3 Assimilating new data

Theorem 3.1 (Bayes linear adjustment). *Let x, y and z be random quantities over which we have a mean and covariance structure initially adjusted by D. The mean of x and the covariance of x and y, further adjusted by z, are*

$$\begin{aligned}\mathsf{E}[x \mid z, D] &= \mathsf{E}[x \mid D] + \mathsf{Cov}[x,\, z \mid D]\,\mathsf{Var}\,[z \mid D]^{-1}\,(z - \mathsf{E}[z \mid D])\\ \mathsf{Cov}[x,\, y \mid z, D] &= \mathsf{Cov}[x,\, y \mid D] - \mathsf{Cov}[x,\, z \mid D]\,\mathsf{Var}\,[z \mid D]^{-1}\,\mathsf{Cov}[z,\, y \mid D]\end{aligned}$$

where '$|$' in this context is to be interpreted as 'adjusted by' (rather than 'conditional upon').

When assimilating a new datum, our objective is to go from beliefs about $\theta_{t-1} \mid D_{t-1}$ to beliefs about $\theta_t \mid D_t$. Remembering that $D_t = (D_{t-1}, y_t)$, the Bayes linear approach requires the mean and covariance structure of $\theta_t, y_t \mid D_{t-1}$. The adjustment process moves the y_t from the left of the bar to the right.

Suppose we currently have beliefs $\theta_{t-1} \mid D_{t-1} \sim \langle m_{t-1},\, C_{t-1}\rangle$. From our model we can infer that $\theta_t \mid D_{t-1} \sim \langle a_t,\, R_t\rangle$, where

$$\begin{aligned}a_t &:= \mathsf{E}[\theta_t \mid D_{t-1}] = G_t\, m_{t-1}\\ R_t &:= \mathsf{Var}\,[\theta_t \mid D_{t-1}] = G_t\, C_{t-1}\, G_t^{\mathsf{T}} + W_t.\end{aligned}$$

In both cases we have introduced θ_{t-1} and used the property $\theta_t \perp\!\!\!\perp D_{t-1} \mid \theta_{t-1}$ from the graph of the DLM.

We also need to compute our beliefs about $y_t \mid D_{t-1}$. As θ_t separates y_t from D_{t-1} in the graph (i.e. $y_t \perp\!\!\!\perp D_{t-1} \mid \theta_t$), it is easy to compute that

$$\mathsf{E}[y_t \mid D_{t-1}] = F_t^{\mathsf{T}} a_t \quad \text{and} \quad \mathsf{Var}\,[y_t \mid D_{t-1}] = F_t^{\mathsf{T}} R_t\, F_t + V_t.$$

Finally, we need the covariance $\mathsf{Cov}[\theta_t,\, y_t \mid D_{t-1}]$. By conditioning on θ_t and using $y_t \perp\!\!\!\perp D_{t-1} \mid \theta_t$ again we find

$$\begin{aligned}\mathsf{Cov}[\theta_t,\, y_t \mid D_{t-1}] &= \mathsf{Cov}[\theta_t,\, \mathsf{E}[y_t \mid \theta_t] \mid D_{t-1}]\\ &= \mathsf{Cov}[\theta_t,\, F_t^{\mathsf{T}}\theta_t \mid D_{t-1}] = R_t\, F_t.\end{aligned}$$

Putting these bits into the Bayes linear expressions,

$$\begin{aligned}\mathsf{E}[\theta_t \mid D_t] &= a_t + (R_t\, F_t)\, Q_t^{-1}(y_t - F_t^{\mathsf{T}} a_t)\\ \mathsf{Var}\,[\theta_t \mid D_t] &= R_t - (R_t\, F_t)\, Q_t^{-1}\,(R_t\, F_t)^{\mathsf{T}}\end{aligned}$$

where, for convenience, I have written $Q_t := \mathsf{Var}\,[y_t \mid D_{t-1}]$.

What if no datum is recorded? It is often the case that y_t records only the fact that the datum for time t is missing. But as $y_t = \emptyset$, so

$$\theta_t \mid D_t = \theta_t \mid (D_{t-1}, \emptyset) = \theta_t \mid D_{t-1} \sim \langle a_t,\, R_t\rangle$$

and so we have simply $m_t = a_t$ and $C_t = R_t$. Note that time-points that have missing data have larger variances, as you might expect. This also gives us a way to handle irregular data: we treat it as lots of missing data on a suitably fine scale.

Exercises

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

$$\begin{aligned}y_t &= \theta_t + \nu_t & \nu_t &\sim \langle 0,\, V\rangle\\ \theta_t &= \theta_{t-1} + \omega_t & \omega_t &\sim \langle 0,\, W\rangle\end{aligned}$$

   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 $\theta_t \mid D_t$.

2. In the generic DLM we can have known *non-zero* 'intercepts' h_t and

g_t in the observation and state equations respectively,

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

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