

8  
2ej.



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

ANALISIS ESTADISTICO DEL EFECTO DE  
ALGUNAS VARIABLES SOBRE PARAMETROS  
ELECTROENCEFALOGRAFICOS.

**T E S I S**

QUE PARA OBTENER EL TITULO DE:  
**M A T E M A T I C O**  
P R E S E N T A :

**THALIA FERNANDEZ HARMONY**

MEXICO, D. F.,

CIUDAD UNIVERSITARIA

1988



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# INDICE

|  |     |
|--|-----|
| Introducción .....   | 1   |
| Capítulo I: Análisis del problema biológico .....  | 7   |
| Capítulo II: Análisis de Correlación Canónica .....  | 27  |
| II.1.- Consideraciones previas .....   | 28  |
| II.2.- Obtención de las variables y las correlaciones canónicas (un punto de vista geométrico) ..... | 40  |
| II.3.- Obtención de las variables y las correlaciones canónicas (un punto de vista algebraico) ..... | 48  |
| II.3.1.- Obtención de la primera pareja de variables canónicas ...                                   | 48  |
| II.3.2.- Caracterización de los eigenvalores de B y sus eigenvectores .....                          | 57  |
| Capítulo III: Discusión de Resultados .....  | 64  |
| Apéndice .....   | 84  |
| Bibliografía .....   | 116 |

# INTRODUCCION

Existe una disciplina relativamente reciente de las ciencias biológicas llamada Neurometría. Esta tiene como objetivo la evaluación de la maduración del Sistema Nervioso y de su integridad anatómica y funcional a través del análisis cuantitativo de la actividad eléctrica cerebral. Una de las formas de registro de esta actividad eléctrica es el electroencefalograma (E.E.G.)

Es conocido que a un paciente se le indica como estudio un electrocardiograma cuando hay sospecha de una alteración cardíaca; así mismo, si se sospecha que un individuo tiene una alteración neurológica, uno de los estudios que se le indica es un E.E.G. Ambos estudios consisten en un conjunto de gráficas continuas del voltaje en función del tiempo y en tales casos, por regla general, el especialista hace un análisis cualitativo por inspección visual de las gráficas y da un diagnóstico basándose en su experiencia.

Sin despreciar en lo absoluto un análisis cualitativo del E.E.G., en este trabajo prestamos atención a un análisis cuantitativo del mismo, pues las características que resultan de interés son ciertas transformaciones de las variables que se obtienen del registro del E.E.G. conocidas como "potencias relativas". También debemos aclarar que los sujetos estudiados no son pacientes, sino un grupo de niños

provenientes de dos escuelas de la zona de Ixtacala y de una escuela de Toluca, seleccionados por tener un examen pediátrico, neurológico y psiquiátrico normal, así como un coeficiente de inteligencia normal.

Además de las variables electroencefalográficas 'potencia relativa' se obtuvieron otras variables (no electroencefalográficas) relativas a los antecedentes biológicos de riesgo, a las características socio-económicas y socioculturales, así como a las características pedagógicas de los individuos, por ejemplo: 'antecedentes del parto', 'estado de la vivienda' y 'evaluación de la lectoescritura'.

Uno de los objetivos de este trabajo es analizar, en caso de que existan, las relaciones lineales entre las variables electroencefalográficas y las variables no electroencefalográficas. Por ejemplo, sería deseable averiguar si existe alguna relación lineal entre las variables electroencefalográficas correspondientes al hemisferio cerebral izquierdo y la evaluación de la lectoescritura.

En este trabajo consideramos que los fenómenos que estudiamos se les puede asociar cierto grado de incertidumbre y además creemos que para un mejor análisis de la información debemos considerar que tienen carácter multivariado. Por estos motivos proponemos utilizar alguna técnica estadística de Análisis Multivariado.

Pero debido al poco conocimiento sobre el problema y a las múltiples especulaciones de los expertos, y tomando también en cuenta que las variables incluidas en el estudio no tienen ninguna distribución conocida y que no hay un esquema de muestreo que permita hacer inferencias, proponemos utilizar alguna técnica del Análisis Exploratorio de Datos.

Debemos hacer notar que todas las variables no electroencefalográficas son cualitativas - nominales u ordinales - y que sus categorías son excluyentes y exhaustivas; por ejemplo, la variable 'estado de la vivienda' se clasifica como 'bueno', 'regular' y 'malo', y la variable 'evaluación de la lectoescritura' se clasifica en 'buena', 'regular', 'mala' y 'pésima' de acuerdo con ciertos criterios que se explican en el capítulo I; un individuo cualquiera pertenece a una, y sólo a una, de las categorías de la variable 'estado de la vivienda' y a una, y sólo a una, de las categorías de la variable 'evaluación de la lectoescritura'.

Supongamos, por ejemplo, que queremos dar una caracterización pedagógica (de acuerdo a la evaluación de la lectoescritura) a partir de la caracterización electroencefalográfica del hemisferio izquierdo, en este caso podría pensarse en el uso del Análisis Discriminante (A.D.) como la técnica más ade-

cuada; pues cada individuo, caracterizado por ciertas variables electroencefalográficas, puede pertenecer a uno, y sólo a uno, de los "grupos" determinados por la variable 'evaluación de la lectoescritura'. Sin embargo, sería absurdo intentar explicar la variable 'estado de la vivienda' a partir de características electroencefalográficas, por eso en este caso no sirve el A.D. Tampoco podríamos, utilizando A.D., dar una caracterización electroencefalográfica a partir de la variable 'evaluación de la lectoescritura' ni de ninguna otra variable no electroencefalográfica.

Por lo tanto, para analizar las posibles relaciones lineales entre variables no electroencefalográficas y variables electroencefalográficas, proponemos el uso del Análisis de Correlación Canónica (A.C.C.), ya que puede verse como una generalización del A.D. con la ventaja de que es simétrico, es decir, no se tiene un conjunto de variables explicativas y otro de variables respuesta, sino que un conjunto de variables puede explicar al otro tanto como a la inversa.

Así, nuestro segundo objetivo es explicar en que consiste el A.C.C., descrito por Hotelling en 1936 por primera vez, dando una interpretación geométrica de los resultados algebraicos obtenidos.

En el Capítulo I de este trabajo, hacemos un



análisis del problema biológico, desde el registro del E.E.G. hasta la obtención de las variables electroencefalográficas que utilizamos, por un lado, y la obtención y definición de las variables no electroencefalográficas por otro.

En el Capítulo II, después de dar algunas consideraciones previas, describimos el A.C.C. y damos una interpretación geométrica del mismo. En el Capítulo III discutimos los resultados de aplicar el A.C.C. al problema biológico y al final del trabajo incluimos un Apéndice con las definiciones y todos los resultados que utilizamos en el Capítulo II.

# CAPITULO I

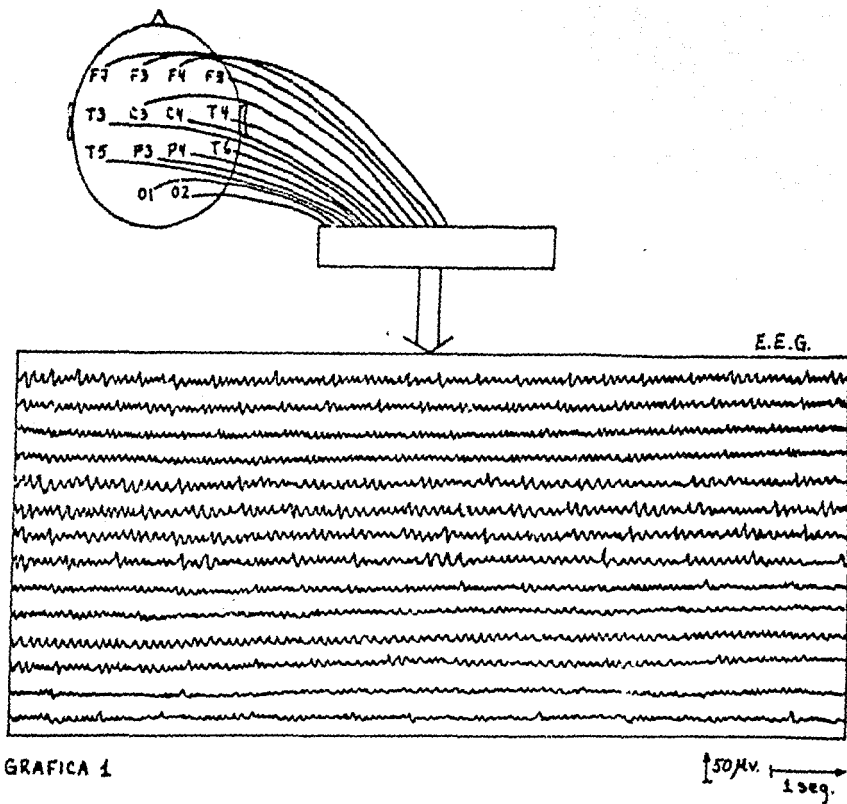
## Análisis del problema biológico

En este capítulo exponemos el problema biológico haciendo una breve explicación de la obtención de los datos electroencefalográficos para poder entender el origen de las variables que ellos generan, así como dando definiciones necesarias para la comprensión del problema. También comentaremos los grupos de variables no electroencefalográficas que constituimos y explicaremos lo que motivó a hacer este estudio.

Como ya dijimos en la introducción, sólo prestaremos atención a una de las dos formas de registro de la actividad eléctrica cerebral: el E.E.G., que es el registro de la actividad eléctrica cerebral espontánea. Este registro se puede hacer en tantas regiones de la cabeza como se quiera, pero en este caso sólo se hizo en las 14 zonas que mostramos en la gráfica 1, y que por convenio se denotan como aparecen representadas ahí<sup>(4)</sup>; las letras F, T, C, P y O significan 'frontal', 'temporal', 'central', 'parietal' y 'occipital' respectivamente. El registro se hace colocándole al individuo un electrodo en cada una de estas zonas, estos electrodos captan señales eléctricas y las conducen al electroencefalógrafo, que es el aparato diseñado para amplificar estas señales y dibujar una gráfica de cada señal; estas gráficas son funciones continuas de la amplitud (microvolts) con respecto al tiempo (milisegundos).

---

<sup>(4)</sup> Sistema 10-20 de Jasper (1959), citado por Harmony, T., 1984.



GRAFICA 1

Estas señales eléctricas existen desde que el individuo nace y dejan de emitirse cuando muere, por eso una pregunta que podríamos hacernos es ; qué intervalos de tiempo debemos escoger? Hay dos consideraciones que tenemos que tomar en cuenta para dar respuesta a esta pregunta:

1) Se necesita que todos los estudios se hagan en condiciones experimentales similares, pues es conocido

que el E.E.G. varía en el tiempo espontáneamente si ocurren cambios en el nivel de vigilia. No es posible descartar todos los estímulos que pueden provocar cambios en el E.E.G., pero si las condiciones experimentales son similares esperamos que produzcan cambios similares en todos los individuos. Además, a juicio del experto que realiza el estudio, todas las mediciones son tomadas sin artefactos, entendiendo por artefacto toda señal que no es de origen cerebral (movimientos musculares, oculares, etc.)

2) No importa en qué instante comiencen a tomarse los registros siempre y cuando se tomen varios intervalos no mayores de 12 segundos y que en total sumen más de 60 segundos.<sup>(2)</sup>

Tomando en cuenta estas consideraciones, los investigadores que registraron los datos de este trabajo tomaron 24 intervalos de 2.56 segundos cada uno, lo que resulta en un total de 61.44 segundos (2.56 no es un número arbitrario,  $256 = 2^8$  permite mayor eficiencia en los cálculos de la computadora).

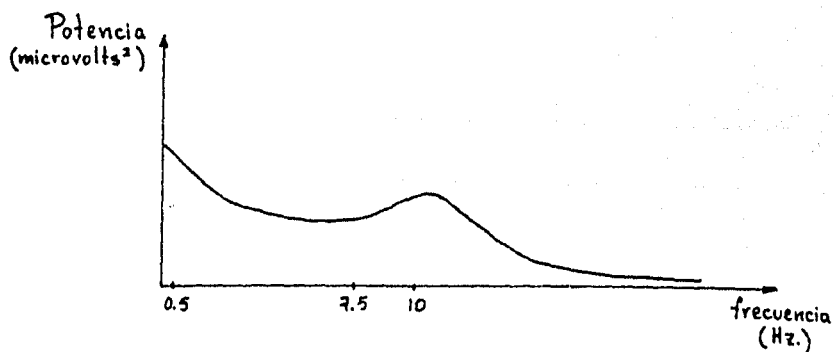
Los datos obtenidos corresponden al E.E.G, pero nosotros no trabajamos directamente con las variables electroencefalográficas registradas, sino con ciertas transformaciones de ellas que se conocen como "potencia relativa". A continuación explicamos cómo se ob-

---

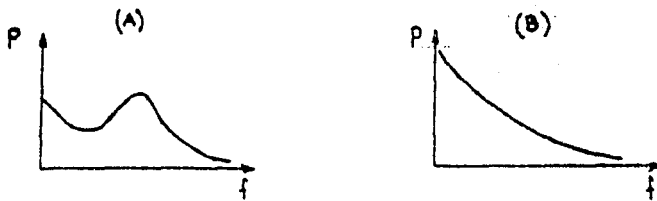
<sup>(2)</sup> El E.E.G. tomado bajo estas restricciones puede ser considerado como un proceso aleatorio estacionario. (Harmony, T, 1984)

tienen estas variables.

Una vez que se tienen las gráficas que constituyen el E.E.G., se consideran solamente los valores de amplitud cada 10 milisegundos y a éstos se les aplica una subrutina de computación<sup>(3)</sup> con lo cual se obtiene el "espectro de potencias", es decir, una gráfica de la potencia en función de la frecuencia como se muestra en la gráfica 2. El espectro de potencias informa sobre la distribución promedio de la potencia de una señal respecto a la frecuencia. Por ejemplo, durante el sueño el E.E.G. consiste fundamentalmente en ondas de gran potencia en bajas frecuencias (de 0.5 Hz a 7.5 Hz) pues la señal varía muy lentamente; mientras que en el caso de un individuo despierto en reposo con los ojos cerrados, el E.E.G. tiene ondas de gran potencia en frecuencias vecinas a 10 hertz (ver gráfica 3).



<sup>(3)</sup> Algoritmo Rápido de Fourier o Transformada Rápida de Fourier (F.F.T)  
(McGille y Avñon, 1981).



- (A) es el espectro de potencias correspondiente a una región posterior de un sujeto normal despierto en reposo con los ojos cerrados.
- (B) es el espectro de potencias en cualquier región durante una fase del sueño llamada sueño lento; también puede verse en individuos con lesión, pero sólo en algunas regiones.

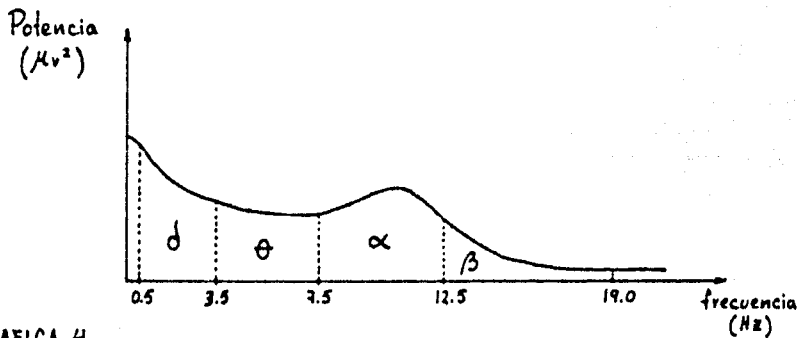
GRAFICA 3.

El análisis que se le hizo a nuestros datos se llama Análisis de Banda Ancha, porque en el espectro de potencias se determinan cuatro intervalos en el eje de las abscisas que definen cuatro áreas o bandas bajo la curva de la potencia. Las bandas se llaman  $\delta$ ,  $\theta$ ,  $\alpha$ ,  $\beta$  y comprenden el área bajo la curva de la potencia en los intervalos de frecuencia  $[0.5, 3.5)$ ,  $[3.5, 7.5)$ ,  $[7.5, 12.5)$  y  $[12.5, 19)$  respectivamente. Estos intervalos no son arbitrarios; originalmente se definieron de acuerdo con la relación que los especialistas observaron que tenían con aspectos funcionales (si el individuo estaba despierto o dormido, si tenía una lesión o no)<sup>(4)</sup> (ver gráfica 4).

Entonces tenemos un espectro de potencias para cada una de las 14 regiones de la cabeza, y en cada uno de ellos tenemos 4 bandas; por lo tan-

<sup>(4)</sup> Posteriormente se demostró estadísticamente que en relación con aspectos funcionales, las bandas obtenidas coinciden con las bandas clásicamente definidas. (Harmony, T., 1954)

to, podemos obtener 56 variables que llamamos "potencia absoluta" y definimos a continuación.



GRAFICA 4.

Definición: Por potencia absoluta de una banda  $\mu$  en una región  $C_i$  entendemos la integral de la potencia con respecto a la frecuencia en la región  $C_i$ , definida en el intervalo en que está definida la banda  $\mu$ .

Por ejemplo, supongamos que se quiere calcular la potencia absoluta de la banda  $\delta$  en la región F7. Llamemos  $P_{F7}(f)$  al espectro de potencias de F7 (lo denotamos así para resaltar que la potencia es función de la frecuencia) y recordemos que la banda  $\delta$  está definida en el intervalo  $[0.5, 3.5)$ , entonces

$$P.A_{\delta, F7} = \int_{0.5}^{3.5} P_{F7}(f) df$$

Como tenemos un amplificador de señales para cada región y ellos pueden establecer pequeñas dife-



rencias en las variables "potencia absoluta, nosotros bajaremos con las "potencias relativas" para evitar que los datos sean dependientes de las diferencias entre los amplificadores.

Definición: La potencia relativa de una banda  $\mu$  en una región  $C_i$  es la potencia absoluta de la banda  $\mu$  en la región  $C_i$  dividida por la suma de las potencias absolutas de todas las bandas en la región  $C_i$ .

Por ejemplo, la potencia relativa de la banda  $\delta$  en la región  $F7$  podemos expresarla como:

$$P.R_{\delta, F7} = \frac{P.A_{\delta, F7}}{P.A_{\delta, F7} + P.A_{\theta, F7} + P.A_{\alpha, F7} + P.A_{\beta, F7}}$$

En total vamos a tener 56 variables de potencia relativa, 14 de cada banda, sin embargo sólo trabajaremos con 42 de ellas: las correspondientes a las bandas  $\delta$ ,  $\theta$  y  $\alpha$ . Descartamos las variables de potencia relativa correspondientes a una de las bandas porque la potencia relativa de cualquier banda en una región  $C_i$  puede obtenerse a partir de la potencia relativa de las bandas restantes en la misma región, ya que  $PR_{\delta, C_i} + PR_{\theta, C_i} + PR_{\alpha, C_i} + PR_{\beta, C_i} = 1$  para toda región  $C_i$ ; y en particular decidimos descartar las de la banda  $\beta$  porque otros estudios han de-

mostrado que son estas las que sufren menos cambios cuando hay alteraciones de las funciones cerebrales (estudios de Matousek y Petersén ; Harmony, 1984)

Es conocido que para el óptimo desarrollo del cerebro y de las funciones mentales es imprescindible una adecuada estimulación sensorial y una adecuada estimulación de la actividad del lenguaje en el momento oportuno ; hay estudios que demuestran que si en determinadas etapas de la vida del niño no se le da la estimulación adecuada , aumenta el riesgo de que sea un individuo con daño cerebral<sup>(5)</sup>. Suponemos que tal estimulación está relacionada con factores sociales y culturales , de ahí nuestro interés en conocer si existe algún tipo de relación entre las variables electroencefalográficas (potencias relativas) y las variables que se refieren a antecedentes socioculturales que describimos a continuación desglosadas en categorías :

ESCOLARIDAD DE LA MADRE , cuyas categorías son:

- analfabeta
- estudió hasta 3° de primaria
- estudió hasta 6° de primaria
- educación media (secundaria o técnica)
- educación superior

ESCOLARIDAD DEL PADRE (iguales categorías que escolaridad de la madre).

---

(5) Cravioto, J. y Arrieta, R., 1982.

#### OCUPACION DE LA MADRE

- ama de casa
- trabajadora eventual
- trabajadora fija
- trabajadora calificada
- profesionista

#### OCUPACION DEL PADRE

- desempleado
- trabajador eventual
- trabajador fijo
- trabajador calificado
- profesionista

Estas cuatro variables fueron consideradas ordinales por los especialistas. Nótese que en las variables de escolaridad, las categorías son excluyentes y exhaustivas, sin embargo en las variables de ocupación no son excluyentes, pues por ejemplo la madre puede ser ama de casa y trabajadora eventual, pero en este caso fue clasificada como trabajadora eventual, pues siempre que un individuo podía ser clasificado en más de una categoría en las variables de ocupación, se le asignó la categoría "más alta" suponiendo que el orden establecido es ascendente, y así hacemos a estas categorías excluyentes.

Las siguientes variables corresponden a los an-

## Antecedentes socioeconómicos:

### INGRESO PERCAPITA (en % de salario mínimo)

- [0% , 25 %)
- [25 % , 50 %)
- [50 % , 100 %)
- 100% ó más de 100%

### ESTADO DE LA VIVIENDA

- bueno
- regular
- malo

Las categorías de esta variable fueron definidas de acuerdo a:

- existencia de los servicios de agua, luz y drenaje.
- existencia de refrigerador, estufa de gas y radio o televisión.
- si la casa está cimentada o no.
- si el niño comparte o no su cama.
- si las habitaciones utilizadas como dormitorio, tienen uso exclusivo como dormitorio o no.
- ventilación determinada por el número de ventanas en relación al número de habitaciones.

Hay múltiples trabajos (Velasco, 1973) que indican que los antecedentes biológicos que mencionamos a continuación, aumentan el riesgo de que el individuo tenga o bien daño cerebral con alteraciones clínicas neu-

rológicas (en nuestra muestra no hay niños con estas características) o bien trastornos en el funcionamiento cerebral como los que se pueden observar en individuos con trastorno en el aprendizaje; por lo tanto es de mucho interés conocer si la presencia de antecedentes biológicos de riesgo, por ella misma, tiene efectos en las variables electroencefalográficas. Los antecedentes biológicos los resumimos en 4 variables, cada una de las cuales tiene varias categorías.

**ANTECEDENTES DE PREPARTO:** Son los problemas que pudo haber tenido la madre durante el embarazo y se supone que pueden haber afectado al hijo. Los clasificamos de acuerdo a las siguientes categorías:

- sin antecedentes
- desnutrición de la madre
- deshidratación de la madre
- amenaza de aborto
- traumatismos de la madre
- uso de fármacos
- madre añosa

**ANTECEDENTES PERINATALES:** Estos antecedentes son los que pudo haber sufrido directamente el niño en el momento del parto e inmediatamente después. Las categorías son las siguientes:

- sin antecedentes perinatales
- anoxia

- convulsiones
- ictericia
- necesitó incubadora
- bajo peso al nacer
- prematuro
- hipermaduro

ANTECEDENTES DEL PARTO : Se refieren a las condiciones en que ocurrió el parto. Las categorías son:

- normal
- no hospitalario
- pélvico
- por cesárea
- aplicación de fórceps
- parto con anestesia general.
- parto prolongado

ANTECEDENTES DE POSTPARTO : Se refiere a los problemas que pudo haber tenido el niño durante su vida, exceptuando la etapa perinatal, hasta el momento en que se le hizo el estudio. Los clasificamos en las siguientes categorías:

- sin antecedentes
- desnutrición
- deshidratación severa
- trauma craneoencefálico con pérdida de conocimiento
- convulsiones
- fiebres altas sostenidas

- enfermedades parasitarias, virales o respiratorias que necesitaron hospitalización.

Las variables de antecedentes biológicos son exhaustivas, pues en caso de no presentar ningún antecedente de riesgo, el individuo se clasifica en la categoría 'sin antecedentes' en cada caso. Sin embargo, así descritas, sus categorías no son excluyentes; para tener un conjunto de categorías excluyentes en cada variable, construimos nuevas categorías en las que se consideraba la presencia de dos o más riesgos, pero con esta nueva clasificación aumentó notablemente el número de categorías. Por ejemplo, si había al menos un niño que tuviera como antecedentes perinatales 'prematureo' y 'bajo peso al nacer', construimos una nueva categoría 'prematureo y bajo peso al nacer' y consideramos que el niño pertenecía a esta última y no pertenecía a ninguna otra, incluidas las dos primeras.

A los neurofisiólogos les interesa conocer las relaciones entre las variables electroencefalográficas y los antecedentes biológicos en general, pues consideran que la información sobre antecedentes biológicos es muy pobre cuando se considera un solo tipo de antecedentes. Debido a que crear nuevas variables para resumir toda la información sobre los antecedentes biológicos resulta en más variables que número de individuos y debido también a limitaciones de los paquetes estadísticos (no encontramos un paquete que tuviera implementado A.C.C. generalizado),

decidimos consultar a dos especialistas para que, analizando las 4 variables de antecedentes, dieran un diagnóstico global para cada individuo; a esta variable le llamamos:

ANTECEDENTES SEGUN EXPERTOS y la clasificamos en:

- normal (sin riesgo)
- riesgos menos graves
- riesgos graves
- riesgos muy graves

Estas variables que contemplan los antecedentes biológicos fueron obtenidas por medio de encuestas a los padres y estos contestaban de acuerdo con lo que recordaban; además en ocasiones las madres sabían, por ejemplo, que el niño al nacer fue puesto en incubadora, pero no conocían el motivo. Por estas razones creemos que estas variables, medidas de este modo, no son muy confiables.

Existen también antecedentes en la literatura acerca de las diferencias electroencefalográficas entre los niños con trastorno de aprendizaje y los niños sin trastorno de aprendizaje (John y cols., 1983). Tomando esto en cuenta, suponemos que la evaluación de la lectoescritura obtenida mediante un conjunto de pruebas pedagógicas (Hinojosa y Rocha, 1985) que se les aplicaron a los niños, puede estar relacionada con las variables electroencefalográficas. La variable



EVALUACION DE LA LECTOESCRITURA , se clasificó en:

- buena : conocimientos de acuerdo a su edad y grado escolar.
- regular: ligeras deficiencias en la lectoes. critura que ameritan ayuda de los padres.
- mala : deficiencias que hacen necesaria una atención especial del maestro y los padres.
- pésima: deficiencias severas que hacen necesaria educación especializada.

Tomando en cuenta que en otros trabajos se ha encontrado que con registros bipolares del E.E.G. (no es nuestro caso) la potencia relativa tiene cambios significativos de acuerdo a la edad (Harmony, 1984), lo primero que comprobamos es que en nuestro estudio las variables electroencefalográficas no están relacionadas linealmente con la edad, por lo cual no las consideramos como funciones de la edad en este trabajo.

Hasta aquí contamos con un grupo de 42 variables electroencefalográficas "potencia relativa" obtenidas a partir del E.E.G. Nos interesa analizar las posibles relaciones lineales entre estas variables y cada una de las siguientes variables:

- escolaridad de la madre
- escolaridad del padre
- ocupación de la madre
- ocupación del padre
- salario per cápita
- estado de la vivienda
- antecedentes según expertos
- evaluación de la lectoescritura.

Para ello utilizamos el A.C.C., técnica del Análisis Exploratorio de Datos. El A.C.C nos permite estudiar las relaciones lineales entre dos "grupos" de variables, que se diferencian por la naturaleza de sus variables. Nótese que las categorías de cada una de las variables no electroencefalográficas se diferencian por su naturaleza de las categorías de las demás variables no electroencefalográficas; por este motivo a cada una de estas variables la consideramos como un "grupo" de variables cuyos elementos son precisamente las categorías de la variable original; estos grupos de variables no electroencefalográficas están descritos en la tabla 1. Las nuevas variables que integran los grupos son variables binarias tales que: si 'y' es una de estas variables en el grupo Y, entonces

$$y = \begin{cases} 1 & \text{si el individuo pertenece a la categoría } y \\ 0 & \text{si no} \end{cases}$$

Por otro lado, debido a que el número de variables electroencefalográficas es muy grande en relación con el número de individuos (161), no podemos considerar en un análisis a todas estas variables simultáneamente, por ello, los especialistas definieron 4 subconjuntos o "grupos" de variables electroencefalográficas que a su modo de ver tienen sentido a la hora de interpretar los resultados (tabla 2)

| GRUPOS DE VARIABLES NO ELECTROENCEFALOGRAFICAS |  |
|--|--|
| NOMBRE DEL GRUPO                               | VARIABLES QUE INTEGRAN AL GRUPO                  |
| 'escolaridad de la madre'                      | analfabeta, . . . , educación superior.          |
| 'escolaridad del padre'                        | analfabeta, . . . , educación superior.          |
| 'ocupación de la madre'                        | ama de casa, t. eventual, . . . , profesionista. |
| 'ocupación del padre'                          | desempleado, t. eventual, . . . , profesionista. |
| 'salario per cápita'                           | [0%, 25%), [25%, 50%), [50%, 100%), 100% ó más   |
| 'estado de la vivienda'                        | bueno, regular, malo.                            |
| 'antecedentes según expertos'                  | normal, riesgos menos graves, graves, muy graves |
| 'evaluación de la lectoescritura'              | bueno, regular, mala, pésima.                    |

TABLA 1

| GRUPOS DE VARIABLES ELECTROENCEFALOGRAFICAS |   |
|---|---|
| NOMBRE DEL GRUPO                            | VARIABLES QUE INTEGRAN AL GRUPO   |
| HEMISFERIO IZQUIERDO                        | RDCI, RDFI, RDFIL, RDTI, RDPI, RDOI, RDTIP,<br>RTCI, RTFI, RTFIL, RTTI, RTPI, RTOI, RTTIP,<br>RACI, RAFI, RAFIL, RATI, RAPI, RAOI, RATIP.                   |
| HEMISFERIO DERECHO                          | RDCD, RDFD, RDFDL, RDTD, RDPD, RDOD, RDTDP,<br>RTCD, RTFD, RTFDL, RTTD, RTPD, RTOD, RTTDP,<br>RACD, RAFD, RAFDL, RATD, RAPD, RAOD, RATDP.                   |
| REGION ANTERIOR                             | RDCI, RDCD, RDFI, RDFD, RDFIL, RDFDL, RDTI, RDTD,<br>RTCI, RTCD, RTFI, RTFD, RTFIL, RTFDL, RTTI, RTTD,<br>RACI, RACD, RAFI, RAFD, RAFIL, RAFDL, RATI, RATD. |
| REGION POSTERIOR                            | RDPI, RPPD, RDOI, RDOD, RDTIP, RDTDP,<br>RTPI, RTPD, RTOI, RTOD, RTTIP, RTTDP,<br>RAPI, RAPD, RAOI, RAOD, RATIP, RATDP.                                     |

TABLA 2 . Esta nueva nomenclatura de las variables se aclara a continuación

RDCI = P.R<sub>δ,C3</sub>

RDCD = P.R<sub>δ,C4</sub>

RDFI = P.R<sub>δ,F3</sub>

RDFD = P.R<sub>δ,F4</sub>

RDFIL = P.R<sub>δ,F7</sub>

RDFDL = P.R<sub>δ,F8</sub>

RDTI = P.R<sub>δ,T3</sub>

RDTD = P.R<sub>δ,T4</sub>

RDPI = P.R<sub>δ,P3</sub>

RDPD = P.R<sub>δ,P4</sub>

RDOI = P.R<sub>δ,O1</sub>

RDOO = P.R<sub>δ,O2</sub>

RDTIP = P.R<sub>δ,T5</sub>

RDTDP = P.R<sub>δ,T6</sub>

Esta nueva nomenclatura se hizo necesaria para denotar a las variables electroencefalográficas en la computadora. La primera letra 'R' es de 'potencia relativa', la segunda letra indica la banda ('D' por  $\delta$ , 'T' por  $\theta$  o 'A' por  $\alpha$ ), la cuarta letra indica si la región es izquierda ('I') ó derecha ('D'); la otra u otras letras indican la región, sin especificar si es izquierda o derecha. Entonces RACI = P.R <sub>$\alpha$ ,C3</sub>, y así con las otras.

El programa que utilizamos para hacer el análisis es CANCORR del SPSS.

Nótese que en cada uno de los grupos de variables no electroencefalográficas, una variable cualquiera queda explicada por las demás variables del mismo grupo, ya que para cada individuo la suma de las observaciones en un grupo es 1. Entonces podemos descartar en cada grupo a una variable. Consideremos por ejemplo, la variable 'evaluación de la lectoescritura' y que en ella descartamos a la categoría 'mala', entonces tenemos un

grupo 'evaluación de la lectoescritura' que incluye las variables 'buena', 'regular', 'mala' y 'pésima'; si en este grupo asignamos a un individuo ceros en las 3 variables, queda claro que el individuo está clasificado en la categoría 'mala'.

En cada una de las "corridas" del programa CANCORR del SPSS consideramos un grupo de la tabla 1 (descartando una variable) y un grupo de la tabla 2. Debe notarse que las variables de un grupo son categorías de una variable nominal y que las del otro grupo son variables continuas; este problema aparece resuelto en la tesis de Saporta, G. (1975)

## CAPITULO II

### Análisis de Correlación Canónica

## II.1.-CONSIDERACIONES PREVIAS

Supongamos que en un conjunto de  $N$  individuos observamos  $n$  características o variables<sup>(6)</sup>, esta información podemos resumirla en una matriz rectangular de orden  $N \times n$ , en donde  $N > n$ , como la siguiente:

$$W = \begin{bmatrix} \omega_1^1 & \omega_1^2 & \dots & \omega_1^n \\ \omega_2^1 & \omega_2^2 & \dots & \omega_2^n \\ \vdots & \vdots & \dots & \vdots \\ \omega_N^1 & \omega_N^2 & \dots & \omega_N^n \end{bmatrix} = [\omega_i^j]$$

donde  $\omega_i^j$  representa la observación correspondiente a la  $j$ -ésima característica en el  $i$ -ésimo individuo. Esta matriz recibe el nombre de 'matriz de datos individuos  $\times$  variables'.

El objetivo del A.C.C es encontrar las mejores relaciones lineales (en el sentido de la correlación) entre dos conjuntos de variables de naturaleza diferente observadas sobre un mismo conjun-

---

<sup>(6)</sup> Clásicamente el A.C.C. supone que estas variables son continuas; sin embargo, si tenemos un grupo de variables continuas y otro grupo de variables binarias puede demostrarse desde un punto de vista más general (Análisis Canónico Funcional, Lancaster, H.O., 1958) que el A.C.C. es el óptimo en el sentido de que maximiza la correlación (Saporta, G., 1975).

to de  $N$  individuos; por lo que, si nos interesa encontrar qué relaciones lineales existen entre las variables electroencefalográficas del hemisferio cerebral izquierdo y las variables de evaluación de la lectoescritura, por ejemplo, podemos tomar a  $\omega$  tal que contenga la información relativa a tales variables.

Así, a cada individuo lo podemos representar como un punto en  $\mathbb{R}^n$  y a cada característica la podemos representar como un punto en  $\mathbb{R}^N$ , obteniendo una "nube de individuos" en  $\mathbb{R}^n$  y una "nube de características" en  $\mathbb{R}^N$ . Entonces, los vectores

$$\underline{\omega}_i = (\omega_i^1, \dots, \omega_i^n) \quad i=1, \dots, N.$$

$$\text{y } \underline{\omega}^j = (\omega_1^j, \dots, \omega_N^j) \quad j=1, \dots, n$$

son vectores en la nube de individuos y en la nube de características respectivamente.

A continuación damos algunas definiciones necesarias para el desarrollo de nuestro trabajo.



Definición 1: La media de la  $j$ -ésima variable la denotamos por

$$\bar{w}^j = \frac{1}{N} \sum_{i=1}^N w_i^j \quad ; \quad j=1, \dots, n.$$

Así, el vector de los valores medios queda expresado como:

$$\bar{w} = (\bar{w}^1, \bar{w}^2, \dots, \bar{w}^n)$$

Definición 2: Denotamos a la covarianza entre la  $j$ -ésima y la  $k$ -ésima variable por:

$$S_{jk} = \frac{1}{N} \sum_{i=1}^N (w_i^j - \bar{w}^j)(w_i^k - \bar{w}^k)$$

y a la varianza de la  $j$ -ésima variable por:

$$S_j^2 = \frac{1}{N} \sum_{i=1}^N (w_i^j - \bar{w}^j)^2$$

Debe notarse que en estas definiciones hemos considerado que todos los individuos tienen la misma importancia, y por ello le asignamos a cada individuo el mismo "peso"  $p_i$  ( $p_i \geq 0$ ,  $\sum_{i=1}^N p_i = 1$ ); en otras palabras  $p_i = \frac{1}{N} \quad \forall i=1, \dots, N$ .

Definición 3: La matriz de varianzas y covarianzas es una matriz cuadrada ( $n \times n$ ) de la forma:

$$R = \begin{bmatrix} S_1^2 & S_{12} & \cdots & S_{1n} \\ S_{12} & S_2^2 & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1n} & S_{2n} & \cdots & S_n^2 \end{bmatrix}$$

Nótese que  $R$  es una matriz simétrica.

Sin pérdida de generalidad, supondremos que las variables están centradas; es decir, que están representadas en un sistema de coordenadas cuyo centro coincide con el vector de valores medios (o centro de gravedad)  $\bar{w} = (\bar{w}^1, \dots, \bar{w}^n)$ . En otras palabras, si  $Z$  es la matriz de datos centrada correspondiente a la matriz de datos  $W$ , la suma de los elementos de cada columna de  $Z$  es igual a cero. El elemento  $z_i^j$  de  $Z$  representa la observación de la  $j$ -ésima variable centrada en el  $i$ -ésimo individuo, es decir,  $z_i^j = (w_i^j - \bar{w}^j)$ . Entonces las definiciones anteriores quedan expresadas como:

$$\bar{z}^j = 0 \quad \forall j = 1, 2, \dots, n$$

$$S_{jk} = \frac{1}{N} \sum_{i=1}^N z_i^j z_i^k \quad ; j, k = 1, 2, \dots, n$$

$$S_j^2 = \frac{1}{N} \sum_{i=1}^N (z_i^j)^2 \quad ; j = 1, 2, \dots, n$$

Supongamos que el conjunto de las  $n$  variables lo podemos expresar como la unión de dos conjuntos de variables diferenciados por la naturaleza de sus elementos. Sean las primeras  $p$  variables las correspondientes a uno de estos conjuntos y las  $q$  variables restantes las correspondientes al otro conjunto ( $n = p + q$ ). En otras palabras, supongamos que se tienen dos "grupos" o submatrices con  $p$  y  $q$  variables respectivamente, observadas sobre un mismo conjunto de  $N$  individuos. En nuestro ejemplo, las  $p$  variables de un grupo pueden corresponder a las variables electroencefalográficas del hemisferio izquierdo y las  $q$  variables del otro grupo, a las variables de evaluación de la lectoescritura, ambas centradas.

A estas matrices que obtenemos al particionar la matriz  $Z$ , las denotamos por  $X (N \times p)$  y  $Y (N \times q)$

$$Z = \left[ \begin{array}{c|c} X & Y \end{array} \right]$$

Esto es equivalente a particionar a cada vector  $\underline{z}_i = (z_i^1, \dots, z_i^n)$  en dos subvectores  $\underline{x}_i$  de las  $p$  variables de un grupo y  $\underline{y}_i$  de las  $q$  variables del otro grupo, es decir  $\underline{z}_i$  puede escribirse como

$$\underline{z}_i = (\underline{x}_i, \underline{y}_i) \quad \text{para toda } i = 1, \dots, N$$

en donde

$$\begin{aligned} \underline{x}_i &= (x_i^1, \dots, x_i^p) = (z_i^1, \dots, z_i^p) \\ \underline{y}_i &= (y_i^1, \dots, y_i^q) = (z_i^{p+1}, \dots, z_i^n) \end{aligned}$$

Entonces:

$$R = \begin{array}{c|ccc|ccc} S_1^2 & S_{12} & \cdots & S_{1p} & S_{1,p+1} & \cdots & S_{1n} \\ S_{12} & S_2^2 & \cdots & S_{2p} & S_{2,p+1} & \cdots & S_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ S_{1p} & S_{2p} & \cdots & S_p^2 & S_{p,p+1} & \cdots & S_{pn} \\ \hline S_{1,p+1} & S_{2,p+1} & \cdots & S_{p,p+1} & S_{p+1}^2 & \cdots & S_{p+1,n} \\ S_{1,p+2} & S_{2,p+2} & \cdots & S_{p,p+2} & S_{p+1,p+2} & \cdots & S_{p+2,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ S_{1n} & S_{2n} & & S_{pn} & S_{p+1,n} & & S_n^2 \end{array}$$

podemos escribirla como:

$$R = \begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array}$$

en donde la matriz  $R_{11}$  de orden  $(p \times p)$  es la matriz de covarianzas de  $X$ , la matriz  $R_{22}$  de orden  $(q \times q)$  es la matriz de covarianzas de  $Y$  y la matriz  $R_{12} = R_{21}'$  de orden  $(p \times q)$  es la matriz de covarianzas entre  $X$  y  $Y$ .

Por otro lado tenemos que la matriz R podemos expresarla como:

$$R = \frac{1}{N} Z' Z$$

$$R = \frac{1}{N} \begin{array}{c|c} \begin{array}{c} \sum_{\lambda=1}^N (z_{\lambda}^1)^2 \\ \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^2 \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^p \\ \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^{p+1} \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^n \end{array} & \begin{array}{c} \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^2 \dots \sum_{\lambda=1}^N z_{\lambda}^1 z_{\lambda}^p \\ \sum_{\lambda=1}^N (z_{\lambda}^2)^2 \dots \sum_{\lambda=1}^N z_{\lambda}^2 z_{\lambda}^p \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^2 z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^2 z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^2 z_{\lambda}^n \end{array} \\ \hline \begin{array}{c} \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N (z_{\lambda}^p)^2 \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \end{array} & \begin{array}{c} \sum_{\lambda=1}^N z_{\lambda}^{p+1} z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N (z_{\lambda}^{p+1})^2 \\ \sum_{\lambda=1}^N z_{\lambda}^{p+1} z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^{p+1} z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^{p+1} z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N (z_{\lambda}^{p+1})^2 \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^{p+1} z_{\lambda}^n \end{array} \\ \hline \begin{array}{c} \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \sum_{\lambda=1}^N z_{\lambda}^n z_{\lambda}^n \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \sum_{\lambda=1}^N z_{\lambda}^n z_{\lambda}^n \dots \sum_{\lambda=1}^N (z_{\lambda}^n)^2 \end{array} & \begin{array}{c} \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \\ \vdots \\ \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^{p+1} \dots \sum_{\lambda=1}^N z_{\lambda}^p z_{\lambda}^n \end{array} \end{array}$$

$$R = \frac{1}{N} \begin{array}{c|c} X'X & X'Y \\ \hline Y'X & Y'Y \end{array}$$

Por lo tanto:

$$R_{11} = \frac{1}{N} X'X$$

$$R_{22} = \frac{1}{N} Y'Y \dots \dots \dots (1)$$

$$R_{12} = \frac{1}{N} X'Y$$

Una forma de caracterizar a los individuos con respecto a la naturaleza de las variables de  $X$  es mediante una combinación lineal de los vectores columna de la matriz  $X$ ; asimismo, una combinación lineal de los vectores columna de la matriz  $Y$  es una forma de caracterizar a los individuos con respecto a la naturaleza de las variables del grupo  $Y$ .

En nuestro ejemplo, una caracterización electroencefalográfica del hemisferio izquierdo para todos los individuos podemos darla a través de:

$$\underline{u} = (u_{(1)}, u_{(2)}, \dots, u_{(N)})' = X\underline{a} \quad ; \underline{a} = (a_1, \dots, a_p)'$$

en donde la  $i$ -ésima componente del vector  $\underline{u}$ :

$$u_{(i)} = a_1 x_{i1} + \dots + a_p x_{ip} = \underline{x}_i \underline{a}$$

es la caracterización electroencefalográfica del hemisferio izquierdo para el  $i$ -ésimo individuo con respecto al vector de coeficientes  $\underline{a}$

Análogamente, una caracterización de todos los individuos con respecto a las variables de  $Y$ , que en nuestro ejemplo serían las variables del grupo 'evaluación de la lectoescritura', podemos darla a través de

$$\underline{y} = Y\underline{b} \quad ; \underline{b} = (b_1, \dots, b_q)'$$

Debe notarse que de estas combinaciones lineales existen una infinidad, es decir, existe un número

infinito de vectores  $\underline{a}$  y  $\underline{b}$  que nos permiten caracterizar, por medio de una combinación lineal, a los individuos con respecto a la naturaleza de las variables de los grupos  $X$  y  $Y$  respectivamente.

A cada una de las variables de  $Z$  la podemos considerar como un vector en  $\mathbb{R}^N$ , entonces cualesquiera dos combinaciones lineales de ellas  $\underline{u}$  y  $\underline{v}$  también son vectores en  $\mathbb{R}^N$ ; además  $\underline{u}$  y  $\underline{v}$  son centradas ya que las variables de  $Z$  son centradas. Demostraremos este resultado sólo para  $\underline{u}$ , pues la demostración para  $\underline{v}$  es análoga.

Demostración:

Demostrar que  $\underline{u}$  es centrada, es equivalente a demostrar que su media es cero.

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N u_i &= \frac{1}{N} \left[ \sum_{i=1}^N a_1 x_i^1 + \dots + \sum_{i=1}^N a_p x_i^p \right] \\ &= a_1 \left( \frac{1}{N} \sum_{i=1}^N x_i^1 \right) + \dots + a_p \left( \frac{1}{N} \sum_{i=1}^N x_i^p \right) \\ &= a_1 \bar{x}^1 + \dots + a_p \bar{x}^p \\ &= a_1 \bar{z}^1 + \dots + a_p \bar{z}^p \\ &= 0 \text{ pues } \bar{z}^j = 0 \quad \forall j=1, \dots, n \end{aligned}$$

L.Q.Q.D.

Sean  $\Psi_x \subset \mathbb{R}^N$  y  $\Psi_y \subset \mathbb{R}^N$  los subespacios generados por las variables (o vectores columna) de  $X$  y de  $Y$  respectivamente. Entonces, cualquier combinación lineal de las variables de  $X$  pertenece al subespacio  $\Psi_x$  y cualquier combinación lineal de las variables de  $Y$  pertenece al subespacio  $\Psi_y$ .

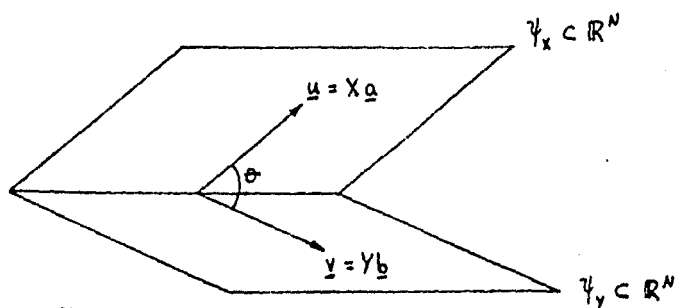
$$\Psi_x = \{ \underline{u} \in \mathbb{R}^N \rightarrow \underline{u} = X\underline{a}, \underline{a} \in \mathbb{R}^p \}$$

$$\Psi_y = \{ \underline{v} \in \mathbb{R}^N \rightarrow \underline{v} = Y\underline{b}, \underline{b} \in \mathbb{R}^q \}$$

Sean  $\underline{u}$  y  $\underline{v}$  dos combinaciones lineales tales que  $\underline{u} \in \Psi_x$  y  $\underline{v} \in \Psi_y$ . Llamemos  $\theta$  al ángulo formado entre  $\underline{u}$  y  $\underline{v}$ .

Como  $\Psi_x$  y  $\Psi_y$  son subespacios, ambos contienen al vector  $\underline{0}$ , esto es,  $\underline{0} \in \Psi_x \cap \Psi_y$ . Las combinaciones lineales que denotamos por  $\underline{u}_0 = X\underline{0}_p$  y  $\underline{v}_0 = Y\underline{0}_q$ , por lo tanto pertenecen a  $\Psi_x \cap \Psi_y$  y se conocen como la solución trivial del A.C.C.; para ellas se satisface  $\theta = 0$

Consideremos entonces  $\underline{u}, \underline{v} \neq \underline{0}$



GRAFICA 5



Nos interesa analizar qué tan "parecidos" son estos dos subespacios, en el sentido de encontrar parejas de vectores  $\underline{u}$  y  $\underline{v}$  que minimicen el ángulo  $\theta$ ; esto es equivalente a encontrar parejas de vectores  $\underline{u}$  y  $\underline{v}$  que maximicen la correlación, pues, como demostramos a continuación

$$r(\underline{u}, \underline{v}) = \cos \theta$$

donde  $r(\underline{u}, \underline{v})$  es la correlación entre  $\underline{u}$  y  $\underline{v}$ .

Demostración

$$r(\underline{u}, \underline{v}) = \frac{\text{cov}(\underline{u}, \underline{v})}{[\text{var}(\underline{u}) \text{var}(\underline{v})]^{1/2}}$$

$$\text{pero } \text{var}(\underline{u}) = \frac{1}{N} \sum_{i=1}^N u_{(i)}^2 = \frac{1}{N} \underline{u}' \underline{u}$$

análogamente

$$\text{var}(\underline{v}) = \frac{1}{N} \underline{v}' \underline{v}$$

$$\text{y } \text{cov}(\underline{u}, \underline{v}) = \frac{1}{N} \sum_{i=1}^N u_{(i)} v_{(i)} = \frac{1}{N} \underline{u}' \underline{v}$$

Por lo tanto

$$r = \frac{\frac{1}{N} \underline{u}' \underline{v}}{[\frac{1}{N} \underline{u}' \underline{u} \frac{1}{N} \underline{v}' \underline{v}]^{1/2}} = \frac{\underline{u}' \underline{v}}{[\underline{u}' \underline{u} \underline{v}' \underline{v}]^{1/2}} = \frac{\langle \underline{u}, \underline{v} \rangle}{[\|\underline{u}\|^2 \|\underline{v}\|^2]^{1/2}} = \cos \theta$$

(2)

L.Q.Q.D.

A estas parejas de vectores  $\underline{u}$  y  $\underline{v}$  se les conoce como variables canónicas.

En el capítulo I decidimos descartar a las variables electroencefalográficas de potencia relativa correspondientes a la banda  $\beta$  y en cada grupo de variables no electroencefalográficas (tabla 1) descartamos una variable, debido a que en ambos casos, las variables descartadas pueden obtenerse como funciones de las restantes variables de su grupo. Nótese que con ésto, las columnas de cada grupo son conjuntos de vectores linealmente independientes, por lo tanto cualesquiera matrices de datos  $X$  y  $Y$  son de rango completo. (\*)

En este trabajo consideramos que  $X$  y  $Y$  son de rango completo, es decir:

$$\text{rango}(X) = p$$

$$\text{rango}(Y) = q$$

Entonces

$$\text{Dim } \Psi_X = p$$

$$\text{Dim } \Psi_Y = q$$

Además, si  $X$  ( $N \times p$ ) y  $Y$  ( $N \times q$ ) son de rango completo,  $X'X$  y  $Y'Y$  son matrices no singulares (teorema A.14)

A continuación, centramos nuestra atención en la búsqueda de la primera pareja de variables  $u$  y  $v$  que minimizan el ángulo  $\theta$  y no son las soluciones triviales. Después buscaremos las  $k-1$  parejas restantes ( $k = \min\{p, q\}$ ). Estas  $k$  parejas de variables las tomamos tales que representen en orden decreciente los "parecidos" entre  $X$  y  $Y$ ; y se buscan tales que  $\{u_1, u_2, \dots, u_k\}$  y  $\{v_1, v_2, \dots, v_k\}$  formen bases ortonormales de  $\Psi_X$  y  $\Psi_Y$  respectivamente.

(\*) Si  $\text{rango}(X) < p$  ó  $\text{rango}(Y) < q$  ver Lebart, Morineau y Warwick (1984)

## II.2.- OBTENCION DE LAS VARIABLES Y LAS CORRELACIONES CANONICAS.

(UN PUNTO DE VISTA GEOMÉTRICO)

El problema consiste en encontrar los vectores  $\underline{u}$  y  $\underline{v}$  tales que  $\theta = \angle(\underline{u}, \underline{v})$  sea mínimo.

$$\underline{u} = X\underline{a} \quad \underline{v} = Y\underline{b}$$

La amplitud del ángulo  $\theta$  no depende de la longitud o norma de los vectores  $\underline{u}$  y  $\underline{v}$ ; por lo tanto, para efectos de unicidad, les imponemos la restricción de que sean tales que

$$\|\underline{u}\| = \|\underline{v}\| = 1$$

Debe notarse que esto implica que:

$$\underline{u}'\underline{u} = \underline{v}'\underline{v} = 1$$

Consideremos en primer lugar los casos extremos:

$$1) \psi_x = \psi_y$$

En este caso los conjuntos de variables son equivalentes en el sentido de que la información contenida en las variables de  $X$  es idéntica a la información contenida en las variables de  $Y$ , es decir, ambos conjuntos de variables están explicando el mismo fenómeno; en tal caso existen dos combinaciones lineales distintas de las triviales tales que  $\theta = 0$ , a saber, cualesquiera dos vectores iguales  $\underline{u}$  y  $\underline{v}$  satisfacen  $\theta = 0$  (ó  $\cos \theta = r(\underline{u}, \underline{v}) = 1$ ).

$$2) \psi_x = \psi_y^\perp = \{ \underline{u} \in \mathbb{R}^N \mid \langle \underline{u}, \underline{v} \rangle = 0 \ \forall \underline{v} \in \psi_y \}$$

(ver definición 10 del apéndice)

En este caso los conjuntos de variables son "lo menos parecidos" posible en el sentido de que no existe ninguna relación lineal entre las variables de  $X$  y las variables de  $Y$ . De acuerdo a la definición de complemento ortogonal, cualquier pareja de vectores  $\underline{u}$  y  $\underline{v}$ ,  $\underline{u} \in \psi_x$  y  $\underline{v} \in \psi_y$ , satisfacen  $\theta = \pm \pi$ , es decir  $\cos \theta = r(\underline{u}, \underline{v}) = 0$ .

Nótese que en el caso 1) la intersección de los subespacios es uno de ellos, mientras que en el caso 2) la intersección de los subespacios  $\psi_x$  y  $\psi_y$  es únicamente el vector cero.

En general no se dan estos casos extremos. Consideremos que  $\psi_x \neq \psi_y$  y que  $\psi_x \neq \psi_y^\perp$ , entonces  $\{0\} \subsetneq \psi_x \cap \psi_y$  y  $\psi_x \cap \psi_y$  es distinta de  $\psi_x$  y de  $\psi_y$ .

Encontrar parejas de variables  $\underline{u}$  y  $\underline{v}$  no triviales tales que  $\theta$  sea mínimo es equivalente a encontrar los tales que  $\cos^2(\underline{u}, \underline{v})$  sea máximo.

Supongamos  $\underline{u}$  conocido.

El vector de  $\psi_y$  "más cercano" a  $\underline{u}$  es la proyección ortogonal de  $\underline{u}$  sobre  $\psi_y$ , a la cual denotamos por  $P_y(\underline{u})$ .

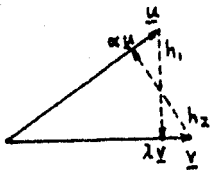
Sea  $\underline{v}$  colineal a  $P_y(\underline{u})$ , entonces existe  $\lambda \in \mathbb{R}$  tal que

$$P_y(\underline{u}) = \lambda \underline{v} \dots \dots \dots (3)$$

Análogamente, la proyección ortogonal de  $y$  sobre  $\psi_x$ , denotada por  $P_x(y)$ , satisface:

$$P_x(y) = \alpha \underline{u} \quad \text{para alguna } \alpha \in \mathbb{R} \dots \dots (4)$$

Nótese que es equivalente maximizar  $\cos(\underline{u}, \underline{y})$  que maximizar  $\cos(\underline{u}, P_y(\underline{u}))$  ó  $\cos(P_x(\underline{y}), \underline{y})$  debido a que el ángulo entre dos vectores no depende de la norma de los vectores.



GRAFICA 6

$$h_1 = \underline{u} - \lambda \underline{y}$$

$$h_2 = \underline{y} - \alpha \underline{u}$$

Para obtener el valor de  $\lambda$  basta resolver la siguiente ecuación:

$$\langle h_1, \underline{y} \rangle = 0$$

$$\begin{aligned} \text{pero } \langle h_1, \underline{y} \rangle &= \langle \underline{u} - \lambda \underline{y}, \underline{y} \rangle = \langle \underline{u}, \underline{y} \rangle - \lambda \langle \underline{y}, \underline{y} \rangle \\ &= \langle \underline{u}, \underline{y} \rangle - \lambda \|\underline{y}\|^2 \end{aligned}$$

$$\therefore \langle h_1, \underline{y} \rangle = 0 \Leftrightarrow \lambda = \frac{\langle \underline{u}, \underline{y} \rangle}{\|\underline{y}\|^2} = \langle \underline{u}, \underline{y} \rangle \dots (5)$$

Análogamente obtenemos:

$$\alpha = \frac{\langle \underline{u}, \underline{y} \rangle}{\|\underline{u}\|^2} = \langle \underline{u}, \underline{y} \rangle \dots (6)$$

Por lo tanto

$$\alpha = \lambda = \langle \underline{u}, \underline{y} \rangle \dots \dots \dots (7)$$

Sean  $\hat{u} = P_Y(\underline{u})$  y  $\hat{v} = P_X(\underline{v})$   
 entonces las ecuaciones (3) y (4) podemos escribirlas  
 como:

$$\hat{u} = \lambda \underline{v} \dots \dots \dots (3)$$

$$\hat{v} = \lambda \underline{u} \dots \dots \dots (4)$$

Hagamos un paréntesis para encontrar las matrices  $A_1$  y  $A_2$  asociadas a las transformaciones lineales  $P_X$  y  $P_Y$  respectivamente. Es decir, buscamos  $A_1$  y  $A_2$  tales que:

$$\hat{u} = A_1 \underline{u}$$

$$\hat{v} = A_2 \underline{v}$$

Como  $\hat{u}$  es colineal a  $\underline{v}$ , entonces:

$$\hat{u} = \sigma \underline{v} = \sigma Y \underline{b} \quad \text{para alguna } \sigma \in \mathbb{R} \dots (10)$$

Por definición

$$\langle \underline{u} - \hat{u}, \underline{v} \rangle = (\underline{u} - \hat{u})' \underline{v} = 0$$

Como cada una de las columnas  $\underline{y}^j$  de  $Y$  ( $j=1, \dots, q$ ) está en el subespacio  $\mathcal{V}_Y$ , se satisface en particular:

$$(\underline{u} - \hat{u})' \underline{y}^j = 0 \quad \forall j=1, \dots, q$$

Entonces

$$\underline{u}' \underline{y}^j = \hat{u}' \underline{y}^j$$

Por lo tanto

$$\underline{u}' Y = \hat{u}' Y \dots \dots \dots (11)$$

Sustituyendo (3) en (9) tenemos:

$$\underline{u}' Y = \sigma \underline{b}' Y' Y \dots \dots \dots (12)$$

pero  $Y$  es de rango completo, entonces  $Y'Y$  es invertible

por lo tanto podemos multiplicar (12) a la derecha por  $(Y'Y)^{-1}$ , de donde:

$$\underline{u}' Y (Y'Y)^{-1} = \sigma \underline{b}'$$

Por lo tanto:

$$\sigma \underline{b} = (Y'Y)^{-1} Y' \underline{u}$$

Entonces:

$$\sigma Y \underline{b} = Y (Y'Y)^{-1} Y' \underline{u}$$

es decir:

$$\hat{u} = Y (Y'Y)^{-1} Y' \underline{u} \quad \dots \quad (13)$$

Por lo tanto:

$$A_1 = Y (Y'Y)^{-1} Y' \underline{u} \quad (14)$$

Análogamente puede obtenerse que:

$$\hat{v} = X (X'X)^{-1} X' \underline{v} \quad \dots \quad (15)$$

de donde

$$A_2 = X (X'X)^{-1} X' \quad \dots \quad (16)$$

Demostremos que  $A_1$  y  $A_2$  son matrices simétricas e idempotentes; por lo tanto  $P_x$  y  $P_y$  son operadores de proyección ortogonal en  $\mathcal{V}_x$  y  $\mathcal{V}_y$  respectivamente.

Demostración:

$$A_1' = [Y (Y'Y)^{-1} Y']' = (Y')' [(Y'Y)']^{-1} Y' = Y (Y'Y)^{-1} Y' = A_1$$

$\therefore A_1$  es simétrica

$$A_1 A_1 = [Y (Y'Y)^{-1} Y'] [Y (Y'Y)^{-1} Y'] = Y (Y'Y)^{-1} \cancel{(Y'Y)} \cancel{(Y'Y)}^{-1} Y' = Y (Y'Y)^{-1} Y' = A_1$$

$\therefore A_1$  es idempotente

Recordemos que  $\hat{u} = A_1 u$  y  $\hat{v} = A_2 v$ , pero

como

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|} \quad \text{y} \quad \|u\| = \|v\| = 1, \text{ entonces}$$

$$\cos \theta = \langle u, v \rangle$$

En (5) obtuvimos que  $\lambda = \langle u, v \rangle$ .

Por lo tanto:

$$\lambda = \cos \theta$$

Sustituyendo este resultado en (8) y (9) resulta:

$$\hat{u} = A_1 u = \cos \theta v \quad \dots \dots \dots (17)$$

$$\hat{v} = A_2 v = \cos \theta u \quad \dots \dots \dots (18)$$

Regresando a la búsqueda de la primera pareja  $u$  y  $v$ , proyectemos ortogonalmente a  $\hat{u}$  y a  $\hat{v}$  sobre  $\psi_x$  y  $\psi_y$  respectivamente, es decir, hagamos  $P_x(\hat{u})$  y  $P_y(\hat{v})$ . Esto es equivalente a multiplicar a la izquierda por  $A_2$  y  $A_1$  a las ecuaciones (17) y (18) respectivamente:

$$A_2 \hat{u} = A_2 \cos \theta v = \cos \theta (A_2 v) = \cos \theta (\cos \theta u) = \cos^2 \theta u$$

$$\text{pero } A_2 \hat{u} = A_2 A_1 u$$

Entonces obtenemos:

$$A_2 A_1 u = \cos^2 \theta u \quad (19)$$

y análogamente:

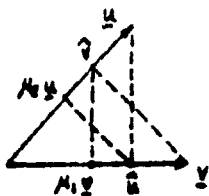
$$A_1 A_2 v = \cos^2 \theta v \quad (20)$$



Al proyectar  $\hat{u}$  sobre  $\Psi_x$  y  $\hat{v}$  sobre  $\Psi_y$ , se satisface que:

$$P_x(\hat{u}) = \mu_1 \underline{u} \quad \text{para alguna } \mu_1 \in \mathbb{R}$$

$$P_y(\hat{v}) = \mu_2 \underline{v} \quad \text{para alguna } \mu_2 \in \mathbb{R}$$



GRAFICA 7

pero  $\mu_1 = \mu_2 = \cos^2 \theta$

Sea  $\mu = \cos^2 \theta$

entonces  $0 \leq \mu \leq 1$

y las ecuaciones (19) y (20) pueden escribirse

como:

$$A_2 A_1 \underline{u} = \mu \underline{u} \quad (21)$$

$$A_1 A_2 \underline{v} = \mu \underline{v} \quad (22)$$

Nótese que  $\underline{u}$  y  $\underline{v}$  son los eigenvectores de las matrices  $A_2 A_1$  y  $A_1 A_2$  respectivamente, asociados ambos al eigenvalor  $\mu$ .

Para encontrar el valor de  $\mu$ , debemos resolver la ecuación:

$$|A_1 A_2 - \mu I| = 0 \quad (\text{ver definición de eigenvalor})$$

La solución de esta ecuación no es única,

hay  $k$  valores de  $\mu$  que la satisfacen.

Sean  $\mu_1, \dots, \mu_k$  los eigenvalores de  $A_1 A_2$  (ó de  $A_2 A_1$ ) ordenados en forma decreciente. Entonces la primera pareja de variables canónicas  $\underline{u}$  y  $\underline{v}$  son los eigenvectores de  $A_2 A_1$  y de  $A_1 A_2$  respectivamente, correspondientes al eigenvalor  $\mu_1$ , que es el mayor. Las siguientes parejas de variables  $\underline{u}_2$  y  $\underline{v}_2$  ( $i=2, \dots, k$ ) son los eigenvectores de  $A_2 A_1$  y  $A_1 A_2$  respectivamente, correspondientes a los eigenvalores  $\mu_2, \dots, \mu_k$ .

En esta sección hemos obtenido las variables canónicas basándonos en el criterio de que estas deben ser tales que minimicen el ángulo  $\theta$ . En la próxima sección haremos un desarrollo similar basándonos en la maximización de la correlación entre ellas. Ambos criterios son equivalentes (p.p. 37) pero desde el punto de vista teórico se presenta una diferencia: geométricamente ni siquiera nos cuestionamos la existencia de un ángulo mínimo, sin embargo algebraicamente no es tan obvia la existencia del máximo de la función de correlación.

Para finalizar nótese que el mayor eigenvalor  $\mu_1$  es el coseno cuadrado del menor ángulo entre una pareja de variables canónicas y equivalentemente es la mayor correlación al cuadrado entre las variables de  $X$  y de  $Y$ .

## II.3.- OBTENCION DE LAS VARIABLES Y LAS CORRELACIONES CANONICAS.

(UN PUNTO DE VISTA ALGEBRAICO)

El problema de obtención de las variables canónicas lo dividimos en dos partes: obtención de la primera pareja de variables canónicas y obtención de las  $k-1$  restantes parejas de variables canónicas.

Para el tratamiento de este problema consideramos que las columnas de  $X$  son linealmente independientes y que las columnas de  $Y$  son linealmente independientes, entonces, por (A.13),  $X'X$  y  $Y'Y$  son invertibles; por lo tanto  $R_{11} = \frac{1}{N} X'X$  y  $R_{22} = \frac{1}{N} Y'Y$  son invertibles.

### II.3.1. - OBTENCION DE LA PRIMERA PAREJA DE VARIABLES CANONICAS.

Podemos plantear el problema como la búsqueda de las combinaciones lineales

$$\underline{u} = X\underline{a} \quad \text{y} \quad \underline{v} = Y\underline{b} \quad \dots \dots \dots (23)$$

en donde  $\underline{a}$  y  $\underline{b}$  deben ser tales que maximicen la correlación entre  $\underline{u}$  y  $\underline{v}$ .

La correlación entre  $\underline{u}$  y  $\underline{v}$  expresada como función de  $\underline{a}$  y  $\underline{b}$ , la podemos obtener sustituyendo

(3) en (2):

$$r = \frac{\frac{1}{N} (\underline{a}' X' Y \underline{b})}{\left[ \frac{1}{N} (\underline{a}' X' X \underline{a}) \frac{1}{N} (\underline{b}' Y' Y \underline{b}) \right]^{1/2}}$$

$$\therefore r = \frac{\underline{a}' \left( \frac{1}{N} X' Y \right) \underline{b}}{\left[ \underline{a}' \left( \frac{1}{N} X' X \right) \underline{a} \quad \underline{b}' \left( \frac{1}{N} Y' Y \right) \underline{b} \right]^{1/2}}$$

Sustituyendo (1) resulta:

$$r = \frac{\underline{a}' R_{12} \underline{b}}{\left[ (\underline{a}' R_{11} \underline{a}) (\underline{b}' R_{22} \underline{b}) \right]^{1/2}} \dots \dots \dots (24)$$

Debido a que los coeficientes de correlación son independientes de la escala de las variables, a los vectores  $\underline{a}$  y  $\underline{b}$  les imponemos la restricción de que sean tales que  $\underline{u}$  y  $\underline{v}$  tengan varianza unitaria, es decir:

$$\text{var}(\underline{u}) = \underline{a}' R_{11} \underline{a} = 1 \dots \dots \dots (25)$$

$$\text{var}(\underline{v}) = \underline{b}' R_{22} \underline{b} = 1 \dots \dots \dots (26)$$

Tomando en cuenta estas restricciones, la ecuación (24) queda expresada como:

$$r = \underline{a}' R_{12} \underline{b} \dots \dots \dots (27)$$

Llamemos  $f(\underline{a}, \underline{b})$  a la función que queremos maximizar sujeta a las restricciones (25) y (26) que escribiremos también como funciones de  $(\underline{a}, \underline{b})$ . Así

el problema se reduce a encontrar  $\underline{a}$  y  $\underline{b}$  tales que maximicen la correlación

$$f(\underline{a}, \underline{b}) = \underline{a}' R_{12} \underline{b} \dots \dots \dots (28)$$

sujeta a las restricciones:

$$g_1(\underline{a}, \underline{b}) = g_1(a_1, \dots, a_p, b_1, \dots, b_q) = \underline{a}' R_{11} \underline{a} - 1 = 0 \dots (29)$$

$$g_2(\underline{a}, \underline{b}) = g_2(a_1, \dots, a_p, b_1, \dots, b_q) = \underline{b}' R_{22} \underline{b} - 1 = 0 \dots (30)$$

En primer lugar demostraremos que la función  $f$ , sujeta a  $g_1$  y  $g_2$  alcanza un máximo.

Demostración:

Como  $f$  es una función bilineal (ver nota 1 en el apéndice) entonces  $f$  es una función continua (por teorema A.1), además  $f$  es de valor real; por lo tanto, si tomamos en cuenta el teorema A.2, el problema se reduce a demostrar que el subconjunto de  $\text{Dom}(f)$  determinado por las restricciones es un compacto, es decir, que es cerrado y acotado (Teorema de Heine-Borel)

Sean  $T_x$  y  $T_y$  transformaciones lineales tales que:

$$T_x: \mathbb{R}^p \rightarrow \mathbb{R}^N \rightarrow \underline{a} \rightarrow X_a \quad T_y: \mathbb{R}^q \rightarrow \mathbb{R}^N \rightarrow \underline{b} \rightarrow Y_b$$

Si nos fijamos en las restricciones

$$\underline{a}' R_{11} \underline{a} = \frac{1}{N} \underline{a}' X' X \underline{a} = \frac{1}{N} \|X \underline{a}\|^2 = 1$$

$$\underline{b}' R_{22} \underline{b} = \frac{1}{N} \underline{b}' Y' Y \underline{b} = \frac{1}{N} \|Y \underline{b}\|^2 = 1$$

resulta que:

$$\|X \underline{a}\| = \sqrt{N} \quad \|Y \underline{b}\| = \sqrt{N}$$

Definimos

$$\mathcal{C} = \{ \underline{c} \in \mathbb{R}^N \text{ t. } \|\underline{c}\| = \sqrt{N} \} \text{ cerrado en } \mathbb{R}^N$$

Entonces la imagen inversa de  $\mathcal{C}$  bajo  $T_x$  que denotamos por  $\mathcal{A} = \{ \underline{a} \in \mathbb{R}^p \text{ t. } \|X \underline{a}\| = \sqrt{N} \}$  y la imagen inversa de  $\mathcal{C}$  bajo  $T_y$  que denotamos por  $\mathcal{B} = \{ \underline{b} \in \mathbb{R}^q \text{ t. } \|Y \underline{b}\| = \sqrt{N} \}$  son cerrados en  $\mathbb{R}^p$  y en  $\mathbb{R}^q$  respectivamente.

Por otro lado tenemos que  $T_x$  y  $T_y$  son transformaciones inyectivas porque

$$T_x(\underline{a}) = X \underline{a} = 0 \Rightarrow \underline{a} = 0 \text{ pues los vectores columna de la matriz } X \text{ son linealmente independientes.}$$

De manera análoga

$$T_y(\underline{b}) = Y \underline{b} = 0 \Rightarrow \underline{b} = 0$$

Por lo tanto, por el teorema A.3, concluimos que  $\exists m_1, m_2 > 0$  t.

$$\|T_x(\underline{a})\| = \|X \underline{a}\| \geq m_1 \|\underline{a}\| \quad \forall \underline{a} \in \mathbb{R}^p$$

$$\|T_y(\underline{b})\| = \|Y \underline{b}\| \geq m_2 \|\underline{b}\| \quad \forall \underline{b} \in \mathbb{R}^q$$

pero como  $\|x_a\| = \|y_b\| = \sqrt{N}$  si  $a \in A$  y  $b \in B$   
entonces

$$\|a\| \leq \frac{\sqrt{N}}{m_1} \quad \forall a \in A$$

$$\|b\| \leq \frac{\sqrt{N}}{m_2} \quad \forall b \in B$$

Con esto queda demostrado que  $A$  y  $B$  son acotados, y como también son cerrados

$A \times B = \{(a, b) \in \mathbb{R}^N : a \in A \wedge b \in B\}$   
es un compacto en  $\mathbb{R}^N$

Nótese que  $A \times B \subset \text{Dom}(f)$  es un compacto en  $\mathbb{R}^N$  y como  $f$  es continua de valor real, entonces existe un punto de  $A \times B$  en donde  $f$  alcanza el máximo. sujeta a las restricciones; sea  $(a, b)$  este punto.

Como se tienen dos restricciones y sabemos que  $f$  sujeta a ellas alcanza un máximo en  $(a, b)$  aplicamos la generalización del teorema A.4, con lo que resulta que existen escalares  $\lambda$  y  $\alpha$  tales que:

$$\nabla f(a, b) = \lambda \nabla g_1(a, b) + \alpha \nabla g_2(a, b) \dots (31)$$

donde  $\lambda$  y  $\alpha$  son multiplicadores de Lagrange.

Por lo tanto, debemos encontrar  $a, b, \lambda$  y  $\alpha$  que satisfagan (31)

Recordemos que:

$$f(\underline{a}, \underline{b}) = \underline{a}' R_{12} \underline{b}$$

$$= b_1 (a_1 S_{1,p+1} + \dots + a_p S_{p,p+1}) + \dots + b_q (a_1 S_{1n} + \dots + a_p S_{pn})$$

$$g_1(\underline{a}, \underline{b}) = \underline{a}' R_{11} \underline{a} - 1$$

$$= a_1 (a_1 S_1^2 + \dots + a_p S_{1p}) + a_2 (a_1 S_{12} + a_2 S_2^2 + \dots + a_p S_{2p}) + \dots$$

$$\dots + a_p (a_1 S_{1p} + \dots + a_p S_p^2) - 1$$

$$g_2(\underline{a}, \underline{b}) = \underline{b}' R_{22} \underline{b} - 1$$

$$= b_1 (b_1 S_{p+1}^2 + \dots + b_q S_{p+1,n}) + b_2 (b_1 S_{p+1,p+2} + b_2 S_{p+2}^2 + \dots + b_q S_{p+2,n}) +$$

$$\dots + b_q (b_1 S_{p+1,n} + \dots + b_q S_n^2) - 1$$

y notemos la analogía de los términos de  $g_1$  y  $g_2$

Calculemos los gradientes indicados en (31):

$$\nabla f(\underline{a}, \underline{b}) = \left( \frac{df}{da_1}, \dots, \frac{df}{da_p}, \frac{df}{db_1}, \dots, \frac{df}{db_q} \right)$$

$$= (b_1 S_{1,p+1} + \dots + b_q S_{1n}, \dots, b_1 S_{p,p+1} + \dots + b_q S_{pn},$$

$$a_1 S_{1,p+1} + \dots + a_p S_{p,p+1}, \dots, a_1 S_{1n} + \dots + a_p S_{pn})$$

$$\therefore \nabla f(\underline{a}, \underline{b}) = (\underline{b}' R_{12}, \underline{a}' R_{12}) \dots \dots \dots (32)$$

$$\nabla g_1(\underline{a}, \underline{b}) = \left( \frac{dg_1}{da_1}, \dots, \frac{dg_1}{da_p}, \frac{dg_1}{db_1}, \dots, \frac{dg_1}{db_q} \right)$$

$$= (2a_1 S_1^2 + 2a_2 S_{12} + \dots + 2a_p S_{1p}, \dots, 2a_1 S_{1p} + 2a_2 S_{2p} + \dots + 2a_p S_p^2,$$

$$0, \dots, 0)$$

$$\therefore \nabla g_1(\underline{a}, \underline{b}) = 2(\underline{a}' R_{11}, 0) \dots \dots \dots (33)$$



Análogamente se calcula:

$$\nabla g_2(\underline{a}, \underline{b}) = 2(\underline{0}, \underline{b}' R_{22}) \dots \dots \dots (34)$$

Sustituyendo (32), (33) y (34) en (31) obtenemos:

$$(\underline{b}' R'_{12}, \underline{a}' R_{12}) = 2\lambda(\underline{a}' R_{11}, \underline{0}) + 2\alpha(\underline{0}, \underline{b}' R_{22})$$

$$\Leftrightarrow (\underline{b}' R'_{12}, \underline{a}' R_{12}) = (2\lambda \underline{a}' R_{11}, 2\alpha \underline{b}' R_{22}) \dots \dots (35)$$

Si igualamos las componentes y tomamos en cuenta las restricciones, obtenemos un sistema de ecuaciones de  $n+2$  ecuaciones y  $n+2$  incógnitas:

$$\left\{ \begin{array}{l} \underline{b}' R'_{12} - 2\lambda \underline{a}' R_{11} = \underline{0} \text{ (1} \times \text{p)} \dots \dots \dots (36) \\ \underline{a}' R_{12} - 2\alpha \underline{b}' R_{22} = \underline{0} \text{ (1} \times \text{q)} \dots \dots \dots (37) \end{array} \right.$$

$$\left\{ \begin{array}{l} \underline{a}' R_{11} \underline{a} = 1 \dots \dots \dots (38) \\ \underline{b}' R_{22} \underline{b} = 1 \dots \dots \dots (39) \end{array} \right.$$

Nótese que  $R_{11} = R'_{11}$  y  $R_{22} = R'_{22}$ , entonces el sistema de ecuaciones anterior es equivalente al siguiente:

$$\left\{ \begin{array}{l} R_{12} \underline{b} - 2\lambda R_{11} \underline{a} = \underline{0} \text{ (p} \times \text{1)} \dots \dots \dots (40) \\ R'_{12} \underline{a} - 2\alpha R_{22} \underline{b} = \underline{0} \text{ (q} \times \text{1)} \dots \dots \dots (41) \end{array} \right.$$

$$\left\{ \begin{array}{l} \underline{a}' R_{11} \underline{a} = 1 \dots \dots \dots (42) \\ \underline{b}' R_{22} \underline{b} = 1 \dots \dots \dots (43) \end{array} \right.$$

Este sistema tiene solución pues  $R_{11}$  y  $R_{22}$  son matrices no singulares.

Multiplicando a la izquierda (40) por  $\underline{a}'$  y (41) por  $\underline{b}'$  resulta:

$$\underline{a}' R_{12} \underline{b} - 2\lambda \underline{a}' R_{11} \underline{a} = 0 \dots\dots\dots (44)$$

$$\underline{b}' R_{12}' \underline{a} - 2\alpha \underline{b}' R_{22} \underline{b} = 0 \dots\dots\dots (45)$$

y sustituyendo (42) en (44) y (43) en (45) obtenemos:

$$\underline{a}' R_{12} \underline{b} - 2\lambda = 0 \dots\dots\dots (46)$$

$$\underline{b}' R_{12}' \underline{a} - 2\alpha = 0 \dots\dots\dots (47)$$

Pero  $\underline{a}' R_{12} \underline{b} \in \mathbb{R}$ , por lo tanto:

$$(\underline{a}' R_{12} \underline{b})' = \underline{a}' R_{12} \underline{b}$$

$$\Leftrightarrow \underline{b}' R_{12}' \underline{a} = \underline{a}' R_{12} \underline{b} \dots\dots\dots (48)$$

y si sustituimos (48) en (47) resulta que:

$$\lambda = \alpha \dots\dots\dots (49)$$

Multiplicando a la izquierda (40) por  $R_{21} R_{11}^{-1}$  y sustituyendo (49) en (41) obtenemos:

$$R_{12}' \underline{a} - 2\lambda R_{22} \underline{b} = \underline{0} \dots\dots\dots (50)$$

$$R_{21} R_{11}^{-1} R_{12} \underline{b} - 2\lambda R_{12}' \underline{a} = \underline{0} \dots\dots\dots (51)$$

La ecuación (50) podemos escribirla como:

$$R_{12}' \underline{a} = 2 \lambda R_{22} \underline{b} \dots \dots \dots (52)$$

Sustituyendo (52) en (51) resulta:

$$(R_{21} R_{11}^{-1} R_{12} - 4 \lambda^2 R_{22}) \underline{b} = \underline{0}_{(q \times 1)}$$

que es equivalente a la ecuación

$$(R_{22}^{-1} R_{12}' R_{11}^{-1} R_{12}) \underline{b} = \beta^2 \underline{b} \dots \dots \dots (53)$$

con  $\beta = 2\lambda$

$$\text{Sea } B = R_{22}^{-1} R_{12}' R_{11}^{-1} R_{12}$$

Nótese que existen  $\beta$  y  $\underline{b} \neq 0$  que satisfacen esta ecuación, puesto que  $f$  alcanza un máximo sujeta a las restricciones; por lo tanto  $\underline{b}$  es un eigenvector de la matriz  $B$  y  $\beta^2$  es el eigenvalor correspondiente al eigenvector  $\underline{b}$ .

Como la matriz  $B$  es una matriz cuadrada ( $q \times q$ ), a lo más va a tener  $q$  eigenvalores distintos de cero.

Haciendo un procedimiento análogo a partir de la ecuación (29) podemos obtener la ecuación

$$(R_{11}^{-1} R_{12} R_{22}^{-1} R_{12}') \underline{a} = \beta^2 \underline{a} \dots \dots \dots (54)$$

No es necesario hacer esto para obtener  $\lambda = \mu$ ,  $\underline{a}$  y  $\underline{b}$  que maximicen  $f$ , pues  $\lambda$  y  $\underline{b}$  se ob-

tienen directamente de (53) y  $\underline{a}$  puede obtenerse sustituyendo  $\underline{b}$  y  $\lambda$  en (40). Sin embargo podemos observar que la matriz  $A = (R_{11}^{-1} R_{12} R_{22}^{-1} R_{12})$  es una matriz cuadrada ( $p \times p$ ), por lo tanto va a tener a lo más  $p$  eigenvalores distintos de cero. Entonces en general, el número de eigenvalores distintos de cero es, a lo más,  $k = \min \{ p, q \}$ .

La forma habitual de resolver la ecuación (53) es calculando las raíces de la ecuación

$$| R_{22}^{-1} R_{12} R_{11}^{-1} R_{12} - \beta^2 I | = 0$$

en donde el determinante se conoce como el polinomio característico de la matriz  $R_{22}^{-1} R_{12} R_{11}^{-1} R_{12}$ .

Nótese que intentando obtener la primera pareja de variables canónicas, hemos llegado a una ecuación que nos permite obtener todas las parejas de variables canónicas. Así, el punto siguiente no será la obtención de las restantes parejas, sino

### II.3.2.- CARACTERIZACION DE LOS EIGENVALORES DE B Y SUS EIGENVECTORES.

Como las columnas de  $X (N \times p)$  son linealmente independientes y las columnas de  $Y (N \times q)$  son linealmente independientes, por el teorema (A.5) tenemos que  $R_{11} = \frac{1}{N} X'X$  y  $R_{22} = \frac{1}{N} Y'Y$  son matrices positivas

definidas (p.d.). Además  $R_{11}$  y  $R_{22}$  son simétricas, entonces (ver nota 2 en el apéndice) podemos afirmar que existen matrices  $P$  ( $p \times p$ ) y  $Q$  ( $q \times q$ ) invertibles y simétricas tales que:

$$PP = R_{11} \quad \text{y} \quad QQ = R_{22}$$

Así:

$$B = R_{22}^{-1} R_{12}' R_{11}^{-1} R_{12} = (QQ)^{-1} R_{12}' (PP)^{-1} R_{12} \dots (35)$$

$$A = R_{11}^{-1} R_{12} R_{22}^{-1} R_{12}' = (PP)^{-1} R_{12} (QQ)^{-1} R_{12}' \dots (36)$$

Si multiplicamos  $B$  a la izquierda por  $Q$  y a la derecha por  $Q^{-1}$  obtenemos la matriz

$$N_2 = Q^{-1} R_{12}' P^{-1} P^{-1} R_{12} Q^{-1}$$

Y si multiplicamos a la matriz  $A$  a la izquierda por  $P$  y a la derecha por  $P^{-1}$ , obtenemos la matriz

$$N_1 = P^{-1} R_{12} Q^{-1} Q^{-1} R_{12}' P^{-1}$$

Sea  $K = P^{-1} R_{12} Q^{-1}$

Entonces

$$N_1 = K K' \quad \text{y} \quad N_2 = K' K$$

Como  $K$  ( $p \times q$ ) y  $K'$  ( $q \times p$ ), por teorema (A.6) tenemos que  $N_1$  y  $N_2$  tienen los mismos eigenvalores distintos de cero; además por el teorema (A.7) resul-

ta que  $A$  y  $N_1$  tienen idénticos eigenvalores, así como  $B$  y  $N_2$ . Por lo tanto  $A, B, N_1$  y  $N_2$  tienen los mismos eigenvalores distintos de cero, y todos ellos son positivos porque  $\beta^2 \geq 0 \quad \forall \beta \in \mathbb{R}$ .

Sea  $k$  el número de eigenvalores distintos de cero, entonces por (A.8)  $k = \text{rango}(K)$ . Pero como  $P$  y  $Q$  son matrices no singulares, por (A.9) tenemos que  $\text{rango}(K) = \text{rango}(R_{12})$

Resumiendo tenemos que

$$k = \text{rango}(K) = \text{rango}(R_{12})$$

Sean los eigenvalores, ordenados en forma decreciente:

$$\beta_1^2 > \beta_2^2 > \dots > \beta_k^2 > 0$$

(A.11) puede escribirse como

$$K = U D V'$$

donde  $U$  y  $V$ , de orden  $p \times k$  y  $q \times k$  respectivamente, son matrices de columnas ortonormales, en particular las columnas de  $U$  son los eigenvectores estandarizados de  $N_1 = K K'$  y las columnas de  $V$  son los eigenvectores estandarizados de  $N_2 = K' K$ . La matriz  $D$  es una matriz diagonal con elementos positivos; más aún, los elementos de la diagonal de  $D$  son las raíces cuadradas de los eigenvalores de  $K' K$ .

Por lo tanto:

$$K = (\alpha_1, \dots, \alpha_k) D (\gamma_1, \dots, \gamma_k)'$$

donde  $\alpha_i$  y  $\gamma_i$  ( $i=1, \dots, k$ ) son los eigenvectores estandarizados de  $N_1$  y  $N_2$  respectivamente y

$$D = \text{diag}(\sqrt{\beta_1^2}, \dots, \sqrt{\beta_k^2})$$

Si los  $\beta_i$  son distintos, los eigenvectores son únicos en cuanto al signo, pues los escogemos de modo que las raíces cuadradas en  $D$  sean positivas.

Además  $N_1$  y  $N_2$  son simétricas, por lo tanto (A.10) sus eigenvectores son ortonormales. Así:

$$\alpha_i' \alpha_j = \delta_{ij} \quad \gamma_i' \gamma_j = \delta_{ij}$$

donde  $\delta_{ij}$  es la Delta de Kronecker

$$\delta_{ij} = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{si no.} \end{cases}$$

Los eigenvectores de  $A$  y  $B$  son los vectores

$$\underline{a}_i = P^{-1} \alpha_i \quad \text{y} \quad \underline{b}_i = Q^{-1} \gamma_i \quad (i=1, \dots, k)$$

Demostración: (sólo para  $\underline{a}_i$ , la de  $\underline{b}_i$  es análoga)

Recordemos que  $N_1 = P A P^{-1}$ , o equivalentemen-

te  $A = P^{-1} N_1 P$ . Entonces:

$$A \underline{a}_i = \beta_i^2 \underline{a}_i \Leftrightarrow P^{-1} N_1 P \underline{a}_i = \beta_i^2 \underline{a}_i$$

Sustituyendo  $\underline{a}_i = P^{-1} \alpha_i$  obtenemos:

$$P^{-1} N_1 P P^{-1} \alpha_i = \beta_i^2 P^{-1} \alpha_i$$

$$\Leftrightarrow N_1 \alpha_i = \beta_i^2 \alpha_i \quad \text{l.a.a.d.}$$

Por lo tanto los eigenvectores  $\underline{a}_i$ ,  $\underline{b}_i$  soluciones de nuestro problema son ortogonales.

- Definición 4:
- 1) Los vectores  $\underline{a}_i$  y  $\underline{b}_i$  son llamados los  $i$ -ésimos vectores canónicos.
  - 2) Las combinaciones lineales  $\underline{u}_i = X\underline{a}_i$  y  $\underline{v}_i = Y\underline{b}_i$  se conocen como las  $i$ -ésimas variables canónicas.
  - 3)  $\rho_i = \sqrt{\beta_i^2}$  recibe el nombre de  $i$ -ésimo coeficiente de correlación canónica entre  $\underline{u}_i$  y  $\underline{v}_i$

Nótese que:

$$\text{corr}(\underline{u}_i, \underline{u}_j) = \underline{a}_i' R_{11} \underline{a}_j = \alpha_i' \alpha_j = d_{ij}$$

$$\text{corr}(\underline{v}_i, \underline{v}_j) = \underline{b}_i' R_{22} \underline{b}_j = \beta_i' \beta_j = d_{ij}$$

Debemos notar la analogía entre las ecuaciones (22) y (53).

En (22) tenemos que:

$$A_1 A_2 \underline{v} = \mathcal{M} \underline{v}$$

Sustituyendo  $A_1$  y  $A_2$  obtenemos:

$$[Y(Y'Y)^{-1}Y'] [X(X'X)^{-1}X'] \underline{v} = \mathcal{M} \underline{v}$$

Por otro lado, en (53) tenemos

$$R_{22}^{-1} R_{12}' R_{11}^{-1} R_{12} \underline{b} = \beta^2 \underline{b}$$



Sustituyendo (4) en (53) y multiplicando a la izquierda por  $Y$  resulta:

$$Y N (Y'Y)^{-1} \frac{1}{N} (Y'X) N (X'X)^{-1} \frac{1}{N} (X'Y) \underline{b} = \beta^2 Y \underline{b}$$

pero  $\underline{y} = Y \underline{b}$  ; por lo tanto:

$$[Y (Y'Y)^{-1} Y'] [X (X'X)^{-1} X'] \underline{y} = \beta^2 \underline{y}$$

Entonces, el mayor eigenvalor  $\mu_1 = \beta_1^2$  es el cuadrado de la correlación entre  $\underline{u}$  y  $\underline{v}$  ó el coseno cuadrado del menor ángulo entre  $\underline{u}$  y  $\underline{v}$  ; y  $\mu_2 = \beta_2^2$  es la correlación cuadrada entre  $\underline{u}_2$  y  $\underline{v}_2$  ó el coseno cuadrado del menor ángulo entre  $\underline{u}_2$  y  $\underline{v}_2$  ( $2=2, \dots, k = \min \{p, q\}$  )

Por último mostramos que si  $\underline{u}$  y  $\underline{v}$  son una pareja de variables canónicas, cada una de las variables es colineal con la proyección de la otra.

Demostración:

De (40) y (50) obtenemos que

$$\underline{a} = \frac{1}{\beta} [R_{11}^{-1} R_{12} \underline{b}]$$

$$\underline{b} = \frac{1}{\beta} [R_{22}^{-1} R_{21} \underline{a}]$$

Multiplicando estas ecuaciones a la izquierda por  $X$  y  $Y$  respectivamente y sustituyendo (1) resulta:

$$X \underline{a} = \frac{1}{\beta} X (X'X)^{-1} X' Y \underline{b}$$

$$Y \underline{b} = \frac{1}{\beta} Y (Y'Y)^{-1} Y' X \underline{a}$$

o equivalentemente:

$$\underline{u} = \frac{1}{\beta} P_X (Y)$$

$$\underline{v} = \frac{1}{\beta} P_Y (X)$$

L.Q.Q.D.

## CAPITULO III

### Discusión de Resultados

La siguiente tabla muestra las correlaciones canónicas de la primera y de la segunda pareja de variables canónicas en todos los análisis efectuados.

| X \ Y                              | HEMISFERIO<br>IZQUIERDO | HEMISFERIO<br>DERECHO | REGION<br>ANTERIOR | REGION<br>POSTERIOR |
|------------------------------------|-------------------------|-----------------------|--------------------|---------------------|
| ESCOLARIDAD<br>DE LA MADRE         | 0.54<br>0.41            | 0.50<br>0.36          | 0.54<br>0.48       | 0.46<br>0.40        |
| ESCOLARIDAD<br>DEL PADRE           | 0.54<br>0.47            | 0.50<br>0.36          | 0.50<br>0.45       | 0.51<br>0.49        |
| OCUPACION<br>DE LA MADRE           | 0.54<br>0.39            | 0.46<br>0.41          | 0.56<br>0.41       | 0.45<br>0.34        |
| OCUPACION<br>DEL PADRE             | 0.47<br>0.40            | 0.43<br>0.38          | 0.45<br>0.39       | 0.45<br>0.34        |
| SALARIO<br>PERCAPITA               | 0.43<br>0.40            | 0.34<br>0.51          | 0.51<br>0.42       | 0.33<br>0.30        |
| ESTADO DE LA<br>VIVIENDA           | 0.46<br>0.42            | 0.42<br>0.34          | 0.44<br>0.42       | 0.43<br>0.38        |
| ANTECEDENTES<br>SEGUN EXPERTOS     | 0.50<br>0.33            | 0.40<br>0.32          | 0.51<br>0.43       | 0.40<br>0.32        |
| EVALUACION DE LA<br>LECTOESCRITURA | 0.72 *<br>0.51          | 0.64<br>0.44          | 0.71 *<br>0.50     | 0.53<br>0.47        |

TABLA 3

X representa a los grupos de variables electroencefalográficas y Y a los grupos de variables no electroencefalográficas

Como una correlación canónica mayor está asociada a una mayor relación lineal entre los grupos de variables correspondientes, centramos nuestra atención en los A.C.C. entre los grupos:

'HEMISFERIO IZQUIERDO' y 'EVALUACION DE LA LECTOESCRITURA'  
'REGION ANTERIOR' y 'EVALUACION DE LA LECTOESCRITURA'.

En la tabla 3 aparecen marcadas con asteriscos las primeras correlaciones canónicas encontradas en ambos análisis.

La tabla 4 muestra los resultados obtenidos en el A.C.C. efectuado entre los grupos 'HEMISFERIO IZQUIERDO' y 'EVALUACION DE LA LECTOESCRITURA'. En la columna denotada por  $x_i$  aparecen las variables del grupo 'HEMISFERIO IZQUIERDO' y en la columna denotada por  $y_j$  aparecen las variables del grupo 'EVALUACION DE LA LECTOESCRITURA' donde L1 corresponde a una evaluación buena, L2 corresponde a una evaluación regular, L3 corresponde a una evaluación mala y L4 corresponde a una evaluación pésima. (ver página 22). Denotamos por  $u_1$  y  $v_1$  a la primera pareja de variables canónicas ( $u_1$  corresponde al grupo de las variables electroencefalográficas del hemisferio cerebral izquierdo y  $v_1$  al grupo de evaluación de la lectoescritura) y denotamos por  $u_2$  y  $v_2$  a la segunda pareja de variables canónicas ( $u_2$  corresponde al grupo de las variables electroencefalográficas del hemisferio izquierdo y  $v_2$  al grupo de

evaluación de la lectoescritura). Así;  $r(u_i, x^j)$ ,  $r(y_i, y^j)$   
denotan la correlación entre la variable original  
-  $x^j$  ó  $y^j$  - y su respectiva variable canónica -  $u_i$   
ó  $y_i$  -.

En la gráfica 3, en (A) representamos los  
pares ordenados  $(r(u_i, x^j), r(u_i, x^j))$  y en (B)  
representamos los pares ordenados  $(r(y_i, y^j), r(y_i, y^j))$   
Esta gráfica es útil en la interpretación de los  
resultados que mostramos en la tabla 4.

| $\underline{x}^j$ | $\underline{y}^j$ | $r(\underline{u}_1, \underline{x}^j)$ | $r(\underline{v}_1, \underline{y}^j)$ | $r(\underline{u}_2, \underline{x}^j)$ | $r(\underline{v}_2, \underline{y}^j)$ |
|-------------------|-------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| RDCI              |                   | - 0.07                                |                                       | 0.15                                  |                                       |
| RDFI              |                   | - 0.17                                |                                       | 0.20                                  |                                       |
| RDFIL             |                   | - 0.28                                |                                       | 0.41 *                                |                                       |
| RDTI              |                   | - 0.05                                |                                       | 0.45 *                                |                                       |
| RDPI              |                   | 0.09                                  |                                       | - 0.11                                |                                       |
| RDOI              |                   | - 0.14                                |                                       | - 0.21                                |                                       |
| RDTIP             |                   | 0.18                                  |                                       | 0.07                                  |                                       |
| RTC1              |                   | - 0.46 *                              |                                       | 0.07                                  |                                       |
| RTFI              |                   | - 0.37 *                              |                                       | 0.17                                  |                                       |
| RTFIL             |                   | - 0.28                                |                                       | - 0.20                                |                                       |
| RTTI              |                   | - 0.37 *                              |                                       | - 0.01                                |                                       |
| RTPI              |                   | - 0.44 *                              |                                       | 0.06                                  |                                       |
| RTOI              |                   | - 0.45 *                              |                                       | 0.12                                  |                                       |
| RTTIP             |                   | - 0.41 *                              |                                       | - 0.02                                |                                       |
| RACI              |                   | 0.38 *                                |                                       | - 0.17                                |                                       |
| RAFI              |                   | 0.43 *                                |                                       | - 0.30                                |                                       |
| RAFIL             |                   | 0.45 *                                |                                       | - 0.28                                |                                       |
| RATI              |                   | 0.21                                  |                                       | - 0.24                                |                                       |
| RAPI              |                   | 0.27                                  |                                       | - 0.01                                |                                       |
| RAOI              |                   | 0.35 *                                |                                       | 0.05                                  |                                       |
| RATIP             |                   | 0.12                                  |                                       | - 0.02                                |                                       |
|                   | L1                |                                       | 0.96 *                                |                                       | - 0.21                                |
|                   | L2                |                                       | - 0.72 *                              |                                       | - 0.59 *                              |
|                   | L3                |                                       | - 0.26                                |                                       | 0.95 *                                |
|                   | L4                |                                       | - 0.16                                |                                       | - 0.05                                |

TABLA 4 'HEMISFERIO IZQUIERDO' y 'EVALUACION DE LA LECTOESCRITURA'





En la tabla 4, las correlaciones más importantes entre una variable original y la variable canónica correspondiente, aparecen marcadas con asteriscos.

Las correlaciones entre la primera variable canónica  $u_1$  y cada una de las variables de potencia relativa en la banda  $\delta$  no las consideramos importantes (en la gráfica podemos ver que, con respecto al eje de las abscisas, las encontramos concentradas alrededor del cero); mientras que con respecto a las variables de potencia relativa en la banda  $\theta$ , las correlaciones de  $u_1$  son altas y negativas (en la gráfica, con respecto al eje de las abscisas, son las que se encuentran más a la izquierda) excepto la correlación entre  $u_1$  y RTFIL que es negativa pero no la consideramos importante. Las correlaciones de  $u_1$  con respecto a las variables de la banda  $\alpha$  son siempre positivas (en la gráfica se encuentran en el primero y cuarto cuadrante) y consideramos que son importantes las correlaciones entre  $u_1$  y las variables RACI, RAFI, RAFIL y RAOI. Esto nos sugiere que menores valores de las variables de la banda  $\theta$  y mayores valores de las variables de la banda  $\alpha$  (en particular RACI, RAFI, RAFIL y RAOI) aparentemente están relacionados a mayores valores de  $u_1$ .

Por otro lado, la correlación entre la primera variable canónica  $v_1$  y la variable L1 es muy alta y positiva (en la gráfica se localiza a la extrema dere-

cha con respecto al eje de las abscisas) y la correlación entre  $y_1$  y  $L_2$  es alta y negativa (en la gráfica se localiza a la izquierda con respecto al eje de las abscisas). Entonces puede pensarse que una evaluación 'buena' en las pruebas pedagógicas de lectoescritura aparentemente está asociada a un mayor valor de  $y_1$  y que una evaluación 'regular' aparentemente está asociada a un menor valor de  $y_1$ .

Estas interpretaciones sugieren que aparentemente una evaluación de la lectoescritura 'buena' puede estar relacionada con un aumento en los valores de la potencia relativa en la banda  $\alpha$  (en particular RACI, RAFI, RAFIL y RAOI) y una disminución en los valores de la potencia relativa en la banda  $\theta$ , con respecto a una evaluación 'regular'. O equivalentemente, que una evaluación 'regular' puede estar relacionada con una disminución de los valores de las variables en la banda  $\alpha$  y un aumento de los valores de las variables en la banda  $\theta$ , con respecto a una evaluación 'buena'.

Analizamos también lo que ocurre con respecto a la segunda pareja de variables canónicas, aunque con mucha mayor precaución pues la correlación canónica es bastante más baja.

Las únicas correlaciones que consideramos importantes son la correlación entre  $u_2$  y RDFIL y la correlación entre  $u_2$  y RDTI (en la gráfica puede observarse que, con respecto al eje de las ordenadas, son las únicas que no están concentradas alrededor del cero), ambas positivas (en el segundo cuadrante de la gráfica): esto nos hace pensar que un aumento en los valores de estas variables pudiera estar asociado a un incremento en los valores de  $u_2$ .

Por otra parte, observamos una correlación alta negativa entre  $L_2$  y  $y_2$  y una correlación alta positiva entre  $L_3$  y  $y_2$ , lo cual sugiere que una evaluación de la lectoescritura 'mala' puede estar asociada a valores mayores de  $y_2$  y que una evaluación 'regular' puede estar asociada a menores valores de  $y_2$ .

Ambas interpretaciones nos llevan a pensar que una evaluación de la lectoescritura 'mala' puede estar relacionada con mayores valores de RDFIL y RDTI que una evaluación 'regular', y viceversa: que una evaluación 'regular' puede estar asociada a menores valores de RDFIL y RDTI que una evaluación 'mala'. Sin embargo esto debemos considerarlo con muchas reservas pues la correlación canónica entre  $u_2$  y  $y_2$  no es muy alta.

Si tratamos de integrar la interpretación que hemos dado con respecto a la primera pareja de variables canónicas y la interpretación que hemos dado con respecto a la segunda pareja de variables canónicas, podemos pensar que ligeras deficiencias en la lectoescritura aparentemente están relacionadas con una disminución en los valores de la banda  $\alpha$  (en particular RACI, RAFI, RAFIL y RAOI) y con un aumento en los valores de la banda  $\theta$  (quizás, excepto RTFIL) con respecto a un buen rendimiento en las pruebas de lectoescritura; pero que si estas deficiencias son más graves (evaluación 'mala') pudieran estar relacionadas con un aumento de RDTI y RDFI (variables de la banda  $\delta$ ) con respecto a deficiencias ligeras.

A continuación mostramos en la tabla 5, los resultados obtenidos en el A.C.C. efectuado a los grupos 'REGION ANTERIOR' y 'EVALUACION DE LA LECTOESCRITURA'. La notación utilizada en la tabla 5 es análoga a la utilizada en la tabla 4.

Para una mejor comprensión de la discusión puede ser útil la gráfica 9.

| $\underline{x}_j$ | $\underline{y}_j$ | $r(\underline{u}_1, \underline{x}_j)$ | $r(\underline{v}_1, \underline{y}_j)$ | $r(\underline{u}_2, \underline{x}_j)$ | $r(\underline{v}_2, \underline{y}_j)$ |
|-------------------|-------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| RDCI              |                   | 0.13                                  |                                       | 0.18                                  |                                       |
| RDCD              |                   | -0.15                                 |                                       | 0.09                                  |                                       |
| RDFI              |                   | 0.22                                  |                                       | 0.19                                  |                                       |
| RDFD              |                   | 0.10                                  |                                       | 0.01                                  |                                       |
| RDFIL             |                   | 0.37 *                                |                                       | 0.25                                  |                                       |
| RDFPL             |                   | 0.12                                  |                                       | -0.06                                 |                                       |
| RDTI              |                   | 0.16                                  |                                       | 0.45 *                                |                                       |
| RDTD              |                   | -0.13                                 |                                       | -0.08                                 |                                       |
| RTCI              |                   | 0.47 *                                |                                       | -0.08                                 |                                       |
| RTCD              |                   | 0.43 *                                |                                       | -0.12                                 |                                       |
| RTFI              |                   | 0.40 *                                |                                       | -0.00                                 |                                       |
| RTFD              |                   | 0.55 *                                |                                       | -0.01                                 |                                       |
| RTFIL             |                   | 0.23                                  |                                       | -0.31                                 |                                       |
| RTFDL             |                   | 0.54 *                                |                                       | 0.07                                  |                                       |
| RTTI              |                   | 0.36 *                                |                                       | -0.15                                 |                                       |
| RTTD              |                   | 0.47 *                                |                                       | 0.12                                  |                                       |
| RACI              |                   | -0.42 *                               |                                       | -0.08                                 |                                       |
| RACD              |                   | -0.26                                 |                                       | -0.00                                 |                                       |
| RAFI              |                   | -0.49 *                               |                                       | -0.14                                 |                                       |
| RAFD              |                   | -0.51 *                               |                                       | -0.03                                 |                                       |
| RAFIL             |                   | -0.51 *                               |                                       | -0.06                                 |                                       |
| RAFPL             |                   | -0.47 *                               |                                       | -0.04                                 |                                       |
| RATI              |                   | -0.27                                 |                                       | -0.20                                 |                                       |
| RATD              |                   | -0.21                                 |                                       | -0.02                                 |                                       |
|                   | L1                |                                       | -0.99 *                               |                                       | 0.06                                  |
|                   | L2                |                                       | 0.47 *                                |                                       | -0.87 *                               |
|                   | L3                |                                       | 0.54 *                                |                                       | 0.69 *                                |
|                   | L4                |                                       | 0.19                                  |                                       | 0.26                                  |

TABLA 5 'REGION ANTERIOR' y 'EVALUACION DE LA LECTOESCRITURA'



Tomando en cuenta las correlaciones más importantes con respecto a  $u_1$ , se puede observar que RDFIL tiene una correlación positiva, al igual que todas las variables de la banda  $\theta$ , a excepción de RTFIL que es positiva, pero no la consideramos muy correlacionada con  $u_1$  (nótese que en la gráfica 9(A) todas estas variables se localizan en el primero o en el cuarto cuadrante). Todas las variables de la banda  $\alpha$  están correlacionadas negativamente con  $u_1$  (se localizan en el tercer cuadrante), pero consideramos más importantes las correlaciones entre  $u_1$  y las variables RACI, RAFI, RAFD, RAFIL y RAFDL. Esto podemos interpretarlo como que un mayor valor de  $u_1$  está relacionado aparentemente con valores mayores de las variables de la banda  $\theta$  (quizás RTFIL es una excepción), menores valores de las variables de la banda  $\alpha$  (en particular RACI, RAFI, RAFD, RAFIL y RAFDL) y un valor mayor de RDFIL.

Por otro lado, si tomamos en cuenta las correlaciones más importantes con respecto a  $v_1$ , observamos que  $v_1$  y L1 tienen una correlación negativa altísima,  $v_1$  y L2 tienen una correlación positiva y  $v_1$  y L3 tienen una correlación positiva más alta que  $v_1$  y L2 (en la gráfica (B)), con respecto al eje de las abscisas, L1 está en

la extrema izquierda y L2 y L3 están muy a la derecha). Esto nos hace pensar que L1 está asociada a menores valores de  $y_1$  que L2 y L3, y que L2 y L3 están asociadas a mayores valores de  $y_1$ .

Como la correlación entre  $u_1$  y  $y_1$  la consideramos importante, las interpretaciones anteriores nos sugieren que una evaluación de lectoescritura 'buena' está asociada aparentemente con mayores valores de  $\alpha$  (en particular RACI, RAFI, RAFO, RAFIL y RAFDL), menores valores de  $\theta$  y menor valor de RDFIL que una evaluación 'regular' o 'mala'; mientras que una evaluación 'regular' o 'mala' está asociada aparentemente con menores valores de  $\alpha$ , mayores valores de  $\theta$  (quizás, excepto RTFIL) y un valor mayor de RDFIL que una evaluación 'buena'.

Las correlaciones entre las variables electroencefalográficas y  $u_2$  no son importantes, con excepción de RDT1, correlacionada positivamente (en la gráfica (A) se encuentra hacia arriba, positiva con respecto al eje de las ordenadas). Esto nos hace pensar que RDT1 puede estar asociada a un valor mayor de  $u_2$ .

Mientras que, con respecto a  $y_2$  se observan correlaciones importantes entre  $y_2$  y L2, que



es negativa y entre  $\nu_2$  y  $L_3$ , que es positiva (en la gráfica  $L_2$  se encuentra por debajo y  $L_3$  por arriba del cero, que con respecto al eje de las ordenadas dan un sentido negativo y positivo respectivamente.). Esto sugiere que  $L_2$  puede estar asociada a disminuciones en  $\nu_2$  y  $L_3$  a aumentos en  $\nu_2$ .

Ambos resultados nos llevan a creer que una evaluación 'regular' en lectoescritura pudiera estar relacionada con disminución de los valores de RDTI con respecto a una evaluación 'mala' y viceversa. Este resultado debe tomarse con más reservas que el del análisis de la primera pareja de variables canónicas pues la correlación canónica de esta segunda pareja de variables es bastante menor.

Si unimos ambas interpretaciones, estas nos sugieren que una evaluación 'buena' aparentemente está asociada a potencias relativas mayores de la banda  $\alpha$  (en particular RACI, RAFI, RAFD, RAFIL y RAFDL), menores en la banda  $\theta$  (a excepción quizás de RTFIL) y menor potencia relativa de RDFIL, con respecto a una evaluación 'regular' o 'mala'. También nos sugieren que una evaluación 'mala' pudiera estar relacionada con mayores valores de RDTI que una evaluación 'regular'. Sin embargo no es claro si la evaluación 'regular' o la evaluación 'mala' están

asociadas cada una de ellas con aumentos o disminuciones de  $\alpha$ ,  $\theta$  ó RDFIL con respecto a la otra.

Haciendo un análisis comparativo de ambas "corridas" nos inclinamos a pensar que los resultados obtenidos son prácticamente equivalentes. En el A.C.C de la región anterior no queda claro si los aumentos de  $\theta$  ó las disminuciones de  $\alpha$  ó el aumento de RDFIL, respecto a una evaluación 'buena', pudieran estar asociados a una evaluación 'regular' o 'mala'; sin embargo, en el A.C.C. del hemisferio izquierdo aparentemente la evaluación 'regular' estaba asociada a aumentos de  $\theta$  y disminuciones de  $\alpha$  respecto a la evaluación 'buena', y la evaluación 'mala' estaba asociada a un aumento de RDFIL respecto a la evaluación 'regular'.

Por lo tanto, consideramos que al parecer, una evaluación "más alta" (de deficiencias ligeras a sin deficiencias) está relacionada con un aumento en los valores de las variables electroencefalográficas de la banda  $\alpha$  (en particular RAC1, RAF1 y RAFIL que se repiten en ambos análisis, y RAO1, RAFO y RAFDL) y una disminución en los valores de las variables de la banda  $\theta$  (quizás excepto RTFIL pues en ambos análisis la correlación entre RTFIL y  $\theta$ , no es importante); mientras que una evaluación de deficiencias que ameritan atención especial del maestro y los padres

aparece relacionada aparentemente con un aumento en el valor de las variables RDTI y RDFIL con respecto a los valores de estas variables cuando la evaluación es 'regular'.

Los resultados obtenidos en los análisis de la variable 'evaluación de la lectoescritura' coinciden con lo descrito en la literatura sobre el estudio de niños con dificultades en la lectoescritura.

Algunos estudios neuropsicológicos del desarrollo (Rutter, M., 1983) han demostrado que en los niños con desarrollo normal los hemisferios cerebrales se van especializando diferenciándose cada vez más: las regiones anatómicas que se relacionan con funciones del lenguaje y todos los procesos de análisis de símbolos se localizan en el hemisferio izquierdo, mientras que las regiones anatómicas que se relacionan con funciones visoespaciales (construcción de rompecabezas, solución de laberintos, etc.) se localizan en el hemisferio derecho. En particular (Luria, A.R., 1982) las regiones frontales izquierdas se relacionan con los aspectos motores del lenguaje hablado o escrito, es decir, con el hablar o el escribir; y en el área temporal izquierda se realiza el análisis del lenguaje, o sea, la comprensión de lo que se habla, lo que se escucha, lo que se lee o lo que se escribe.

Por estos motivos, en los A.C.C. que efectuamos llama la atención que las correlaciones canónicas entre los grupos de variables electroencefalográficas y la evaluación de la lectoescritura, hayan sido mucho más altas en la región anterior que en la posterior (recordemos que las zonas frontales y temporales se encuentran en la región anterior) y en el hemisferio izquierdo que en el derecho.

Por otra parte Sklar, Hanley y Simmons (1973) describieron valores mayores de las potencias relativas de la banda  $\theta$  en niños disléxicos que en niños normales; Rebert, Wexler y Sproul (1973) encontraron valores mayores de las potencias relativas de la banda  $\theta$  en el hemisferio izquierdo en niños disléxicos que en niños normales y John y cols. (1983) encontraron que una gran muestra de niños con trastorno de aprendizaje se caracterizaba por tener valores menores de potencia relativa en la banda  $\alpha$  con respecto a los niños sin trastorno de aprendizaje.

Un incremento de la actividad  $\theta$  ó una disminución de la actividad  $\alpha$ , o ambas cosas simultáneamente son considerados signos de inmadurez o de lesión cerebral (Velasco, M. y cols, 1973), por lo tanto, el hallazgo de estos signos en la región anterior y en el hemisferio izquierdo se pue-

de interpretar como una alteración funcional de estas regiones, debida posiblemente a inmadurez, que explicaría las alteraciones conductuales observadas en los niños con dificultades en la lecto-escritura.

Un incremento en los valores de la banda  $\delta$  en el E.E.G es considerado como un signo, no de inmadurez, sino de lesión cerebral (Velasco, M. y cols., 1973). Esto ha sido descrito para pacientes con afasia producida por infartos o tumores cerebrales, en los cuales están presentes síntomas neurológicos bien definidos como la dificultad en la comprensión del lenguaje. El hallazgo de que un incremento de los valores de RDTI ó RDFIL pudiera estar asociado a una evaluación 'mala' de la lectoescritura puede interpretarse como una lesión en estas regiones que implicaría serias dificultades en el análisis del lenguaje hablado o escrito. En nuestra muestra, los niños seleccionados no presentaron evidencias clínicas que reflejaran algún trastorno neurológico; además, cualquier resultado en este sentido dijimos que lo tomaríamos con reservas; no obstante queda la duda de si el signo electroencefalográfico RDTI ó RDFIL alta corresponde exclusivamente a una lesión o si puede presentarse en una alteración grave de la lectoescritura aso-

ciada a inmadurez sin presencia de lesión. Actualmente a los niños de la muestra se les están haciendo estudios de tomografía computarizada de cabeza (Fernández-Bouzas, A.; comunicación personal) que podrían ayudar a aclarar esta duda.

## APENDICE

En este apéndice presentamos algunas de las definiciones que utilizamos en el Capítulo II de este trabajo y los teoremas más importantes que usamos en el mismo capítulo. Todos los teoremas aparecen demostrados en este apéndice; en la recopilación y en la demostración de algunos de ellos recurrimos a Bartle, 1982; Friedberg, 1982; Mardia, 1979 y Marsden, 1976. A continuación damos un índice de las definiciones:

- 1) orden de una matriz.
- 2) matriz simétrica.
- 3) conjunto linealmente independiente.
- 4) rango de una matriz.
- 5) matriz no singular o invertible.
- 6) matriz ortogonal.
- 7) producto interior.
- 8) norma de un vector.
- 9) vector unitario, vectores ortogonales, subconjunto ortogonal, subconjunto ortonormal.
- 10) complemento ortogonal.
- 11) matriz positiva definida y matriz positiva semidefinida.
- 12) eigenvector y eigenvalor
- 13) forma bilineal.
- 14) función inyectiva.
- 15) conjunto cerrado
- 16) conjunto compacto.
- 17) matriz idempotente.



Definición 1: Si una matriz rectangular  $A$  tiene  $n$  filas y  $p$  columnas, decimos que  $A$  es de orden  $n \times p$ .

Definición 2:  $A$  es una matriz simétrica si  $A' = A$ .

Definición 3: Un subconjunto  $S = \{x_1, \dots, x_n\}$  de un espacio vectorial  $V$  es linealmente independiente si  $a_1 x_1 + \dots + a_n x_n = 0$  implica  $a_1 = \dots = a_n = 0$ .

Definición 4: El rango de una matriz  $A$  ( $n \times p$ ) se define como el máximo número de filas (o columnas) linealmente independientes en  $A$ .

Definición 5: La matriz cuadrada  $A$  es no singular o invertible si existe una matriz  $B$  (única) que satisface

$$AB = BA = I$$

La matriz  $B$ , de igual orden que  $A$ , se conoce como la inversa de  $A$  y se denota por  $A^{-1}$ .

Definición 6: Una matriz cuadrada  $A$  es ortogonal si  $AA' = I$

Definición 7: Sea  $V$  un espacio vectorial sobre  $\mathbb{R}$ . Un producto interior en  $V$  es una función que asigna a cada par de vectores  $x, y$  en  $V$  un escalar en  $\mathbb{R}$ , representado como  $\langle x, y \rangle$ , tal que para todo  $x, y, z$  en  $V$  y toda  $c$  en  $\mathbb{R}$  se tiene que:

$$\langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$$

$$\langle cx, y \rangle = c \langle x, y \rangle$$

$$\langle x, y \rangle = \langle y, x \rangle$$

$$\langle x, x \rangle > 0 \text{ si } x \neq 0$$

Definición 8: Sea  $V$  un espacio con producto interior. Para  $x \in V$  definimos la norma (o longitud) de  $x$  mediante

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Definición 9: Sea  $V$  un espacio con producto interior. Un vector  $x$  en  $V$  es un vector unitario si  $\|x\| = 1$ . Los vectores  $x$  y  $z$  son ortogonales si  $\langle x, z \rangle = 0$ . Un subconjunto  $S$  de  $V$  es ortogonal si cualquier par de elementos distintos de  $S$  son ortogonales. Finalmente, un subconjunto  $S$  de  $V$  es ortonormal si  $S$  es ortogonal y está formado únicamente de vectores unitarios.

Definición 10: Sea  $V$  un espacio con producto interior y sea  $S$  un subconjunto de  $V$ . Definimos a  $S^\perp$  como el conjunto de todos aquellos vectores de  $V$  que son ortogonales a todos los vectores de  $S$ ; esto es,  $S^\perp = \{x \in V \cdot \langle x, y \rangle = 0 \ \forall y \in S\}$ . A  $S^\perp$  se le llama complemento ortogonal de  $S$ .

Definición 11: Una matriz simétrica  $A$  es llamada positiva definida (semidefinida), si la forma cuadrática  $Q(x) = x'Ax$  es positiva definida (semidefinida), es decir,  $Q(x) > 0$  ( $Q(x) \geq 0$ ) para toda  $x$  distinta de cero.

NOTA:

Una caracterización de una matriz positiva definida (semidefinida) es que todos sus eigenvalores sean positivos (no negativos).

Definición 12: Si  $A$  es una matriz de orden  $n \times n$ , un elemento no nulo  $x$  se denomina eigenvector de  $A$ , si existe un escalar  $\lambda$  tal que  $Ax = \lambda x$ . Al escalar  $\lambda$  se le llama eigenvalor correspondiente al eigenvector  $x$ .

NOTA:

Los eigenvalores son los valores de  $\lambda$  que satisfacen la ecuación  $|A - \lambda I| = 0$ . El eigenvector asociado a determinada  $\lambda$  es el vector  $x \neq 0$  tal que  $Ax - \lambda I = 0$

Definición 13: Sea  $V$  un espacio vectorial en un campo  $F$ . Una función  $H$  que va del conjunto  $V \times V$  de pares ordenados de vectores en  $V$  a  $F$ , se llama forma bilineal en  $V$  si  $H$  es lineal en cada variable cuando la otra variable se mantiene fija, esto es, si

$$H(ax_1 + x_2, y) = aH(x_1, y) + H(x_2, y) \text{ para toda } x_1, x_2, y \in V \text{ y } a \in F.$$

$$H(x, ay_1 + y_2) = aH(x, y_1) + H(x, y_2) \text{ para toda } x, y_1, y_2 \in V \text{ y } a \in F.$$

Definición 14: Sea  $f$  una función con dominio  $\mathcal{D}(f)$  en  $A$ . Se dice que  $f$  es inyectiva si  $f(a) = f(a')$  implica  $a = a' \forall a, a' \in A$ .

NOTA:

Si  $f$  es lineal  $f(a) = f(a') \Leftrightarrow f(a - a') = 0$   
entonces la condición de inyectividad puede escribirse como

$$f(a) = 0 \text{ implica } a = 0 \forall a \in A.$$

Definición 15: Se dice que un conjunto  $F$  en  $\mathbb{R}^p$  es cerrado en  $\mathbb{R}^p$  cuando su complemento  $\mathcal{B}(F) = \mathbb{R}^p \setminus F$  es abierto en  $\mathbb{R}^p$ , es decir, si para cada punto  $x \in \mathcal{B}(F)$  existe un número real  $r > 0$  tal que todo punto  $y$  en  $\mathbb{R}^p$  que satisface  $\|x - y\| < r$  también pertenece al conjunto  $\mathcal{B}(F)$ .

Definición 16: Se dice que un conjunto  $K$  es compacto si siempre que está contenido en la unión de una colección  $\mathcal{G} = \{G_\alpha\}$  de conjuntos abiertos también está contenido en la unión de algún número finito de conjuntos en  $\mathcal{G}$ .

Definición 17: Se dice que una matriz cuadrada  $A$  es idempotente si  $AA = A$ .

NOTA 1:  $f(\underline{a}, \underline{b}) = \underline{a}' R_{12} \underline{b}$  es una forma bilineal.

Demostración:

Supongamos  $p < q$ , sin pérdida de generalidad.

Sean:

$$\underline{\hat{a}}' = (a_1, \dots, a_p, 0, \dots, 0)_{1 \times q}$$

$$\underline{b}' = (b_1, \dots, b_p, b_{p+1}, \dots, b_q)_{1 \times q}$$

$$\hat{R}_{12} = \begin{array}{|cccc|} \hline S_{1,p+1} & \dots & S_{1,p+q} & \\ \vdots & & & \\ S_{p,p+1} & \dots & S_{p,p+q} & \\ \hline 0 & \dots & 0 & \\ \vdots & & \vdots & \\ 0 & \dots & 0 & \\ \hline \end{array} \quad q \times q$$

Nótese que  $\underline{a}' R_{12} \underline{b} = \underline{\hat{a}}' \hat{R}_{12} \underline{b}$

Definimos  $\hat{f}: \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}$  .j.

$$\hat{f}(\underline{\hat{a}}, \underline{b}) = \underline{\hat{a}}' \hat{R}_{12} \underline{b}$$

$$\text{p.d.} \begin{cases} \hat{f}(\alpha \underline{\hat{a}}_1 + \underline{\hat{a}}_2, \underline{b}) = \alpha \hat{f}(\underline{\hat{a}}_1, \underline{b}) + \hat{f}(\underline{\hat{a}}_2, \underline{b}) \quad \forall \underline{\hat{a}}_1, \underline{\hat{a}}_2 \in \mathbb{R}^p \\ \hat{f}(\underline{\hat{a}}, \beta \underline{b}_1 + \underline{b}_2) = \beta \hat{f}(\underline{\hat{a}}, \underline{b}_1) + \hat{f}(\underline{\hat{a}}, \underline{b}_2) \quad \forall \underline{b}_1, \underline{b}_2 \in \mathbb{R}^q \end{cases}$$

Sean  $\underline{\hat{a}}_1, \underline{\hat{a}}_2 \in \mathbb{R}^p$  cualesquiera

$$\begin{aligned} \hat{f}(\alpha \underline{\hat{a}}_1 + \underline{\hat{a}}_2, \underline{b}) &= (\alpha \underline{\hat{a}}_1 + \underline{\hat{a}}_2)' \hat{R}_{12} \underline{b} \\ &= \alpha \underline{\hat{a}}_1' \hat{R}_{12} \underline{b} + \underline{\hat{a}}_2' \hat{R}_{12} \underline{b} \\ &= \alpha \hat{f}(\underline{\hat{a}}_1, \underline{b}) + \hat{f}(\underline{\hat{a}}_2, \underline{b}) \end{aligned}$$

$$\begin{aligned}
 \hat{f}(\hat{\alpha}, \beta \hat{b}_1 + \hat{b}_2) &= \hat{\alpha}' \hat{R}_{12} (\beta \hat{b}_1 + \hat{b}_2) \\
 &= \beta \hat{\alpha}' \hat{R}_{12} \hat{b}_1 + \hat{\alpha}' \hat{R}_{12} \hat{b}_2 \\
 &= \beta \hat{f}(\hat{\alpha}, \hat{b}_1) + \hat{f}(\hat{\alpha}, \hat{b}_2)
 \end{aligned}$$

L.O.Q.D.

Teorema A.1. : Toda función bilineal es continua.

Para demostrar este teorema necesitamos el siguiente resultado:

Teorema A.1.1: Sea  $f(x, y)$  de  $\mathbb{R}^n \times \mathbb{R}^n$  en  $\mathbb{R}$  una forma bilineal. Entonces existe  $M > 0$  tal que  $\|f(x, y)\| \leq M \|x\| \|y\|$  para toda  $x, y \in \mathbb{R}^n$ .

Demostración:

Sea  $f(x, y) = x' R y$  y sea  $z = R y$ , entonces:

$$\|f(x, y)\| = |x' R y| = |x' z| \leq \|x\| \|z\| \quad \forall x, z \in \mathbb{R}^n$$

pero  $\|z\| = \|R y\|$

Por definición  $\|R\| = \max_{y \neq 0} \frac{\|R y\|}{\|y\|}$

Entonces:

$$\|R\| \geq \frac{\|R y\|}{\|y\|} \quad \forall y \in \mathbb{R}^n, y \neq 0.$$

$$\therefore \|z\| \leq \|R\| \|y\| \quad \forall y \in \mathbb{R}^n, y \neq 0$$

es decir, existe  $M > 0$  tal que

$$\|z\| \leq M \|y\| \quad \forall y \in \mathbb{R}^n, y \neq 0$$

Por lo tanto existe  $M > 0$  tal que

$$\|f(x, y)\| \leq \|x\| \|z\| \leq M \|x\| \|y\| \quad \text{L.O.Q.D.}$$

### Demostración:

Sea  $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  una función bilineal

Sea  $\epsilon > 0$ ,  $\epsilon \in \mathbb{R}$ . y sea  $(a, b) \in \mathbb{R}^n \times \mathbb{R}^n$ .

Debemos probar que existe  $d(\epsilon) > 0$  tal que

si  $\|(x, y) - (a, b)\| < d(\epsilon)$ , entonces

$$\|f(x, y) - f(a, b)\| < \epsilon$$

Consideremos dos casos:

1<sup>er</sup> caso:  $a \neq 0$  y  $b \neq 0$

Nótese que

$$\begin{aligned}\|f(x, y) - f(a, b)\| &= \|f(x, y) - f(x, b) + f(x, b) - f(a, b)\| \\ &= \|f(x, y - b) + f(x - a, b)\| \\ &= \|f(x, y - b) - f(a, y - b) + f(x - a, b) + f(a, y - b)\| \\ &= \|f(x - a, y - b) + f(x - a, b) + f(a, y - b)\| \\ &\leq \|f(x - a, y - b)\| + \|f(x - a, b)\| + \|f(a, y - b)\|\end{aligned}$$

Por lo tanto, por el resultado A.1.1. tenemos que:

$$\|f(x, y) - f(a, b)\| \leq M \|x - a\| \|y - b\| + M \|x - a\| \|b\| + M \|a\| \|y - b\|$$

para alguna  $M > 0$

$$\text{Sea } d(\epsilon) = \min \left\{ \sqrt{\frac{\epsilon}{3M}}, \frac{\epsilon}{3M \|b\|}, \frac{\epsilon}{3M \|a\|} \right\} \rightarrow \|(x, y) - (a, b)\| < d(\epsilon)$$

$$\text{Como } \|x - a\| \leq \|(x - a, y - b)\| = \|(x, y) - (a, b)\| < d(\epsilon)$$

$$\text{y } \|y - b\| \leq \|(x, y) - (a, b)\| < d(\epsilon)$$





Teorema A.2: Teorema del valor máximo y mínimo  
 Sea  $K \subset \text{Dom}(f)$  compacto en  $\mathbb{R}^N$  y sea  $f$  función continua de valor real. Entonces existen puntos  $x^*$  y  $x_*$  en  $K$  tales que:

$$f(x^*) = \sup \{ f(x) \mid x \in K \}$$

$$f(x_*) = \inf \{ f(x) \mid x \in K \}$$

Demostración: (Bartle, p.p. 180)

En la demostración utilizaremos, sin demostrar, el siguiente

Teorema (CONSERVACIÓN DE COMPACIDAD): Si  $K \subseteq \text{Dom}(f)$  es compacto y  $f$  es continua en  $K$ , entonces  $f(K)$  es compacto.

Dado que  $K$  es compacto en  $\mathbb{R}^P$ , del teorema anterior se deduce que  $f(K)$  está acotada en  $\mathbb{R}$ .

Sean  $M = \sup f(K)$  y  $(x_n)$  una sucesión en  $K$  tal que  $f(x_n) \geq M - \frac{1}{n}$ ,  $n \in \mathbb{N}$

Por el teorema de Bolzano-Weierstrass, alguna subsucesión  $(x_{n(k)})$  converge a un límite  $x^* \in K$ .

Dado que  $f$  es continua en  $x^*$ , se debe tener

$$f(x^*) = \lim (f(x_{n(k)})) = M.$$

La demostración de la existencia de  $x_*$  es análoga. L.Q.D.P.

Corolario A.2.1: Sea  $f$  una función de  $\text{Dom}(f) \subseteq \mathbb{R}^P$  a  $\mathbb{R}^q$  y sea  $K \subset \text{Dom}(f)$  compacto. Si  $f$  es continua en  $K$ , existen puntos  $x^*$  y  $x_*$  en  $K$  tales

que

$$\|f(x^*)\| = \sup \{ \|f(x)\| \mid x \in K \}$$

$$\|f(x_*)\| = \inf \{ \|f(x)\| \mid x \in K \}$$

Teorema A.3: Sea  $f: \mathbb{R}^p \rightarrow \mathbb{R}^n$  una función lineal, entonces  $f$  es inyectiva si y sólo si existe  $m > 0$  tal que

$$\|f(x)\| \geq m \|x\| \text{ para toda } x \in \mathbb{R}^p.$$

Demostración: (Bartle, p.p. 181)

( $\Rightarrow$ ) Supongamos que  $f$  es inyectiva y sea  $S = \{x \in \mathbb{R}^p \mid \|x\| = 1\}$ . Por el corolario A.2.1, existe  $x_0 \in S$  tal que  $\|f(x_0)\| = m = \inf \{ \|f(x)\| \mid x \in S \}$ . Dado que  $f$  es inyectiva,  $m = \|f(x_0)\| > 0$ .  
Por lo tanto

$$\|f(x)\| \geq m > 0 \text{ para toda } x \in S.$$

Ahora, si  $u \in \mathbb{R}^p$ ,  $u \neq 0$ , entonces  $\frac{u}{\|u\|} \in S$  y por la linealidad de  $f$  se tiene:

$$\frac{1}{\|u\|} \|f(u)\| = \left\| f\left(\frac{u}{\|u\|}\right) \right\| \geq m,$$

por lo que se deduce que  $\|f(u)\| \geq m \|u\| \forall u \in \mathbb{R}^p$  (ya que el resultado es trivial para  $u=0$ ).

( $\Leftarrow$ ) Supongamos  $\|f(x)\| \geq m \|x\|$  para toda  $x \in \mathbb{R}^p$ . Si  $f(x_1) = f(x_2)$  se tiene:

$$0 = \|f(x_1) - f(x_2)\| = \|f(x_1 - x_2)\| \geq m \|x_1 - x_2\|,$$

que implica  $x_1 = x_2$

Por lo tanto,  $f$  es inyectiva. L.Q.Q.D.

Teorema A.4: Sean  $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  y  $g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$  funciones suaves dadas. Sea  $x_0 \in U$  tal que  $g(x_0) = c_0$  y sea  $S$  el conjunto nivel de  $g$  con valor  $c_0$  (recordar que este es el conjunto de puntos  $x \in \mathbb{R}^n$  con  $g(x) = c_0$ ). Supongamos  $\nabla g(x_0) \neq 0$ . Si  $f|_S$ , que denota  $f$  restringida a  $S$ , tiene un máximo o un mínimo en  $x_0$ , entonces existe  $\lambda \in \mathbb{R}$  tal que:

$$\nabla f(x_0) = \lambda \nabla g(x_0) \quad (*)$$

Demostración: (Marsden, p.p. 191)

Recordemos que el espacio tangente a  $S$  en  $x_0$  se define como el espacio ortogonal a  $\nabla g(x_0)$ . Esta definición considera que si  $c(t)$  es una trayectoria en  $S$  y  $c(0) = x_0$ , entonces  $c'(0)$  es un vector tangente a  $S$  en  $x_0$ ; pero

$$\frac{d}{dt} g(c(t)) = \frac{d}{dt} c_0 = 0$$

y además por la regla de la cadena,

$$\left. \frac{d}{dt} g(c(t)) \right|_{t=0} = \nabla g(x_0) \cdot c'(0)$$

entonces

$\nabla g(x_0) \cdot c'(0) = 0$ ; es decir,  $c'(0)$  es ortogonal a  $\nabla g(x_0)$ .

Si  $f|_S$  tiene un máximo en  $x_0$ , entonces  $f(c(t))$  tendrá, con certeza un máximo en  $t=0$ . Por cálculo de una variable

---

(\*) el símbolo  $\nabla$  denota al gradiente.

$$\left. \frac{d}{dt} f(c(t)) \right|_{t=0} = 0$$

De aquí, por la regla de la cadena

$$0 = \left. \frac{d}{dt} f(c(t)) \right|_{t=0} = \nabla f(x_0) \cdot c'(0)$$

Así,  $\nabla f(x_0)$  es perpendicular a la tangente de cada curva en  $S$  y también es perpendicular al espacio tangente de  $S$  en  $x_0$ . De aquí, como el espacio perpendicular a este espacio tangente es unidimensional,  $\nabla f(x_0)$  y  $\nabla g(x_0)$  son paralelos. Como  $\nabla g(x_0) \neq 0$ , se sigue que  $\nabla f(x_0)$  es un múltiplo de  $\nabla g(x_0)$ ; esto es:

Existe  $\lambda \in \mathbb{R}$  tal que

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

L.Q.Q.D.

#### GENERALIZACION DEL TEOREMA A.4:

Si una superficie  $S$  está definida por  $k$  restricciones

$$g_1(x_1, \dots, x_n) = c_1$$

$$\vdots$$

$$g_k(x_1, \dots, x_n) = c_k$$

y si  $f$  tiene un máximo o un mínimo en  $x_0 \in S$ , existen constantes  $\lambda_1, \dots, \lambda_k$  tales que:

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \dots + \lambda_k \nabla g_k(x_0)$$

Teorema A.5: Si  $A$  es una matriz  $n \times m$  de rango  $m < n$ , entonces  $A'A$  es positiva definida.

Demostración:

Por definición  $A'A$  es positiva definida si y sólo si  $Q(\underline{x}) = \underline{x}'A'A\underline{x} > 0$  para toda  $\underline{x} \neq 0$

Pero

$$Q(\underline{x}) = \underline{x}'A'A\underline{x} = \|A\underline{x}\|^2 \geq 0$$

Supongamos  $Q(\underline{x}) = 0$ , entonces

$$\|A\underline{x}\| = 0 \Rightarrow A\underline{x} = 0$$

Sean  $N(A) = \{ \underline{x} \in \mathbb{R}^m \mid A\underline{x} = 0 \}$

y  $M(A)$  el subespacio vectorial generado por las columnas de  $A$  (nótese que  $\text{rango}(A) = \dim M(A)$ )

Como

$$\dim N(A) + \dim M(A) = m$$

entonces

$$\dim N(A) = m - \dim M(A) = m - \text{rango}(A) = 0$$

Por lo tanto

$$A\underline{x} = 0 \Rightarrow \underline{x} = 0 \quad ; \text{CONTRADICCIÓN!}$$

$\underline{x} \neq 0$  por hipótesis

$$\therefore Q(\underline{x}) = \underline{x}'A'A\underline{x} > 0$$

$\therefore A'A$  es positiva definida.

L.Q.Q.D.

NOTA 2:

Teorema: Sea  $A$  una matriz real de  $p \times p$ . Entonces  $A$  es simétrica si y sólo si  $A$  es ortogonalmente equivalente a una matriz real diagonal.

Demostración:

( $\Rightarrow$ ) Para una matriz real simétrica existe  $\beta$ , base ortonormal para  $F^p$  formada por eigenvectores de  $A$ . Por lo tanto  $A$  es similar a una matriz diagonal  $D$ . Utilizando otro resultado que no demostraremos, la matriz  $C$ , cuyas columnas son los vectores de  $\beta$  es tal que

$$D = CAC^{-1}$$

Pero como las columnas de  $C$  son una base ortonormal para  $F^p$ , se infiere que  $C$  es ortogonal  $\therefore A$  es ortogonalmente equivalente a  $D$ .

( $\Leftarrow$ ) Sea  $CAC^{-1}$  una matriz diagonal, entonces:

$$CAC^{-1} = (CAC^{-1})' = CA'C^{-1}$$

pero  $C$  ortogonal, por lo tanto  $C^{-1} = C'$ ; así:

$$CAC^{-1} = CA'C^{-1}$$

de donde

$$A = A', \text{ es decir, } A \text{ es simétrica.}$$

*l.a.q.d.*

Esto significa que si  $A (p \times p)$  es simétrica, existe  $C$ , matriz ortogonal, tal que:

$$D = CAC^{-1} \text{ es una matriz diagonal.}$$

$D$  puede escribirse como

$$D = \text{diag} (\rho_1, \dots, \rho_p)$$

Si además  $A$  es positiva definida ( $p_i > 0 \forall i=1, \dots, p$ ) podemos definir

$$D^{1/2} = \text{diag}(\sqrt{p_1}, \dots, \sqrt{p_p})$$

$$\text{y } A^{1/2} = CD^{1/2}C'$$

Sea  $P = CD^{1/2}C'$ . Entonces:

$$PP = (CD^{1/2}C')(CD^{1/2}C') = (CD^{1/2})(C^{-1}C)(D^{1/2}C') = CDC' = A$$

$C' = C^{-1}$  porque  $C$  es ortogonal.

Teorema A.6: Para dos matrices  $A (n \times p)$  y  $B (p \times n)$ , los eigenvalores de  $AB$  y  $BA$  son iguales.

Demostración:

Utilizaremos el siguiente resultado:

$$\begin{aligned} \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} &= |A_{11}| |A_{22} - A_{21}A_{11}^{-1}A_{12}| \\ &= |A_{22}| |A_{11} - A_{12}A_{22}^{-1}A_{21}| \dots \dots \dots (*) \end{aligned}$$

Notemos que las siguientes matrices son semejantes: ( $\lambda \neq 0$ )

$$\begin{bmatrix} -I_n & -A \\ B & \lambda I_p \end{bmatrix} \sim \begin{bmatrix} -\lambda I_n & -\lambda A \\ B & \lambda I_p \end{bmatrix} \sim \begin{bmatrix} -\lambda I_n & -A \\ B & I_p \end{bmatrix}$$

multiplicando los  $n$  primeros renglones por  $\lambda$ 
multiplicando las últimas  $p$  columnas por  $\frac{1}{\lambda}$

Por lo tanto

$$\begin{vmatrix} -I_n & -A \\ B & \lambda I_p \end{vmatrix} = \begin{vmatrix} -\lambda I_n & -A \\ B & I_p \end{vmatrix}$$



pero por resultado (i) tenemos que

$$\begin{vmatrix} -I_n & A \\ B & \lambda I_p \end{vmatrix} = |-I_n| |\lambda I_p + B(-I_n)^{-1}A| \\ = |BA - \lambda I_p|$$

$$\text{y } \begin{vmatrix} -\lambda I_n & -A \\ B & I_p \end{vmatrix} = |I_p| |-\lambda I_n + A(I_p)^{-1}B| \\ = |AB - \lambda I_n|$$

Por lo tanto

$$|BA - \lambda I_p| = |AB - \lambda I_n|$$

Por definición  $\lambda \neq 0$  es eigenvalor de  $BA$  si y sólo si  
 $|BA - \lambda I_p| = 0$ , pero esto ocurre si y sólo si  
 $|AB - \lambda I_n| = 0$  si y sólo si  $\lambda \neq 0$  es eigenvalor de  $AB$ .

L.Q.Q.D.

Teorema A.7: Los eigenvalores de  $A$  son idénticos a los eigenvalores de  $CAC^{-1}$ ,  $C$  invertible.

Demostración:

Sea  $P = CAC^{-1} - \lambda I$

$\lambda$  es eigenvalor de  $CAC^{-1}$  si y sólo si  $|P| = 0$

Como  $C$  es invertible,  $C \neq 0$ . Por lo tanto tenemos que

$$|P| = 0 \text{ si y sólo si } |C^{-1}||P||C| = 0$$

pero  $|C^{-1}||P||C| = |C^{-1}PC| = |A - \lambda I| = 0$  si y sólo si  $\lambda$  es eigenvalor de  $A$ .

L.Q.Q.D.

Teorema A.8: Para cualquier matriz  $A$

$$\text{rango}(A) = \text{rango}(A'A) = \text{rango}(AA')$$

Demostración:

Definimos  $N(A) = \{ \underline{x} \in \mathbb{R}^p \mid A\underline{x} = 0 \}$  el espacio nulo de  $A$ , y

$M(A)$  el subespacio generado por las columnas de  $A$ .

Demostremos: i)  $\dim M(A) = \dim M(A'A)$

ii)  $\dim M(A) = \dim M(AA')$

i) Sea  $\underline{x} \in N(A)$ , entonces  $A\underline{x} = 0$

$$\text{pero } A\underline{x} = 0 \Rightarrow A'A\underline{x} = 0 \Rightarrow \underline{x} \in N(A'A)$$

Sea  $\underline{x} \in N(A'A)$ , entonces  $A'A\underline{x} = 0$

$$\text{pero } A'A\underline{x} = 0 \Rightarrow \underline{x}'A'A\underline{x} = \|A\underline{x}\|^2 = 0 \Leftrightarrow A\underline{x} = 0 \\ \Rightarrow \underline{x} \in N(A)$$

Por lo tanto

$$\underline{x} \in N(A'A) \Leftrightarrow \underline{x} \in N(A)$$

es decir:

$$N(A) = N(A'A)$$

Nótese que el número de columnas de  $A'A$  es  $p$ , igual al número de columnas de  $A$ .

Entonces, como  $\dim N(A) + \dim M(A) = p$

$$\text{y } \dim N(A'A) + \dim M(A'A) = p$$

resulta:

$$\dim M(A) = \dim M(A'A)$$

ii) Primero nótese que la dimensión del subespacio generado por las filas o columnas de  $A$  es igual a la dimensión del subespacio generado por las filas o columnas de  $A'$ , es decir:

$\dim M(A) = \dim M(A')$ , o en otras palabras:

$$\text{rango}(A) = \text{rango}(A')$$

Por lo tanto es equivalente demostrar ii) que demostrar:

$$\dim M(A') = \dim N(AA')$$

El procedimiento es análogo al de i), aunque nótese que el número de columnas de  $A'$  y  $AA'$  es  $n$ . En este caso

$$N(A') = \{x \in \mathbb{R}^n : A'x = 0\}$$

$$N(AA') = \{x \in \mathbb{R}^n : AA'x = 0\}$$

y el resultado lo escribimos como:

$$\dim M(A) + \dim N(A') = n$$

l.q.q.d.

Teorema A.9: Sea  $A$  una matriz de orden  $n \times p$ . Si  $B$  ( $n \times n$ ) y  $C$  ( $p \times p$ ) son matrices no singulares, entonces

$$\text{rango}(BAC) = \text{rango}(A)$$

### Demstración:

Definimos  $N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$

$$N(BA) = \{x \in \mathbb{R}^n \mid BAx = 0\}$$

$$N(BA)' = \{x \in \mathbb{R}^n \mid (BA)'x = 0\}$$

$$N(BAC)' = \{x \in \mathbb{R}^n \mid (BAC)'x = 0\}$$

Demostremos: i)  $x \in N(A)$  si y sólo si  $x \in N(BA)$

ii)  $x \in N(BA)'$  si y sólo si  $x \in N(BAC)'$

$$i) \quad x \in N(A) \Leftrightarrow Ax = 0$$

$$\Rightarrow BAx = 0 \Leftrightarrow x \in N(BA)$$

$$x \in N(BA) \Leftrightarrow BAx = 0$$

$$\Rightarrow B^{-1}BAx = Ax = 0 \Leftrightarrow x \in N(A)$$

$$ii) \quad x \in N(BA)' \Leftrightarrow A'B'x = 0$$

$$\Rightarrow C'A'B'x = 0 \Leftrightarrow x \in N(BAC)'$$

$$x \in N(BAC)' \Leftrightarrow C'A'B'x = 0$$

$$\Rightarrow C'^{-1}C'A'B'x = A'B'x = 0 \Leftrightarrow x \in N(BA)'$$

De i) tenemos  $N(A) = N(BA)$  y como  $A$  y  $BA$  tienen el mismo número de columnas, tenemos:

$$\dim M(A) = \dim M(BA)$$

es decir,

$$\text{rango}(A) = \text{rango}(BA)$$

De ii) tenemos  $N(BA)' = N(BAC)'$ , por lo tanto:

$$\text{rango}(BA)' = \text{rango}(BAC)'$$

Pero  $\text{rango}(BA)' = \text{rango}(BA)$  y  $\text{rango}(BAC)' = \text{rango}(BAC)$

Por lo tanto:

$$\text{rango}(A) = \text{rango}(BAC) \quad \text{L.Q.E.D.}$$

Teorema A.10: Teorema de la descomposición espectral:  
Una matriz simétrica  $A$  ( $p \times p$ ) puede escribirse como

$$A = \Gamma \Lambda \Gamma' = \sum \lambda_i v_{(i)} v_{(i)}'$$

donde  $\Lambda$  es una matriz diagonal de los eigenvalores  $\lambda_i$  de  $A$  y  $\Gamma$  es una matriz ortogonal cuyas columnas  $v_{(i)}$  son eigenvectores estandarizados.

Demostración: (Mardia)

Supongamos que podemos encontrar vectores ortogonales  $v_{(1)}, \dots, v_{(p)}$  tales que  $A v_{(i)} = \lambda_i v_{(i)}$  para alguna  $\lambda_i \in \mathbb{R}$ . Entonces:

$$v_{(i)}' A v_{(j)} = \lambda_j v_{(i)}' v_{(j)} = \begin{cases} \lambda_i & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases}$$

o en la forma matricial

$$\Gamma' A \Gamma = \Lambda$$

Multiplicando a la izquierda y a la derecha por  $\Gamma$  y  $\Gamma'$  respectivamente, obtenemos

$$A = \Gamma \Lambda \Gamma'$$

Por A.7,  $A$  y  $\Lambda$  tienen los mismos eigenvalores, entonces los eigenvalores de  $A$  son los elementos de la diagonal de  $\Lambda$ .

Por lo tanto, debemos encontrar una base ortonormal de eigenvectores.

Nótese que si  $\lambda_i \neq \lambda_j$  son eigenvalores distintos con eigenvectores  $x$ ,  $y$  respectivamente, entonces

$$\lambda_i x'y = x' Ay = y' Ax = \lambda_j y'x$$

por lo tanto  $y'x = 0$ . Así, para una matriz simétrica, los eigenvectores correspondientes a distintos eigenvalores son ortogonales uno a otro

Supongamos que hay  $k$  eigenvalores distintos de  $A$  con sus correspondientes eigenespacios  $H_1, \dots, H_k$  de dimensiones  $r_1, \dots, r_k$ . Sea

$$r = \sum_{j=1}^k r_j$$

Como eigenespacios diferentes son ortogonales, existe un conjunto ortonormal de vectores  $e_1, \dots, e_r$  tales que los vectores denotados

$$\sum_{i=1}^{r_1} e_i, \dots, \sum_{i=1}^{r_k} e_i$$

forman una base para  $H_j$ .

Tomando en cuenta el siguiente

Teorema: Sea  $\lambda_i$  un eigenvalor de  $A$  ( $n \times n$ ), con eigenespacio  $H$  de dimensión  $r$ . Si  $k$  denota la multiplicidad de  $\lambda_i$ , entonces  $1 \leq r \leq k$ .

$r_j$  es menor o igual que la multiplicidad de su eigenvalor correspondiente. Entonces, reordenando los eigenvalores  $\lambda_i$  si es necesario, supongamos

$$A e_i = \lambda_i e_i \quad i=1, \dots, r$$

y  $r \leq p$ . Si todos los eigenvalores son distintos,  $r=p$  y en tal caso hacemos  $x_{(1)} = e_1$ . Mostramos a continuación que la situación  $r < p$  lleva a una contradicción.

Sin pérdida de generalidad supongamos que todos los eigenvalores de  $A$  son positivos (si no fuera así podemos sustituir  $A$  por  $A + \alpha I$  para alguna  $\alpha$ , pues  $A$  y  $A + \alpha I$  tienen los mismos eigenvalores). Sea

$$B = A - \sum_{i=1}^r \lambda_i e_i e_i'$$

Entonces

$$\text{traza } B = \text{traza } A - \sum_{i=1}^r \lambda_i (e_i' e_i) = \sum_{i=r+1}^p \lambda_i > 0$$

ya que  $r < p$ . Entonces  $B$  tiene al menos un eigenvalor distinto de 0, llámese  $\theta$ . Sea  $x \neq 0$  un eigenvector correspondiente. Así, para  $1 \leq j \leq r$ ,

$$\theta e_j' x = e_j' Bx = \left\{ \lambda_j e_j' - \sum_{i=1}^r \lambda_i (e_j' e_i) e_i' \right\} x = 0$$

por lo tanto  $x$  y  $e_j$  son ortogonales;  $j=1, \dots, r$ .

Por lo tanto

$$\theta x = Bx = \left( A - \sum \lambda_i e_i e_i' \right) x = Ax - \sum \lambda_i (e_i' x) e_i = Ax$$

y  $x$  es también un eigenvector de  $A$ .

Así  $\theta = \lambda_i$  para alguna  $i$ , y  $x$  es una combinación lineal de alguna  $e_i$ , lo cual contradice la ortogonalidad entre  $x$  y  $e_i$

L.Q.Q.D.

Teorema A.11: Teorema de descomposición del valor singular:

Si  $A$  es una matriz  $(n \times p)$  de rango  $r$ , entonces  $A$  puede escribirse como

$$A = ULV'$$

donde  $U(n \times r)$  y  $V(p \times r)$  son matrices de columnas ortonormales ( $U'U = V'V = I_r$ ) y  $L$  es una matriz diagonal con elementos positivos

Demostración: (Mardia,

Como  $A'A$  es una matriz simétrica de rango  $r$ , podemos usar el Teorema A.10 y escribir:

$$A'A = V\Lambda V'$$

donde  $V(p \times r)$  es una matriz de columnas ortonormales de eigenvectores de  $A'A$  y  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$  contiene los eigenvalores distintos de cero.

Todos los  $\lambda_i$  son positivos porque

$$\lambda_i = v'(\lambda) A'A v(\lambda) = \|A v(\lambda)\|^2 > 0 \quad , v(\lambda) \text{ es la } i\text{-ésima columna de } V$$

Sea

$$l_i = \lambda_i^{1/2} \quad i=1, \dots, r$$

y sea

$$L = \text{diag}(l_1, \dots, l_r)$$

Definimos  $U(n \times r)$  por

$$u(\lambda) = l_i^{-1} A v(\lambda) \quad ; \quad \lambda=1, \dots, r$$

Entonces

$$u(\lambda) u(\lambda) = l_i^{-1} l_j^{-1} v'(\lambda) A'A v(\lambda) = \lambda_i l_i^{-1} l_j^{-1} v'(\lambda) v(\lambda) = \begin{cases} 1, & \lambda=j \\ 0, & \lambda \neq j \end{cases}$$



Así  $U$  también es una matriz de columnas ortogonales.

Cualquier vector  $x \in \mathbb{R}^p$  puede escribirse como

$$x = \sum \alpha_i v(i) + y$$

donde  $y \in N(A)$ , el espacio nulo de  $A$ . Nótese que  $N(A) = N(A^t A)$  es el eigenspacio de  $A^t A$  para el eigenvalor cero, por lo tanto  $y$  es ortogonal a los eigenvectores  $v(i)$

Sea  $e_i$  el vector  $(1 \times r)$  con 1 en la  $i$ -ésima componente y 0 en las demás. Entonces:

$$ULV^t x = \sum \alpha_i UL e_i + 0$$

$$= \sum \alpha_i u_i + 0$$

$$= \sum \alpha_i A v(i) + Ay = Ax$$

Como la ecuación se satisface para toda  $x \in \mathbb{R}^p$ , entonces

$$ULV^t = A$$

L.Q.Q.D.

NOTA: Las columnas de  $U$  son eigenvectores de  $AA^t$  y las columnas de  $V$  son eigenvectores de  $A^t A$ . Además, por A.6, los eigenvalores de  $AA^t$  y  $A^t A$  son los mismos

Teorema A.12: Sea  $V$  un espacio vectorial con producto interior y sea  $T$  un operador lineal en  $V$ . Entonces  $T$  es una proyección ortogonal si y sólo si  $T^2 = T = T^*$

Demostración:

Para demostrar este teorema utilizamos el siguiente

Teorema: Sea  $V$  un espacio con producto interior de dimensión finita y sea  $T$  un operador lineal en  $V$ . Entonces existe un operador lineal único  $T^*$  en  $V$  tal que

$$\langle T(x), y \rangle = \langle x, T^*(y) \rangle \quad \forall x, y \in V.$$

$$\text{Sean } N(T) = \{ x \in \mathbb{R}^n \mid T(x) = 0 \}$$

$$M(T) = \{ y \in \mathbb{R}^n \mid T(y) = y \}$$

$$V = \mathbb{R}^n$$

( $\Rightarrow$ ) Demostraremos i)  $T^2 = T$   
ii)  $T = T^*$

i) Sea  $x \in \mathbb{R}^n$

entonces  $T(x) \in M(T) = W_1$

Por definición  $T: V \rightarrow V$  es una proyección sobre  $W_1$  si existe  $W_2$  tal que  $W_1 \oplus W_2 = V$

y  $T(y) = y \quad \forall y \in W_1$ . Por lo tanto

$$T(T(x)) = T(x)$$

$$\therefore T^2 = T$$

ii) Si  $T$  es una proyección ortogonal, por definición

$$V = M(T) \oplus N(T) \quad \text{y} \quad M(T)^\perp = N(T)$$

donde

$$M(T)^\perp = \{x \in V \mid \forall m \in M(T), m \perp x\}$$

Sean  $x, y \in V$ , entonces

$$x = x_1 + x_2 \quad y = y_1 + y_2$$

$$\text{donde } x_1, y_1 \in M(T) \quad ; \quad x_2, y_2 \in N(T)$$

Por lo tanto

$$\begin{aligned} \langle x, T(y) \rangle &= \langle x_1 + x_2, y_1 \rangle = \langle x_1, y_1 \rangle + \langle x_2, y_1 \rangle \\ &= \langle x_1, y_1 \rangle \end{aligned}$$

$$\begin{aligned} \langle x, T^*(y) \rangle &= \langle T(x), y \rangle = \langle x_1 + x_2, y_1 + y_2 \rangle \\ &= \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle \\ &= \langle x_1, y_1 \rangle \end{aligned}$$

Así:

$$\langle x, T(y) \rangle = \langle x, T^*(y) \rangle \quad \forall x, y \in V$$

Por lo tanto:

$$T = T^*$$

( $\Leftarrow$ ) Demostraremos i)  $T$  es una proyección

ii)  $T$  es ortogonal.

1) Sea  $x \in V$   $\nabla$ .  $TT(x) = T(x) \Leftrightarrow T(x - T(x)) = 0$   
por linealidad

Sea  $y = x - T(x) \in N(T)$ . Entonces

$$x = y + T(x) \quad \text{con } y \in N(T), T(x) \in M(T)$$

Sea  $z \in N(T) \cap M(T)$ , entonces

$$\text{existe } y \in V \text{ tal que } z = T(y)$$

$$\therefore 0 = T(z) = TT(y) = T(y) = z$$

Por lo tanto  $V = M(T) \oplus N(T)$

ii) Sean  $x \in M(\tau)$ ,  $y \in N(\tau)$ . Entonces:

$$\underline{x} = T(x) = T^*(x), \quad y \text{ así:}$$

$$\langle x, y \rangle = \langle T^*(x), y \rangle = \langle x, T(y) \rangle = \langle x, 0 \rangle = 0$$

Por lo tanto  $x \in N(\tau)^\perp$

de donde se tiene  $M(\tau) \subseteq N(\tau)^\perp$

Sea  $y \in N(\tau)^\perp$ . Debemos demostrar que

$y \in M(\tau)$ , esto es, que  $T(y) = y$

Ahora bien,

$$\begin{aligned} \|y - T(y)\|^2 &= \langle y - T(y), y - T(y) \rangle \\ &= \langle y, y - T(y) \rangle - \langle T(y), y - T(y) \rangle \end{aligned}$$

Como  $y - T(y) \in N(\tau)$  y  $y \in N(\tau)^\perp$

$$\langle y, y - T(y) \rangle = 0$$

Pero también  $\langle T(y), y - T(y) \rangle$

$$= \langle y, T^*(y - T(y)) \rangle$$

$$= \langle y, T(y - T(y)) \rangle = \langle y, 0 \rangle = 0$$

Por lo tanto  $y - T(y) = 0$ , esto es

$$y = T(y) \in M(\tau)$$

$$\therefore M(\tau) = N(\tau)^\perp$$

L.Q.Q.D.

NOTA: Este teorema puede entenderse de la siguiente forma:

"Sea  $V$  un espacio de dimensión  $n$  con producto interior, sea  $A$  ( $n \times n$ ) una matriz y sea  $T$  su operador lineal asociado en  $V$ . Entonces  $T$  es una proyección ortogonal si y sólo si  $A$  es simétrica e idempotente"

Teorema A.13:  $A (n \times n)$  tiene inversa si y sólo si  
 $\text{rango}(A) = n$

Demostración:

( $\Leftarrow$ ) Sea  $A (n \times n)$  tal que  $\text{rango}(A) = n$   
entonces el número de columnas linealmente  
independientes de  $A$  es  $n$ , de  
donde  $|A| \neq 0$   
 $\therefore A$  invertible.

( $\Rightarrow$ ) Supongamos  $A (n \times n)$  tal que existe una  
matriz  $A^{-1} (n \times n)$   $\cdot$   $A^{-1}A = AA^{-1} = I_n$   
Notese que  $\text{rango}(I_n) = \text{rango}(A^{-1}A) = n$

Sean  $N(A) = \{ \underline{x} \in \mathbb{R}^n \cdot A\underline{x} = 0$

y  $N(A^{-1}A) = \{ \underline{x} \in \mathbb{R}^n \cdot A^{-1}A\underline{x} = 0 \}$

$\underline{x} \in N(A)$  ssi  $A\underline{x} = 0$

$\Rightarrow A^{-1}A\underline{x} = 0$  ssi  $\underline{x} \in N(A^{-1}A)$

$\underline{x} \in N(A^{-1}A)$  ssi  $A^{-1}A\underline{x} = 0$

$\Rightarrow AA^{-1}A\underline{x} = 0$  ssi  $\underline{x} \in N(A)$

$\therefore N(A) = N(A^{-1}A)$

Como  $A$  y  $A^{-1}A$  tienen ambas  $n$  columnas

$\dim M(A) = \dim M(A^{-1}A)$

$\therefore \text{rango}(A) = \text{rango}(A^{-1}A) = n$

L.Q.Q.P.

Teorema A.14: Si  $A$  es una matriz  $(n \times p)$ ,  $p < n$ , de rango  $p$ ; entonces  $A'A$  es invertible

Demostración:

$A'A$  es de orden  $(p \times p)$

Por A.13 tenemos que  $A'A$  es invertible si y sólo si  $\text{rango}(A'A) = p$

pero

$\text{rango}(A'A) = \text{rango}(A)$  por A.8

$\therefore \text{rango}(A'A) = p$

Por lo tanto

$A'A$  invertible L.A.A.D.

## BIBLIOGRAFIA

- Bartle, R. Introducción al Análisis Matemático. Limusa S.A., México D.F., 1982.
- Başar, E. E.E.G. - Brain Dynamics. Relation between EEG and brain evoked potentials. Elsevier / North - Holland. Biomedical Press, Amsterdam, 1980.
- Cravioto, J. y Arrieta, R. Nutrición, desarrollo mental, conducta y aprendizaje. UNICEF, México, 1982.
- Friedberg, S.H., Insel, A.J. y Spence, L.E. Algebra Lineal. Publicaciones Cultural, S.A., México, 1982.
- Gittins, R. Ecological Applications of Canonical Analysis en Multivariate Methods in Ecological Work. L.Orloci, C.R. Rao y W.M. Stiteler (eds.) p.p. 309 - 535. International Co-operative Publishing House, Fairland, Maryland, 1979.
- Harmony, T. Functional Neuroscience. Volume III. Neurometric Assessment of Brain Dysfunction in Neurological Patients. Lawrence Erlbaum Associates Publishers, London, 1984.

- Hinojosa, G. y Rocha, C. Evaluación de la lectura en niños de primaria ; reporte preliminar. V Coloquio Interno de Investigación, ENEP - Iztacala , 1985.
- John, E. R. , Prichep, L. , Ahn, H. , Easton, P. , Fridman, J. y Kaye, H. Neurometric evaluation of cognitive dysfunctions and neurological disorders in children. Progress in Neurobiology. Vol 21 p.p. 239-290 , 1983.
- Kotz, S. y Johnson, N. L. (eds.) Encyclopedia of Statistical Sciences. Vol 1 p.p 355-365, 1982.
- Lancaster, H. O. The structure of bivariate distributions. Annals - Math. Stat. 29 p.p. 719-736, 1958.
- Lebart, L. , Morineau, A. y Warwick, K. M. Multivariate Descriptive Statistical Analysis. John Wiley & Sons, Inc. , New York, 1984.
- Luria, A. R. Las funciones corticales superiores del hombre. Editorial Científico-Técnica, La Habana , 1982.



Mardia, K.V., Kent, J.T. y Bibby, J.M. Multivariate Analysis. Academic Press, London, 1979.

Marsden, J.E., Tromba, A.J. Cálculo vectorial. Fondo Educativo Interamericano, S.A., México D.F., 1976.

McGillem, C.D. y Auñón, J.I. Signal Processing of event related brain potentials. Purdue University, West Lafayette, Indiana, 1981.

Paniagua, G.M. El Análisis de Correspondencias, un método de Análisis Multivariado Descriptivo. (tesis de licenciatura en actuaría. ITAM.) México D.F., 1986.

Rebert, Ch.S., Wexler, B.N. y Sproul, A. EEG asymmetry in educationally handicapped children. Electroencephalography and Clinical Neurophysiology. Vol 45 pp. 436-442, 1978.

Rutter, M. (ed.) Developmental Neuropsychiatry. Guilford Press, New York, 1983.

- Saporta, G. Liaisons entre plusieurs ensembles de variables et codage de données qualitatives. (tesis de doctorado 3er ciclo). L'Universite Pierre et Marie Curie, 1975.
- Sklar, B., Hanley, J. y Simmons, W.W. An EEG experiment aimed toward identifying dyslexic children. Nature (London) Vol 240 p.p. 414 - 416, 1972.
- Velasco, M., López, M., Rojas-Ramírez, J. A. y Velasco, F. Valor del electroencefalograma en el diagnóstico neurológico y psiquiátrico: estudio correlativo de 10,000 pacientes. Archivos de investigación médica. Vol 4 p.p. 125-138, 1973.