

# ANALISIS COMPUTACIONAL DE LA INTERACCION GENOTIPO-AMBIENTE CON EL MODELO AMMI

## COMPUTATIONAL ANALYSIS OF THE GENOTYPE-ENVIRONMENT INTERACTION WITH THE AMMI MODEL

Roberto Cruz Medina<sup>1</sup> y Arturo Hernández Jasso<sup>2</sup>

### RESUMEN

El modelo multiplicativo, popularizado por Gauch en la literatura especializada como el modelo AMMI (Modelo de efectos principales aditivos e interacción multiplicativa) es actualmente uno de los modelos más utilizados a nivel internacional para el estudio de la interacción genotipo-ambiente y si bien este modelo puede ser de gran utilidad para el fitomejorador, en México se le desconoce o se carece de programas computacionales para utilizarlo. En este trabajo se describe la aplicación del modelo AMMI con el paquete estadístico Statgraphics. Asimismo se proporciona un programa en SAS (Statistical Analysis System) con el mismo fin y se describen las operaciones matriciales necesarias para su aplicación. El análisis se ilustra con los datos de una investigación en algodonero (*Gossypium hirsutum L.*).

### PALABRAS CLAVE ADICIONALES

Ajuste bilineal, modelo multiplicativo, análisis computacional, statgraphics, SAS.

### SUMMARY

The multiplicative model spread over the breeder's world as the additive main effect and multiplicative interaction method (AMMI) by Gauch is one of the approaches most commonly used to study the genotype-environment interaction. This approach has proved to be useful, but in México, plant breeders do not know it or lack programs to use it. In this paper the AMMI model is applied with the Statgraphics computational package. A computational program in SAS (Statistical Analysis System) is

provided as well, and matrix operations are described in order to apply it. This analysis is exemplified with cotton (*Gossypium hirsutum L.*) research data.

### ADDITIONAL INDEX WORDS

Bilinear fit, multiplicative model, computational analysis, statgraphics, SAS.

### INTRODUCCION

La interacción genotipo-ambiente en los programas de mejoramiento genético se manifiesta, generalmente, en la dificultad para la identificación del "mejor" genotipo. Esta situación, reconocida desde principios de siglo, se puede puntualizar en las siguientes palabras de Engleow citado por Hill (1975): "Ya no podemos esperar más, como lo hicieron alguna vez los fitomejoradores, por el genotipo que sobresale dondequiera y a través de los años, nuestra aspiración es producir para cada localidad el genotipo mejor adaptado al ambiente que ésta ofrece".

La idea anterior interpretada en forma extrema indicaría que a cada localidad se le puede asignar un genotipo diferente, sin embargo, también es cierto que existen localidades distantes geográficamente pero con características ambientales semejantes que permiten la recomendación del mismo genotipo [véase por ejemplo los análisis del programa de mejoramiento de maíz del CIMMYT en Crossa (1990)]. Varios de los métodos estadísticos disponibles para el análisis de la interacción genotipo-ambiente tratan precisamente de identificar a los genotipos con características afines y a los

<sup>1</sup> Instituto Tecnológico de Sonora. Apdo. Postal 541, C.P. 85000 Cd. Obregón, Son.

<sup>2</sup> Programa de Algodonero del Campo Experimental Valle del Yaqui, Apdo. Postal 515, C.P. 85000 Cd. Obregón, Son.

conjuntos de localidades en los que éstos manifiestan su máximo potencial, sin embargo, la principal dificultad que enfrenta el fitomejorador para la utilización de estos métodos, es la carencia de programas computacionales necesarios en su aplicación.

Este trabajo tiene como objetivo presentar los programas estadísticos en Statgraphics™ [este programa se presenta con más detalle debido a que es un paquete pequeño (3 Megabytes) y económico] y en SAS™ (un paquete más grande y más difundido) para la aplicación de uno de los métodos más utilizados para el análisis de la interacción genotipo-ambiente, el modelo AMMI. Se proporcionan también algunas indicaciones para la interpretación de resultados.

## REVISION DE LITERATURA

Gran parte de los métodos más utilizados en la actualidad para el análisis de la interacción genotipo-ambiente, se derivan o son generalizaciones del modelo multiplicativo proporcionado por Yates y Cochran (1938), quienes al analizar series de experimentos con lugares, años y variedades incluyeron un análisis de regresión entre el rendimiento de cada variedad con el rendimiento promedio de cada localidad. Es interesante observar que en el ajuste de estos modelos generalizados juegan un papel muy importante los componentes principales.

De las generalizaciones de este modelo, una de las más divulgadas actualmente es la versión popularizada por Gauch (1988) como el método AMMI. En este método, la matriz de residuales de medias por ambientes después de ajustar por genotipos y ambientes se aproxima por medio de componentes principales, el número de componentes utilizados dependerá de la significancia de éstos y del porcentaje de variabilidad explicada. Sin embargo, si se desea una interpretación adecuada de la variabilidad explicada se recomienda utilizar sólo los

componentes susceptibles de ser interpretados.

Los programas presentados en este trabajo permiten utilizar cualquier número de componentes principales, pero se enfatiza la utilización del primero si resulta significativo. Como puede suceder que aún cuando el primer componente principal resulte significativo no explique un porcentaje adecuado (digamos más del 50%) de la interacción genotipo- ambiente, se pueden utilizar más componentes sacrificando la explicación de la agrupación de genotipos obtenida o bien se pueden utilizar otros métodos como los que propone Lin (1990) para el análisis de la interacción. No se tratarán estos casos debido a que el objetivo principal es proporcionar programas de computo para la aplicación del modelo AMMI.

## METODOLOGIA

Se tratará el caso más frecuente en México, cuando  $g$  genotipos se estudian en  $a$  ambientes en un diseño en bloques al azar con  $r$  repeticiones. Sin embargo, sólo se analizará la descomposición de la interacción genotipo-ambiente utilizando el modelo AMMI, suponiendo, por supuesto, que la prueba de la interacción genotipo-ambiente resultó significativa en el cuadro de análisis de varianza del modelo completo. Esta prueba con algunas suposiciones usuales sobre los efectos del modelo (fijos o aleatorios) las presenta Cruz (1990).

### Modelo AMMI

Para la aplicación del método AMMI (si la interacción genotipo-ambiente resulta significativa) se utiliza la nomenclatura siguiente:

- $\mu$  Media general
- $A_i$  Efecto del ambiente  $i$
- $G_j$  Efecto del genotipo  $j$

$Y_{ij}$	Rendimiento medio de genotipo j en el ambiente i
$Y_{i\cdot}$	Rendimiento medio en el ambiente i
$Y_{\cdot j}$	Rendimiento medio del genotipo j
$W$	matriz de orden a x g cuyos elementos son los residuales $W_{ij}$ de las medias de ambientes por genotipos al ajustar $\mu$ , $A_i$ y $G_j$
$W_{ij}$	$= Y_{ij} - Y_{i\cdot} - Y_{\cdot j} + Y_{\cdot\cdot}$

El modelo AMMI utilizando dos términos multiplicativos (dos componentes principales) postula que:

$$Y_{ij} = \mu + A_i + G_j + \lambda_1 h_{i1} f_{j1} + \lambda_2 h_{i2} f_{j2} \quad (1)$$

Donde:

$d_{ij}$  es la desviación de la respuesta lineal del genotipo j en el ambiente i.

$f'_k$  =  $(f_{1k}, f_{2k}, \dots, f_{gk})^3$  con  $k=1,2$ ; son los vectores estandarizados de las ponderaciones de los genotipos ("PCA scores de genotipos" o vectores característicos de  $W'W$  correspondientes a  $\lambda_k^2$ ).

$h'_k$  =  $(h_{1k}, h_{2k}, \dots, h_{ak})$  con  $k=1,2$ ; son los vectores estandarizados de las ponderaciones de los ambientes ("PCA scores de ambientes" o vectores característicos de  $WW'$  correspondientes  $\lambda_k^2$ ).

$\lambda_k^2$  con  $k=1,2$ ; son los mayores valores característicos de las matrices  $W'W$  o  $WW'$  [estas matrices tienen los mismos valores característicos (Gabriel, 1978)] y  $\lambda_1$  y  $\lambda_2$  (valores característicos del k-ésimo componente principal) son las raíces cuadradas positivas de los valores  $\lambda_k^2$ .

## Pruebas de significancia de los componentes principales

### Prueba Conservadora

Para probar la significancia de los componentes principales se puede utilizar la prueba conservadora (resultan "pocos" componentes significativos para un nivel de significancia  $\alpha$  predeterminado) propuesta por Gollob (1968), en la cual se asocian  $a+g-2k-1$  grados de libertad (g.l.) a la suma de cuadrados del componente principal k, que se contrasta con una prueba F con el cuadrado medio del error del análisis de varianza conjunto.

### Prueba exacta

Para la prueba exacta del primer componente principal se pueden utilizar las tablas que presentan Johnson y Graybill (1972) del estadístico  $u_1 = \lambda_1 / \sum \lambda_j$ , y para la prueba del segundo componente las que proporcionan Krishnaiah y Schuurmann (1974) del estadístico  $u_2 = \lambda_2 / \sum \lambda_j$ . Desafortunadamente las tablas no presentan las estadísticas de prueba para los valores de genotipos y ambientes usuales en los ensayos de rendimiento.

### Representación gráfica

Si los dos componentes principales explican una parte importante de la interacción y resultan significativos, se puede ilustrar las diferencias entre genotipos graficando al primer componente principal en el eje de las X y al segundo en el eje de las Z. Las coordenadas del genotipo j, para  $j = 1, 2, \dots, g$ ; serán:

<sup>3</sup>  $f'$  indica la traspuesta del vector o matriz f.

$$X_j = \sum_{i=1}^a h_{1i} W_{ij} \quad (2)$$

$$Z_j = \sum_{i=1}^a h_{2i} W_{ij}$$

Sin embargo, como usualmente el segundo componente principal es difícil de interpretar, puede ser de mayor utilidad en el análisis de resultados la gráfica en la que se representa la media del genotipo en el eje de las X y el primer componente principal en el eje de las Z (Zobel *et al.*, 1988).

### EJEMPLO NUMERICO

En este ejemplo se usará la información de seis genotipos del programa de algodone-ro del CEVY - INIFAP y dos testigos co-merciales evaluados en 11 ambientes de prueba en el período 1987-1992 (Cuadro 1) en las principales zonas productoras de algodón del norte de México, la variable a evaluar es el rendimiento en algodón pluma en kg/ha (Cuadro 2).

Cuadro 1. Ambientes y Genotipos en el estudio de la interacción genotipo-ambiente en algodone-ro en México.

Genotipos	Ambientes
1. CIANO Alamos-92	1. Valle del Yaqui 1987 (VY-87)
2. CIANO Arivechi-90	2. Valle del Yaqui 1988 (VY-88)
3. CIANO Cocorim-92	3. Valle del Yaqui 1989 (VY-89)
4. CIANO Cubachi-86	4. Valle del Yaqui 1990 (VY-90)
5. Deltapine 80	5. Valle del Yaqui 1991 (VY-91)
6. Deltapine 90	6. Valle del Yaqui 1992 (VY-92)
7. CIANO Tajimaroa-92	7. Valle de Juárez 1990 (JUA-90)
8. CIANO yaquimi-86	8. Valle de Juárez 1991 (JUA-91)
	9. Valle de Mexicali 1990 (MEX-90)
	10. Valle de Mexicali 1991 (MEX-91)
	11. Costa de Hermosillo 1991 (HIL-91)

Cuadro 2. Rendimientos medios de algodón pluma (kg/ha) de ocho genotipos evaluados en 11 ambientes.

Genotipo	A m b i e n t e s											Promedio
	VY-87	VY-88	VY-89	VY-90	VY-91	VY-92	JUA-90	JUA-91	MEX-90	MEX-91	HIL-91	
Alamos	1854	1433	1784	1616	1145	1126	2436	837	999	1543	1834	1510
Arivechi	1935	1328	1432	1849	1171	1121	2165	809	1093	1419	1190	1410
Cocorim	1983	1204	1539	1831	1132	1074	2822	1026	1339	1576	1955	1588
Cubachi	1689	1193	1644	1386	1213	1069	2547	954	1404	1425	1790	1483
Deltapine 80	1590	1270	1488	1410	1129	998	2413	1065	1416	1623	1570	1452
Deltapine 90	1628	1355	1482	1781	1191	952	2185	930	1195	1531	1505	1430
Tajimaroa	2035	1330	1551	1746	1105	948	2296	918	1258	1463	1624	1479
Yaquimi-86	1601	1093	1461	1427	1139	1037	2294	956	1521	1387	1433	1395

## Programa de análisis con el Statgraphics

Para la obtención del Análisis de Varianza correspondiente a la Interacción genotipo-ambiente, se necesita crear un conjunto de datos con el nombre de, por ejemplo, ALGODON, con las variables G (genotipos), A (ambientes) y Y (rendimiento en kg/ha), semejante al que se presenta en el cuadro 3.

Cuadro 3. Conjunto de datos ALGODON.

Programa de Algodonero del  
CEVY - INIFAP.

DATO	A	G	Y
1	1	1	1854
2	1	2	1935
3	1	3	1983
.	.	.	.
8	1	8	1601
9	2	1	1433
.	.	.	.
.	.	.	.
88	11	8	1433

La variable DATO no aparecerá en el archivo, se incluyó en este cuadro sólo para señalar el número total de datos, las instrucciones para la creación de este archivo se presentan a continuación.

### Archivo de datos para el Análisis de Varianza

El Statgraphics consta de 6 menús principales, para la activación de cualquier procedimiento se selecciona con el cursor y se presiona la tecla ENTER, dentro de los procedimientos existen usualmente varias opciones, para cambiar de opción se utiliza el tabulador y en algunas ocasiones la barra espaciadora. Para ejecutar un procedimiento

el programa usualmente indica si se utiliza la tecla F6 o la ENTER, para abandonar un procedimiento y pasar a un menú superior se utiliza la tecla ESC.

Para la creación del archivo ALGODON seleccione A. DATA MANAGEMENT, dentro de este menú active la opción 2. FILE OPERATIONS y siga las siguientes instrucciones:

- Ponga el nombre del archivo (ALGODON) en el campo para el nombre (FILE NAME).
- Presione el tabulador para cambiar al campo de operaciones (DESIRED OPERATIONS).
- Ponga B en el campo de operación.
- Presione F6.

Una vez creado el archivo se procede a editarlo con las siguientes instrucciones:

- El nombre del archivo (ALGODON) permanece en el campo para el nombre.
- Presione el tabulador para cambiar al campo de operaciones y teclee C para editarlo:
- El cursor se sitúa en NAME, escriba G para definir a la variable G (genotipos), si en el siguiente campo aparece N (variable numérica) presione F6 para crear a la variable, si no aparece N cambie al siguiente campo con el tabulador y presione la barra espaciadora hasta que aparezca N y presione F6; en forma similar defina a las variables A (ambientes) y Y (rendimiento), presione la tecla ESC para editar a las variables (escribir sus valores correspondientes), al introducir el último dato se presiona F6 y se selecciona la opción grabar y salir (SAVE AND EXIT), y la tecla ESC tres veces para volver al menú principal, de esta forma se obtiene un conjunto de datos semejante al del cuadro 3.

### Análisis de Varianza

Para obtener las sumas de cuadrados correspondientes a los factores: Ambientes, Genotipos e Interacción Genotipo-Ambiente se analizará a la variable Y del conjunto de datos ALGODON con un modelo de dos factores: A y G , con este propósito pase al menú<sup>4</sup> correspondiente a J. Analysis of Variance, seleccione: 2. Multifactor Analysis of Variance y siga las instrucciones:

- En la opción DATA ponga el nombre de la variable ALGODON.Y (se especifica también el nombre del archivo), esto se puede efectuar tecleando F7 (VARIABLES) y seleccionando con ENTER la variable ALGODON.Y.
- Pase a la opción de factores del modelo (FACTORS).
- Teclee como primer factor en el campo A: ALGODON.A y pase al segundo factor B: con las flechas, teclee ALGODON.G (los factores tambien se pueden seleccionar con F7).
- Presione el tabulador cinco veces para llegar a la opción correspondiente a las interacciones (INTERACTIONS) y escriba 1 (ignorar interacciones mayores a uno).

- Presione F6 y se obtendrá un menú del que se selecciona con ENTER la primera opción ANOVA TABLE para obtener el Cuadro de Análisis de Varianza que se presenta en el Cuadro 4.

- Presione la tecla ESC para obtener de nuevo el menú, seleccione ahora SAVE RESIDUALS para grabar los residuales del modelo que posteriormente se utilizarán. El cursor aparece en FILE: en donde se escribirá el nombre del archivo teclee ALGODON, en la siguiente opción VARIABLE: teclee RESID y presione F6, presione la tecla ESC tantas veces como sea necesario para volver al menú principal.

En el Cuadro 4, ignore las pruebas de significancia, sólo serán útiles las sumas de cuadrados que al multiplicarlas por el número de repeticiones (cuatro) proporcionan las sumas de cuadrados del cuadro 5 que es parte del análisis conjunto en donde se tienen los factores A, R(A) (repeticiones dentro de ambientes), G, GA, ERROR y TOTAL, la suma de cuadrados del residual corresponde a la suma de cuadrados de la interacción genotipo-ambiente. la cual se prueba con ERROR (error conjunto) [Cruz (1991)]. Estos resultados se presentan en el Cuadro 5. En el error conjunto se omitió la suma de

Cuadro 4. Cuadro de Análisis de Varianza para la obtención de la suma de cuadrados de la Interacción genotipo-ambiente.

FACTORES	G.L.	S.C	C.M.	Fc.
G	7	295504	42215	
A	10	13226344	1322634	
RESIDUAL	70	1497966	21399	
TOTAL	87	15019814		

<sup>4</sup> Las instrucciones del paquete computacional se presentan en Mayúsculas y Subrayadas.

Cuadro 5. Prueba de Significancia de la Interacción Genotipo-Ambiente.

FACTORES	G.L.	S.C	C.M.	Fc.	$F_{\alpha=0.05}$
G	7	1182016	168859	3.75*	2.01
GA	70	5991860	85598	1.90*	1.30
ERROR	210	9448110	44991		

cuadrados del error correspondiente a la localidad JUA-90, debido a su gran magnitud (representaba más del 50 % de la suma de cuadrados del error de todos los ambientes).

#### Archivo de datos para el Modelo AMMI

Para aplicar al modelo AMMI en el Statgraphics se necesita un archivo de datos, por ejemplo ALGOCOM, que contenga como variables a los residuales de las observaciones al ajustar genotipos, esto es, a  $Y_{ij} - Y_j$  agrupadas por ambientes (debido a que el procedimiento de Componentes Principales del Statgraphics en forma automática ajustará por ambientes y operará con los residuales  $W_{ij} = Y_{ij} - Y_{i.} - Y_{.j} + Y_{..}$ ).

Para formar el archivo ALGOCOM se puede crear un archivo semejante al que se presenta en el Cuadro 1 (sean  $Y_1$  a  $Y_{11}$  las medias por ambientes y MEDIA la última variable con las medias de genotipos por ambientes) con instrucciones semejantes a las del apéndice 1, los residuales  $Y_{ij} - Y_j$  se pueden obtener con la opción 2. FILE OPERATIONS del DATA MANAGEMENT con las instrucciones:

File name: ALGOCOM

OPERACION: J. UPDATE y presione F6

Sitúe el cursor sobre la primera variable  $Y_1$  seleccione A (ASIGNE), ENTER  
ASSIGMENT: Y1-MEDIA

ENTER, el programa pedirá un comentario (COMMENT) que se puede evitar con ESC, efectúe la misma operación con las siguientes variables y por último elimine a la variable MEDIA con la opción E (ERASE), de esta forma se obtiene el archivo necesario para aplicar el modelo AMMI.

El archivo ALGOCOM se puede formar también actualizando al archivo ALGODON para no introducir nuevamente a las medias con las siguientes instrucciones:

Copiar el archivo ALGODON con el nombre ALGOCOM, en la opción 2. FILE OPERATIONS, el archivo de datos del Cuadro 1 se formará eliminado observaciones de la variable Y por ejemplo, para definir a la variable Y1 (los rendimientos en el ambiente 1) se eliminarán todas las observaciones numeradas de la 9 en adelante, para la variable Y2 se eliminarán de la 1 a la 8 y de la 17 en adelante etc., con las instrucciones:

File name: ALGODON

Operación: A. COPY

Entrar nuevamente la opción 2. FILE OPERATION con:

FILE NAME: ALGOCOM

OPERACION: J. UPDATE

Sitúe el cursor sobre la última variable Y presione N (crear una nueva variable), nombre: Y1

ASSIGMENT: -80 DROP Y (elimin. las últimas 80 observaciones), presione ESC para evitar el comentario.

Cree la variable Y2, ASSIGMENT: -72 DROP Y , evite el comentario, como se puede observar falta eliminar a las primeras 8 observaciones, esto se obtiene seleccionando a la variable Y2, seleccione A = ASSIGN.

ASSIGMENT: 8 DROP Y2 (se eliminan las primeras ocho observaciones para tener solo a las del ambiente 2.

Proceda en forma similar para generar las 9 variables de ambientes, elimine a las variables G, A y Y, e introduzca una nueva variable con el nombre de MEDIA en el módulo de edición de archivos, situando el cursor en la variable Y9 y oprimiendo F5 seleccione ADD NEW COLUMN (Añadir una Nueva Columna), cambie al módulo de actualizar archivos y a cada variable Y réstelle la MEDIA (como se indicó anteriormente), una vez obtenidas todas las diferencias elimine la variable MEDIA y se tendrá el archivo necesario para el procedimiento de Componentes Principales.

### Análisis con el modelo AMMI

Pase a la opción Q del menú: Multivariate Methods del Capítulo ADVANCED PROCEDURES:

Seleccione 4. PRINCIPAL COMPONENTS y pase a la opción:

DATA VECTORS OR FILENAME: teclee ALGOCOM

STANDARDIZE: NO, presione F6 y aparecerán los porcentajes de la interacción genotipo-ambiente explicadas por los diversos componentes (Cuadro 6), con ESC aparecen diversas opciones, seleccione SAVE EIGENVALUES (los eigenvalores se presentan también en el Cuadro 6) y SAVE PRINCIPAL COMPONENTS (los compo-

nentes principales se presentan en el Cuadro 7) con los nombres RES.EIGEN y RES.COMP respectivamente, se puede obtener una gráfica de los genotipos utilizando los dos mayores componentes principales (Scatterplot), sin embargo, debido a que en este procedimiento las variables se estandarizan automáticamente y a que usualmente el segundo componente principal es difícil de interpretar no se presenta esta gráfica.

Cuadro 6. Eigenvalores y Porcentajes de Variación.

Número	EIGEN-valor	Porcentaje
1	95147.7	44.46
2	61613.4	28.80
3	30190.3	14.11
4	14000.9	6.54
5	7736.1	3.62
6	4656.0	2.18
7	615.7	0.29
8	0.99	0.00
9	0.01	0.00
10	$2.71 \times 10^{-3}$	0.00
11	$1.2 \times 10^{-4}$	0.00

### Gráfica de Genotipos

Para obtener la gráfica de genotipos que usualmente se presenta al aplicar el modelo AMMI se necesita multiplicar a la traspuesta de la matriz de componentes principales por la matriz de residuales al ajustar genotipos y ambientes (que contenga a genotipos como columnas).

Para efectuar esta multiplicación es necesario definir en el archivo ALGODON una nueva variable RM (matriz de residuales), con la definición 11 8 RESHAPE RESID (en la opción J. UPDATE). Para efectuar la multiplicación matricial es necesario cargar los programas de operación con matrices, esto se logra tecleando F8, aparecerá el

Cuadro 7. Eigenvectores de la matriz analizada

A	EV1	EV2	EV3	EV4	EV5	EV6	EV7	EV8	EV9	EV10	EV11
1	0.3529	0.2712	0.3631	-0.4739	-0.4021	0.3650	-0.1204	-0.1822	0.2793	0.1559	-0.0246
2	0.2893	-0.0408	-0.4701	0.1741	-0.1433	0.2051	-0.2925	-0.3111	-0.5156	0.3692	-0.2466
3	-0.0238	0.1091	-0.503	-0.3120	-0.1396	-0.1594	0.2677	0.3589	0.2640	0.1610	-0.5467
4	0.5521	0.0878	0.3841	0.3604	0.1266	-0.4629	0.2361	-0.0175	0.0570	0.2269	-0.2626
5	0.0598	-0.2464	-0.1567	-0.1011	0.1654	-0.3854	-0.6574	-0.2358	0.4864	0.0229	-0.0156
6	0.0756	-0.1223	-0.1370	-0.4062	0.5548	0.0455	0.4069	-0.3605	0.0140	0.3575	0.2493
7	-0.3915	0.3303	0.3251	-0.0195	0.4940	0.1983	-0.2785	0.0432	-0.0748	0.1995	-0.4727
8	-0.1494	-0.2738	0.0818	0.2570	-0.1200	0.1851	-0.0593	0.3997	0.2358	0.7030	0.2613
9	-0.3673	-0.5720	0.3216	-0.1406	-0.3217	-0.1512	0.1781	-0.3352	-0.1174	0.0744	-0.3536
10	0.0095	-0.0993	-0.1443	0.4688	0.0779	0.4890	0.2256	-0.3267	0.5083	-0.1864	-0.2315
11	-0.4080	0.5554	-0.1180	0.1916	-0.2815	-0.3146	0.1054	-0.4179	0.1145	0.2436	0.1951

mensaje **ENTER COMMAND** teclee LOAD y seleccione **MATRIX OPERATIONS**, seleccione **READ**. A continuación actualice el archivo RES creando una nueva variable con el nombre TRAS con la definición TRANSPOSE COMP (la matriz TRAS será la traspuesta de la matriz de vectores característicos), para obtener las coordenadas de genotipos se crea una nueva variable COORD con la definición:

#### TRAS MATMULT ALGODON.RM

Las fileras de la matriz COORD así obtenida contienen a las coordenadas de los genotipos. En el Cuadro 8 se presentan como columnas las dos primeras fileras que corresponden a los dos componentes principales mayores, graficando las coordenadas correspondientes al primer componente principal en el eje Z con las medias de genotipos en el eje X se obtiene la gráfica usual para el modelo AMMI (Figura 1). Algunos paquetes estadísticos presentan a estas coordenadas ajustadas con media igual a cero (las coordenadas del primer componente ajustadas por su media se presentan en la tercera columna del Cuadro 8). Con este ajuste la distribución de los genotipos en la gráfica no se altera.

Cuadro 8. Coordenadas de los genotipos en estudio.

Primera	Segunda	Primera ajustada
71.9527	346.2940	-5.98890
545.5920	-132.1910	467.6504
-181.9570	336.5120	-259.8986
-372.5080	26.3184	-450.4496
-255.0920	-213.1490	-333.0336
203.9570	-122.3850	126.0154
199.9560	91.9561	122.0144
-211.9000	-333.3550	-289.8416

#### Programa de análisis con el SAS

En este programa se supondrá que el usuario tiene los conocimientos necesarios para la definición de archivos y para la utilización del procedimiento ANOVA, si el lector carece de estos conocimientos se recomienda la consulta del trabajo de Martínez (1983). El cuadro 4 se obtiene con un análisis en "bloques al azar" definiendo a Genotipos como tratamientos y a Ambientes como bloque, con el procedimiento **PROC ANOVA**, para la obtención del modelo

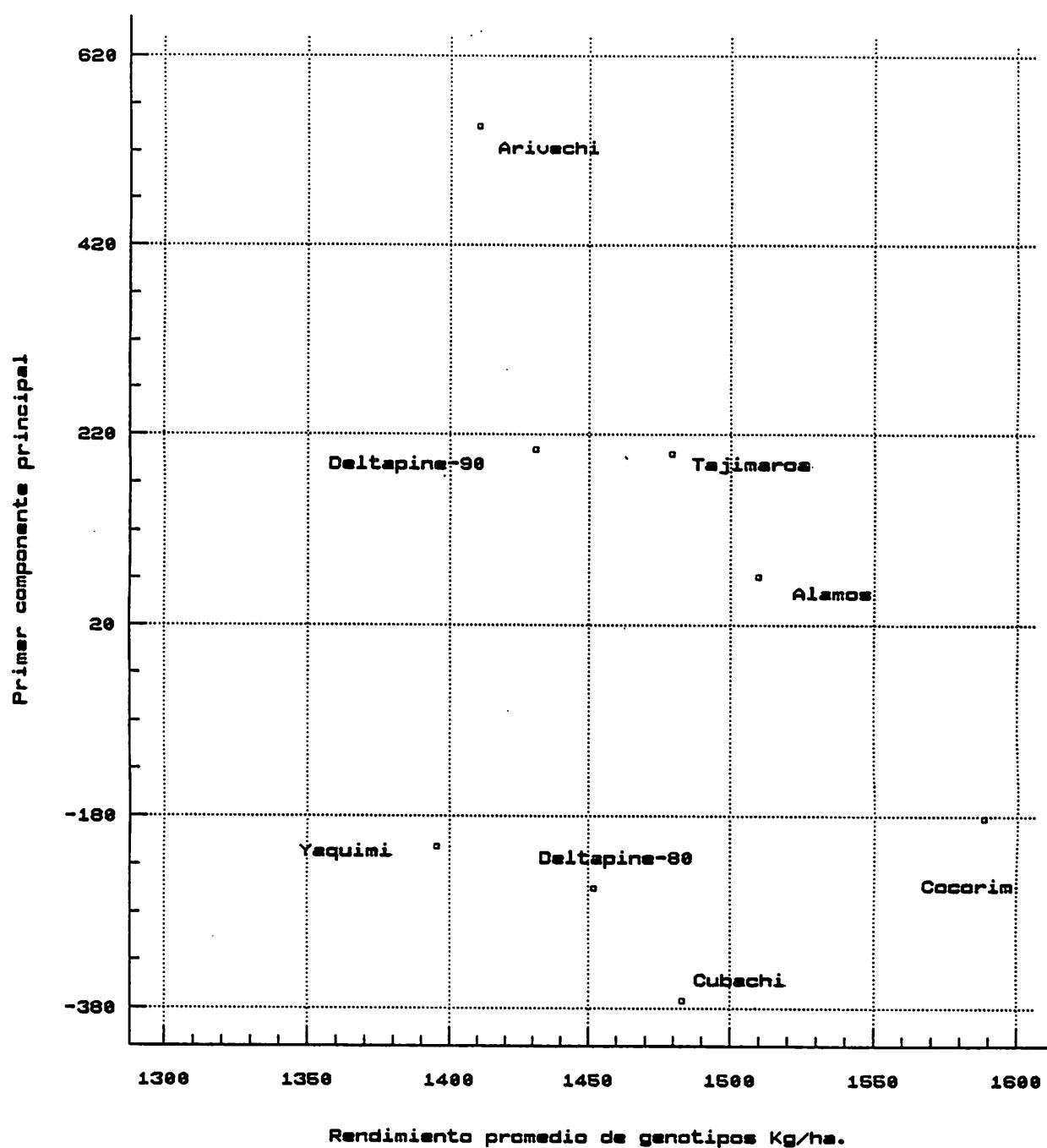


Figura 1. Gráfica de genotipos con el primer componente principal.

AMMI se necesita un archivo que defina a ambientes como variables, esto es, un conjunto de datos semejante al del Cuadro 2.

### Programa para la obtención del modelo AMMI

```
DATA ALGODON;
INPUT V1-V12;
V1=V1-V12;V2=V2-V12;V3=V3-V12;V4=V4-
V12;V5=V5-V12;
V6=V6-V12;V7=V7-V12;V8=V8-V12;V9=V9-
V12;
V10=V10-V12;V11=V11-V12;
CARDS;
1854 1433 1784 1616 1145 1126 2436 837 999 1543
1834 1510
1935 1328 1432 1849 1171 1121 2165 809 1093
1419 1190 1410

1601 1093 1461 1427 1139 1037 2294 956 1521
1387 1433 1395
PROC PRINCOMP OUT=B COVARIANCE N=2;
VAR V1-V11;
PROC PRINT DATA=B;
PROC PLOT;PLOT PRIN1*V12='*';RUN;
```

En este programa se han introducido 12 variables, la última variable representa a las medias de genotipos. Después de la definición de variables, los datos se ajustan por las medias de genotipos ( $V_i = V_i - V_{12}$  para  $i=1,2..11$ ) porque para obtener el modelo AMMI se requiere trabajar con los residuales  $W_{ij}$ , y como el procedimiento de componentes principales **PROC PRINCOMP** ajusta automáticamente por las medias de las variables (ambientes), solo se requiere el ajuste por medias de genotipos.

En el programa de componentes principales **PROC PRINCOMP** se deben solicitar las coordenadas de ambientes en un archivo de nombre B con la opción **OUT=B**. Con la instrucción **COVARIANCE** se pide que el análisis se efectúe con la matriz de covarianzas, con el fin de obtener las sumas de cuadrados asociadas a cada componente principal. La instrucción **N=2** solicita que

sólo se calculen dos componentes (esta opción puede eliminarse). Con la instrucción **VAR V1-V11** se pide que en el análisis sólo se incluyan las 11 primeras variables (la variable V12 sólo se utiliza para ajustar a los datos por las medias de genotipos). Con la última instrucción se solicita la gráfica de la Figura 1.

En los resultados se cumplirá la relación:

$$SCGA = r(g-1)(\text{varianza total}) = 4 \times 7 \times 213571.875 = 598012.5 \text{ y}$$

los eigenvalores serán aproximadamente iguales a los obtenidos con el Statgraphics). Los eigenvectores pueden tener signos opuestos pero con los mismos valores absolutos debido a que en el producto matricial de la descomposición singular de la matriz W se pueden intercambiar los signos de los factores, por ejemplo:  $W = H \Lambda F' = (-H) \Lambda (-F')$  (vea el párrafo siguiente). Las coordenadas de los genotipos las proporciona el SAS como PRIN1, PRIN2, etc.

### Operaciones matriciales para la obtención del modelo AMMI

Una vez obtenida la matriz W de residuales obtenga la descomposición singular de la matriz digamos:

$$W = H \Lambda F'$$

donde:

H es una matriz de orden  $a \times r(W)$ <sup>5</sup> que tiene como columnas a los vectores  $h_k$  definidos en la ecuación (1).

$\Lambda$  es una matriz de orden  $r(W) \times r(W)$  que tiene como elementos diagonales a los vectores característicos  $\lambda_k$  con  $k=1,2,\dots,r(W)$ .

F es una matriz de orden  $r(W) \times g$  que tiene como columnas a los vectores  $f_k$ .

<sup>5</sup>  $r(W)$  indica el rango de matriz W que es el mínimo de  $(a-1, g-1)$ .

Si el paquete matricial que se utiliza no tiene a la función para la descomposición singular de la matriz obtenga las matrices  $WW'$  y  $W'W$ .

Los valores característicos diferentes de cero de  $WW'$  y  $W'W$  (estas matrices tienen los mismos valores característicos) serán  $\lambda_k^2$  con  $k=1,2,\dots,r(W)$ .

La matriz  $H$  tendrá como columnas los  $r(W)$  vectores característicos de  $WW'$  asociados con los valores característicos  $\lambda_k^2$ .

La matriz  $F$  tendrá como columnas los  $r(W)$  vectores característicos de  $W'W$  asociados con los valores característicos  $\lambda_k^2$ .

Las coordenadas de los genotipos son las hileras de la matriz  $C$  obtenida con el producto:  $C = W'H$ .

Se notará que los valores característicos  $\lambda_k^2$  son  $(g-1)$  veces los valores característicos obtenidos con el SAS o el STATGRAPHICS ya que estos paquetes obtienen los valores característicos de la matriz de covarianzas que es  $[1/(g-1)]WW'$ .

## RESULTADOS Y DISCUSION

Como el modelo AMMI no sería de utilidad si el primer componente principal no resultara significativo son necesarias las:

### Pruebas de Significancia

Suponiendo que ambientes es un factor aleatorio, el primer componente principal se puede probar con la estadística:

$$U_1 = \frac{\lambda_1}{\sum \lambda_j} = 0.4446$$

El valor de este cociente se presenta como porcentaje en el Cuadro 6 (porcentaje

explicado por el primer componente). Como el valor crítico para  $\alpha = 0.05$  es  $U_{\alpha} = 0.4334$  [Este valor se obtuvo por interpolación de la Tabla 1 de Johnson y Graybill (1972)], y como  $U > U_{\alpha}$ , se acepta la significancia del primer componente principal; Las tablas que presentan Krishnaiah y Schurmann (1974) para la prueba del segundo componente principal resultan incompletas para este caso.

Otra forma de efectuar las pruebas de significancia de los componentes principales es utilizando las pruebas aproximadas de Gollob (1968) en las que al componente principal  $k$  se le asignan  $a + g - 2k - 1$  grados de libertad y se le contrasta con el error combinado [esta prueba es la utilizada por Gauch(1988)].

### Interpretación de los componentes

Para la utilización del modelo AMMI en la clasificación de genotipos es necesaria la interpretación de los componentes principales. Con este fin es necesario analizar cuidadosamente los elementos de los eigenvectores.

Observe que el primer componente se puede interpretar como un contraste (la suma de los elementos del vector es igual a cero) entre el ambiente del Valle del Yaqui y los ambientes restantes de prueba ya que los signos de los primeros seis términos del primer eigenvector EV1 (Cuadro 7) resultaron positivos a excepción del tercer término (VY-89) que es negativo pero cercano a cero y los cinco últimos términos del vector correspondientes a los ambientes fuera del Valle del Yaqui resultaron negativos a excepción del décimo (MEX-91) que resultó positivo pero cercano a cero.

Por esta razón en la gráfica usual del modelo AMMI, en la que el promedio general de cada genotipo se representa en el eje X y el primer componente en el eje Z, los

genotipos más alejados del eje Z serán los de mayor rendimiento general y los más distantes, hacia arriba del eje X tendrán una mayor adaptación (mayor rendimiento) en el Valle del Yaqui con respecto a los ambientes restantes.

El segundo componente principal también es un contraste pero resulta difícil de interpretar.

El modelo AMMI utilizado como menciona Cruz (1992), siempre explica una suma de cuadrados mayor a la del método de ajuste por regresión, con la desventaja de introducir un mayor número de parámetros que en algunas ocasiones pueden ser de difícil interpretación en este ejemplo, sin embargo, el primer componente principal sí permite una interpretación clara.

## BIBLIOGRAFIA

- Crossa, J., H. G. Gauch, and R. W. Zobel. 1990.** Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. *Crop Science* 30:493-500.
- Cruz Medina, R. 1990.** Análisis estadístico de la interacción genotipo-ambiente. ITA 21, DGTAM. SEP. México. 210 pp.
- \_\_\_\_\_. 1992. Generalización de modelos para el análisis de la interacción genotipo-ambiente. *Revista Fitotecnia Mexicana*. 15:149-158.
- Gabriel, K. R. 1978.** Least squares approximation of matrices by additive and multiplicative models. *J. Royal Stat. Soc.(B)* 2:186-196.
- Gauch, H. G. 1988.** Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.

**Gollob, H. F. 1968.** A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 33:73-116.

**Hill, J. 1975.** Genotype-environment interactions a challenge for plant breeding. *J. Agric. Sci.* 85:447-493.

**Johnson, D. E., and F. A. Graybill. 1972.** An Analysis of a two-way model with interaction and no replication. *J. Amer. Stat. Assoc.* Vol. 67: 862-869.

**Krishnaiah, P. R., and F. J. Schuurmann. 1979.** On the exact distribution of the ratios of the individual roots to the trace of the complex and real Wishart matrices. *Aerospace Research Laboratories* 74-0026 USAF.

**Lin, C. S., and G. Butler. 1990.** Cluster analysis for analyzing two-way classification data. *Agron. J.* 82:344-348.

**Martínez G., A. 1983.** Introducción al SAS. Colegio de Postgraduados CEC. México. 168 pp.

**SAS User's Guide Version 5. 1987.** SAS Institute Inc. N.C. USA.

**Statgraphics User's Guide Version 5. 1991.** Statistical Graphics Corporation. USA.

**Yates, F., and W. G. Cochran. 1938.** The analysis of group of experiments *J. Agric. Sci.* 23: 556-580.

**Zobel, W.R., M. J. Madison and H. G. Gauch. 1988.** Statistical analysis of a yield trial. *Agron. J.* 80:388-393.

**En el proceso editorial del Vol. 17 Núm. 1 de la Revista Fitotecnia Mexicana colaboraron las siguientes personas:**

**TIPOGRAFIA COMPUTARIZADA:**

**Rosario Sabbagh Estrada**

**DIBUJO:**

**Moisés Aguilar Castillo**

**COORDINACION GENERAL:**

**Dr. José Luis Rodríguez Ontiveros**