Web-based Supplementary Materials for "A Bayesian Model for Sparse Functional Data" by W.K. Thompson and O. Rosen

December 3, 2006

Web Appendix

Here, we give a detailed account of the distributions and sampling scheme outlined in sections 2 and 3. Let $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_n')'$ be the vector of responses for all n subjects. Let $m = \sum_{i=1}^n m_i$ be the total number of observations across all subjects. With the prior distributions specified in Section 2.2, we have

$$
\begin{aligned}
p(\boldsymbol{\beta}_\gamma, \boldsymbol{b}_\gamma, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{\gamma}, \pi \mid \boldsymbol{y}) \quad &\propto \quad p(\boldsymbol{y} \mid \boldsymbol{b}_\gamma, \sigma_\epsilon^2, \boldsymbol{\gamma}) p(\boldsymbol{b}_\gamma \mid \boldsymbol{\beta}_\gamma, \Sigma_{b,\gamma}, \boldsymbol{\gamma}) p(\boldsymbol{\beta}_\gamma \mid \Sigma_{b,\gamma}, \boldsymbol{\gamma}) \\
&\quad \times p(\sigma_\epsilon^2) p(\Sigma_{b,\gamma} \mid \boldsymbol{\gamma}) p(\boldsymbol{\gamma} \mid \pi) p(\pi) \\
&\propto \quad \sigma_\epsilon^{-m} \exp\{-\frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^n (\boldsymbol{y}_i - X_{\gamma,i}\boldsymbol{b}_{\gamma,i})'(\boldsymbol{y}_i - X_{\gamma,i}\boldsymbol{b}_{\gamma,i})\} \\
&\quad \times \mid \Sigma_{b,\gamma} \mid^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^n (\boldsymbol{b}_{\gamma,i} - \boldsymbol{\beta}_\gamma)' \Sigma_{b,\gamma}^{-1} (\boldsymbol{b}_{\gamma,i} - \boldsymbol{\beta}_\gamma)\} \\
&\quad \times c^{-q_\gamma/2} \exp\{-\frac{1}{2c} \boldsymbol{\beta}_\gamma' \boldsymbol{\beta}_\gamma\} \; (\sigma_\epsilon^2)^{-(c_\epsilon+1)} \exp\{-d_\epsilon/\sigma_\epsilon^2\} \\
&\quad \times \mid \Sigma_{b,\gamma} \mid^{-(\eta_b+q_\gamma+1)/2} \exp\{-\frac{\eta_b}{2} \mathrm{trace}(S_\gamma \Sigma_{b,\gamma}^{-1})\} \\
&\quad \times \pi^{q_\gamma}(1-\pi)^{L-q_\gamma} \pi^{c_\pi - 1}(1-\pi)^{d_\pi - 1}.
\end{aligned}
$$

Step 1: Sample s distinct values $\boldsymbol{l} = \{l_1, \ldots, l_s\}$ (for predetermined $1 \leq s \leq L$) from $\{1, \ldots, L\}$ without replacement such that each such vector is equally probable.

Step 2: Sample $\left\{\gamma_l, \boldsymbol{\beta}_\gamma, \boldsymbol{b}_\gamma\right\}$ conditional on $\{\boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}\}$. This is done in three substeps, using the factorization

$$
\begin{aligned}
p(\boldsymbol{\gamma}_l, \boldsymbol{\beta}_\gamma, \boldsymbol{b}_\gamma \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}) \quad &\propto \quad p(\boldsymbol{\gamma}_l \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y})\, p(\boldsymbol{\beta}_\gamma \mid \boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}) \\
&\quad \times p(\boldsymbol{b}_\gamma \mid \boldsymbol{\gamma}, \boldsymbol{\beta}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}).
\end{aligned}
$$

Step 2(a): Draw γ_l; the posterior conditional probability that $\boldsymbol{\gamma}_l = \boldsymbol{u}$ (where $\boldsymbol{u}$ ranges over all zero-one vectors of dimension s) is given by

$$
\begin{aligned}
p(\boldsymbol{\gamma}_l = \boldsymbol{u} \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}) \quad &\propto \quad p(\boldsymbol{y} \mid \boldsymbol{l}, \boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}) \\
&\quad \times p(\boldsymbol{\gamma}_l = \boldsymbol{u} \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}), \qquad \text{(A.1)}
\end{aligned}
$$

where

$$
\begin{aligned}
p(\boldsymbol{y} \mid \boldsymbol{l}, \boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_b) \quad &= \quad \int_{\boldsymbol{b}_\gamma} p(\boldsymbol{y} \mid \boldsymbol{l}, \boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}, \boldsymbol{b}, \sigma_\epsilon^2, \Sigma_{b,\gamma}) \\
&\quad \times \{ \int_{\boldsymbol{\beta}_\gamma} p(\boldsymbol{b}_\gamma \mid \boldsymbol{l}, \boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}, \boldsymbol{\beta}_\gamma, \sigma_\epsilon^2, \Sigma_{b,\gamma}) \\
&\quad \times p(\boldsymbol{\beta}_\gamma \mid \boldsymbol{l}, \boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}) \mathrm{d}\boldsymbol{\beta}_\gamma \} \mathrm{d}\boldsymbol{b}_\gamma. \qquad \text{(A.2)}
\end{aligned}
$$

Let $q_{\gamma(u)}$ be the dimension of the model selected by $\boldsymbol{\gamma}(\boldsymbol{u}) = \{\boldsymbol{\gamma}_l = \boldsymbol{u}, \boldsymbol{\gamma}_{(l)}\}$. Denote the right side of (A.2) by L_u. Then it can be shown that L_u is equal to

$$
\begin{aligned}
L_u \quad &= \quad c^{-q_{\gamma(u)}/2} |\Sigma_{b,\gamma(u)}|^{-n/2} |n\Sigma_{b,\gamma(u)}^{-1} + \frac{1}{c} I_{\gamma(u)}|^{-1/2} |A_{\gamma(u)}|^{-1/2} \\
&\quad \times \exp\left\{ \frac{1}{2\sigma_\epsilon^4} \boldsymbol{y}' (X_{\gamma(u)}^D) A_{\gamma(u)}^{-1} (X_{\gamma(u)}^D)' \boldsymbol{y} \right\}, \qquad \text{(A.3)}
\end{aligned}
$$

where $\Sigma_{b,\gamma(u)}$ is the $q_{\gamma(u)} \times q_{\gamma(u)}$ covariance matrix corresponding to the breakpoints selected by $\boldsymbol{\gamma}(\boldsymbol{u})$, $X_{\gamma(u)}^D = \mathrm{diag}\{X_{\gamma(u),i}\}$ is an $m \times nq_{\gamma(u)}$ block-diagonal matrix and

$$
\begin{aligned}
A_{\gamma(u)} \quad &= \quad \mathrm{diag}\left(\frac{1}{\sigma_\epsilon^2} X'_{\gamma(u),i} X_{\gamma(u),i} + \Sigma_{b,\gamma(u)}^{-1} \right) - \\
&\quad (\boldsymbol{1} \otimes I_{q_{\gamma(u)}}) \left(\Sigma_{b,\gamma(u)}^{-1} \left(n\Sigma_{b,\gamma(u)}^{-1} + \frac{1}{c} I_{q_{\gamma(u)}} \right)^{-1} \Sigma_{b,\gamma(u)}^{-1} \right) (\boldsymbol{1} \otimes I_{q_{\gamma(u)}})',
\end{aligned}
$$

where $\mathbf{1}$ is the n-dimensional vector of all ones, $I_{q_{\gamma(u)}}$ is the $q_{\gamma(u)} \times q_{\gamma(u)}$ identity matrix, and $\otimes$ denotes the Kronecker product. While calculation of L_u involves taking the inverse and determinant of the $nq_{\gamma(u)} \times nq_{\gamma(u)}$ matrix $A_{\gamma(u)}$, because of the special form of this matrix, these can be simplified through the use of Woodbury's formula (Lange, 1999, pp. 86-87), so that only the determinant and inverse of $q_{\gamma(u)} \times q_{\gamma(u)}$ matrices are necessary.

Let $\Theta_u = p(\gamma_l = \boldsymbol{u} \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(bfl)}, \sigma_\epsilon^2, \Sigma_{b,\gamma})$. Then

$$\begin{aligned} \Theta_u &= \int_\pi p(\boldsymbol{\gamma}_l = \boldsymbol{u} \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \pi) p(\pi) \mathrm{d}\pi \\ &= \mathrm{B}(c_\pi + q_{\gamma(u)}, d_\pi + L - q_{\gamma(u)}) / \mathrm{B}(c_\pi + q_{\gamma(0)}, d_\pi + L - s - q_{\gamma(0)}), \end{aligned} \tag{A.4}$$

where $\mathrm{B}(\cdot, \cdot)$ is the beta function, L is the number of interior breakpoints, and $q_{\gamma(0)}$ is the dimension of the model selected by $\boldsymbol{\gamma}(\mathbf{0}) = \{\boldsymbol{\gamma}_l = \mathbf{0}, \boldsymbol{\gamma}_{(l)}\}$. Using L_u and Θ_u as defined in (A.3) and (A.4), we have that (A.1) is

$$p(\boldsymbol{\gamma}_l = \boldsymbol{u} \mid \boldsymbol{l}, \boldsymbol{\gamma}_{(l)}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}) = \frac{\Theta_u L_u}{\sum_{u'} \Theta_{u'} L_{u'}},$$

where the sum in the denominator of the right side is taken over all zero-one vectors of dimension s.

Step 2(b): Draw $\boldsymbol{\beta}_\gamma$ conditional on $\{\boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}\}$. The posterior conditional is

$$\begin{aligned} p(\boldsymbol{\beta}_\gamma \mid \boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma}, \boldsymbol{y}) \;\propto\; & \int_{\boldsymbol{b}_\gamma} \left\{ p(\boldsymbol{y} \mid \boldsymbol{b}_\gamma, \boldsymbol{\beta}_\gamma, \boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma})\, p(\boldsymbol{b}_\gamma \mid \boldsymbol{\beta}_\gamma, \boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma}) \right\} \mathrm{d}\boldsymbol{b}_\gamma \\ & \times p(\boldsymbol{\beta}_\gamma \mid \boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_{b,\gamma}). \end{aligned}$$

After integration, we see that $\boldsymbol{\beta}_\gamma \mid \{\boldsymbol{\gamma}, \sigma_\epsilon^2, \Sigma_b, \boldsymbol{y}\} \sim \mathrm{MVN}(\boldsymbol{\mu}_{\boldsymbol{\beta}|\cdot}\,, \Sigma_{\boldsymbol{\beta}|\cdot})$, where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\beta}|\cdot} &= \Sigma_{\boldsymbol{\beta}|\cdot} \sum_{i=1}^n \Sigma_{b,\gamma}^{-1} \left(\frac{1}{\sigma_\epsilon^2} X'_{\gamma,i} X_{\gamma,i} + \Sigma_{b,\gamma}^{-1} \right)^{-1} \frac{1}{\sigma_\epsilon^2} X'_{\gamma,i} \boldsymbol{y}_i \quad \text{and} \\ \Sigma_{\boldsymbol{\beta}|\cdot} &= \left\{ n\Sigma_{b,\gamma}^{-1} + \frac{1}{c} I_\gamma - \sum_{i=1}^n \Sigma_{b,\gamma}^{-1} \left(\frac{1}{\sigma_\epsilon^2} X'_{\gamma,i} X_{\gamma,i} + \Sigma_{b,\gamma}^{-1} \right)^{-1} \Sigma_{b,\gamma}^{-1} \right\}^{-1}. \end{aligned} \tag{A.5}$$

Step 2(c): Draw the $\{\boldsymbol{b}_{\gamma,i}\}_{i=1}^{n}$ conditional on $\{\boldsymbol{\gamma}, \boldsymbol{\beta}_{\gamma}, \sigma_{\epsilon}^{2}, \Sigma_{b,\gamma}, \boldsymbol{y}\}$. The conditional posterior distributions of the $\{\boldsymbol{b}_{\gamma,i}\}_{i=1}^{n}$ are independent MVN$((\boldsymbol{\mu}_{\boldsymbol{b}_{i}|\cdot}\ , \Sigma_{\boldsymbol{b}_{i}|\cdot})$, where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{b}_{i}|\cdot} &= \Sigma_{\boldsymbol{b}_{i}|\cdot}\left(\frac{1}{\sigma_{\epsilon}^{2}}X_{\gamma,i}'\boldsymbol{y}_{i} + \Sigma_{b,\gamma}^{-1}\boldsymbol{\beta}_{\gamma}\right) \quad \text{and} \\ \Sigma_{\boldsymbol{b}_{i}|\cdot} &= \left(\frac{1}{\sigma_{\epsilon}^{2}}X_{\gamma,i}'X_{\gamma,i} + \Sigma_{b,\gamma}^{-1}\right)^{-1}. \end{aligned} \tag{A.6}$$

Step 3: Draw σ_{ϵ}^{2} conditional on $\{\boldsymbol{\gamma}, \boldsymbol{\beta}_{\gamma}, \boldsymbol{b}_{\gamma}, \Sigma_{b,\gamma}, \boldsymbol{y}\}$. The posterior conditonal of σ_{ϵ}^{2} is inverse gamma, IG$(c_{\epsilon|\cdot}, d_{\epsilon|\cdot})$, with

$$\begin{aligned} c_{\epsilon|\cdot} &= c_{\epsilon} + \frac{m}{2}, \quad \text{and} \\ d_{\epsilon|\cdot} &= d_{\epsilon} + \frac{1}{2}\sum_{I=1}^{n}(\boldsymbol{y}_{i} - X_{\gamma,i}\boldsymbol{b}_{\gamma,i})'(\boldsymbol{y}_{i} - X_{\gamma,i}\boldsymbol{b}_{\gamma,i}). \end{aligned} \tag{A.7}$$

The parameters c_{ϵ} and d_{ϵ} may be chosen to obtain a vague prior on σ_{ϵ}^{2}.

Step 4: Draw $\Sigma_{b,\gamma}$ conditional on $\{\boldsymbol{\gamma}, \boldsymbol{\beta}_{\gamma}, \boldsymbol{b}_{\gamma}, \sigma_{\epsilon}^{2}, \boldsymbol{y}\}$. The posterior conditional distribution of $\Sigma_{b,\gamma}$ is inverse Wishart, IW$(\eta_{b|\cdot}\ S_{b|\cdot}, \eta_{b|\cdot})$, where

$$\begin{aligned} \eta_{b|\cdot} &= \eta_{b} + n, \quad \text{and} \\ S_{b|\cdot} &= \frac{1}{\eta_{b|\cdot}}\left(\eta_{b}S_{\gamma} + \sum_{i=1}^{n}\left(\boldsymbol{b}_{\gamma,i} - \boldsymbol{\beta}_{\gamma}\right)\left(\boldsymbol{b}_{\gamma,i} - \boldsymbol{\beta}_{\gamma}\right)'\right). \end{aligned} \tag{A.8}$$

Let $\boldsymbol{t}$ be an ordered vector of τ time values containing all the unique time points in the data. Let X_{τ} denote the $\tau \times K$ design matrix of B-splines evaluated at all the points of $\boldsymbol{t}$ and let $X_{\tau,\boldsymbol{\gamma}}$ be the $\tau \times q_{\gamma}$ design matrix corresponding to the breakpoints selected in $\boldsymbol{\gamma}$. Furthermore, define $\Phi_{\gamma} = (X_{\tau}'X_{\tau})^{-1}X_{\tau}'X_{\tau,\boldsymbol{\gamma}}$. Since models created by removing breakpoints are linear subspaces of the full model, we have that $X_{\tau}\Phi_{\gamma} = X_{\tau,\boldsymbol{\gamma}}$ and hence we also have $X_{\tau,\boldsymbol{\gamma}}\boldsymbol{\beta}_{\gamma} = X_{\tau}\Phi_{\gamma}\boldsymbol{\beta}_{\gamma}$. Thus, going from the full model to the model given by $\boldsymbol{\gamma}$ implies the linear restrictions on $\boldsymbol{\beta}$ (and similarly for the $\boldsymbol{b}_{i}$'s) given by $\boldsymbol{\beta} = \Phi_{\gamma}\boldsymbol{\beta}_{\gamma}$ and hence

$\boldsymbol{\beta}_\gamma = \left(\Phi'_\gamma\Phi_\gamma\right)^{-1}\Phi'_\gamma\boldsymbol{\beta} = \Omega'_\gamma\boldsymbol{\beta}$, where $\Omega_\gamma = \Phi_\gamma(\Phi'_\gamma\Phi_\gamma)^{-1}$. This implies that the covariance matrix $\Sigma_{b,\gamma}$ can be obtained from Σ_b by using the relationship $\Sigma_{b,\gamma} = \Omega'_\gamma\Sigma_b\Omega_\gamma$. Additionally, let $P_{\Omega_\gamma} = \Omega_\gamma(\Omega'_\gamma\Omega_\gamma)^{-1}\Omega'_\gamma$. Then,

$$\begin{aligned}\Sigma_b &= P_{\Omega_\gamma}\Sigma_b P_{\Omega_\gamma} + (I - P_{\Omega_\gamma})\Sigma_b(I - P_{\Omega_\gamma}) \\ &= \Omega_\gamma(\Omega'_\gamma\Omega_\gamma)^{-1}\Sigma_{b,\gamma}(\Omega'_\gamma\Omega_\gamma)^{-1}\Omega'_\gamma + (I - P_{\Omega_\gamma})\Sigma_b(I - P_{\Omega_\gamma}). \qquad \text{(A.9)}\end{aligned}$$

At iteration $r+1$ of the MCMC algorithm, let $\Sigma_b^{(r)}$ be the value of Σ_b generated in the previous iteration. Step 4 consists of drawing $\Sigma_{b,\gamma}^{(r+1)}$ according to the $\text{IW}(S_{b|\cdot}, \eta_{b|\cdot})$ as detailed above and then using (A.9), with $\Sigma_{b,\gamma}^{(r+1)}$ in place of $\Sigma_{b,\gamma}$ and $\Sigma_b^{(r)}$ in place of Σ_b on the right side of this equation to obtain the new draw of the full covariance matrix $\Sigma_b^{(r+1)}$.