

## Exercise sheet 2

### Solutions

**Exercise 1.** Assuming that the two directions of motion are independent, the state equations for the different components of  $\boldsymbol{\theta}_k$  can be expressed as follows:

$$\begin{aligned}\mathbf{x}_k^{(i)} &= \mathbf{x}_{k-1}^{(i)} + \Delta \dot{\mathbf{x}}_{k-1}^{(i)} + \mathbf{u}_{k,i} \\ \dot{\mathbf{x}}_k^{(i)} &= \dot{\mathbf{x}}_{k-1}^{(i)} + \mathbf{u}_{k,i+2},\end{aligned}$$

for any  $i \in \{1, 2\}$ , which can be expressed more concisely in a matrix form (following the same approach for the noise) as  $\boldsymbol{\theta}_k = F\boldsymbol{\theta}_{k-1} + \mathbf{u}_k$  with  $\mathbf{u}_k \sim N(\cdot; 0, U)$  and with

$$F = \begin{pmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} \Delta^4/4 & 0 & \Delta^3/2 & 0 \\ 0 & \Delta^4/4 & 0 & \Delta^3/2 \\ \Delta^3/2 & 0 & \Delta^2 & 0 \\ 0 & \Delta^3/2 & 0 & \Delta^2 \end{pmatrix}.$$

Similarly, assuming that the position is observed, the observation equation is  $\mathbf{y}_k = H\boldsymbol{\theta}_k + \mathbf{v}_k$  with the observation matrix  $H$  defined as

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

and the noise can be assumed to be of the form  $\mathbf{v}_k \sim N(\cdot; 0, \sigma^2 I_2)$ .

**Exercise 2.** 1. The state and observation equations are  $\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \mathbf{u}_k$  and  $\mathbf{y}_k = \boldsymbol{\theta}_k + \mathbf{v}_k$  with  $(\mathbf{u}_k)_k$  and  $(\mathbf{v}_k)_k$  i.i.d. sequences of random variables distributed according to  $N(\cdot; 0, U)$  and  $N(\cdot; 0, V)$  respectively.

2. The prediction of the Kalman filter becomes  $m_k = \hat{m}_{k-1}$  and  $P_k = \hat{P}_{k-1} + U$ , the innovation, Kalman gain and covariance of the innovation become

$$z_k = y_k - m_k, \quad K_k = \frac{P_k}{S_k}, \quad \text{and} \quad S_k = P_k + V$$

so that the update step can be simplified to

$$\begin{aligned}\hat{m}_k &= m_k + K_k z_k = \frac{V}{P_k + V} m_k + \frac{P_k}{P_k + V} y_k \\ \hat{P}_k &= (1 - K_k) P_k = \frac{V P_k}{P_k + V}\end{aligned}$$

This is indeed what was found in Section 1.3.1 in the case of an unknown mean and a known variance.

3. The predictive distribution of  $\mathbf{y}_k$  given  $\mathbf{y}_{0:k-1} = y_{0:k-1}$  is normal (as a linear combination of normally-distributed random variables) with mean  $m_k$  and covariance  $P_k + V$ . The latter is the same as the covariance of the innovation, which is not surprising since the (random) innovation  $z_k$  is simply  $\mathbf{y}_k$  shifted by a constant ( $-m_k$ ), so they have the same covariance matrix.

**Exercise 3.** 1. From the previous exercise, we easily compute that

$$\begin{aligned}\hat{P}_k &= \frac{V P_k}{P_k + V} = \frac{V \hat{P}_{k-1} + V U}{\hat{P}_{k-1} + U + V} \\ K_k &= \frac{P_k}{P_k + V} = \frac{\hat{P}_{k-1} + U}{\hat{P}_{k-1} + U + V}\end{aligned}$$

Since it holds that  $\hat{P}_{k-1} = VK_{k-1}$ , it follows that

$$K_k = \frac{VK_{k-1} + U}{VK_{k-1} + U + V}.$$

These equations can be expressed more concisely as

$$\begin{aligned}\hat{P}_k^{-1} &= (\hat{P}_{k-1} + U)^{-1} + V^{-1} \\ K_k^{-1} &= \left(K_{k-1} + \frac{U}{V}\right)^{-1} + 1\end{aligned}$$

2. Denoting  $R$  the ratio  $U/V$ , it holds that

$$\begin{aligned}K_k - K_{k-1} &= K_k K_{k-1} (K_{k-1}^{-1} - K_k^{-1}) \\ &= K_k K_{k-1} ((K_{k-2} + R)^{-1} - (K_{k-1} + R)^{-1}) \\ &= \frac{K_k K_{k-1} (K_{k-1} - K_{k-2})}{(K_{k-1} + R)(K_{k-2} + R)}\end{aligned}$$

so that

$$\begin{aligned}\frac{K_k - K_{k-1}}{K_{k-1} - K_{k-2}} &= \frac{K_k K_{k-1}}{(K_{k-1} + R)(K_{k-2} + R)} \\ &= K_k K_{k-1} (K_k^{-1} - 1)(K_{k-1}^{-1} - 1) \\ &= (1 - K_k)(1 - K_{k-1}) \in (0, 1)\end{aligned}$$

since  $K_k \in (0, 1)$  for any  $k \geq 0$ . The lower bound implies that the sequence is monotonic and the upper bound implies that the increment becomes smaller with  $k$ , so that the sequence has a unique limit  $K$ .

**Exercise 4.** The distribution of  $\boldsymbol{\theta}_{k+1}$  given  $\mathbf{y}_{0:k-1} = y_{0:k-1}$  is also normal with mean

$$\mathbb{E}(\boldsymbol{\theta}_{k+1} | \mathbf{y}_{0:k-1} = y_{0:k-1}) = \mathbb{E}(\boldsymbol{\theta}_k | \mathbf{y}_{0:k-1} = y_{0:k-1}) = \hat{m}_{k-1}$$

and variance

$$\text{var}(\boldsymbol{\theta}_{k+1} | \mathbf{y}_{0:k-1} = y_{0:k-1}) = \text{var}(\boldsymbol{\theta}_k | \mathbf{y}_{0:k-1} = y_{0:k-1}) + U = \hat{P}_{k-1} + 2U$$

In these conditions, the Kalman gain at time step  $k+1$  without update at time  $k$  becomes

$$\tilde{K}_{k+1} = \frac{\hat{P}_{k-1} + 2U}{\hat{P}_{k-1} + 2U + V},$$

This alternative Kalman gain is greater than what would have been  $K_k$  if  $y_k$  had been available, so that the influence of  $y_{k+1}$  will be greater. This is to be expected since there is more uncertainty when observations are missing.

**Exercise 5.** To answer this question, it is convenient to rewrite the state  $\boldsymbol{\theta}_k$  as

$$\boldsymbol{\theta}_{k+i} = \boldsymbol{\theta}_k + \sum_{j=1}^i \mathbf{u}_{k+j}$$

for any  $i > 0$ , so that

$$\mathbf{y}_{k+i} = \boldsymbol{\theta}_{k+i} + \mathbf{v}_{k+i} = \boldsymbol{\theta}_k + \mathbf{v}_{k+i} + \sum_{j=1}^i \mathbf{u}_{k+j}$$

The aggregate future sales  $A_{k,\delta}$  can then be expressed as

$$\begin{aligned}
A_{k,\delta} &= \sum_{i=1}^{\delta} \mathbf{y}_{k+i} \\
&= \sum_{i=1}^{\delta} \left( \boldsymbol{\theta}_k + \mathbf{v}_{k+i} + \sum_{j=1}^i \mathbf{u}_{k+j} \right) \\
&= \delta \boldsymbol{\theta}_k + \sum_{i=1}^{\delta} \mathbf{v}_{k+i} + \sum_{i=1}^{\delta} \sum_{j=1}^i \mathbf{u}_{k+j} \\
&= \delta \boldsymbol{\theta}_k + \sum_{i=1}^{\delta} \mathbf{v}_{k+i} + \sum_{i=1}^{\delta} (\delta - i + 1) \mathbf{u}_{k+i}.
\end{aligned}$$

Since  $A_{k,\delta}$  is defined as a linear combination of normally-distributed random variables, it also follows a normal distribution with mean

$$\mathbb{E}(A_{k,\delta} | \mathbf{y}_{0:k} = y_{0:k}) = \delta \mathbb{E}(\boldsymbol{\theta}_k | \mathbf{y}_{0:k} = y_{0:k}) = \delta \hat{\boldsymbol{\mu}}_k$$

and variance

$$\begin{aligned}
\text{var}(A_{k,\delta} | \mathbf{y}_{0:k} = y_{0:k}) &= \delta^2 \text{var}(\boldsymbol{\theta}_k | \mathbf{y}_{0:k} = y_{0:k}) + \sum_{i=1}^{\delta} \text{var}(\mathbf{v}_{k+i}) + \sum_{i=1}^{\delta} (\delta - i + 1)^2 \text{var}(\mathbf{u}_{k+i}) \\
&= \delta^2 \hat{P}_k + \delta V + U \sum_{i=1}^{\delta} i^2 \\
&= \delta^2 \hat{P}_k + \delta V + \frac{1}{6} \delta (\delta + 1) (2\delta + 1) U.
\end{aligned}$$