

El programa CANON para IBM 360(*)

por CARLOS CUADRAS-AVELLANA **, JUAN A. CAMPÁ-VIÑETA ***, y JOAQUÍN MONTORIOL-POUS ***,

RESUMEN

El programa CANON, escrito en lenguaje Fortran IV para IBM 360, lleva a cabo un análisis canónico completo sobre un conjunto de hasta 15 variables y 30 grupos. Puede admitir 99.999 observaciones para cada variable.

RÉSUMÉ

Le programme CANON, écrit en langage Fortran IV pour IBM 360, effectue un analyse canonique complet sur un ensemble jusqu'à 15 variables et 30 groupes. Il peut admettre 99.999 observations pour chaque variable.

INTRODUCCIÓN

El programa CANON realiza un análisis canónico completo sobre un conjunto de hasta 15 variables y 30 grupos. Cada variable puede admitir un número máximo de 99.999 observaciones, que puede ser ampliado con sólo cambiar el formato de entrada. Esta notable capacidad es debida a que los datos son leídos observación a observación y solamente uno de ellos está almacenado en memoria en un momento dado. Los datos se almacenan en un fichero en disco, para ser leídos posteriormente en el programa.

El programa utiliza 11 subrutinas, 5 de ellas procedentes del S.S.P. (1) y las restantes originales del primero que suscribe. Están escritas en lenguaje FORTRAN IV, por su carácter más universal y por disponer de las rutinas necesarias en este lenguaje. Está preparado para el sistema IBM 360 necesitando 64K de memoria y almacenamiento auxiliar de cinta o disco.

* Este trabajo ha sido realizado, en parte, con la Ayuda para el Fomento de la Investigación en la Universidad.

** Laboratorio de Cálculo. Facultad de Ciencias. Universidad de Barcelona.

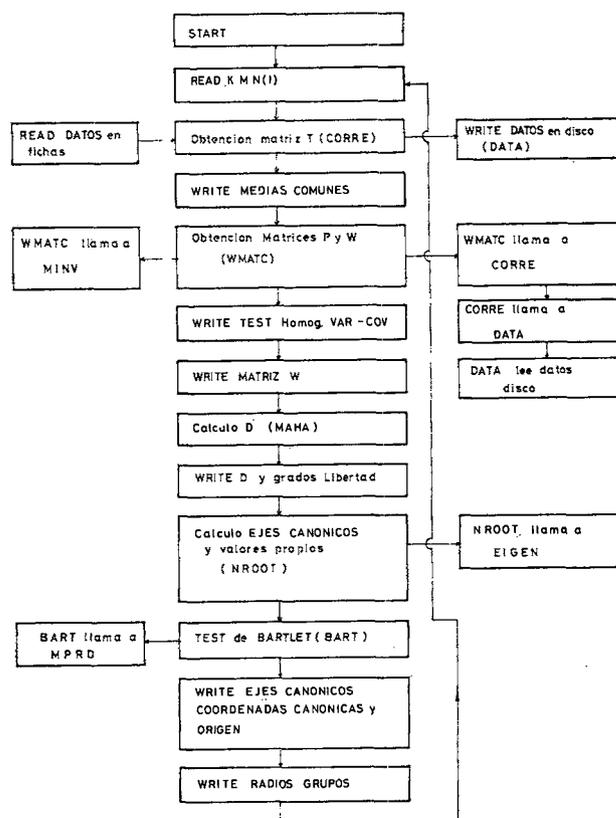
*** Departamento de Cristalografía y Mineralogía, Universidad de Barcelona. Sección de Mineralogía, Instituto "Jaime Almera", C. S. de I. C. Barcelona.

(1) IBM. Scientific Subroutine Package. 360A-CM-03X, versión III. Programmer's Manual H20-0205-3.

ENTRADA Y SALIDA DE DATOS

Como datos de entrada exige el número M de variables, número K de grupos y el número de observaciones de las M variables en cada uno de los grupos.

A continuación siguen las observaciones de las M variables de cada grupo, es decir: primer indivi-



duo, primer grupo, ídem, 2.º individuo, 1.º grupo,, ídem 1.º individuo, 2.º grupo, etc.

El programa obtiene las medias comunes, las medias de grupos, el test de homogeneidad de varianzas-

covarianzas, la pooled matriz de dispersión, la distancia D^2 de Mahalanobis, los ejes canónicos, las coordenadas canónicas de los grupos, el test de significación de los ejes canónicos y los radios de los grupos.

FUNCIONAMIENTO DEL PROGRAMA

El programa empieza leyendo los datos y quedándose en disco, al tiempo que calcula la matriz de varianza-covarianza total (subrutina CØRRE que llama a DATA)

$$T = (t_{ie}); t_{ie} = \sum_{h=1}^N (X_{ijh} - \bar{X}_i) (X_{ejh} - \bar{X}_e)$$

Seguidamente llama la subrutina WMATC, que vuelve a leer los datos, guardados en disco, calculando las matrices de varianza-covarianza de cada grupo y a partir de éstas la pooled dispersión matriz

$$W = (W_{ij}); W_{ij} = \sum_{j=1}^K \left(\sum_{h=1}^{n_j} (X_{ijh} - \bar{X}_{ij}) (X_{ejh} - \bar{X}_{ej}) \right)$$

al tiempo que efectúa el test de homogeneidad, calculando el valor ji-cuadrado

$$\chi^2 = -2 \left[-1 - \left\{ \sum_{i=1}^k \frac{1}{n_i - 1} - \frac{1}{N - k} \right\} \frac{2m^2 + 3m - 1}{6(m+1)(k-1)} \right]$$

$$L_n \left\{ \frac{\prod_{i=1}^k |S_i| \frac{n_i - 1}{2}}{|S| \frac{N - k}{2}} \right\}$$

A continuación llama a la subrutina MAHA y obtiene la distancia (al cuadrado) de Mahalanobis.

$$D^2 = \sum_{i=1}^m \sum_{e=1}^m d_{ij} \sum_{h=1}^n h_k (\bar{X}_{ih} - \bar{X}_i) (\bar{X}_{eh} - \bar{X}_e)$$

$$(d_{ij}) = S^{-1}$$

Calcula luego, la matriz de varianzas-covarianzas "entre grupos"

$$P = T - W$$

y llama a la subrutina NRØØT, para resolver la ecuación

$$|P - \phi W| = 0$$

y poder determinar los ejes canónicos. Una vez obtenidos éstos, se obtienen las coordenadas canónicas y el origen. Finalmente se llama a la subrutina BART para efectuar el test de significación de los ejes canónicos, calculándose los valores

$$\chi_i^2 = \left\{ (N - 1) - (m + i)/2 \right\}$$

$$L_n \left\{ \prod_{j=i+1}^m (1 + \phi_j) \right\}$$

y sus correspondientes grados de libertad. Antes de leer los datos del siguiente problema, el programa principal, calcula y escribe los radios $1.645 / \sqrt{n_j}$ de los grupos.

PROGRAMA Y SUBRUTINAS

El subprograma principal denominado CANØN así como las subrutinas DATA, WMATC, MAHA y BART, han sido escritas y programadas por el primero que suscribe, usándose además, las subrutinas del S.S. P. (1):

- CØRRE
- NRØØT
- EIGEN (valores y vectores propios de una matriz simétrica)
- MINV (inversión de matrices)
- LØC (localización de elementos en una matriz)
- MPRD (multiplica dos matrices)
- GMTRA (transposición de matrices)

(1) IBM. Scientific Subroutine Package, 360-A-CM-03X, versión III. Programmer's Manual H20-0205-3.

EJEMPLO

NÚMERO DE GRUPOS	4				
NÚMERO DE VARIABLES	6				
TAM. MUESTRAS					
GRUPO					
1	8				
2	7				
3	7				
4	8				
MEDIAS COMUNES					
7.66667	7.96667	7.33333	7.90000	19.39999	10.13333
TEST HOMOGENEIDAD					
V. JI C.	26.0541	G. LIB.	63		

GRUPO 1 MEDIAS	7.87500	7.50000	4.62500	7.25000	18.50000	8.87500
GRUPO 2 MEDIAS	7.14286	8.57143	9.57143	7.85714	20.14285	12.57143
GRUPO 3 MEDIAS	7.85714	7.85714	8.85714	9.28571	17.42856	10.14286
GRUPO 4 MEDIAS	7.75000	8.00000	6.75000	7.37500	21.37500	9.25000

POOLED DISPERSIÓN MATRIZ

FILA 1	19.61879	— 11.16207	— 5.21497	— 6.09890	— 22.2860	— 9.54052
FILA 2	— 11.16207	11.94504	5.61813	1.91758	22.60986	10.66757
FILA 3	— 5.21497	5.61813	39.45937	3.93681	16.23486	9.34546
FILA 4	— 6.09890	1.91758	3.93681	9.83310	4.62157	3.83791
FILA 5	— 22.74860	22.60986	16.23486	4.62157	62.78635	30.18266
FILA 6	— 9.54052	10.66757	9.34546	3.83791	30.18266	29.57483

D-CUADRADO MAHALANOBIS GENERALIZADÀ 12.78068 G. LIB 18

EJES CANÓNICOS

EJE 1	0.00291	— 0.22832	— 0.08114	— 0.07226	0.20904	— 0.16088
EJE 2	— 0.15136	— 0.16365	0.02925	— 0.30181	0.00347	0.14475
EJE 3	0.14647	— 0.03984	0.10092	0.10376	0.14238	— 0.08320

COORDENADAS CANÓNICAS

GRUPO 1	— 0.14927	— 3.12329	3.96923
GRUPO 2	— 1.09252	— 2.68561	4.30786
GRUPO 3	— 1.14932	— 3.48985	4.33269
GRUPO 4	0.09540	— 3.09751	4.53657

ORIGEN

— 0.53746 — 3.09982 4.28434

EJE V. JI. C. G.LIB.

1	4.1415	18
2	1.2670	10
3	0.4717	4

RADIOS GRUPOS

1	0.5816
2	0.6218
3	0.6218
4	0.5816