Getting started with bayesammi

Muhammad Yaseen

2026-11-04

Table of contents

1 Abstract **1**

2 The data **1**

3 The least-squares AMMI fit **2**

4 The Bayesian fit **3**

5 Traces and the biplot **5**

6 Interpretation **8**

7 Where to go next **10**

8 References **10**

1 Abstract

bayesammi fits the AMMI model of a multi-environment trial by Bayesian methods (Crossa et al. 2011). This vignette goes from the plot data of the maize trial `Maiz` to posterior summaries, traces, the biplot with highest posterior density (HPD) regions and a written interpretation. Each later vignette takes one part further.

2 The data

`Maiz` holds the grain yield (kg/ha) of 9 maize genotypes in 20 sites, a randomized complete block design with 4 replicates in every site, one row per plot.

```
data(Maiz)
str(Maiz)
```

```
'data.frame':   720 obs. of  7 variables:
 $ site  : int  1 1 1 1 1 1 1 1 1 1 ...
 $ entry : int  1 1 1 1 2 2 2 2 3 3 ...
 $ fold10: int  1 1 1 1 2 2 2 2 3 3 ...
 $ rep   : int  1 2 3 4 1 2 3 4 1 2 ...
 $ y     : num  3700 2948 3745 4093 3801 ...
 $ fold5 : int  1 1 1 1 2 2 2 2 3 3 ...
 $ fold3 : int  1 1 1 1 2 2 2 2 3 3 ...
```

```
m <-
  ge_mean(
      .data = Maiz
    , .y    = y
    , .gen  = entry
    , .env  = site
    )
round(m$ge_means[, 1:6])
```

```
     1    2    3    4    5    6
1 3622 3728 5554 4566 4380 6437
2 3426 3919 4937 4963 5201 6036
3 3446 4082 5117 5136 4178 6459
4 3720 4539 4542 6030 5672 6678
5 3165 4079 6173 5831 5414 6882
6 4116 4878 5205 5980 5591 6916
7 3354 4767 5389 4342 4277 6745
8 4529 3393 5248 4442 4476 4986
9 3136 4500 3780 5781 5407 5610
```

```
m$grand_mean
```

```
[1] 4858.138
```

3 The least-squares AMMI fit

The classical AMMI analysis decomposes the interaction effects by their singular values. `bayes_ammi()` starts its chains there.

```
ls_fit <-
  ge_ammi(
      .data = Maiz
    , .y    = y
    , .gen  = entry
    , .env  = site
    , .rep  = rep
    )
round(ls_fit$lambdas)
```

```
[1] 5923 3070 2552 2320 1759 1313  793  757
```

```
ls_fit$sigma2
```

```
[1] 602844.9
```

4 The Bayesian fit

Two chains of 3000 kept iterations after a burn-in of 500 take a few seconds. The first two bilinear terms are fitted, as in Crossa et al. (2011).

```
fit <-
  bayes_ammi(
      .data   = Maiz
    , .y      = y
    , .gen    = entry
    , .env    = site
    , .rep    = rep
    , .nIter  = 3000
    , .k      = 2
    , .burnin = 500
    , .thin   = 1
    , .chains = 2
    , .priors = 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
        )
    , .level  = 0.95
    , .seed   = 1
    , .align  = TRUE
    , .model  = "AMMI"
    , .het    = FALSE
    , .waic   = TRUE
    )
fit
```

```
Bayesian AMMI with 2 multiplicative term(s): 9 genotypes, 20 environments, 4 replicates
2 chain(s) x 3000 kept draws (iterations 3000 after a burn-in of 500, thinning 1); 2.2 s

Posterior means, standard deviations and 95% HPD intervals
    parameter      mean       sd     lower      upper  ess   rhat
```

```
 1:        mu   4858.07    31.89   4795.72   4920.000 6000 1.0000
 2:    sigma2 762170.00 43547.08 680191.01 849256.062 3485 1.0004
 3:    tau[1]   -242.76    91.46   -423.86    -66.666 6000 1.0000
 4:    tau[2]   -253.40    93.64   -436.19    -77.744 6000 1.0008
 5:    tau[3]    -33.52    91.18   -218.04    136.232 6000 0.9999
 6:    tau[4]    359.55    93.27    172.14    539.121 6000 0.9998
 7:    tau[5]    388.98    92.32    208.37    567.004 6096 1.0007
 8:    tau[6]    473.31    91.75    301.12    657.735 6421 1.0010
 9:    tau[7]   -187.34    91.91   -373.11     -9.187 5933 1.0001
10:    tau[8]   -574.59    92.03   -761.32   -400.553 6000 1.0004
11:    tau[9]     69.76    92.85   -111.00    256.875 6103 0.9999
12:  delta[1]  -1246.13   142.02  -1525.49   -962.924 6000 1.0002
13:  delta[2]   -649.12   141.57   -940.63   -384.929 6000 1.0007
14:  delta[3]    245.13   139.84    -19.09    531.615 6000 1.0003
15:  delta[4]    371.33   142.61     90.07    646.447 6000 0.9999
16:  delta[5]     95.88   142.33   -172.80    386.005 6000 1.0006
17:  delta[6]   1446.97   143.88   1162.49   1721.442 6000 1.0016
18:  delta[7]  -1630.75   141.67  -1907.90  -1353.769 6205 1.0008
19:  delta[8]   -828.82   141.55  -1108.06   -555.506 6000 1.0011
20:  delta[9]    114.57   142.26   -171.87    378.237 6000 1.0006
21: delta[10]  -1918.93   142.67  -2197.75  -1640.002 6000 1.0000
22: delta[11]    450.84   140.72    171.60    723.433 5780 1.0000
23: delta[12]   2657.39   142.28   2391.09   2945.944 5955 1.0006
24: delta[13]   1472.04   140.92   1198.52   1747.738 6000 1.0007
25: delta[14]   1191.53   139.66    932.72   1473.889 8040 1.0000
26: delta[15]    193.01   143.01    -84.20    468.691 6000 1.0000
27: delta[16]    548.12   143.53    260.90    819.786 6463 0.9999
28: delta[17]     16.18   138.64   -258.01    288.866 6000 1.0004
29: delta[18]   -306.32   142.38   -587.56    -31.154 6000 1.0001
30: delta[19]  -2206.97   141.22  -2486.94  -1932.614 6000 1.0000
31: delta[20]    -15.96   141.43   -291.97    261.856 6000 1.0012
32: lambda[1]   5459.38   462.21   4526.54   6305.934 4977 1.0007
33: lambda[2]   1695.05   770.70     59.23   2891.761  657 1.0000
    parameter      mean       sd     lower      upper  ess   rhat

Scores of the bilinear terms: x$summary (blocks alpha, gamma); biplot coordinates: x$scores

Sources: Crossa et al. (2011); Spiegelhalter et al. (2002); Watanabe (2010).
Full references: bam_cite(x).
```

Each row gives the posterior mean, the posterior standard deviation and the 95% HPD interval of one parameter. `ess` is the effective sample size and `rhat` the potential scale reduction factor. Values of `rhat` near 1 show that the two chains agree. The scores of the bilinear terms are in `fit$summary` too:

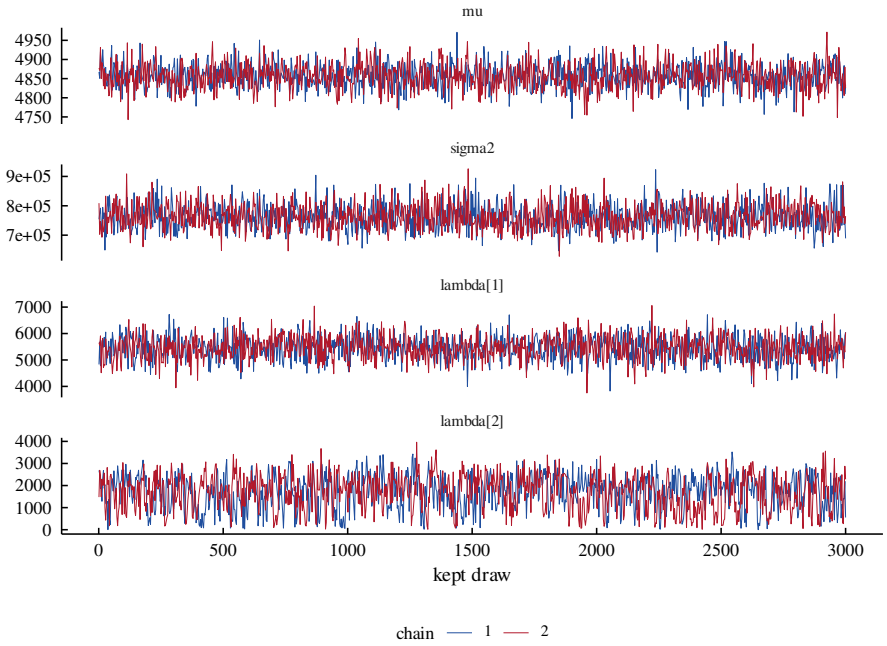
```
head(x = fit$summary[block == "alpha"], n = 9)
```

```
    parameter block label axis       mean         sd       mc_se      lower
1: alpha[1,1] alpha     1    1 -0.3706505 0.07529579 0.001658143 -0.5197187
2: alpha[2,1] alpha     2    1 -0.2396904 0.07800751 0.001611312 -0.3893679
3: alpha[3,1] alpha     3    1 -0.1586555 0.07599431 0.001311272 -0.3067989
4: alpha[4,1] alpha     4    1  0.5005975 0.06431033 0.000963502  0.3806173
5: alpha[5,1] alpha     5    1  0.2767790 0.07816070 0.001827806  0.1224777
6: alpha[6,1] alpha     6    1  0.2627733 0.07516035 0.001326350  0.1131133
7: alpha[7,1] alpha     7    1 -0.1162922 0.08445246 0.001973271 -0.2817327
8: alpha[8,1] alpha     8    1 -0.4685018 0.07829858 0.002869252 -0.6235357
9: alpha[9,1] alpha     9    1  0.3136405 0.07533103 0.001318671  0.1648137
         upper       ess      rhat
1: -0.22338869 2062.0425 1.0004630
2: -0.08263156 2343.7636 1.0007061
3: -0.01273444 3358.7402 0.9999796
4:  0.63240744 4455.0876 1.0011760
5:  0.42313429 1828.5907 1.0003721
6:  0.41014001 3211.1550 1.0009135
7:  0.04936470 1831.6872 1.0003807
8: -0.32031922  744.6813 1.0001463
9:  0.45908885 3263.4317 1.0000082
```

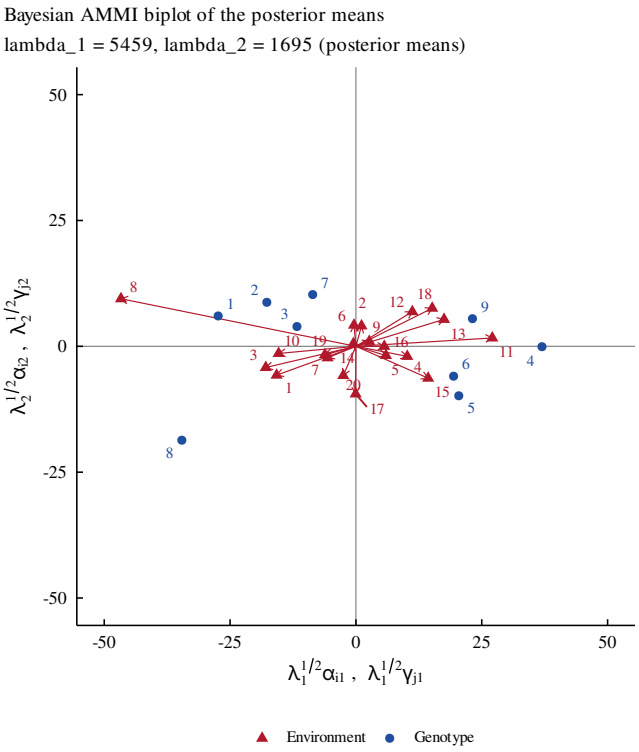
5 Traces and the biplot

```
plot(x = fit, which = "trace")
```

Traces of the Markov chains

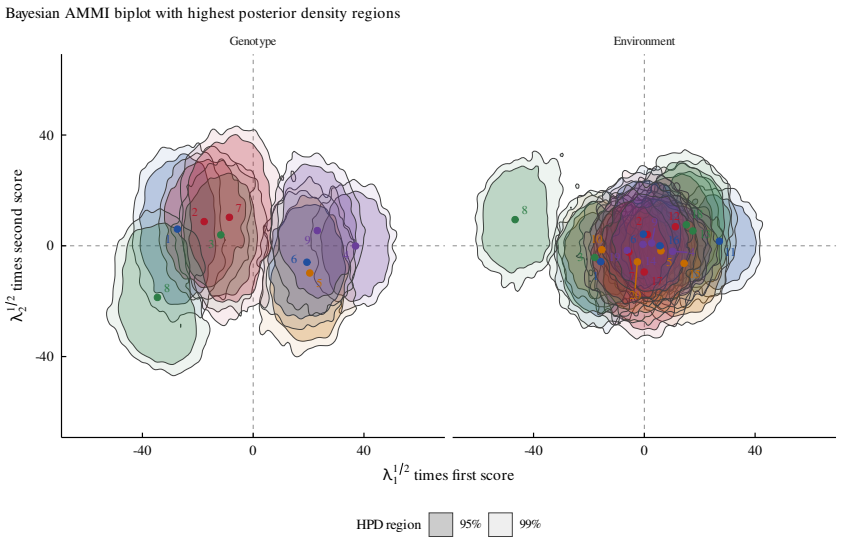



```
plot(x = fit, which = "biplot")
```



The biplot of the posterior means looks like a classical AMMI biplot. `biplots()` adds a region around every point that holds 95% (and 99%) of its posterior. A region that contains the origin marks a genotype or environment whose interaction is not different from zero.

```
bp <-
  biplots(
      model          = fit
    , burnin         = 0
    , thin           = 1
    , pb             = c(0.05, 0.01)
    , plot_stable    = TRUE
    , plot_unstable  = TRUE
    , ncolors        = 5
    , .which         = "both"
    , .grid          = 100
    )
bp$biplot
```




```
bp$biplot_data[byplot == "Genotype"]
```

```
     byplot ID name          x           y stable axis1_lower axis1_upper
1: Genotype  1    1 -27.368323   6.03321209  FALSE  -38.071788 -15.9090724
2: Genotype  2    2 -17.702243   8.74538197  FALSE  -29.061165  -5.9653409
3: Genotype  3    3 -11.710424   3.92986954   TRUE  -22.559185  -0.8968565
4: Genotype  4    4  36.962380  -0.07527436  FALSE   27.219588  46.8022130
5: Genotype  5    5  20.432367  -9.82290682  FALSE    9.285344  31.6261960
6: Genotype  6    6  19.404674  -5.93418865  FALSE    8.102189  30.2750278
7: Genotype  7    7  -8.591164  10.28368104   TRUE  -20.023364   4.2222777
8: Genotype  8    8 -34.581324 -18.65154864  FALSE  -45.676116 -22.8148948
9: Genotype  9    9  23.154057   5.49177384  FALSE   11.564476  33.7506718
   axis1_significant           group
1:              TRUE        negative
2:              TRUE        negative
3:              TRUE        negative
4:              TRUE        positive
5:              TRUE        positive
6:              TRUE        positive
7:             FALSE not significant
8:              TRUE        negative
9:              TRUE        positive
```

6 Interpretation

```
bam_interpret(
    x             = fit
  , higher_better = TRUE
  , top           = 5
  )
```

```
Bayesian AMMI analysis

2 bilinear term(s): lambda_1 = 5,459, lambda_2 = 1,695; the chains have
converged.

Findings:
- The chains have converged by the criteria of Crossa et al. (2011) (largest
  potential scale reduction factor 1.004); the smallest effective sample size
  of the main parameters is 657.
- 4 of 132 Heidelberger-Welch stationarity and half-width tests failed (about
  5% fail by chance at level 0.05).
- Overall mean 4,858 (95% HPD 4,796 to 4,920); error variance 762,200, a
  coefficient of variation of 18.0%.
- Genotypes significantly above average (HPD of the effect excludes zero): 4,
  5 and 6; below average: 1, 2, 7 and 8.
- Environments with significantly high means: 4, 6, 11, 12, 13 and 2 more;
  low: 1, 2, 7, 8, 10 and 2 more.
- Singular values with 95% HPD intervals: lambda_1 = 5,459 (4,527 to 6,306);
  lambda_2 = 1,695 (59.23 to 2,892).
- On the first bilinear term, genotypes with significantly positive scores:
  4, 5, 6 and 9; negative: 1, 2, 3 and 8.
- Environments with significantly positive first scores: 11, 12, 13, 15 and
  18; negative: 1, 3, 8 and 10.
- Genotypes and environments with significant first scores of the same sign
  interact positively: the genotype yields more there than its main effect
  predicts; of opposite sign, less.

Recommendations:
- Recommend the genotypes with significant first scores for the environments
  whose scores have the same sign (biplots() shows the regions).
- Draw biplots() to see which genotypes and environments form homogeneous
  groups (overlapping HPD regions).

Cautions:
- The HPD interval of lambda_2 reaches down to near zero: the term is weakly
  supported and its scores are uncertain.

Sources:
- 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
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society B, 64(4), 583-639. doi:10.1111/1467-9868.00353
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research, 11, 3571-3594.

7 Where to go next

- Vignette 2 derives the model, the priors and the sampler.
- Vignette 3 reproduces the published analysis of these data.
- Vignettes 4 to 6 cover the biplot regions, convergence, and the choice of priors and number of terms.
- Vignette 7 fits unbalanced trials and environment-specific error variances.
- Vignette 8 fits the GGE (sites regression) model and finds which genotype wins where.
- Vignette 9 turns the posterior into decisions: probabilities of superior performance, ranks, stability, predictions, and the choice and checking of models.
- Vignette 10 fits the Bayesian Eberhart-Russell, Finlay-Wilkinson and Shukla stability models.
- Vignette 11 gives the posterior distributions of classical and AMMI-based stability indices.
- Vignette 12 describes the application started by `bam_app()`, which runs all of this without code.
- Vignette 13 lists the corrections to version 0.3.0 and the checks of the sampler.
- Vignette 14 explains the interpretation.
- Vignette 15 lists the references of every method; `citation("bayesammi")` and `bam_cite()` give them too.

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