ALGEBRAIC FORMULAS TO COMPUTE SUMS OF SQUARES IN THE ANALYSIS II OF GARDNER AND EBERHART

by

A. ORDAS

Misión Biológica de Galicia, C.S.I.C., Pontevedra (Spain)

INTRODUCTION

Plant breeders have often used the diallel cross to investigate the genetical properties of a group of populations. Among the numerous, different kinds of analyses that have been proposed for the diallel cross, the nonorthogonal analysis (Analysis II) suggested by Gardner and Eberhart (1966) allows to estimate heterosis effects from a fixed set of varieties when the parent populations and their crosses (without reciprocals) are included in the analysis. The partition of the entry sum of squares is obtained by least squares analysis and involves the use of an electronic computer in order to invert several matrices which are needed to calculate the sums of squares.

Gardner (1967) developed some shortcut formulas which could be used to compute the needed sums of squares in the analysis using an ordinary desk calculator. In this paper I present some algebraic formulas that simplify those presented by Gardner (1967).
DESCRIPTION OF THE FORMULAS

In the Analysis II of Gardner and Eberhart (1966) the total sum of squares among population means is subdivided as presented in Table 1.

Table 1. Nonorthogonal partition of the entry sum of squares in the Analysis II of Gardner and Eberhart (1966)

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Degrees of freedom</th>
<th>Sums of squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>( \frac{n(n+1)}{2} - 1 )</td>
<td>((B'G)_1 - CF)</td>
</tr>
<tr>
<td>Varieties</td>
<td>(n - 1)</td>
<td>((B'G)_1 - CF)</td>
</tr>
<tr>
<td>Heterosis</td>
<td>(n\left(\frac{n-1}{2}\right))</td>
<td>((B'G)_1 - (B'G)_3)</td>
</tr>
<tr>
<td>Average</td>
<td>1</td>
<td>((B'G)_2 - (B'G)_1)</td>
</tr>
<tr>
<td>Variety</td>
<td>(n - 1)</td>
<td>((B'G)_3 - (B'G)_2)</td>
</tr>
<tr>
<td>Specific</td>
<td>(n\left(\frac{n-3}{2}\right))</td>
<td>((B'G)_4 - (B'G)_2)</td>
</tr>
</tbody>
</table>

The formulas that I present to compute \((B'G)_i\) are as follows:

\[
(B'G)_1 = \frac{4}{n+2} \sum_j G_j^2 - \frac{2}{(n+1)(n+2)} V_j^2.
\]

\[
(B'G)_2 = \frac{4}{n+2} \sum_j G_j^2 + \frac{1}{n} V_j^2 + \frac{2}{n(n-1)} C_j^2 - \frac{4}{n(n+2)} V_j^2.
\]

\[
(B'G)_3 = \sum_j V_j^2 + \frac{1}{n-2} \sum_j C_j^2 - \frac{2}{(n-1)(n-2)} C_j^2.
\]

\[
(B'G)_4 = \sum_j V_j^2 + \sum_{j<j'} C_{jj'}.
\]

where:

- \(n\) = number of varieties in the diallel
- \(V_j\) = mean of the jth variety
- \(C_{jj'}\) = mean of the cross of varieties \(j\) and \(j'\)
- \(C_j = \sum_{j'} C_{jj'}\)
\[ G_j = V_j + \frac{1}{2} C_j. \]

\[ V_\cdot = \sum V_j \]

\[ C_\cdot = \sum_{j<j'} C_{jj'} \]

\[ Y_\cdot = V_\cdot + C_\cdot \]

As an illustration, the formulas presented above are applied to the data used by Gardner and Eberhart (1966). For the sake of convenience to the reader, these data are presented in Table 2 with some small modifications.

**Table 2. Mean grain yield of six open pollinated varieties of corn and one set of all possible single crosses among them (from Gardner and Eberhart (1966)) and some parameters obtained from them**

<table>
<thead>
<tr>
<th>Variety</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>( C_j^* )</th>
<th>( G_j^{**} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--- bu/a ---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>91.0</td>
<td>98.8</td>
<td>91.1</td>
<td>95.3</td>
<td>93.5</td>
<td>100.7</td>
<td>479.4</td>
<td>330.70</td>
</tr>
<tr>
<td>2</td>
<td>88.8</td>
<td>91.7</td>
<td>95.7</td>
<td>97.1</td>
<td>94.1</td>
<td>105.4</td>
<td>488.1</td>
<td>335.75</td>
</tr>
<tr>
<td>3</td>
<td>91.1</td>
<td>92.7</td>
<td>87.9</td>
<td>101.3</td>
<td>91.6</td>
<td>103.3</td>
<td>480.0</td>
<td>327.90</td>
</tr>
<tr>
<td>4</td>
<td>95.3</td>
<td>97.1</td>
<td>101.3</td>
<td>96.6</td>
<td>95.4</td>
<td>102.7</td>
<td>491.8</td>
<td>342.50</td>
</tr>
<tr>
<td>5</td>
<td>99.5</td>
<td>94.1</td>
<td>91.6</td>
<td>93.4</td>
<td>91.3</td>
<td>101.6</td>
<td>476.2</td>
<td>329.40</td>
</tr>
<tr>
<td>6</td>
<td>100.7</td>
<td>105.4</td>
<td>103.3</td>
<td>102.7</td>
<td>101.6</td>
<td>96.2</td>
<td>513.7</td>
<td>353.05</td>
</tr>
</tbody>
</table>

* \( C_j = \sum_{j<j'} C_{jj'} \)*

** \( G_j = V_j + \frac{1}{2} C_j \).**

In this example \( V_\cdot = 554.7 \), \( C_\cdot = 1,464.6 \), \( Y_\cdot = 2,019.3 \), \( CF = (2,019.3)^2/21 = 194,170.1186 \).

Applying the formulas for the computation of \((B'G)_n\), we obtain:

\[(B'G)_1 = 194,404.3486\]
\[(B'G)_2 = 194,619.7890\]
\[(B'G)_3 = 194,579.5090\]
\[(B'G)_4 = 194,643.2900\]
With these values we finally obtain the sums of squares presented in Table 3.

**Table 3. Sums of squares of six varieties and their 15 variety crosses**

<table>
<thead>
<tr>
<th>Source</th>
<th>d. f.</th>
<th>s. s.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>20</td>
<td>413.17</td>
</tr>
<tr>
<td>Varieties</td>
<td>5</td>
<td>234.23</td>
</tr>
<tr>
<td>Heterosis</td>
<td>15</td>
<td>238.94</td>
</tr>
<tr>
<td>Average</td>
<td>1</td>
<td>116.44</td>
</tr>
<tr>
<td>Variety</td>
<td>5</td>
<td>59.72</td>
</tr>
<tr>
<td>Specific</td>
<td>9</td>
<td>63.78</td>
</tr>
</tbody>
</table>

**DISCUSSION**

To obtain the sums of squares for the least squares Analysis II of Gardner and Eberhart (1966), it is necessary to invert some $A_k$ matrices according to the terminology of Eberhart (1964).

To have an idea of the magnitude of the task, these $A_k$ matrices will have an order of 7, 8, 14, and 29 for models 1, 2, 3, and 4, respectively, for a diallel of only six varieties. For ten varieties, these figures go up to 11, 12, 22, and 67. Obviously, electronic computer facilities are needed to make such computations.

The approach followed by Gardner (1967) is based on developing algebraic formulas to compute directly the sums of squares due to each of the sources of variation described in Table 1. With such an approach we lose the "flavor" of the least squares analysis, in which the sums of squares $(B'G)_i$ (for $i = 1, \ldots, 4$) are due to fitting successively more complicated models.

The formulas presented in this paper allows the investigator to calculate the $(B'G)_i$ sums of squares with a little algebraic computation that can be carried out quickly with only the help of a desk electronic calculator saving, then, costly computer time.
In addition, the structure of the least squares method of analysis is maintained.

RESUMEN

Para obtener la partición de la suma de cuadrados debida a poblaciones en el Análisis II de Gardner y Eberhart es preciso inverting varias matrices. En este trabajo se presentan unas fórmulas algebraicas que permiten calcular dichas sumas de cuadrados con una simple calculadora de mesa, evitando, así, el empleo de un ordenador para llevar a cabo la inversión matricial.

SUMMARY

In the Analysis II of Gardner and Eberhart several matrices should be inverted in order to obtain the partition of the entry sum of squares. Algebraic formulas were developed to calculate these sums of squares with a simple desk calculator overcoming, thus, the necessity of a computer to carry out matrix inversion.

REFERENCES

Eberhart, S. A.

Gardner, C. O.

Gardner, C. O., and Eberhart, S. A.