Supplement to

“Spectral-based noncentral F mixed effect models, with application to otoacoustic emissions”

Lai Wei\(^1\), Peter F. Craigmile\(^2,\ast\), Wayne M. King\(^3\)

\(^1\) Biostatistics Center, The Ohio State University, Columbus. OH 43210.

\(^2\) Dept. of Statistics, The Ohio State University, Columbus. OH 43210.

\(^3\) Signal Processing and Communications, The Mathworks, Inc., 3 Apple Hill, Natick, MA 01760

\ast\) email: pfc@stat.osu.edu

The MCMC algorithm used to fit the Bayesian DPOAE noncentral F mixed model

We use \(\theta \setminus \{\kappa\}\) to denote the collection of parameters contained in \(\theta\), but excluding the parameter \(\kappa\). The MCMC algorithm below assumes that \(K_j = K\) for all subjects \(j\). (More carefully subsetting and indexing is needed in the unbalanced case.)

**Update** \(\{\log r_{j,k}\}\): Fix a subject \(j = 1, \ldots, J\), and let \(\log r_{j,*} = (\log r_{j,1}, \ldots, \log r_{j,K})^T\). Then

\[
\pi(\log r_{j,*} | z, \theta \setminus \{\log r_{j,*}\}) \propto \prod_{k=1}^{K} \left[ \prod_{l=1}^{L} f_{Z_{j,k,l}}(z_{j,k,l} | e^{N \Delta r_{j,k}}) \right] n(\log r_{j,k} | \alpha_k + G_j \beta_k, \tau_k^2).
\]

We use a Metropolis-Hastings symmetric random walk update. Supposing we are at \(\log r_{j,*}\), we propose \(\log r_{j,*}^{new}\) from a \(K\)-variate normal with mean \(\log r_{j,*}\) and covariance \(\Sigma_j\), for some \(K \times K\) positive definite matrix \(\Sigma_j\) (in practice, we base \(\Sigma_j\) on the estimated covariance matrix of the maximum likelihood estimate of \(\log r_{j,*}\), calculated using only the data for
subject \(j\), scaled to obtain an acceptance probability of around 0.4. Then we accept the new value, \(\log r_{j,\star}^{new}\), with probability \(\min(e^q, 1)\), where

\[
q = \left\{ \sum_{k=1}^K \sum_{l=1}^L \log f_{Z_{j,k,l}}(z_{j,k,l} \mid e^{N\Delta r_{j,k}^{new}}) + \sum_{k=1}^K \log n(\log r_{j,k}^{new} | \alpha_k + G_j \beta_k, \tau_k^2) \right\} - \left\{ \sum_{k=1}^K \sum_{l=1}^L \log f_{Z_{j,k,l}}(z_{j,k,l} \mid e^{N\Delta r_{j,k}}) + \sum_{k=1}^K \log n(\log r_{j,k} | \alpha_k + G_j \beta_k, \tau_k^2) \right\} ;
\]

otherwise we remain at \(\log r_{j,\star}\).

**Update \{\alpha_k\} and \{\beta_k\}:** Fixing a \(k = 1, \ldots, K\) we sample \(\alpha_k\) and \(\beta_k\) jointly, conditional on the data and other parameters. First note that

\[
\pi(\alpha_k, \beta_k | z, \theta \setminus \{\alpha_k, \beta_k\}) \propto \left\{ \prod_{j=1}^J n(\log r_{j,k} | \alpha_k + G_j \beta_k, \tau_k^2) \right\} n(\alpha_k | \mu_\alpha, \sigma_\alpha^2) n(\beta_k | \mu_\beta, \sigma_\beta^2).
\]

Let \(X\) be a \(J \times 2\) design matrix with first column all ones, and second column \(G_j\) \((j = 1, \ldots, J)\). Then \(\beta_{\star,k} = (\beta_1, \ldots, \log r_{j,k})^T\), conditional on \(\alpha_k, \beta_k\) and \(\tau_k^2\), is \(J\)-variate normal with mean \(X(\alpha_k, \beta_k)^T\) and covariance \(\tau_k^2 I_J\), where \(I_J\) is the \(J \times J\) identity matrix. This is a Bayesian regression model. Hence, letting

\[
V = \begin{bmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix},
\]

our sample from \((\alpha_k, \beta_k)^T\), conditional on the data and other parameters, is a bivariate normal draw with a mean \(\Sigma^{-1} c\) and covariance \(\Sigma^{-1}\) where \(\Sigma = X^T X / \tau_k^2 + V^{-1}\) and \(c = X^T \beta_{\star,k} / \tau_k^2 + V^{-1}(\mu_\alpha, \mu_\beta)^T\).
Update $\{\tau_k^2\}$: For each $k = 1, \ldots, K$ we have that

$$
\pi(\tau_k^2| \mathbf{z}, \theta \setminus \{\tau_k^2\}) \propto \left\{ \prod_{j=1}^J n(\log r_{j,k}|\alpha_k + G_j \beta_k, \tau_k^2) \right\} \text{ig}(\tau_k^2|s_\tau, r_\tau),
$$

leading us to sample $\tau_k^2$, conditional on the data and other parameters, from an inverse gamma distribution with shape $s$ and rate $r$, where

$$
s = s_\sigma + J/2, \quad \text{and} \quad r = r_\sigma + \frac{1}{2} \sum_{j=1}^J (\log r_{j,k} - \alpha_k - G_j \beta_k)^2.
$$

Update $\mu_\alpha$: (The update for $\mu_\beta$ is similar.)

$$
\pi(\mu_\alpha| \mathbf{z}, \theta \setminus \{\mu_\alpha\}) \propto \left\{ \prod_{k=1}^K n(\alpha_k|\mu_\alpha, \sigma_\alpha^2) \right\} n(\mu_\alpha|m_\alpha, v_\alpha),
$$

and hence we sample $\mu_\alpha$, conditional on the data and other parameters, from a normal distribution with mean $m/p$ and variance $1/p$ where

$$
m = \sum_{k=1}^K \alpha_k/\sigma_\alpha^2 + m_\alpha/v_\alpha, \quad \text{and} \quad p = K/\sigma_\alpha^2 + 1/v_\alpha.
$$

Update $\sigma_\alpha^2$: (The update for $\sigma_\beta^2$ is similar.)

$$
\pi(\sigma_\alpha^2| \mathbf{z}, \theta \setminus \{\sigma_\alpha^2\}) \propto \left\{ \prod_{k=1}^K n(\alpha_k|\mu_\alpha, \sigma_\alpha^2) \right\} \text{ig}(\sigma_\alpha^2|s_\sigma, r_\sigma),
$$

and hence we sample $\sigma_\alpha^2$, conditional on the data and other parameters, from an inverse gamma distribution with shape $s$ and rate $r$, where

$$
s = s_\tau + K/2, \quad \text{and} \quad r = r_\tau + \frac{1}{2} \sum_{k=1}^K (\alpha_k - \mu_\alpha)^2.
$$