

REPRESENTACION CANONICA EN MANOVA: APLICACION A UNA CLASE DE DISEÑO ANIDADO

J. M. OLLER, C. M. CUADRAS

Se aplican ciertas técnicas de decisión propias del análisis multivariante para resolver y presentar la variabilidad de un diseño experimental de uso frecuente. Se plantea el modelo matemático asociado al diseño y se resuelven con detalle los problemas de estimación de parámetros y pruebas de significación. Los niveles de los factores del diseño se interpretan como funciones paramétricas estimables y son representados por utilización de análisis canónico. Los resultados obtenidos se aplican a la comparación y representación de tratamientos en una experiencia agrícola.

1. INTRODUCCION.

En experiencias agrícolas y en farmacología es frecuente estudiar el efecto de determinados tratamientos sobre distintos individuos y en períodos diferentes, en base a algunas variables observables.

La experiencia agrícola que ha motivado este trabajo es una investigación realizada sobre el efecto que 6 tratamientos (uno de ellos - testigo) tenían sobre la recogida de la aceituna. Cada tratamiento se aplicó a 3 árboles distintos y se estudió la caída del fruto a lo largo de 4 semanas. El efecto de los tratamientos consistía en disminuir la sujeción del fruto en el árbol, para lo cual se midieron dos variables:

Y_1 = fuerza de abscisión,
 Y_2 = porcentaje de aceituna caída.

La disposición de los datos se encuentra en la tabla I. Véase /4/.

Por la forma en que se tomaron los datos, es necesario considerar tres factores cualitativos: tratamiento, individuo y semana. Debe introducirse un efecto individual, puesto que podemos considerar de entrada, una notable variabilidad entre árboles frente a un mismo tratamiento. Como además existe una

clara dependencia entre las variables Y_1 , Y_2 , se estudian conjuntamente en el modelo que se expone a continuación.

2. MODELO UTILIZADO.

Consideremos el modelo lineal normal bivariante

$$Y_{ijhq} = \mu_q + \alpha_{iq} + \beta_{ijq} + \lambda_{hq} + e_{ijhq} \quad (1)$$

$$(i=1, \dots, n; j=1, \dots, m; h=1, \dots, p; q=1, 2)$$

con las siguientes $2(n+2)$ restricciones lineales:

$$\alpha_{.q} = \beta_{i.q} = \lambda_{.q} = 0$$

$$q = 1, 2 \quad i = 1, \dots, n \quad (2)$$

Se trata de un modelo de tres factores. En nuestro caso concreto μ_q será la media general de la variable q , el efecto del tratamiento vendrá caracterizado por los α_{iq} , con $n=6$, el efecto individual estará determinado por los β_{ijq} , con $m=3$, y el efecto semana vendrá dado por los λ_{hq} , con $p=4$. El factor individuo se halla anidado respecto del factor tratamiento y el factor semana está cruzado respecto a ambos.

- J.M.Oller - Universitat de Barcelona - Fac. de Biología-Bioestadística - Barcelona
- Carles M. Cuadras - Universitat de Barcelona - Fac. de Biología-Bioestadística -Av. Diagonal 637 -Barcelona
- Article rebut el Gener de 1983.

El modelo puede expresarse matricialmente - como:

$$Y = X\theta + E \quad (3)$$

donde Y es una matriz de nmp filas y 2 columnas, ocupando el elemento y_{ijhq} la fila $---$ $(i-1)mp + (j-1)p + h$, columna q. θ es una matriz de $nm+p-1$ filas y 2 columnas, ocupando μ_q la primera fila, columna q, α_{iq} la $i+1$ fila, columna q, β_{ijq} la $n+(m-1)(i-1)+j$ fila, columna q y λ_{hq} la $nm+h$ fila, columna q. E es una matriz de nmp filas y 2 columnas, de errores aleatorios, normales e independientemente distribuidos por filas. X es una matriz de nmp filas y $nm+p-1$ columnas, la matriz del diseño. Para manejar con cierta comodidad la matriz X, la fraccionaremos, por columnas, - en m+3 matrices, M, A, B_1, \dots, B_m, C , definidas de la siguiente forma:

M (nmp, 1)

$$m_{k1} = 1 \quad k=1, \dots, nmp$$

A (nmp, n-1)

$$a_{kq} = \begin{cases} 1 & \text{si } (q-1)mp < k \leq qmp \\ -1 & \text{si } (n-1)mp < k \\ 0 & \text{en los demás casos} \end{cases}$$

B_i (nmp, m-1) $i=1, \dots, n$

$$b_{kq}^i = \begin{cases} 1 & \text{si } (i-1)mp + p(q-1) < k \leq (i-1)mp + pq \\ -1 & \text{si } (im-1)p < k \leq imp \\ 0 & \text{en los demás casos} \end{cases}$$

C (nmp, p-1)

$$c_{kq} = \begin{cases} 1 & \text{si } k \in \{\dot{p}+q\} \\ -1 & \text{si } k \in \{\dot{p}\} \\ 0 & \text{en los demás casos} \end{cases}$$

X es entonces:

$$X = (M, A, B_1, \dots, B_n, C)$$

$$X'X = \begin{pmatrix} M'M & 0 & & & 0 \\ 0 & A'A & & & \\ \cdot & & B_1'B_1 & & \\ \cdot & & & \cdot & \\ \cdot & & & & B_n'B_n \\ 0 & \dots & \dots & \dots & C'C \end{pmatrix} \quad (5)$$

verificándose, si llamamos $R=M'M$, $S=A'A$, --- $T_k=B_k'B_k$ y $V=C'C$,

$$R(1,1) \quad r_{11} = nmp$$

$$S(n-1, n-1) \quad s_{ij} = \begin{cases} 2mp & i=j \\ mp & i \neq j \end{cases} \quad (6)$$

$$T_k(m-1, m-1) \quad t_{ij} = \begin{cases} 2p & i=j \\ p & i \neq j \end{cases} \quad k=1, \dots, n$$

$$V(p-1, p-1) \quad v_{ij} = \begin{cases} 2nm & i=j \\ nm & i \neq j \end{cases}$$

Es inmediato comprobar que una matriz Z definida por:

$$z_{ij} = 1 + \delta_{ij} \quad i, j = 1 \dots n \quad (7)$$

donde δ_{ij} son los deltas de Kronecker, verifica:

$$|Z| = n+1 \quad (8)$$

$$Z^{-1} = \left(\delta_{ij} - \frac{1}{n+1} \right)$$

por tanto el rango de X, igual al rango de $X'X$, es máximo, es decir, igual al número de parámetros:

$$\begin{aligned} \text{rango}(X'X) &= \text{rango}(R) + \text{rango}(S) + \text{rango}(T_1) + \dots + \\ &+ \text{rango}(T_n) + \text{rango}(V) = 1 + (n-1) + n(m-1) + (p-1) = \\ &= nm + p-1 \end{aligned} \quad (9)$$

3. ESTIMACION POR MINIMOS CUADRADOS DE LOS PARAMETROS.

A partir de X, como paso previo para el planteo de las ecuaciones normales /5/, hemos de obtener $X'X$ resultando ser:

c) Significación del efecto semana

$$H_C : \lambda_{hq} = 0 \quad h=1, \dots, p \\ q=1, 2 \quad (\text{hipótesis nula})$$

$$H : \exists \lambda_{rs} \neq 0 \quad (\text{hipótesis alternativa})$$

El rango de la matriz asociada a esta hipótesis lineal es $p-1$. La matriz de dispersiones bajo H_C , viene dada por:

$$c_{qs} = \sum_{ijh} (y_{ijhq} - y_{ij,q})(y_{ijhs} - y_{ij,s}) \quad (19) \\ q, s=1, 2$$

El criterio de decisión se basa en el valor que toma el estadístico:

$$F_C = \frac{(p-1)(nm-1)-1}{p-1} \left[\sqrt{\frac{c_{11} \cdot c_{22} - c_{12}^2}{r_{11} \cdot r_{22} - r_{12}^2}} - 1 \right] \quad (20)$$

que se distribuye según una distribución F de Fisher-Snedecor con $2(p-1)$ y $2((p-1)(nm-1)-1)$ grados de libertad.

5. REPRESENTACION CANONICA DE FUNCIONES PARAMÉTRICAS ESTIMABLES.

Cuando los niveles de determinados factores de un diseño factorial son significativamente diferentes, tiene interés analizar las diferencias por representación canónica de ciertas funciones paramétricas estimables. Esta técnica ha sido formulada por Cuadras /2/. Véanse aplicaciones (genética, sistemática, medicina, etc.) en /3/.

Suponiendo que H_A y H_C son hipótesis que han sido rechazadas, consideraremos las funciones paramétricas

$$\psi_{iq} = \mu_q + \alpha_{iq} \quad q=1, 2 \quad i=1, \dots, n \quad (21)$$

$$\phi_{hq} = \mu_q + \lambda_{hq} \quad q=1, 2 \quad h=1, \dots, p$$

que dan cuenta del efecto tratamiento y el efecto semana respectivamente.

Las estimaciones de Gauss-Markov son

$$\hat{\psi}_{iq} = \hat{\mu}_q + \hat{\alpha}_{iq} = y_{i..q} \quad i=1, \dots, n \\ \hat{\phi}_{hq} = \hat{\mu}_q + \hat{\lambda}_{hq} = y_{..hq} \quad h=1, \dots, p \quad q=1, 2 \quad (22)$$

Para efectuar la representación canónica de

dichas funciones paramétricas tendremos que definir

$$B_\psi = \begin{pmatrix} y_{1..1} - \bar{y}_{...1} & \dots & y_{1..2} - \bar{y}_{...2} \\ \vdots & & \vdots \\ y_{n..1} - \bar{y}_{...1} & \dots & y_{n..2} - \bar{y}_{...2} \end{pmatrix} \quad (23)$$

$$\Sigma_\psi = B'_\psi B_\psi \quad (24)$$

Σ_ψ refleja la covariabilidad entre las funciones paramétricas ψ_{iq} , en el sentido de Cuadras /2/. Análogamente:

$$B_\phi = \begin{pmatrix} y_{..11} - \bar{y}_{...1} & \dots & y_{..12} - \bar{y}_{...2} \\ \vdots & & \vdots \\ y_{..p1} - \bar{y}_{...1} & \dots & y_{..p2} - \bar{y}_{...2} \end{pmatrix} \quad (25)$$

$$\Sigma_\phi = B'_\phi B_\phi \quad (26)$$

Además /3/:

$$\hat{\Sigma} = \frac{1}{(nm-1)(p-1)} \mathbb{R} \quad (27)$$

es una estimación insesgada de Σ (matriz de varianzas y covarianzas de las dos variables observables y_1, y_2), siendo \mathbb{R} la matriz definida en (14). A continuación deberemos encontrar una base ortogonal respecto a Σ_ψ (o a Σ_ϕ) y además ortonormal respecto a $\hat{\Sigma}$. El problema se resuelve determinando los valores y vectores propios de $\Sigma_\psi (\delta \Sigma_\phi)$ respecto a $\hat{\Sigma}$. Tendremos pues:

$$|\Sigma_\psi - \lambda_i \hat{\Sigma}| = 0 \quad \begin{cases} \Sigma_\psi v_i = \lambda_i \hat{\Sigma} v_i \\ v_i' \hat{\Sigma} v_j = \delta_{ij} \end{cases} \quad (28)$$

y análogamente:

$$|\Sigma_\phi - \lambda_i \hat{\Sigma}| = 0 \quad \begin{cases} \Sigma_\phi v_i = \lambda_i \hat{\Sigma} v_i \\ v_i' \hat{\Sigma} v_j = \delta_{ij} \end{cases} \quad (29)$$

Los vectores propios así obtenidos definen las direcciones de máxima variabilidad respecto a la métrica definida por $\hat{\Sigma}$.

Si llamamos V_ψ a la matriz formada por los vectores propios obtenidos en (28) (dispuestos en columnas) y análogamente llamamos V_ϕ

a la matriz formada por los vectores propios obtenidos en (29), estas matrices definen -- las transformaciones de coordenadas buscadas, de forma que el primer eje absorberá la máxima variabilidad. Las matrices con las nuevas coordenadas de las funciones paramétricas -- vendrán dadas por:

$$\tilde{B}_\psi = V_\psi B_\psi \quad \tilde{B}_\phi = V_\phi B_\phi \quad (30)$$

6. REGIONES CONFIDENCIALES PARA LAS FUNCIONES PARAMÉTRICAS ESTIMABLES.

Es conveniente, a continuación, determinar -- las regiones confidenciales de las estimaciones de las funciones paramétricas estimables. Para ello hay que obtener previamente la expresión de las mismas en función de las medias de las casillas del diseño:

$$\hat{\psi}_{iq} = y_{i..q} = \sum_{j=1}^m \sum_{h=1}^p \frac{1}{mp} y_{ijhq} \quad (31)$$

$$\hat{\phi}_{hq} = y_{..hq} = \sum_{i=1}^n \sum_{j=1}^m \frac{1}{nm} y_{ijhq}$$

Las regiones confidenciales, con coeficiente de confianza $1-\alpha$, vendrán dadas, /2/, por:

$$||\hat{\psi}_{i..q} - \psi_{i..q}|| < \frac{R_\alpha}{\sqrt{mp}} \quad ||\hat{\phi}_{i..h} - \phi_{i..h}|| < \frac{R_\alpha}{\sqrt{nm}} \quad (32)$$

siendo

$$(R_\alpha)^2 = F_\alpha \frac{2(nm-1)(p-1)}{(nm-1)(p-1)-1} \quad (33)$$

$$\text{Prob} \left[F > F_\alpha \right] = \alpha \quad (34)$$

donde F sigue una distribución F con 2 y --- $(nm-1)(p-1)-1$ grados de libertad. Cada región confidencial (32) está representada por un -- círculo referido a las variables canónicas -- (ver Fig. 1 y Fig. 2).

7. CORRELACION VARIABLES OBSERVABLES CON EJES CANONICOS.

Es interesante hallar las correlaciones de -- las variables observables, en nuestro ejemplo fuerza de abscisión y porcentaje de aceituna -- caída, con las variables canónicas, para una posible interpretación de las mismas.

Si llamamos:

$$D^{-\frac{1}{2}} = \text{diag} \left(\frac{1}{\sqrt{\hat{\sigma}_{11}^2}}, \frac{1}{\sqrt{\hat{\sigma}_{22}^2}} \right) \quad (35)$$

siendo $\hat{\sigma}_{ii}^2$ los elementos de la diagonal principal de $\hat{\Sigma}$, la matriz

$$A = D^{-\frac{1}{2}} \hat{\Sigma} V \quad (36)$$

contiene las correlaciones de las variables observables con las variables canónicas.

8. APLICACION.

Aplicando los anteriores resultados a nuestro caso concreto, obtenemos las siguientes matrices de dispersiones:

Sin restricciones:

$$R = \begin{pmatrix} 92861.00 & -3921.28 \\ -3921.28 & 1280.46 \end{pmatrix} \quad (37)$$

Bajo la hipótesis nula H_A (no hay efecto de -- los tratamientos)

$$A = \begin{pmatrix} 602112.56 & -20360.42 \\ -20360.42 & 2334.97 \end{pmatrix} \quad (38)$$

Bajo la hipótesis nula H_B (no hay efecto individuo).

$$B = \begin{pmatrix} 145640.31 & -5855.41 \\ -5855.41 & 1734.69 \end{pmatrix} \quad (39)$$

Finalmente, bajo la hipótesis nula H_C (no hay efecto semana)

$$C = \begin{pmatrix} 189623.00 & -13972.11 \\ -13972.11 & 3544.50 \end{pmatrix} \quad (40)$$

Los valores de los estadísticos sobre los que se toma la decisión, para cada una de estas -- tres hipótesis, son respectivamente:

$$\begin{aligned} F_A &= 21.36 && \text{con 10 y 100 grados de libertad} \\ F_B &= 1.88 && \text{con 24 y 100 grados de libertad} \\ F_C &= 19.10 && \text{con 6 y 50 grados de libertad} \end{aligned} \quad (41)$$

Por lo que rechazamos la hipótesis nula, in-

cluso para niveles de significación tan bajos como 1% en el primer y tercer caso y para niveles de significación del orden del 5% para el segundo.

En cuanto a la representación canónica de las funciones paramétricas estimables, se obtiene como estimación de $\hat{\Sigma}$:

$$\hat{\Sigma} = \begin{pmatrix} 1820.80 & -76.89 \\ -76.89 & 25.11 \end{pmatrix} \quad (42)$$

Para el efecto tratamiento, tendremos:

$$B_{\psi} = \begin{pmatrix} 167.61 & -3.11 \\ 22.69 & -3.29 \\ -104.97 & 7.42 \\ -14.64 & -2.56 \\ -42.31 & -0.75 \\ -28.39 & 2.28 \end{pmatrix} \quad (43)$$

$$\hat{\Sigma}_{\psi} = \begin{pmatrix} 42437.67 & -1369.94 \\ -1369.94 & 87.88 \end{pmatrix}$$

y para el efecto semana:

$$B_{\phi} = \begin{pmatrix} 62.28 & -5.34 \\ -9.44 & -3.60 \\ -23.94 & -0.23 \\ -28.89 & 9.18 \end{pmatrix} \quad (44)$$

$$\hat{\Sigma}_{\phi} = \begin{pmatrix} 5375.61 & -558.38 \\ -558.38 & 125.78 \end{pmatrix}$$

En el proceso de diagonalización, para el efecto tratamiento, resultan los valores propios:

$$\lambda_1 = 23.52 \quad \lambda_2 = 1.98 \quad (45)$$

cuyos vectores propios, de norma unitaria -- respecto a $\hat{\Sigma}$, resultan ser, por columnas:

$$V_{\psi} = \begin{pmatrix} 0.0242 & 0.0067 \\ 0.0211 & 0.2128 \end{pmatrix} \quad (46)$$

Para el efecto semana se obtiene:

$$\lambda_1 = 5.24 \quad \lambda_2 = 1.75 \quad (47)$$

$$V_{\phi} = \begin{pmatrix} 0.0065 & 0.0243 \\ -0.1731 & 0.1256 \end{pmatrix} \quad (48)$$

Efectuando las transformaciones de coordena-

das, para el efecto tratamiento, obtenemos:

$$\tilde{B}_{\psi} = \begin{pmatrix} 3.99 & 0.46 \\ 0.48 & -0.55 \\ -2.38 & 0.88 \\ -0.41 & -0.64 \\ -1.05 & -0.44 \\ -0.64 & 0.30 \end{pmatrix} \quad (49)$$

y para el efecto semana:

$$\tilde{B}_{\phi} = \begin{pmatrix} 1.33 & 0.84 \\ 0.56 & -0.68 \\ -0.12 & -0.61 \\ -1.77 & 0.45 \end{pmatrix} \quad (50)$$

El radio de los intervalos de confianza, para regiones confidenciales al 90%, es:

$$\text{Efecto tratamiento: } r_{\psi} = 0.643 \quad (51)$$

$$\text{Efecto semana: } r_{\phi} = 0.525$$

En cuanto a las correlaciones de las variables con los ejes canónicos debido a que:

$$D^{-\frac{1}{2}} \hat{\Sigma} = \begin{pmatrix} 42.67 & -1.80 \\ -15.34 & 5.01 \end{pmatrix} \quad (52)$$

resulta:

$$A_{\psi} = D^{-\frac{1}{2}} \hat{\Sigma} V_{\psi} = \begin{pmatrix} v_1 & v_2 \\ 0.995 & -0.099 \\ -0.266 & 0.964 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \quad (53)$$

y para el efecto semana:

$$A_{\phi} = D^{-\frac{1}{2}} \hat{\Sigma} V_{\phi} = \begin{pmatrix} v_1 & v_2 \\ 0.587 & 0.809 \\ -0.966 & 0.257 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \quad (54)$$

9. CONCLUSIONES.

En función de los datos de la tabla 1, podemos aceptar que hay diferencias entre tratamientos, entre semanas y una variabilidad -- significativa entre los árboles.

Para visualizar las diferencias entre tratamientos, recurrimos a la representación canónica de las funciones paramétricas ψ_1 . Examinando la Fig. 1, se aprecia que el tratamiento 1 (testigo) se diferencia claramente de los demás, así como el tratamiento 3 (ETHREL 2.500 p.p.m.), que queda en la zona opuesta. Los demás tratamientos (ETHREL 15.000 p.p.m.,

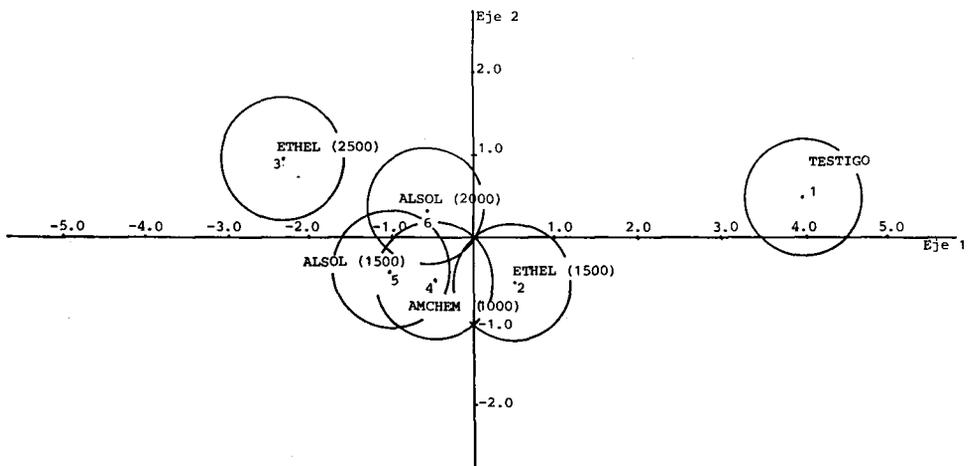


Fig. 1.- Representación canónica de los tratamientos.

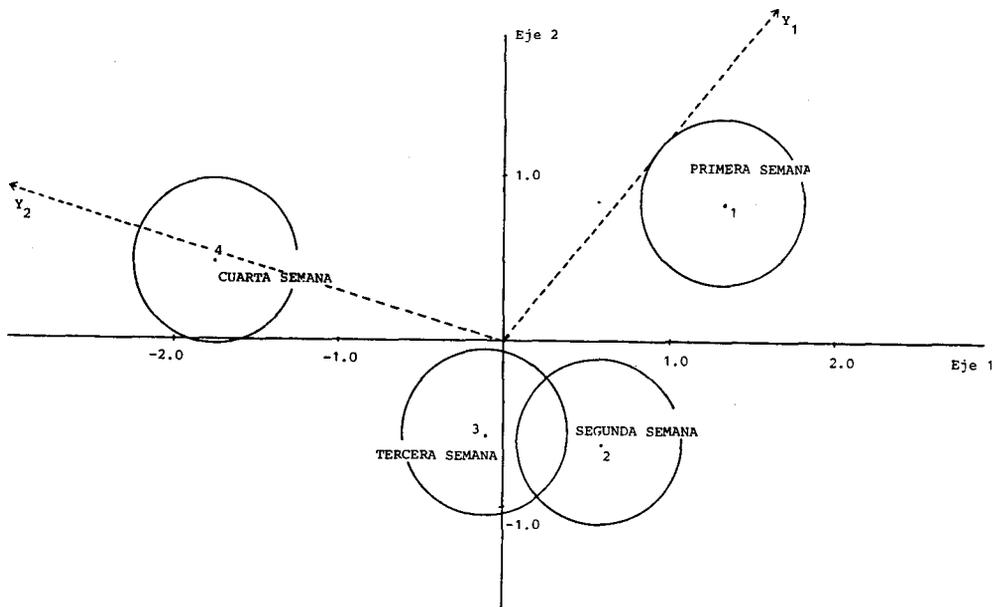


Fig. 2.- Representación canónica de las semanas.

TABLA 1.: FUERZA DE ABCISION Y PORCENTAJE

Tratamiento	Arbol	S e m a n a s			
		I	II	III	IV
Testigo	1	(467 3.05)	(495 4.45)	(488 3.98)	(575 7.32)
	2	(550 1.91)	(450 2.08)	(450 2.84)	(497 9.31)
	3	(556 4.10)	(556 3.28)	(575 5.01)	(497 14.59)
ETHEL (1.500)	1	(400 3.15)	(400 3.04)	(390 3.67)	(357 9.62)
	2	(438 1.44)	(342 1.49)	(357 2.62)	(335 10.43)
	3	(470 1.23)	(331 1.71)	(282 3.98)	(315 17.33)
ETHEL (2.500)	1	(289 4.56)	(195 6.84)	(162 13.24)	(145 44.74)
	2	(386 1.74)	(282 2.17)	(295 4.95)	(187 29.06)
	3	(284 4.74)	(241 16.8)	(207 32.6)	(212 27.07)
AMCHEM (1.000)	1	(429 1.75)	(339 1.86)	(315 3.38)	(288 15.8)
	2	(381 3.16)	(277 4.56)	(275 7.70)	(275 17.27)
	3	(505 1.18)	(333 1.19)	(292 1.83)	(260 8.76)
ALSOL (1.500)	1	(302 3.40)	(250 3.84)	(216 9.36)	(215 18.23)
	2	(467 4.20)	(308 5.90)	(274 10.5)	(337 12.6)
	3	(369 2.24)	(287 3.84)	(302 7.9)	(310 8.22)
ALSOL (2.000)	1	(403 3.11)	(353 4.71)	(299 12.46)	(212 30.20)
	2	(323 5.89)	(301 12.27)	(295 12.76)	(355 19.96)
	3	(319 1.74)	(307 3.9)	(312 5.89)	(325 13.70)

ALSOL 1.500 y 2.000 p.p.m., AMCHEM 1.000 --- p.p.m.) forman un grupo más o menos compacto y están poco diferenciados entre sí.

Puesto que todos los tratamientos difieren del testigo, tienen efectos significativos. Observando la matriz (53), podemos identificar el primer eje con la fuerza de abscisión (Y_1), el segundo eje con el porcentaje de -- aceituna caída (Y_2). Se concluye que ETHEL (2.500) es el que más disminuye Y_1 , facilitando la caída de la aceituna, pero también el porcentaje de aceituna perdida es mayor. De los demás tratamientos, cuya eficacia respecto a Y_1 es parecida, AMCHEM (1.000) es el que goza de una menor pérdida de aceituna, -- por lo que fué considerado como tratamiento más idóneo.

Con la representación canónica de las funciones paramétricas ϕ_i , visualizamos las diferencias temporales. Observando la matriz --- (54), podemos relacionar los ejes canónicos con las variables observables, tal como hacemos en la Fig. 2. La fuerza de abscisión (Y_1) desciende notablemente a partir de la primera semana, y la variación es pequeña para -- las siguientes semanas, mientras que el porcentaje de aceituna caída (Y_2) es mínimo a -- la primera semana y alcanza un valor algo -- más alto a la segunda y tercera semana, ---- aumentando de forma apreciable a la cuarta -- semana. Se consideró entonces, como momento apropiado de recogida, la segunda semana, -- pues se tiene una fuerza de abscisión baja, -- sin haber una pérdida grande de aceituna por caída espontánea.

Tal como ha sido planteado, este modelo así como la técnica de representación utilizada, puede aplicarse a cualquier diseño anidado - p-variable con $p > 2$.

10. BIBLIOGRAFIA.

- /1/ ANDERSON, T.W. "An introduction to multivariate Statistical analysis". Wiley, New York, (1958).
- /2/ CUADRAS, C.M. "Análisis discriminante de funciones paramétricas estimables". Trab.Est. Inv. Op., 25(3), 3-31, (1.973).
- /3/ CUADRAS, C.M. "Métodos de análisis mul-

tivariante". Ed. Eunibar, Barcelona, (1.981).

- /4/ JORBA, J. "Investigaciones sobre el empleo de productos químicos aptos para facilitar la recolección de la aceituna". Tesina de la Fac. de Biología, -- Univ. de Barcelona (inédito), (1.977).
- /5/ SCHEFFE, H. "The Analysis of Variance". Wiley, New York, (1.959).