10 "Answers" to the exerises

Exercises

1. Prove Theorem 2.1 using elementary probability (for simplicity, do not bother to condition everything on D).

   Answer: This is first-year book-work!

$$\begin{aligned}\mathsf{E}[x] &= \sum_i p(x_i)\, x_i \\ &= \sum_i \left(\sum_j p(x_i \mid z_j)\, p(z_j) \right) x_i \\ &= \sum_j p(z_j) \sum_i x_i\, p(x_i \mid z_j) \\ &= \mathsf{E}[\mathsf{E}[x_i \mid z]]\end{aligned}$$

   as required.

2. Prove $\mathsf{Cov}[x,\, y \mid D] = \mathsf{Cov}[x,\, \mathsf{E}[y \mid x, D] \mid D]$.

   Answer: We introduce x itself. In this case $\mathsf{Cov}[x,\, y \mid x, D] = 0$ and its expectation is also 0, so we are left with $\mathsf{Cov}[x,\, y \mid D] = \mathsf{Cov}[\mathsf{E}[x \mid x, D]\,,\, \mathsf{E}[y \mid x, D] \mid D] = \mathsf{Cov}[x,\, \mathsf{E}[y \mid x, D] \mid D]$, as required.

3. (a) Show that (equation numbers relate to hand-out sheet 2)

$$\begin{aligned}\mathsf{E}[\theta_{t+k} \mid D_t] &= a_t(k) := G_{t+k}\, a_t(k-1) \\ \mathsf{Var}\,[\theta_{t+k} \mid D_t] &= R_t(k) := G_{t+k}\, R_t(k-1)\, G_{t+k}^{\mathsf{T}} + W_{t+k}\end{aligned} \tag{4}$$

   subject to the initial values $a_t(0) := m_t$ and $R_t(0) := C_t$, and that, for any $0 \le j < k$,

$$\mathsf{Cov}[\theta_{t+j},\, \theta_{t+k} \mid D_t] = C_t(j,k) := C_t(j, k-1)\, G_{t+k}^{\mathsf{T}} \tag{5}$$

   subject to the initial values $C_t(j,j) = R_t(j)$.

   Answer: We prove these by induction. Equations (4) are obviously true for $k = 0$. Assume they are true for $k-1$. At time k we have, for the mean,

$$\begin{aligned}a_t(k) &:= \mathsf{E}[\theta_{t+k} \mid D_t] \\ &= \mathsf{E}[\mathsf{E}[\theta_{t+k} \mid \theta_{t+k-1}, D_t] \mid D_t] \\ &= \mathsf{E}[\mathsf{E}[\theta_{t+k} \mid \theta_{t+k-1}] \mid D_t] \\ &= \mathsf{E}[G_{t+k}\, \theta_{t+k-1} \mid D_t] \\ &= G_{t+k}\, a_t(k-1)\end{aligned}$$

   using the graph of the DLM in the usual fashion. The variance is similar.

   We can do the covariance a bit differently. We have, introducing θ_{t+j},

$$\mathsf{Cov}[\theta_{t+j},\, \theta_{t+k} \mid D_t] = \mathsf{Cov}[\theta_{t+j},\, \mathsf{E}[\theta_{t+k} \mid \theta_{t+j}] \mid D_t]\,.$$

   Then it is easy to see that, as above,

$$\begin{aligned}\mathsf{E}[\theta_{t+k} \mid \theta_{t+j}] &= \mathsf{E}[\mathsf{E}[\theta_{t+k} \mid \theta_{t+k-1}] \mid \theta_{t+j}] \\ &= G_{t+k}\, \mathsf{E}[\theta_{t+k-1} \mid \theta_{t+j}] \\ &= \cdots \\ &= G_{t+k}\, G_{t+k-1} \cdots G_{t+j+1}\, \theta_{t+j}\,.\end{aligned}$$

   Substsituting this into the covariance expression gives

$$\begin{aligned}\mathsf{Cov}[\theta_{t+j},\, \theta_{t+k} \mid D_t] &= \mathsf{Cov}[\theta_{t+j},\, \theta_{t+j} \mid D_t]\, (G_{t+k}\, G_{t+k-1} \cdots G_{t+j+1})^{\mathsf{T}} \\ &= R_t(j)\, G_{t+j+1}^{\mathsf{T}} G_{t+j+2}^{\mathsf{T}} \cdots G_{t+k}^{\mathsf{T}}\end{aligned}$$

   which is the explicit form of the recursive expression above.

   (b) Give explicit (i.e. non-recursive) expressions for (4) and (5) in the case where $\{F_{t+k}, V_{t+k}, G_{t+k}, W_{t+k}\} = \{F, V, G, W\}$ for all $k \geq 1$.

Answer: By back-substitution we must have

$$a_t(k) = G^k m_t$$
$$C_t(k) = \sum_{i=0}^{k-1} (G^i) W (G^\top)^i + G^k C_t (G^\top)^k$$

4. Consider the 'random walk with noise' DLM,

$$y_t = \theta_t + \nu_t \qquad \nu_t \sim \langle 0,\, V\rangle$$
$$\theta_t = \theta_{t-1} + \omega_t \quad \omega_t \sim \langle 0,\, W\rangle$$

where θ_t is a scalar, and both V and W are time-invariant. Starting from $\theta_{t-1} \sim \langle m_{t-1},\, C_{t-1}\rangle$, compute the mean and variance for $\theta_t \mid D_t$.

Answer: In this DLM we have $F = G = 1$. So we find that

$$a_t = \mathsf{E}[\theta_t \mid D_{t-1}] = \mathsf{E}[\mathsf{E}[\theta_t \mid \theta_{t-1}] \mid D_{t-1}] = \mathsf{E}[\theta_{t-1} \mid D_{t-1}] = m_{t-1}$$
$$R_t = \mathsf{Var}\,[\theta_t \mid D_{t-1}] = \cdots = C_{t-1} + W.$$

Then it follows (from $F = 1$) that $y_t \mid D_{t-1} \sim \langle a_t,\, R_t + V\rangle$ and $\mathsf{Cov}[\theta_t,\, y_t \mid D_{t-1}] = R_t$. This gives

$$\begin{aligned}
\mathsf{E}[\theta_t \mid D_t] &= m_{t-1} + \frac{C_{t-1} + W}{C_{t-1} + W + V}(y_t - m_{t-1}) \\
&= w_{t-1}\, y_t + (1 - w_{t-1})\, m_{t-1} \\
\mathsf{Var}\,[\theta_t \mid D_t] &= C_{t-1} + W - \frac{(C_{t-1} + W)^2}{C_{t-1} + W + V} \\
&= w_{t-1}\, V
\end{aligned}$$

where $w_{t-1} := (C_{t-1} + W)/(C_{t-1} + W + V)$ and we must have $0 < w_{t-1} < 1$ if $V > 0$ and $W > 0$.

5. In the generic DLM we can have known *non-zero* 'intercepts' h_t and g_t in the observation and state equations respectively,

$$y_t = h_t + F_t^\top \theta_t + \nu_t \qquad \nu_t \sim \langle 0,\, V_t\rangle$$
$$\theta_t = g_t + G_t\, \theta_{t-1} + \omega_t \quad \omega_t \sim \langle \mathbf{0},\, W_t\rangle$$

How does the presence of these two intercepts affect the updating of our beliefs about θ_t by the datum y_t?

Answer: It simply changes the values of $\mathsf{E}[\theta_t \mid D_{t-1}]$ and $\mathsf{E}[y_t \mid D_{t-1}]$. Thus $\mathsf{E}[\theta_t \mid D_t]$ will be different but $\mathsf{Var}\,[\theta_t \mid D_t]$ will be unchanged.

6. Consider the case of an object in 2-dimensional Euclidean space falling freely (but not necessarily vertically) under gravity. Denote the location of this object by (x_t, y_t), and denote the general state vector as θ_t.

   (a) What is an appropriate state equation for this object? How do you interpret the elements of the state vector?

   Answer: It's falling freely under gravity, so the only force it experiences is a constant downward force. Neglecting friction, this suggests bolting two plynomial growth models together:

$$\theta_t = \begin{pmatrix} L_2 & \mathbf{0} \\ \mathbf{0} & L_3 \end{pmatrix} \theta_{t-1} + \omega_t \qquad \omega_t \sim \langle \mathbf{0},\, W\rangle$$

   where $\theta_t = (x_t, \dot{x}_t, y_t, \dot{y}_t, \ddot{y}_t)$ In other words, x 'velocity' ($\dot{x}_t$) is 'constant' (i.e. not systematically increasing or decreasing), and y 'acceleration' ($\ddot{y}_t$) is 'constant'.

   (b) What factors might you consider in determining the error variance for your state equation?

Answer: Remember with the simple polynomial growth model of order n that

$$W = L_n D (L_n)^\intercal$$

where D is a diagonal variance matrix. If we bolt two together, we get

$$W = \begin{pmatrix} L_2 D_x (L_2)^\intercal & \mathbf{0} \\ \mathbf{0} & L_3 D_y (L_3)^\intercal \end{pmatrix}$$

where D_x and D_y are diagonal variance matrices for the horizontal and vertical displacement terms, respectively. So we need one variance scalar per term of θ_t.

Friction affects the vertical displacement (and the horizontal displacement if $\dot{x}_t$ is large), in conjunction with the object's shape and rotation. Winds affect the horizontal displacement. I might be tempted to put most of my uncertainty into the evolution of the $\dot{x}_t$ and $\ddot{y}_t$ terms, with only a bit in the others to account for the truncation error that must be present because we are working with discrete time steps Δt rather than dt.

(c) Every five seconds I record the angle α_t between the ground and the object from my location $(\bar{x}, 0)$. What is the observation equation?

Answer: By simple trigonometry we have $\tan \alpha_t = y_t/(x_t - \bar{x})$, or

$$\alpha_t = \arctan\big(y_t/(x_t - \bar{x})\big) + \nu_t \qquad \nu_t \sim \langle 0,\, V\rangle$$

where V is the variance of the measurement error.

(d) Given $\theta_t \mid D_{t-1} \sim \langle a_t,\, R_t\rangle$, how do I compute the mean and variance of $\alpha_t \mid D_{t-1}$, and the covariance $\mathsf{Cov}[\theta_t,\, \alpha_t \mid D_{t-1}]$?

Answer: I linearise the observation equation around a_t, the prior mean for θ_t. In general, if $\alpha_t = f(\theta_t) + \nu_t$, then

$$\alpha_t \approx f(a_t) + \nabla f(a_t)(\theta_t - a_t) + \nu_t$$

(treating $\nabla f(a_t)$ as a row vector). So, after conditioning on θ_t,

$$\begin{aligned} \mathsf{E}[\alpha_t \mid D_{t-1}] &= f(a_t) \\ \mathsf{Var}\,[\alpha_t \mid D_{t-1}] &= (F_t)^\intercal R_t F_t + V \\ \mathsf{Cov}[\theta_t,\, \alpha_t \mid D_{t-1}] &= R_t F_t \end{aligned}$$

where $(F_t)^\intercal = \nabla f(a_t)$.

In this particular case we have

$$f(\theta_t) = \arctan\big(y_t/(x_t - \bar{x})\big) = \arctan\big(\theta_{t3}/(\theta_{t1} - \bar{x})\big).$$

This gives us $f(a_t)$ directly, leaving only F_t to be found. Remembering that $\frac{d}{dx}\arctan x = 1/(1+x^2)$, we have

$$F_t = \frac{1}{1 + \big(a_{t3}/(a_{t1} - \bar{x})\big)^2} \begin{pmatrix} -a_{t3}/(a_{t1} - \bar{x})^2 \\ 0 \\ 1/(a_{t1} - \bar{x}) \\ 0 \\ 0 \end{pmatrix}$$

which completes the calculation.