Model, priors and the Gibbs sampler

Muhammad Yaseen

2026-11-04

Table of contents

1 Abstract **1**

2 Model **2**

3 Priors **2**

4 Conditional distributions **3**
4.1 Overall mean and main effects 3
4.2 Singular values 3
4.3 Score vectors 4
4.4 Error variance 4

5 Signs and identifiability **4**

6 The sampler **4**

7 Extensions **5**

8 References **5**

1 Abstract

This vignette states the AMMI model, the prior distributions and the conditional posterior distributions that `bayes_ammi()` samples. It follows Crossa et al. (2011), with each conditional written out for the balanced design. Each formula is followed by a check against the package.

2 Model

For genotype $i = 1, \dots, G$ in replicate $l = 1, \dots, B$ of environment $j = 1, \dots, E$,

$$y_{ijl} = \mu + \tau_i + \delta_j + \sum_{k=1}^{K} \lambda_k \alpha_{ik} \gamma_{jk} + e_{ijl}, \qquad e_{ijl} \sim N(0, \sigma^2),$$

with $\sum_i \tau_i = \sum_j \delta_j = 0$ and $\lambda_1 \geq \cdots \geq \lambda_K > 0$. The score vectors $\boldsymbol{\alpha}_k = (\alpha_{ik})$ and $\boldsymbol{\gamma}_k = (\gamma_{jk})$ have unit length and are orthogonal to $\mathbf{1}$ and to the other vectors of their kind. With $n = BGE$ observations, the cell means $\mathbf{Y} = (\bar{y}_{ij})$ and the within-cell sum of squares $\text{SSW} = \sum_{ijl}(y_{ijl} - \bar{y}_{ij})^2$ are sufficient statistics when the design is balanced:

$$\sum_{ijl}(y_{ijl} - \eta_{ij})^2 = \text{SSW} + B\sum_{ij}(\bar{y}_{ij} - \eta_{ij})^2.$$

3 Priors

`bam_priors()` sets the vague but proper priors of the paper:

$$\mu \sim N(\mu_0, \sigma_\mu^2), \quad \boldsymbol{\tau} \sim N(\mathbf{m}_\tau, \sigma_\tau^2 \mathbf{I}), \quad \boldsymbol{\delta} \sim N(\mathbf{m}_\delta, \sigma_\delta^2 \mathbf{I}), \quad \lambda_k \sim N^+(\lambda_0, \sigma_\lambda^2),$$

$$\sigma^2 \sim \text{Inv-}\chi^2(\nu, S_0^2),$$

with uniform priors for $\boldsymbol{\alpha}_k$ and $\boldsymbol{\gamma}_k$ on their constrained spheres. The variances are 10^8 and $\nu = 1$ by default. S_0^2 defaults to the least-squares residual mean square, so the prior carries the information of one observation whatever the unit of the trait. Hobert and Casella (1996) show why ν and S_0^2 must be positive.

```
bam_priors(
    mu_mean     = 0
  , mu_var      = 1e+8
  , tau_mean    = 0
  , tau_var     = 1e+8
  , delta_mean  = 0
  , delta_var   = 1e+8
  , lambda_mean = 0
  , lambda_var  = 1e+8
  , nu          = 1
  , s2          = NULL
  )
```

```
Priors of the Bayesian AMMI model
  mu            ~ N(0e+00, 1e+08)
  tau_i         ~ N(0e+00, 1e+08), sum zero
  delta_j       ~ N(0e+00, 1e+08), sum zero
  lambda_k      ~ N+(0e+00, 1e+08), ordered
  sigma^2       ~ Inv-chi^2(nu = 1e+00, s2 = least-squares residual mean square)
  alpha_k, gamma_k uniform on the constrained unit spheres
```

4 Conditional distributions

Write $c_x = \sigma^2/\sigma_x^2$ for the ratio of the error variance to a prior variance.

4.1 Overall mean and main effects

The main effects are written in orthonormal contrasts, $\boldsymbol{\tau} = \mathbf{K}_G^\top \mathbf{t}$ and $\boldsymbol{\delta} = \mathbf{K}_E^\top \mathbf{d}$, where $\mathbf{K}_n$ is the $(n-1) \times n$ matrix of `matrix_k()`, $\mathbf{K}_n \mathbf{K}_n^\top = \mathbf{I}$ and $\mathbf{K}_n \mathbf{1} = \mathbf{0}$. The interaction terms have zero row and column sums. In a balanced design μ, $\mathbf{t}$ and $\mathbf{d}$ are therefore conditionally independent:

$$\mu \sim N\left(\frac{n\bar{y} + c_\mu \mu_0}{n + c_\mu}, \frac{\sigma^2}{n + c_\mu}\right), \qquad \mathbf{t} \sim N\left(\frac{BE\,\mathbf{K}_G \bar{\mathbf{y}}_G + c_\tau \mathbf{K}_G \mathbf{m}_\tau}{BE + c_\tau}, \frac{\sigma^2}{BE + c_\tau}\mathbf{I}\right),$$

and $\mathbf{d}$ in the same way, with BG and the environment means $\bar{\mathbf{y}}_E$. This is equation 5 of the paper with $(\mathbf{W}^\top \mathbf{W} + \mathbf{D})$ evaluated.

```
K <-
  matrix_k(
      n = 4
    )
round(x = K %*% t(K), digits = 12)
```

```
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
```

```
round(x = K %*% rep(1, 4), digits = 12)
```

```
     [,1]
[1,]    0
[2,]    0
[3,]    0
```

4.2 Singular values

Since $\boldsymbol{\alpha}_k \perp \mathbf{1}, \boldsymbol{\alpha}_{k^*}$ and $\boldsymbol{\gamma}_k \perp \mathbf{1}, \boldsymbol{\gamma}_{k^*}$, the data enter the conditional of λ_k only through $\boldsymbol{\alpha}_k^\top \mathbf{Y} \boldsymbol{\gamma}_k$, and $\mathbf{X}_3^\top \mathbf{X}_3 = B$ in the paper's notation:

$$\lambda_k \sim N\left(\frac{B\,\boldsymbol{\alpha}_k^\top \mathbf{Y} \boldsymbol{\gamma}_k + c_\lambda \lambda_0}{B + c_\lambda}, \frac{\sigma^2}{B + c_\lambda}\right) \quad \text{truncated to } (\lambda_{k+1}, \lambda_{k-1}),$$

with $\lambda_0 = \infty$ above the first and $\lambda_{K+1} = 0$ below the last (equation 7). The truncation keeps the order, and a draw that would leave the interval is drawn from its tail by the inverse distribution function.

4.3 Score vectors

As a function of $\boldsymbol{\alpha}_k$ the likelihood is $\exp\{(B\lambda_k/\sigma^2)\,\boldsymbol{\alpha}_k^\top \mathbf{Y}\boldsymbol{\gamma}_k\}$. This is a von Mises-Fisher density on the unit sphere of the subspace orthogonal to $\mathbf{1}$ and to the other genotype vectors (equations 8 and 9). With $\mathbf{H}_k$ an orthonormal basis of that subspace ($G-K$ columns), $\boldsymbol{\alpha}_k = \mathbf{H}_k\mathbf{a}$ and

$$\mathbf{a} \sim \text{vMF}\left(\frac{B\lambda_k}{\sigma^2}\mathbf{H}_k^\top \mathbf{Y}\boldsymbol{\gamma}_k\right),$$

drawn by the algorithm of Wood (1994). $\boldsymbol{\gamma}_k$ is drawn in the same way from $\mathbf{Y}^\top\boldsymbol{\alpha}_k$.

```
set.seed(seed = 1)
X <- t(replicate(n = 2000, expr = bayesammi:::.bam_rvmf(v = c(10, 0,
 ↪  0))))
c(mean_first = mean(X[, 1]),
  theory     = besselI(x = 10, nu = 1.5) / besselI(x = 10, nu = 0.5))
```

```
mean_first     theory
 0.8983546  0.9000000
```

4.4 Error variance

$$\sigma^2 \sim \frac{\nu S_0^2 + \text{SSW} + B\sum_{ij}(\bar{y}_{ij} - \hat{\eta}_{ij})^2}{\chi^2_{\nu+n}} \qquad \text{(equation 10)}.$$

5 Signs and identifiability

Replacing $(\boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k)$ by $(-\boldsymbol{\alpha}_k, -\boldsymbol{\gamma}_k)$ leaves the model unchanged, so the posterior has 2^K mirror images. Within one iteration the sign of $\boldsymbol{\alpha}_k$ follows that of $\boldsymbol{\gamma}_k$. Over a long run, though, a weak term can move from one mirror image to the other, and its posterior means then shrink towards zero. With `.align = TRUE` every draw takes the orientation that agrees with the least-squares vectors. This chooses one of the mirror images and leaves the model unchanged.

6 The sampler

Every iteration draws μ, $\boldsymbol{\tau}$, $\boldsymbol{\delta}$; then, for $k = 1, \ldots, K$, λ_k, $\boldsymbol{\alpha}_k$ and $\boldsymbol{\gamma}_k$; and finally σ^2. The first chain starts at the least-squares fit and the other chains at dispersed values, so that the Gelman-Rubin diagnostic can detect a chain that has not mixed.

The sampler is written in C++ (**Rcpp** and **RcppArmadillo**) and takes its random numbers from R, so `.seed` or `set.seed()` reproduces a run. A plain R version of the same sampler is kept in the package as a reference; on a balanced trial with one error variance both return the same draws (vignette 13).

7 Extensions

The same scheme carries over to three extensions, each treated in its own vignette.

- **Unbalanced trials** (vignette 7). Missing plots and empty cells are drawn from $N(\eta_{ij}, \sigma_j^2)$ at every iteration (data augmentation; Tanner and Wong 1987), after which the conditionals above apply to the completed data.
- **Environment-specific variances** (vignette 7). With $e_{ijl} \sim N(0, \sigma_j^2)$, every σ_j^2 has a scaled inverse chi-squared conditional, μ and $\boldsymbol{\delta}$ are drawn jointly, and the conditional of $\boldsymbol{\gamma}_k$ becomes a Bingham-von Mises-Fisher distribution (Hoff 2009), drawn by slice sampling one coordinate at a time.
- **The GGE model** (vignette 8). The genotype effects are dropped and $\boldsymbol{\gamma}_k$ is no longer orthogonal to $\mathbf{1}$ (Perez-Elizalde et al. 2012).

The pointwise log-likelihood of every plot is accumulated during the run, which gives DIC and WAIC without storing the draws (vignette 9).

8 References

Crossa, J., Perez-Elizalde, S., Jarquin, D., Cotes, J. M., Viele, K., Liu, G. and Cornelius, P. L. (2011). Bayesian estimation of the additive main effects and multiplicative interaction model. *Crop Science*, 51(4), 1458–1469. doi:10.2135/cropsci2010.06.0343

Viele, K. and Srinivasan, C. (2000). Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. *Journal of Statistical Planning and Inference*, 84(1–2), 201–219.

Wood, A. T. A. (1994). Simulation of the von Mises Fisher distribution. *Communications in Statistics - Simulation and Computation*, 23(1), 157–164.

Hobert, J. P. and Casella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91(436), 1461–1473.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley, Chichester.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398), 528–540.

Hoff, P. D. (2009). Simulation of the matrix Bingham-von Mises-Fisher distribution, with applications to multivariate and relational data. *Journal of Computational and Graphical Statistics*, 18(2), 438–456.

Perez-Elizalde, S., Jarquin, D. and Crossa, J. (2012). A general Bayesian estimation method of linear-bilinear models applied to plant breeding trials with

genotype x environment interaction. *Journal of Agricultural, Biological, and Environmental Statistics*, 17, 15–37. doi:10.1007/s13253-011-0063-9