Haul-out Site Estimation using a Truncation Approximation of a Dirichlet Process Mixture Model

Brian M. Brost

13 August 2015

Implementation

The file haulouts.1.sim.R simulates data according to the model statement presented below, and haulouts.1.mcmc.R contains the MCMC algorithm for model fitting. Dirichlet process mixture model implementation follows the blocked Gibbs sampler truncation approximation of Ishwaran and James (2001) and Gelman et al. (2014).

Description

A Dirichlet process mixture model for haul-out site estimation. Unlike the model implemented in haulout.dp.mixture.2.mcmc.R, which has the same specification, $\tilde{\mathcal{S}}$ is represented as a raster in this version. Consequently, $\tilde{\mathcal{S}}$ can be linear in nature like the shoreline support of actual harbor seal haul-out sites. As such, the update for $\boldsymbol{\mu}_{0,h}$ proceeds using Metropolis-Hastings, which is useful for generating proposals $\boldsymbol{\mu}^*_{0,h}$ automatically on their support $\tilde{\mathcal{S}}$.

Model statement

Let $\mathbf{s}(t) = (s_1(t), s_2(t))'$ and $\boldsymbol{\mu}(t) = (\mu_1(t), \mu_2(t))'$, be observed and true locations of a single individual at some time t, respectively. Also let $\boldsymbol{\mu}_{0,h} = (\mu_{0,1,h}, \mu_{0,2,h})'$, for $h = 1, \ldots, H$, be the locations of haul-out sites (i.e., cluster centroids), where H is the maximum number of haul-outs allowed per the truncation approximation of the Dirichlet process mixture model. The latent indicator variable $z(t)$ denotes when location $\mathbf{s}(t)$ is on a haul-out site ($z(t) = 1$) or not ($z(t) = 0$). When an individual is hauled-out, note that $\boldsymbol{\mu}(t) = \boldsymbol{\mu}_{0,h_t}$, where h_t acts as a classification variable that identifies the $\boldsymbol{\mu}_{0,h}$ associated with each $\boldsymbol{\mu}(t)$. Furthermore, let $\mathcal{S}$ be the support of the movement process and $\tilde{S}$ be the support of the haul-out sites (i.e., the Dirichlet process and all possible $\boldsymbol{\mu}_{0,h}$). Note that $\mathcal{S}$ and $\tilde{S}$ overlap, i.e., $\tilde{\mathcal{S}} \subset \mathcal{S}$. The domain defined by $\mathcal{S}$ therefore represents at-sea locations or locations of the individual while milling adjacent to the haul-out site. Also note that $\bar{\mathcal{S}}$, the complement of $\mathcal{S}$, represents inaccessible locations (i.e., terrestrial sites that are not haul-outs).

Information pertaining to the wet/dry status of the individual is available from a second data source, $y(\tilde{t})$ (e.g., records from a *.SEA file). Note that records in y are observed at times $\tilde{t}$, which may not be the same as the times t at which locations are collected. The binary data in $y\left(\tilde{t}\right)$ are modeled using an auxiliary variable $\mathbf{v}$, which itself is modeled semiparametrically as a function of $\boldsymbol{\beta}$, the 'fixed' effects that provide inference on covariates of biological interest in the matrix $\mathbf{X}$, and the 'random' effects $\boldsymbol{\alpha}$ that describe non-linear trend or dependence in wet/dry status through the basis expansion $\mathbf{W}$. The auxilliary variable $\mathbf{v}$ describes a continuous (in time) haul-out process, and links the wet/dry status of the individual with its telemetry location.

$$
\begin{aligned}
\mathbf{s}(t) &\sim \begin{cases} \mathcal{N}(\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}), & z(t)=1 \\ \mathcal{N}(\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}+\sigma_\mu^2\mathbf{I}), & z(t)=0 \end{cases} \\
z(t) &\sim \begin{cases} 0, & v(t)\le 0 \\ 1, & v(t)>0 \end{cases} \\
y\left(\tilde{t}\right) &\sim \begin{cases} 0, & v\left(\tilde{t}\right)\le 0 \\ 1, & v\left(\tilde{t}\right)>0 \end{cases} \\
\mathbf{v} &\sim \mathcal{N}\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{W}\boldsymbol{\alpha}, \mathbf{I}\right) \\
h_t &\sim \text{Cat}\left(\frac{\pi_h}{\sum_{\tilde{h}=1}^{H}\pi_{\tilde{h}}}\right) \\
\pi_h &\sim \text{Stick}(\theta) \\
\boldsymbol{\mu}_{0,h} &\sim \text{Unif}(\tilde{\mathcal{S}}) \\
\theta &\sim \text{Gamma}(r_\theta, q_\theta) \\
\boldsymbol{\beta} &\sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2\mathbf{I}) \\
\boldsymbol{\alpha} &\sim \mathcal{N}(\mathbf{0}, \sigma_\alpha^2\mathbf{I}) \\
\sigma &\sim \text{Unif}(l_\sigma, u_\sigma) \\
\sigma_\mu &\sim \text{Unif}(l_{\sigma_\mu}, u_{\sigma_\mu}) \\
\sigma_\alpha^2 &\sim \text{IG}(r_\sigma, q_\sigma)
\end{aligned}
$$

The concentration parameter θ affects the clustering in the Dirichlet process mixture: smaller values yield fewer clusters with more observations per cluster, whereas larger values yield more clusters with fewer observations per cluster. Note that the lines in this model statement pertaining to h_t, π_h, and $\boldsymbol{\mu}_{0,h}$ comprise the stick-breaking representation of the Dirichlet process mixture model, i.e.,

$$
\begin{aligned}
\boldsymbol{\mu}_{0,h} &\sim \mathbf{G} \\
\mathbf{G} &\sim \text{DP}(\theta, \mathbf{G}_0) \\
\mathbf{G}_0 &\sim \text{Unif}(\tilde{\mathcal{S}})
\end{aligned}
$$

Full conditional distributions

Haul-out site locations ($\boldsymbol{\mu}_{0,h}$):

$$
\begin{aligned}
[\boldsymbol{\mu}_{0,h}|\cdot] &\propto \prod_{\{t:h_t=h\}} [\mathbf{s}(t)|\boldsymbol{\mu}_{0,h},\sigma^2]^{z(t)}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h},\sigma^2,\sigma^2_\mu]^{1-z(t)}[\boldsymbol{\mu}_{0,h} \mid \tilde{\mathcal{S}}] \\
&\propto \prod_{\{t:h_t=h\}} \mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h},\sigma^2\mathbf{I})^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h},\sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})^{1-z(t)}1_{\{\boldsymbol{\mu}_{0,h}\in\tilde{\mathcal{S}}\}}
\end{aligned}
$$

The update for $\boldsymbol{\mu}_{0,h}$ proceeds using Metropolis-Hastings and proposals are sampled from $\tilde{\mathcal{S}}$ with probability proportional to the proposal distribution, e.g., $\mathcal{N}(\boldsymbol{\mu}^\star_{0,h}|\boldsymbol{\mu}_{0,h},\sigma^2_{tune}\mathbf{I})$. Note that the product is over all $\mathbf{s}(t)$ that are members of $\boldsymbol{\mu}_{0,h}$. For each of the unoccupied haul-out locations (i.e., clusters with no members), $\boldsymbol{\mu}_{0,h}$ is sampled directly from the prior.

Probability mass for haul-out location $\boldsymbol{\mu}_{0,h}$ (π_h):

The stick-breaking representation of a Dirichlet process mixture consists of two components, namely a cluster weight and a cluster probability. Let η_h denote the weight assigned to cluster h, where $\eta_h \sim \text{Beta}(1,\theta)$. The associated full-conditional is

$$
[\eta_h|\cdot] \sim \text{Beta}\left(1+n_h, \theta + \sum_{\tilde{h}=h+1}^{H} n_{\tilde{h}}\right), \text{ for } h=1,\ldots,H-1,
$$

and $\eta_H = 1$. The parameter n_h denotes the number of observations allocated to cluster h. Note that η_h is sampled in order of decreasing n_h, i.e., n_h is sorted largest to smallest and η_h is sampled in sequence. The cluster probabilities (π_h) are deterministic and calculated as

$$
\pi_h = \eta_h \prod_{\tilde{h}<h}(1-\eta_{\tilde{h}}).
$$

The probabilities π_h are also calculated in order of decreasing n_h. See page 553 in Gelman et al. (2014) and Section 5.2 in Ishwaran and James (2001).

Dirichlet process concentration parameter (θ):

$$
[\theta|\cdot] \propto \text{Gamma}(r_\theta + H - 1, q_\theta - \sum_{h=1}^{H-1}\log(1-v_h)).
$$

See page 553 in Gelman et al. (2014). Also see Escobar and West (1995) and West (1997?, white paper) for alternative full-conditionals for θ.

Haul-out classification variable (h_t):

$$
\begin{aligned}
[h_t|\cdot] &\sim [\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}}, z(t), \sigma^2, \sigma^2_\mu][h_t \mid \pi_h] \\
&\sim \text{Cat}\left(\frac{\pi_h[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, z(t), \sigma^2, \sigma^2_\mu]}{\sum_{\tilde{h}=1}^{H}\pi_{\tilde{h}}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}}, z(t), \sigma^2, \sigma^2_\mu]}\right) \\
&\sim \text{Cat}\left(\frac{\pi_h[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t},\sigma^2]^{z(t)}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t},\sigma^2,\sigma^2_\mu]^{1-z(t)}}{\sum_{\tilde{h}=1}^{H}\pi_{\tilde{h}}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}},\sigma^2]^{z(t)}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}},\sigma^2,\sigma^2_\mu]^{1-z(t)}}\right) \\
&\sim \text{Cat}\left(\frac{\pi_h\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t},\sigma^2)^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t},\sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})^{1-z(t)}}{\sum_{\tilde{h}=1}^{H}\pi_{\tilde{h}}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}},\sigma^2)^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,\tilde{h}},\sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})^{1-z(t)}}\right).
\end{aligned}
$$

This update proceeds just as in multinomial sampling; see page 552 in Gelman et al. (2014).

Latent haul-out indicator variable for telemetry locations ($z(t)$):

$$
\begin{aligned}
[z(t)|\cdot] &\propto [\mathbf{s}(t)|\boldsymbol{\mu}_{0,h}, z(t), \sigma^2, \sigma^2_\mu][z(t) \mid v(t)] \\
&\propto [\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2]^{z(t)}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma^2_\mu]^{1-z(t)}[z(t)|v(t)] \\
&\propto \mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I})^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})^{1-z(t)}\text{Bern}(z(t)|v(t)) \\
&\propto \mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I})^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})^{1-z(t)}p(t)^{z(t)}(1-p(t))^{1-z(t)} \\
&\propto \left(p(t)\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I})\right)^{z(t)}\left((1-p(t))[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma^2_\mu]\right)^{1-z(t)} \\
&= \text{Bern}(\widetilde{p(t)}),
\end{aligned}
$$

where $v(t) = \mathbf{x}'(t)\boldsymbol{\beta} + \mathbf{w}'(t)\boldsymbol{\alpha}$, $p(t) = \Phi(v(t))$, and

$$
\widetilde{p(t)} = \frac{p(t)\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I})}{p(t)\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}) + (1-p(t))\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I}+\sigma^2_\mu\mathbf{I})}.
$$

Auxiliary variable for continuous haul-out process ($\mathbf{v}$):

$$
\begin{aligned}
[\mathbf{v}|\cdot] &\propto \prod_{\{t,\tilde{t}\}} [z(t), y(\tilde{t}) \mid \mathbf{v}][\mathbf{v} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}] \\
&\propto \prod_t \{[z(t) \mid v(t)][v(t) \mid \boldsymbol{\alpha}, \boldsymbol{\beta}]\} \prod_{\tilde{t}} \left\{\left[y(\tilde{t})|v(\tilde{t})\right]\left[v(\tilde{t}) \mid \boldsymbol{\alpha}, \boldsymbol{\beta}\right]\right\} \\
&\propto \prod_t \left\{\left(1_{\{z(t)=0\}}1_{\{v(t)\le 0\}} + 1_{\{z(t)=1\}}1_{\{v(t)>0\}}\right) \times \mathcal{N}\left(v(t) \mid \mathbf{x}(t)'\boldsymbol{\beta} + \mathbf{w}(t)'\boldsymbol{\alpha}, \mathbf{1}\right)\right\} \times \\
&\qquad \prod_{\tilde{t}} \left\{\left(1_{\{y(\tilde{t})=0\}}1_{\{v(\tilde{t})\le 0\}} + 1_{\{y(\tilde{t})=1\}}1_{\{v(\tilde{t})>0\}}\right) \times \mathcal{N}\left(v(\tilde{t}) \mid \mathbf{x}(\tilde{t})'\boldsymbol{\beta} + \mathbf{w}(\tilde{t})'\boldsymbol{\alpha}, \mathbf{1}\right)\right\}
\end{aligned}
$$

For wet/dry data (y) observed at times $\tilde{t}$,

$$
\begin{aligned}
\left[v(\tilde{t})|\cdot\right] &\propto \left[y(\tilde{t})|v(\tilde{t})\right]\left[v(\tilde{t}) \mid \boldsymbol{\alpha}, \boldsymbol{\beta}\right] \\
&\propto \left(1_{\{y(\tilde{t})=0\}}1_{\{v(\tilde{t})\le 0\}} + 1_{\{y(\tilde{t})=1\}}1_{\{v(\tilde{t})>0\}}\right) \times \mathcal{N}\left(v(\tilde{t}) \mid \mathbf{x}(\tilde{t})'\boldsymbol{\beta} + \mathbf{w}(\tilde{t})'\boldsymbol{\alpha}, \mathbf{1}\right) \\
&= \begin{cases} \mathcal{TN}\left(\mathbf{x}(\tilde{t})'\boldsymbol{\beta} + \mathbf{w}(\tilde{t})'\boldsymbol{\alpha}, \mathbf{1}\right)_{-\infty}^{0}, & y(\tilde{t}) = 0 \\ \mathcal{TN}\left(\mathbf{x}(\tilde{t})'\boldsymbol{\beta} + \mathbf{w}(\tilde{t})'\boldsymbol{\alpha}, \mathbf{1}\right)_{0}^{\infty}, & y(\tilde{t}) = 1 \end{cases}
\end{aligned}
$$

For the latent wet/dry status (z) of telemetry locations at times t,

$$
\begin{aligned}
[v(t)|\cdot] &\propto [z(t) \mid v(t)][v(t) \mid \boldsymbol{\alpha}, \boldsymbol{\beta}] \\
&\propto \left(1_{\{z(t)=0\}}1_{\{v(t)\le 0\}} + 1_{\{z(t)=1\}}1_{\{v(t)>0\}}\right) \times \mathcal{N}\left(v(t) \mid \mathbf{x}(t)'\boldsymbol{\beta} + \mathbf{w}(t)'\boldsymbol{\alpha}, \mathbf{1}\right) \\
&= \begin{cases} \mathcal{TN}\left(\mathbf{x}(t)'\boldsymbol{\beta} + \mathbf{w}(t)'\boldsymbol{\alpha}, \mathbf{1}\right)_{-\infty}^{0}, & z(t) = 0 \\ \mathcal{TN}\left(\mathbf{x}(t)'\boldsymbol{\beta} + \mathbf{w}(t)'\boldsymbol{\alpha}, \mathbf{1}\right)_{0}^{\infty}, & z(t) = 1 \end{cases}
\end{aligned}
$$

Haul-out probability coefficients ($\boldsymbol{\beta}$):

$$\begin{aligned}
[\boldsymbol{\beta}|\cdot] &\propto [\mathbf{v}|\boldsymbol{\beta},\boldsymbol{\alpha}][\boldsymbol{\beta} \mid \boldsymbol{\mu}_\beta, \sigma^2_\beta] \\
&\propto \mathcal{N}\left(\mathbf{v} \mid \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha}, \mathbf{1}\right)\mathcal{N}(\boldsymbol{\mu}_\beta \mid \mathbf{0}, \sigma^2_\beta \mathbf{I}) \\
&\propto \exp\left\{-\frac{1}{2}\left(\mathbf{v} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha})\right)'\left(\mathbf{v} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha})\right)\right\} \times \\
&\qquad \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)'\left(\sigma^2_\beta \mathbf{I}\right)^{-1}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left((\mathbf{v} - \mathbf{W}\boldsymbol{\alpha}) - \mathbf{X}\boldsymbol{\beta}\right)'\left((\mathbf{v} - \mathbf{W}\boldsymbol{\alpha}) - \mathbf{X}\boldsymbol{\beta}\right)\right\} \times \\
&\qquad \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)'\left(\sigma^2_\beta \mathbf{I}\right)^{-1}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(-2\left(\left(\mathbf{v} - \mathbf{W}\boldsymbol{\alpha}\right)\mathbf{X} + \boldsymbol{\mu}_\beta\left(\sigma^2_\beta \mathbf{I}\right)^{-1}\right)\boldsymbol{\beta} + \boldsymbol{\beta}'\left(\mathbf{X}'\mathbf{X} + \left(\sigma^2_\beta \mathbf{I}\right)^{-1}\right)\boldsymbol{\beta}\right)\right\} \\
&= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{X}'\mathbf{X} + \left(\sigma^2_\beta \mathbf{I}\right)^{-1}$ and $\mathbf{b}' = \left(\mathbf{v} - \mathbf{W}\boldsymbol{\alpha}\right)'\mathbf{X} + \boldsymbol{\mu}'_\beta\left(\sigma^2_\beta \mathbf{I}\right)^{-1}$. Note that the matrices $\mathbf{X}$ and $\mathbf{W}$ contain covariates pertaining to times associated with both $\mathbf{s}(t)$ and $y\left(\tilde{t}\right)$.

Random 'effects' for non-linear trend/dependence ($\boldsymbol{\alpha}$):

$$\begin{aligned}
[\boldsymbol{\alpha}|\cdot] &\propto [\mathbf{v}|\boldsymbol{\beta},\boldsymbol{\alpha}]\left[\boldsymbol{\alpha} \mid \mathbf{0}, \sigma^2_\alpha\right] \\
&\propto \mathcal{N}\left(\mathbf{v} \mid \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha}, \mathbf{1}\right)\mathcal{N}(\boldsymbol{\alpha} \mid \mathbf{0}, \sigma^{-1}_\alpha \mathbf{I}) \\
&\propto \exp\left\{-\frac{1}{2}\left(\mathbf{v} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha})\right)'\left(\mathbf{v} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\alpha})\right)\right\} \times \\
&\qquad \exp\left\{-\frac{1}{2}\left(\boldsymbol{\alpha} - \mathbf{0}\right)'\left(\sigma^2_\alpha \mathbf{I}\right)^{-1}\left(\boldsymbol{\alpha} - \mathbf{0}\right)\right\} \\
&= \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}),
\end{aligned}$$

where $\mathbf{A} = \mathbf{W}'\mathbf{W} + \left(\sigma^2_\alpha \mathbf{I}\right)^{-1}$ and $\mathbf{b}' = \left(\mathbf{v} - \mathbf{X}\boldsymbol{\beta}\right)'\mathbf{W}$. Note that the matrices $\mathbf{X}$ and $\mathbf{W}$ contain covariates pertaining to times associated with both $\mathbf{s}(t)$ and $y\left(\tilde{t}\right)$.

Error in the observation process (σ):

$$\begin{aligned}
[\sigma|\cdot] &\propto \prod_t [\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2]^{z(t)}[\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma^2_\mu]^{1-z(t)}[\sigma] \\
&\propto \prod_t \mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I})^{z(t)}\mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I} + \sigma^2_\mu)^{1-z(t)}
\end{aligned}$$

The update for σ proceeds using Metropolis-Hastings.

Homerange dispersion parameter (σ_μ):

$$\begin{aligned}
[\sigma_\mu|\cdot] &\propto \prod_{\{t:z(t)=0\}} [\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2, \sigma^2_\mu][\sigma_\mu] \\
&\propto \prod_{\{t:z(t)=0\}} \mathcal{N}(\mathbf{s}(t)|\boldsymbol{\mu}_{0,h_t}, \sigma^2\mathbf{I} + \sigma^2_\mu \mathbf{I})
\end{aligned}$$

The update for σ_μ proceeds using Metropolis-Hastings. Note that the product is over all t for which $z(t) = 0$.

Variance of random effects (σ^2_α):

$$
\begin{aligned}
[\sigma^2_\alpha|\cdot] &\propto [\boldsymbol{\alpha}|\mathbf{0}, \sigma^2_\alpha][\sigma^2_\alpha \mid r_\sigma, q_\sigma] \\
&\propto \mathcal{N}(\boldsymbol{\alpha}|\mathbf{0}, \sigma^2_\alpha \mathbf{I})\mathrm{IG}(\sigma^2_\alpha \mid r_\sigma, q_\sigma) \\
&\propto |\sigma^2_\alpha \mathbf{I}|^{-1/2} \exp\left\{ -\frac{1}{2} \left((\boldsymbol{\alpha} - \mathbf{0})' \left(\sigma^2_\alpha \mathbf{I}\right)^{-1} (\boldsymbol{\alpha} - \mathbf{0}) \right) \right\} \times \\
&\qquad \left(\sigma^2_\alpha\right)^{-(q_\sigma+1)} \exp\left\{ -\frac{1}{r_\sigma \sigma^2_\alpha} \right\} \\
&\propto \left(\sigma^2_\alpha\right)^{-M/2} \exp\left\{ -\frac{1}{2\sigma^2_\alpha} \boldsymbol{\alpha}'\boldsymbol{\alpha} \right\} \\
&\qquad \left(\sigma^2_\alpha\right)^{-(q_\sigma+1)} \exp\left\{ -\frac{1}{r_\sigma \sigma^2_\alpha} \right\} \\
&\propto \left(\sigma^2_\alpha\right)^{-(M/2+q+1)} \exp\left\{ -\frac{1}{\sigma^2_\alpha} \left(\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2} + \frac{1}{r_\sigma} \right) \right\} \\
&= \mathrm{IG}\left(\left(\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2} + \frac{1}{r_\sigma} \right)^{-1}, \frac{M}{2} + q_\sigma \right)
\end{aligned}
$$

where M is the length of $\boldsymbol{\alpha}$ (or column dimension of $\mathbf{W}$).

References

Escobar, M.D., and M. West. 1995.Bayesian density estimation and inference using mixtures. Journal of the American Statistical Association, 90:577–588.

Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. 2014. Bayesian data analysis. CRC Press.

Ishwaran, H., and L.F. James. 2001. Gibbs sampling methods for stick-breaking priors. Journal of the American Statistical Association 96: 161–173.

West, M. 1997. Hyperparameter estimation in Dirichlet process mixture models. Unpublished?