Validation and corrections

Muhammad Yaseen

2026-11-04

Table of contents

1 Abstract **1**

2 Corrections to version 0.2.0 **1**

3 The compiled sampler and the R reference **2**

4 The pair moves **3**

5 Simulated data **4**

6 The published analysis **6**

7 References **6**

1 Abstract

Version 1.0.0 replaces the sampler of version 0.2.0. This vignette lists what changed and why. It then checks the new sampler on its building blocks, against a plain R version of itself, on simulated data, and on the published analysis (vignette 3).

2 Corrections to version 0.2.0

Version 0.2.0	Now
`bayes_ammi()` read the earlier trial from objects in the workspace (`n0`, `mu0`, `tau0`, `lambdas0`, …) and stopped with "object 'n0' not found" without them	the earlier trial is the argument `.prior_data`; vague priors without it

Version 0.2.0	Now
the singular values were computed once and never sampled	drawn every iteration from their truncated normal conditional (eq. 2.15)
`tau1` held a precision drawn from a residual matrix computed once before the loop	`tau1` holds the draws of σ^2, from the current residuals
the scores were drawn without their constraints and orthogonalised afterwards	von Mises-Fisher draws on the constrained spheres, with pair moves when all terms are fitted
all 11 terms, no burn-in, one chain, no seed	`.k`, `.burnin`, `.thin`, `.chains`, `.seed`; summaries, HPD intervals, diagnostics, DIC and WAIC
AMMI only	AMMI and SREG (`.model`)
columns of `delta1` named for the genotypes	`delta1`, …, `delta25`
help page of `cultivo2009` described another data set	900 rows: site, rep, entry, y

The tables and effects (`bys_ls_means()`, `bys_ls_var()`, `bys_ls_gen_effects()`, `bys_ls_env_effects()`, `bys_ls_ge_effects()`, `bys_ls_ammi()`, `bys_contrasts()`, `bys_orthonormal()`) return the numbers of version 0.2.0 (test `test-helpers.R`).

3 The compiled sampler and the R reference

```
ns <- asNamespace(ns = "baystability")
ex <- list(y = quote(y), gen = quote(entry), env = quote(site), rep =
 ↪  quote(rep))
cd <-
  ns$.bys_prep(
      "check"
    , cultivo2009
    , ex$y
    , ex$gen
    , ex$env
    , ex$rep
    , environment()
    , allow_missing = TRUE
    )
K  <- 11
ls <- ns$.bys_ls(cd$Y, K)
tp <- ns$.bys_trial_prior("check", cultivo2008, ex, environment(), cd, K,
 ↪  FALSE, NULL)
pr <-
```

```
  bys_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
    )
pr$nu <- tp$a
pr$s2 <- tp$s2
pl <- ns$.bys_prior_list(pr, 12, 25, K, tp)
st <- ns$.bys_start(ls, tp$s2, 1)
set.seed(seed = 9)
r  <- ns$.bys_chain(cd$Y, 3, cd$ssw, K, pl, st, list(alpha = ls$alpha),
 ↪  5, 1:5)
set.seed(seed = 9)
cc <- ns$.bys_run_chain(cd, K, FALSE, FALSE, pr, st, ls$alpha, 5, rep(1L,
 ↪  5), TRUE, FALSE, tp)
c(scalars = max(abs(cc$scalar - r$scalar) / abs(r$scalar)), alpha =
 ↪  max(abs(cc$alpha - r$alpha)),
  gamma = max(abs(cc$gamma - r$gamma)))
```

```
     scalars        alpha        gamma
3.739605e-13 1.665335e-14 1.132427e-14
```

The two agree to rounding error, including the pair moves.

4 The pair moves

Without the pair moves the genotype scores of an 11-term fit stay at their starting values. With them they mix:

```
fm <-
  bayes_ammi(
      .data       = cultivo2009
    , .y          = y
    , .gen        = entry
    , .env        = site
    , .rep        = rep
    , .nIter      = 1000
    , .prior_data = cultivo2008
    , .n0         = NULL
    , .model      = "AMMI"
    , .k          = NULL
```

```
    , .burnin     = 200
    , .thin       = 1
    , .chains     = 2
    , .priors     = bys_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
        )
    , .level      = 0.95
    , .seed       = 3
    , .align      = TRUE
    , .waic       = TRUE
    )
fm$summary[block == "alpha", list(
    min_sd      = min(sd)
  , median_rhat = stats::median(rhat)
  , max_rhat    = max(rhat)
  )]
```

```
       min_sd median_rhat max_rhat
1: 0.08049804    1.000215 1.007371
```

5 Simulated data

```
set.seed(seed = 8)
G <- 8
E <- 12
B <- 3
A  <- qr.Q(qr(cbind(1, matrix(data = rnorm(G * 2), nrow = G))))[, 2:3]
Gm <- qr.Q(qr(cbind(1, matrix(data = rnorm(E * 2), nrow = E))))[, 2:3]
tau <- as.vector(scale(x = rnorm(n = G, mean = 0, sd = 6), scale =
 ↪  FALSE))
del <- as.vector(scale(x = rnorm(n = E, mean = 0, sd = 10), scale =
 ↪  FALSE))
M <- 100 + outer(X = tau, Y = del, FUN = "+") + A %*% (c(80, 40) * t(Gm))
sim <- expand.grid(rep = 1:B, gen = 1:G, env = 1:E)
sim$y <- M[cbind(sim$gen, sim$env)] + rnorm(n = nrow(sim), mean = 0, sd =
 ↪  5)
fs <-
  bayes_ammi(
```

```
      .data       = sim
    , .y          = y
    , .gen        = gen
    , .env        = env
    , .rep        = rep
    , .nIter      = 3000
    , .prior_data = NULL
    , .n0         = NULL
    , .model      = "AMMI"
    , .k          = 2
    , .burnin     = 500
    , .thin       = 1
    , .chains     = 2
    , .priors     = bys_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
        )
    , .level      = 0.95
    , .seed       = 9
    , .align      = TRUE
    , .waic       = TRUE
    )
s <- fs$summary
c(tau_covered   = mean(s[block == "tau", lower] <= tau & tau <= s[block
 ↪  == "tau", upper]),
  delta_covered = mean(s[block == "delta", lower] <= del & del <= s[block
   ↪  == "delta", upper]))
```

```
  tau_covered delta_covered
        0.875         1.000
```

```
s[block %in% c("sigma2", "lambda"), list(parameter, truth = c(25, 80,
 ↪  40), mean, lower, upper)]
```

```
   parameter truth     mean    lower    upper
1:    sigma2    25 27.22189 22.25513 32.24195
2: lambda[1]    80 80.55573 74.65189 86.69097
3: lambda[2]    40 41.83578 35.73567 47.56598
```

6 The published analysis

Vignette 3 compares the 38 linear effects of Table B.1, σ, the singular values and the significant scores of Table 2. The test `test-perez2012.R` repeats the comparison of Table B.1 and σ.

```
bys_reference(
    source   = "Perez"
  , evaluate = TRUE
  , level    = 1
  )[, list(quantity, value, computed, agrees)]
```

```
1: posterior mean of mu with the first year as prior and equal weights: the average of the t
2:                                          posterior mean of alpha_7: the average of the ye
3:                                        posterior mean of beta_5: the average of the yearl
   value  computed agrees
1:  3848 3848.2749   TRUE
2:   499  499.1293   TRUE
3:  1536 1535.5777   TRUE
```

7 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.