Analysis of GxE Interactions

With Clone Selector

Raúl Eyzaguirre

International Potato Center
CIP

June 28, 2011

Outline

1. GxE interaction
2. Linear Regression
3. AMMI and GGE Biplots
4. Tai stability analysis

Outline

1. GxE interaction
2. Linear Regression
3. AMMI and GGE Biplots
4. Tai stability analysis

Outline

Outline

What is GxE interaction and how it is described

We have GxE interaction when the performance of a genotype depends on the environment.

There is a component for GxE interaction in any statistical model. For the RCBD it is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \gamma_{k(j)} + \epsilon_{ijk}$$

If we have g genotypes and e environments, we have $g \times e$ interaction effects.

Example

```
     env1   env2   env3  env4   env5
g01 23.00  8.350 16.500 15.85  7.000
g02 16.27 15.750 21.000 17.48  3.625
g03 20.45 11.275 17.500 14.15  3.825
g04 21.68  7.425 22.000 12.90  9.700
g05 11.38  3.625 12.500 15.95  4.450
g06 14.97  1.900 10.750 15.30  9.500
g07 21.25  6.225 16.750 24.07  3.175
g08 12.45  5.625 24.500  9.50  1.975
g09 17.85 12.825 30.500 22.93  8.800
g10 14.25  6.150 12.500 18.93  4.250
g11 23.12  4.425  7.250  9.25  1.575
g12 11.25 17.500 14.500 14.50  3.375
g13 27.68  7.000 19.250 18.50 10.025
g14  7.65  2.225  4.375 17.65  2.525
```

It has $14 \times 5 = 90$ interaction effects and it needs $(14 - 1) \times (5 - 1) = 52$ degrees of freedom to be explained.

Main idea

It uses simple linear regression to explain the interaction.

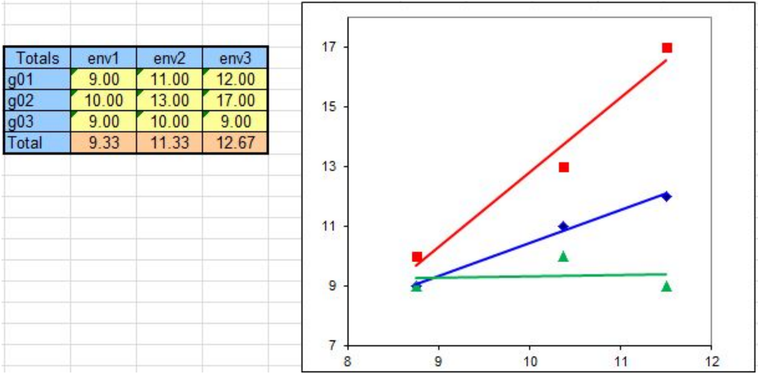
For each genotype, a simple linear regression of its individual mean value (y) is fitted on the mean of all genotypes for each environment (x).

In the previous example, you need 52 degrees of freedom to explain interaction, but just 13 to fit 14 simple linear regressions.

Example

```
              Df Sum Sq Mean Sq F value    Pr(>F)  LSD5
G             13 2599.1  199.93  5.6673 9.198e-09 3.704
E              4 7079.7 1769.92 27.7848 8.823e-07 3.215
R:E           15  955.5   63.70  1.8057 3.604e-02    NA
GxE           52 4208.8   80.94  2.2944 2.304e-05 8.283
- Het.Regr.G  13  535.7   41.21  0.4375 9.451e-01    NA
- Dev.Regr.G  39 3673.1   94.18  2.6698 4.955e-06    NA
- Het.Regr.E   4  515.9  128.98  1.6764 1.709e-01    NA
- Dev.Regr.E  48 3692.9   76.94  2.1809 1.022e-04    NA
Residuals    195 6879.1   35.28      NA        NA    NA
```

An ilustrative case



Totals	env1	env2	env3
g01	9.00	11.00	12.00
g02	10.00	13.00	17.00
g03	9.00	10.00	9.00
Total	9.33	11.33	12.67

Interpretation

- The slope of the regression line is used as a stability parameter.
- Genotypes with slope around 1 have average stability over all environments.
- Genotypes with slope greater than 1 have below average stability (they are very sensitive to changes in the environments), and hence are suitable for high-yielding environments.
- Genotypes with slope less than 1 have above average stability (they are very insensitive to changes in the environments), and hence, they could be suitable for low-yielding environments.

Main idea

It uses a singular value decomposition to explain the interaction.

Let $\mathbf{\Delta}$ be the matrix $g \times e$ of interaction effects. The rank of this matrix must be $S = \min\{g-1, e-1\}$.

This matrix can be factored in the following way:

$$\mathbf{\Delta} = \mathbf{G\Lambda E} = \sum_{s=1}^{S} \lambda_s \mathbf{g}_s \mathbf{e}_s^T$$

$\mathbf{g}_s$ and $\mathbf{e}_s$ are the singular vectors associated to the genotypes (rows of the $\mathbf{\Delta}$ matrix) and environments (columns of the $\mathbf{\Delta}$ matrix).

The AMMI model

It uses a singular value decomposition to rewrite

$$y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \gamma_{k(j)} + \epsilon_{ijk}$$

as

$$y_{ijk} = \mu + \alpha_i + \beta_j + \sum_{s=1}^{S} \lambda_s g_{si} a_{sj}^T + \gamma_{k(j)} + \epsilon_{ijk}$$

but it uses only $N < S$ components

$$y_{ijk} \approx \mu + \alpha_i + \beta_j + \sum_{s=1}^{N} \lambda_s g_{si} a_{sj}^T + \gamma_{k(j)} + \epsilon_{ijk}$$

Here the name AMMI: Additive Main effect Multiplicative Interaction.

Utility of AMMI

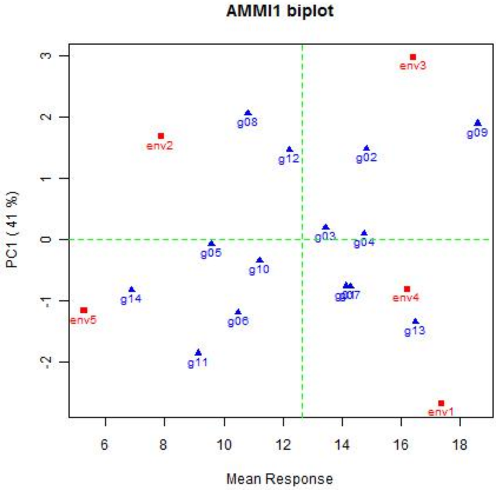
- If the number of components to explain interaction is two, you can get a nice plot.
- You can explain the interaction with less degrees of freedom (a more parsimonious model).

Example

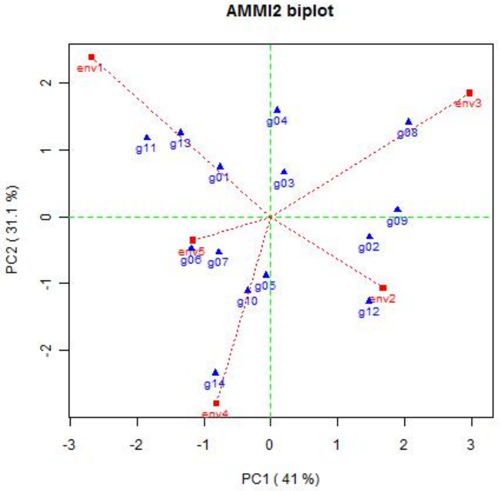
```
     env1   env2   env3  env4   env5
g01 23.00  8.350 16.500 15.85  7.000
g02 16.27 15.750 21.000 17.48  3.625
g03 20.45 11.275 17.500 14.15  3.825
g04 21.68  7.425 22.000 12.90  9.700
g05 11.38  3.625 12.500 15.95  4.450
g06 14.97  1.900 10.750 15.30  9.500
g07 21.25  6.225 16.750 24.07  3.175
g08 12.45  5.625 24.500  9.50  1.975
g09 17.85 12.825 30.500 22.93  8.800
g10 14.25  6.150 12.500 18.93  4.250
g11 23.12  4.425  7.250  9.25  1.575
g12 11.25 17.500 14.500 14.50  3.375
g13 27.68  7.000 19.250 18.50 10.025
g14  7.65  2.225  4.375 17.65  2.525
```

```
   PC   Cont Cont_Acum Df Sum Sq Mean Sq F value    Pr(>F)
1 PC1 41.017     41.02 16 1726.3  107.89   3.058 0.0001299
2 PC2 31.081     72.10 14 1308.1   93.44   2.649 0.0014853
3 PC3 18.870     90.97 12  794.2   66.18   1.876 0.0393038
4 PC4  9.032    100.00 10  380.2   38.02   1.078 0.3811887
```

Biplot example



Biplot example

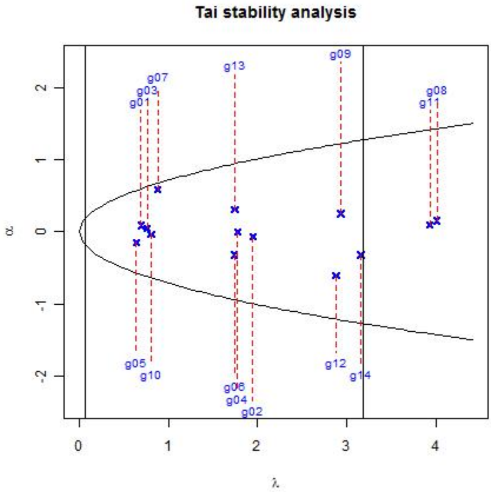


Example

It computes two stability parameters, α and λ.

```
     Slope     SE  MSdev MSentry MSinteract      PC1      PC2     alpha lambda
g01 1.0849 0.2446  7.564   42.87      5.901 -0.75663  0.74105  0.088089 0.6917
g02 0.9317 0.4102 21.267   43.39     16.098  1.47397 -0.30546 -0.070824 1.9470
g03 1.0459 0.2565  8.320   40.82      6.306  0.19427  0.66038  0.047653 0.7616
g04 0.9941 0.3916 19.389   45.77     14.543  0.08947  1.58214 -0.006136 1.7757
g05 0.8577 0.2358  7.032   28.53      5.914 -0.07546 -0.87986 -0.147582 0.6411
g06 0.6928 0.3890 19.134   29.52     17.333 -1.18480 -0.48304 -0.318636 1.7387
g07 1.5622 0.2827 10.100   84.71     17.564 -0.77473 -0.53085  0.583177 0.8795
g08 1.1439 0.5884 43.775   74.19     33.486  2.05290  1.40810  0.149319 4.0060
g09 1.2385 0.5040 32.113   72.57     25.883  1.89075  0.09664  0.247436 2.9328
g10 0.9666 0.2636  8.787   36.12      6.625 -0.34657 -1.10583 -0.034676 0.8046
g11 1.0889 0.5826 42.908   69.66     32.431 -1.85028  1.17439  0.092214 3.9285
g12 0.4129 0.5028 31.965   29.36     34.868  1.46290 -1.26646 -0.609012 2.8778
g13 1.2924 0.3892 19.147   67.15     17.063 -1.34661  1.24597  0.303359 1.7412
g14 0.6873 0.5231 34.592   40.87     29.035 -0.82918 -2.33720 -0.324382 3.1540
```

Example


Tai stability analysis
α
λ
2
1
0
-1
-2
0
1
2
3
4
g01
g02
g03
g04
g05
g06
g07
g08
g09
g10
g11
g12
g13
g14

Thank you for your attention.