\begin{aligned}
m(t) &= e^A x_0 \\
&= x_0 \sum_{n=0}^{\infty} \frac{A^n}{n!} \\
&= x_0 \sum_{n=0}^{\infty} \frac{\begin{bmatrix} 0 & t - t_0 \\ \xi^2(t_0 - t) & 2\xi(t_0 - t) \end{bmatrix}^n}{n!}
\end{aligned}$$

Unfortunately the matrix F isn't nilpotent nor diagonalizable, so we need to use jordan normal form. F can therefore be written as $F = VJV^{-1}$, where J is the jordan normal form with $J = D + N$, with the diagonal matrix of eigenvalues D and a nilpotent matrix N. This leads to $J = \begin{bmatrix} -\xi & 1 \\ 0 & -\xi \end{bmatrix}$ for our example. We can now rewrite $e^F = e^{VJV^{-1}} = Ve^J V^{-1} = V(e^D e^N)V^{-1}$ which shows that all $n \geq 2$ for the sum notation of e^F cancel. This leads to:

$$\begin{aligned}
m(t) &= x_0 \left(\frac{(Ft - Ft_0)^1}{1!} + \frac{(Ft - Ft_0)^0}{0!} \right) \\
&= x_0 \left(\begin{bmatrix} 0 & t - t_0 \\ \xi^2(t_0 - t) & 2\xi(t_0 - t) \end{bmatrix} + I \right) \\
&= \begin{bmatrix} 1 & t - t_0 \\ \xi^2(t_0 - t) & 2\xi(t_0 - t) + 1 \end{bmatrix} x_0
\end{aligned}$$