Model, priors from an earlier trial and the Gibbs sampler

Muhammad Yaseen

2026-11-04

Table of contents

1 Abstract **1**

2 Models **1**

3 Priors from an earlier trial **2**

4 Conditional distributions **2**

5 Drawing the score matrices **2**

6 Signs and starting values **3**

7 The sampler **3**

8 References **3**

1 Abstract

This vignette states the linear-bilinear models, the conjugate priors that carry an earlier trial into the analysis, the conditional posterior distributions and the sampler of `bayes_ammi()`, following Perez-Elizalde et al. (2012). One step is new: when all bilinear terms are fitted, the genotype scores can only move two columns at a time.

2 Models

With $\mathbf{Y}$ the $G \times E$ table of cell means and B replicates,

$$\mathbf{Y} = \mu \mathbf{1}_G \mathbf{1}_E^\top + \boldsymbol{\tau} \mathbf{1}_E^\top + \mathbf{1}_G \boldsymbol{\delta}^\top + \mathbf{U}\mathbf{D}\mathbf{V}^\top + \bar{\mathbf{E}}$$

(AMMI), with $\mathbf{1}^\top\boldsymbol{\tau} = \mathbf{1}^\top\boldsymbol{\delta} = 0$, $\mathbf{D} = \mathrm{diag}(\lambda_1 \geq \cdots \geq \lambda_K)$, $\mathbf{U}^\top\mathbf{U} = \mathbf{V}^\top\mathbf{V} = \mathbf{I}$ and $\mathbf{U}^\top\mathbf{1} = \mathbf{V}^\top\mathbf{1} = \mathbf{0}$. Dropping $\boldsymbol{\tau}$ and the constraint $\mathbf{V}^\top\mathbf{1} = \mathbf{0}$ gives the sites regression model (SREG, GGE).

3 Priors from an earlier trial

With $\mathbf{Y}_0$ the cell means of the earlier trial, n_0 its replicates, μ_0, $\boldsymbol{\tau}_0$, $\boldsymbol{\delta}_0$ its effects and $\mathbf{U}_0\mathbf{D}_0\mathbf{V}_0^\top$ the singular value decomposition of its interaction, the priors are conjugate and scale with σ^2:

$$\mu \sim N\left(\mu_0, \frac{\sigma^2}{n_0 GE}\right), \quad \mathbf{K}_G\boldsymbol{\tau} \sim N\left(\mathbf{K}_G\boldsymbol{\tau}_0, \frac{\sigma^2}{n_0 E}\mathbf{I}\right), \quad \mathbf{K}_E\boldsymbol{\delta} \sim N\left(\mathbf{K}_E\boldsymbol{\delta}_0, \frac{\sigma^2}{n_0 G}\mathbf{I}\right),$$

$$\lambda_k \sim N\left(\lambda_{0k}, \frac{\sigma^2}{n_0}\right) \text{ truncated to the order}, \quad p(\mathbf{U}) \propto \mathrm{etr}\left(\frac{n_0}{\sigma^2}\mathbf{Y}_0\mathbf{V}_0\mathbf{D}_0\mathbf{U}^\top\right), \quad p(\mathbf{V}) \propto \mathrm{etr}\left(\frac{n_0}{\sigma^2}\mathbf{Y}_0^\top\mathbf{U}_0\mathbf{D}_0\mathbf{V}^\top\right),$$

and $aS_0^2/\sigma^2 \sim \chi^2_a$ with $a = n_0 - 1$ and S_0^2 the within-cell mean square of the earlier trial (paper, eq. 2.7-2.11). The prior counts as n_0 observations per cell: with $n_0 = B$ the earlier trial weighs as much as the new one.

Without an earlier trial, `bys_priors()` gives vague, proper priors that do not depend on σ^2 (Crossa et al. 2011).

4 Conditional distributions

Write $n = BGE$. Each conditional mixes the prior and the data in proportion $n_0 : B$:

$$\mu \mid \cdot \sim N\left(\frac{n\bar{y} + n_0 GE\mu_0}{n + n_0 GE}, \frac{\sigma^2}{n + n_0 GE}\right), \qquad \lambda_k \mid \cdot \sim N\left(\frac{B\,\boldsymbol{\alpha}_k^\top\mathbf{Y}\boldsymbol{\gamma}_k + n_0\lambda_{0k}}{B + n_0}, \frac{\sigma^2}{B + n_0}\right)$$

on $(\lambda_{k+1}, \lambda_{k-1})$, and the same for $\boldsymbol{\tau}$ and $\boldsymbol{\delta}$ (eq. 2.15). The score matrices have von Mises-Fisher conditionals (eq. 2.13-2.14),

$$p(\mathbf{U} \mid \cdot) \propto \mathrm{etr}\left\{\frac{1}{\sigma^2}(B\mathbf{Y}\mathbf{V}\mathbf{D} + n_0\mathbf{Y}_0\mathbf{V}_0\mathbf{D}_0)\mathbf{U}^\top\right\},$$

and σ^2 is scaled inverse chi-squared with $a + n$ degrees of freedom and scale $aS_0^2 + \mathrm{SSW} + B\sum_{ij}(\bar{y}_{ij} - \hat{y}_{ij})^2$. As in the paper, the dependence of the priors on σ^2 is not carried into its conditional: it would involve the normalising constants of the von Mises-Fisher priors.

5 Drawing the score matrices

The conditional of one column $\boldsymbol{\alpha}_k$ given the others is von Mises-Fisher on the unit sphere of the subspace orthogonal to $\mathbf{1}$ and to the other columns (Hoff

2009), drawn by the algorithm of Wood (1994). That subspace has dimension $G - K$. When all terms are fitted with fewer genotypes than environments, $K = G - 1$ and the subspace has dimension one: each column is fixed by the others up to its sign, and a column-wise sampler never moves. This is what happens with the 12 genotypes and 11 terms of the maize trial.

`bayes_ammi()` adds a move for every pair of columns. Given the other $K - 2$ columns, the pair $(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_l)$ lies in a plane with orthonormal basis $\mathbf{N}$:

$$\boldsymbol{\alpha}_k = \mathbf{N}(\cos\theta, \sin\theta)^\top, \qquad \boldsymbol{\alpha}_l = \epsilon\,\mathbf{N}(-\sin\theta, \cos\theta)^\top, \qquad \epsilon = \pm 1.$$

With von Mises-Fisher parameters $\mathbf{c}_k$ and $\mathbf{c}_l$, $\mathbf{a} = \mathbf{N}^\top\mathbf{c}_k$ and $\mathbf{b} = \mathbf{N}^\top\mathbf{c}_l$, the density is $\exp\{\kappa_\epsilon \cos(\theta - \phi_\epsilon)\}$ with $\kappa_\epsilon \cos\phi_\epsilon = a_1 + \epsilon b_2$ and $\kappa_\epsilon \sin\phi_\epsilon = a_2 - \epsilon b_1$. So ϵ is drawn with probability proportional to $I_0(\kappa_\epsilon)$, then θ from the von Mises distribution (Best and Fisher 1979). The move is an exact Gibbs step. It runs whenever a score matrix fills its constrained space, for the genotype scores ($K = G - 1$) or the environment scores ($K = E - 1$, or $K = E$ for SREG).

```
set.seed(seed = 1)
th <- replicate(n = 5000, expr = baystability:::.bys_rvonmises(m = 0.5,
 ↪  kap = 3))
c(simulated = mean(cos(th - 0.5)), theory = besselI(x = 3, nu = 1) /
 ↪  besselI(x = 3, nu = 0))
```

```
simulated    theory
0.8114549 0.8099853
```

6 Signs and starting values

$(\boldsymbol{\alpha}_k, \boldsymbol{\gamma}_k)$ and $(-\boldsymbol{\alpha}_k, -\boldsymbol{\gamma}_k)$ give the same model. Every draw keeps the orientation of the earlier trial's vectors (`.align = TRUE`), oriented themselves like the least-squares vectors of the trial. With an earlier trial the first chain starts at its least-squares fit, the others at dispersed values.

7 The sampler

Every iteration draws μ, $\boldsymbol{\tau}$ and $\boldsymbol{\delta}$; then, for $k = 1, \ldots, K$, λ_k, $\boldsymbol{\alpha}_k$ and $\boldsymbol{\gamma}_k$; then the pair moves when needed; and finally σ^2. The sampler is compiled (C++) and takes its random numbers from R. A plain R version is kept in the package, and both give the same draws from the same seed (vignette 5).

8 References

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

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

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.

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

Best, D. J. and Fisher, N. I. (1979). Efficient simulation of the von Mises distribution. *Journal of the Royal Statistical Society C (Applied Statistics)*, 28(2), 152–157.