# Exercise Sheet #8

**Sample Solutions**

December 13, 2018

In Exercise 1, we will use the following lemma, which was not explicitly given in the lecture, but follows from the material presented in Lecture 3. You are free to use it in the following exercises and in the exam!

**Lemma 1.** *If random variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ have the Gaussian probability distributions*

$$p(x) = \mathcal{N}(x; \mu, \Sigma), \tag{1}$$
$$p(y|x) = \mathcal{N}(y; Ax + b, \Lambda), \tag{2}$$

*then the joint distribution of $(x, y)$ is given by*

$$p(x, y) = \mathcal{N}\left(\begin{pmatrix} x \\ y \end{pmatrix}; \begin{pmatrix} \mu \\ A\mu + b \end{pmatrix} \begin{pmatrix} \Sigma & \Sigma A^\intercal \\ A\Sigma & A\Sigma A^\intercal + \Lambda \end{pmatrix}\right). \tag{3}$$

## 1 Exercise 1

Since $p(x_0) = \mathcal{N}(x_0; m_0^-, P_0^-)$, it is left to show that

$$p(x_t|y_{0:t}) = \mathcal{N}\left(x_t; m_t^- - K_t(y_t - Hm_t^-), (I - K_tH)P_t^-\right), \quad \text{and} \tag{4}$$
$$p(x_t|y_{0:t-1}) = \mathcal{N}\left(x_t; Am_{t-1}, AP_{t-1}A^\intercal + Q\right), \quad \text{and} \tag{5}$$
$$p(x_t|Y) = \mathcal{N}\left(x_t; m_t + G_t(m_{t+1}^s - m_{t+1}^-), P_t + G_t(P_{t+1}^s - P_{t+1}^-)G_t^\intercal\right), \tag{6}$$

with *Kalman gain*

$$K_t = P_t^- H^\intercal (HP_t^- H^\intercal + R)^{-1}, \tag{7}$$

and *RTS gain*

$$G_t = P_t A^\intercal (P_{t+1}^{-1})^{-1}. \tag{8}$$

*Proof.* To show eq. (4), we assume that $p(x_t|y_{0:t-1}) = \mathcal{N}\left(x_t; m_t^-, P_t^-\right)$. Note by Lemma 1 that, given $y_{0:t-1}$, $(x_t, y_t)$ are jointly Gaussian distributed:

$$\begin{aligned}
p\left(x_t, y_t|y_{0:t-1}\right) &= \underbrace{p(y_t|x_t, y_{0:t-1})}_{=p(y_t|x_t)} p(x_t|y_{0:t-1}) = \mathcal{N}(y_t; Hx_t, R)\mathcal{N}(x_t; m_t^-, P_t^-) \\
&= \mathcal{N}\left(\begin{pmatrix} x_t \\ y_t \end{pmatrix}; \begin{pmatrix} m_t^- \\ Hm_t^- \end{pmatrix}, \begin{pmatrix} P_t^- & P_t^- H^\intercal \\ HP_t^- & HP_t^- H^\intercal + R \end{pmatrix}\right). \qquad (9)
\end{aligned}$$

Now, by the theorem from Slide 16 (Lecture 03), eq. (4) follows:

$$\begin{aligned}
p(x_t|y_{0:t}) &= \mathcal{N}\left(x_t; m_t^- + P_t^- H^\intercal (HP_t^- H^\intercal + R)^{-1}(y_t - Hm_t^-), P_t^- - P_t^- H^\intercal (HP_t^- H^\intercal + R)^{-1} HP_t^-\right) \\
&\overset{eq.\ (7)}{=} \mathcal{N}\left(x_t; m_t^- + K_t(y_t - Hm_t^-), P_t^- + K_t HP_t^-\right). \qquad (10)
\end{aligned}$$

To show eq. (5), we assume that $p(x_{t-1}|y_{0:t-1}) = \mathcal{N}(x_{t-1}; m_{t-1}, P_{t-1})$. Note that by Lemma 1, given $y_{0:t-1}$, $(x_{t-1}, x_t)$ are jointly Gaussian distributed:

$$\begin{aligned}
p(x_{t-1}, x_t|y_{0:t-1}) &= \underbrace{p(x_t|x_{t-1}, y_{0:t-1})}_{=p(x_t|x_{t-1})} p(x_{t-1}|y_{0:t-1}) = \mathcal{N}(x_t; Ax_{t-1}, Q)\mathcal{N}(x_{t-1}; m_{t-1}, P_{t-1}) \\
&= \mathcal{N}\left(\begin{pmatrix} x_{t-1} \\ x_t \end{pmatrix}; \begin{pmatrix} m_{t-1} \\ Am_{t-1} \end{pmatrix}, \begin{pmatrix} P_{t-1} & P_{t-1}A^\intercal \\ AP_{t-1} & AP_{t-1}A^\intercal + Q \end{pmatrix}\right). \qquad (11)
\end{aligned}$$

Now, by 'marginals of Gaussians are Gaussians' (slide 17, lecture 03), eq. (5) follows immediately.
To show the remaining equation eq. (6), we assume that $p(x_{t+1}|Y) = \mathcal{N}(x_{t+1}; m_{t+1}^s, P_{t+1}^s)$. Note that

$$p(x_{t+1}, x_t|Y) = \underbrace{p(x_t|x_{t+1}, Y)}_{=p(x_t|x_{t+1}, y_{0:t})} \underbrace{p(x_{t+1}|Y)}_{=\mathcal{N}(x_{t+1}; m_{t+1}^s, P_{t+1}^s)}, \qquad (12)$$

where $p(x_t|x_{t+1}, Y) = p(x_t|x_{t+1}, y_{0:t})$ holds due to the Markov property. Now, in order to derive the missing term $p(x_t|x_{t+1}, y_{0:t})$, we compute the joint distribution by Lemma 1:

$$\begin{aligned}
p(x_t, x_{t+1}|y_{0:t}) &= p(x_{t+1}|x_t, y_{0:t})p(x_t|y_{0:t}) = p(x_{t+1}|x_t)\mathcal{N}(x_t; m_t, P_t) \\
&= \mathcal{N}(x_{t+1}; Ax_t, Q)\mathcal{N}(x_t; m_t, P_t) \\
&= \mathcal{N}\left(\begin{pmatrix} x_t \\ x_{t+1} \end{pmatrix}; \begin{pmatrix} m_t \\ Am_t \end{pmatrix}, \begin{pmatrix} P_t & P_tA^\intercal \\ AP_t & AP_tA^\intercal + Q \end{pmatrix}\right). \qquad (13)
\end{aligned}$$

Now, by the theorem from Slide 16 (Lecture 03), we have

$$\begin{aligned}
p(x_t|x_{t+1}, y_{0:t}) &= \mathcal{N}\left(x_t; m_t + P_tA^\intercal(AP_tA^\intercal + Q)^{-1}(x_{t+1} - Am_t), P_t - P_tA^\intercal(AP_tA^\intercal + Q)^{-1}AP_t\right) \\
&\overset{eq.\ (5)}{=} \mathcal{N}\left(x_t; m_t + P_tA^\intercal(P_{t+1}^-)^{-1}(x_{t+1} - Am_t), P_t - P_tA^\intercal(P_{t+1}^-)^{-1}AP_t\right) \\
&\overset{eq.\ (8)}{=} \mathcal{N}\left(x_t; m_t + G_t(x_{t+1} - Am_t), P_t - G_tP_{t+1}^-G_t^\intercal\right). \qquad (14)
\end{aligned}$$

Insertion of eq. (14) into eq. (12) yields

$$
\begin{aligned}
p(x_{t+1}, x_t|Y) &= \mathcal{N}\left(x_t; m_t + G_t(x_{t+1} - Am_t), P_t - G_t P^{-}_{t+1} G_t^\intercal\right) \mathcal{N}(x_{t+1}; m^s_{t+1}, P^s_{t+1}) \\
&= \mathcal{N}\left(\begin{pmatrix} x_{t+1} \\ x_t \end{pmatrix}; \begin{pmatrix} m^s_{t+1} \\ m_t + G_t(m^s_{t+1} - Am_t) \end{pmatrix}, \begin{pmatrix} P^s_{t+1} & P^s_{t+1} G_t^\intercal \\ G_t P^s_{t+1} & G_t P^s_{t+1} G_t^\intercal + P_t - G_t P^{-}_{t+1} G_t^\intercal \end{pmatrix}\right).
\end{aligned} \tag{15}
$$

By taking the marginal of $x_t$ (compare 'marginals of Gaussians are Gaussians', slide 17, lecture 03), we obtain eq. (6)

$$
p(x_t|Y) = \mathcal{N}\Big(x_t; m_t + G_t(m^s_{t+1} - Am_t), \underbrace{G_t P^s_{t+1} G_t^\intercal + P_t - G_t P^{-}_{t+1} G_t^\intercal}_{= P_t + G_t(P^s_{t+1} - P^{-}_{t+1}) G_t^\intercal}\Big), \tag{16}
$$

which concludes the whole proof. □

**Remark 1.** *If you had trouble with this exercise, it may be due to the fact that you did not write down the joint distributions eqs.* (10), (11), (13) *and* (15) *by Lemma 1.*

## 2 Exercise 2)

We are using the notation $a \wedge b := \min(a, b)$.

### 2.1 Exercise 2a)

Since $e^0 = I$, we can compute

$$
m(t) = e^0 x_0 = x_0, \tag{17}
$$

and

$$
k(t_a, t_b) = \int_{t_0}^{t_a \wedge t_b} e^{0(t_a - \tau)} \theta^2 e^{0^\intercal (t_b - \tau)} d\tau = \int_{t_0}^{t_a \wedge t_b} \theta^2 d\tau = \theta^2 (t_a \wedge t_b - t_a). \tag{18}
$$

### 2.2 Exercise 2b)

Define

$$
A := \begin{pmatrix} 0 & (t - t_0) \\ 0 & 0 \end{pmatrix}. \tag{19}
$$

Note that $A^2 = 0$. Hence,

$$
\begin{aligned}
\exp(F(t - t_0)) &= \exp(A) \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & (t - t_0) \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & (t - t_0) \\ 0 & 1 \end{pmatrix}.
\end{aligned} \tag{20}
$$

and

$$m(t) = \begin{pmatrix} 1 & t - t_0 \\ 0 & 1 \end{pmatrix} x_0. \tag{21}$$

Furthermore,

$$\begin{aligned}
k(t_a, t_b) &= \int_{t_0}^{t_a \wedge t_b} \exp\left( \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} (t_a - \tau) \right) \begin{pmatrix} 0 \\ \theta \end{pmatrix} \begin{pmatrix} 0 & \theta \end{pmatrix} \exp\left( \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} (t_b - \tau) \right) d\tau \\
&= \int_{t_0}^{t_a \wedge t_b} \begin{pmatrix} 1 & (t_a - \tau) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \theta \end{pmatrix} \begin{pmatrix} 0 & \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ (t_b - \tau) & 1 \end{pmatrix} d\tau \\
&= \theta^2 \int_{t_0}^{t_a \wedge t_b} \begin{pmatrix} t_a - t \\ 1 \end{pmatrix} \begin{pmatrix} t_b - t & 1 \end{pmatrix} d\tau \\
&= \theta^2 \int_{t_0}^{t_a \wedge t_b} \begin{pmatrix} (t_a - \tau)(t_b - \tau) & (t_a - \tau) \\ (t_b - \tau) & 1 \end{pmatrix} d\tau \\
&= \theta^2 \left[ \begin{pmatrix} \frac{\tau^3}{3} - \frac{(t_a + t_b)\tau^2}{2} + t_a t_b \tau & -\frac{\tau^2}{2} + t_a \tau \\ -\frac{\tau^2}{2} + t_b \tau & \tau \end{pmatrix} \right]_{\tau = t_0}^{\tau = t_a \wedge t_b} \\
&= \theta^2 \begin{pmatrix} \frac{(t_a \wedge t_b)^3 - t_0^3}{3} - \frac{(t_a + t_b)((t_a \wedge t_b)^2 - t_0^2)}{2} + t_a t_b (t_a \wedge t_b - t_0) & -\frac{(t_a \wedge t_b)^2 - t_0^2}{2} + t_a (t_a \wedge t_b - t_0) \\ -\frac{(t_a \wedge t_b)^2 - t_0^2}{2} + t_b (t_a \wedge t_b - t_0) & (t_a \wedge t_b - t_0) \end{pmatrix}
\end{aligned} \tag{22}$$

### 2.3 Exercise 2c)

Since $e^F(t - t_0) = e^{-\xi(t - t_0)}$, we can compute

$$m(t) = e^{-\xi(t - t_0)} x_0, \tag{23}$$

and

$$\begin{aligned}
k(t_a, t_b) &= \int_{t_0}^{t_a \wedge t_b} e^{-\xi(t_a - \tau)} \theta^2 e^{-\xi(t_b - \tau)} d\tau \\
&= \theta^2 \int_{t_0}^{t_a \wedge t_b} e^{-\xi(t_a + t_b - 2\tau)} d\tau \\
&= \frac{\theta^2}{2\xi} \left[ e^{-\xi(t_a + t_b - 2\tau)} \right]_{\tau = t_0}^{\tau = t_a \wedge t_b} \\
&= \frac{\theta^2}{2\xi} \left( e^{-\xi(t_a + t_b - 2(t_a \wedge t_b))} - e^{-\xi(t_a + t_b - 2t_0)} \right) \\
&= \frac{\theta^2}{2\xi} \left( e^{-\xi|t_a - t_b|} - e^{-\xi(t_a + t_b - 2t_0)} \right).
\end{aligned} \tag{24}$$

### 2.4 Exercise 2d)

$$F = \begin{pmatrix} 0 & 1 \\ -\xi^2 & -2\xi \end{pmatrix}; \qquad L = \begin{pmatrix} 0 \\ \theta \end{pmatrix} \tag{25}$$

We wish to decompose the matrix to be able to exponentiate it. To this end, we compute the eigenvalues of $F$.

$$\det\left(\begin{pmatrix}\lambda & -1\\ \xi^2 & \lambda+2\xi\end{pmatrix}\right) = (\lambda+\xi)^2 = 0, \tag{26}$$

if and only if $\lambda = -\xi$, which is the only eigenvalue (with algebraic multiplicity 2). To compute the eigenvectors, let $v_1 := (a, b)^\intercal$ and

$$(F - \lambda I_2)v_1 = \begin{pmatrix}\xi & 1\\ -\xi^2 & -\xi\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = 0, \tag{27}$$

if and only if $b = -\xi a$. Hence, $v_1 := (1, -\xi)^\intercal$. Thus, $F$ is not diagonalizable. Instead, we compute the Jordan normal decomposition of $F$ to bring it into a form $F = VJV^{-1}$ with

$$J = \begin{pmatrix}-\xi & 1\\ 0 & -\xi\end{pmatrix}, \tag{28}$$

which is the Jordan block for the eigenvalue of $F$. Let $N = J + \xi I_2$, which has eigenvalues 0 and is thus *nilpotent*. To construct the matrix $V$, we need to find the *generalized eigenvector* $v_2$ that satisfies $Nv_2 = v_1$ (and note that $Nv_1 = 0$). The second column of $N$ is the $v_1$ and thus we can choose $v_2 = \begin{pmatrix}0\\ 1\end{pmatrix}$. Hence

$$V = \begin{pmatrix}v_1 & v_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ -\xi & 1\end{pmatrix}, \quad \text{and} \quad V^{-1} = \begin{pmatrix}1 & 0\\ \xi & 1\end{pmatrix}. \tag{29}$$

Decomposing $N$ in terms of $V$ yields

$$F = V\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}V^{-1} \tag{30}$$

and $F$ decomposes into a diagonal matrix and a nilpotent one $N$ as

$$\begin{aligned} F = N - \xi I_2 &= -\xi I_2 V V^{-1} + V\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}V^{-1} \\ &= V\begin{pmatrix}-\xi & 1\\ 0 & -\xi\end{pmatrix}V^{-1} = VJV^{-1}. \end{aligned} \tag{31}$$

Hence we have brought $F$ into the Jordan normal form and decomposed $J$ into a diagonal matrix $D_J = -\xi I_2$ and a nilpotent matrix $N_J = \left(\begin{smallmatrix}0 & 1\\ 0 & 0\end{smallmatrix}\right)$.

Using this decomposition, we can write

$$e^{F(t-t_0)} = e^{VJV^{-1}(t-t_0)} = Ve^{J(t-t_0)}V^{-1}, \tag{32}$$

and, by $J = D_J + N_J$ due to and $N_J^2 = 0$, and using that $e^{D_J+N_J} = e^{D_J}e^{N_J}$ as $D_J$ and $N_J$ commute,

$$\begin{aligned} e^{J(t-t_0)} &= e^{\left(-\xi I_2+N_J\right)(t-t_0)} \\ &= e^{-\xi I_2(t-t_0)}e^{N_J(t-t_0)} \\ &= e^{-\xi(t-t_0)}I_2(I_2+N_J(t-t_0)) \\ &= e^{-\xi(t-t_0)}\begin{pmatrix}1 & t-t_0\\ 0 & 1\end{pmatrix} \end{aligned} \tag{33}$$

and analogously

$$e^{J(t_a-\tau)} = e^{-\xi(t_a-\tau)}\begin{pmatrix}1 & t_a-\tau\\ 0 & 1\end{pmatrix}, \qquad \text{and} \tag{34}$$

$$e^{J^\intercal(t_b-\tau)} = e^{-\xi(t_b-\tau)}\begin{pmatrix}1 & 0\\ t_b-\tau & 1\end{pmatrix}. \tag{35}$$

If we insert eqs. (33) to (35) into the provided formulas for $m(t)$ and $k(t_a, t_b)$, we obtain the results:

$$\begin{aligned} m(t) &= V\begin{pmatrix}e^{-\xi(t-t_0)} & (t-t_0)e^{-\xi(t-t_0)}\\ 0 & e^{-\xi(t-t_0)}\end{pmatrix}V^{-1}x_0, \\ &= e^{-\xi(t-t_0)}\begin{pmatrix}1 & 0\\ -\xi & 1\end{pmatrix}\begin{pmatrix}1 & t-t_0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \xi & 1\end{pmatrix}x_0 \\ &= e^{-\xi(t-t_0)}\begin{pmatrix}1+\xi(t-t_0) & t-t_0\\ -\xi^2(t-t_0) & 1-\xi(t-t_0)\end{pmatrix}x_0. \end{aligned} \tag{36}$$

and

$$\begin{aligned} k(t_a,t_b) &= \int_{t_0}^{t_a\wedge t_b} V\begin{pmatrix}e^{-\xi(t_a-\tau)} & (t_a-\tau)e^{-\xi(t_a-\tau)}\\ 0 & e^{-\xi(t_a-\tau)}\end{pmatrix}V^{-1}\begin{pmatrix}0 & 0\\ 0 & \theta^2\end{pmatrix}V\begin{pmatrix}e^{-\xi(t_b-\tau)} & 0\\ (t_b-\tau)e^{-\xi(t_b-\tau)} & e^{-\xi(t_b-\tau)}\end{pmatrix}V^{-1}d\tau \\ &= \theta^2 V\int_{t_0}^{t_a\wedge t_b} e^{-\xi(t_a+t_b-2\tau)}\begin{pmatrix}(t_a-\tau)(t_b-\tau) & t_a-\tau\\ t_b-\tau & 1\end{pmatrix}d\tau\, V^{-1} \end{aligned} \tag{37}$$

for which we need the following integrals (for which we only provide the ansatz to solve them analytically; you could also use Wolfram Alpha):

$$\int_{t_0}^{t_a\wedge t_b} e^{2\xi\tau}d\tau = \frac{1}{2\xi}(e^{2\xi(t_a\wedge t_b)} - e^{2\xi t_0}), \tag{38}$$

$$\begin{aligned} \int_{t_0}^{t_a\wedge t_b} \tau e^{2\xi\tau}d\tau &= \frac{1}{2}\frac{\partial}{\partial\xi}\int_{t_0}^{t_a\wedge t_b} e^{2\xi\tau}d\tau \\ &= \frac{1}{4\xi^2}\left(e^{2\xi t_0}(1-2\xi t_0) + e^{2\xi(t_a\wedge t_b)}(2\xi(t_a\wedge t_b)-1)\right) \end{aligned} \tag{39}$$

and

$$\begin{aligned}\int_{t_0}^{t_a \wedge t_b} \tau^2 e^{2\xi\tau} d\tau &= \frac{1}{4} \frac{\partial^2}{\partial \xi^2} \int_{t_0}^{t_a \wedge t_b} e^{2\xi\tau} d\tau \\ &= \frac{1}{4\xi^3} \left( e^{2\xi(t_a \wedge t_b)} (2\xi(t_a \wedge t_b)(\xi(t_a \wedge t_b) - 1) + 1) - e^{2\xi t_0} (2\xi t_0 (\xi t_0 - 1) + 1) \right) .\end{aligned} \tag{40}$$

What remains is to plug these integrals into eq. (37) and simplify. This is straightforward but quite tedious, so it is omitted here.