Exercise sheet 3
Solutions

Exercise 1. The state and observation equations of the DLM with states $(\boldsymbol{\theta}'_k)_{k\geq 0}$ and observations $(\boldsymbol{y}_k)_{k\geq 0}$ can be expressed as

$$\begin{aligned} S_k^{-1}\boldsymbol{\theta}'_k &= F_k S_{k-1}^{-1}\boldsymbol{\theta}'_{k-1} + \boldsymbol{u}_k \\ \boldsymbol{y}_k &= H_k S_k^{-1}\boldsymbol{\theta}'_k + \boldsymbol{v}_k \end{aligned}$$

with $\boldsymbol{u}_k \sim \mathrm{N}(\cdot\,; 0, U_k)$ so that

$$\begin{aligned} \boldsymbol{\theta}'_k &= S_k F_k S_{k-1}^{-1}\boldsymbol{\theta}'_{k-1} + \boldsymbol{u}'_k \\ \boldsymbol{y}_k &= H_k S_k^{-1}\boldsymbol{\theta}'_k + \boldsymbol{v}_k, \end{aligned}$$

with $\boldsymbol{u}'_k \sim \mathrm{N}(\cdot\,; 0, S_k U_k S_k^{\intercal})$. It also holds that

$$\begin{aligned} \mathbb{E}(\boldsymbol{\theta}'_k \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= S_k \mathbb{E}(\boldsymbol{\theta}_k \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) = S_k \hat{m}_k \\ \mathrm{var}(\boldsymbol{\theta}'_k \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= S_k \,\mathrm{var}(\boldsymbol{\theta}_k \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) S_k^{\intercal} = S_k \hat{P}_k S_k^{\intercal}. \end{aligned}$$

Exercise 2. Before looking at the specific models, we can express $\boldsymbol{y}_{k+\delta}$ in a convenient form by first looking at $\boldsymbol{\theta}_{k+\delta}$ and finding that

$$\begin{aligned} \boldsymbol{\theta}_{k+\delta} &= F\boldsymbol{\theta}_{k+\delta-1} + \boldsymbol{u}_{k+\delta} \\ &= F(F\boldsymbol{\theta}_{k+\delta-2} + \boldsymbol{u}_{k+\delta-1}) + \boldsymbol{u}_{k+\delta} \\ &= F^{\delta}\boldsymbol{\theta}_k + \sum_{i=0}^{\delta-1} F^i \boldsymbol{u}_{k+\delta-i} \end{aligned}$$

so that

$$\boldsymbol{y}_{k+\delta} = H\boldsymbol{\theta}_{k+\delta} + \boldsymbol{v}_{k+\delta} = HF^{\delta}\boldsymbol{\theta}_k + \sum_{i=0}^{\delta-1} HF^i \boldsymbol{u}_{k+\delta-i} + \boldsymbol{v}_{k+\delta}.$$

The distribution of $\boldsymbol{y}_{k+\delta}$ given $\boldsymbol{y}_{0:k} = y_{0:k}$ is therefore normal as soon as the distribution of $\boldsymbol{\theta}_{k+\delta}$ given $\boldsymbol{y}_{0:k} = y_{0:k}$ is also normal (as a linear combination of normally distributed random variables). The result can now be used for both cases:

1. A second-order polynomial trend model has transition and observation matrix

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad H = e_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

so that

$$HF^{\delta} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \delta \end{pmatrix}$$

for any $\delta \geq 0$. Denoting $\boldsymbol{\theta}_k = \big(\boldsymbol{x}_k, \dot{\boldsymbol{x}}_k\big)^{\intercal}$, we deduce that

$$\boldsymbol{y}_{k+\delta} = \boldsymbol{x}_k + \delta\dot{\boldsymbol{x}}_k + \sum_{i=0}^{\delta-1} \big(\boldsymbol{\epsilon}_{k+\delta-i,1} + (1+i)\boldsymbol{\epsilon}_{k+\delta-i,2}\big) + \boldsymbol{v}_{k+\delta}.$$

The mean and variance of $\boldsymbol{y}_{k+\delta}$ given $\boldsymbol{y}_{0:k} = y_{0:k}$ can then be found as

$$\begin{aligned}
\mathbb{E}(\boldsymbol{y}_{k+\delta} \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= \hat{m}_{k,1} + \delta \hat{m}_{k,2} \\
\mathrm{var}(\boldsymbol{y}_{k+\delta} \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= \begin{pmatrix} 1 & \delta \end{pmatrix} \hat{P}_k \begin{pmatrix} 1 & \delta \end{pmatrix}^{\intercal} + \delta\sigma_1^2 + \sum_{i=0}^{\delta-1} (1+i)^2 \sigma_2^2 + V_{k+\delta} \\
&= \hat{P}_{k,(1,1)} + 2\delta \hat{P}_{k,(1,2)} + \delta^2 \hat{P}_{k,(2,2)} + \delta\sigma_1^2 + \frac{\delta(\delta+1)(2\delta+1)}{6}\sigma_2^2 + V_{k+\delta}
\end{aligned}$$

2. The superposition of a first-order polynomial trend model with a first-harmonic Fourier-form seasonal model has for transition and observation matrix

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & \sin(\omega) \\ 0 & -\sin(\omega) & \cos(\omega) \end{pmatrix} \qquad \text{and} \qquad H = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}.$$

Given that F is a block-diagonal matrix with the second block being a rotation matrix, it holds that

$$F^{\delta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\delta\omega) & \sin(\delta\omega) \\ 0 & -\sin(\delta\omega) & \cos(\delta\omega) \end{pmatrix},$$

so that

$$HF^{\delta} = \begin{pmatrix} 1 & \cos(\delta\omega) & \sin(\delta\omega) \end{pmatrix} \doteq A.$$

For similar reasons, it holds that $F^{\intercal} = F^{-1}$ and it follows from the form of U that $FUF^{\intercal} = U$. Therefore, the mean and variance of $\boldsymbol{y}_{k+\delta}$ given $\boldsymbol{y}_{0:k} = y_{0:k}$ can then be deduced to be

$$\begin{aligned}
\mathbb{E}(\boldsymbol{y}_{k+\delta} \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= \hat{m}_{k,1} + \cos(\delta\omega)\hat{m}_{k,2} + \sin(\delta\omega)\hat{m}_{k,3} \\
\mathrm{var}(\boldsymbol{y}_{k+\delta} \,|\, \boldsymbol{y}_{0:k} = y_{0:k}) &= A\hat{P}_k A^{\intercal} + \sum_{i=0}^{\delta-1} HF^i U (F^i)^{\intercal} H^{\intercal} + V_{k+\delta} \\
&= A\hat{P}_k A^{\intercal} + \delta(\sigma^2 + \sigma'^2) + V_{k+\delta}
\end{aligned}$$

Exercise 3. 1. This result can be verified easily by induction. It is obviously true for $\delta = 1$ so assuming it holds for a given δ, we want to prove it also holds for $\delta + 1$:

$$F^{\delta+1} = F^{\delta}F = \begin{pmatrix} \lambda a_1 & a_1 + \lambda a_2 & \dots & a_{d-1} + \lambda a_d \\ 0 & \lambda a_1 & \dots & a_{d-1} + \lambda a_{d-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda a_1 \end{pmatrix}.$$

Considering the $j+1$-th term on the first line, $j > 0$, we have

$$\begin{aligned}
a_j + \lambda a_{j+1} &= \binom{\delta}{j-1}\lambda^{\delta-j+1} + \lambda\binom{\delta}{j}\lambda^{\delta-j} \\
&= \left(\binom{\delta}{j-1} + \binom{\delta}{j}\right)\lambda^{(\delta+1)-j} \\
&= \binom{\delta+1}{j}\lambda^{(\delta+1)-j}
\end{aligned}$$

where the last line is due to Pascal's rule. This is the correct term for $\delta + 1$ which completes the proof by induction.

2. When $d = 3$, the forecast function g_0 is of the form

$$\begin{aligned}
g_0(\delta) &= \lambda^{\delta} m_{0,1} + \delta\lambda^{\delta-1} m_{0,2} + \frac{\delta(\delta-1)}{2}\lambda^{\delta-2} m_{0,3} \\
&= \lambda^{\delta}\Big(m_{0,1} + \delta\Big(\frac{m_{0,2}}{\lambda} - \frac{m_{0,3}}{2\lambda^2}\Big) + \delta^2 \frac{m_{0,3}}{2\lambda^2}\Big)
\end{aligned}$$

Since we assume that $g_0(\delta) = A\delta^2 \exp(-\rho\delta)$, we can identify the following relations:

$$\lambda = \exp(-\rho), \qquad A = \frac{m_{0,3}}{2\lambda^2}, \qquad m_{0,1} = 0 \qquad \text{and} \qquad \frac{m_{0,2}}{\lambda} = \frac{m_{0,3}}{2\lambda^2}$$

with the third and last equality following from the fact that the constant terms and the terms in λ must be equal to 0 (since there are no such terms in the assumed expression of g_0). To use the information regarding when the maximum is attained, we consider the derivative of $\log g_0$ as follows

$$\frac{\mathrm{d}}{\mathrm{d}\delta} \log g_0(\delta) = \frac{2}{\delta} - \rho = 0 \implies \delta = \frac{2}{\rho} = 4 \qquad \text{and} \qquad \rho = 1/2,$$

from which it also follows that $\lambda = \exp(-1/2)$. We also know that the maximum value is $30,000$ so that

$$g_0(4) = 4^2 A \exp(-2) = 30,000 \implies A = 1875 \exp(2) \approx 13,854.$$

The values of $m_{0,2}$ and $m_{0,3}$ can also be found to be

$$m_{0,3} = 2\lambda^2 A \approx 10,194$$
$$m_{0,2} = \frac{m_{0,3}}{2\lambda} \approx 8,403.$$

Since λ equals to $\exp(-1/2) \in (-1, 1)$, it is true that the sales will be modelled as tapering off.