Spatio – Temporal Methods in Environmental Epidemiology: Supplementary Material for Chapter 10

Gavin Shaddick and James V Zidek

Sampling from the full conditional of the overall mean parameter, β

The full condition for parameter β can be written

$$\begin{aligned} p(\beta|y_{\cdot},\theta_{\cdot},\sigma_u^2) &\propto \left\{\prod_{t=1}^{T} N(\beta+\theta_t,\sigma_u^2)\right\} \times N(\mu_\beta,\sigma_\beta^2) \\ &\propto \exp\left\{-\frac{\sum_{t=1}^{T}(y_t-(\beta+\theta_t))^2}{2\sigma_u^2}\right\} \times \exp\left\{-\frac{(\beta-\mu_\beta)^2}{2\sigma_\beta^2}\right\} \\ &\propto N(m_\beta,s_\beta) \text{ where,} \end{aligned}$$

$$m_\beta = \frac{\frac{1}{T}\sum_{t=1}^{T}(y_t-\theta_t)\sigma_\beta^2 + \mu_\beta\sigma_u^2/T}{\sigma_u^2/T+\sigma_\beta^2} \text{ and } s_\beta = \frac{\sigma_\beta^2\sigma_u^2/T}{\sigma_u^2/T+\sigma_\beta^2}$$

In the absence of prior knowledge on β, a vague prior could be arranged taking the form $N(0,\sigma_\beta^2)$ with σ_β^2 very large. Therefore the above simplifies to

$$m_\beta = \frac{1}{T}\sum_{t=1}^{T}(y_t-\theta_t) \text{ and } s_\beta = \frac{\sigma_u^2}{T}$$

0.0.0.1 Sampling from the full conditional of the precision of the random error, σ_u^{-2}

A Gamma prior distribution is chosen for the parameter σ_u^{-2} with relatively small mean and variance. The parameterisation of the gamma distribution used in this thesis is

$$Gam(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp^{\beta-1} \text{ for } x \geq 0 \tag{1}$$

$$\begin{aligned} p(\sigma_u^{-2}|y_{\cdot},\theta_{\cdot},\beta) &\propto \left\{\prod_{t=1}^{T} N(\beta+\theta_t,\sigma_u^2)\right\} \times Gam(a_u,b_u) \\ &\propto \exp\left\{-\frac{\sum_{t=1}^{T}(y_t-(\beta+\theta_t))^2}{2\sigma_u^2}\right\} \times \sigma_u^{-2(a_u-1)}\exp\{-b_u\sigma_u^{-2}\} \\ &\propto Gam\left(\frac{T}{2}+a_u, b_u+\frac{1}{2}\sum_{t=1}^{T}(y_t-(\beta+\theta_t))^2\right) \end{aligned}$$

0.0.0.2 Sampling from the full conditional of the precision of the temporal variation, σ_w^{-2}

Sampling from $p(\sigma_w^{-2}|\theta_{\cdot},\alpha)$ works in a similar way to that of sampling from the posterior of σ_u^{-2}. A Gamma distribution can be chosen with relatively small mean and variance. The full conditional is as follows.

$$\begin{aligned} p(\sigma_w^{-2}|\theta_{\cdot},\alpha) &\propto \left\{\prod_{t=2}^{T} N(\alpha\theta_{t-1},\sigma_w^2)\right\} \times Gam(a_w,b_w) \\ &\propto \exp\left\{-\frac{\sum_{t=2}^{T}(\theta_t-\alpha\theta_{t-1})^2}{2\sigma_w^2}\right\} \times \sigma_u^{-2(a_w-1)}\exp\{-b_w\sigma_w^{-2}\} \\ &\propto Gam\left(\frac{T-1}{2}+a_w, b_w+\frac{1}{2}\sum_{t=2}^{T}(\theta_t-\alpha\theta_{t-1}))^2\right) \end{aligned}$$

0.0.0.3 Sampling from the full conditional of the temporal effects, $\theta_.$

The full conditional for each θ_t, $t = 1, ..., T$, is a normal distribution

$$
\begin{aligned}
p(\theta_t | y_., \beta, \sigma_u^2, \sigma_w^2, \alpha) \quad & \propto \left\{ \prod_{t=1}^{T} N(\beta + \theta_t, \sigma_u^2) \right\} \times \left\{ \prod_{t=2}^{T} N(\alpha\theta_{t-1}, \sigma_w^2) \right\} \\
& \propto \exp\left\{ -\frac{\sum_{t=1}^{T} (y_t - (\beta + \theta_t))^2}{2\sigma_u^2} \right\} \times \exp\left\{ -\frac{\sum_{t=2}^{T-1} (\theta_t - \alpha(\frac{\theta_{t-1}+\theta_{t+1}}{1+\alpha^2}))^2}{2\sigma_w^2/(1+\alpha^2)} \right\} \\
& \propto N(\mu_{\theta_t}, s_{\theta_t})
\end{aligned}
$$

where

$$
s_{\theta_t} = \frac{1}{\frac{1}{\sigma_u^2} + \frac{\alpha^2}{\sigma_w^2}} \tag{2}
$$

$$
\frac{1}{\frac{1}{\sigma_u^2} + \frac{1+\alpha^2}{\sigma_w^2}} \qquad \text{and} \tag{3}
$$

$$
\frac{1}{\frac{1}{\sigma_u^2} + \frac{1}{\sigma_w^2}} \tag{4}
$$

$$
\mu_{\theta_t} = \left(\frac{y_t - \beta}{\sigma_u^2} + \frac{\alpha\theta_{t+1}}{\sigma_w^2} \right) s_{\theta_t} \quad \text{for } t = 1, \tag{5}
$$

$$
\left(\frac{y_t - \beta}{\sigma_u^2} + \frac{\alpha(\theta_{t-1} + \theta_{t+1})}{\sigma_w^2} \right) s_{\theta_t} \quad \text{for } t = 2, ..., T-1, \tag{6}
$$

$$
\left(\frac{y_t - \beta}{\sigma_u^2} + \frac{\alpha\theta_{t-1}}{\sigma_w^2} \right) s_{\theta_t} \quad \text{for } t = T, \tag{7}
$$

0.0.0.4 Sampling from the full condition of the α parameter

In the AR(1) case, the α parameter, having a Uniform distribution as a prior, will not have a full conditional of closed form. In such cases, a Metropolis-Hasting step can be used with proposed value of a accepted with probability

$$
c = min\left[1, \frac{p(\theta_. | \alpha', \sigma_w)}{p(\theta_. | \alpha^c, \sigma_w)} \right]
$$

where α' is the proposed value and α^c the current value of α. For the proposed value, α', a random value generated from a $Unif(-1, 1)$, is used.

0.0.1 Implemention using WinBUGS

The methods for temporal estimation and prediction can all be fit using MCMC and here we discuss their implementation using WinBUGS. Following on from the material in Chapter **??** and the implmentation of the air pollution example in Chapter **??** we show how a random walk model can be embedded in a measurement–process model formulation. For clarity, the material presented within this chapter on classical time series methodology was presented as though there was no measurement error present, i.e. $Y_t = Z_t$ in the formulation seen in Equation **??** but here we revert to the full measurement–process.

- **Stage One, Observed Data Model:**

  $$Y_t = \beta_1 X_t + m_t + v_t,$$

  where v_t are i.i.d. $N(0, \sigma_v^2)$ and β_1 is the effect of covariate X_1.

- **Stage Two, Temporal Model:**

  $$m_t = \rho m_{t-1} + w_t$$

  with w_t i.i.d. as $N(0, \sigma_w^2)$ and, in this case, we will use $\rho = 1$, i.e. a first order random walk.

Recall that from a Bayesian perspective, the second (temporal) stage may be viewed as a prior distribution for $m' = (m_1, ..., m_T)$, and that $p(m|\sigma_w^2)$. In this case, the duality between a random walk and CAR prior (as seen in Chapter **??**) can be expressed as

$$p(m_t|m_{-t}, \sigma_w^2) \sim \begin{cases} N(m_{t+1}, \sigma_w^2) & \text{for } t = 1, \\ N\left(\frac{m_{t-1}+m_{t+1}}{2}, \frac{\sigma_w^2}{2}\right) & \text{for } t = 2, ..., T-1, \\ N(m_{t-1}, \sigma_w^2) & \text{for } t = T. \end{cases}$$

where m_{-t} represents the vector of m's with m_t removed. It is noted that σ_w^2 is a *conditional* variance and so it is not comparable to σ_v^2.

References