



UNIVERSIDAD DE BILBAO.

FACULTAD DE CIENCIAS ECONOMICAS Y EMPRESARIALES.

"GENERALIZACIONES EN EL ANALISIS CANONICO".

Tesis Doctoral presentada por Antonio Jesús Baigorri Matamala, dirigida por el Catedrático de la Universidad de Bilbao Dr. D. Antonio Fz. de Trocóniz.

BILBAO, SEPTIEMBRE DE 1.979.

A mis padres.

## CON MI AGRADECIMIENTO

Deseo expresar mi agradecimiento a D. Antonio Fz. de Trocóniz por su generosa ayuda y estímulo, tanto en la dirección de esta Tesis como a lo largo de toda mi vida académica.

Igualmente a mis compañeros Maribel García del Valle, Anjeles Iztueta, Karmele Fz. Aguirre, Carmen Balzategui, Reyes Isusi, M<sup>a</sup> José Mijangos, Roberto Escuder, Julio Grafe, Fernando Tusell, José M<sup>a</sup> Piris, José E. Jiménez, Alberto Marínez, Juan Ig. Iñurrieta, José M<sup>a</sup> Madurga y Jesús Alvarez por sus atenciones y aclaraciones que me han prestado. También estoy en deuda con el Padre Chacón por la amabilidad con que atendió mis consultas. No quiero olvidar tampoco a José M<sup>a</sup> Caridad y Juan - Urrutia, profesores que orientaron mi vocación. Por último deseo agradecer a Paquita Martín su cuidadosa mecanografía.

Antonio J. Baigorri Matamala .

INDICE. -

	página
INTRODUCCION .....	1

### PARTE I

#### EL MODELO MATEMATICO

#### CAPITULO I. EL MODELO CANONICO DE DOS VECTORES DE VARIABLES ALEATORIAS

1.1. Introducción .....	15
1.2. Definición del modelo canónico .....	18
1.3. La existencia de los vectores canónicos. Su determinación y significado .....	20
1.3.1. Consideraciones previas .....	20
1.3.2. Prueba de la existencia de los vectores canónicos .....	21
1.3.3. Determinación de los vectores canónicos .....	26
1.3.4. Descripción de los vectores canónicos .....	30
1.4. El modelo canónico no normalizado .....	32
1.5. Propiedades del modelo canónico .....	35
1.6. Relajación de hipótesis del modelo .....	52
1.7. La interpretación en el modelo canónico ..	55
1.8. Distribución de los vectores canónicos poblacionales .....	59

#### CAPITULO II. LOS MODELOS CANONICOS PARCIALES. DOS POSIBLES EXTENSIONES.

2.1. Introducción .....	61
2.2. Modelos parciales .....	65
2.2.1. Modelo parcial $L(3)-C(3)$ .....	65
2.2.2. Modelo parte .....	68
2.2.3. Modelo biparcial .....	70

2. 2. 4. Modelo $G_1$ .....	72
2. 2. 5. Modelo $G_2-L(5)-C(5)$ .....	75
2. 3. Dos generalizaciones en los modelos canónicos parciales .....	80
2. 3. 1. Modelos parciales $L(7)$ y $C(7)$ .....	83
2. 3. 1. 1. Modelo $L(7)$ .....	84
2. 3. 1. 2. Modelo $C(7)$ .....	88
2. 3. 2. Modelos parciales $L(2n+1)$ y $C(2n+1)$ ..	95
2. 3. 2. 1. Modelo $L(2n+1)$ .....	95
2. 3. 2. 2. Modelo $C(2n+1)$ .....	99
2. 4. Los modelos parciales en las distribuciones - normales multivariantes .....	101

### CAPITULO III. LOS MODELOS CANONICOS GENERALI- ZADOS. MODELOS PARCIALES GENE- RALIZADOS.

3. 1. Introducción .....	104
3. 2. Imposibilidad de generalización del modelo ca- nónico manteniéndose las principales propieda- des del modelo de dos vectores de variables ...	107
3. 3. Método de máxima correlación .....	109
3. 4. Método de aproximación a una matriz de rango uno .....	115
3. 5. Comparación de ambos procedimientos .....	120
3. 6. Modelos canónicos parciales generalizados ...	121
3. 6. 1. Modelo parcial generalizado .....	121
3. 6. 2. Modelo parte generalizado .....	123
3. 6. 3. Modelo biparcial generalizado .....	124
3. 6. 4. Modelo $G_1$ generalizado .....	124

3.6.5. Modelo $G_2$ generalizado .....	125
--	-----

PARTE II

EL MODELO MUESTRAL

CAPITULO IV. EL MODELO CANONICO MUESTRAL DE  
DOS VECTORES DE VARIABLES ALEA-  
TORIAS .

4.1. Introducción .....	131
4.2. Correlaciones y vectores canónicos mues- trales .....	135
4.3. Estimadores de máxima verosimilitud de las correlaciones y vectores de coeficien- tes canónicos .....	138
4.4. Distribuciones en el muestreo de las co- rrelaciones y vectores de coeficientes - canónicos muestrales .....	141
4.4.1. Consideraciones preliminares ...	141
4.4.2. Distribución en el muestreo de - las correlaciones canónicas mues- trales .....	151
4.4.3. Distribución en el muestreo de los vectores de coeficientes canónicos muestrales .....	153
4.5. Contrastes de hipótesis en el modelo ca- nónico .....	156
4.5.1. Distribuciones utilizadas en los contrastes de hipótesis relativos al modelo canónico .....	157

4.5.1.1.	Criterio de Wilks .....	158
4.5.1.1.1.	Distribución aproximada .....	160
4.5.1.1.2.	Distribución exacta .....	161
4.5.1.2.	Criterios de Pillai .....	167
4.5.1.3.	Criterio de Hotelling .....	171
4.5.1.4.	Criterio de Roy .....	172
4.5.2.	Contrastes de hipótesis relativos a las correlaciones canónicas .....	174
4.5.2.1.	Contraste de la hipótesis de que todas las correlaciones canónicas poblacionales son nulas .....	174
4.5.2.2.	Contraste de la hipótesis de que parte de las correlaciones canónicas poblacionales son nulas .....	178
4.5.3.	Contrastes de hipótesis relativos a los coeficientes del modelo de regresión lineal multivariante de X sobre Y .....	180
4.5.3.1.	Contraste de la hipótesis de que la matriz de coeficientes de regresión $B_a$ es una matriz prefijada $B_a^*$ .....	181
4.5.3.2.	Contraste de la hipótesis de que la matriz de coeficientes de regresión $B_a = [0]$ .....	182

## CAPITULO V. LOS MODELOS MUESTRALES DE LOS MODELOS CANONICOS PARCIALES

5.1.	Introducción .....	185
5.2.	El modelo muestral del análisis canónico parcial de Rao .....	188
5.2.1.	Consideraciones preliminares .....	188
5.2.2.	Definición de las correlaciones y vectores	

canónicos parciales muestrales .....	189
5. 2. 3. Estimadores de máxima verosimilitud de las - correlaciones y vectores canónicos parciales	190
5. 2. 4. Distribuciones en el muestreo de las correla- ciones y vectores canónicos parciales muestr <u>a</u> les .....	191
5. 2. 4. 1. Introducción .....	191
5. 2. 4. 2. Distribución en el muestreo de cier- tas matrices de variables aleatorias.	192
5. 2. 4. 3. Distribución en el muestreo de las - correlaciones canónicas parciales - muestrales .....	194
5. 2. 4. 4. Distribución en el muestreo de los vectores de coeficientes canónicos parciales muestrales .....	195
5. 2. 5. Pruebas de hipótesis en el modelo canóni- co parcial de Rao .....	196
5. 2. 5. 1. Contrastes de hipótesis relati- vos a las correlaciones canóni- cas parciales .....	196
5. 2. 5. 1. 1. Contraste de la hipótesis - de que todas las correlacio- nes canónicas parciales son nulas .....	196
5. 2. 5. 1. 2. Contraste de la hipótesis de que parte de las correlacio- nes canónicas parciales son nulas .....	198
5. 3. El modelo muestral en el resto de modelos par-- ciales .....	198

5.4. Referencia al modelo muestral de los modelos canónicos generalizados .....	200
CONCLUSIONES .....	203
BIBLIOGRAFIA CONSULTADA .....	209

INTRODUCCION. -

El estudio de las relaciones y dependencias entre variables aleatorias ocupa un lugar importante dentro de la teoría estadística. Siguiendo a M. G. Kendall el tratamiento de las interdependencias de tales variables conduce a la teoría de la correlación mientras que el de las dependencias lo hace a la teoría de la regresión. Dentro de las relaciones de dependencia han sido las relaciones lineales las que han alcanzado un mayor desarrollo en su estudio concretándose en la teoría de la regresión lineal.

Desde los primeros trabajos debidos a Pearson y Yule, la teoría de la correlación y regresión entre variables aleatorias ha experimentado un gran avance tanto en sus modelos matemáticos como muestrales hasta llegar a su última expresión que es el análisis canónico. Tal análisis constituye la generalización natural del modelo de regresión múltiple al caso en que el regresando lo constituya un vector de variables aleatorias y es debido a Harold Hotelling. Este autor, en 1935, expuso el procedimiento de obtener, dados dos vectores de variables aleatorias, aquellas combinaciones lineales del primer y segundo vector que tengan máxima correlación, denominando a tal correlación, coeficiente de correlación canónica entre los vectores aleatorios de referencia. Con posterioridad, en 1936, Hotelling completa el procedimiento extendiéndolo a sucesivas combinaciones lineales incorreladas con las obtenidas previamente y que tengan máxima correlación siendo el número de parejas de combinaciones lineales y de coeficientes de correlación canónica el rango de la matriz de covarianzas inter-vectores. En tal procedimiento se originan dos nuevos vectores de variables aleatorias a los que denominaremos vectores canónicos.<sup>(1)</sup>

---

(1) Véase 1.2.

Un nuevo avance dentro del análisis canónico lo da Paul Horst en 1961, al establecer cuatro criterios de generalizar el análisis canónico de Hotelling a más de dos vectores de variables aleatorias si bien en tales casos se pierden alguna de las propiedades del modelo de dos vectores. En estas generalizaciones que en lo sucesivo nos referiremos a ellas como "análisis canónico generalizado" no se logra la incorrelación entre componentes de los nuevos vectores originados, con distinto ordinal.

R. B. Rao, en 1969, expone una nueva línea de estudio en los modelos canónicos como es el análisis canónico parcial. Este análisis consiste en generalizar a vectores aleatorios el concepto de correlación parcial entre dos variables aleatorias. Esta línea de trabajo no acabó aquí; otros autores como Timm y Carlson en 1976 generalizan los coeficientes de correlación parte y biparcial entre dos variables aleatorias a vectores aleatorios, caso este último que es más general que el estudiado por Rao. Con posterioridad, Sik-Yum-Lee (1978) presenta dos nuevos modelos con las mismas características anteriores, pero con estructuras de dependencia entre los vectores de variables que intervienen que incluyen como casos particulares a los modelos de Rao y Timm y Carlson. Nos referiremos genéricamente a todo este conjunto de modelos como "modelos canónicos parciales" por ser su contenido el análisis canónico entre vectores aleatorios en los que se han eliminado las influencias lineales de otros.

Hasta este punto hemos situado los diferentes campos de desarrollo del análisis canónico. Fijadas estas ideas señalaremos los objetivos que hemos querido llevar a cabo en esta memoria.

1. - Generalizar los modelos parciales a estructuras de dependencia cuyos casos particulares son los modelos de Sik-Yum-Lee y por tanto el resto de modelos parciales.

2. - Ampliar los modelos canónicos generalizados en la línea de los modelos parciales.

3. - Obtener los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos parciales, así como sus distribuciones en el muestreo.

4. - Separación del modelo matemático y modelo muestral en el conjunto de desarrollos que contiene este trabajo.

Esta serie de objetivos constituyen las diferentes aportaciones originales en cada uno de los campos del análisis canónico que contiene esta memoria; nos referiremos con detalle a estos cuatro puntos.

Por lo que respecta al primero de ellos expondremos el contenido de las generalizaciones a efectuar. El modelo canónico parcial más general que hemos encontrado en la literatura estadística ha sido el modelo  $G_2$ -biparcial de Sik-Yum-Lee. En él se obtienen las correlaciones y vectores canónicos entre dos vectores de variables aleatorias previa eliminación de las influencias lineales de otros tres, cuya relación con los dos primeros supone una estructura de dependencias más general que la de los modelos de Rao y Timm y Carlson. Los modelos canónicos parciales  $L(7)$  y  $C(7)$  que aquí presentaremos constituyen dos nuevos modelos que generalizan a  $G_2$ .

En ellos se consideran las correlaciones y vectores canónicos entre dos vectores de variables aleatorias en las que se han eliminado las influencias lineales de otros cinco ligados a los anteriores bajo estructuras de dependencia de las que son casos particulares el modelo  $G_2$ . Además el modelo C(7) incluye como caso particular al modelo L(7).

Tanto los modelos de Rao,  $G_2$ , L(7). (Igualmente Rao,  $G_2$ , C(7)) pueden estudiarse en forma genérica para cualquier número impar. El resultado son los nuevos modelos L(2n+1) y C(2n+1) Basados en estos modelos genéricos justificaremos como L(3) y C(3) constituyen el modelo de Rao, L(5) y C(5), el modelo de Sik-Yum-Lee y para  $n > 2$  los modelos L(2n+1) y C(2n+1) presentan estructuras de dependencia diferentes estando siempre la del primero incluida en la del segundo .

En relación al segundo objetivo diremos lo siguiente: Los modelos canónicos generalizados tienen por objeto el estudio de la interdependencia entre más de dos vectores de variables aleatorias. En este trabajo introducimos la idea de los modelos parciales en los modelos canónicos generalizados. Los modelos canónicos resultantes de tal consideración presentan el estudio de las interdependencias entre más de dos vectores aleatorios cuando se han eliminado las influencias lineales de otros. Según la estructura de dependencia supuesta entre los vectores que intervienen hemos obtenido los modelos a los que hemos denominado: "modelo canónico de Rao generalizado", "modelos parte y bipolar generalizados", "modelos  $G_1$  y  $G_2$  generalizados. Nos referiremos genéricamente a todos ellos como "modelos canónicos parciales generalizados".

Por lo que respecta al tercer objetivo precisaremos los siguientes puntos. Las distribuciones en el muestreo de los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos en el modelo de Hotelling se han estudiado para el denominado "null case". En este trabajo hemos obtenido en un primer paso los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos en el análisis canónico parcial de Rao. En base a un resultado probado por T. W. Anderson en relación a las distribuciones en el muestreo de ciertas matrices de variables aleatorias, hemos llegado a las distribuciones en el muestreo también en el "null case" de los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos parciales; las funciones de densidad de estas distribuciones presentan una estructura análoga a las de sus similares en el modelo de Hotelling, diferenciándose únicamente en sus parámetros. Estos resultados completan el modelo muestral del análisis canónico parcial de Rao cuyo estudio fue comenzado por su autor y Timm y Carlson.

Finalmente, refiriéndonos al último objetivo señalaremos que es simplemente metodológico. Suele ser frecuente al estudiar ciertas técnicas de análisis multivariante el hacerlo en forma descriptiva, haciendo uso de las muestras como si de colectivos se tratara sin ninguna aplicación de la inferencia estadística; otras veces, aunque no se mencione expresamente, los modelos están estudiados en las respectivas poblaciones. Estas características comunes a la mayor parte de técnicas multivariantes son señaladas por A. Aznar, al referirse a los mo-

delos Factorial y de Componentes Principales. También el análisis canónico está afectado de esta problemática; salvo muy raras excepciones no se establece una distinción "a priori" de cuando dicho análisis está referido a la población o a muestras de ella. Este ha sido el motivo por el cual hemos dividido en dos partes este trabajo. La primera de ellas está dedicada a la exposición de los modelos matemáticos de los diferentes enfoques que ya hemos comentado del análisis canónico, mientras que la segunda parte contiene los respectivos modelos muestrales.

Expuestos los objetivos que pretende esta memoria entraremos a detallar el contenido de los diferentes capítulos.

Los tres primeros que constituyen la parte dedicada al análisis canónico teórico contienen las diferentes líneas de desarrollo que éste ha recibido. En el capítulo primero exponemos el análisis canónico de dos vectores aleatorios. En él hemos incluido numerosas modificaciones del modelo que expuso Hotelling y que lo han ido completando a lo largo del tiempo. El método seguido en la exposición no es el habitual de obtener las correlaciones y vectores canónicos como resultado de optimizar las funciones coeficiente de correlación entre las sucesivas combinaciones lineales de los vectores iniciales, con las restricciones ya aludidas. Frente a este enfoque hemos escogido el basado en la descomposición singular de una matriz arbitraria. Este planteamiento del análisis canónico de dos vectores aleatorios de referencia permite dar un tratamiento unitario a la solución de otra serie de modelos canónicos como son los modelos parciales, por lo cual lo consideramos básico para los sucesivos

capítulos de esta memoria.

El contenido de este primer capítulo, aparte de ser una recopilación exhaustiva de los diferentes estudios acerca del análisis canónico de Hotelling constituye el soporte teórico en que se basan las generalizaciones que se presentan en este trabajo.

El capítulo II lo hemos dedicado a los modelos canónicos parciales, y en él hemos distinguido dos secciones. En la primera de ellas exponemos los diferentes modelos parciales publicados en la literatura dedicada a estos temas; tales son: el modelo parcial de Rao, los modelos parte y biparcial de Timm y Carlson y los modelos  $G_1$  y  $G_2$ -biparcial de Sik-Yum-Lee. En ellos hemos introducido grafos para señalar las dependencias entre vectores pudiéndose así obtener una idea visual de cómo unos modelos generalizan a otros. Este conjunto de modelos constituye el estado actual del análisis canónico parcial.

La segunda sección la hemos dedicado a exponer los nuevos modelos canónicos parciales que generalizan al modelo  $G_2$ -biparcial. A tales modelos generalizados les hemos denominado extensión lineal y circular del modelo de Rao. (Esta denominación está basada en la estructura de dependencias de los grafos asociados a tales modelos y en que dichos modelos constituyen los diferentes puntos de dos cadenas cuyos dos primeros puntos son en ambas el modelo de Rao y  $G_2$ -biparcial). Dentro de esta segunda sección hemos distinguido dos partes. En la primera estudiamos los modelos  $L(7)$  y  $C(7)$  (tercer punto de ambas cadenas) y en la segunda los modelos que hemos denominado  $L(2n+1)$  y  $C(2n+1)$  los cuales hacen referencia en forma genérica

a las dos cadenas en que puede extenderse el modelo de Rao.

En el tercer capítulo estudiamos el análisis canónico generalizado a más de dos vectores aleatorios. También lo hemos dividido en dos secciones. En la primera exponemos los diferentes enfoques bajo los que pueden obtenerse las correlaciones y vectores canónicos generalizados. Tales modelos canónicos constituyen los últimos resultados que de esta línea del análisis tenemos constancia.

En la segunda sección exponemos los nuevos modelos resultantes de aplicar tales enfoques en vectores en los que ya se han eliminado influencias lineales de otros. Al generalizar las diferentes estructuras de dependencia de los modelos parciales a más de dos vectores de variables aleatorias hemos llegado a los modelos parciales generalizados que anteriormente hemos comentado. Los nuevos modelos  $L(2n+1)$  y  $C(2n+1)$  también admiten esta generalización considerando bloques de vectores.

En la segunda parte estudiamos el modelo muestral de los diferentes enfoques del análisis canónico. Es bien sabida la dificultad que presenta el desarrollo de la teoría muestral de cualquier modelo multivariante; por referirnos al análisis canónico de dos vectores de variables aleatorias, diremos que existen cuestiones dentro de esta teoría aún no resueltas. En gran parte de este tipo de análisis los contrastes de hipótesis suelen referirse a los denominados en la literatura estadística "null cases"; en general estos contrastes están fundamentados en distribuciones de Wishart centradas que conducen a estadísticos cuya distribución en el muestreo es la  $\Lambda$  de

Wilks,  $T_0^2$  de Hotelling, etc.; en cualquier otro caso las distribuciones de tales estadísticos o bien no son conocidas o únicamente se pueden presentar de forma aproximada. Hemos querido introducir este comentario para justificar de alguna manera las fuertes restricciones bajo las que se ha desarrollado la teoría muestral en el análisis canónico.

Centrándonos ya en el contenido de esta segunda parte, en ella hemos distinguido dos capítulos, el primero está dedicado al modelo muestral del análisis canónico de Hotelling, así como a ciertas cuestiones de carácter general relativas a las distribuciones de ciertos estadísticos de amplio uso en los contrastes de hipótesis multivariante. Dicho capítulo, sumamente extenso, contiene el procedimiento de obtener estimadores de máxima verosimilitud para las correlaciones y vectores de coeficientes canónicos, así como las distribuciones en el muestreo de tales estimadores en el "null case". También hemos incluido una serie de contrastes de hipótesis en relación con las correlaciones canónicas y los coeficientes del modelo de regresión lineal multivariante asociado.

Podría extrañar la no dedicación de sendos capítulos a los modelos muestrales del análisis canónico parcial y del generalizado tal como se ha hecho en el modelo matemático. El motivo puede encontrarse en el comentario que hemos incluido al hablar de los modelos muestrales. Los modelos parciales han tenido muy poco desarrollo en este campo siendo nulo el de los generalizados. Ante esta situación hemos incluido en el último capítulo una serie de resultados en torno a estos dos tipos de modelos. El contenido fundamental lo constituye el modelo muestral del análisis canónico parcial de Rao. Este autor en la exposición de tal análisis, junto al modelo matemático, desarrolla el

procedimiento de obtener estimadores de máxima verosimilitud para las correlaciones canónicas. Posteriormente Timm y Carlson construyen una prueba de hipótesis para tales correlaciones basada en el criterio  $\Lambda$  de Wilks. En este trabajo hemos completado en varios puntos el modelo muestral de Rao, como son los siguientes:

- 1º Obtención de los estimadores de máxima verosimilitud para los vectores de coeficientes canónicos.
- 2º Obtención de la distribución en el muestreo de los estimadores de máxima verosimilitud de las correlaciones canónicas parciales, en el "null case".
- 3º Obtención de la distribución en el muestreo de los estimadores de máxima verosimilitud de los vectores de coeficientes canónicos parciales en el "null case".
- 4º Construcción de pruebas de hipótesis en relación a las correlaciones canónicas parciales basadas en los criterios de Pillai Hotelling y Roy.

En relación con el resto de modelos parciales diremos que no hemos entrado en el estudio de sus modelos muestrales por lo que únicamente expondremos algunos tests de hipótesis debidos a sus autores acerca de las correlaciones canónicas. Igualmente en los modelos canónicos generalizados incluimos como cuestión ligada a su modelo muestral el contraste de hipótesis de independencia entre  $m > 2$  vectores de variables aleatorias.

Resumen en breves líneas, las aportaciones fundamentales que contiene esta memoria, citaremos las siguientes:

- 1º Desarrollo de los modelos parciales  $L(2n+1)$  y  $C(2n+1)$  representaciones generales de las dos cadenas en que puede generalizarse el modelo de Rao, de las cuales dicho modelo y el  $G_2$  bi-parcial son los dos primeros puntos en ambas. Cuando  $n > 2$  la extensión circular incluye a la lineal.
- 2º Introducción de un nuevo tipo de modelos canónicos como son los parciales generalizados. En ellos se generaliza a más de dos vectores aleatorios los modelos canónicos parciales.
- 3º Obtención de los estimadores de máxima verosimilitud para los vectores de coeficientes canónicos en el análisis canónico parcial de Rao, así como la distribución en el muestreo de tales vectores y la de los estimadores de máxima verosimilitud de las correlaciones canónicas parciales ambas en el "null case".

CAPITULO I: El modelo canónico de dos vectores de variables aleatorias.

PARTE PRIMERA:

EL MODELO MATEMATICO.

## 1.1. Introducción

La relación entre dos vectores de variables aleatorias fue estudiada por primera vez por Harold Hotelling<sup>(1)</sup>. Este autor en 1935 en su célebre artículo "The most predictable criterion" expone el procedimiento de obtener aquellas combinaciones lineales de cada vector de variables que presenten máxima correlación.

Desde esa fecha el denominado análisis canónico de dos vectores de variables ha tenido una amplia sistematización. Prueba de ello son los trabajos posteriores de Hotelling<sup>(1)</sup>, Anderson<sup>(2)</sup>, T.W., Morrison<sup>(3)</sup> D., Kshirsagar<sup>(4)</sup>, etc.

Este profuso desarrollo del modelo canónico ha venido motivado en parte, por el hecho de ser un marco general dentro del que se engloban un elevado número de técnicas del análisis multivariante.

El análisis canónico poblacional referido a dos vectores de variables que vamos a desarrollar en este capítulo presenta una serie de diferencias en relación al de Hotelling debidas unas al método seguido y otras a las extensiones que de él se han realizado en algunos campos.

La línea que seguiremos en la exposición del modelo ha sido la adoptada por Horst<sup>(5)</sup> y Kshirsagar utilizando la descompo-

---

(1) Hotelling, H. [ 22 ] pág. 139-142 [ 23 ] pág. 321-377.

(2) Anderson, T.W. [ 2 ] pág. 272-287.

(3) Morrison, D. [ 33 ] pág. 213-218.

(4) Kshirsagar, A. [ 28 ] pág. 247-285.

(5) Horst, P. [ 14 ] pág. 130-132.

sición singular de una matriz. Hemos preferido este tipo de enfoque frente al presentado por Anderson, Morrison y otros autores, por poder dar un tratamiento unitario a todos los modelos que a lo largo de este trabajo se expondrán.

Dado que este enfoque no presenta la claridad del citado en segundo lugar hemos incluido en este capítulo un amplio estudio de las propiedades del modelo que dan luz sobre el sentido y significado de las correlaciones y vectores canónicos.

Otra cuestión en que se pone especial énfasis es en la de señalar cómo el modelo canónico es el modelo de regresión múltiple generalizado y en qué consiste dicha generalización.

Hemos juzgado interesante incluir algunas aportaciones recientes que desarrollan el modelo poblacional. En este capítulo únicamente haremos referencia a las relativas al modelo de dos vectores de variables y no parcial. Entre ellas destacaremos las dos siguientes:

- (1) Extensión del modelo canónico de dos vectores de variables al caso en que estos sean de distribución conjunta degenerada; este desarrollo es debido a Khatri, G.G.<sup>(1)</sup>. Este autor obtiene las correlaciones canónicas en base al concepto de inversa generalizada de una matriz.
- (2) Transformaciones de las variables canónicas que ayudan a su interpretación. Esta línea de trabajo poco desarrollada en el análisis canónico tiene gran importancia por facilitar su aplicación práctica. Otras técnicas multivariantes como el análisis en com-

---

(1) Khatri, G.G. [27], pág. 465-471.

ponentes principales y análisis factorial han tenido un amplio tratamiento en este sentido logrando una serie de métodos que constituye una ayuda importante para la interpretación de las nuevas variables obtenidas. En el análisis canónico propiamente dicho, únicamente hemos encontrado un trabajo en este sentido; nos estamos refiriendo al artículo de Cliff, N. y Kruss, F. <sup>(1)</sup> sobre la interpretación de los vectores canónicos. En base a él justificaremos el empleo para este menester de los métodos de rotación ortogonal utilizados en el análisis factorial.

La terminología que se introduce en este capítulo para designar a los distintos elementos del modelo canónico ha sido sugerida por el Director de esta Memoria. Debemos señalar que algunas denominaciones que aquí empleamos son distintas a las que habitualmente se utilizan en el modelo canónico y que incluso una aun coincidiendo en su expresión gramatical con otra de uso común en el análisis canónico presenta diferente significado y por tanto debe tomarse en el sentido que aquí se da. El motivo de esta modificación está justificado por la mayor precisión que permite dar a las referencias a los distintos elementos del modelo canónico.

Otra precisión a hacer es la relativa a la numeración de las expresiones que aparecen en este trabajo. La numeración por la derecha corresponde a la que se da a la expresión por sí es necesario hacer referencia a ella en otras ocasiones, mientras que la numeración por la izquierda indica la de las expresiones que es útil consultar para obtener la que aquí se hace referencia.

---

(1) Cliff, N. y Kruss, F., [7], pág. 35-43.

## 1.2. Definición del modelo canónico

Sea  $\vec{Z} = (\vec{X}^T, \vec{Y}^T)$  un supervector de variables aleatorias con distribución conjunta no degenerada de parámetros  $(\vec{0}, \Sigma)$ . Denominaremos a  $\vec{X}$  e  $\vec{Y}$  vectores aleatorios de referencia.

Si  $p$  y  $q$  :  $p \leq q$  son las dimensiones de  $\vec{X}$  e  $\vec{Y}$  la matriz de covarianzas <sup>(1)</sup> de  $\vec{Z}$  particionada en base a tales dimensiones la expresaremos:

$$\Sigma_{\vec{Z}} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{matrix} p \\ q \end{matrix} \quad (1)$$

Se denominan vectores canónicos entre  $\vec{X}$  e  $\vec{Y}$  a los vectores de variables  $\vec{U}$  y  $\vec{V}$  que satisfacen (2), (3), (4) y (5)

$$\begin{aligned} \vec{U} &= L^T \vec{X} \\ \vec{V} &= M^T \vec{Y} \end{aligned} \quad (2) \quad \equiv \quad \begin{bmatrix} \vec{U} \\ \vec{V} \end{bmatrix} = \begin{bmatrix} L^T & 0 \\ 0 & M^T \end{bmatrix} \begin{bmatrix} \vec{X} \\ \vec{Y} \end{bmatrix} \quad (3)$$

$$\Sigma_{\vec{U}\vec{V}} = \begin{bmatrix} L^T \Sigma_{11} L & L^T \Sigma_{12} M \\ M^T \Sigma_{21} L & M^T \Sigma_{22} M \end{bmatrix} = \begin{bmatrix} I_p & P \\ P^T & I_q \end{bmatrix} \quad (4)$$

$$P = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_r & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \begin{matrix} \uparrow \uparrow \\ \vdots \\ \downarrow \downarrow \\ \leftarrow p \quad \rightarrow q-p \end{matrix} \quad : \lambda_1 > \lambda_2 > \dots > \lambda_r > 0 \quad (5)$$

en donde  $r = \text{rango } \Sigma_{12}$

(1) Otra posibilidad de desarrollar el modelo canónico es en términos de la matriz de correlaciones. Cooley, W. [8] pág. 168.

Denominaremos a los componentes de  $\vec{U}$  y  $\vec{V}$  variables canónicas entre  $\vec{X}$  e  $\vec{Y}$  y en particular a las parejas de componentes de  $\vec{U}$  y  $\vec{V}$  con el mismo ordinal:  $(u_1, v_1), (u_2, v_2), \dots, (u_r, v_r)$  primeras, segundas... r-ésimas variables canónicas entre  $\vec{X}$  e  $\vec{Y}$ . El resto de componentes de  $\vec{U}$  y  $\vec{V}$  únicamente presentan interés en determinado tipo de desarrollos teóricos. Como se desprende de (4) las variables canónicas están normalizadas en varianza.

A los elementos no nulos de  $P, \lambda_1, \lambda_2, \dots, \lambda_r$  les denominaremos primera, segunda... r-ésima correlación canónica entre  $\vec{X}$  e  $\vec{Y}$  y a las columnas de las matrices L y M vectores de coeficientes canónicos.

En lo sucesivo suprimiremos las flechas en las notaciones referidas a vectores de variables.

### 1.3. La existencia de los vectores canónicos. Su determinación y significado.

#### 1.3.1. Consideraciones previas.

Justificaremos la existencia de los vectores canónicos  $U$  y  $V$  de  $X$  e  $Y$  mediante el siguiente teorema conocido como "descomposición singular de una matriz arbitraria".

Si  $C$  es una matriz de elementos reales de orden  $p \times q$  y de rango  $r \leq \min(p, q)$  dicha matriz puede descomponerse en el producto:

$$C = A P B^T \quad (6)$$

en donde las matrices  $A$ ,  $P$  y  $B$  cumplen las siguientes propiedades:

- $A$  es una matriz ortogonal de orden  $p \times p$ .
- $B$  es una matriz ortogonal de orden  $q \times q$ .
- $P$  es una matriz de orden  $p \times q$  con la estructura señalada en (5) cumpliéndose que sus elementos no nulos son las raíces cuadradas con signo positivo de las raíces características de  $CC^T$  y  $C^T C$  <sup>(1)</sup>

Se demuestra que las  $r$  primeras columnas de las matrices  $A$  y  $B$  son los vectores característicos asociados a dichas raíces en las matrices  $CC^T$  y  $C^T C$ ; el resto de columnas hasta  $p$  y  $q$  pueden completarse mediante las condiciones de ortogonalidad <sup>(2)</sup>.

(1) Las matrices  $CC^T$  y  $C^T C$  tienen  $r$  raíces características positivas y además iguales las de una y otra. Fz. Trocóniz [13] pág. (1.16.21) y (1.18.11).

(2) La prueba de este teorema puede verse en Kshirsagar, A. [28] pág. 247-248.

Dada la estructura de la matriz P la igualdad (6) también puede expresarse

$$C = \lambda_1 a_1 b_1^T + \lambda_2 a_2 b_2^T + \dots + \lambda_r a_r b_r^T \quad (7)$$

siendo  $\lambda_1, \lambda_2, \dots, \lambda_r$  los elementos no nulos de P y  $a_i$  y  $b_i$  las columnas que ocupan el lugar i en las matrices A y B. Una propiedad importante de la descomposición (7) es la siguiente: si  $\lambda_1 > \lambda_2 > \dots > \lambda_r > 0$  y si únicamente se toman  $k \leq r$  sumandos la matriz resultante constituye la mejor aproximación a C a través de una matriz de rango k según el criterio mínimo cuadrático <sup>(1)</sup>.

### 1.3.2. Prueba de la existencia de los vectores canónicos.

Consideremos la siguiente matriz

$$C = t_1^{-1} \Sigma_{12} (t_2^T)^{-1} \quad (8)$$

en donde  $t_1$  y  $t_2$  son matrices triangulares inferiores <sup>(2)</sup> que satisfacen

$$\Sigma_{11} = t_1 t_1^T \quad (9)$$

$$\Sigma_{22} = t_2 t_2^T \quad (10)$$

y  $\Sigma_{11}, \Sigma_{22}, \Sigma_{12}$  los bloques de partición de (1).

La matriz C definida en (8) es de orden  $p \times q$  y su rango r.

$$\text{ran } C = \min [\text{ran } (t_1^{-1}), \text{ran } \Sigma_{12}, \text{ran } (t_2^T)^{-1}] = \min (p, q, r) = r \quad (11)$$

(1) Véase Good, I. J. [15] pág. 823.

(2) El procedimiento de obtener las matrices triangulares  $t_1$  y  $t_2$  que satisfacen (9) y (10) puede verse en Fz. Trocóniz [13] pág. 1. 10.2.

Por el teorema expuesto en 1.3.1.  $C$  puede descomponerse en producto de tres matrices con las características señaladas en (6).

$$(6) \quad C = t_1^{-1} \Sigma_{12} (t_2^{-1})^{-1} = A P B^T \quad (12)$$

$$A^T A = A A^T = I_p \quad (13)$$

$$B^T B = B B^T = I_q \quad (14)$$

siendo  $P$  una matriz de la estructura indicada en (5).

Las matrices  $CC^T$  y  $C^T C$  para este caso concreto adoptan la forma

$$CC^T = t_1^{-1} \Sigma_{12} (t_2^{-1})^{-1} t_2^{-1} \Sigma_{21} (t_1^{-1})^T = t_1^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} (t_1^{-1})^T \quad (15)$$

$$C^T C = t_2^{-1} \Sigma_{21} (t_1^{-1})^{-1} t_1^{-1} \Sigma_{12} (t_2^{-1})^T = t_2^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} (t_2^{-1})^T \quad (16)$$

$A$  y  $B$  por estar formadas por los vectores característicos de  $CC^T$  y  $C^T C$  son las soluciones de los sistemas

$$\left[ t_1^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} (t_1^{-1})^T - \lambda^2 I \right] A = 0 \quad (17)$$

$$\left[ t_2^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} (t_2^{-1})^T - \lambda^2 I \right] B = 0 \quad (18)$$

Las raíces características de  $CC^T$  y  $C^T C$  se designan  $\lambda^2$ . Como antes se indicó los elementos no nulos de  $P$  son las raíces cuadradas con signo positivo de dichas raíces características.

Operando en las ecuaciones (17) y (18) (premultiplicar por  $t_1$  y  $t_2$  respectivamente) se obtiene:

$$[\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 t_1 t_1^T] (t_1^T)^{-1} A = 0 \quad (19)$$

$$[\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 t_2 t_2^T] (t_2^T)^{-1} B = 0 \quad (20)$$

Denominaremos L y M a las matrices de vectores asociados en las ecuaciones anteriores pudiéndose expresar finalmente <sup>(1)</sup> (19) y (20)

$$[\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}] L = 0 \quad (21) \quad : \quad L = (t_1^T)^{-1} A \quad (23)$$

$$[\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 \Sigma_{22}] M = 0 \quad (22) \quad : \quad M = (t_2^T)^{-1} B \quad (24)$$

Cuando las matrices L y M que satisfacen (23) y (24) se toman como matrices de transformación de los vectores aleatorios de referencia X e Y los vectores de variables resultantes tienen una matriz de covarianzas:

$$\Sigma_{uv} = \begin{bmatrix} L^T \Sigma_{11} L & L^T \Sigma_{12} M \\ M^T \Sigma_{21} L & M^T \Sigma_{22} M \end{bmatrix}$$

con las condiciones exigidas en (4) y (5)

En efecto se tiene:

$$(9) (23) \quad L^T \Sigma_{11} L = A^T t_1^T \Sigma_{11} (t_1^T)^{-1} A = A^T t_1^T (t_1^T)^T A = A^T A = I_p \quad (25)$$

$$(10) (24) \quad M^T \Sigma_{22} M = B^T t_2^T \Sigma_{22} (t_2^T)^{-1} B = B^T t_2^T (t_2^T)^T B = B^T B = I_q \quad (26)$$

$$(12) \quad L^T \Sigma_{12} M = A^T [t_1^T \Sigma_{12} (t_2^T)^T] B = A^T [A P B] B = P$$

(1) Los sistemas (17), (18) y (19), (20) tienen las mismas soluciones en  $\lambda^2$  pues

$|\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}| = 0$  ;  $|t_1^T| |\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}| |t_1^T|^T = |t_1^T \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} (t_1^T)^T - \lambda^2 \Sigma_{11}| = 0$   
No ocurre lo mismo con sus vectores asociados, estando relacionados por las ecuaciones (23) y (24).

De esta forma ha sido probado que si los vectores  $U$  y  $V$  se obtienen como transformación lineal <sup>(1)</sup> de los vectores de referencia  $X$  e  $Y$  a través de las expresiones

$$U = L^T X \quad \text{siendo} \quad L = (t_i^1)^T A \quad (27)$$

$$V = M^T Y \quad \text{"} \quad M = (t_i^2)^T B \quad (28)$$

con  $A$ ,  $B$ ,  $t_1$ ,  $t_2$  definidas en (9), (10), (13) y (14), su matriz de covarianzas adopta la forma:

$$(4) \quad \Sigma_{u,v} = \left[ \begin{array}{c|c} I_p & P \\ \hline P^T & I_q \end{array} \right] \quad (29)$$

donde  $P$  tiene la estructura de (5)

El enfoque que aquí hemos expuestos supone una adaptación al modelo poblacional del realizado por Paul Horst <sup>(2)</sup> en matrices de datos.

Un procedimiento similar para obtener los vectores canónicos  $U$  y  $V$  es el expuesto por Kshirsagar, A. <sup>(3)</sup> En él se aplica el teorema expuesto en 1.3.1. a la matriz:

$$C = \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} \quad (30)$$

- (1) Siguiendo a Horst P. [20] pág. 333; en realidad se realizan dos transformaciones, la primera constituye un paso de los vectores  $X$  y  $Y$  a  $\phi_1$  y  $\phi_2$ :  $\phi_1 = t_i^1 X$ ;  $\phi_2 = t_i^2 Y$  cuya matriz de covarianzas conjunta es:

$$\begin{bmatrix} I_p & t_i^1 \Sigma_{12} (t_i^2)^T \\ t_i^2 \Sigma_{21} (t_i^1)^T & I_q \end{bmatrix}$$

La segunda:  $u = A^T \phi_1$ ;  $v = B^T \phi_2$  convierten los elementos no diagonales de la matriz anterior en la matriz  $P$ .

- (2) Horst, P. [19], pág. 130-132.  
 (3) Kshirsagar, A. [28], pág. 250. Otra forma de lograr que los vectores de variables en la primera transformación a que se se meten tengan matriz de covarianzas, unidad dentro de cada grupo es escogiendo ésta de la forma:  $\rho = \Sigma_{11}^{-1/2} X$ ;  $\beta = \Sigma_{22}^{-1/2} Y$ .

Consideremos su descomposición singular

$$C = \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1/2} = E R F^T \quad (31)$$

$$E^T E = E E^T = I_p \quad (32)$$

$$F^T F = F F^T = I_q \quad (33)$$

siendo R con la estructura (5). Las matrices E y F vienen determinadas por las ecuaciones

$$\left[ \Sigma_{11}^{-1/2} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1/2} - \lambda^2 I \right] E = 0 \quad (34)$$

$$\left[ \Sigma_{22}^{-1/2} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1/2} - \lambda^2 I \right] F = 0 \quad (35)$$

transformables en las siguientes

$$\left[ \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right] \Sigma_{11}^{-1/2} E = 0 \quad (36)$$

$$\left[ \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 \Sigma_{22} \right] \Sigma_{22}^{-1/2} F = 0 \quad (37)$$

Es inmediato demostrar que si las matrices L y M de la expresión (2) son de la forma

$$L = \Sigma_{11}^{-1/2} E \quad (38)$$

$$M = \Sigma_{22}^{-1/2} F \quad (39)$$

los nuevos vectores de variables U y V tienen una matriz de covarianzas según se expuso en (4) cumpliéndose por tanto, las condiciones exigidas a los vectores canónicos.

Una comparación de los dos procedimientos de obtención de los vectores canónicos nos permite establecer las siguientes equivalencias:

- Las matrices de (12) y de (31) son iguales puesto que según la nota (1) de la pág. 23, (17) y (18) tienen las mismas soluciones en  $\lambda^2$  que

(21) y (22) y por tanto que (36) y (37).

- Las expresiones  $A^T t_1^{-1}$  y  $E^T \Sigma_{11}^{-1/2}$  son iguales (lo mismo  $B^T t_2^{-1}$  y  $F^T \Sigma_{22}^{-1/2}$ ) puesto que L y M son únicas y vienen determinadas por los vectores asociados a las matrices  $[\Sigma_{12} \Sigma_{11}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}]$  y  $[\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 \Sigma_{22}]$  con sus respectivas condiciones de normalización.

La obtención de los vectores canónicos puede fundamentarse tanto en la descomposición singular de la matriz (8) como (30) ya que éstas conducen a los mismos resultados, no obstante, en lo sucesivo nos referiremos preferentemente a la segunda fórmula puesto que supone ciertas ventajas al hacer referencia a ella.

### 1.3.3. Determinación de los vectores canónicos.

La resolución del modelo canónico requiere hallar las correlaciones canónicas entre los vectores de referencia X e Y y las matrices de coeficientes canónicos L y M. Como se demostró en el apartado anterior tales parámetros satisfacen las ecuaciones

$$(21) \quad [\Sigma_{12} \Sigma_{11}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}] L = 0 \quad : \quad L^T \Sigma_{11} L = I_p$$

$$(22) \quad [\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 \Sigma_{22}] M = 0 \quad : \quad M^T \Sigma_{22} M = I_q$$

Para resolver los sistemas anteriores es necesario hallar las raíces de  $|\Sigma_{12} \Sigma_{11}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11}| = 0$  y seguidamente sus vectores asociados en (21) y (22) con las respectivas condiciones de normalización.

Si  $p$  no es pequeño, el procedimiento descrito anteriormente es muy engorroso puesto que requiere hallar las raíces de polinomios de grado elevado. Para paliar este inconveniente T.W. Anderson<sup>(1)</sup> propone el siguiente procedimiento iterativo para determinar las columnas de  $L$  así como las correlaciones canónicas  $\lambda_1, \lambda_2, \dots, \lambda_v$ <sup>(2)</sup>.

a) Descripción del procedimiento iterativo.

La ecuación (21) puede expresarse para la primera columna:

$$\sum_{12} \sum_{22}^{-1} \sum_{21} p_2 = \lambda_1^2 \sum_{11} p_1$$

siendo equivalente a la siguiente:

$$\sum_{11}^{-1} \sum_{12} \sum_{22}^{-1} \sum_{21} p_2 = \lambda_1^2 p_1 \quad (40)$$

que también puede escribirse<sup>(3)</sup>

$$L \Lambda^2 L^{-1} p_1 = \lambda_1 p_1 \quad : \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \quad (41)$$

en donde  $L = [p_1, \dots, p_p]$  y  $L^{-1} = [p_1^T, \dots, p_p^T]$

Si denominamos  $E = L \Lambda^2 L^{-1}$  y si partimos de una solución inicial  $\delta^0$  para  $I_1$ , el proceso:

$$\begin{aligned} E \delta^0 &\propto p^1 \\ E p^1 &\propto p^2 \\ \vdots & \\ E p^{k-1} &\propto p^k \end{aligned} \quad : p^i \text{ tipificados} \quad (42)$$

conduce a la expresión

$$E^k \delta^0 \propto p^k \quad (43)$$

convergiendo  $p^k$  hacia  $I_1$ , si  $\lambda_1 > \lambda_2 > \dots > \lambda_v > 0$

(1) Anderson, T.W., [2], pág. 301-302.

(2) En la práctica no es necesario resolver el segundo sistema puesto que:  $\mu_i^2 = \frac{1}{\lambda_i} \sum_{12}^{-1} \sum_{21} h_i^2$

Veáse propiedad 1.5.

(3) Esta equivalencia se desprende de las igualdades  $L^T \Sigma_{11} L = I$ ,  $L^T \Sigma_{21} M_1 = \Lambda$ ,  $M_1^T \Sigma_{21} M_1 = J$  y únicamente tiene importancia teórica para probar la convergencia del procedimiento. ( $M = \begin{bmatrix} M_{11} & M_{12} \\ p & p \end{bmatrix}$ )

Similar procedimiento es aplicable en la obtención de  $l_2 \dots l_r$

b) Prueba de la convergencia. (1)

La expresión (43) puede expresarse como igualdad:

$$p^k = K_R E^k \delta_0$$

y teniendo en cuenta, que  $E^k = L \Lambda^{2k} L^{-1}$

$$p^k = K_R \lambda_1^{2k} \left[ L \left( \frac{\Lambda}{\lambda_1} \right)^{2k} L^{-1} \right] \quad (44)$$

Al estar  $p^k$  tipificado:

$$\left( K_R \lambda_1^{2k} \right)^2 \delta_0^T \left[ L \left( \frac{\Lambda}{\lambda_1} \right)^{2k} L^{-1} \right]^T \sum_{11} \left[ L \left( \frac{\Lambda}{\lambda_1} \right)^{2k} L^{-1} \right] = 1 \quad (45)$$

pasando al límite cuando  $k \rightarrow \infty$  y teniendo en cuenta que

si  $\lambda_1 > \lambda_2 > \dots > \lambda_r > 0$

$$\lim_{k \rightarrow \infty} L \left( \frac{\Lambda}{\lambda_1} \right)^{2k} L^{-1} = L \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} L^{-1} = P_1 P_1^T \quad (46)$$

$$\lim_{k \rightarrow \infty} \left( K_R \lambda_1^{2k} \right)^2 \delta_0^T P_1 P_1^T \sum_{11} P_1 P_1^T \delta_0 = \lim_{k \rightarrow \infty} \left( K_R \lambda_1^{2k} \right)^2 \delta_0^T P_1 P_1^T \delta_0 = 1 \quad (47)$$

$$\lim_{k \rightarrow \infty} \left( K_R \lambda_1^{2k} \right)^2 = \frac{1}{P_1^T \delta_0} \quad (48)$$

Hallando el límite en (44) y utilizando las igualdades (46) y

(48) se tiene

$$\lim_{k \rightarrow \infty} p^k = \frac{1}{P_1^T \delta_0} P_1 P_1^T \delta_0 = P_1 \quad \text{c. q. d.}$$

De la misma forma puede probarse la convergencia del proceso de obtención de  $l_2$  teniendo en cuenta la igualdad

$$E_1 = E - \lambda_1^2 P_1 P_1^T = L \Lambda_1^2 L^{-1} : \Lambda_1 = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_r & 0 \\ 0 & 0 & \dots & 0 & b \end{bmatrix} \quad (49)$$

Podría continuarse hasta la etapa  $r$

(1) La demostración de la convergencia del proceso iterativo que aquí exponemos ha sido realizada por el Director de esta Memoria.

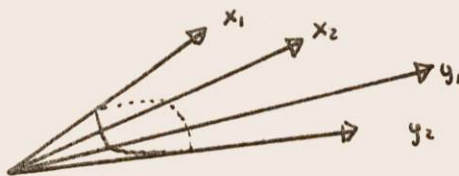


### 1.3.4. Descripción de los vectores canónicos.

El examen de la matriz (4) permite caracterizar los vectores canónicos  $U$  y  $V$ . Dichos vectores tienen las siguientes propiedades:

- 1.- Las variables canónicas  $u_1, u_2, \dots, u_r$  son incorreladas entre sí.
- 2.- Las variables canónicas  $v_1, v_2, \dots, v_r$  también son incorreladas entre sí.
- 3.- Las variables canónicas  $u_i$  y  $v_j$  son incorreladas salvo en el caso  $i = j$ . Cuando ésto es así la correlación entre  $u_i, v_i$  es el elemento diagonal  $\lambda_i$  de  $P$ . Así tendremos que  $\lambda_1$  es la correlación entre  $u_1$  y  $v_1$ ;  $\lambda_2$  es la correlación entre  $u_2$  y  $v_2$  ...  $\lambda_r$  es la correlación entre  $u_r$  y  $v_r$ .

Si en un sistema gráfico representamos las variables iniciales  $Z = (X, Y)$  asociando un vector a cada uno de los componentes en donde el módulo sea su varianza y el coseno del ángulo que forme con las demás componentes las correlaciones con ellas obtenemos la figura<sup>(1)</sup> (consideraremos el caso  $x_1, x_2, y_1, y_2$ )



$$\begin{aligned} \text{Mod } z_i &= \sqrt{\text{Var } z_i} \\ \cos z_i z_j &= \rho_{z_i z_j} \\ z_i &\equiv x_i \\ z_j &\equiv y_j \end{aligned}$$

Fig. 1.

(1) La representación gráfica únicamente pretende dar una idea de la relación entre las nuevas variables, no correspondiéndose al caso real que en general ocurre  $\lambda_1 \neq \lambda_2$  lo cual da lugar a que  $u_1, u_2, v_1, v_2$  únicamente sean representables en  $\mathbb{R}^4$ . Si la distribución no es degenerada, también la primera figura sólo es representable en  $\mathbb{R}^4$ .

Las variables canónicas  $u_1, u_2, v_1, v_2$  tendrán la estructura que se señala en la figura II. <sup>(1)</sup>

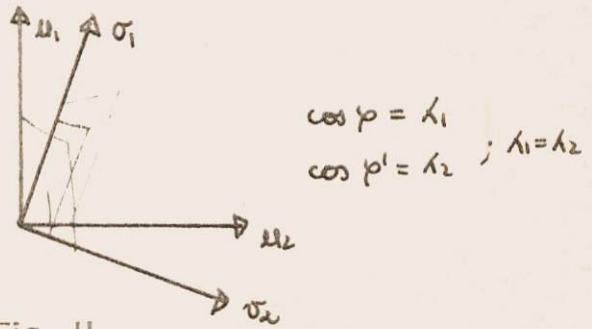


Fig. II

---

(1) Véase nota (1) de la página anterior.

#### 1.4. El modelo canónico no normalizado.

Otra derivación alternativa del modelo canónico es la que sustituye las condiciones de normalización en varianza de las variables canónicas por las de ortogonalidad de las matrices de transformación de cada uno de los grupos. Esta variación respecto a lo ya estudiado es propuesta por Van de Geer<sup>(1)</sup>. Lo desarrollaremos en términos - de la matriz de covarianza de los vectores aleatorios de referencia  $X, Y$ :  $X = (x_1, x_2 \dots x_p)$  y  $Y = (y_1, y_2 \dots y_q)$ , estando tales vectores distribuidos conjuntamente con parámetros  $(0 \Sigma)$ :  $p \leq q$  y  $\text{rang} \Sigma = p+q$ .

Si particionamos la matriz  $\Sigma$  en base a las dimensiones de  $X$  e  $Y$  obtenemos

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

en donde supondremos que  $\text{rang} \Sigma_{12} = r \leq p$ .

Se denominan vectores canónicos a los vectores  $U, V$ :

$$U = L^T X \quad (53)$$

$$V = M^T Y \quad (54)$$

que satisfacen (55), (56), (57) y (58).

$$L^T L = L L^T = I_p \quad (55)$$

$$M^T M = M M^T = I_q \quad (56)$$

$$\Sigma_{00} = \begin{bmatrix} L^T \Sigma_{11} L & L^T \Sigma_{12} M \\ M^T \Sigma_{21} L & M^T \Sigma_{22} M \end{bmatrix} \quad (57)$$

$$P = L^T \Sigma_{12} M. \quad (58)$$

siendo (58) una matriz de la estructura (5).

(1) Van de Geer [50] pág. 169-170.

La prueba de la existencia de tales vectores la basaremos en la descomposición singular de la matriz  $\Sigma_{12}$ .

$$\Sigma_{12} = L P M^T : L^T L = I_p ; M^T M = I_q$$

donde L y M son las soluciones de los sistemas

$$\begin{aligned} [\Sigma_{12} \Sigma_{21} - \lambda^2 I] L &= 0 \\ [\Sigma_{21} \Sigma_{12} - \lambda^2 I] M &= 0 \end{aligned} \quad (59)$$

y  $P = L^T \Sigma_{12} M$  una matriz de la forma (5) cuyos elementos no nulos son la raíz cuadrada con signo positivo de las soluciones en  $\lambda^2$  de los determinantes

$$|\Sigma_{12} \Sigma_{21} - \lambda^2 I| = 0 ; |\Sigma_{21} \Sigma_{12} - \lambda^2 I| = 0$$

Si L y M satisfacen los sistemas anteriores se tiene:

$$L^T \Sigma_{12} M = P ; L^T L = I_p ; M^T M = I_q$$

Podría demostrarse que los vectores canónicos  $U = L^T X$  y  $V = M^T Y$  son de todos los obtenidos por transformación ortogonal de X e Y los que tienen covarianza máxima entre componentes de U y V con igual subíndice y que son incorrelados con los de distinto subíndice (en 1.5.1. realizaremos la demostración análoga a la que aquí señalamos para el modelo definido en 1.2.).

En este modelo el giro de eje realizado es menos simple que el anterior ya que no logra incorrelación en cada uno de los vectores canónicos aunque conserva las varianzas de los vectores de referencia y las correlaciones dentro de cada vector.

Si lo detallamos gráficamente del conjunto inicial  $(x_1, x_2, y_1, y_2)$

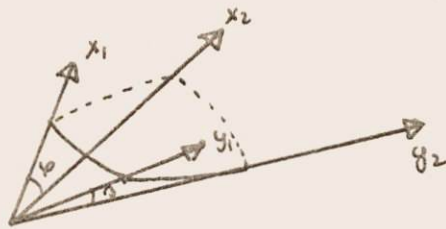


Fig. III

en el modelo canónico no normalizado se pasa a la siguiente estructura

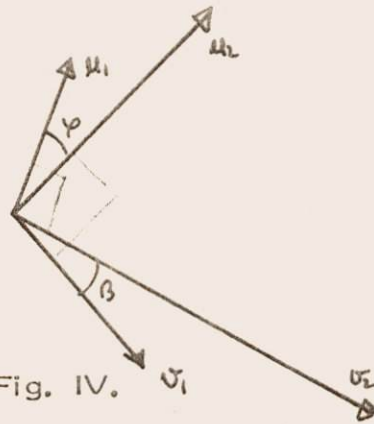


Fig. IV.

De la misma forma que en el modelo definido en 1.2. la representación gráfica que aquí exponemos únicamente pretende dar una idea de la transformación realizada. Una representación en la que  $u_1$  y  $u_2$  conserve el mismo ángulo que  $x_1$  y  $x_2$  (lo mismo para  $v_1$  y  $v_2$  y  $y_1$  y  $y_2$  y que además  $u_1 \perp v_2$ ,  $v_1 \perp u_2$  en general requiere ser representado en  $\mathbb{R}^4$

En la práctica el modelo más difundido ha sido el que hemos expuesto en 1.2. es decir, el de variables canónicas normalizadas en varianza. Nos referiremos únicamente a él en el resto de este trabajo.

### 1.5. Propiedades del modelo canónico.

1.5.1. 1) Las primeras variables canónicas  $u_1, v_1$  definidas en 1.2. son de todas las combinaciones lineales de X e Y normalizadas en varianza las de mayor correlación siendo ésta la primera correlación canónica.

En efecto: a) Sean 
$$u_1 = p_1^T X \quad (60)$$

$$v_1 = w_1^T Y \quad (61)$$

dos combinaciones lineales de X e Y normalizadas en varianza es decir, tales que:

$$E u_1^2 = p_1^T E X X^T p_1 = p_1^T \Sigma_{11} p_1 = 1 \quad (62)$$

$$E v_1^2 = w_1^T E Y Y^T w_1 = w_1^T \Sigma_{22} w_1 = 1 \quad (63)$$

La expresión del coeficiente de correlación entre  $u_1$  y  $v_1$  viene dada por:

$$\rho_{u_1, v_1} = E u_1 v_1 = p_1^T E X Y^T w_1 = p_1^T \Sigma_{12} w_1 \quad (64)$$

ya que  $u_1$  y  $v_1$  son de varianza 1.

Si

$$\psi_1 = p_1^T \Sigma_{12} w_1 - \frac{1}{2} \lambda [p_1^T \Sigma_{11} p_1 - 1] - \frac{1}{2} \mu [w_1^T \Sigma_{22} w_1 - 1] \quad (65)$$

es la función auxiliar de Lagrange para optimizar la función coeficiente de correlación la derivación de tal función da lugar al siguiente sistema de ecuaciones

$$\frac{\partial \psi_1}{\partial p_1} = \Sigma_{12} w_1 - \lambda \Sigma_{11} p_1 = 0 \quad (66)$$

$$\frac{\partial \psi_1}{\partial w_1} = \Sigma_{21} p_1 - \mu \Sigma_{22} w_1 = 0$$

que operando en él:

$$p_1^T \Sigma_{12} w_1 - \lambda p_1^T \Sigma_{11} p_1 = 0$$

$$w_1^T \Sigma_{21} p_1 - \mu w_1^T \Sigma_{22} w_1 = 0$$

cumpléndose

$$(62) \quad (63) \quad \rho_1^T \Sigma_{12} \omega_1 = \omega_1^T \Sigma_{21} \rho_1 = \lambda = \mu.$$

con lo cual las condiciones de óptimo de (66) se reducen a los siguientes sistemas:

$$\begin{aligned} -\lambda \Sigma_{11} \rho_1 + \Sigma_{12} \omega_1 &= 0 \\ \Sigma_{21} \rho_1 - \lambda \Sigma_{22} \omega_1 &= 0 \end{aligned} \quad (67)$$

cuya solución no trivial exige

$$\begin{vmatrix} -\lambda \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda \Sigma_{22} \end{vmatrix} = 0 \quad (68)$$

Como  $\lambda = \mu = \rho_1^T \Sigma_{12} \omega_1$ , utilizaremos la mayor raíz de (68) puesto que  $\rho_1^T \Sigma_{12} \omega_1$  es la función a maximizar.

b) Vamos a probar que esa raíz máxima es la mayor correlación canónica entre X e Y y que las soluciones  $(l_1, m_1)$  correspondientes en el sistema (67) son los primeros vectores de coeficientes canónicos.

b-1) Para probar la primera parte utilizaremos la siguiente propiedad de los determinantes de matrices particionadas <sup>(1)</sup>

Si:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; \quad |A| = \left| A_{11} - A_{12} A_{22}^{-1} A_{21} \right| \left| A_{22} \right|$$

(1) Esta propiedad está probada en Anderson, T. W. [2] pág. 343 y 344 en base a una matriz  $B = \begin{bmatrix} I & A_{12} A_{22}^{-1} \\ 0 & I \end{bmatrix}$  cumpliéndose  $|A| = |BAB^T| = \left| A_{11} - A_{12} A_{22}^{-1} A_{21} \right| \left| A_{22} \right|$ .

Aplicando esta propiedad a (68):

$$\begin{vmatrix} -\lambda \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda \Sigma_{22} \end{vmatrix} = \left| -\lambda \Sigma_{11} - \Sigma_{12} [-\lambda \Sigma_{22}^{-1}] \Sigma_{21} \right| \left| -\lambda \Sigma_{22} \right|$$

Por ser  $\Sigma_{22}$  definida positiva,  $|\Sigma_{22}| \neq 0$  con lo cual

$$\left| -\lambda^2 \Sigma_{11} + \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right| = 0$$

Este determinante coincide con el del sistema (21) por lo tanto mayor raíz de (68) es la primera correlación canónica entre X e Y (recuérdese  $\lambda_1 > \lambda_2 > \dots > \lambda_n > 0$ ).

b-2) Probaremos que  $l_1$  y  $m_1$  soluciones asociadas en (67) a la mayor raíz  $\lambda_1$  de (68) son los vectores de coeficientes de las primeras variables canónicas.

Considerando dicha raíz  $\lambda_1$  el sistema (67) puede expresarse

$$\Sigma_{12} m_1 - \lambda_1 \Sigma_{11} l_1 = 0 \quad (69)$$

$$\Sigma_{21} l_1 - \lambda_1 \Sigma_{22} m_1 = 0 \quad (70)$$

Una sencilla transformación en (69) (70) permite pasar a los sistemas (21) y (22) cuyas primeras columnas de las matrices L y M son los vectores de coeficientes canónicos asociados a las primeras variables canónicas. Premultiplicando (70) por  $\Sigma_{22}^{-1}$  obtenemos

$$\Sigma_{22}^{-1} \Sigma_{21} l_1 - \lambda_1 m_1 = 0 \quad (71)$$

premultiplicando esta expresión por  $\Sigma_{12}$  llegamos a:

$$\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} l_1 - \lambda_1 \Sigma_{12} m_1 = 0 \quad (72)$$

y finalmente sustituyendo (72) en (69) da lugar a

$$(21) \quad \sum_{12} \sum_{12}^{-1} \sum_{21} p_1 - \lambda_1^2 \sum_{11} p_1 = 0$$

expresión que se corresponde con el sistema (21) en el que habíamos definido los vectores canónicos.

De forma totalmente paralela escogiendo convenientemente la matriz B y transformando los sistemas (69) y (70) podría probarse que  $\lambda_1^2$  es también la mayor raíz del sistema (22) y  $m_1$  su vector de coeficientes canónicos.

II) Las segundas variables canónicas  $u_2, v_2$  son de todas las combinaciones lineales de X e Y normalizadas en varianza e incorreladas con las primeras variables canónicas las que tienen correlación máxima, siendo ésta la segunda correlación canónica.

En efecto: Sean

$$u_2 = l_2^T X$$

$$v_2 = m_2^T Y$$

dos variables obtenidas por combinación lineal de X y de Y y que supondremos que son de varianza 1 e incorreladas con  $u_1$  y  $v_1$ ; es decir, cumplen:

$$\begin{aligned} l_2^T \Sigma_{11} l_2 &= 1 & ; & & l_2^T \Sigma_{11} l_1 &= 0 \\ m_2^T \Sigma_{22} m_2 &= 1 & ; & & m_2^T \Sigma_{22} m_1 &= 0 \end{aligned} \quad (73)$$

Si

$$\mathcal{L}_2 = l_2^T \Sigma_{12} m_2 - \frac{1}{2} \lambda [l_2^T \Sigma_{11} l_2 - 1] - \frac{1}{2} \mu [m_2^T \Sigma_{22} m_2 - 1] + \nu [l_2^T \Sigma_{11} l_1] + \theta [m_2^T \Sigma_{22} m_1]$$

es la función auxiliar de Lagrange para la optimización de la función coeficiente de correlación; la derivación de esta función conduce a

$$\frac{\partial \mathcal{K}_2}{\partial \rho_2} = \sum_{12} \omega_2 - \lambda \sum_{11} \rho_2 + \nu \sum_{11} \rho_1 = 0 \quad (74)$$

$$\frac{\partial \mathcal{K}_2}{\partial \omega_1} = \sum_{21} \rho_2 - \mu \sum_{22} \omega_2 + \theta \sum_{22} \omega_1 = 0 \quad (75)$$

En la función a optimizar no están incluidas las funciones  $u_1$  incorrelada con  $v_2$  y  $v_1$  incorrelada con  $u_2$ ; las tendremos en cuenta al operar <sup>(1)</sup> en las ecuaciones (74) y (75). Premultiplicando tales ecuaciones por  $l_1^T$  y  $m_1^T$ :

$$l_1^T \sum_{12} \omega_2 - \lambda l_1^T \sum_{11} \rho_2 + \nu l_1^T \sum_{11} \rho_1 = 0 \quad (76)$$

$$m_1^T \sum_{21} \rho_2 - \mu m_1^T \sum_{22} \omega_2 + \theta m_1^T \sum_{22} \omega_1 = 0 \quad (77)$$

e incluyendo las condiciones (62), (63) y (73) se tiene

$$\begin{aligned} \nu \cdot 1 &= 0 \\ \theta \cdot 1 &= 0 \end{aligned} \Rightarrow \nu = \theta = 0$$

con lo cual el sistema (74) y (75) se transforma en

$$\begin{aligned} \sum_{12} \omega_2 - \lambda \sum_{11} \rho_2 &= 0 \\ \sum_{21} \rho_2 - \lambda \sum_{22} \omega_1 &= 0 \end{aligned}$$

de donde se sigue que  $\lambda = m_1^T \sum_{21} \rho_2$

Este sistema es análogo a (67) y su solución con las restricciones del enunciado son los vectores de coeficientes  $(l_2, m_2)$  asociados a la segunda mayor raíz de (68) puesto que  $m_2^T \sum_{21} \rho_2$  es la función a maximizar.

III) Probada la equivalencia entre los sistemas (67) y (74), y (75) podría continuarse hasta la etapa  $r$  supuesto que es el rango de la matriz  $\sum_{12}$ ; además en base a esta equivalencia queda justificado que (68) tiene  $2r$  raíces;  $r$  de la forma  $\lambda_i^+$ ;  $i=1, 2, \dots, r$  y  $r$   $-\lambda_i^-$ ;  $i=1, 2, \dots, r$

- (1) El máximo de  $\mathcal{K}_2$  y de esta misma función con las condiciones adicionales  $l_1^T \sum_{12} \omega_2$  y  $m_1^T \sum_{21} \rho_2$  coinciden. Para ello sólo es preciso probar que las soluciones obtenidas al optimizar  $\mathcal{K}_2$  también satisfacen las nuevas restricciones. Sean  $\rho_2^0$  y  $\omega_2^0$  los vectores que hacen óptima a  $\mathcal{K}_2$ ; dichos vectores cumplen:

$$\begin{aligned} \sum_{12} \omega_2^0 - \lambda \sum_{11} \rho_2^0 &= 0 \\ \sum_{21} \rho_2^0 - \lambda \sum_{22} \omega_2^0 &= 0 \end{aligned}$$

de donde se deduce:  $l_1^T \sum_{12} \omega_2^0 - \lambda l_1^T \sum_{11} \rho_2^0 = 0 \Rightarrow l_1^T \sum_{12} \omega_2^0 = 0$   
 $m_1^T \sum_{21} \rho_2^0 - \lambda m_1^T \sum_{22} \omega_2^0 = 0 \Rightarrow m_1^T \sum_{21} \rho_2^0 = 0$  c.q.d.

La mayor parte de los autores presentan el modelo canónico a través de las propiedades anteriores, citaremos entre otros: Cooley W. <sup>(1)</sup>, Anderson T. W. <sup>(2)</sup>, Morrison D. <sup>(3)</sup>, Joshtons J. <sup>(4)</sup>; otros como P. Horst <sup>(5)</sup> y Kshirsagar A. <sup>(6)</sup>, lo consideran como una propiedad de las variables definidas en (2), en cualquier caso ésta no es la única propiedad a partir de la cual pueden obtenerse las variables y coeficientes de correlación canónicos tal como expondremos a continuación.

1.5.2. De todas las combinaciones lineales de X e Y normalizadas en varianza, las primeras variables canónicas  $U_1, V_1$  son aquellas que tienen una matriz de covarianza que mejor se aproxima en el sentido minimocuadrático a una matriz de correlaciones de rango 1.

Sea  $\Sigma_{u,v}$  la matriz de correlaciones entre  $u = l_1^T x$  y  $v = m^T y$   
 :  $l_1^T \Sigma_{11} l_1 = 1$  y  $m^T \Sigma_{22} m = 1$

$$\Sigma_{u,v} = \begin{bmatrix} 1 & l_1^T \Sigma_{12} m \\ m^T \Sigma_{21} l_1 & 1 \end{bmatrix}$$

---

(1) Cooley [ 8 ] pág. 168.

(2) Anderson T. W. [ 2 ] pág. 272-287.

(3) Morrison, D. [33] pág. 213-218.

(4) Joshton J. [ 25 ] pág. 352-355.

(5) Horsts P. [ 14 ] pág. 130-132.

(6) Kshirsagar [ 28 ] pág. 247-285.

Cuando  $l_1$  y  $m_1$  son los coeficientes de las primeras variables canónicas,  $l_1^T \Sigma_{12} u_1 = u_1^T \Sigma_{21} l_1 = \lambda_1$ ; en donde  $\lambda_1$  es la mayor correlación canónica.

$$\text{Luego } \Sigma_{uv} = \begin{bmatrix} 1 & \lambda_1 \\ \lambda_1 & 1 \end{bmatrix}$$

Por la propiedad 1.5.1. cualquier otras variables  $u'_i, v'_i$  normalizadas en varianza tendrán un coeficiente  $\lambda'_i$  igual o menor.

Por tanto  $2(1-\lambda_1)^2 \leq 2(1-\lambda'_i)^2$  con lo cual queda probada la propiedad<sup>(1)</sup>.

Razonando de forma inversa podrían obtenerse las variables y correlaciones canónicas en base a exigir que  $l_1$  y  $m_1$  sean de tal forma que las variables  $u_1 = l_1^T X$  y  $v_1 = m_1^T Y$  tengan una matriz de correlaciones que mejor se aproxime a una matriz de rango 1, según el criterio mínimo cuadrático.

Igualmente podrían obtenerse el resto de variables canónicas exigiendo que las variables  $u_2 = l_2^T X$  y  $v_2 = m_2^T Y$  además de ser incorreladas con  $u_1$  y  $v_1$  ya obtenidas tengan una matriz de correlaciones que mejor se aproxime a una matriz de rango 1 en el sentido mínimo cuadrático. El proceso puede continuarse hasta la etapa  $r$ .

---

(1) Igualmente las segundas variables canónicas tienen una matriz de correlaciones que de todas las combinaciones lineales normalizadas en varianza y ortogonales a  $u_1, v_1$ , es la que mejor se aproxima a una matriz de rango 1 en el sentido mínimo cuadrático.

Puede continuarse hasta  $r = \text{ran } \Sigma_{12}$

1.5.3. 1) De todas las combinaciones lineales de X e Y, normalizadas en varianza, las primeras variables canónicas son las que tienen una matriz de correlaciones cuya suma de elementos no diagonales es máxima.

Sean  $u_1 = \rho_1^T X$  y  $v_1 = \omega_1^T Y$  dos combinaciones lineales de X e Y normalizadas en varianza y sea  $\Sigma_{u,v}$  su matriz de correlaciones

$$(9), (10) \quad \Sigma_{u,v} = \begin{bmatrix} 1 & \rho_1^T \Sigma_{12} \omega_1 \\ \omega_1^T \Sigma_{21} \rho_1 & 1 \end{bmatrix} \quad (23), (24) \quad \begin{bmatrix} 1 & a_1^T t_1^T \Sigma_{12} (t_2^T)^{-1} b_1 \\ b_1^T t_2^T \Sigma_{21} (t_1^T)^{-1} a_1 & 1 \end{bmatrix}$$

La suma de los elementos no diagonales de  $\Sigma_{u,v}$  puede expresarse<sup>(1)</sup>

$$\psi = [1, 1] \begin{bmatrix} a_1^T & 0 \\ 0 & b_1^T \end{bmatrix} \begin{bmatrix} 1 & t_1^T \Sigma_{12} (t_2^T)^{-1} \\ t_2^T \Sigma_{21} (t_1^T)^{-1} & 1 \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} - [1, 1] \begin{bmatrix} a_1^T & 0 \\ 0 & b_1^T \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

que en forma abreviada la denotaremos:

$$\psi = 1^T D^T R D 1 - 1^T D^T D 1$$

en donde

$$D = \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} \quad ; \quad R = \begin{bmatrix} 1 & t_1^T \Sigma_{12} (t_2^T)^{-1} \\ t_2^T \Sigma_{21} (t_1^T)^{-1} & 1 \end{bmatrix}$$

Si añadimos la condición de ortogonalidad a la función anterior y escribimos  $P = R - I$ , tenemos la función a derivar

$$\Theta = 1^T D^T P D 1 - 1^T D^T D 1 \quad ; \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

Efectuando la derivación respecto a  $1^T D$

$$\frac{\partial \Theta}{\partial 1^T D} = P D - D \lambda = 0 \quad (78)$$

$$(1) \quad [1 \ 1] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = a+b+c+d.$$

Ecuación que desarrollándola se transforma en

$$\begin{aligned} t_1^{-1} \Sigma_{12} (t_2^T)^{-1} b_1 &= a_1 \lambda_1 & ; & & a_1^T t_1^{-1} \Sigma_{12} (t_2^T)^{-1} b_1 &= \lambda_1 \\ t_2^{-1} \Sigma_{21} (t_1^T)^{-1} a_1 &= b_1 \lambda_2 & ; & & b_1^T t_2^{-1} \Sigma_{21} (t_1^T)^{-1} a_1 &= \lambda_2 \end{aligned} \Rightarrow \lambda_1 = \lambda_2$$

El sistema (78) teniendo en cuenta que  $\Sigma_{11} = t_1 t_1^T$  y  $\Sigma_{22} = t_2 t_2^T$  se transforma en el sistema (17) y (18) y seguidamente en el (21) y (22) siendo

$$(23) \quad l_1 = (t_1^{-1})^T a_1$$

$$(24) \quad m_1 = (t_2^{-1})^T b_1$$

los coeficientes de las primeras variables canónicas; y  $\lambda_1 = \lambda_2$  la mayor correlación canónica entre las variables  $X$  e  $Y$ .

De forma análoga (con las restricciones a añadir en su caso) se probaría que la segunda correlación canónica entre  $X$  e  $Y$  es la que hace máxima la suma de los elementos no diagonales de la matriz de la forma  $\Sigma_{u,v}$ , siendo los vectores  $a$  y  $b$  que la integran ortogonales a los anteriores.

El procedimiento puede seguirse hasta la etapa  $r = \text{rang } \Sigma_{12}$ .

También esta propiedad puede tomarse como procedimiento de obtención de las variables canónicas.

Haremos un breve comentario en relación a las propiedades 1.5.1., 1.5.2., y 1.5.3.

Estas propiedades ciertamente constituyen la expresión de una misma idea siendo triviales las demás cuando se ha probado una de

---

(1) Horst P. [ 19 ] pág. 129-149 y [ 20 ] pág. 331-347.

ellas debido a que operamos con matrices de correlaciones y por tanto simétricas.

En esta exposición han sido mostradas como propiedades diferentes debido a que en la generalización del modelo canónico, 1.5.2. y 1.5.3. presentan dos criterios de obtención de los vectores canónicos generalizados. Estas cuestiones serán expuestas con más detalle en el capítulo III.

1.5.4. Si  $u_1$  y  $v_1$  son las primeras variables canónicas las regresiones de  $u_1$  sobre  $Y$  y de  $v_1$  sobre  $X$  tienen unos coeficientes de regresión que son proporcionales a los vectores canónicos  $m_1$  y  $l_1$  (respectivamente) asociados a tales variables siendo el factor de proporcionalidad la primera correlación canónica.

Prueba:

$$\text{Sea } u_1 = p_{11}x_1 + p_{21}x_2 + \dots + p_{p1}x_p = X^T p_1$$

La regresión de  $u_1$  sobre  $y_1, y_2, \dots, y_q$  conduce a los siguientes resultados:

Si denominamos <sup>(1)</sup>

$$u^* = b_{11}y_1 + b_{21}y_2 + \dots + b_{q1}y_q = Y^T b_1$$

al hiperplano de regresión, sus coeficientes son los que hacen mínima la función

$$\begin{aligned} E(u - u^*)^2 &= E(X^T p_1 - Y^T b_1)^T (X^T p_1 - Y^T b_1) \\ &= p_1^T \Sigma_{11} p_1 - 2b_1^T \Sigma_{12} p_1 + b_1^T \Sigma_{11} b_1 \end{aligned}$$

(1) En la notación empleada en Fz. Trocóniz, A. [12] pág. 16.5.  $u^*$  tiene el carácter de estimación lineal de  $u$  por  $y_1, y_2, \dots, y_q$ .

Al anular la primera derivada respecto a  $b_1$  se obtiene

$$\frac{\partial E(u-u^*)^2}{\partial b_1^T} = -2 \sum_{12} l_1 + 2 \sum_{22} b_1 = 0$$

$$; b_1 = \sum_{22}^{-1} \sum_{21} l_1 \quad (79)$$

De forma análoga si  $\sigma^* = c_{11}x_1 + c_{21}x_2 + \dots + c_{p1}x_p = X^T C_1$  la regresión de  $v_1$  sobre  $x_1, x_2, \dots, x_p$  conduce a:

$$C_1 = \sum_{11}^{-1} \sum_{12} u_1 \quad (80)$$

Al ser  $l_1$  y  $m_1$  los primeros vectores de coeficientes canónicos; por 1.5.1. satisfacen las ecuaciones

$$\sum_{12} u_1 = \lambda_1 \sum_{11} l_1$$

$$\sum_{21} l_1 = \lambda_1 \sum_{22} u_1$$

Estas igualdades nos llevan a los siguientes resultados para los coeficientes de regresión

$$b_1 = \sum_{22}^{-1} \sum_{21} l_1 = \lambda_1 \sum_{22}^{-1} \sum_{22} u_1 = \lambda_1 u_1$$

$$c_1 = \sum_{11}^{-1} \sum_{12} u_1 = \lambda_1 \sum_{11}^{-1} \sum_{11} l_1 = \lambda_1 l_1 \quad \text{c. q. d.}$$

Otro tanto ocurre con las segundas variables canónicas  $u_2$  y  $v_2$ . Si se considera la regresión de  $u_2$  sobre  $y_1, y_2, \dots, y_q$  y  $v_2$  sobre  $x_1, x_2, \dots, x_p$  cuando  $u_2$  y  $v_2$  sean las segundas variables canónicas los coeficientes son proporcionales a los vectores canónicos  $m_2$  y  $l_2$  respectivamente, siendo el factor de proporcionalidad la segunda correlación canónica,  $\lambda_2$

Pueden continuarse las regresiones hasta la etapa  $r = \text{rang } \Sigma_{12}$ .

1.5.5. Si  $u_1 = \rho_1^T X$  y  $v_1 = \mu_1^T Y$  son combinaciones de  $X$  e  $Y$ , normalizadas en varianza, las regresiones de  $u_1$  sobre  $Y$  y de  $v_1$  sobre  $X$  tienen un error de regresión mínimo cuando  $l_1$  y  $m_1$  son los primeros vectores de coeficientes canónicos.

En efecto. El error en ambas expresiones viene dado por el mínimo de las respectivas funciones optimizadas en 1.5.4.

$$E (u_1 - u_1^*)^2 = \rho_1^T \Sigma_{11} \rho_1 - 2 \rho_1^T \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} + \rho_1^T \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{22} \Sigma_{22}^{-1} \Sigma_{21} \rho_1 \\ = 1 - \rho_1^T \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

$$E (\sigma_1 - \sigma_1^*)^2 = \mu_1^T \Sigma_{22} \mu_1 - 2 \mu_1^T \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} + \mu_1^T \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{11} \Sigma_{11}^{-1} \Sigma_{12} \mu_1 \\ = 1 - \mu_1^T \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$$

Si  $l_1$  y  $m_1$  son los primeros vectores de coeficientes canónicos satisfacen las ecuaciones:

$$(69) \quad \Sigma_{12} \mu_1 = \lambda_1 \Sigma_{11} \rho_1$$

$$(70) \quad \Sigma_{21} \rho_1 = \lambda_1 \Sigma_{22} \mu_1$$

de las cuales pueden obtenerse:

$$\rho_1^T \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \rho_1 = \lambda_1^2 \mu_1^T \Sigma_{22} \Sigma_{22}^{-1} \Sigma_{22} \mu_1 = \lambda_1^2$$

$$\mu_1^T \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \mu_1 = \lambda_1^2 \rho_1^T \Sigma_{11} \Sigma_{11}^{-1} \Sigma_{11} \rho_1 = \lambda_1^2$$

y por tanto

$$E (u_1 - u_1^*)^2 = 1 - \lambda_1^2$$

$$E (\sigma_1 - \sigma_1^*)^2 = 1 - \lambda_1^2$$

Por 1.5.1. sabemos que  $\lambda_1$  es el coeficiente de correlación entre  $u_1$  y  $v_1$ , que es máximo cuando dichas variables son las primeras canónicas. Por tanto  $E (u_1 - u_1^*)^2 = E (\sigma_1 - \sigma_1^*)^2$  es mínimo si  $l_1$  y  $m_1$  satisfacen (69) y (70)

Igual propiedad cumplen el segundo par de regresiones, en relación al error de regresión, pudiéndose continuar el proceso hasta la etapa  $r = \text{rang. } \Sigma_{12}$

1.5.6. Si alguno de los vectores de referencia  $X, Y$  consta de un único elemento entonces existe un solo coeficiente de correlación canónica siendo éste, el coeficiente de correlación múltiple en la regresión del vector con un único elemento sobre el otro.

Consideremos el caso en que el vector  $X$  consta de un único elemento. Sea éste  $x_1$ ; la matriz de covarianzas de  $(x_1, y_1, \dots, y_q)$  particionada adopta la forma

$$\Sigma = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1q+1} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{q+1,1} & \lambda_{q+1,2} & \dots & \lambda_{q+1,q+1} \end{bmatrix} \quad ; \quad \Sigma_{12} = [\lambda_{12} \dots \lambda_{1q+1}]$$

$$; \quad \text{var } \Sigma_{12} = 1$$

con lo cual el determinante del sistema (25):  $|\lambda^2 \Sigma_{11} + \Sigma_{12} \Sigma_{12}^{-1} \Sigma_{21}| = 0$  se transforma en:  $|\lambda^2 \lambda_{11} + \Sigma_{12} \Sigma_{12}^{-1} \Sigma_{21}| = 0$

que por ser un escalar

$$\lambda^2 = \frac{\Sigma_{12} \Sigma_{12}^{-1} \Sigma_{21}}{\lambda_{11}} \quad (81)$$

Una expresión del coeficiente de correlación múltiple <sup>(1)</sup> vendría dada por

$$\lambda = \frac{\text{cov}(x_1, x_1^*)}{\sqrt{\text{var } x_1 \cdot \text{var } x_1^*}} \quad (82)$$

Calculando el numerador y denominador de (82), se obtiene:

$$\text{cov}(x_1, x_1^*) = \lambda_{11} - \frac{A}{A_{11}} \quad ; \quad \text{var } x_1^* = \lambda_{11} - \frac{A}{A_{11}} \quad ; \quad \text{var } x_1 = \lambda_{11}$$

en donde  $A$  es el determinante de  $\Sigma$  y  $A_{11}$  el adjunto del elemento  $\lambda_{11}$

Llevadas estas expresiones a (82) y elevando al cuadrado el resultado, se obtiene (81) cuadrado de la única correlación canónica existente.

(1) Véase Fz. de Trocóniz, A. [12], pág. 16.5.

1.5.7. Si alguno de los vectores de referencia  $X, Y$  consta de un único elemento, la regresión de este elemento con el otro vector de referencia tiene unos coeficientes de regresión proporcionales al vector de coeficientes canónicos de la variable canónica correspondiente.

$$\text{Sea } X = x_1 \quad ; \quad Y = (y_1 y_2 \dots y_p)$$

Las primeras variables canónicas  $u_1$  y  $v_1$  son de la forma

$$u_1 = l_1 x_1 \\ v_1 = m_{11} y_1 + m_{12} y_2 + \dots + m_{1q} y_q = Y^T m_1$$

siendo  $l_1$  y  $m_1$

$$(23) \quad l_1 = (t_1^{-1}) a \quad : \quad a^T a = 1$$

$$(24) \quad m_1 = (t_1^{-1}) b \quad : \quad b^T b = 1$$

Como en este caso  $\Sigma_{11}$  se reduce a un único elemento, la descomposición en el producto de dos matrices triangulares da lugar a :

$$\lambda_{11} = t_1 t_1^T \quad ; \quad t_1 = \sqrt{\lambda_{11}} \quad - \quad l_1 = (\sqrt{\lambda_{11}})^{-1} a.$$

utilizando ( 23 )

$$a^T a = l_1^T \lambda_{11}^{1/2} \lambda_{11}^{1/2} l_1 = \lambda_{11} l_1^T l_1 = 1 \quad ; \quad l_1 = \frac{1}{\sqrt{\lambda_{11}}} \quad ; \quad a = \frac{1}{\sqrt{\lambda_{11}}} x_1$$

y por la propiedad 1.5.4.

$$b = \lambda_1 m_1 \quad ( 83 )$$

en donde  $b$  son los coeficientes de regresión de  $u_1$  sobre  $y_1 y_2 \dots y_q$ . Si en lugar de la regresión anterior se efectúa la de  $u \sqrt{\lambda_{11}}$  sobre  $y_1 y_2 \dots y_q$  el resultado ( 83 ) se transforma en el siguiente

$$b = \lambda_1 \cdot \sqrt{\lambda_{11}} m_1$$

$$m_1 = \frac{b}{\lambda_1 \cdot \sqrt{\lambda_{11}}} \quad \text{c. q. d.} \quad ( 84 )$$

De forma totalmente paralela se probaría la equivalencia para  $X = (x_1 x_2 \dots x_p)$  e  $Y = y_1$  sin más que una reordenación de la

matriz de covarianzas de  $(x_1 x_2 \dots x_p y_1)$  refiriéndola al caso en que las variables estén en la forma  $(y_1 x_1 \dots x_p)$ .

1.5.8. Las correlaciones canónicas son invariantes ante transformaciones lineales no singulares de los vectores de referencia.

Sean  $c = CX$ :  $|c| \neq 0$

$d = DY$ :  $|d| \neq 0$

La matriz de covarianzas particionada de los vectores  $c, d$  puede expresarse

$$\Sigma_{cd} = \begin{bmatrix} C \Sigma_{11} C^T & C \Sigma_{12} D^T \\ \hline D \Sigma_{21} C^T & D \Sigma_{22} D \end{bmatrix}$$

Las correlaciones canónicas satisfacen los determinantes igualados a cero de los sistemas (21) y (22)

Tomando uno de ellos

$$(21) \quad \left| \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right| = 0$$

Las correlaciones canónicas de los vectores  $c$  y  $d$  satisfacerán

$$\begin{aligned} & \left| C \Sigma_{12} D^T [D \Sigma_{22} D^T]^{-1} D \Sigma_{21} C^T - \lambda^2 C \Sigma_{11} C^T \right| = 0 \\ & |c| \left| \Sigma_{12} D^T (D^T)^{-1} \Sigma_{22}^{-1} D \Sigma_{21} - \lambda^2 \Sigma_{11} \right| |c^T| = 0 \end{aligned}$$

y como  $C$  es no singular

$$\left| \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right| = 0 \quad \text{c. q. d.}$$

1.5.9. Los vectores canónicos de los vectores de variables  $c$  y  $d$  transformados lineales no singulares de  $X$  e  $Y$  tienen unos coeficientes  $(C)^{-1}L$  y  $(D)^{-1}M$  siendo  $L$  y  $M$  los coeficientes canónicos correspondientes a  $X$  e  $Y$ .

Si  $L$  y  $M$  son los coeficientes canónicos de  $X$  e  $Y$  satisfacen los sistemas

$$\left[ \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right] L = 0$$

$$\left[ \Sigma_{31} \Sigma_{11}^{-1} \Sigma_{12} - \lambda^2 \Sigma_{22} \right] M = 0$$

Teniendo en cuenta su matriz de covarianzas los vectores canónicos de  $c$  y  $d$  tendrán unos coeficientes que cumplirán los sistemas

$$\left[ -\lambda^2 C \Sigma_{11} C^T + C \Sigma_{12} D^T (D \Sigma_{22} D^T)^{-1} D \Sigma_{21} C^T \right] R = 0$$

$$\left[ -\lambda^2 D \Sigma_{22} D^T + D \Sigma_{21} C^T (C \Sigma_{11} C^T)^{-1} C \Sigma_{12} D^T \right] S = 0$$

Sistemas que pueden ser expresados

$$\left[ -\lambda^2 \Sigma_{11} + \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \right] C^T R = 0$$

$$\left[ -\lambda^2 \Sigma_{22} + \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right] D^T S = 0$$

Con lo cual

$$L = C^T R$$

$$R = (C^T)^{-1} L$$

$$M = D^T S$$

$$S = (D^T)^{-1} M \quad \text{c.q.d.}$$

1.5.10. Las correlaciones canónicas son invariantes frente a cambios en la unidad de medida de los vectores de referencia.

Utilizaremos el mismo procedimiento que sigue Drhymes <sup>(1)</sup> en el modelo de componentes principales para probar la dependencia de la unidad de medida de las raíces características. (en este caso concreto independencia).

Sea  $X^*$   $Y^*$  los nuevos vectores resultantes de cambiar la unidad de medida de los iniciales  $X$   $Y$ .

$$\begin{bmatrix} X^* \\ Y^* \end{bmatrix} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \quad ; \quad D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

$$X^* = D_1 X \quad ; \quad Y^* = D_2 Y \quad (85)$$

Como  $D$  es una matriz diagonal de elementos no nulos

$$|D| \neq 0 \quad \Rightarrow \quad |D_1| \neq 0 \quad \wedge \quad |D_2| \neq 0$$

con lo cual las transformaciones (85) son transformaciones lineales no singulares y según 1.5.8. conservan las correlaciones canónicas.

Teniendo en cuenta 1.5.9. los vectores canónicos de  $X^*$   $Y^*$  vienen dados por

$$u^* = D_1^{-1} L x$$

$$v^* = D_2^{-1} M y$$

---

(1) Drhymes [11] pág. 58. Este autor recoge la relación entre la unidad de medida entre las distintas variables en una matriz diagonal  $D$  que por su definición es de elementos no nulos.

### 1.6. Relajación de hipótesis del modelo

1.6.1. Las correlaciones canónicas no satisfacen  $\lambda_1 > \lambda_2 \dots > \lambda_r > 0$

Si la ecuación determinante  $|\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \lambda^2\Sigma_{11}| = 0$  posee  $t \leq r$  raíces múltiples  $\lambda^2 = \lambda_t^2$  (que sin pérdida de generalidad supondremos que son las últimas; las  $t$  parejas de variables canónicas asociadas a tales raíces están unívocamente determinadas excepto por la multiplicación por la derecha por sendas matrices ortogonales.

La prueba de la afirmación anterior requiere demostrar que si  $L$  y  $M$  son los vectores de coeficientes canónicos correspondientes a tales variables;  $LC$  y  $MD : C^T C = I; D^T D = I$  también son soluciones de los sistemas

$$\begin{aligned} [\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \lambda_t^2\Sigma_{11}]L &= 0 \\ [+ \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \lambda_t^2\Sigma_{22}]M &= 0 \end{aligned}$$

Utilizaremos el siguiente teorema <sup>(1)</sup>: Si  $S_{p,p}$  es una matriz definida positiva y simétrica con  $\nu$  soluciones iguales de  $|S - \lambda I| = 0$   $\lambda_{r+1} = \dots = \lambda_{r+\nu} = \lambda^*$ ;  $[S - \lambda^* I]$  es de rango  $p - \nu$  y sus vectores asociados están unívocamente determinados excepto por la postmultiplicación por una matriz ortogonal.

Si  $S = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$  tiene las  $t$  últimas raíces características iguales a  $\lambda^c$  por el teorema anterior los vectores  $E^*$  del sistema

$$[\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2} - \lambda^c I]E^* = 0$$

están unívocamente determinados salvo la postmultiplicación por una matriz ortogonal. Sea  $C$  tal matriz; por (34) y (38) los vectores canónicos correspondientes a las  $r$  raíces iguales satisfacen

$$L^* = \Sigma_{11}^{-1/2} E^* C$$

(1) Anderson, T.W., [ 2 ] página 277.

De la misma forma se probaría:  $M^* = \Sigma_{22}^{-1/2} F^* D$

1.6.2. Los vectores de referencia  $X$  e  $Y$  tienen una distribución conjunta degenerada ( $\text{ran } \Sigma < p+q$ )

Si  $\text{ran } \Sigma < p+q$  no puede evaluarse en forma general la influencia de esta hipótesis en el rango de  $\Sigma_{12}$  pero sí en el de las matrices  $\Sigma_{11}$  y  $\Sigma_{22}$  ya que alguna de ellas será singular y por tanto no existen las correlaciones y vectores canónicos en el sentido que hasta aquí se ha dado.

C. G. Khatri<sup>(1)</sup> ha desarrollado un procedimiento de obtención de las correlaciones canónicas en distribuciones degeneradas en base al concepto de inversa generalizada de una matriz<sup>(2)</sup>. Si  $\Sigma$  es singular y  $\Sigma^-$  es la g-inversa de  $\Sigma$  las correlaciones canónicas entre  $X$  e  $Y$  vienen dadas por las soluciones en  $\lambda$  de la ecuación

$$\begin{aligned} \text{tr}_p \left[ \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right] &= 0 \\ = \text{tr}_p \left[ \Sigma_{12} \Sigma_{11}^+ \Sigma_{21} - \lambda^2 \Sigma_{22} \right] &= 0 \end{aligned}$$

(1) C. G. Khatri [ 27 ] pág. 465-471.

(2) Si  $A$  es una matriz de orden  $m \times b$ :  $\text{rg } A = k \leq \min(m, b)$  se dice que  $A^+$  es la inversa generalizada de  $A$  si 1)  $AA^+A = A$   
2)  $A^+AA^+ = A^+$ ; 3)  $(AA^+)^+ = (AA^+)$ ; 4)  $(A^+A)^+ = A^+A$ ; a este tipo de inversa aunque no existe unidad en la terminología, en la última referencia bibliográfica que hemos podido consultar, Green P. [ 16 ] pág. 330 se le denomina Moore-Penrose inversa. Otro tipo de inversa generalizada es la propuesta por Rao, R. [ 41 ] pág. 16; que suele denotarse  $A^-$ . A esta matriz se le denomina g-inversa y su relación con  $A$  es:  $AA^-A = A$ . En general  $A^-$  no es única pero  $A^+$  sí. El cálculo de  $\Sigma^+$  puede obtenerse en base a su matriz unificándola. Si  $F$  es:  $F^T \Sigma_{11} F = I_p$ ;  $\Sigma_{11} = F(F^T F)^{-2} F^T$  cumple las condiciones 1, 2, 3 y 4 de Moore-Penrose inversa.

en donde  $\text{tr}_p A \equiv$  Suma de los menores principales de orden  $p$   
 en  $A$ :  $p \leq p$ ;  $\text{ran } \Sigma_{11} = p$

Igualmente se demuestra que hay  $(p-t)$  correlaciones canónicas iguales a 1 siendo  $t$  el rango de

$$\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{+} \Sigma_{21}$$

Si la distribución no es degenerada  $\Sigma^{+} = \Sigma_{22}^{-1}$  y

$$\text{tr}_p \left[ \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right] = \left| \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \lambda^2 \Sigma_{11} \right| = 0$$

cuyas raíces son el cuadrado de las correlaciones canónicas.

### 1. 7. La interpretación en el modelo canónico

Una de las principales críticas que ha tenido el modelo canónico es la dificultad de interpretación de los vectores canónicos. En tal sentido se manifiesta Kendall M. G. <sup>(1)</sup>, Lebart et Fenelon <sup>(2)</sup>, Cooley <sup>(3)</sup> y otros autores. Este problema inherente en varias de las técnicas del análisis multivariante ha tenido adecuado tratamiento en los modelos de Componentes Principales y Análisis Factorial desarrollándose un elevado número de procedimientos que permiten esclarecer el significado de las variables resultantes del análisis efectuado; citaremos por ser los más conocidos los métodos "Vari-max", "Oblimax", "Quartimax", "Quartimin". Todos ellos permiten obtener una nueva estructura de coeficientes que ponga de manifiesto las dependencias existentes entre las variables iniciales y las nuevas. En síntesis estos procedimientos son transformaciones tanto ortogonales como oblicuas que no desvirtúan el objetivo del análisis y ponen de manifiesto dependencias que pudieran no ser detectables en una primera etapa.

Por lo que respecta al modelo canónico este tipo de análisis secundario ha sido totalmente relegado, únicamente hemos encontrado en la literatura sobre el tema dos referencias a este problema y una de ellas referida a un caso restringido del modelo canónico como es

- 
- (1) Kendall, M.G. [ 26 ] pág. 305. Destaca la importancia desde el punto de vista teórico el modelo canónico unido a su dificultad de interpretación.
- (2) Lebart et Fenelon [32 ] pág. 275. Se manifiesta en el mismo sentido que Kendall afirmando su importancia teórica.
- (3) Cooley W. [ 8 ] pág. 185. Señala la dificultad de interpretación de las variables canónicas proponiendo en lugar de utilizar el análisis canónico en el análisis de datos los siguientes procedimientos:
- Análisis factorial conjunto de X e Y.
  - Análisis factorial por separado de cada grupo y tomar como vectores de referencia los factores obtenidos.

el análisis de varianza multivariante.

En un trabajo de Cliff, N. y Krus, J. <sup>(1)</sup> se plantea un procedimiento para obtener una nueva estructura de coeficientes de los vectores canónicos que permita identificarlos.

En síntesis este método consiste en hallar una transformación ortogonal común de los coeficientes de tales vectores que de acuerdo a un determinado criterio dé lugar a cierto contraste en las columnas de las matrices que contienen los nuevos coeficientes y que se conserven en cierta medida las principales propiedades del modelo canónico.

$$\text{Si} \quad \begin{array}{l} \mu = L^T x \quad ; \quad L_{r \times p} \\ \sigma = M^T y \quad ; \quad M_{r \times q} \end{array}$$

son las  $r$  parejas de variables canónicas asociadas a las  $r$  elementos no nulos de  $P$ , se trata de hallar las variables

$$\begin{array}{l} z_1 = T L^T x \\ z_2 = T M^T y \end{array} \quad (86) \quad \begin{array}{l} : T_{r \times r} ; \\ T^T T = I_r \end{array}$$

de forma que las matrices  $T L^T$  y  $T M^T$  presenten en sus filas contraste entre sus elementos <sup>(2)</sup>.

Las  $r$  parejas de variables canónicas  $(\mu, \sigma)$  poseen una matriz de correlaciones de la forma:

$$\Sigma_{\mu, \sigma} = \begin{bmatrix} I_r & : & P_r \\ \vdots & & \vdots \\ P_r & : & I_r \end{bmatrix}$$

en donde  $P_r$  es una matriz diagonal cuyos elementos no nulos son los correspondientes a la matriz  $P$  de (4).

(1) Cliff, N. y Krus, J. [ 7 ] pág. 35-43.

(2) Algunos de los coeficientes serán elementos próximos a cero mientras que otros tendrán valores elevados.

No ocurre lo mismo con la matriz de correlaciones de los vectores  $z_1$  y  $z_2$  siendo ésta

$$\Sigma_{z_1, z_2} = \begin{bmatrix} I_v & TP_r T^T \\ \vdots & \vdots \\ T^T P_r T & I_v \end{bmatrix} \quad (87)$$

El bloque  $TP_r T^T$  en general no es una matriz diagonal; no obstante las variables transformadas  $z_1, z_2$  poseen una importante propiedad como es la conservación de la suma de las correlaciones canónicas.

En efecto, si denominamos  $R$  a la matriz:

$$R = T P T^T$$

$P$  es la matriz diagonalizada de  $R$  siendo  $T$  la matriz ortogonal de vectores asociados de  $R$

Una de las propiedades de la diagonalización es que se conserva el rango y la traza <sup>(1)</sup>, luego la suma de correlaciones canónicas permanece constante.

---

(1)  $\text{rang } T P T^T = \text{rang } R$ ; como  $T \neq 0$  rango  $T P T^T = \text{rang } P = \text{rang } R$

$\text{tr } T R T^T = \text{tr } P T^T T = \text{tr } P = \text{tr } R$ .

Veáse Fz. Trocóniz [13] pág. 1. 16. 23.

Como conclusión, el procedimiento produce incorrelación dentro de cada vector de variables pero no incorrelación entre las parejas de componentes de cada vector con distinto ordinal.

A nuestro juicio este segundo resultado no desvirtúa el objetivo perseguido por el modelo canónico por lo cual la interpretación que en base a él se obtenga es válida.

Hasta aquí no se ha hecho referencia a la forma de determinar  $T$ . Cualquier procedimiento que persiga un contraste entre los coeficientes es válido con tal que la transformación sea ortogonal, en este sentido no vemos ningún inconveniente a la utilización de los procedimientos de rotación ortogonal del Análisis Factorial en las dos matrices de coeficientes de las variables canónicas consideradas en conjunto. Esta última cuestión es importante ya que una rotación ortogonal por separado dará lugar a dos matrices  $T$  y  $T'$  sin ninguna propiedad en relación a las correlaciones canónicas de la matriz  $TRT'$ .

Entre los procedimientos de obtención de  $T$  citaremos:

a) Varimax<sup>(1)</sup>

Consiste en obtener la matriz de transformación  $T$  de forma que sus coeficientes hagan máxima la función

$$V = \frac{1}{p} \sum_i \sum_j t_{ij}^4 - \frac{1}{p^2} \sum_j \left( \sum_i t_{ij}^2 \right)^2$$

b) Quartimax

En este procedimiento la matriz de transformación  $T$  se determina haciendo máxima la función:

$$V^1 = \frac{1}{p^2} \sum_i \sum_j t_{ij}^4 - (\bar{t}^2)^2 \quad : \quad \bar{t}^2 = \frac{1}{p^2} \sum_i \sum_j t_{ij}^2$$

---

(1) Harman, H. [17] pág. 283 y 290.

1. 8. Distribución de los vectores canónicos poblacionales.

Sea  $(x^T, y^T) \in N_{p+q}(\mu, \Sigma)$

Las distribuciones de los vectores canónicos  $u = L^T x$  y  $v = M^T y$  pueden obtenerse por transformación lineal de las componentes marginales  $X$  e  $Y$ .

Las distribuciones marginales de  $X$  e  $Y$  son normales <sup>(1)</sup> de parámetros:

$$\begin{aligned} x &\in N_p(\mu_1, \Sigma_{11}) \\ y &\in N_q(\mu_2, \Sigma_{22}) \end{aligned} \quad : \mu = (\mu_1, \mu_2) \text{ y } \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Las transformaciones lineales de variables normales también son normales <sup>(2)</sup>.

Estas dos propiedades permiten obtener las funciones de densidad de  $u$  y  $v$ .

$$\begin{aligned} f(u) &= \frac{1}{(\sqrt{2\pi})^p \sqrt{|L^T \Sigma_{11} L|}} \cdot e^{-\frac{(u-\mu_u)^T [L^T \Sigma_{11} L]^{-1} (u-\mu_u)}{2}} = \frac{1}{(\sqrt{2\pi})^p} \cdot e^{-\frac{(u-\mu_u)^T (u-\mu_u)}{2}} \\ f(v) &= \frac{1}{(\sqrt{2\pi})^q \sqrt{|M^T \Sigma_{22} M|}} \cdot e^{-\frac{(v-\mu_v)^T [M^T \Sigma_{22} M]^{-1} (v-\mu_v)}{2}} = \frac{1}{(\sqrt{2\pi})^q} \cdot e^{-\frac{(v-\mu_v)^T (v-\mu_v)}{2}} \end{aligned}$$

Si suponemos que  $(X \ Y)$  es de vector de medios  $\mu = 0$  las distribuciones de los vectores canónicos son normales reducidas.

(1) La prueba puede verse en Fz. Trocóniz, A. [13] pág. 29.4. utilizando el teorema de Unicidad e Inversión.

(2) Igualmente para la prueba de esta afirmación nos remitiremos a Fz. Trocóniz, A. [13] pág. 29.5.

CAPITULO II: Los modelos canónicos parciales. Dos posibles  
extensiones.

## 2. 1. Introducción

1) Bajo el nombre genérico de modelos canónicos parciales incluiremos un conjunto de generalizaciones al modelo desarrollado por Hotelling<sup>(1)</sup> que extienden los conceptos de correlación parcial, biparcial, y otros de la regresión múltiple a vectores de variables.

El estudio de estos modelos es relativamente reciente; haremos una breve referencia cronológica. El primer trabajo en esta línea ha sido el realizado por Rao<sup>(2)</sup> R. B. Este autor, en 1969, obtiene las correlaciones entre dos vectores de variables - cuando se ha eliminado la influencia lineal de un tercero denominándolos correlaciones canónicas parciales.

En 1976 Timm, N. y Carlson, J.<sup>(3)</sup> presentan dos variaciones sobre el modelo de Rao, R. B. que denominan modelos parte y biparcial. El primero de ellos se refiere al caso en que el tercer vector de variables sea incorrelado con alguno de los dos primeros; a las correlaciones canónicas derivadas de este modelo las denominan correlaciones canónicas parte. El segundo modelo presenta un caso más general que el desarrollado por Rao, R. B.; en él se considera la correlación entre dos vectores cuando se ha eliminado la influencia lineal sobre ellos de otros dos; a los coeficientes de correlación así obtenidos los denominan coeficientes de correlación canónica biparciales.

---

(1) Hotelling, H. [ 22 ] pág. 129-142.

(2) Rao, R. B. [ 42 ] pág. 211-219.

(3) Timm, N. y Carlson, J. [ 48 ] pág.

Una última aportación en este campo es la debida a Sik - Yum - Lee (1978).<sup>(1)</sup> Este autor propone dos nuevos modelos a los que denomina  $G_1$  y  $G_2$ . El modelo  $G_1$  constituye una generalización tanto del de Rao como de los de Timm y Carlson. En él establece una cadena de posibles dependencias entre cinco vectores de variables obteniendo las correlaciones entre dos de ellos cuando se ha eliminado la influencia lineal de los otros. El modelo  $G_2$  supone una generalización de  $G_1$ .

Dividiremos este capítulo en dos partes: La primera de ellas estará dedicada a la exposición de los trabajos de Rao, Timm y Carlson y Sik-Yum-Lee, haciendo mención de como unos -- generalizan a los otros. En dicha exposición realizaremos una serie de modificaciones de los artículos originales en orden a seguir la misma línea metodológica adoptada en el capítulo I, es decir, no se obtendrán las correlaciones y vectores canónicos analizados en cada caso optimizando una función, sino que se probará la existencia de dichas variables en base a la descomposición singular de una matriz adecuada a cada uno de los modelos a analizar; dicha descomposición singular conducirá a las ecuaciones que permiten obtener las correlaciones y vectores canónicos.

La segunda parte estará dedicada al desarrollo de dos modelos canónicos parciales originales que generalizan a  $G_2$  y por tanto a todos los anteriores. Dichos modelos han sido obtenidos por un procedimiento inductivo y les hemos denominado modelos canónicos parciales circular y lineal en base a la estructura del grafo de dependencias entre vectores que en ellos se suponen. En notación abreviada nos referiremos a ellos denominándoles  $C(2n+1)$  y  $L(2n+1)$  siendo  $n$  un número natural.

---

(1) Sik-Yum-Lee, [47] pág.427-431.

Probaremos en esta segunda parte que  $C(2n+1)$  coincide con  $L(2n+1)$  si  $n = 1, \text{ ó } 2$  obteniéndose en estos casos los modelos de Rao y  $G_2$ . Para  $n = 3, 4, \dots$   $C(2n+1)$  y  $L(2n+1)$  presentan estructuras de dependencia diferentes.

Con la notación aquí adoptada nos referiremos al modelo de Rao como  $L(3)$  o bien  $C(3)$ . Igualmente el modelo  $G_2$  será denominado  $L(5)$  o  $C(5)$ .

En la exposición de los métodos que contiene este capítulo haremos frecuente uso de la regresión entre dos vectores de variables, por lo cual comenzamos exponiendo el sentido que damos a este concepto.

Si  $X$  e  $Y$  son dos vectores de variables aleatorias definidas conjuntamente de dimensiones  $p$  y  $q$ , de parámetros  $(\theta, \Sigma)$  denominaremos modelo de regresión lineal múltiple, multivariante del vector  $Y$  sobre el vector  $X$  al sistema de ecuaciones:

$$y = Ax + \varepsilon \quad (1)$$

en donde la matriz  $A$  recoge en sus filas los coeficientes de las regresiones múltiples de cada componente de  $Y$  sobre  $X$ .

Dichas filas de coeficientes pueden obtenerse haciendo mínimas las funciones

$$E (y_i - \alpha_i^T x)^2 \quad i = 1, 2, \dots, q$$

siendo los elementos de  $\alpha_i$  que hacen mínima la anterior expresión<sup>(1)</sup>

---

(1) Fz. de Trocóniz, A. [ 12 ] cap. 16. 3.

$$a_i^T = (\lambda_{p+i,1}, \dots, \lambda_{p+i,p}) \Sigma_{11}^{-1} \quad i = 1, \dots, q \quad (2)$$

en donde  $\Sigma_{11}$  y  $\lambda_{p+i,1} \dots \lambda_{p+i,p}$  son la matriz de covarianzas de las componentes del vector  $X_j$ , y las covarianzas de la componente  $y_i$  con todas las componentes del vector  $X$ .

Si particionamos  $\Sigma$  en base a las dimensiones de  $X$  e  $Y$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (3)$$

teniendo en cuenta (2) y el sentido que hemos dado a la matriz  $A$  ésta puede expresarse en términos de (3):

$$A = \Sigma_{21} \Sigma_{11}^{-1} \quad (4)$$

Designaremos  $Y^*$  al vector de estimaciones lineales m. c. de  $Y$  por el vector  $X$ .

$$y^* = Ax = \Sigma_{21} \Sigma_{11}^{-1} x$$

$$\varepsilon_x = y - y^* = y - \Sigma_{21} \Sigma_{11}^{-1} x \quad (5)$$

Otra cuestión a tener en cuenta en el desarrollo de este capítulo es el significado del vector residuo aleatorio. Siguiendo a Rao R. B. (1) dicho residuo muestra al vector  $Y$  después que se ha eliminado la influencia real del vector  $X$ .

Si se considera la regresión de  $X$  sobre  $Y$  se probaría de forma análoga que:

$$B = \Sigma_{12} \Sigma_{22}^{-1} \\ x^* = \Sigma_{12} \Sigma_{22}^{-1} y \quad ; \quad \varepsilon_x = x - \Sigma_{12} \Sigma_{22}^{-1} y.$$

Hechas estas precisiones pasamos a exponer los modelos canónicos parciales.

(1) Rao, R. B. [42] pág. 213.

## 2. 2. Modelos parciales

Siguiendo la pauta señalada en la introducción expondremos en esta sección los modelos canónicos que han sido denominados por sus autores: parcial, parte, biparcial,  $G_1$  y  $G_2$ . En todos ellos nos ha parecido conveniente señalar las posibles dependencias entre los diferentes vectores de variables a través de un grafo, evitando así una extensa descripción que por otra parte no tendría la claridad del grafo. En los primeros modelos que expondremos dicho grafo parece innecesario debido a que las relaciones entre vectores de variables que en ellos se suponen son muy simples, no obstante permite obtener una representación visual de cómo unos modelos generalizan a otros.

### 2. 2. 1. Modelo parcial - C(3) - H(3).

La publicación en 1969 del artículo "Partial canonical correlations" debido a Rao, R. B. <sup>(1)</sup> abre un nuevo campo de desarrollo del análisis canónico. En dicho artículo su autor extiende el concepto de correlación parcial entre dos componentes de un vector de tres variables aleatorias al de correlación canónica parcial entre dos vectores de variables de un supervector de tres vectores.

Expondremos dicho modelo con las modificaciones señaladas en la introducción.

Sean  $( X_1^T, X_2^T, X_3^T )$  tres vectores de variables aleatorias definidas conjuntamente con parámetros  $(0, \Sigma)$  y de dimensiones  $n_1, n_2, n_3, : n_1 \leq n_2$  .

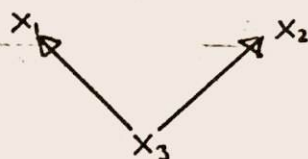
---

(1) Rao, R. B. [42] pág. 211-219.

La matriz de covarianzas de  $(X_1^T, X_2^T, X_3^T)$  particionada según las dimensiones de los vectores  $X_1, X_2, X_3$  es:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}$$

Consideremos que el vector  $X_3$  influye sobre  $X_1$  y  $X_2$ . Estas dependencias las expresaremos según el siguiente grafo:



G-1

La matriz de covarianzas entre los vectores  $X_1$  y  $X_2$  después de eliminar las influencias lineales de  $X_3$  en ambos vectores viene dada por la expresión <sup>(1)</sup>:

$$\Sigma_{.3} = \begin{bmatrix} \Sigma_{11.3} & \Sigma_{12.3} \\ \Sigma_{21.3} & \Sigma_{22.3} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} \\ \Sigma_{21} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{31} & \Sigma_{22} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{32} \end{bmatrix} \quad (7)$$

Suponemos que el rango de  $\Sigma_{33}$  es  $r \leq n_3$ .

Se denominan vectores canónicos parciales entre los vectores  $X_1$  y  $X_2$  a los vectores  $U, V$ :

$$\begin{aligned} U &= L^T \varepsilon_{13} \\ V &= M^T \varepsilon_{23} \end{aligned} \quad (8)$$

(1) Si  $X_1^* = \Sigma_{13} \Sigma_{33}^{-1} X_3$   
 $X_2^* = \Sigma_{23} \Sigma_{33}^{-1} X_3$ .

$E(X_1 - X_1^*)(X_1 - X_1^*)^T = \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31}$   
 $E(X_1 - X_1^*)(X_2 - X_2^*)^T = \Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32}$   
 $E(X_2 - X_2^*)(X_2 - X_2^*)^T = \Sigma_{22} - \Sigma_{23} \Sigma_{33}^{-1} \Sigma_{32}$

en donde:

$$\mathcal{E}_{13} = X_1 - X_1^* = X_1 - \sum_{13} \sum_{33}^{-1} X_3$$

$$\mathcal{E}_{23} = X_2 - X_2^* = X_2 - \sum_{23} \sum_{33}^{-1} X_3$$

cuya matriz de covarianzas es de la forma:

$$1.2.(4). \quad \Sigma_{4\sigma} = \begin{bmatrix} I_{n_1} & P \\ P^T & I_{n_2} \end{bmatrix} \quad (9)$$

siendo  $P$  una matriz con la estructura 1.2.(5).

La nota (1) de la página 66 prueba que la matriz de covarianzas entre  $\mathcal{E}_{13}$  y  $\mathcal{E}_{23}$  es la matriz (7). Esto permite trasladar el problema de obtener los vectores canónicos parciales entre dos vectores de variables  $X_1$  y  $X_2$  al de obtener los vectores canónicos entre sus residuos. Formularemos esta cuestión de la siguiente forma: Dados dos vectores de variables que en este caso son  $\mathcal{E}_{13}$  y  $\mathcal{E}_{23}$  cuya matriz de covarianzas es la expresada en (7) se trata de obtener los nuevos vectores canónicos  $U$  y  $V$  con matriz de covarianzas de la forma (9). La descomposición singular de una matriz apropiada prueba la existencia de los vectores  $U$  y  $V$  con las características señaladas.

Sea  $C = \sum_{11.3}^{-1/2} \sum_{17.3} \sum_{23.3}^{-1/2}$  Tal como se probó en 1.3.2. su descomposición singular conduce a la obtención de las matrices  $L$  y  $M$  de las transformaciones (8) siendo éstas las soluciones de los sistemas

$$\left[ \sum_{12.3} \sum_{22.3}^{-1} \sum_{21.3} - \lambda^2 \sum_{11.3} \right] L = 0 \quad : L^T \sum_{11.3} L = I_{n_1} \quad (10)$$

$$\left[ \sum_{21.3} \sum_{11.3}^{-1} \sum_{12.3} - \lambda^2 \sum_{22.3} \right] M = 0 \quad : M^T \sum_{22.3} M = I_{n_2} \quad (11)$$

Cuando  $L$  y  $M$  satisfacen (10) y (11);  $\mu = L^T e_{13}$  y  $\sigma = M^T e_{23}$  tienen una matriz de covarianzas de la forma (9) en donde los  $r$  elementos no nulos de la matriz  $P$  son las raíces cuadradas con signo positivo de las soluciones de las ecuaciones determinantes

$$\begin{aligned} & \left| \sum_{12,3} \sum_{22,3}^{-1} \sum_{21,3} - \lambda^2 \sum_{11,3} \right| = 0 \\ & \left| \sum_{21,3} \sum_{11,3}^{-1} \sum_{12,3} - \lambda^2 \sum_{22,3} \right| = 0 \end{aligned} \quad (12)$$

Los vectores  $U$  y  $V$  se denominan vectores canónicos parciales de los vectores  $X_1$  y  $X_2$  y los  $r$  elementos no nulos de  $P$  correlaciones canónicas parciales entre  $X_1$  y  $X_2$ . Dichas correlaciones son un índice de la interdependencia entre los vectores  $X_1$  y  $X_2$  cuando se ha eliminado el efecto lineal del vector  $X_3$ .

### 2.2.2. El modelo canónico parte.

Un caso particular del modelo parcial es el denominado por Timm y Carlson <sup>(1)</sup> modelo canónico parte.

Cuando en el grafo  $G-I$  se supone que  $X_3$  es incorrelado con alguno de los vectores  $X_1$  o  $X_2$  [ bien  $\Sigma_{13} = 0$  o bien  $\Sigma_{23} = 0$  ] llegamos al modelo canónico parte que expondremos a continuación.

Sean  $(x_1^T, x_2^T, x_3^T)$  tres vectores de variables aleatorias definidas conjuntamente de dimensiones  $n_1, n_2$  y  $n_3$ :  $n_1 \leq n_2$  y de parámetros  $(\theta, \Sigma)$  y supongamos que  $X_3$  influye linealmente en  $X_1$



G-II

(1) Timm y Carlson [ 48 ] pág. 165-174.

La matriz de covarianzas entre el vector  $X_1$  y el vector  $X_2$  después de eliminada la influencia lineal de  $X_3$  vendrá dada por la expresión <sup>(1)</sup>:

$$\sum_{\varepsilon_{13} X_1} = \begin{bmatrix} \Sigma_{11.3} & \Sigma_{12.3} \\ \Sigma_{21.3} & \Sigma_{22} \end{bmatrix} \quad (13)$$

Sea  $r \leq n_1$  el rango de  $\Sigma_{11.3}$ . Se denominan vectores canónicos parte entre los vectores  $X_1$  y  $X_2$  a los vectores

$$\begin{aligned} \mu &= L^T \varepsilon_{13} \\ \sigma &= M^T X_2. \end{aligned} \quad (14)$$

cuya matriz de covarianzas es de la forma (9).

La descomposición singular de la matriz  $\Sigma_{11.3}^{-1/2} \Sigma_{12.3} \Sigma_{22}^{-1/2}$  permite probar la existencia de tales vectores deduciéndose según 1.3.2. que si  $L$  y  $M$  son las soluciones de los sistemas

$$\left[ \Sigma_{12.3} \Sigma_{22}^{-1} \Sigma_{21.3} - \lambda^2 \Sigma_{11.3} \right] L = 0 \quad : \quad L^T \Sigma_{11.3} L = 0 \quad (15)$$

$$\left[ \Sigma_{21.3} \Sigma_{11.3}^{-1} \Sigma_{12.3} - \lambda^2 \Sigma_{22} \right] M = 0 \quad M^T \Sigma_{22} M = 0 \quad (16)$$

$U = L^T \varepsilon_{13}$  y  $V = M^T$  y tienen una matriz de covarianzas de la forma (9) en donde los elementos no nulos de la matriz  $P$  son la raíz - con signo positivo en las  $r$  soluciones comunes a las ecuaciones determinantes:

$$\begin{aligned} \left| \Sigma_{12.3} \Sigma_{22}^{-1} \Sigma_{21.3} - \lambda^2 \Sigma_{11.3} \right| &= 0 \\ \left| \Sigma_{21.3} \Sigma_{11.3}^{-1} \Sigma_{12.3} - \lambda^2 \Sigma_{22} \right| &= 0 \end{aligned} \quad (17)$$

(1) Véase la nota (1) de la pág. 66.

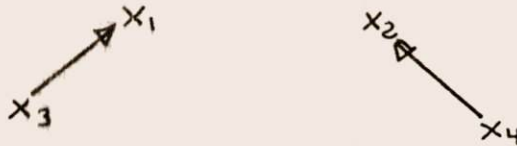
En este caso por ser  $\Sigma_{23} = [0]$  se tiene  $\Sigma_{22.3} = \Sigma_{22}$

A los vectores (14) se les denomina vectores canónicos parte entre  $X_1$  y  $X_2$  y a los  $r$  elementos no nulos de  $P$  correlaciones canónicas parte. Estas correlaciones son una medida de la interdependencia existente entre  $X_1$  y  $X_2$  cuando en  $X_1$  se ha eliminado la influencia lineal de  $X_3$ .

### 2.2.3. Modelo biparcial

Este modelo canónico desarrollado también por Timm y Carlson hace referencia a un caso más general que el modelo parcial de Rao, R. B. En él se supone que los vectores que influyen sobre  $X_1$  y  $X_2$  son diferentes.

Sean  $(x_1^T, x_2^T, x_3^T, x_4^T)$  vectores de variables aleatorias distribuidas conjuntamente de dimensiones  $n_1, n_2, n_3$  y  $n_4$ :  $n_1 \leq n_2$  y de parámetros  $(0 \Sigma)$  y supongamos que  $X_3$  influye en  $X_1$  y  $X_4$  en  $X_2$



G-III

La matriz de covarianzas particionada en base a las dimensiones de tales vectores de variables la expresaremos :

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{bmatrix} \quad (18)$$

La matriz de covarianzas entre  $X_1$  y  $X_2$  después de eliminar las influencias lineales expresadas en G-III vendrá dada por:

$$\Sigma_{\varepsilon_{13} \varepsilon_{24}} = \begin{bmatrix} \Sigma_{11.3} & \Sigma_{12}^* \\ \Sigma_{21}^* & \Sigma_{22.4} \end{bmatrix} \quad (19)$$

en donde los elementos  $\Sigma_{11.3}$ ,  $\Sigma_{12}^*$  y  $\Sigma_{22.4}$  adoptan las siguientes formas <sup>(1)</sup>

$$\Sigma_{11.3} = E \varepsilon_{13} \varepsilon_{13}^T = \Sigma_{11} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} = \Sigma_{11.3}$$

$$\Sigma_{22.4} = E \varepsilon_{24} \varepsilon_{24}^T = \Sigma_{22} - \Sigma_{24} \Sigma_{44}^{-1} \Sigma_{42} = \Sigma_{22.4}$$

$$\Sigma_{12}^* = E \varepsilon_{13} \varepsilon_{24}^T = \Sigma_{12} - \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{32} - \Sigma_{14} \Sigma_{44}^{-1} \Sigma_{43} + \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{34} \Sigma_{44}^{-1} \Sigma_{42}$$

Sea  $r \leq n_1$  el rango de  $\Sigma_{12}^*$ . Se denominan vectores canónicos biparciales a los obtenidos a través de la transformación

$$\begin{aligned} \mu &= L^T \varepsilon_{13} \\ \nu &= M^T \varepsilon_{24} \end{aligned} \quad (20)$$

cuya matriz de covarianzas es de la forma (9).

En este caso la descomposición singular de  $\Sigma_{11.3}^{-1/2} \Sigma_{12}^* \Sigma_{22.4}^{-1/2}$  prueba la existencia de los vectores (20) con las características indicadas, siendo L y M las soluciones de los sistemas:

$$\begin{aligned} [\Sigma_{12}^* \Sigma_{22.4}^{-1} \Sigma_{21}^* - \lambda^2 \Sigma_{11.3}] L &= 0 & : & \quad L^T \Sigma_{11.3} L = I_{n_1} \\ [\Sigma_{21}^* \Sigma_{11.3}^{-1} \Sigma_{12}^* - \lambda^2 \Sigma_{22.4}] M &= 0 & : & \quad M^T \Sigma_{22.4} M = I_{n_2} \end{aligned} \quad (21)$$

y los elementos no nulos de P las raíces cuadradas con signo positivo de las soluciones de las ecuaciones determinantes derivadas de los sistemas (21). A dichos elementos no nulos Timm y Carlson los denominan correlaciones canónicas biparciales y son una medida de las interrelaciones existentes entre los vectores  $X_1$  y  $X_2$  cuando se han eliminado las influencias lineales de  $X_3$  y  $X_4$ .

(1) Estas expresiones pueden obtenerse teniendo en cuenta que:

$$\begin{aligned} \varepsilon_{13} &= X_1 - \Sigma_{13} \Sigma_{33}^{-1} X_3 \\ \varepsilon_{24} &= X_2 - \Sigma_{24} \Sigma_{44}^{-1} X_4. \end{aligned}$$

Este modelo constituye un caso más general que los que hasta aquí se han presentado. Si alguno de los vectores  $X_3$  y  $X_4$  está incorrelado con  $X_1$  o  $X_2$  se obtiene el modelo canónico parte.

Si  $X_3$  coincide con  $X_4$  el resultado es el modelo canónico parcial.

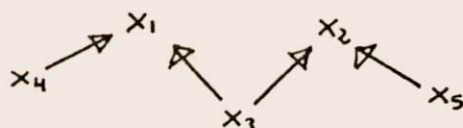
#### 2.2.4. Modelo canónico $G_1$

Un nuevo avance en este campo de los modelos canónicos parciales lo da Sik - Yum - Lee<sup>(1)</sup>. Este autor en la parte correspondiente a notas y comentarios de la revista *Psychometrika* correspondiente al mes de Septiembre de 1978 desarrolla dos modelos que generalizan todos los resultados expuestos hasta este punto. Dedicaremos este apartado a estudiar el que su autor a denominado  $G_1$ .

Sean  $(x_1^T, x_2^T, x_3^T, x_4^T, x_5^T)$  vectores de variables aleatorias definidas conjuntamente de dimensiones  $n_1, n_2, n_3, n_4$  y  $n_5$ :  $n_1 \leq n_2$  y sean  $(0, \Sigma)$  sus parámetros.

Denotaremos  $\Sigma = [\Sigma_{ij}]_{i,j=1..5}$  la matriz de covarianzas particionada en base a las dimensiones de los cinco vectores.

Consideremos las dependencias señaladas en el grafo adjunto



G-IV

(1) Sik - Yum - Lee [ 47 ] pág. 427-431.

Si se eliminan las influencias lineales sobre  $X_1$  y  $X_2$  de los vectores señalados en G-IV la matriz de covarianzas resultante será expresada

$$\sum_{\epsilon_{134}, \epsilon_{235}} = \begin{bmatrix} \sum_{11}^{***} & \sum_{12}^{***} \\ \sum_{21}^{***} & \sum_{22}^{***} \end{bmatrix} \quad (23)$$

en donde

$$\sum_{11}^{***} = \sum_{11.3} - \sum_{14.3} \sum_{44}^{-1} \sum_{41.3} \quad (24)$$

$$\sum_{22}^{***} = \sum_{22.3} - \sum_{25.3} \sum_{55}^{-1} \sum_{52.3} \quad (25)$$

$$\begin{aligned} \sum_{12}^{***} &= \sum_{12.3} - \sum_{14.3} \sum_{44}^{-1} \sum_{42.3} - \sum_{15.3} \sum_{55}^{-1} \sum_{51.3} + \\ &+ \sum_{14.3} \sum_{44}^{-1} \sum_{45} \sum_{55}^{-1} \sum_{52.3} \end{aligned} \quad (26)$$

Las expresiones (24), (25) y (26) han sido obtenidas en base a la realización de un conjunto de regresiones que permiten eliminar en  $X_1$  y  $X_2$  las influencias de las dependencias admitidas. Expondremos a continuación el procedimiento para llegar a ellas.

$$\text{Si } x_1^* = \sum_{13} \sum_{33}^{-1} x_3 ; \quad \epsilon_{13} = x_1 - \sum_{13} \sum_{33}^{-1} x_3.$$

la regresión de  $\epsilon_{13}$  sobre  $X_4$  vendrá dada por

$$(5) \quad \epsilon_{13}^* = E(\epsilon_{13} x_4^T) \cdot [E(x_4 x_4^T)]^{-1} x_4 = \sum_{14.3} \sum_{44}^{-1} x_4$$

siendo el residuo de dicha regresión

$$\epsilon_{134} = \epsilon_{13} - \sum_{14.3} \sum_{44}^{-1} x_4 \quad (27)$$

De la misma forma si  $x_2^* = \sum_{23} \sum_{33}^{-1} x_3 ; \quad \epsilon_{23} = x_2 - \sum_{23} \sum_{33}^{-1} x_3$  la regresión de  $\epsilon_{23}$  sobre  $X_5$  da lugar a ;

$$(5) \quad \epsilon_{23}^* = E(\epsilon_{23} x_5^T) [E(x_5 x_5^T)]^{-1} x_5 = \sum_{25.3} \sum_{55}^{-1} x_5$$

con lo cual

$$\varepsilon_{235} = \varepsilon_{23} - \sum_{25.3} \sum_{55}^{-1} x_5 \quad (28)$$

Las expresiones (27) y (28) muestran los vectores  $X_1$  y  $X_2$  desprovistos de las Influencias lineales supuestas en G-IV. Ambas nos permiten llegar a (24), (25) y (26).

$$\begin{aligned} \Sigma_{11}^{**} &= E (\varepsilon_{13} - \sum_{14.3} \sum_{44}^{-1} x_4) (\varepsilon_{13} - \sum_{14.3} \sum_{44}^{-1} x_4)^T \\ &= \Sigma_{11.3} - \sum_{14.3} \sum_{44}^{-1} \Sigma_{41.3} \quad \text{que prueba (24)} \end{aligned}$$

$$\begin{aligned} \Sigma_{22}^{**} &= E (\varepsilon_{23} - \sum_{25.3} \sum_{55}^{-1} x_5) (\varepsilon_{23} - \sum_{25.3} \sum_{55}^{-1} x_5)^T \\ &= \Sigma_{22.3} - \sum_{25.3} \sum_{55}^{-1} \Sigma_{52.3} \quad \text{que prueba (25)} \end{aligned}$$

$$\begin{aligned} \Sigma_{12}^{**} &= E (\varepsilon_{13} - \sum_{14.3} \sum_{44}^{-1} x_4) (\varepsilon_{23} - \sum_{25.3} \sum_{55}^{-1} x_5)^T = \Sigma_{12.3} - \sum_{15.3} \sum_{55}^{-1} \Sigma_{523} \\ &\quad - \sum_{14.3} \sum_{44}^{-1} \Sigma_{42.3} + \sum_{14.3} \sum_{44}^{-1} \sum_{45} \sum_{55}^{-1} \Sigma_{52.3} \quad \text{que prueba (26)} \end{aligned}$$

Una vez obtenida la matriz covarianzas entre  $X_1$  y  $X_2$  sin sus influencias lineales en cada caso el proceso es similar al seguido en los modelos anteriores.

Sea  $r \leq n_1$  el rango de  $\Sigma_{12}^{**}$

Se denominan vectores canónicos  $G_1$  a los obtenidos a través de la siguiente transformación:

$$\begin{aligned} \mu &= L^T \varepsilon_{134} \\ \nu &= M^T \varepsilon_{235} \end{aligned} \quad (29)$$

cuya matriz de covarianzas es de la forma (9).

La matriz cuya descomposición singular prueba la existencia de las variables (29) es en este caso

$$\sum_{11}^{**} - \frac{1}{2}, \quad \sum_{12}^{**}, \quad \sum_{22}^{**} - \frac{1}{2}. \quad (30)$$

siendo L y M soluciones de sistemas análogos a los expuestos en modelos anteriores sin más que sustituir las particiones de la matriz de covarianzas por (24), (25) y (26).

Se denominan correlaciones canónicas  $G_1$  entre  $X_1$  y  $X_2$  a la  $r$  de elementos no nulos de la matriz P obtenidos en forma análoga a como se indicó con anterioridad.

Destacaremos como  $G_1$  generaliza a todos los modelos parciales que se han expuesto hasta este momento:

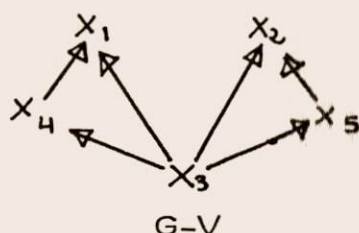
Cuando  $X_4$  es incorrelado con  $X_1$  ( $\Sigma_{41} = [0]$ ) y  $X_5$  lo es con  $X_2$  ( $\Sigma_{52} = [0]$ ) el modelo que se obtiene es el parcial de Rao.

Si  $X_3$  está incorrelado con  $X_1$  y  $X_2$  ( $\Sigma_{13} = [0] \wedge \Sigma_{23} = [0]$ ) el modelo resultante es el biparcial.

#### 2.2.5. Modelo $G_2$ - L(5) - C(5) -

Finalmente dentro de este conjunto de modelos parciales expondremos el modelo canónico  $G_2$  debido también a Sik - Yum - Lee. Las características de los vectores de variables que en él in-

tervienen son iguales que en  $G_1$  y la estructura de dependencias entre vectores viene dada a través del siguiente grafo



Análogamente a los casos anteriores se pretende hallar las correlaciones canónicas entre  $X_1$  y  $X_2$  cuando en estos vectores se han eliminado las influencias lineales supuestas en  $G-V$ .

Señalaremos el procedimiento de eliminar tales influencias en los vectores anteriores.

a) Eliminación de efectos lineales en  $X_1$ .

$$\text{Sea } X_4^* = \sum_{43} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{43} = X_4 - \sum_{43} \sum_{33}^{-1} X_3$$

$$\text{y } X_1^* = \sum_{13} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{13} = X_1 - \sum_{13} \sum_{33}^{-1} X_3.$$

El error de la regresión de  $\varepsilon_{13}$  sobre  $\varepsilon_{53}$  presenta al vector  $X_1$  desprovisto de las influencias de  $X_3$  y  $X_4$

$$\varepsilon_{13}^* = E \varepsilon_{13} \varepsilon_{43}^T [E \varepsilon_{43} \varepsilon_{43}^T]^{-1} \varepsilon_{43} = \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43}$$

$$\varepsilon_{134} = \varepsilon_{13} - \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43}. \quad (31)$$

b) Eliminación de efectos lineales en  $X_2$

$$\text{Sea } X_5^* = \sum_{53} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{53} = X_5 - \sum_{53} \sum_{33}^{-1} X_3$$

$$\text{y } X_2^* = \sum_{23} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{23} = X_2 - \sum_{23} \sum_{33}^{-1} X_3$$

de la misma forma el error de regresión de  $\varepsilon_{13}$  sobre  $\varepsilon_{53}$  presenta al vector  $X_2$  desprovisto de las influencias lineales de  $X_3$  y  $X_5$

$$\varepsilon_{23}^* = E(\varepsilon_{23} \varepsilon_{53}^T) [E(\varepsilon_{53} \varepsilon_{53}^T)]^{-1} \varepsilon_{53} = \sum_{25.3} \sum_{55.3}^{-1} \varepsilon_{53}$$

$$\varepsilon_{235} = \varepsilon_{23} - \sum_{25.3} \sum_{55.3}^{-1} \varepsilon_{53}. \quad (32)$$

Las expresiones (31) y (32) permiten obtener la matriz de covarianzas entre  $\varepsilon_{134}$  y  $\varepsilon_{235}$

$$\sum_{\varepsilon_{134}, \varepsilon_{235}} = \begin{bmatrix} \sum_{11}^{***} & \sum_{12}^{***} \\ \sum_{21}^{***} & \sum_{22}^{***} \end{bmatrix} \quad (33)$$

en donde<sup>(1)</sup>

$$\begin{aligned} \sum_{11}^{***} &= E(\varepsilon_{134} \varepsilon_{134}^T) = E(\varepsilon_{13} - \sum_{14.3}^{-1} \sum_{44.3}^{-1} \varepsilon_{43}) (\varepsilon_{13} - \sum_{14.3}^{-1} \sum_{44.3}^{-1} \varepsilon_{43})^T \\ &= E(\varepsilon_{13} \varepsilon_{13}^T) - \sum_{14.2} \sum_{44.3}^{-1} \cdot E(\varepsilon_{43} \varepsilon_{13}^T) - [E(\varepsilon_{13} \varepsilon_{43}^T)] \sum_{44.3}^{-1} \sum_{41.3} \\ &+ \sum_{14.3} \sum_{44.3}^{-1} (E \varepsilon_{43} \varepsilon_{43}^T) \sum_{44.3}^{-1} \sum_{41.3} = \sum_{11.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{41.3} \\ &- \sum_{14.3} \sum_{44.3}^{-1} \sum_{41.3} + \sum_{14.2} \sum_{44.3}^{-1} \sum_{44.2} \sum_{44.3}^{-1} \sum_{41.3} = \sum_{11.3.4} \quad (2) \end{aligned} \quad (34)$$

$$\begin{aligned} \sum_{22}^{***} &= E(\varepsilon_{235} \varepsilon_{235}^T) = E(\varepsilon_{23} - \sum_{25.3} \sum_{55.3}^{-1} \varepsilon_{53}) (\varepsilon_{23} - \sum_{25.3} \sum_{55.3}^{-1} \varepsilon_{53})^T \\ &= E(\varepsilon_{23} \varepsilon_{23}^T) - \sum_{25.3} \sum_{55.3}^{-1} E(\varepsilon_{53} \varepsilon_{23}^T) - [E(\varepsilon_{23} \varepsilon_{53}^T)] \sum_{55.3}^{-1} \sum_{52.3} \\ &+ \sum_{25.3} \sum_{55.3}^{-1} [E(\varepsilon_{53} \varepsilon_{53}^T)] \sum_{55.3}^{-1} \sum_{52.3} = \sum_{22.3} - \sum_{25.3} \sum_{55.3}^{-1} \sum_{52.3} \\ &- \sum_{25.3} \sum_{55.3}^{-1} \sum_{52.3} + \sum_{25.3} \sum_{55.3}^{-1} \sum_{55.3} \sum_{55.3}^{-1} \sum_{52.3} = \sum_{22.3.5} \end{aligned} \quad (35)$$

(1) La notación \*\*\* podrá ser sustituida por C(5) o L(5) según se justificará en los apartados siguientes.

(2) La notación  $\sum_{11.3.4}$  y siguientes, viene justificada por la adoptada en (7)

$$\begin{aligned}
\Sigma_{12}^{***} &= E (\varepsilon_{13} - \Sigma_{14.3} \Sigma_{44.3}^{-1} \varepsilon_{43}) (\varepsilon_{23} - \Sigma_{25.3} \Sigma_{55.3}^{-1} \varepsilon_{53})^T \\
&= E (\varepsilon_{13} \varepsilon_{23}^T) - \Sigma_{14.3} \Sigma_{44.3}^{-1} [E (\varepsilon_{43} \varepsilon_{23}^T)] - [E (\varepsilon_{13} \varepsilon_{53}^T)] \Sigma_{55.3}^{-1} \Sigma_{52.3} \\
&+ \Sigma_{14.3} \Sigma_{44.3}^{-1} [E (\varepsilon_{43} \varepsilon_{53}^T)] \Sigma_{55.3}^{-1} \Sigma_{52.3} = \Sigma_{12.3} - \Sigma_{14.3} \Sigma_{44.3}^{-1} \Sigma_{42.3} \\
&- \Sigma_{15.3} \Sigma_{55.3}^{-1} \Sigma_{52.3} + \Sigma_{14.3} \Sigma_{44.3}^{-1} \Sigma_{45.3} \Sigma_{55.3}^{-1} \Sigma_{52.3} \quad (36)
\end{aligned}$$

Sea  $r \leq n_1$  el rango de  $\Sigma_{12}^{***}$ . Se denominan vectores canónicos  $G_2$  a los transformados:

$$\begin{aligned}
\mu &= L^T \varepsilon_{134} \\
\nu &= M^T \varepsilon_{235}
\end{aligned} \quad (37)$$

cuya matriz de covarianzas es de la forma señalada en (9).

La descomposición singular de:

$$\Sigma_{11.3.4}^{-1/2} \Sigma_{12}^{***} \Sigma_{22.3.5}^{-1/2} \quad (38)$$

prueba la existencia de tales vectores pudiéndose obtener  $L$  y  $M$  sin más que sustituir en cualquiera de los sistemas de los modelos anteriores las particiones de sus matrices de covarianzas por las de (33).

Los determinantes de dichos sistemas conducen al cálculo de los  $r$  elementos no nulos de la matriz  $P$  como se expuso en los anteriores apartados.

Se denominan correlaciones canónicas  $G_2$  entre  $X_1$  y  $X_2$  a los elementos no nulos de  $P$ . Dichas correlaciones muestran el grado de interdependencia entre los vectores  $X_1$  y  $X_2$  desprovistos de las influencias lineales supuestas en  $G-V$ .

Justificaremos cómo  $G_2$  generaliza a todos los modelos anteriores. Si  $X_4$  es incorrelado con  $X_1$ , y  $X_5$  lo es con  $X_2$  el modelo resultante es  $G_1$ . Como en 2.2.4. se puso de manifiesto que  $G_1$  incluía los modelos parte, parcial y biparcial;  $G_2$  es el modelo más general de los estudiados hasta este punto incluyendo como casos particulares a todos los ya mencionados.

### 2.3. Dos generalizaciones en los modelos canónicos parciales.

Todos los modelos canónicos englobados bajo el epígrafe común de "parciales" y que han sido expuestos en 2.2. tienen la característica común de ser modelos canónicos entre vectores de variables depurados de las influencias lineales de otros. No cabe duda de que este tipo de modelos no se agota en el más general de los estudiados que ha sido  $G_2$ . Otro tipo de influencias pueden ser admitidas en  $X_1$  y  $X_2$  siendo posible su eliminación por un procedimiento similar al expuesto en 2.2.5.

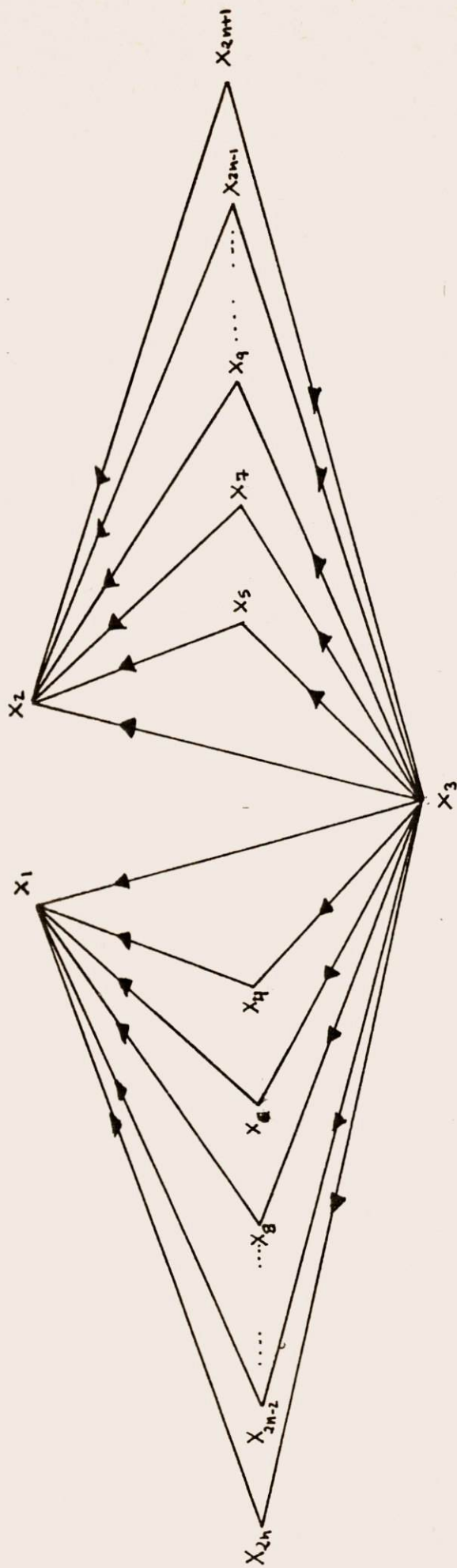
En este apartado estudiaremos dos nuevos modelos canónicos parciales que serán denominados extensión lineal y circular del modelo de Rao R. B. Adoptaremos previamente las siguientes definiciones:

Diremos que un modelo canónico parcial es la extensión lineal del modelo de Rao, R. B. cuando la estructura de dependencia entre los vectores de variables que lo integran es de la forma señalada en G-VI.

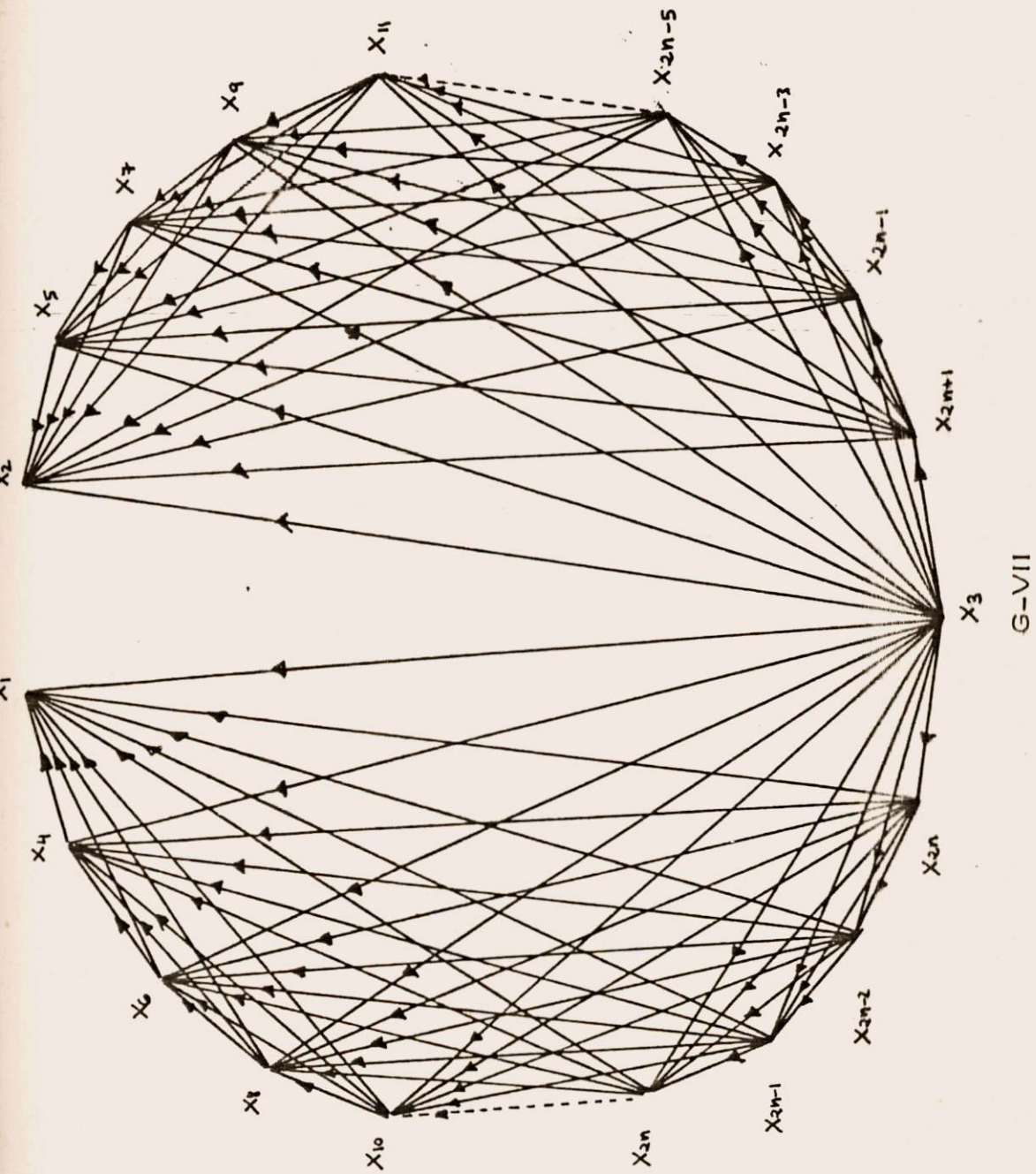
Similarmenete, diremos que un modelo canónico es la extensión circular del modelo de Rao, R. B. si las dependencias entre vectores de variables son de la forma indicada en G-VII.

A los modelos canónicos parciales derivados de G-VI y G-VII los denominaremos  $L(2n+1)$  y  $C(2n+1)$  respectivamente, en donde  $n$  es un número natural.

Las definiciones anteriores nos permiten realizar las siguien-



G-VI



G-VII

tes afirmaciones:

1. - Si  $n=1$  tanto  $L(3)$  como  $C(3)$  coinciden con el modelo canónico parcial de Rao, R. B.
2. - Si  $n=2$ ,  $L(5)$  también coincide con  $C(5)$  siendo ambos el modelo  $G_2$  de Sik - Yum - Lee.
3. - Si  $n > 2$ . Los modelos  $L(2(n+1))$  y  $C(2n+1)$  son diferentes ya que lo son sus grafos asociados. Además  $L(2n+1)$  es un caso particular de  $C(2n+1)$ .

Distinguiremos en esta sección dos apartados a desarrollar, el primero de ellos estará dedicado a la exposición de los modelos  $L(7)$  y  $C(7)$ ; el segundo contendrá una generalización de los resultados obtenidos para cualquier  $n \in \mathbb{N}$ . En esta generalización se aplicará el método inductivo, tomando como base los modelos  $L(3)$ ,  $L(5)$  y  $L(7)$  y por otra parte  $C(3)$ ,  $C(5)$  y  $C(7)$ .

### 2. 3. 1. Modelos canónicos parciales $L(7)$ y $C(7)$

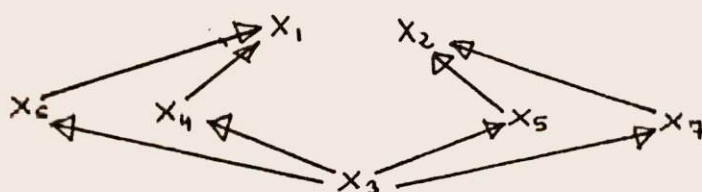
Señalada la diferencia entre estos dos modelos expondremos en primer lugar el modelo  $L(7)$  y a continuación su caso más general  $C(7)$ .

En ambos consideraremos un conjunto de siete vectores de variables aleatorias  $(x_1^T, x_2^T, x_3^T, x_4^T, x_5^T, x_6^T, x_7^T)$  definidas conjuntamente de dimensiones  $n_1, n_2, n_3, n_4, n_5, n_6$  y  $n_7$ :  $n_1 \leq n_2$  y de parámetros  $(0, \Sigma)$

Denotaremos  $\Sigma = [\Sigma_{ij}]$   $ij=1,2,\dots,7$  a la matriz de covarianzas particionadas en base a las dimensiones de los vectores considerados,

2. 3. 1. 1. Modelo L (7)

Cuando en G-VI  $n=3$  la dependencia entre los siete vectores de variables que intervienen da lugar al siguiente grafo:



G-VIII

Con objeto de obtener  $X_1, X_2$  desprovistos de los efectos lineales expuestos en G-VIII seguiremos las siguientes etapas:

1. Eliminación de efectos lineales en  $X_1$ a) Eliminación de efectos de  $X_3$  sobre  $X_4$ 

$$X_4^* = \sum_{43} \sum_{33}^{-1} X_3 ; \quad \varepsilon_{43} = X_4 - \sum_{43} \sum_{33}^{-1} X_3$$

b) Eliminación de efectos de  $X_3$  en  $X_1$ 

$$X_1^* = \sum_{13} \sum_{33}^{-1} X_3 ; \quad \varepsilon_{13} = X_1 - \sum_{13} \sum_{33}^{-1} X_3$$

c) Eliminación de efectos de  $\varepsilon_{43}$  en  $\varepsilon_{13}$ 

$$\varepsilon_{13}^* = \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43} ; \quad \varepsilon_{134} = \varepsilon_{13} - \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43}$$

d) Eliminación de efectos de  $X_3$  sobre  $X_6$ 

$$X_6^* = \sum_{63} \sum_{33}^{-1} X_3 ; \quad \varepsilon_{63} = X_6 - \sum_{63} \sum_{33}^{-1} X_3$$

e) Eliminación de efectos de  $\varepsilon_{134}$  en  $\varepsilon_{63}$ 

$$\varepsilon_{134}^* = \left( E \varepsilon_{134} \varepsilon_{63}^T \right) \sum_{66.3}^{-1} \varepsilon_{63} ; \quad \varepsilon_{1346}^{H(4)} = \varepsilon_{134} - E(\varepsilon_{134} \varepsilon_{63}^T) \sum_{66.3} \varepsilon_{63} \quad (39)$$

Desarrollando (39)

$$\begin{aligned} E \varepsilon_{134} \varepsilon_{63}^T &= E \left[ \varepsilon_{13} - \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43} \right] \varepsilon_{63}^T \\ &= \sum_{16.3} - \sum_{14.3}^{-1} \sum_{44.3}^{-1} \sum_{46.3} = \sum_{16.3.4}. \end{aligned}$$

que da lugar a

$$\varepsilon_{1346}^{H(4)} = \varepsilon_{134} - \sum_{16.3.4} \sum_{66.3}^{-1} \varepsilon_{63} \quad (40)$$

## 2. Eliminación de efectos lineales en $X_2$

Un razonamiento análogo al del apartado anterior conduce al siguiente resultado

$$\varepsilon_{2357}^{H(4)} = \varepsilon_{235} - \sum_{27.3.5} \sum_{77.3}^{-1} \varepsilon_{73} \quad (41)$$

Las expresiones (40) y 41 muestran los vectores  $X_1$  y  $X_2$  desprovistos de los efectos lineales considerados en G-VIII. Antes de definir las variables canónicas entre estos vectores obtendremos su matriz de covarianzas particionada  $\sum_{11}^{H(4)}, \sum_{26}^{H(4)}, \sum_{12}^{H(4)}$

$$\begin{aligned} E \varepsilon_{1346}^{H(4)} \varepsilon_{1346}^{H(4)T} &= E \left( \varepsilon_{134} - \sum_{16.3.4} \sum_{66.3}^{-1} \varepsilon_{63} \right) \left( \varepsilon_{134} - \sum_{16.3.4} \sum_{66.3}^{-1} \varepsilon_{63} \right)^T \\ &= E \varepsilon_{134} \varepsilon_{134}^T - \sum_{16.3.4} \sum_{66.3}^{-1} \left( E \varepsilon_{63} \varepsilon_{134}^T \right) - \left( E \varepsilon_{134} \varepsilon_{63}^T \right) \sum_{66.3}^{-1} \sum_{61.3.4} \\ &\quad + \sum_{16.3.4} \sum_{66.3}^{-1} \left[ E \left( \varepsilon_{63} \varepsilon_{63}^T \right) \right] \sum_{66.3}^{-1} \sum_{61.3.4}. \quad (42) \end{aligned}$$

$$\begin{aligned} E \left( \varepsilon_{134} \varepsilon_{134}^T \right) &= E \left( \varepsilon_{13} - \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43} \right) \left( \varepsilon_{13} - \sum_{14.3} \sum_{44.3}^{-1} \varepsilon_{43} \right)^T \\ &= \sum_{11.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{41.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{41.3} + \sum_{14.3} \sum_{44.3}^{-1} \sum_{44.3} \sum_{44.3}^{-1} \sum_{41.3} \\ &= \sum_{11.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{41.3} = \sum_{11.3.4}. \quad (43) \end{aligned}$$

sustituyendo (43) y teniendo en cuenta el valor del resto de los sumandos se obtiene :

$$\begin{aligned} \sum_{11}^{H(7)} &= E \mathcal{E}_{1346}^{H(7)} \mathcal{E}_{1346}^T = \sum_{11.3.4} - \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{61.3.4} - \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{61.3.4} \\ &+ \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{66.3} \sum_{66.3}^{-1} \sum_{61.3.4} = \sum_{11.3.4} - \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{61.3.4} \quad (44) \end{aligned}$$

De la misma forma repitiendo el proceso para  $\mathcal{E}_{2357}$

$$\sum_{22}^{H(7)} = E \mathcal{E}_{2357}^{H(7)} \mathcal{E}_{2357}^T = \sum_{22.3.5} - \sum_{27.3.5} \sum_{77.3}^{-1} \sum_{72.3.5} \quad (45)$$

Finalmente obtendremos  $\sum_{12}^{H(7)}$

$$\begin{aligned} \sum_{12}^{H(7)} &= E \mathcal{E}_{1346}^{H(7)} \mathcal{E}_{2357}^T = E (\mathcal{E}_{134} - \sum_{16.3.4} \sum_{66.3}^{-1} \mathcal{E}_{63}) (\mathcal{E}_{235} - \sum_{27.3.5} \sum_{77.3}^{-1} \mathcal{E}_{73})^T \\ &= E \mathcal{E}_{134} \mathcal{E}_{235}^T - \sum_{16.3.4} \sum_{66.3}^{-1} [E(\mathcal{E}_{63} \mathcal{E}_{235}^T)] - E[\mathcal{E}_{134} \mathcal{E}_{73}^T] \sum_{77.3}^{-1} \sum_{72.3.5} \\ &+ \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{72.3.5}. \quad (46) \end{aligned}$$

calculando los sumandos de (46)

$$\begin{aligned} E \mathcal{E}_{134} \mathcal{E}_{235}^T &= E (\mathcal{E}_{13} - \sum_{14.3} \sum_{44.3}^{-1} \mathcal{E}_{43}) (\mathcal{E}_{23} - \sum_{25.3} \sum_{55.3}^{-1} \mathcal{E}_{53})^T \\ &= \sum_{12.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{42.3} - \sum_{15.3} \sum_{85.3} \sum_{92.3} + \sum_{14.3} \sum_{44.3}^{-1} \sum_{45.3} \sum_{55.3}^{-1} \sum_{52.3} \\ &= \sum_{14.3.4} + \sum_{12.3.5} - \sum_{12.3} + \sum_{14.3} \sum_{44.3}^{-1} \sum_{45.3} \sum_{55.3}^{-1} \sum_{52.3} \quad (47) \end{aligned}$$

$$E \mathcal{E}_{63} \mathcal{E}_{235}^T = E [\mathcal{E}_{63} (\mathcal{E}_{23} - \sum_{25.3} \sum_{55.3}^{-1} \mathcal{E}_{53})] = \sum_{62.3} - \sum_{66.3} \sum_{55.3}^{-1} \sum_{52.3} = \sum_{62.3.5} \quad (48)$$

$$E \mathcal{E}_{134} \mathcal{E}_{73}^T = E [(\mathcal{E}_{13} - \sum_{14.3} \sum_{44.3}^{-1} \mathcal{E}_{43}) \mathcal{E}_{73}^T] = \sum_{17.3} - \sum_{14.3} \sum_{44.3}^{-1} \sum_{47.3} = \sum_{17.3.4} \quad (49)$$

y sustituyendo en esta expresión los resultados anteriores

$$\begin{aligned} \sum_{12}^{H(7)} &= \sum_{12.3.4} + \sum_{12.3.5} - \sum_{12.3} + \sum_{14.3} \sum_{44.3}^{-1} \sum_{45.3} \sum_{55.3}^{-1} \sum_{52.3} \\ &- \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{67.3.5} - \sum_{17.3.4} \sum_{77.3}^{-1} \sum_{72.3.5} + \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{72.3.5} \quad (50) \end{aligned}$$

Las expresiones (44), (45) y (50) permiten obtener la matriz de covarianzas entre los residuos  $\Sigma_{1346}^{H(7)}$  y  $\Sigma_{2357}^{H(7)}$

$$\Sigma^{H(7)} = \begin{bmatrix} \Sigma_{11}^{H(7)} & \Sigma_{12}^{H(7)} \\ \Sigma_{21}^{H(7)} & \Sigma_{22}^{H(7)} \end{bmatrix} \quad (51)$$

Sea  $r \leq n_1$  el rango de  $\Sigma_{12}^{H(7)}$ . Se denominan vectores canónicos  $L(7)$  a las transformadas de  $\Sigma_{1346}^{H(7)}$  y  $\Sigma_{2357}^{H(7)}$

$$\begin{aligned} \mu &= L^T \Sigma_{1346}^{H(7)} \\ \nu &= M^T \Sigma_{2357}^{H(7)} \end{aligned} \quad (52)$$

cuya matriz de covarianzas es de la forma (9).

La descomposición singular de  $\Sigma_{12}^{H(7)-1/2} \Sigma_{12}^{H(7)} \Sigma_{22}^{H(7)-1/2}$  prueba la existencia de los vectores canónicos  $U$  y  $V$  con tales características siendo  $L$  y  $M$  las soluciones de los sistemas

$$\left[ \Sigma_{12}^{H(7)} \Sigma_{22}^{H(7)-1} \Sigma_{21}^{H(7)} - \lambda^2 \Sigma_{11}^{H(7)} \right] L = 0 : L^T \Sigma_{11}^{H(7)} L = 0 \quad (53)$$

$$\left[ \Sigma_{21}^{H(7)} \Sigma_{11}^{H(7)-1} \Sigma_{12}^{H(7)} - \lambda^2 \Sigma_{22}^{H(7)} \right] M = 0 : M^T \Sigma_{22}^{H(7)} M = 0 \quad (54)$$

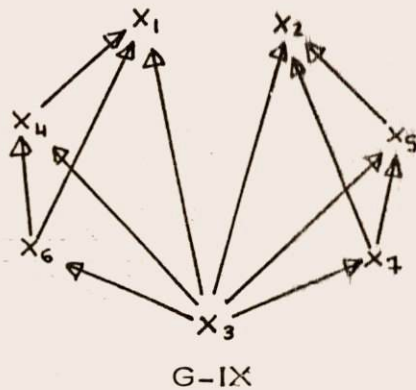
y los elementos no nulos de  $P$  las raíces cuadradas con signo positivo de las ecuaciones determinantes

$$\begin{aligned} & \left| \Sigma_{12}^{H(7)} \Sigma_{22}^{H(7)-1} \Sigma_{21}^{H(7)} - \lambda^2 \Sigma_{11}^{H(7)} \right| = 0 \\ & \left| \Sigma_{21}^{H(7)} \Sigma_{11}^{H(7)-1} \Sigma_{12}^{H(7)} - \lambda^2 \Sigma_{22}^{H(7)} \right| = 0 \end{aligned} \quad (55)$$

A dichos elementos no nulos de  $P$  les denominaremos correlaciones canónicas  $L(7)$  entre  $X_1$  y  $X_2$  siendo éstos una medida de las interrelaciones existentes entre  $X_1$  y  $X_2$  después de haber eliminado las influencias lineales supuestos en G-VIII.

### 2.3.1.2. Modelo C(7),

Cuando en G-VII  $n=3$  la dependencia entre los siete vectores de variables que intervienen da lugar al siguiente grafo:



El objeto del modelo es obtener las correlaciones y vectores canónicos entre  $X_1$  y  $X_2$  cuando se han eliminado las influencias lineales de los vectores señalados en G-IX. Procederemos como en el modelo anterior.

#### 1. Eliminación de efectos lineales en $X_1$ .

a) Eliminación en  $X_1$  del efecto de  $X_3$

$$X_1^* = \sum_{13} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{13} = X_1 - \sum_{13} \sum_{33}^{-1} X_3 \quad (56)$$

b) Eliminación en  $X_6$  del efecto lineal de  $X_3$

$$X_6^* = \sum_{63} \sum_{33}^{-1} X_3 \quad ; \quad \varepsilon_{63} = X_6 - \sum_{63} \sum_{33}^{-1} X_3 \quad (57)$$

c) Eliminación en  $\varepsilon_{13}$  del efecto lineal de  $\varepsilon_{63}$

$$\varepsilon_{13}^* = \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63} \quad ; \quad \varepsilon_{136} = \varepsilon_{13} - \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63} \quad (58)$$

Probaremos este resultado: Los coeficientes de la regresión de  $\varepsilon_{13}$  sobre  $\varepsilon_{63}$  vienen dados por la expresión:

$$\varepsilon_{13}^* = (E \varepsilon_{13} \varepsilon_{63}^T) (E \varepsilon_{63} \varepsilon_{63}^T)^{-1} \varepsilon_{63}$$

Operando en ella obtenemos:

$$\begin{aligned} E (\varepsilon_{63} \varepsilon_{63}^T) &= E (x_6 - \sum_{63} \sum_{33}^{-1} x_3) (x_6 - \sum_{63} \sum_{33}^{-1} x_3)^T \\ &= \sum_{66} - \sum_{63} \sum_{33}^{-1} \sum_{36} = \sum_{66.3} \end{aligned}$$

por otra parte

$$E \varepsilon_{13} \varepsilon_{63}^T = E (x_1 - \sum_{13} \sum_{33}^{-1} x_3) (x_6 - \sum_{63} \sum_{33}^{-1} x_3) = \sum_{16} - \sum_{13} \sum_{33}^{-1} \sum_{36} = \sum_{16.3}$$

con lo cual

$$\varepsilon_{13}^* = \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63} \quad \text{c q d}$$

- d) Eliminación en  $X_4$  del efecto lineal de  $X_3$

$$x_4^* = \sum_{43} \sum_{33}^{-1} x_3 \quad ; \quad \varepsilon_{43} = x_4 - \sum_{43} \sum_{33}^{-1} x_3 \quad (59)$$

- e) Eliminación en  $\varepsilon_{43}$  del efecto lineal de  $\varepsilon_{63}$

$$\varepsilon_{43}^* = E (\varepsilon_{43} \varepsilon_{63}^T) [E \varepsilon_{63} \varepsilon_{63}^T]^{-1} \varepsilon_{63}$$

por un razonamiento análogo al expuesto en c)

$$\varepsilon_{43}^* = \sum_{46.3} \sum_{66.3}^{-1} \varepsilon_{63}$$

$$\varepsilon_{436} = \varepsilon_{43} - \sum_{46.3} \sum_{66.3}^{-1} \varepsilon_{63} \quad (60)$$

- f) Eliminación en  $\varepsilon_{136}$  del efecto lineal de  $\varepsilon_{436}$

$$\varepsilon_{136}^* = E (\varepsilon_{136} \varepsilon_{436}^T) [E (\varepsilon_{436} \varepsilon_{436}^T)]^{-1} \varepsilon_{436} \quad (61)$$

$$\begin{aligned} E (\varepsilon_{136} \varepsilon_{436}^T) &= E (\varepsilon_{13} - \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63}) (\varepsilon_{43} - \sum_{46.3} \sum_{66.3}^{-1} \varepsilon_{63})^T \\ &= \sum_{14.3} - \sum_{16.3} \sum_{66.3}^{-1} \sum_{64.3} = \sum_{14.3.6} \quad (62) \end{aligned}$$

$$\begin{aligned}
 E \Sigma_{436} \Sigma_{436}^T &= E \left( \Sigma_{43} - \Sigma_{46.3} \Sigma_{66.3}^{-1} \Sigma_{63} \right) \left( \Sigma_{43} - \Sigma_{46.3} \Sigma_{66.3}^{-1} \Sigma_{63} \right)^T \\
 &= \Sigma_{44.3} - \Sigma_{46.3} \Sigma_{66.3}^{-1} \Sigma_{64.3} = \Sigma_{44.3.6} \quad (63)
 \end{aligned}$$

sustituyendo finalmente (62) y (63) en (61);

$$\Sigma_{136}^* = \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \Sigma_{436}$$

siendo

$$\Sigma_{1364}^{c(7)} = \Sigma_{136} - \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \Sigma_{436} \quad (64)$$

La expresión (64) muestra al vector  $X_1$  desprovisto de las influencias lineales contenidas en el grafo G-IX.

## 2. Eliminación de efectos lineales en $X_1$ .

Seguiremos las mismas etapas que en el caso anterior

$$a) \quad X_2^* = \Sigma_{23} \Sigma_{33}^{-1} X_3 \quad ; \quad \Sigma_{23} = X_2 - \Sigma_{23} \Sigma_{33}^{-1} X_3 \quad (65)$$

$$b) \quad X_7^* = \Sigma_{73} \Sigma_{33}^{-1} X_3 \quad ; \quad \Sigma_{73} = X_7 - \Sigma_{73} \Sigma_{33}^{-1} X_3 \quad (66)$$

$$c) \quad \Sigma_{23}^* = \Sigma_{27.3} \Sigma_{77.3}^{-1} \Sigma_{73} \quad ; \quad \Sigma_{237} = \Sigma_{23} - \Sigma_{27.3} \Sigma_{77.3}^{-1} \Sigma_{73} \quad (67)$$

$$d) \quad X_5^* = \Sigma_{53} \Sigma_{33}^{-1} X_3 \quad ; \quad \Sigma_{53} = X_5 - \Sigma_{53} \Sigma_{33}^{-1} X_3 \quad (68)$$

$$e) \quad \Sigma_{53}^* = \Sigma_{57.3} \Sigma_{77.3}^{-1} \Sigma_{73} \quad ; \quad \Sigma_{573} = \Sigma_{53} - \Sigma_{57.3} \Sigma_{77.3}^{-1} \Sigma_{73} \quad (69)$$

$$\begin{aligned}
 f) \quad \Sigma_{237}^* &= \Sigma_{25.3.7} \Sigma_{55.3.7}^{-1} \Sigma_{537} \\
 \Sigma_{23745}^{c(7)} &= \Sigma_{237} - \Sigma_{25.3.7} \Sigma_{55.3.7}^{-1} \Sigma_{537} \quad (70)
 \end{aligned}$$

La expresión (70) muestra el vector  $X_2$  desprovisto de las influencias lineales contenidas en el grafo G- IX.

Denominaremos a la matriz de covarianzas entre  $\varepsilon_{1364}^{(7)}$ ,  $\varepsilon_{2336}^{(7)}$  :  $\Sigma^{(7)}$

Esta matriz particionada en base a las dimensiones de estos vectores la denotaremos:

$$\Sigma^{(7)} = \begin{bmatrix} \Sigma_{11}^{(7)} & \Sigma_{12}^{(7)} \\ \Sigma_{21}^{(7)} & \Sigma_{22}^{(7)} \end{bmatrix} \quad (71)$$

$\Sigma_{11}^{(7)}$   $\Sigma_{12}^{(7)}$   $\Sigma_{22}^{(7)}$  pueden obtenerse teniendo en cuenta (64) y (70):

$$\begin{aligned} \Sigma_{11}^{(7)} &= E \varepsilon_{1364} \varepsilon_{1364}^T = E (\varepsilon_{136} - \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \varepsilon_{436}) (\varepsilon_{136} - \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \varepsilon_{436})^T \\ &= E \varepsilon_{136} \varepsilon_{136}^T - \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} [E \varepsilon_{436} \varepsilon_{136}^T] - [E \varepsilon_{136} \varepsilon_{436}^T] \Sigma_{44.3.6}^{-1} \Sigma_{41.3.6} \\ &+ \Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} [E \varepsilon_{436} \varepsilon_{436}^T] \Sigma_{44.3.6}^{-1} \Sigma_{41.3.6} \quad (72) \end{aligned}$$

Calcularemos los sumandos de la expresión (72)

$$E \varepsilon_{136} \varepsilon_{136}^T = \Sigma_{11.3} - \Sigma_{16.3} \Sigma_{66.3}^{-1} \Sigma_{61.3} = \Sigma_{11.3.6} \quad (73)$$

$$E \varepsilon_{436} \varepsilon_{436}^T = \Sigma_{41.3} - \Sigma_{46.3} \Sigma_{66.3}^{-1} \Sigma_{61.3} = \Sigma_{41.3.6}$$

y sustituyendo este resultado en el segundo sumando de (72) obtenemos:

$$\Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \Sigma_{41.3.6} \quad (74)$$

$$E \varepsilon_{136} \varepsilon_{436}^T = \Sigma_{14.3} - \Sigma_{16.3} \Sigma_{66.3}^{-1} \Sigma_{64.3} = \Sigma_{14.3.6}$$

que sustituido en el tercer sumando de (72) da lugar a:

$$\Sigma_{14.3.6} \Sigma_{44.3.6}^{-1} \Sigma_{41.3.6} \quad (75)$$

finalmente:

$$E \varepsilon_{436} \varepsilon_{436}^T = \sum_{44.2} - \sum_{46.3} \sum_{66.3}^{-1} \sum_{64.3} = \sum_{44.3.6} \quad (76)$$

resultado que sustituido en el cuarto sumando de (72) da lugar a:

$$\sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{41.3.6} \quad (77)$$

Llevando (73), (74), (75) y (77) a (72)

$$\begin{aligned} \sum_{11}^{c(7)} &= \sum_{11.3.6} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{41.3.6} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{41.3.6} + \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{41.3.6} \\ &= \sum_{11.3.6} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{41.3.6} = \sum_{11.3.6.4} \quad (78) \end{aligned}$$

La simetría de G-IX permite escribir

$$\sum_{22}^{c(7)} = \sum_{22.3.7} - \sum_{25.3.7} \sum_{55.3.7}^{-1} \sum_{52.3.7} = \sum_{22.3.7.5} \quad (79)$$

Por último calcularemos  $\sum_{12}^{c(7)}$

$$\begin{aligned} \sum_{12}^{c(7)} &= E \varepsilon_{1364}^{c(7)} \varepsilon_{2357}^{T c(7)} = E \left( \varepsilon_{136} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \varepsilon_{476} \right) \left( \varepsilon_{237} - \sum_{25.3.7} \sum_{56.3.7} \varepsilon_{537} \right)^T \\ &= E \varepsilon_{136} \varepsilon_{237}^T - \left[ E \varepsilon_{136} \varepsilon_{537}^T \right] \sum_{55.3.7}^{-1} \sum_{52.3.7} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \left[ \right. \\ &\quad \left. \left[ E \varepsilon_{436} \varepsilon_{237}^T \right] + \sum_{14.3.6} \sum_{44.3.6}^{-1} \left[ E \left( \varepsilon_{436} \varepsilon_{537}^T \right) \right] \sum_{55.3.7} \sum_{52.3.7} \right] \quad (80) \end{aligned}$$

Detallaremos el cálculo de cada uno de los cuatro sumandos de (80)

$$\begin{aligned}
 E \varepsilon_{136} \varepsilon_{237}^T &= E (\varepsilon_{13} - \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63}) (\varepsilon_{23} - \sum_{27.3} \sum_{77.3}^{-1} \varepsilon_{63})^T \\
 &= \sum_{12.3} - \sum_{16.3} \sum_{66.3}^{-1} \sum_{62.3} - \sum_{17.3} \sum_{77.3}^{-1} \sum_{72.3} \\
 &\quad + \sum_{16.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{72.3} \quad (81)
 \end{aligned}$$

$$\begin{aligned}
 E \varepsilon_{136} \varepsilon_{537}^T &= E (\varepsilon_{13} - \sum_{16.3} \sum_{66.3}^{-1} \varepsilon_{63}) (\varepsilon_{53} - \sum_{57.3} \sum_{77.3}^{-1} \varepsilon_{73})^T \\
 &= \sum_{15.3} - \sum_{17.3} \sum_{77.3}^{-1} \sum_{75.3} - \sum_{16.3} \sum_{66.3}^{-1} \sum_{65.3} \\
 &\quad + \sum_{16.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{75.3}. \quad (82)
 \end{aligned}$$

$$\begin{aligned}
 E \varepsilon_{436} \varepsilon_{237}^T &= E (\varepsilon_{43} - \sum_{46.3} \sum_{66.3}^{-1} \varepsilon_{63}) (\varepsilon_{23} - \sum_{27.3} \sum_{77.3}^{-1} \varepsilon_{73}) \\
 &= \sum_{42.3} - \sum_{46.3} \sum_{66.3}^{-1} \sum_{67.3} - \sum_{47.3} \sum_{77.3}^{-1} \sum_{72.3} \\
 &\quad + \sum_{46.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{72.3}. \quad (83)
 \end{aligned}$$

$$\begin{aligned}
 E \varepsilon_{436} \varepsilon_{537}^T &= E (\varepsilon_{43} - \sum_{46.3} \sum_{66.3}^{-1} \varepsilon_{63}) (\varepsilon_{53} - \sum_{57.3} \sum_{77.3}^{-1} \varepsilon_{73})^T \\
 &= \sum_{45.3} - \sum_{46.3} \sum_{66.3}^{-1} \sum_{65.3} - \sum_{47.3} \sum_{77.3}^{-1} \sum_{75.3} \\
 &\quad + \sum_{46.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{75.3} \quad (84)
 \end{aligned}$$

(81) puede expresarse abreviadamente

$$\sum_{12.3.6} + \sum_{12.3.7} + \sum_{16.3} \sum_{66.3}^{-1} \sum_{62.3} \sum_{77.3}^{-1} \sum_{72.3} - \sum_{12.3} \quad (85)$$

igualmente (82)

$$\sum_{15.3.6} + \sum_{15.3.7} + \sum_{16.3} \sum_{66.3}^{-1} \sum_{62.3} \sum_{77.3}^{-1} \sum_{75.3} - \sum_{15.3} \quad (86)$$

(83) da lugar a

$$\sum_{42.3.6} + \sum_{42.3.7} + \sum_{46.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{72.3} - \sum_{42.3} \quad (87)$$

y finalmente (84)

$$\sum_{45.3.6} + \sum_{45.3.7} + \sum_{46.3} \sum_{66.3}^{-1} \sum_{67.3} \sum_{77.3}^{-1} \sum_{75.3} - \sum_{45.3} \quad (88)$$

Sustituyendo (85), (86), (87), y (88) en (80) se obtiene la expresión de  $\sum_{12}^{C(7)}$

$$\begin{aligned} \sum_{12}^{C(7)} &= \sum_{12.3.6} + \sum_{12.3.7} + \sum_{16.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{17.3} \\ &\quad - \sum_{14.3.6} \sum_{44.3.6}^{-1} \left[ \sum_{47.3.6} + \sum_{42.3.7} + \sum_{46.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{42.3} \right] \\ &\quad - \left[ \sum_{18.3.7} + \sum_{15.3.6} + \sum_{16.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{18.3} \right] \sum_{55.3.7} \sum_{52.3.7} \\ &\quad + \sum_{14.3.6} \sum_{44.3.6} \left[ \sum_{45.3.6} + \sum_{48.3.7} + \sum_{46.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{45.3} \right] \sum_{55.3.7} \sum_{52.3.7} \\ &= \sum_{12.3.6.4} + \sum_{12.3.7.5} - \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{42.3.7} - \sum_{15.3.6} \sum_{56.3.7}^{-1} \sum_{88.3.7} + \sum_{14.3.6} \sum_{44.3.6}^{-1} \\ &\quad \sum_{49.3.6} \sum_{57.3.7}^{-1} \sum_{52.3.7} + \sum_{14.3.6} \sum_{44.3.6}^{-1} \sum_{45.3.7} \sum_{58.3.7}^{-1} \sum_{52.3.7} + \sum_{14.3.6} \sum_{44.3.6} \left[ \sum_{46.3} \sum_{CC.3}^{-1} \sum_{C7.3} \right. \\ &\quad \left. \sum_{77.3}^{-1} \sum_{77.3} - \sum_{45.3} \right] \sum_{55.3.7} \sum_{52.3.7} - \left[ \sum_{46.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{45.3} \right] \sum_{56.3.7} \sum_{52.3.7} - \\ &\quad - \sum_{14.3.6} \sum_{44.3.6}^{-1} \left[ \sum_{46.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{42.3} \right] + \sum_{16.3} \sum_{CC.3}^{-1} \sum_{C7.3} \sum_{77.3}^{-1} \sum_{77.3} - \sum_{18.3} \end{aligned}$$

Sea  $r \leq n_1$  el rango de  $\sum_{12}^{C(7)}$ . Se denominan vectores canónicos  $C(7)$  a las transformadas :

$$\begin{aligned} \mu &= L^T \varepsilon_{1364}^{C(7)} \\ \nu &= M^T \varepsilon_{2374}^{C(7)} \end{aligned} \quad (89)$$

cuya matriz de covarianzas es de la forma (9).

La descomposición singular de  $\sum_{11}^{C(7)} \sum_{12}^{C(7)} \sum_{22}^{C(7)}$  prueba la existencia de tales vectores con las características señaladas siendo  $L$  y  $M$  las soluciones de los siguientes sistemas:

$$\left[ \sum_{12}^{C(7)} \sum_{22}^{C(7)}^{-1} \sum_{21}^{C(7)} - \lambda^2 \sum_{11}^{C(7)} \right] L = 0 \quad : \quad L^T \sum_{11} L = I_{n_1} \quad (90)$$

$$\left[ \sum_{21}^{C(7)} \sum_{11}^{C(7)}^{-1} \sum_{12}^{C(7)} - \lambda^2 \sum_{22}^{C(7)} \right] M = 0 \quad : \quad M^T \sum_{22} M = I_{n_2} \quad (91)$$

y los elementos no nulos de  $P$  las soluciones de las ecuaciones determinantes

$$\begin{aligned} & \left| \begin{matrix} c^{(7)} & c^{(7)} & c^{(7)} \\ \Sigma_{12} & \Sigma_{22}^{-1} & \Sigma_{21} \end{matrix} - \lambda^2 \Sigma_{11} \right| = 0 \\ & \left| \begin{matrix} c^{(7)} & c^{(7)} & c^{(7)} \\ \Sigma_{21} & \Sigma_{11}^{-1} & \Sigma_{12} \end{matrix} - \lambda^2 \Sigma_{22} \right| = 0 \end{aligned} \quad (92)$$

A dichos elementos no nulos de  $P$  los denominaremos correlaciones canónicas  $C(7)$  entre  $X_1$  y  $X_2$  y miden la interdependencia entre estos vectores cuando se han eliminado las influencias lineales supuestas en  $G-IX$

### 2. 3. 2. Modelos canónicos parciales $L(2n+1)$ y $C(2n+1)$

Como final de este capítulo presentaremos dos modelos canónicos parciales que constituyen dos posibles extensiones de todos los modelos incluidos en 2. 2.

En ambos casos consideraremos  $2n+1$  vectores de variables aleatorias  $(x_1^T, x_2^T, \dots, x_{2n+1}^T)$  distribuidas conjuntamente de dimensiones  $n_1, n_2, \dots, n_{2n+1}$  :  $n_1 \leq n_2$  y de parámetros  $(0, \Sigma)$ .

La matriz de covarianzas  $\Sigma$  puede expresarse particionada en base a las dimensiones de los vectores considerados

$$\Sigma = \left[ \Sigma_{ij} \right] \quad ij = 1, 2, \dots, 2n+1 \quad (93)$$

#### 2. 3. 2. 1. Modelo canónico $L(2n+1)$

Consideremos que las dependencias entre los  $2n+1$  vectores de variables vienen indicadas en el grafo  $G-VI$ .

Como en todos los casos anteriores obtendremos las correlaciones y variables canónicas entre  $X_1$  y  $X_2$  después de eliminar las influencias lineales señaladas en dicho grafo.

Obtenida la estructura de los residuos en  $L(3)$ ,  $L(5)$  y  $L(7)$  estos resultados nos permiten conseguir la de  $L(2n-1)$  y en base a ellos la de  $L(2n+1)$ .

Supongamos dados los residuos de  $L(2n-1)$

$$\begin{aligned} \varepsilon_{1346 \dots (2n-4)(2n-2)}^{H(2n-1)} &= \varepsilon_{1346 \dots (2n-4)} - \left[ \sum_{1(2n-2).3.4} - \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{6(2n-2).3} \right. \\ &- \sum_{18.3.4} \sum_{88.3}^{-1} \sum_{8(2n-2).3} - \dots - \sum_{1(2n-4).3.4} \sum_{(2n-4)(2n-4).3}^{-1} \sum_{(2n-4)(2n-7).3} \\ &\left. + \sum_{16.3.4} \sum_{66.3}^{-1} \sum_{68.3} \sum_{88.3}^{-1} \sum_{810.3} \dots \sum_{(2n-4)(2n-4)}^{-1} \sum_{(2n-4)(2n-2)} \right] \sum_{(2n-2)(2n-2).3}^{-1} \varepsilon_{(2n-2)3} \quad (94) \end{aligned}$$

$$\begin{aligned} \varepsilon_{1357 \dots (2n-3)(2n-1)}^{H(2n-1)} &= \varepsilon_{1357 \dots (2n-3)} - \left[ \sum_{1(2n-1).3.5} - \sum_{17.3.5} \sum_{77.3}^{-1} \sum_{7(2n-1).3} \right. \\ &- \sum_{19.3.5} \sum_{99.3}^{-1} \sum_{9(2n-1).3} - \dots - \sum_{1(2n-3).3.5} \sum_{(2n-3)(2n-3).3}^{-1} \sum_{(2n-3)(2n-1).3} \\ &\left. + \sum_{17.3.5} \sum_{77.3}^{-1} \sum_{79.3} \sum_{99.3}^{-1} \sum_{911.3} \dots \sum_{(2n-3)(2n-3).3}^{-1} \sum_{(2n-3)(2n-1).3} \right] \sum_{(2n-1)(2n-1).3}^{-1} \varepsilon_{(2n-1)3} \quad (95) \end{aligned}$$

A partir de (94) y (95) deduciremos las expresiones de los residuos de  $L(2n+1)$ .

Dichos vectores de variables se obtienen eliminando de (94) y (95) los efectos lineales de  $\varepsilon_{2n3}$  y  $\varepsilon_{(2n+1)3}$  respectivamente.

$$E_{1346 \dots 2n}^{H(2n+1)} = E_{1346 \dots 2n-2} - E_{1346 \dots 2n-2}^* \quad (96)$$

$$\text{si } E_{2n,3} = X_{2n} - \sum_{2n,3}^{-1} X_3$$

$$\begin{aligned} E_{1346 \dots (2n-2)}^* &= E[E_{1346 \dots (2n-2)} E_{2n,3}^T] [E E_{2n,3} E_{2n,3}^T]^{-1} E_{2n,3} \\ &= \left[ \sum_{1(2n),3,4} - \sum_{16,3,4} \sum_{66,3}^{-1} \sum_{6(2n),3} \dots - \sum_{1(2n-2),3,4} \sum_{(2n-2)(2n-1),3}^{-1} \sum_{(2n-2)(2n-1),3} \right. \\ &\quad \left. + \sum_{16,3,4} \sum_{66,3}^{-1} \sum_{68,3} \dots \sum_{(2n-2)(2n-1)}^{-1} \sum_{(2n-2)(2n-1)} \right] \\ &\quad \left. \sum_{(2n)(2n)}^{-1} E_{(2n),3} \right] \quad (97) \end{aligned}$$

Sustituyendo (97) en (96) se llega a  $E_{1346 \dots 2n}^{H(2n+1)}$

De forma totalmente análoga obtendríamos  $E_{1357 \dots (2n+1)}$

$$\text{Si } E_{1357 \dots 2n+1}^{H(2n+1)} = E_{1357 \dots 2n-1} - E_{1357 \dots 2n-1}^* \quad (98)$$

$$E_{(2n+1),3} = X_{2n+1} - \sum_{(2n+1),3}^{-1} X_3$$

$$E_{1345 \dots 2n-1}^* = E[E_{1357 \dots 2n-1} E_{(2n+1),3}^T] [E[E_{2n+1,3} E_{2n+1,3}^T]]^{-1} E_{2n+1,3}$$

que conduce a

$$\begin{aligned} E_{1345 \dots 2n-1}^* &= \sum_{1(2n+1),3,5} - \sum_{17,3,5} \sum_{77,3}^{-1} \sum_{7(2n+1),3} \dots - \sum_{1(2n-1),3,5} \sum_{(2n-1)(2n-1),3}^{-1} \sum_{(2n-1)(2n-1),3} \\ &\quad + \sum_{17,3,5} \sum_{77,3}^{-1} \sum_{79,3} \sum_{99,3}^{-1} \dots \sum_{(2n-1)(2n-1)}^{-1} \sum_{(2n-1)(2n-1)} \left] \sum_{(2n+1)(2n+1)}^{-1} E_{(2n+1),3} \right] \quad (99). \end{aligned}$$

expresión que llevada a (98) permite obtener  $E_{1357 \dots 2n+1}^{H(2n+1)}$

Las expresiones (96) y (98) previa sustitución en ellas de (97) y (99) permiten calcular la matriz de covarianzas entre  $X_1$  y  $X_2$  desprovistos de los efectos lineales supuestos en G-VI.

$$\Sigma_{11}^{H(2n+1)} = E E_{1346 \dots 2n}^{H(2n+1)} E_{1346 \dots 2n}^{H(2n+1)} \quad (100)$$

$$\sum_{12}^{H(2n+1)} = E E_{1246 \dots 2n}^{H(2n+1)} E_{2357 \dots 2n+1}^T H(2n+1) \quad (101)$$

$$\sum_{22}^{H(2n+1)} = E E_{2357 \dots 2n+1}^{H(2n+1)} E_{2357 \dots 2n+1}^T H(2n+1) \quad (102)$$

Sea  $r \leq n_1$  el rango de  $\sum_{12}^{H(2n+1)}$ . Denominaremos vectores canónicos  $L(2n+1)$  a las obtenidas mediante las siguientes transformaciones

$$\begin{aligned} \mu &= L^T E_{1346 \dots 2n}^{H(2n+1)} \\ \nu &= M^T E_{2357 \dots 2n+1}^{H(2n+1)} \end{aligned} \quad (103)$$

cuya matriz de covarianzas es de la forma (9).

La existencia de tales variables con las características señaladas puede probarse mediante la descomposición singular de la matriz:

$$\sum_{11}^{H(2n+1)} - \frac{1}{2} \sum_{12}^{H(2n+1)} \sum_{22}^{H(2n+1)} - \frac{1}{2} \quad (104)$$

siendo  $L$  y  $M$  soluciones de sistemas análogos a los casos anteriores y  $P$  la matriz cuyos elementos no nulos son la raíz cuadrada de las soluciones de las ecuaciones determinantes derivadas de los sistemas anteriores. A dichos elementos no nulos de  $P$  les denominaremos correlaciones canónicas  $L(2n+1)$  entre  $X_1$  y  $X_2$ ; dichas correlaciones son una medida de la interdependencia entre tales vectores después de eliminadas las influencias lineales de los vectores considerados en G-VI.

### 2.3.2.2. Modelo canónico C(2n+1).

Otro tipo de dependencias que pueden suponerse en los vectores  $X_1$  y  $X_2$  es la expresada en el grafo G-VII.

Teniendo en cuenta la fórmula (64) el vector  $X_1$  después de eliminadas las influencias lineales admitidas lo denotaremos -

$$E_{13(2n)(2n-2) \dots 10.8.6.4}^{C(2n+1)} \quad (105)$$

Igualmente, teniendo en cuenta (70) el vector  $X_2$  después de eliminadas las influencias lineales lo expresaremos:

$$E_{23(2n+1)(2n-1) \dots 9.7.5}^{C(2n+1)} \quad (106)$$

Las fórmulas (64) y (70) permiten extender sus resultados para el cálculo de los residuos anteriores

$$E_{13(2n)(2n-2) \dots 10.8.6} = E_{13(2n)(2n-1) \dots 8.6} - E_{13(2n)(2n-2) \dots 8.6}^*$$

$$E_{13(2n)}^* = 8.6 \left[ \sum_{14.3(2n)} - 8.6 \sum_{44.3 \dots 8.6}^{-1} \right] E_{43(2n)} - 8.6 \quad (107)$$

De la misma forma se calcularían  $E_{13(2n)} - 10.8.6$  y  $E_{43(2n)} - 8.6$  hasta obtener los últimos residuos en función únicamente de los vectores  $X_1 \dots X_{2n+1}$  y de los elementos de su matriz de covarianzas particionada. Igual proceso se seguiría respecto a  $E_{23(2n+1)(2n-1) \dots 9.7.5}$

La matriz de covarianzas entre (105) y (106) en forma particionada viene dada por las siguientes expresiones:

$$\sum_{11}^{C(2n+1)} = E \varepsilon_{13}^{C(2n+1)} \dots 10.8.6.4 \varepsilon_{13}^T \dots 10.8.6.4 = \sum_{11.3.(2n)} \dots 10.8.6.4. \quad (108)$$

$$\sum_{22}^{C(2n+1)} = E \varepsilon_{23}^{C(2n+1)} \dots 11.9.7.5 \varepsilon_{23}^T \dots 11.9.7.5 = \sum_{22.3.(2n+1)} \dots 11.9.7.5. \quad (109)$$

$$\sum_{12}^{C(2n+1)} = E \varepsilon_{13} \dots 10.8.6.4 \varepsilon_{23}^T \dots 11.9.7.5 \quad (110)$$

(110) puede evaluarse teniendo en cuenta (107) y la relación análoga para el residuo en  $X_2$ .

Sea  $r \leq n_1$  el rango de  $\sum_{12}^{C(2n+1)}$ . Denominaremos vectores canónicos  $C(2n+1)$  entre los vectores  $X_1$  y  $X_2$  a las obtenidas a través de las siguientes transformaciones

$$\begin{aligned} \mu &= L^T \varepsilon_{13.(2n)} \dots 10.8.6.4 \\ \nu &= M^T \varepsilon_{23.(2n+1)} \dots 11.9.7.5 \end{aligned} \quad (111)$$

cuya matriz de covarianzas es de la forma (9).

La descomposición singular de la matriz

$$\sum_{11}^{C(2n+1)} \sum_{12}^{C(2n+1)} \sum_{22}^{C(2n+1)} \quad (112)$$

permite probar la existencia de tales vectores con las características ya señaladas siendo las matrices  $L$  y  $M$  soluciones de sistemas análogos a los de los apartados anteriores y los elementos no nulos de la matriz  $P$  las raíces cuadradas con signo positivo de las ecuaciones determinantes asociadas a dichos sistemas. A dichos elementos no nulos de  $P$  los denominaremos correlaciones canónicas  $C(2n+1)$  entre los vectores  $X_1$  y  $X_2$ .

#### 2.4. Los modelos parciales en las distribuciones normales multivariantes.

Una de las hipótesis utilizadas en todos los modelos canónicos expuestos en este capítulo ha sido la de suponer que los vectores de variables aleatorias que en ellos intervienen poseen una distribución de probabilidades conjunta en la que implícitamente se ha supuesto la existencia de momentos de primero y segundo orden.

Cuando la distribución de probabilidades es normal multivariante los resultados obtenidos adquieren un especial significado en términos de los parámetros de las distribuciones condicionadas.

Citaremos el siguiente párrafo de Cramer, H. <sup>(1)</sup> relativo al coeficiente de correlación parcial entre dos variables "el coeficiente de correlación parcial posee una importante propiedad para la distribución normal. No solamente pone de manifiesto la correlación entre los residuos sino también la correlación entre dos variables para valores fijos del resto".

Esta propiedad es trasladable a vectores de variables aleatorias pudiéndose dar idéntico significado a las correlaciones canónicas de los modelos parciales en términos de las correlaciones canónicas en distribuciones condicionadas. Así, en el primer modelo estudiado, las correlaciones canónicas parciales son las correlaciones canónicas entre  $x_1/x_3$  y  $x_2/x_3$ . Igual propiedad poseen el resto de modelos estudiados en relación a los coeficientes de correlación canónica entre vectores condicionados.

---

(1) Cramer, H. [ 9 ] pág. 364.

En base a esta idea daremos una interpretación de las matrices de covarianzas utilizadas en tales modelos:

$\Sigma_{11.2.(2n).(2n-2) \dots 2.6.4}$  nos indica la matriz de covarianzas de la distribución de  $X_1 / X_3 X_{2n} \dots X_6 X_4$

$\Sigma_{22.3.(2n+1) \dots 4.7.5}$  nos indica la matriz de covarianzas de la distribución de  $X_2 / X_3 X_{2n+1} \dots X_7 X_5$ . Idéntico significado tienen todas las expresiones análogas utilizadas en este capítulo.

CAPITULO III: Los modelos canónicos generalizados. Modelos par-  
ciales generalizados.

### 3. 1. Introducción

Dedicaremos este capítulo a exponer una serie de criterios en base a los cuales puede generalizarse a más de dos vectores de variables el modelo canónico expuesto en el capítulo 1.

Una primera cuestión a tener en cuenta y que probaremos a continuación es que el modelo canónico definido en 1. 2. no puede extenderse cumpliendo unos requisitos análogos a los exigidos en el caso de dos vectores de variables a la matriz de correlaciones de los vectores canónicos. En general, no puede lograrse que los vectores canónicos considerados individualmente sean de matriz de correlación unidad y que al considerarlos a pares únicamente sean no nulas las correlaciones correspondientes a componentes de igual ordinal.

Fijado este punto, en los modelos canónicos generalizados trataremos de encontrar los vectores canónicos con matriz de correlaciones que mejor se aproximen a la situación descrita en el párrafo anterior.

El desarrollo de este tipo de generalizaciones son debidas casi en su totalidad a Paul Horst. Este autor, en 1961 publicó dos artículos en los cuales propone cuatro procedimientos de llevar a cabo una generalización del modelo canónico con las características ya señaladas. En el primero de ellos <sup>(1)</sup> "Relations among m sets of measures", basándose en la propiedad 1. 5. 3. del modelo de Hotelling expone el criterio de obtención de los vectores canónicos haciendo cumplir a éstas una propiedad análoga pero referida a m vectores de variables; al método de obtención de estos vectores lo denomi-

---

(1) Horst, P. [19], pág. 129-149.

na "Método de máxima correlación". En un segundo artículo <sup>(1)</sup> desarrolla tres nuevos procedimientos de obtener las correlaciones y vectores canónicos generalizados siendo denominados "Método de aproximación a una matriz de rango 1", "Método oblicuo de máxima varianza" y "Método ortogonal de máxima varianza". En este capítulo justificaremos algunas de tales denominaciones.

Este autor ha incluido en una publicación posterior sobre análisis factorial <sup>(2)</sup> estos cuatro criterios de generalizar el modelo canónico bajo el epígrafe común "análisis factorial de grupos múltiples", situándose así en la línea doctrinal que considera al análisis canónico como un análisis factorial en varios vectores de variables; la cuestión es polémica puesto que ambos modelos persiguen objetivos diferentes en la obtención de las variables transformadas; únicamente dejaremos constancia aquí de esta línea de opinión en torno al análisis canónico.

Este capítulo estará dedicado a exponer dos de las generalizaciones de Paul Horst, así como a comparar los resultados que de ellas se derivan. El motivo que no incluyamos el resto es debido a que su exposición originaria está realizada en términos de matrices de datos no siendo posible trasladar tales criterios a matrices de covarianzas. También introduciremos una extensión de estos modelos a la que denominaremos análisis canónico parcial generalizado. Esta consiste en obtener las correlaciones y vectores canónicos entre más de dos vectores de variables cuando en ellos se han eliminado las influencias lineales de otros. Estos modelos que aquí exponemos suponen una nueva línea de trabajo en los modelos canónicos generalizados cerrando con ello las extensiones posibles del análisis canónico.

(1) Horst, P. [ 20 ] pág. 331-347.

(2) Horst, P. [ 21 ] pág. 565-595.

En las tres primeras secciones de este capítulo supondremos los vectores de variables aleatorias  $(X_1, X_2, \dots, X_m)$  definidas conjuntamente y de parámetros  $(0, \Sigma)$ ; suponemos que las dimensiones de tales vectores son  $n_1, n_2, \dots, n_m$ . La matriz de covarianzas particionada en base a tales dimensiones la denotaremos

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1m} \\ \Sigma_{21} & \Sigma_{22} & \dots & \Sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{m1} & \Sigma_{m2} & \dots & \Sigma_{mm} \end{bmatrix} \quad (1)$$

3. 2. Imposibilidad de generalización del modelo canónico manteniéndose las principales propiedades del modelo de dos vectores de variables.

Una extensión de la definición 1. 2. nos llevaría a buscar vectores canónicos ;

$$\begin{aligned} \mu &= L^T x_1 \\ \nu &= M^T x_2 \\ &\vdots \\ w &= N^T x_w \end{aligned} \quad (2)$$

cuya matriz de covarianzas fuese de la siguiente estructura:

$$\Sigma_{\nu\nu\dots w} = \begin{bmatrix} I_{n_1} & P_{12} & \dots & P_{1w} \\ P_{21} & I_{n_2} & \dots & P_{2w} \\ \vdots & \vdots & \ddots & \vdots \\ P_{w1} & P_{w2} & \dots & I_{n_w} \end{bmatrix} \quad (3)$$

en donde las matrices  $P_{ij}$  adoptasen la forma 1. 2. (5)

Vamos a probar que en general no existen los vectores (2) con matriz de covarianzas (3). Lo haremos para el caso de tres vectores de variables  $x_1, x_2, x_3$ .

Sean  $(x_1, x_2, x_3)$  vectores de v. a. de parámetros  $(0, \Sigma)$  definidos conjuntamente de dimensiones  $n_1, n_2, n_3$ . En cada pareja de vectores pueden obtenerse separadamente las correlaciones y vectores canónicos según se expuso en 1. 3. Consideremos  $x_1$  y  $x_2$

$$1. 3. (23) \quad \mu = A^T t_1^T x_1$$

$$1. 3. (24) \quad \nu = B^T t_2^T x_2 \quad : \quad t_1^T \Sigma_{12} (t_2^T)^T = A P_{12} B^T \quad (4)$$

siendo A y B matrices ortogonales.

Obteniendo las variables canónicas en  $x_1$  y  $x_3$

$$\begin{aligned} \mu &= A^T t_1^{-1} x_1 \\ W &= C^T t_3^{-1} x_3 \end{aligned} \quad : \quad t_1^{-1} \Sigma_{13} (t_3^{-1})^T = A P_{13} C^T \quad (5)$$

La matriz  $A$  de (4) y (5) no puede satisfacer a la vez los sistemas de ecuaciones

$$\begin{aligned} & [ t_1^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} (t_1^{-1})^T - \lambda^2 I ] A = 0 \quad (6) \\ 1.3.(17) & [ t_1^{-1} \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} (t_1^{-1})^T - \lambda^2 I ] A = 0 \end{aligned}$$

ya que implicaría:

$$\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = \Sigma_{13} \Sigma_{33}^{-1} \Sigma_{31} \quad (7)$$

que en general no se cumple.

Ante la imposibilidad de obtener vectores transformados (2) con matriz de covarianzas (3) en los modelos canónicos generalizados se tratará de hallar nuevas variables con una matriz de covarianzas cuya diagonal principal coincida con la de (3) y las matrices particionadas no diagonales se aproximen a las de ésta.

Según los criterios de obtener tales aproximaciones desarrollaremos dos procedimientos de generalizar el modelo canónico. Ambos coinciden cuando se aplican a dos vectores de referencia.

### 3.3. Método de máxima correlación,

En la propiedad 1.5.3. del modelo de Hotelling hemos demostrado cómo las primeras variables canónicas tienen una matriz de covarianzas cuya suma de elementos no diagonales es máxima en relación a la matriz de otra pareja de variables transformadas lineales y con varianza unidad. Asimismo, las segundas variables canónicas tienen una matriz de covarianzas cuya suma de elementos no diagonales es máxima con relación a la matriz de otra pareja de variables transformadas incorreladas con las primeras variables canónicas y de varianza 1. Análoga propiedad cumple las  $r$ -ésimas variables canónicas.

Paul Horst basándose en esta propiedad ha propuesto un criterio de generalización del modelo canónico. Según este criterio las primeras variables canónicas de un supervector de variables aleatorias  $(X_1, X_2, \dots, X_m)$  definidas conjuntamente de parámetros  $(0, \Sigma)$  y de dimensiones  $n_1, n_2, \dots, n_m$  son las variables transformadas:

$$\begin{aligned} U_1 &= l_1^T X_1 \\ v_1 &= m_1^T X_2 \\ &\vdots \\ W_1 &= n_1^T X_m \end{aligned} \quad (8)$$

normalizadas en varianza que poseen una matriz de covarianzas cuya suma de elementos no diagonales es máxima.

Las segundas variables canónicas son las variables transformadas

$$\begin{aligned} U_2 &= l_2^T X_1 \\ v_2 &= m_2^T X_2 \\ &\vdots \\ W_2 &= n_2^T X_m \end{aligned} \quad (9)$$

normalizadas en varianza e incorreladas con las anteriores dentro de cada nuevo vector  $(U_1, U_2, \dots, U_r, V_1, V_2, \dots, V_r, W_1, W_2, \dots, W_r)$  cuya matriz de co

varianzas tiene unos elementos no diagonales cuya suma es un máximo con relación a la matriz de covarianzas de otras variables obtenidas por transformación de  $(x_1, x_2, \dots, x_m)$  e incorreladas con las primeras variables canónicas, dentro de cada grupo.

Igualmente las  $r$ -ésimas variables canónicas son las transformadas

$$\begin{aligned} \mu_r &= p_r^T x_1 \\ v_r &= w_r^T x_2 \\ &\vdots \\ w_r &= n_r^T x_m \end{aligned} \quad (10)$$

con varianza unidad e incorreladas con las anteriores dentro de cada nuevo vector  $(\mu_1, \mu_2, \dots, \mu_r, \dots, w_1, w_2, \dots, w_r)$  cuya matriz de covarianzas tiene unos elementos no diagonales cuya suma es un máximo.

Antes de obtener (8), (9) y (10) es conveniente realizar en  $(x_1, x_2, \dots, x_m)$  una transformación previa con objeto de que los nuevos vectores sean de matriz de covarianza unidad.

Sean

$$\begin{aligned} \mu^* &= (t_1^T)^{-1} x_1 & : t_i \cdot t_i^T &= \Sigma_{11} \\ v^* &= t_2^{-1} x_2 & : t_i \cdot t_i^T &= \Sigma_{22} \\ &\vdots & & \vdots \\ w^* &= t_m^{-1} x_m & : t_m \cdot t_m^T &= \Sigma_{mm} \end{aligned} \quad (11)$$

La matriz de covarianzas entre  $\mu^*, v^*, \dots, w^*$  es de la forma:

$$R = \begin{bmatrix} I & t_1^T \Sigma_1(t_1)^T \dots & t_1^T \Sigma_m(t_m)^T \\ t_2^T \Sigma_2(t_2)^T & I & \dots & t_2^T \Sigma_m(t_m)^T \\ \vdots & \vdots & \ddots & \vdots \\ t_m^T \Sigma_m(t_m)^T & t_m^T \Sigma_m(t_m)^T & \dots & I \end{bmatrix} \quad (12)$$

Todas las transformaciones ortogonales de (11) en cada vector de variables conservan la diagonal principal de (12). Escogeremos de todas ellas las que en cada etapa den lugar a matrices de covarianzas (entre variables con el mismo subíndice) con suma de elementos no diagonales máxima.

Si  $a_i, b_i, \dots, c_i$  son las primeras columnas de las matrices de transformación ortogonal de cada vector de (11)

$$a_i^T a_i = 1; \quad b_i^T b_i = 1 \dots \dots c_i^T c_i = 1 \quad (13)$$

los comentarios anteriores justifican el tomar como primeras variables canónicas generalizadas las expresiones

$$\begin{aligned} \mu_1 &= a_1^T t_1^{-1} x_1 = \rho_1^T x_1 \\ \nu_1 &= b_1^T t_2^{-1} x_2 = \omega_1^T x_1 \\ &\vdots \\ \omega_1 &= c_1^T t_m^{-1} x_m = \eta_1^T x_m \end{aligned} \quad (14) \quad : \sum_{i=1}^m \omega_i = \begin{bmatrix} I & \dots & a_1^T t_1^{-1} \Sigma_m(t_m)^T c_1 \\ \vdots & & \vdots \\ c_1^T t_m^{-1} \Sigma_m(t_m)^T & \dots & I \end{bmatrix} \quad (15)$$

cuando su matriz de covarianzas tenga unos elementos no diagonales cuya suma sea máxima.

Procederemos a la obtención de  $a_1, b_1, \dots, c_1$ .

Sea

$$\varphi_1 = 1^T D_1^T R D_1 1 - 1^T D_1^T D_1 \lambda_1 - 1^T D_1^T D_1 1 \quad (16)$$



en donde :

$$\lambda_2 = \begin{bmatrix} \lambda_{12} \\ \lambda_{21} \\ \vdots \\ \lambda_{m2} \end{bmatrix}; \quad p_{21} = \begin{bmatrix} p_{11}^{21} \\ p_{21}^{21} \\ \vdots \\ p_{m1}^{21} \end{bmatrix}; \quad D_2 = \begin{bmatrix} q_2 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_2 \end{bmatrix}$$

La optimización de  $\mathcal{L}_2$  respecto a:  $\mathbf{1}^T \mathbf{D}$  conduce a:

$$P_1 D_2 \mathbf{1} - D_2 \lambda_2 - D_1 p_{21} = 0 \quad (22)$$

expresión que premultiplicada por  $D_1$  da lugar a:

$$D_1^T P_1 D_2 \mathbf{1} = p_{21} \mathbf{1}$$

y sustituida en la condición de óptimo:

$$(\mathbf{I} - D_1 D_1^T) P_1 D_2 \mathbf{1} = D_2 \lambda_2$$

finalmente, utilizando  $D_1^T D_2 = 0$

$$(\mathbf{I} - D_1 D_1^T) P_1 (\mathbf{I} - D_1 D_1^T) D_2 \mathbf{1} = D_2 \lambda_2 \quad (23)$$

Si denominamos  $P_2$  a  $(\mathbf{I} - D_1 D_1^T) P_1 (\mathbf{I} - D_1 D_1^T)$  la expresión (23) se reduce a:

$$P_2 D_2 \mathbf{1} = D_2 \lambda_2 \quad (24)$$

Este sistema es análogo a (18) pudiendo resolverse por el mismo procedimiento iterativo.

Razonando de forma similar puede llegarse a la obtención de

$$D_r = \begin{bmatrix} a_r & 0 & \dots & 0 \\ 0 & b_r & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_r \end{bmatrix} \quad (25)$$

como solución del sistema: (1)

$$P_r D_r \mathbf{1} = D_r \lambda_r$$

(1) El número  $r$  es función de los rangos de las particiones de (12) en cualquier caso como máximo, coincidirá con el número de componentes del vector que menos tiene.

en donde:

$$\lambda_v = \begin{bmatrix} \lambda_{1v} \\ \lambda_{2v} \\ \vdots \\ \lambda_{mv} \end{bmatrix}$$

y siendo

$$\rho_v = (I - D_1 D_1^T \dots - D_{r-1} D_{r-1}^T) P_i (I - D_1 D_1^T \dots - D_{r-1} D_{r-1}^T) \quad (26)$$

Comentaremos los resultados obtenidos: Los vectores canónicos

$$\begin{array}{lll} \mu = A^T t_1^1 x_1 & : & A \quad r \times n_1 \\ \sigma = B^T t_2^1 x_2 & : & B \quad r \times n_2 \\ \vdots & & \\ \omega = C^T t_m^1 x_m & : & C \quad r \times n_m \end{array}$$

no tienen una matriz de covarianzas como se señaló en (5) pero sí se aproximan a esta situación.  $(\mu, \sigma, \dots, \omega)$  tienen dentro de cada grupo sus componentes incorrelados; además las correlaciones entre componentes con el mismo subíndice han sido obtenidas en cada etapa de forma que su suma fuese máxima, siendo dichas correlaciones los elementos que ocupan el lugar  $ii$ :  $i=1, 2, \dots, r$  de  $P_{ij}$   $\forall i, j: i, j=1, 2, \dots, m$ . Este método conduce a matrices  $P_{ij}$  cuyos elementos distintos de los que ocupan los lugares  $ii$  en general no son nulos.

### 3.4. Método de aproximación a una matriz de rango uno.

Otra forma de lograr unas nuevas variables obtenidas por transformación de  $(x_1, x_2, \dots, x_m)$  cuya matriz de covarianzas se aproxime a la matriz (5), es a través de un criterio basado en la generalización de la propiedad 1.5.2. Recordaremos dicha propiedad para el modelo canónico de dos vectores de variables: Las primeras variables canónicas tienen una matriz de correlaciones, que de todas las de las transformadas lineales normalizadas en varianza de los vectores  $X_1$  y  $X_2$  es la que mejor se aproxima a una matriz de rango 1, en el sentido mínimo cuadrático. Propiedades análogas cumplen el resto de variables canónicas con condiciones adicionales de incorrelación con respecto a las variables previamente obtenidas.

Esta propiedad del modelo canónico de dos vectores de variables ha servido a Paul Horst para formular un nuevo criterio de obtención de las correlaciones y variables canónicas generalizadas.

Como en el caso anterior, consideraremos una transformación previa de cada vector de variables que dé lugar a nuevos vectores con matriz de covarianza unidad.

Sea  $R$  la matriz de covarianzas entre dichas variables transformadas. Se trata de hallar unas nuevas variables obtenidas por transformación ortogonal de las anteriores, cuya matriz de covarianzas en cada etapa se aproxime mejor a una matriz de rango 1 - en el sentido mínimo cuadrático.

Consideremos un primer grupo de variables (correspondientes a los primeros componentes de cada nuevo vector) según se indicó en (14). Determinaremos  $a_1, b_1, \dots, c_1$  exigiendo a la matriz de covarianzas de las variables (14) con las condiciones (13) el requisito expuesto en el párrafo anterior.

Sea  $D_1$  una matriz con las características señaladas en (17). La matriz de covarianzas entre las variables anteriores puede expresarse:

$$D_1^T R D_1 \quad (27)$$

Por ser una matriz definida positiva y simétrica existe su descomposición canónica<sup>(1)</sup> siendo el primer sumando de esta descomposición la matriz de rango 1 que mejor se aproxima a la dada según el criterio mínimo cuadrático<sup>(2)</sup>.

Sea  $\lambda_1$  la mayor raíz característica de (27) y  $p_1$  su vector asociado

$$D_1^T R D_1 p_1 = \lambda_1 p_1 \quad : \quad p_1^T p_1 = 1 \quad (28)$$

Por las propiedades anteriores  $\lambda_1 p_1 p_1^T$  es la matriz de rango 1 que mejor se aproxima a (27) según el criterio mínimo cuadrático. Además:

$$p_1^T D_1^T R D_1 p_1 = \lambda_1 \quad (29)$$

(1) La descomposición canónica de una matriz  $A$  definida, positiva y simétrica, viene dada por la expresión  $A = \lambda_1 a_1 a_1^T + \dots + \lambda_v a_v a_v^T$  en donde  $\lambda_1 > \lambda_2 > \dots > \lambda_v$  son las raíces características de  $A$  y  $a_1, a_2, \dots, a_v$  sus vectores asociados, siendo  $v = \text{rang} A$  Jorekög [24] pág. 51.

(2) Esta propiedad está demostrada en Good, I. [15] pág. 823.

Puesto que las raíces características de (27) dependen de  $D_1$  y esta matriz constituye las incógnitas a determinar, obtendremos  $D_1$  de forma que la expresión (29) sea máxima.

Otra forma alternativa de (29) es la siguiente:

$$Q_1^T R Q_1 = \lambda_1 \quad (30)$$

Siendo

$$Q_1 = D_1 p_1 \quad (31)$$

$$y \quad Q_1^T Q_1 = 1 \quad (32)$$

De estas expresiones deducimos que (30) es máximo cuando  $\lambda_1$  es la mayor raíz característica<sup>(1)</sup> de  $R$ . Si  $Q_1^*$  es el vector asociado a dicha raíz. (30), (31) y (32) permiten despejar  $D_1$  en función de  $Q_1^*$  (2)

$$p_1^T D_1^T D_1 p_1 = 1^T D_{Q_1} D_{Q_1} 1; \quad p_1^T p_1 = 1^T D_{Q_1}^T D_{Q_1} 1; \quad p_1^T p_1 = 1 p_1^2$$

$$1^T p_1^2 = 1^T D_{Q_1}^T D_{Q_1} 1; \quad p_1 = (D_{Q_1}^T D_{Q_1})^{1/2} 1 \quad (33)$$

$$D_1 (D_{Q_1}^T D_{Q_1})^{1/2} 1 = D_{Q_1} 1$$

$$D_1 = D_{Q_1} (D_{Q_1}^T D_{Q_1})^{1/2} \quad (34)$$

(1) Anderson, T.W. [ 2 ] pág. 274. La función  $Q_1^T R Q_1$  alcanza el máximo cuando  $Q_1$  es el vector asociado a la mayor raíz de  $R$ , y coincide con dicha raíz.

(2) Expresaremos  $Q_1$  a través de una matriz de la forma "D" ya utilizada en (17) y que contenga la diagonal principal las particiones de este vector. Denotaremos tal matriz:  $D_{Q_1} \Rightarrow Q_1 = D_{Q_1} 1$

La expresión (34) permite obtener los coeficientes  $a_1, b_1, \dots, c_1$  de la matriz (17) según el criterio considerado en este apartado.

Los coeficientes de las segundas variables canónicas generalizadas se obtendrán de forma similar añadiendo las condiciones de incorrelación ( $a_1 \perp a_2, b_1 \perp b_2, \dots, c_1 \perp c_2$ ). En forma analítica las expresaremos

$$D_2^T D_1 = 0 \quad (35)$$

En donde  $D_2$  es una matriz que contiene la diagonal principal, los vectores  $a_2, b_2, \dots, c_2$

Sea  $D_2^T R D_2$  la matriz de correlaciones entre las variables  $u_2, v_2, \dots, w_2$  definidas en (9). La mejor aproximación a  $D_2^T R D_2$  través de una matriz de rango 1 con la condición  $D_2^T D_1 = 0$  se logra con la matriz  $\lambda_2 p_2 p_2^T$  en donde  $\lambda_2$  es la segunda mayor raíz de  $R$  y  $p_2$  su vector asociado<sup>(1)</sup>.

Un razonamiento análogo empleado en el caso anterior nos lleva a los siguientes resultados

$$D_2^T R D_2 p_2 = \lambda_2 p_2 \quad : \quad p_2^T p_2 = 1 \quad (36)$$

$$\lambda_2 = p_2^T D_2^T R D_2 p_2 \quad (37)$$

y tomando :

$$Q_2 = D_2 p_2 \quad (38)$$

(1) Véase Good, I. [ 15 ] pág. 823.

obtenemos

$$\begin{aligned} Q_2^T R Q_2 &= \lambda_2 \\ R Q_2 &= \lambda_2 Q_2 \end{aligned} \quad (39)$$

La función (39) alcanza el óptimo cuando  $\lambda_2$  es la mayor raíz de  $R^{(1)}$ .

A partir de aquí puede despejarse  $D_2$  en función de  $D_{Q_2}$  siendo esta última matriz la que contiene en la diagonal principal al vector  $Q_2$  particionado. El procedimiento es totalmente análogo al empleado para obtener (34). En base a él se obtienen las siguientes expresiones:

$$P_2 = (D_{Q_2}^T D_{Q_2})^{1/2} \mathbf{1} \quad (40)$$

$$D_2 = D_{Q_2} (D_{Q_2}^T D_{Q_2})^{-1/2} \quad (41)$$

El proceso de obtención de variables y correlaciones canónicas generalizadas podría continuarse hasta la etapa  $r^{(2)}$  pudiéndose hallar los coeficientes  $a_r, b_r, \dots, c_r$  como soluciones de:

$$P_r = (D_{Q_r}^T D_{Q_r})^{1/2} \mathbf{1} \quad (42)$$

$$D_r = D_{Q_r} [D_{Q_r}^T D_{Q_r}]^{-1/2} \quad (43)$$

(1) Anderson, T.W. [ 2 ] pág. 274 y 275. El máximo de  $Q_2^T R Q_2 : Q_2^T Q_2 = 1$  se alcanza cuando  $Q_2$  es el vector asociado a la segunda mayor raíz de  $R$  coincidiendo el óptimo de la función con dicha raíz.

(2) Cuando todos los vectores constan de igual número de componentes y las particiones de  $R$  son matrices no singulares,  $r$  coincide con la dimensión común a tales vectores.

### 3. 5. Comparación de ambos procedimientos.

Tanto el procedimiento de máxima correlación como el procedimiento de aproximación a una matriz de rango uno, permiten obtener los coeficientes de los vectores canónicos generalizados así como los coeficientes de correlación canónica. Ya señalamos en 3 . 2. como en general, no es posible que las nuevas variables tengan matriz de covarianzas intergrupos de la forma expuesta en 1. 5. Consideraremos que un procedimiento es más adecuado que otro en la medida que las matrices de covarianza intergrupos se acerquen más a matrices de la forma considerada en 1. 5.

Una manera de determinar tal acercamiento es evaluando la suma de los cuadrados de los elementos que ocupen lugares descritos con sub-índices distintos en las matrices de covarianzas intergrupos. P. Horst<sup>(1)</sup> ha mostrado en una aplicación de los dos métodos al mismo conjunto de datos que los resultados obtenidos son bastante similares, no obstante, señala la posibilidad de llegar a resultados más dispares en el caso de emplear otra serie de datos iniciales. De cualquier forma, precisa que es necesaria una investigación sobre los resultados a producirse ante diferente tipo de datos experimentales.

---

(1) Horst, P. [20 ], página 330.

### 3.6. Modelos canónicos parciales generalizados.

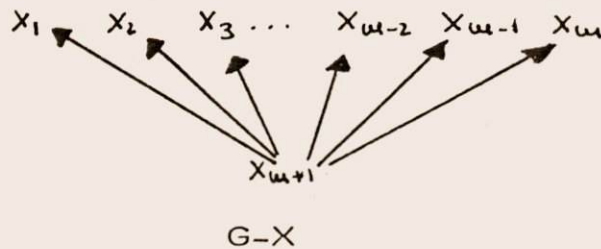
Denominaremos modelos canónicos parciales generalizados aquellos cuyo objeto es estudiar la interdependencia entre más de dos vectores de variables aleatorias cuando se han eliminado las influencias lineales de otros.

En este apartado vamos a desarrollar un conjunto de generalizaciones de los modelos parciales expuestos en 2.2. En ellos determinaremos las correlaciones y variables canónicas generalizadas entre  $m$  vectores de residuos obtenidos en cada caso eliminando las influencias lineales supuestas.

En todos los modelos supondremos que los vectores de variables aleatorias que intervienen están distribuidos conjuntamente siendo sus parámetros  $(0, \Sigma)$ . No haremos mención al número de componentes de cada vector; en cualquier caso el número de variables canónicas parciales generalizadas será como máximo el mínimo del número de componentes de los vectores que intervengan en cada modelo.

#### 3.6.1. Modelo parcial generalizado.

Si consideramos los vectores de variables aleatorias  $(X_1, \dots, X_{u+1})_{:u}$  una extensión del grafo  $G-I$  a  $m$  vectores de variables da lugar a la siguiente estructura de dependencias



Denotaremos  $\varepsilon_{1,u+1}, \varepsilon_{2,u+1}, \dots, \varepsilon_{u,u+1}$  a los vectores  $x_1, x_2, \dots, x_u$  después de eliminadas las influencias lineales supuestas en  $G-X$ . Estos pueden obtenerse como diferencias entre  $x_1, x_2, \dots, x_u$  y sus estimaciones lineales mínimo-cuadráticas

$$\varepsilon_{i,u+1} = x_i - \sum_{j=1}^i \sum_{u+1, u+1}^{-1} x_{u+1} \quad i=1, \dots, u \quad (44)$$

Denominaremos correlaciones y vectores canónicos parciales generalizadas entre  $(x_1, x_2, \dots, x_u)$  a las correlaciones y vectores canónicos generalizados entre  $(\varepsilon_{1,u+1}, \dots, \varepsilon_{u,u+1})$

Los procedimientos expuestos en 3.3. y 3.4. requieren para el cálculo de tales vectores conocer la matriz de covarianzas entre los residuos  $\varepsilon_{1,u+1}, \dots, \varepsilon_{u,u+1}$ . Evaluando las expresiones  $E \varepsilon_{i,u+1} \varepsilon_{j,u+1}^T$   $i, j=1, \dots, u$  a través de (44) dicha matriz adopta la siguiente forma;

$$\sum_{\varepsilon_{1,u+1}, \dots, \varepsilon_{u,u+1}} \begin{bmatrix} \sum_{11,u+1} & \sum_{12,u+1} & \dots & \sum_{1u,u+1} \\ \sum_{21,u+1} & \sum_{22,u+1} & \dots & \sum_{2u,u+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{u1,u+1} & \sum_{u2,u+1} & \dots & \sum_{uu,u+1} \end{bmatrix} \quad (45)$$

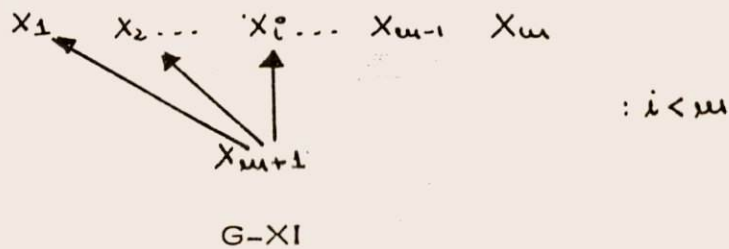
La sustitución de las particiones  $\sum_{ij,u+1}$  en lugar de  $\sum_{ij}$  en las matrices  $P_1, P_2, \dots, P_r$  de (18), (24) y (26) en el procedimiento de "máxima correlación" permite obtener las correlaciones y vectores canónicos parciales generalizados entre  $\varepsilon_{1,u+1}, \dots, \varepsilon_{u,u+1}$

Si el procedimiento que se aplica para obtener tales variables es el de "aproximación a una matriz de rango 1", este puede ser utilizado sin más que sustituir en la matriz  $R(12)$  las particiones  $\sum_{ij}$  por  $\sum_{ij,u+1}$ . La evaluación de (34), (41) y (43) después de esta sustitución conduce a las correlaciones y vectores canónicos ya mencionadas.

En el resto de modelos canónicos que expandremos en este capítulo únicamente nos referiremos al grafo de dependencia entre los vectores de variables que intervienen y a las matrices de covarianzas de los residuos resultantes de eliminar las influencias lineales admitidas en cada caso. La aplicación de los procedimientos señalados en el párrafo anterior conduce a la obtención de las correlaciones y vectores canónicos de cada modelo a considerar.

### 3.6.2. Modelo parte generalizado.

Una extensión del grafo G-II a  $m$  vectores de variables da lugar al siguiente tipo de dependencias entre  $(x_1, x_2, \dots, x_{m+1}) : m > 2$ .

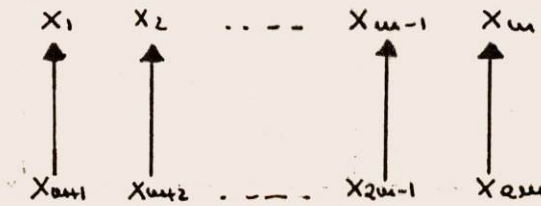


La matriz de covarianzas entre  $x_1, x_2, \dots, x_m$  después de eliminadas las influencias lineales propuestas en G-XI adopta la forma

$$\sum_{\epsilon_{1,m+1} \dots \epsilon_{i,m+1} \dots} x_{12} = \begin{bmatrix} \Sigma_{11,m+1} & \dots & \Sigma_{1i,m+1} & \dots & \Sigma_{1m} \\ \Sigma_{21,m+1} & \dots & \Sigma_{2i,m+1} & \dots & \Sigma_{2m} \\ \vdots & & \vdots & & \vdots \\ \Sigma_{i1,m+1} & \dots & \Sigma_{ii,m+1} & \dots & \Sigma_{im} \\ \Sigma_{(i+1)1,m+1} & \dots & \Sigma_{(i+1)i,m+1} & \dots & \Sigma_{(i+1)m} \\ \vdots & & \vdots & & \vdots \\ \Sigma_{m1} & \dots & \Sigma_{m2} & \dots & \Sigma_{mm} \end{bmatrix} \quad (46)$$

3.6.3. Modelo biparcial generalizado.

La generalización del modelo expuesto en el 2.2.3. requiere la consideración de los vectores  $(x_1, x_2, \dots, x_{2u})$ ;  $u > 2$  cuyas relaciones vienen definidas a través del siguiente grafo



G-XII

Si  $\varepsilon_{1, u+1}, \varepsilon_{2, u+2}, \dots, \dots, \dots, \varepsilon_{u, 2u}$  representan a los vectores  $x_1, x_2, \dots, x_{2u}$  desprovistos de las influencias lineales de G-XII, su matriz de covarianzas es de la siguiente forma:

$$\begin{bmatrix} \Sigma_{11, u+1} & \Sigma_{12}^* & \dots & \Sigma_{1, 2u}^* \\ \Sigma_{21}^* & \Sigma_{11, u+2} & \dots & \Sigma_{2, 2u}^* \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{u1}^* & \Sigma_{u2}^* & \dots & \Sigma_{u, 2u}^* \end{bmatrix} \quad (47)$$

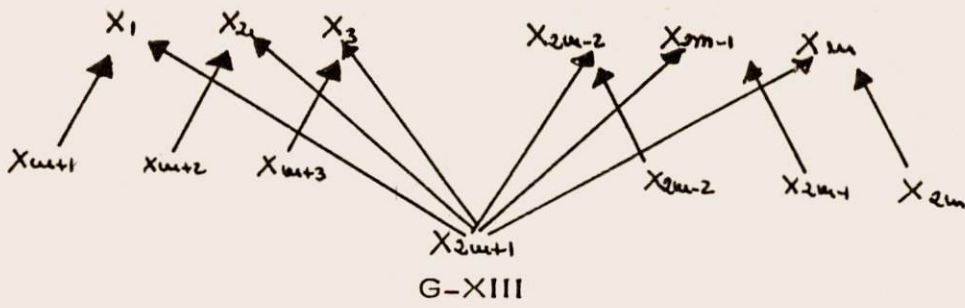
en donde <sup>(1)</sup>  $\Sigma_{ij}^* = \Sigma_{ij} - \Sigma_{i, (u+1)} \Sigma_{(u+1), (u+1)}^{-1} \Sigma_{(u+1), j} - \Sigma_{i, (u+2)} \Sigma_{(u+2), (u+2)}^{-1} \Sigma_{(u+2), j} - \dots - \Sigma_{i, (u-1)} \Sigma_{(u-1), (u-1)}^{-1} \Sigma_{(u-1), j} + \Sigma_{i, (u+1)} \Sigma_{(u+1), (u+1)}^{-1} \Sigma_{(u+1), (u+2)} \Sigma_{(u+2), (u+2)}^{-1} \Sigma_{(u+2), (u+1)} \Sigma_{(u+1), j} - \dots$

3.6.4. Modelo  $G_1$  generalizado.

En este modelo estudiaremos las interdependencias entre los vectores de variables aleatorias cuando se suponen las siguientes dependencias:

(1) Esta expresión puede calcularse como desarrollo de

$$E \varepsilon_i (u+1) \varepsilon_j^T (u+2) : \varepsilon_i = x_i - \Sigma_{i, (u+1)} \Sigma_{(u+1), (u+1)}^{-1} x_{(u+1)} \dots$$



El vector  $x_i^*$  :  $i \leq m$  después de eliminadas las Influencias señaladas en G-XIII lo denotaremos <sup>(1)</sup> :

$$\varepsilon_{i(2m+1)} z_{u+1} = \varepsilon_{i(2m+1)} - \sum_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} X_{u+1} \quad (48)$$

Siendo la matriz de covarianzas entre los residuos

$$\varepsilon_{i(2m+1)(u+1)} : i = 1, 2, \dots, m.$$

$$\sum_{\varepsilon_{i(2m+1)(u+1)} \dots \varepsilon_{m(2m+1)} z_{u+1}} = \begin{bmatrix} \Sigma_{11}^{**} & \Sigma_{12}^{**} & \dots & \Sigma_{1m}^{**} \\ \Sigma_{21}^{**} & \Sigma_{22}^{**} & \dots & \Sigma_{2m}^{**} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{m1}^{**} & \Sigma_{m2}^{**} & \dots & \Sigma_{mm}^{**} \end{bmatrix} \quad (49)$$

en donde

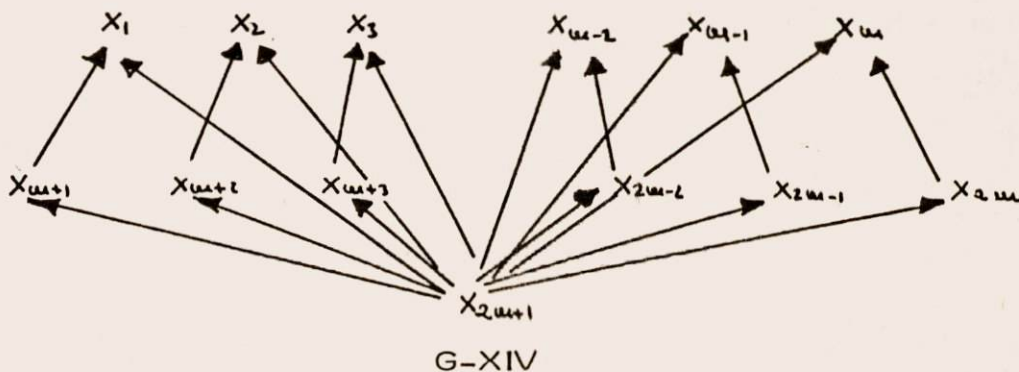
$$\Sigma_{ij}^{**} = \begin{cases} \Sigma_{i(2m+1)} - \Sigma_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} \Sigma_{(u+1)(u+1)} \Sigma_{(u+1)i} & \text{Si } i=j \\ \Sigma_{i(2m+1)} - \Sigma_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} \Sigma_{(u+1)(u+1)} \Sigma_{(u+1)j} - \Sigma_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} \Sigma_{(u+1)j} \\ \Sigma_{i(u+1), 2m+1} + \Sigma_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} \Sigma_{(u+1)(u+1)} \Sigma_{(u+1)j} - \Sigma_{i(u+1), 2m+1} \Sigma_{(u+1)(u+1)}^{-1} \Sigma_{(u+1)j} \end{cases} \text{ Si } i \neq j$$

3.6.5. Modelo G<sub>2</sub> generalizado.

Si se consideran los vectores de variables aleatorias  $x_1, x_2, \dots, x_{2m}$  la generalización del grafo G-V cuando se pretende

(1) Véase (27) y (28) del capítulo II.

estudiar las interdependencias de m de ellos conduce al siguiente:



Los vectores  $X_i : i = 1, 2, \dots, u$  después de eliminadas las influencias lineales supuestas en G-XIV dan lugar a las siguientes expresiones (1)

$$E_{i(2u+1)(u+i)} = E_{i(2u+1)} - \sum_{i(u+i), 2u+1} \sum_{(u+i)(u+i), 2u+1}^{-1} X_{u+i} \quad i=1, \dots, u \quad (50)$$

siendo la matriz de covarianzas entre ellos

$$\sum_{E_{1(2u+1)(u+1)} \dots E_{u(2u+1)2u}} = \begin{bmatrix} \sum_{11, (2u+1), (u+1)}^{***} & \sum_{12}^{***} & \dots & \sum_{1u}^{***} \\ \sum_{21}^{***} & \sum_{22, (2u+1)(u+2)}^{***} & \dots & \sum_{2u}^{***} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{u1}^{***} & \sum_{u2}^{***} & \dots & \sum_{uu, (2u+1)2u}^{***} \end{bmatrix} \quad (51)$$

en donde (2)

$$\begin{aligned} \sum_{ij}^{***} &= \sum_{j, (2u+1)} - \sum_{i(u+i), (2u+1)} \cdot \sum_{(u+i)(u+i), (2u+1)} \sum_{(u+i)^2} j_{\circ(2u+1)} - \\ &- \sum_{i(u+i), 2u+1} \sum_{(u+i)(u+i), (2u+1)}^{-1} \sum_{(u+i)j} j_{\circ(2u+1)} + \\ &+ \sum_{i(u+i), 2u+1} \sum_{(u+i)(u+i), 2u+1} \sum_{(u+i)(u+i), 2u+1}^{-1} \sum_{(u+i)(u+i)j} j_{\circ(2u+1)} \end{aligned}$$

(1) Véase (31) y (32) del capítulo II.

(2) Esta expresión es el resultado de  $E E_{i(2u+1)(u+i)} E_{j(2u+1)(u+i)}^T$  cuando estos residuos son de la forma (50).

### Comentario final

Los modelos canónicos expuestos en 3.6. poseen la característica común de estar referidos a  $m$  vectores de variables aleatorias en los que se han eliminado las influencias lineales de otros. El orden en la exposición ha sido de menos a más general de tal forma que el modelo  $G_2$  generalizado incluye como caso particular a  $G_1$ ;  $G_1$  al modelo biparcial, etc.

Estos modelos pueden estudiarse en forma genérica extendiendo a  $m$  vectores de variables los modelos  $L(2n+1)$  y  $C(2n+1)$  expuestos en el capítulo II. Esta extensión no ofrece dificultades de tipo estadístico ya que en síntesis se reduce a obtener la matriz de covarianzas entre  $m$  vectores de la forma (96) o bien de la forma (105)<sup>(1)</sup>. Dado que en  $L(2n+1)$  y  $C(2n+1)$  no hemos obtenido de forma explícita todas las particiones de sus matrices de covarianzas tampoco lo haremos aquí; únicamente dejamos constancia del método a seguir en caso a proceder a tal generalización.

Puede aplicarse este tipo de modelos la misma crítica que Timm y Carlson realizan a sus modelos "parte" y "biparcial" señalando que las dificultades de su desarrollo son más bien de tipo pragmático, ya que es difícil encontrar fenómenos que respondan a tal tipo de dependencia. Por otra parte, la interpretación de los coeficientes de correlación y vectores canónicos se hace mucho más oscura de lo que por sí ya es. El motivo de haber presentado en este trabajo este tipo de generalizaciones ha sido el de completar esa parte del análisis canónico que son los modelos parciales.

---

(1) Las notaciones (96) y (105) corresponden al capítulo II.

En todo este capítulo no hemos hecho referencia a la distribución conjunta de los vectores de variables que en cada modelo intervienen. Cuando dicha distribución es normal multivariante, las matrices de covarianzas que aparecen en 3.6. pueden interpretarse como matrices de covarianzas de determinadas distribuciones condicionadas. Esta característica ya fue mencionada en 2.4. por lo cual para este tipo de consideraciones nos remitimos a este apartado.

PARTE SEGUNDA:

EL MODELO MUESTRAL.

CAPITULO IV: El modelo canónico muestral de dos vectores de variables aleatorias.

#### 4. 1. Introducción.

Todos los modelos canónicos que hemos estudiado en la primera parte de este trabajo han sido desarrollados en el ámbito de colectivos teóricos representados por vectores aleatorios  $X, Y, \dots$ . Asociados a estos modelos teóricos aparecen los correspondientes modelos muestrales que deben permitir contrastar las aproximaciones o estimaciones obtenidas en estos con las facilitadas por los modelos teóricos.

La teoría muestral en los métodos de análisis multivariante ha tenido un desarrollo con dificultades muy fuertes que a veces se han obviado al coste de limitarse a resultados asintóticos. Así una característica que también señalaremos es que todos los modelos muestrales de las distintas técnicas multivariantes han sido estudiados únicamente para muestras procedentes de colectivos normales. El modelo canónico no ha sido una excepción a esta tónica general aunque quizás el hecho de ser un marco genérico para un conjunto de estas técnicas ha despertado un mayor interés en un núcleo de expertos de la estadística teórica; prueba de ello es la amplia gama de artículos dedicados a estos temas aparecidos en las décadas de los cincuenta y sesenta en las revistas "Biometrika" y "The Annals of Mathematical Statistics". En la actualidad puede afirmarse que el modelo canónico es de todos los modelos multivariantes aquel en que más se ha avanzado en el desarrollo de su teoría muestral.

El objeto de este capítulo dentro del modelo muestral es análogo al del capítulo I en el modelo poblacional; ofrece una síntesis de los

resultados que a lo largo del tiempo se han ido obteniendo en esta materia que serán utilizados en el capítulo siguiente como base de desarrollo de los modelos muestrales del análisis canónico parcial. En el contenido específico de este capítulo distinguimos tres partes:

La primera está dedicada al estudio de los estimadores de las correlaciones y vectores de coeficientes canónicos. Presentaremos el procedimiento de obtención de los estimadores de máxima verosimilitud de tales correlaciones y vectores. En la segunda parte se estudia el modo de obtener las distribuciones en el muestreo de los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos. Toda la problemática de estas distribuciones gira en torno a la de las raíces características y vectores asociados de ciertas matrices de variables aleatorias. En primer lugar estudiaremos un caso genérico como es la obtención de las distribuciones de las raíces características y vectores asociados del sistema  $[A - \lambda B] Y = 0$  cuando A y B son matrices de variables aleatorias independientes con distribución de Wishart de parámetro común  $\Sigma$ ; en base a este caso obtendremos las distribuciones en el muestreo de las correlaciones y vectores de coeficientes canónicos muestrales cuando éstos se han obtenido de muestras procedentes de dos vectores aleatorios independientes y con distribución de probabilidades conjunta normal.

La tercera parte de este capítulo la hemos dedicado a fundamentar una serie de contrastes de hipótesis relativos a las correlaciones y vectores de coeficientes canónicos. Hemos creído conveniente introducir en un primer apartado una serie de distribucio-

nes de probabilidades que bajo ciertas restricciones siguen los estadísticos que emplearemos para el contraste de tales hipótesis. El segundo apartado lo dedicaremos a la exposición de los tests que serán utilizados en el contraste de las hipótesis a las que anteriormente nos hemos referido.

Introduciremos aquí algunas cuestiones de carácter general necesarias en la exposición de los modelos muestrales que contiene este capítulo. En lo sucesivo designaremos  $(X^T, Y^T)$  a dos vectores de variables aleatorias (vectores aleatorios de referencia) con distribución conjunta normal  $N(m, \Sigma)$  de dimensiones  $p$  y  $q$ :  $p \leq q$  y  $\text{rang } \Sigma = p+q$  de los cuales se ha tomado una muestra de tamaño  $N = n+1 \gg p+q$ . Dicha muestra puede expresarse en la siguiente matriz

$$Z = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \dots & x_{pN} \\ \hline y_{11} & y_{12} & \dots & y_{1N} \\ y_{21} & y_{22} & \dots & y_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{q1} & y_{q2} & \dots & y_{qN} \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} \quad (1)$$

Si utilizamos las matrices

$$E_{NN} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad (2)$$

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (3)$$

$$E_{N1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (4)$$

en base a ellas puede obtenerse la matriz de covarianzas muestral:

$$S = \frac{1}{N} Z (I - \frac{E_{NN}}{N}) Z^T \quad (5)$$

Denominaremos  $D$  a la matriz:

$$D = Z (I - \frac{E_{NN}}{N}) Z^T \quad (6)$$

(1) La expresión de la matriz de covarianzas  $S$  en función de  $Z$ , es introducida por Kshirsagar [28] cap. II.

Mediante (5) y (6) llegamos a la siguiente expresión:

$$S = \frac{1}{N} D \quad (7)$$

Haciendo uso de la partición (1) se obtiene

$$(6) \quad D = \begin{bmatrix} \check{X} \\ \check{Y} \end{bmatrix} \left[ I - \frac{E_{NN}}{N} \right] \begin{bmatrix} \check{X}^T & \check{Y}^T \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ \dots & \dots \\ D_{21} & D_{22} \end{bmatrix} \quad (8)$$

y finalmente por (7)

$$\begin{bmatrix} S_{11} & S_{12} \\ \dots & \dots \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} D_{11}/N & D_{12}/N \\ \dots & \dots \\ D_{21}/N & D_{22}/N \end{bmatrix} \quad (9)$$

#### 4.2. Correlaciones y vectores canónicos muestrales.

Las correlaciones y vectores de coeficientes canónicos muestrales se definen de forma totalmente similar a las del modelo poblacional sin más que sustituir las particiones de la matriz de covarianzas poblacional de las expresiones (21) y (22) por las particiones de la matriz de covarianzas muestral presentada en (9). Si efectuamos tal sustitución denominaremos vectores canónicos muestrales a los obtenidos a través de la transformación:

$$u^* = G^T x \quad ; \quad v^* = H^T y$$

en donde las matrices G y H son las soluciones de los sistemas

$$[S_{12} \ S_{22}^{-1} \ S_{21} - c^2 S_{11}] G^* = 0 \quad : \quad G^{T*} S_{11} G^* = I \quad (10)$$

$$[S_{21} \ S_{11}^{-1} \ S_{12} - c^2 S_{22}] H^* = 0 \quad : \quad H^{T*} S_{22} H = I \quad (11)$$

Igualmente denominaremos correlaciones canónicas muestrales a las raíces cuadradas con signo positivo ordenadas de forma decreciente de las soluciones de las ecuaciones determinantes

$$|S_{12} \ S_{22}^{-1} \ S_{21} - c^2 S_{11}| = 0 \quad ; \quad |S_{21} \ S_{11}^{-1} \ S_{12} - c^2 S_{22}| = 0 \quad (12)$$

Suele ser frecuente definir las correlaciones y vectores canónicos muestrales, bien en la forma anterior o como a continuación se expone:

Se denominan vectores canónicos muestrales a los obtenidos a través de la siguiente transformación en la que la matriz de covarianzas S, de (5) se sustituye a efectos de transformación del sistema por su proporcional D; es decir :

$$\mu = G^T X \quad ; \quad \sigma = H^T Y$$

en donde G y H son soluciones de los sistemas

$$[D_{12} D_{22}^{-1} D_{21} - c^2 D_{11}] G = 0 \quad : \quad G^T D_{11} G = I \quad (13)$$

$$[D_{21} D_{11}^{-1} D_{12} - c^2 D_{22}] H = 0 \quad : \quad H^T D_{22} H = I \quad (14)$$

En este caso se denominan correlaciones canónicas muestrales a las raíces cuadradas con signo positivo ordenadas de forma decreciente, de las ecuaciones determinantes

$$|D_{12} D_{22}^{-1} D_{21} - c^2 D_{11}| = 0 \quad ; \quad |D_{21} D_{11}^{-1} D_{12} - c^2 D_{22}| = 0 \quad (15)$$

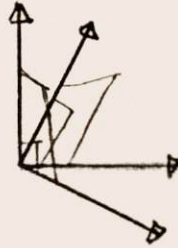
Como tendremos ocasión de ver la sustitución de la expresión (5) por (6) en las ecuaciones anteriores tiene una serie de ventajas ya que las particiones de la matriz D que intervienen en (13) y (14) siguen determinadas distribuciones en el muestreo lo cual facilita la obtención de las distribuciones de las correlaciones y vectores de coeficientes canónicos.

Comentaremos ambas definiciones. Por lo que respecta a las correlaciones canónicas muestrales los determinantes (12) y (15) tienen las mismas raíces<sup>(1)</sup> por lo que las dos definiciones coinciden. No ocurre lo mismo con los vectores de coeficientes canónicos; las columnas de las matrices de los sistemas (10), (11) y (13), (14) son proporcionales por lo que las variables canónicas en la primera y segunda definición presentan la misma estructura aunque a diferente escala.

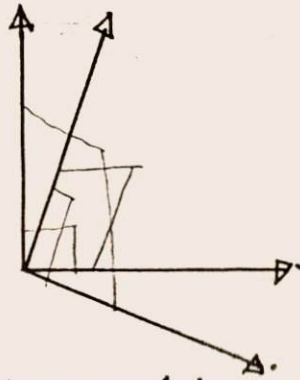
---

(1)  $|D_{12} D_{22}^{-1} D_{21} - c^2 D_{11}| = \frac{1}{N^p} |S_{12} S_{22}^{-1} S_{21} - c^2 S_{11}| = 0$

Gráficamente;



Vectores canónicos según (10) y (11)



Vectores canónicos según (13) y (14)

Las correlaciones y vectores canónicos muestrales poseen una serie de propiedades similares a las estudiadas en el modelo poblacional por lo que no serán repetidas aquí; únicamente haremos dos precisiones:

- 1) Las matrices  $S$  y  $D$  son definidas positivas con probabilidad uno.
- 2)  $S_{12}$  y  $D_{12}$  son de rango  $p = \min(p, q)$  con probabilidad uno, aunque se hayan obtenido con muestras procedentes de dos vectores con sólo  $r < p$  correlaciones canónicas poblacionales no nulas.

#### 4.3. Estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos

En lo sucesivo nos referiremos extensamente a las matrices  $D$ ,  $G$ ,  $H$  y  $C$  asociadas en la correspondencia modelo teórico, modelo muestral a las matrices  $N\Sigma$ ,  $L$ ,  $M$  y  $P$ .

Al establecer el modelo teórico se demostraron las siguientes relaciones

$$\begin{aligned} L^T \Sigma_{11} L &= I_p & M_2^T \Sigma_{22} M_2 &= I_{q-p} & : M &= \begin{bmatrix} M_1 & | & M_2 \\ \hline p & & q-p \end{bmatrix}_p \\ L^T \Sigma_{12} M_1 &= P^* & L^T \Sigma_{12} M_2 &= 0 & : P &= \begin{bmatrix} P^* & | & 0 \\ \hline p & & q-p \end{bmatrix}_p \end{aligned} \quad (16)$$

Análogamente entre las matrices del modelo muestral se tiene:

$$\begin{aligned} G^T D_{11} G &= I_p & H_2^T D_{22} H_2 &= I_{q-p} & ; H &= \begin{bmatrix} H_1 & | & H_2 \\ \hline p & & q-p \end{bmatrix}_p \\ G^T D_{12} H_1 &= C^* & G^T D_{12} H_2 &= 0 & ; C &= \begin{bmatrix} C^* & | & 0 \\ \hline p & & q-p \end{bmatrix}_p \\ H^T D_{22} H &= I_q & H_1^T D_{22} H_1 &= 0 & C &= \begin{bmatrix} c_{11} & 0 & \dots & 0 \\ 0 & c_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_{pp} \end{bmatrix}_p \end{aligned} \quad (17)$$

Una proposición fundamental para el desarrollo que a continuación exponemos es la siguiente:

Teorema fundamental. Si la matriz  $D$  es un estimador de máxima verosimilitud de  $N\Sigma$  las matrices  $C$ ,  $G$  y  $H$  constituyen estimadores de máxima verosimilitud de las matrices poblacionales  $P$ ,  $L$  y  $M$  cuando se obtienen como soluciones de los sistemas (13), (14) y (15).

En la demostración de este teorema emplearemos los siguientes lemas:

Lema 1. Dada una muestra de tamaño  $N$  de una variable aleatoria normal  $(p+q)$ -variante con distribución no degenerada y de pará-

metros  $(m, \Sigma)$ , el estimador de máxima verosimilitud de  $N\Sigma$  viene dada por la matriz  $D^{(1)}$  definida en (6).

Lema 2. Si  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$  son estimadores de máxima verosimilitud de  $\theta_1, \theta_2, \dots, \theta_m$  entonces  $\psi_1(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m), \dots, \psi_m(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m)$  son los estimadores de máxima verosimilitud de  $\psi_1(\theta_1, \theta_2, \dots, \theta_m), \dots, \psi_m(\theta_1, \theta_2, \dots, \theta_m)$  cuando  $\psi_1, \psi_2, \dots, \psi_m$  es una transformación "uniquely". Si los estimadores de  $\theta_1, \theta_2, \dots, \theta_m$  son únicos, también lo son los de  $\psi_1, \dots, \psi_m$  (2).

Los sistemas expuestos en (17) suponen una transformación de las variables aleatorias contenidas en la matriz  $D$  a las variables de las matrices  $C, G$  y  $H$ . Demostraremos que en esta transformación coinciden el número de ecuaciones con el número de elementos de  $C, G$  y  $H$ .

El número de tales ecuaciones es:

$$\left[ p + \frac{p(p-1)}{2} \right] + p^2 + 2(q-p)p + \left[ \frac{p(p-1)}{2} + p \right] + (q-p) + (q-p)(q-p-1)/2 = \frac{p^2}{2} + \frac{q^2}{2} - \frac{p}{2} - \frac{q}{2} + pq$$

Al no existir ninguna relación que ligue a  $H_2$  con las correlaciones canónicas los elementos de esta matriz están indeterminados. Cualquier multiplicación por la izquierda por una matriz ortogonal de orden  $q-p$  conduce a otra matriz que también satisface (17).

(1) Expondremos aquí una síntesis de tal demostración. Una prueba más detallada puede verse en Morrison, D. [ 33 ] pág. 95.

La función de verosimilitud de la muestra la expresaremos

$$L(\omega, \Sigma) = 2\pi^{-\frac{N(p+q)}{2}} |\Sigma|^{-\frac{N}{2}} \exp. \left[ -\frac{1}{2} \omega + \Sigma^{-1} \omega \right]$$

Derivando respecto a  $\Sigma$  obtenemos

$$-\frac{N}{2} [\Sigma^{-1}]^T - \frac{1}{2} [-\Sigma^{-1} \omega \Sigma^{-1}]^T = 0 ; \hat{\Sigma} = \frac{\omega}{N} ; D = N \hat{\Sigma}.$$

Las reglas de esta derivación pueden consultarse en Fz. Trocóniz A. [ 13 ] A-6.

(2) La prueba de este teorema puede verse en Anderson, T.W. [ 2 ] pág. 48.

Entre las formas de subsanar dicha indeterminación Anderson T.W. propone la exigencia de que las últimas  $q-p$  filas de  $H_2$  sean una matriz triangular. Esto es equivalente a añadir  $(q-p)(q-p-1)/2$  nuevas ecuaciones que unidas a las anteriores da lugar al siguiente resultado.

$$\frac{p^2}{2} + \frac{q^2}{2} + pq - \frac{q}{2} - \frac{p}{2} + (q-p)(q-p-1)/2 = p^2 + q^2 + p.$$

Este número coincide con el de elementos de G, H y C ; es decir,

$$\begin{array}{ccc} G & H & C \\ \downarrow & \downarrow & \downarrow \\ p^2 + & q^2 + & p. \end{array} \quad (18)$$

Este resultado nos permite afirmar que los elementos C, G y H de la transformación (17) cuando se añaden las restricciones del párrafo anterior acerca de  $H_2$  están únivocamente determinados y por tanto satisfacen al lema 2. Puesto que las matrices  $D_{11}$ ,  $D_{12}$ ,  $D_{22}$  utilizadas en (13) y (14) son los estimadores de máxima verosimilitud de  $N\Sigma_{11}$ ,  $N\Sigma_{12}$ ,  $N\Sigma_{22}$  (según se desprende del lema 1) Las matrices C, G y H son los estimadores de máxima verosimilitud de P, L y M.

#### 4.4. Distribuciones en el muestreo de las correlaciones y vectores de coeficientes canónicos muestrales.

##### 4.4.1. Consideraciones preliminares

El estudio de las distribuciones en el muestreo de las correlaciones y vectores de coeficientes canónicos muestrales ha sido un tema de gran interés dentro de la estadística matemática. Numerosos autores han dedicado sus trabajos a este tipo de distribuciones abarcando la numerosa casuística que se puede presentar; algunos casos dada su extrema dificultad están aún sin resolver y en otros únicamente se ha llegado a resultados asintóticos. Dedicaremos esta sección a la obtención de las distribuciones de las correlaciones y vectores de coeficientes canónicos en el caso más simple que puede presentarse, es decir, cuando éstas han obtenido de muestras de dos vectores de variables aleatorias independientes con distribución conjunta normal multivariante; este caso suele ser conocido en la literatura estadística como "null case".<sup>(1)</sup>

Con objeto de llegar a tales distribuciones a continuación presentaremos una serie de resultados previos que facilitarán la cuestión.

a) Si  $A^*$  y  $B^*$  son dos matrices de orden  $p \times p$  con distribuciones de Wishart<sup>(2)</sup> independientes

$$A^* \in W_p [\Sigma, \omega] ; B^* \in W_p [\Sigma, \omega] \quad (19)$$

(1) Esta denominación está justificada por referirse al caso en que se obtienen las distribuciones de las correlaciones canónicas muestrales cuando las correlaciones canónicas poblacionales son nulas.

(2) Si  $A = \sum_{n=1}^N (x_n - \bar{x})(x_n - \bar{x})^T$ ;  $N = n+1$  y  $x_n$  son muestras aleatorias simples procedentes de  $N_p(\omega, \Sigma)$ . La matriz  $A$  se dice que sigue la distribución de Wishart siendo su función de densidad:

$$f(A) = \frac{|A|^{1/2} (n-p-1)! \exp(-1/2 \text{tr } A \Sigma^{-1})}{2^{np/2} \pi^{1/2} \Gamma(p-1) |\Sigma|^{1/2} n! \Gamma(p-1) \Gamma(n/2)}$$

suele indicarse abreviadamente  $A \in W_p(\Sigma, \omega)$ . Al parámetro  $n$  se le denomina grados de libertad de la distribución.

las matrices  $A$  y  $B$  obtenidas de  $A^*$  y  $B^*$  según la transformación

$$\begin{aligned} A &= C A^* C^T \\ B &= C B^* C^T \end{aligned} \quad (20)$$

en donde  $C$  es la matriz unificadora de  $\Sigma$  :  $C^T \Sigma C = I_p$  , (21)

tienen las distribuciones independientes<sup>(1)</sup> :

$$A \in W_p [I, u] ; B \in W_p [I, u] \quad (22)$$

b) Las raíces del determinante  $|A^* - \lambda B^*| = 0$  coinciden con las raíces de  $|A - \lambda B| = 0$

En efecto,

$$|A - \lambda B| = |C A^* C^T - \lambda C B^* C^T| = |C [A^* - \lambda B^*] C^T| = |C| |A^* - \lambda B^*| |C^T| = 0$$

como  $|C| \neq 0$  por ser matriz unificadora.  $|A - \lambda B| = |A^* - \lambda B^*| = 0$

c) Los vectores asociados a las raíces de  $|A - \lambda B| = 0$  y  $|A^* - \lambda B^*| = 0$  en los sistemas  $[A - \lambda B]x = 0$  y  $[A^* - \lambda B^*]x^* = 0$  guardan la siguiente relación

$$x^* = C^T x$$

En efecto, si partimos del sistema  $[A - \lambda B]x = 0$

$$C^{-1} [A - \lambda B] x = 0$$

$$C^{-1} [C A^* C^T - \lambda C B^* C^T] x = 0 ; [A^* - \lambda B^*] C^T x = 0 ; x^* = C^T x \quad (23)$$

d) Los sistemas  $[A - \lambda B]x = 0$  y  $[A - \lambda(B + \lambda^{-1}A)]y = 0$  tienen raíces características ligadas por la expresión  $\lambda = \frac{\lambda}{1 + \lambda}$  y vectores asociados iguales o proporcionales.

(1) Morrison, D. [ 33 ] pág. 98. prop. 3.5.

En efecto, del primer sistema se obtiene

$$\left| A - \frac{t}{1-t} B \right| = 0 ; \quad \lambda = \frac{t}{1-t} ; \quad t = \frac{\lambda}{1+\lambda}$$

Por lo que respecta a los vectores asociados

$$(A - \lambda B)x = 0 ;$$

$$\left( A - \frac{\lambda}{1+\lambda} B \right) y = 0 ; \quad (A + \lambda A - \lambda A - \lambda B)y = 0 ; \quad (A - \lambda B)y = 0 \quad \text{c.q.d.}$$

e) Si  $Y$  es la matriz de vectores asociados del sistema  $[A - t(A+B)]y = 0 \quad \therefore Y^{-1} = E$  las ecuaciones:

$$\begin{aligned} A+B &= E^T E & ; & & A &= E^T F E & \quad (1) \\ A &= E^T F E & ; & & B &= E^T (I-F) E & \quad (24) \end{aligned}$$

definen únivocamente las matrices  $E$  y  $F$  en función de  $A$  y  $B$ , cuando

$$F = \begin{bmatrix} t_1 & 0 & \dots & 0 \\ 0 & t_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & t_p \end{bmatrix} \quad \text{y} \quad Y^{-1} = E : e_{ij} > 0 \quad \forall i$$

$$t_1 > t_2 > \dots > t_p > 0$$

siendo  $e_{ij}$  elementos de la matriz  $E$ .

(1) Estas ecuaciones pueden obtenerse a partir de las soluciones del sistema  $[A - t(A+B)]y = 0$ . Estas satisfacen:

$$AY = (A+B)YF \quad (25)$$

$$Y^T (A+B)Y = I \quad (26)$$

conteniendo esta última ecuación las condiciones de normalización. Operando en (25) y (26)

$$Y^T AY = Y^T (A+B)YF ; \quad \Delta = (Y^T)^{-1} F Y^{-1} = E^T F E$$

$$(Y^T)^{-1} Y^T (A+B) Y Y^{-1} = [Y^T]^{-1} Y^{-1} ; \quad A+B = E^T E$$

El número de variables distintas en las matrices A y B viene dado por:

$$\frac{p(p-1)}{2} + p + \frac{p(p-1)}{2} + p = p^2 + p.$$

expresión que coincide con el número de ecuaciones de (24) y el de variables de las matrices E y F; como consecuencia la transformación (24) es biunívoca

f) La función de densidad conjunta de las variables de las matrices E y F viene dada por la expresión:

$$f(E, F) = 2^p k |E^T E|^{1/2(u+n-p)} e^{-1/2 tr E^T E} \prod_{i=1}^p t_i^{1/2(u-p-1)} \prod_{u_1} (1+t_i)^{1/2(u+n-p)} \prod_{i < j} (t_i - t_j) \quad (27)$$

$$\text{en donde } k = \frac{1}{2^{1/2(u+n)p} \prod_{i=1}^p \Gamma[1/2(u+1-i)] \Gamma[1/2(u+1-i)]}$$

En efecto, si consideramos la transformación definida en (24) la función de densidad conjunta de E y F puede obtenerse en base a la función de densidad conjunta de A y B aplicando el procedimiento general de transformación <sup>(1)</sup>; basándonos en él

$$f(E, F) = f[A(E, F), B(E, F)] \frac{D(A, B)}{D(E, F)} \quad (2)$$

En efecto, puesto que A y B son matrices de variables aleatorias con distribuciones independientes:  $A \in W_p(I_p, u)$ ;  $B \in W_p(I_p, u)$ .

su distribución conjunta viene dada por la expresión

$$f(A, B) = k |A|^{1/2(u-p-1)} |B|^{1/2(u-p-1)} e^{-1/2 tr [A+B]} \quad (29)$$

(1) Véase Fz. Trocóniz, A. [ 12 ] pág. 9.7 y 9.8.

(2) En la notación simbólica de (28)  $f(E, F)$  indica la función de densidad conjunta de las variables E y F,  $f[A(E, F), B(E, F)]$  indica la función de densidad conjunta de las variables de A y B puestas en función de las de E y F, y finalmente  $\frac{D(A, B)}{D(E, F)}$  es el jacobiano de la transformación.

siendo el jacobiano de la transformación<sup>(1)</sup>:

$$\left| \frac{\partial(\Delta, \beta)}{\partial(E, F)} \right| = 2^p |E|^{p+2} \prod_{i=1}^p (t_i - t) \quad (30)$$

Utilizando las expresiones (24) y (29) y (30) se llega a la función:

$$f(E, F) = k |E^T F E|^{1/2(u-p-1)} |E^T(I-F)E|^{1/2(n-p-1)} e^{-1/2 \text{tr} E^T E} \cdot 2^p |E|^{p+2} \prod_{i=1}^p (t_i - t)$$

Finalmente, teniendo en cuenta

$$|E^T F E| = |E^T| |F| |E| = |E^T E| \prod_{i=1}^p t_i \quad (32)$$

$$|E^T(I-F)E| = |I-F| |E^T E| = |E^T E| \prod_{i=1}^p (1-t_i) \quad (33)$$

y sustituyendo estas expresiones en (30) obtenemos (27) c. q. d.

g) La función de densidad marginal  $f(E)$  adopta la forma

$$f(E) = \frac{\prod_{i=1}^p \Gamma[1/2(p+1-i)] 2^p}{2^{1/2(u+n-p)p} \prod_{i=1}^p \Gamma[1/2(u+n+1-i)] (2\pi)^{p/2}} |E^T E|^{1/2(u+n-p)} e^{-1/2 \text{tr} E^T E} \quad (34)$$

Prueba:

Dada la estructura multiplicativa de (27) la función de densidad marginal  $f(E)$  será:

$$f(E) = k_1 \cdot |E^T E|^{1/2(u+n-p)} e^{-1/2 \text{tr} E^T E} \quad (35)$$

por lo que únicamente es necesario determinar su constante. Procederemos a ello.

(1) Deemer and Olkin [ 10 ] pág. 345 a 367.

Para hallar  $k_1$  es preciso calcular la integral :

$$2^p \int_0^{+\infty} \int_0^{+\infty} \dots \int_0^{+\infty} \int_{-\infty}^{+\infty} |E^T E|^{\frac{1}{2}(\mu+n-p)} e^{-\frac{1}{2} \text{tr} E^T E} dE \quad (36)$$

en la cual, (dado que se ha hecho la restricción  $\theta_{ii} > 0 \forall i$  para que la transformación sea "biunívoca") p límites inferiores son iguales a 0. Como la función bajo el signo integral es simétrica respecto al origen, la sustitución de los límites inferiores 0 por  $-\infty$  equivale a multiplicar la expresión (36) por  $2^p$ . Con lo cual se obtiene finalmente:

$$(2\pi)^{\frac{1}{2}p^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |E^T E|^{\frac{1}{2}(\mu+n-p)} \left[ \frac{1}{(2\pi)^{\frac{1}{2}p^2}} e^{-\frac{\sum_i \sum_j e_{ij}^2}{2}} d\mathbf{e}_{ij} \right] \quad (37)$$

Esta expresión constituye el momento de orden  $\frac{1}{2}(\mu+n-p)$  de la varianza generalizada <sup>(1)</sup> de  $E^T E$  cuando esta matriz tiene la distribución de Wishart de parámetros  $W[\mathbf{I}, p]$ . La nota (1) nos permite obtener el resultado de (37) siendo éste:

$$(2\pi)^{\frac{1}{2}p^2} \frac{2^{\frac{1}{2}p(\mu+n-p)} \prod \left[ \frac{1}{2}(\mu+n+1-i) \right]}{\prod \Gamma \left[ \frac{1}{2}(p+1-i) \right]} \quad (38)$$

Utilizando (38) es fácil llegar al valor de  $k_1$  puesto que

$$(35) \quad \int \dots \int \int f(E) dE = 1.$$

y en base a (34)

$$\int_0^{+\infty} \int_0^{+\infty} \dots \int_0^{+\infty} \int_{-\infty}^{+\infty} k_1 |E^T E|^{\frac{1}{2}(\mu+n-p)} e^{-\frac{1}{2} \text{tr} E^T E} = 1 \quad (39)$$

Finalmente, teniendo en cuenta (39)

(1) Anderson, T.W., [ 2 ] pág. 171. El momento de orden  $h$  de la varianza generalizada de una matriz  $A : A \in W[\Sigma, n]$  viene dado por :

$$|\Sigma^h| \prod_{i=1}^p \left[ \frac{2^h \Gamma \left[ \frac{1}{2}(N-i)+h \right]}{\Gamma \left[ \frac{1}{2}(N-i) \right]} \right]$$

Al aplicar esta fórmula a (38) téngase presentes estas sustituciones:

$$\begin{aligned} N &= p+1 \\ h &= \frac{1}{2}(\mu+n-p) \end{aligned}$$

$$(38), (39) \quad k_1 = \frac{\prod \Gamma [\frac{1}{2}(p+1-i)]}{2^{\frac{1}{2}(u+n-2)p} \prod \Gamma [\frac{1}{2}(u+n+1-i)] \prod \frac{1}{2} p^2} \text{ c. q. d. (40)}$$

h) La función de densidad marginal  $f(E)$  adopta la forma

$$f(F) = \frac{\prod \frac{1}{2} p \prod_{i=1}^p \Gamma [\frac{1}{2}(u+u+1-i)] \prod_{i=1}^p f_i^{\frac{1}{2}(u-p-1)} \prod_{i=1}^p (1-f_i)^{\frac{1}{2}(u-p-1)} \prod_{i=1}^p (f_i - f_{i+1})}{\prod [\Gamma [\frac{1}{2}(u+1-i)] \Gamma [\frac{1}{2}(u+1-i)] \Gamma [\frac{1}{2}(p+1-i)]]} \quad (41)$$

en donde las variables  $f_i$  satisfacen  $1 > f_1 > f_2 > \dots > f_p > 0$

La prueba es inmediata teniendo en cuenta que  $f(E, F) = f(E) \cdot f(F)$  en donde  $f(E)$  y  $f(F)$  han sido obtenidas en los apartados f) y g). En cuanto a la condición  $1 > f_1 > f_2 \dots > f_p > 0$  esta viene delimitada por  $f_i = \frac{\lambda_i}{1 + \lambda_i}$

i) Si  $A^* \in W_p[\Sigma, u]$  y  $B^* \in W_p[\Sigma, n]$  (independientes) la función de densidad marginal de  $E^*$  en donde

$$A^* + B^* = E^* E$$

$$A^* = E^* F E$$

$$\therefore E^* = Y^{-1} ; A^* Y^* = (A^* + B^*) Y F \text{ y } Y^{*T} (A^* + B^*) Y = I \quad (42)$$

adopta la forma

$$f(E^*) = \frac{\prod \Gamma [(p+1-i)] \Sigma^{-\frac{(u+n)}{2}} \prod \frac{1}{2} p^2 |E^* E|^{-\frac{1}{2}(u-p)} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*}}{2^{\frac{1}{2}(u+n-2)p} \prod \Gamma [\frac{1}{2}(u+n+1-i)]} \quad (43)$$

El procedimiento de obtención de tal función es totalmente análogo al empleado en el apartado g)

Si partimos de la función de densidad conjunta de  $A^*$  y  $B^*$

$$f(A^* B^*) = p(p, n) p(p, m) |\Sigma|^{-\frac{(m+n)}{2}} |A^*|^{\frac{1}{2}(m-p-1)} |B^*|^{\frac{1}{2}(n-p-1)} \cdot e^{-\frac{1}{2} \text{tr} \Sigma (A^* + B^*)} \quad (44)$$

puede obtenerse la función de densidad conjunta de  $E^*$  y  $F$  sin más que sustituir en (44) las matrices  $A^*$  y  $B^*$  en función de  $E^*$  y  $F$  y multiplicar por el jacobiano de la transformación (1)

$$f(E^*, F) = \frac{2^p p(p, n) p(p, m)}{|\Sigma|^{\frac{m+n}{2}}} |E^{*T} E^*|^{\frac{m+n-p}{2}} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*} \prod_{i=1}^m t_i^{\frac{m-p-1}{2}} \prod_{i=1}^n (t_i + 1)^{\frac{1}{2}(n-p-1)} \quad (45)$$

en donde  $p(p, n)$  y  $p(p, m)$  son las constantes de las funciones de densidad de  $A^*$  y  $B^*$ . Dada la estructura multiplicativa de (45) el hallar la función de densidad marginal de  $E^*$  únicamente requiere hallar la constante de tal función.

Sea

$$f(E^*) = k \cdot |\Sigma|^{-\frac{(m+n)}{2}} |E^{*T} E^*|^{\frac{1}{2}(m+n-p)} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*} \quad (46)$$

la función de densidad marginal de  $E^*$ . De nuevo el momento de orden  $h$  de  $|E^{*T} E^*|$  nos permite obtener la constante  $k$

$$\begin{aligned} 2^p \int_0^\infty \dots \int_0^\infty \int_{-\infty}^{+\infty} |E^{*T} E^*|^{\frac{m+n-p}{2}} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*} dE^* &= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |E^{*T} E^*|^{\frac{m+n-p}{2}} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*} dE^* \\ &= (2\pi)^{\frac{1}{2} p^2} \sum \rho_{1/2} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |E^{*T} E^*|^{\frac{m+n-p}{2}} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} E^{*T} E^*} \cdot \frac{1}{(2\pi)^{\frac{1}{2} p^2} \sum \rho_{1/2}} dE^* \end{aligned}$$

y según la nota (1) de la página 146:

$$= (2\pi)^{\frac{1}{2} p^2} \sum \frac{1}{2} \rho_{1/2} \frac{2^{\frac{1}{2}(m+n-p)} p(p, m+n-p)}{\prod_{i=1}^m \Gamma[\frac{1}{2}(m+n+1-i)] \prod_{i=1}^n \Gamma[\frac{1}{2}(p+1-i)]} \quad (47)$$

(1) El jacobiano de la transformación (42) viene dado por la expresión:

$$\frac{\partial(A^* B^*)}{\partial(E^* F)} = 2^p |E^*|^{p+2} \prod_{i=1}^n (t_i - f_i^2)$$

Ver Deemer and Olkin [ 10 ] pág. 345-367.

Utilizando este resultado en la integral

$$\int_0^{\infty} \int_0^{\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} k \cdot |E^T E|^{-\frac{m+n-p}{2}} e^{-\frac{1}{2} W \Sigma^{-1} E^T E} dE = 1$$

$$k \cdot (2\pi)^{\frac{1}{2} p^2} \sum \frac{w+w}{2} 2^{\frac{1}{2} p(m+n-p)} \frac{\prod \Gamma[\frac{1}{2}(m+n+1-i)]}{2^p \prod \Gamma[p+1-i]} = 1$$

$$\text{y finalmente } k = \frac{\prod \Gamma[p+1-i] \sum^{-\frac{(m+n)}{2}} \prod^{-\frac{1}{2} p^2}}{2^{\frac{1}{2}(m+n-2)} \prod \Gamma[\frac{1}{2}(m+n+1-i)]} \quad \text{c. q. d.}$$

j) Distribuciones condicionadas de variables aleatorias normales  $p+q$ -variantes.

Presentaremos aquí una serie de resultados relativos a las distribuciones condicionadas de variables normales:

- 1º) Las distribuciones condicionadas de una distribución normal no singular que resultan de asignar valores fijos a cualquier conjunto de componentes son también normales <sup>(1)</sup>.
- 2º) Si  $(X, Y)$  es una variable aleatoria  $p+q$ -variante con distribución normal de parámetros  $m$  y  $\Sigma$  la distribución condicionada de  $x$  con relación a  $y$  en donde estas últimas variables son fijas, es normal de parámetros <sup>(1)</sup>:

(1) La demostración puede verse en Fz. Trocóniz [12] pág. 29.8 y 29.9.

$$E(X/Y) = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (y - \mu_2) \quad (48)$$

$$\Sigma_{X/Y} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \quad (49)$$

3º) La expresión (48) constituye la regresión lineal del vector de variables  $X$  sobre  $Y$  cuando  $Y$  es fijo. Además (49) es la matriz de errores de la regresión anterior.

4º) Si los vectores  $X$  e  $Y$  son independientes  $\Sigma_{12} = 0$ ;  $\Sigma_{12} \Sigma_{22}^{-1} = B$  (1)

k) Las matrices

$$Q = D_{12} D_{22}^{-1} D_{21} \quad (50)$$

$$D_{11.2} = D_{11} - D_{12} D_{22}^{-1} D_{21} \quad (51)$$

obtenidas a partir de las particiones de la matriz presentada en (6) siguen distribuciones de Wishart independientes con parámetros

$$Q \in W_p [\Sigma_{11}, q] \quad (52)$$

$$D_{11.2} \in W_p [\Sigma_{11}, n-q] \quad (53)$$

cuando se cumplen las siguientes restricciones: (2)

- a)  $Y$  es un vector de variables no aleatorias y por tanto  $\check{Y}$  es una matriz no aleatoria.
- b) cada columna de la matriz  $\check{X}$  es una variable aleatoria normal  $p$ -variante con la misma matriz de covarianzas siendo las columnas variables aleatorias independientes.
- c)  $E(X/Y) = \rho E_{1n} + BY$  en donde  $B$  es la matriz de parámetros poblacionales en la regresión del vector  $X$  sobre  $Y$  que supondremos nula (3)

(1)  $B$  representa la matriz de coeficientes de regresión del vector  $X$  sobre  $Y$ .

(2) La prueba de este Teorema puede verse en Khirsagar [28] pág. 263

(3) Si  $B=0$  se cumple  $E(X/Y) = E(X)$  por tanto  $Y$  es independiente de  $X$ .  $\rho$  es un vector columna de  $p$  componentes.

$$d) \quad N \gg p+q+1$$

#### 4.4.2. Distribución en el muestreo de las correlaciones canónicas muestrales.

Los resultados obtenidos en la sección 4.4.1. nos permiten llegar a la distribución de las correlaciones canónicas muestrales éstas fueron definidas como las raíces cuadradas con signo positivo de las raíces de :

$$(15) \quad | D_{12} D_{22}^{-1} D_{21} - c^2 D_{11} | = 0 \quad (54)$$

En el apartado 4.4.1. j) se expuso que las matrices

$$(50) \quad Q = D_{12} D_{22}^{-1} D_{21}$$

$$(51) \quad D_{11.2} = D_{11} - D_{12} D_{22}^{-1} D_{21}$$

son independientes y siguen las distribuciones

$$Q \in W_p [\Sigma_{11}, q] ; D_{11.2} \in W_p [\Sigma_{11}, n-q]$$

de aquí se sigue la siguiente distribución de su suma<sup>(1)</sup>

$$Q + D_{11.2} = D_{11} \in W_p [\Sigma_{11}, n] \quad (55)$$

La ecuación (54) también puede escribirse

$$| Q - c^2 (Q + D_{11.2}) | = 0 \quad (56)$$

La expresión (56) satisface las condiciones bajo las cuales ha sido obtenida la distribución de las raíces características de  $|A^* - \lambda B^*| = 0$  en el apartado 4.4.1. h). Por tanto, adaptando a este caso concreto la función (41) llegamos a la distribución en el muestreo de los cuadrados de las correlaciones canónicas muestrales

$$f(c_1^2, \dots, c_p^2) = \frac{\prod_{i=1}^p \Gamma[\frac{1}{2}(n-i+1)] \prod_{i=1}^p c_i^{2 \cdot \frac{1}{2}(q-p-1)} \prod_{i=1}^p (1-c_i^2)^{n-q-p-1}}{\prod_{i=1}^p \Gamma[\frac{1}{2}(n-q-i+1)] \prod_{i=1}^p \Gamma[\frac{1}{2}(p+1-i)] \prod_{i=1}^p \Gamma[\frac{1}{2}(q+1-i)]} \prod_{i < j} (c_i^2 - c_j^2) \quad (57)$$

(1) Véase Morrison, D. [33] pág. 98. prop. 3.5.

Recordaremos que esta función de densidad ha sido obtenida con las restricciones impuestas en j) y por tanto constituye un caso sometido a fuertes limitaciones.

Cuando el vector  $X$  consta de un único componente la expresión (57) nos permite obtener la función de densidad del cuadrado del coeficiente de correlación múltiple en la regresión de  $X$  sobre  $Y$  siendo ésta

$$f(c^2) = \frac{P[\frac{1}{2}n]}{P[\frac{1}{2}(n-q)] P[\frac{1}{2}q]} c^{2\frac{1}{2}(q-2)} (1-c^2)^{\frac{1}{2}(n-q-3)} \quad (58)$$

---

(1) Véase 1.5.6.

#### 4.4.3. Distribución en el muestreo de los vectores de coeficientes canónicos muestrales.

Antes de obtener la distribución en el muestreo de los vectores de coeficientes canónicos muestrales es preciso hacer algunas consideraciones previas:

1. - Unicamente obtendremos la distribución de la matriz de coeficiente  $G$ . Para los sistemas

$$(13) \quad [D_{12} \ D_{22}^{-1} \ D_{21} - c^2 D_{11}] G = 0$$

$$(14) \quad [D_{21} \ D_{11}^{-1} \ D_{12} - c^2 D_{22}] H = 0$$

ya hemos visto que las matrices  $D_{11}$  y  $D_{12} D_{22}^{-1} D_{21}$  siguen distribuciones de Wishart independientes. No ocurre lo mismo con las matrices  $D_{22}$  y  $D_{21} D_{11}^{-1} D_{12}$  ya que esta última es singular motivo por el que no sigue la distribución de Wishart. Esto impide aplicar los métodos expuestos en 4.4.1. para la obtención de la distribución de la matriz  $H$ .

2. - La distribución de  $G$  será obtenida con las restricciones utilizadas en el apartado 4.4.1. j) trasladadas a este lugar.
3. - Los resultados expuestos en 4.4.1. en los que se ha obtenido la distribución de la matriz  $E^*$  son directamente aplicables a la obtención de la distribución de  $G$  aunque para ello es preciso realizar una serie de modificaciones que vamos a exponer a continuación.

Tomando como base los resultados de 4.4.1. k) cuando se cumplen las restricciones allí expuestas son independientes las matrices  $Q$  y  $D_{11.2}$  y se tiene:

$$Q = D_{12} D_{22}^{-1} D_{21} \in W_p [\Sigma_{11}, q]$$

$$D_{11.2} = D_{11} - D_{12} D_{22}^{-1} D_{21} \in W_p [\Sigma_{11}, n-q]$$

La matriz  $G$  de vectores canónicos es solución del sistema

$$[Q - c^2(Q + D_{11.2})]G = 0$$

Puesto que las matrices  $Q$  y  $D_{11.2}$  satisfacen las condiciones exigidas en 4.4.1. la función de densidad de  $E^* = G^{-1}$  viene dada por la expresión

$$f(E^*) = \kappa |E^{*T} E^*|^{\frac{1}{2}(u-p)} e^{-\frac{1}{2} \text{tr} \Sigma_{11}^{-1} E^{*T} E^*} \quad (59)$$

$$; E^* = G^{-1} ; \kappa = \frac{\pi^p \Gamma[(p+1) \frac{u}{2}] \Sigma^{-\frac{(u+p)}{2}}}{2^{\frac{1}{2}(u+p-2)} \pi^p \Gamma[\frac{1}{2}(u+p+1)]} \eta^{-\frac{1}{2} p^2}$$

A partir de (59) es inmediato obtener la función de densidad de los vectores canónicos  $G$ . El jacobiano de transformación de  $G^{-1}$  a  $G$  adopta la forma<sup>(1)</sup>

$$\frac{\partial G^{-1}}{\partial G} = |G|^{-2p} \quad (60)$$

Utilizando esta última expresión puede llegarse a la función de densidad de  $G$  siendo ésta:

$$f(G) = \kappa |G^{-T} G^{-1}|^{\frac{1}{2}(u-p)} e^{-\frac{1}{2} \text{tr} \Sigma_{11}^{-1} G^{-T} G^{-1}}$$

$$= \kappa |G^{-1}|^{-(u+p)} e^{-\frac{1}{2} \text{tr} \Sigma_{11}^{-1} G^{-T} G^{-1}} \quad (61)$$

(1) Véase Deemer and Olkin [10] pág.

en donde 
$$K = \frac{\prod [p(p+1-p)]}{2^{(n-1)p} p^{[1/2(u+1-p)]}} \sum_{ii}^{-n/2} p^{-1/2 p^2}$$

Recordaremos que esta función de densidad ha sido obtenida teniendo en cuenta la restricción impuesta a la matriz E que hace que la transformación empleada sea biunívoca.

#### 4.5. Contrastes de hipótesis en el modelo canónico.

Dos son los tipos de bloques de parámetros en torno a los cuales suelen realizarse contrastes de hipótesis en el modelo canónico teórico

- a) Las correlaciones canónicas
- b) Los vectores de coeficientes canónicos.

En relación a los primeros diremos que los contrastes de hipótesis acerca de las correlaciones canónicas entre dos vectores de variables aleatorias de referencia están íntimamente ligados a los contrastes de independencia de tales vectores bajo las hipótesis de normalidad en su distribución conjunta. Son mutuamente implicables las hipótesis "los vectores  $X$  e  $Y : X, Y \in N_p(\mu, \Sigma)$  son independientes" y "las correlaciones canónicas entre  $X$  e  $Y$  son nulas". Este es el motivo por el cual nos referiremos a algunos contrastes de hipótesis relativos a la independencia de dos vectores con distribución conjunta normal.

Por lo que respecta al segundo tipo de contrastes es preciso hacer algunas aclaraciones previas. El modelo canónico como ya expusimos en el capítulo I constituye la generalización del modelo de regresión lineal. En tal generalización se originan dos nuevos vectores, a los que hemos denominado vectores canónicos, a través de combinaciones lineales de los vectores aleatorios de referencia. - Cuando en dichos vectores se hacen hipótesis de dependencia lineal de unos sobre otros obtenemos los modelos lineales multivariantes. En este tipo de modelos existe un planteamiento previo de causa-efecto que puede ser objeto de contraste. Es bien sabida la relación que liga al modelo canónico con los modelos lineales multivariantes.

Trasladaremos algunos de los contrastes de hipótesis sobre vectores canónicos a los contrastes sobre los coeficientes de regresión; este cambio viene motivado por el mejor conocimiento de las distribuciones de los estadísticos que se utilizan para tales contrastes en el modelo lineal.

En las pruebas estadísticas unilaterales suele ser frecuente el siguiente proceso:

- 1º.- Toma de una muestra y en base a ella construir un estadístico que contenga el parámetro o parámetros sobre los que va a realizarse el contraste.
- 2º.- Obtener la distribución en el muestreo de tal estadístico cuando se cumple la hipótesis del contraste.
- 3º.- Establecer una región crítica con un nivel de significación prefijado.

Este proceso será utilizado en los contrastes de hipótesis a realizar en el modelo canónico si bien las distribuciones teóricas que siguen tales estadísticos utilizados presentan ciertas características que juzgamos conveniente comentar a continuación.

#### 4.5.1. Distribuciones utilizadas en el contraste de hipótesis relativas al modelo canónico.

Estas distribuciones o criterios como comúnmente son denominadas responden a una serie de generalizaciones al caso multivariante de resultados de la estadística univariante. Reproduciremos el siguiente comentario de Ingram Olkin y Herman Rubin: <sup>(1)</sup> Si  $X$  e  $Y$  son dos variables aleatorias independientes con distribuciones

(1) Olkin I. y Rubin, H., [ 36 ] pág. 261.

$\chi^2$  de parámetros  $m$  y  $n$ ; salvo constantes  $\frac{x}{y}$  y  $\frac{x}{x+y}$  están distribuidas como variables  $F$  y  $\beta$ . En el análisis multivariante la distribución de Wishart juega el mismo papel que la distribución  $\chi^2$  en el caso univariante; ahora bien, sin embargo, no hay una generalización natural única de los cocientes anteriores". Señaladas estas características expondremos aquí una serie de distribuciones de probabilidades obtenidas a partir de distribuciones Wishart independientes según la lógica incluida en el comentario anterior.

#### 4. 5. 1. 1. Criterio $\mathcal{L}$ de Wilks

Una de las distribuciones de probabilidades que serán utilizadas en las pruebas de hipótesis a realizar en el modelo canónico es la distribución  $\mathcal{L}$  de Wilks. Si  $A$  y  $B$  son dos matrices de variables aleatorias con distribuciones Wishart independientes

$$\begin{aligned} A &\in W_p [\Sigma, n-q] \\ B &\in W_p [\Sigma, q] \end{aligned} \quad (62)$$

la variable aleatoria  $W$  de la distribución  $\mathcal{L}$  se define por la relación:

$$W = \frac{|A|}{|A+B|} \quad (63)$$

En notación abreviada si la variable  $W$  se obtiene según la expresión (63) indicaremos:

$$W \in \mathcal{L}(p, q, n)$$

En la expresión anterior los parámetros  $p, q, n$  del criterio son respectivamente el orden común de  $A$  y  $B$  los grados de libertad de  $B$  y los grados de libertad de  $A+B$ .

La utilización de esta distribución en los contrastes de hipótesis requiere conocer para cada terna de valores  $p$ ,  $q$  y  $n$  el área de las colas que en cada valor de la variable  $W$  deja su función de densidad.

La primera solución dada a este problema fue a través de aproximaciones a distribuciones tales como  $\chi^2$  y  $F$ . Con posterioridad en la década de los sesenta la obtención de la función de densidad de la variable de la distribución  $\mathcal{L}$  se revela como una cuestión de gran interés ya que son varias las pruebas de hipótesis basadas en estadísticos con distribución  $\mathcal{L}$ . Martin - Schatzoff<sup>(1)</sup> (1966) fue quien primero presentó un procedimiento de obtención de la función de densidad de  $W$  aunque no da esta en forma explícita. Este autor ha realizado una tabulación de los factores de conversión a valores en la distribución  $\chi^2(p, q)$  (para diferentes niveles de significación  $\alpha$ ), de los valores de  $W$ . Con posterioridad K. C. Sreedharan Pillai y Arjun K. Gupta<sup>(2)</sup> (1969) han obtenido la función de densidad de la variable  $W$  ampliando la tabulación hecha por Schatzoff. Dedicaremos el resto de la sección 4.5.1.1.a estudiarla problemática en torno a la distribución  $\mathcal{L}$  de Wilks, exponiendo en primer lugar las aproximaciones que se han efectuado a tal distribución y a continuación la derivación de la función de densidad de la variable  $W$ .

---

(1) Schatzoff, M. [45] págs. 347-358.

(2) Pillai, K. C. y Gupta, A., [39] págs. 109-118.

#### 4.5.1.1.1. Distribución aproximada del criterio $\Lambda$ de Wilks.

Son varios los autores que han obtenido las distribuciones a las que se aproxima la distribución de la variable  $w$  de Wilks. Citaremos las debidas a Bartlett y Rao.

Bartlett<sup>(1)</sup> obtuvo el siguiente resultado

$$- [n - \frac{1}{2}(p+q+1)] \lg_e w$$

se distribuye aproximadamente como  $\chi^2(p \times q)$ . Esta aproximación es suficientemente buena para las aplicaciones prácticas, cuando los parámetros de la distribución  $\Lambda$  satisfacen

$$\frac{w}{3} \geq p^2 + q^2 : w = u - \frac{(p+q+1)}{2}$$

Rao<sup>(2)</sup> ha dado esta aproximación a la distribución  $\Lambda$  :

$$\frac{ws + 2\lambda}{2v} \cdot \frac{1 - w^{1/3}}{w^{1/3}}$$

se distribuye aproximadamente como  $F[2v, ws + 2\lambda]$

en donde

$$v = pq/2 ; \lambda = -(pq-2)/4 ; s = (p^2q^2 - 4)^{1/2} / (p^2 + q^2 - 5)^{1/2}$$

Estas aproximaciones, en su día jugaron un papel importante dentro del análisis multivariante por facilitar la aplicación de una serie de tests basados en el criterio  $\Lambda$ . En la actualidad al estar tabulada tal distribución su uso es menor; no obstante la aproximación debida a Bartlett es de gran precisión.<sup>(3)</sup>

(1) Bartlett, M. S. [ 4 ] pág. 180.

(2) Rao, C. R. [ 40 ] pág. 177.

(3) Khsirsagar, R. [ 28 ] pág. 301.

#### 4.5.1.1.2. Distribución exacta.

La obtención de la función de densidad de las variables  $w$  de Wilks definida en (63) está basada en un conjunto de proposiciones previas que constituyen los puntos de partida de los procedimientos ideados por Schatzoff y Pillai-Gupta para llegar a tal función. Estas proposiciones son las que en forma de teorema exponemos a continuación:

Teorema 1. La variable  $w$  del criterio  $\Lambda$  de Wilks se distribuye como :

$$W = X_1 \cdot X_2 \cdot \dots \cdot X_p \quad (64)$$

en donde  $x_i$  son independientes y tienen distribuciones

$$x_i \in \beta \left[ \frac{n-q+1-i}{2}, \frac{q}{2} \right]$$

Teorema 2.  $W \in \Lambda(p, q, n)$  y  $W \in \Lambda(q, p, nq-p)$  tienen la misma distribución (2)

Teorema 3. La variable  $W \in \Lambda(2r, q, n)$  se distribuye como

$$W = Y_1^2 \cdot Y_2^2 \cdot \dots \cdot Y_r^2$$

en donde  $Y_i; i = 1, 2, \dots, r$  son independientes con distri

(1) La prueba de este Teorema puede verse en Kshirsagar, A. [28] págs. 292-y siguientes. Expondremos aquí una síntesis:

Si se considera la transformación:

$A+B = CC^T, A = CLC^T$  en donde  $C$  es la matriz triangular "lower" de la descomposición de  $A+B$ , la variable  $W$  puede expresarse en función de las variables de la matriz  $L$  como sigue:

$$w = \frac{|A|}{|A+B|} = \frac{|C||L||C^T|}{|C||C^T|} = |L|$$

Un primer paso lo constituye el obtener la función de densidad conjunta de  $CC^T$  y  $L$  y a continuación la marginal de  $L$ . Obtenida esta última una nueva transformación  $L = TT^T$  en donde  $T$  es una matriz triangular "lower" permite demostrar:

$$w = |L| = |TT^T| = \prod_{i=1}^p t_{ii}^2 \quad ; \quad t_{ii} \in \beta \left[ \frac{n-q+1-i}{2}, \frac{q}{2} \right]$$

en donde  $t_{ii}$  son los elementos diagonales de  $T$

(2) Anderson, T.W., [ 2 ] pág. 193.

buciones  $\beta [n-q+1-2i, q]$

La variable  $W \in \mathcal{L}(2s+1, q, n)$  se distribuye como

$$W = z_1^2 \cdot z_2^2 \cdot \dots \cdot z_s^2 \cdot z_{s+1}$$

en donde  $z_i: i=1 \dots s$  están independientemente distribuidas como  $\beta [n-q+1-2i, q]$  y  $z_{s+1}$  está independientemente distribuida como  $\beta [\frac{1}{2}(n-q+1-p), q/2]$  (1)

Teorema 4. Si  $s_1$  y  $s_2$  son dos variables aleatorias:  $s_1 > 0$ ;  $s_2 > 0$  y sus funciones de densidad salvo constantes son de la forma  $s_1^k e^{as_1}$  y  $e^{bs_2}$  la función de densidad de la variable suma  $v = s_1 + s_2$  viene dada por la expresión

$$f(v) = e^{bv} \int_0^v u^k e^{(a-b)u} du$$

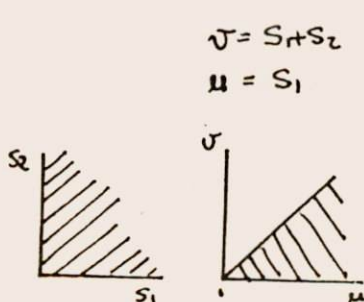
Según los valores de los parámetros  $a$  y  $b$  la función de densidad anterior tiene las siguientes expresiones (2)

$$a=b \quad f(v) = e^{bv} \frac{v^{k+1}}{k+1}$$

$$a \neq b \quad f(v) = e^{av} \left[ \sum_{r=1}^{k+1} (-1)^{r+1} \frac{k!}{(k-r+1)!} \frac{v^{k-r+1}}{(a-b)^r} + e^{bv} \left( \frac{1}{b-a} \right)^{k+1} k! \right] \quad (65)$$

(1) Véase Anderson, T.W. [ 2 ] pág. 195.

(2)  $f(v)$  puede obtenerse a través de una variable auxiliar según el siguiente proceso



$$f(u, v) = u^k \cdot e^{av} e^{(a-b)u} \cdot \frac{1}{v}$$

$$f(v) = \int_0^v u^k e^{bv} e^{(a-b)u} du$$

$$f(v) = e^{bv} \int_0^v u^k e^{(a-b)u} du$$

## a) Procedimiento de Schatzoff.

El procedimiento ideado por este autor para obtener la función de densidad de la variable  $W \in \mathcal{L}(p, q, n)$  está basado en los resultados de los teoremas 1 y 4 de este apartado. En síntesis consiste en transformar la expresión (60) en una suma de variables y a partir de ahí sucesivas aplicaciones del resultado obtenido en el teorema 4, permiten llegar a la función de densidad de  $W$ . Aplicando este proceso Schatzoff ha llegado por inducción a la siguiente conclusión. La función de densidad de la variable  $W$  de la distribución  $\mathcal{L}$  adopta la forma

$$f(w)^{(p)} = \left[ \prod_{i=1}^p K_i \right] \sum_{j=1}^m c_j W^{\frac{1}{2}(n-l_j)} (-\lg w)^{k_j} \quad (66)$$

en donde  $K_i = \frac{1}{\beta \left[ \frac{1}{2}(n-q-l_i+1), \frac{1}{2}q \right]}$

y las constantes  $m$ ,  $l_j$  y  $k_j$  son enteros determinados por  $p$ ,  $q$  y  $n$ . Detallaremos las etapas seguidas para obtener  $f(w)$ .

Cuando  $p=1$  el teorema 1 permite afirmar que  $f(w)$  es la función de densidad de la variable de la distribución  $\beta$ .

Supongamos que (66) se cumple para  $p=N$

Hallaremos la función de densidad  $f(w)$  para  $p=N+1$ . Si en (64) se realiza la transformación:

$$y_i = -\lg x_i \quad ; \quad x_i = e^{-y_i} \\ -\lg w = y_1 + y_2 + \dots + y_p$$

y para este caso concreto

$$-\lg W = y_1 + y_2 + \dots + y_{N+1} \quad (67)$$

(1) Emplearemos el superíndice  $p$  para referirnos a la función de densidad de  $V$  cuando ésta es producto de  $p$  variables aleatorias de la forma (64).

como 
$$f(x_i) = \frac{1}{\beta [\frac{1}{2}(n-q-i+1), \frac{q}{2}]} x^{\frac{1}{2}(n-q-i+1)-1} (1-x)^{\frac{q}{2}-1}$$

la transformación  $y_i = -\log x_i$  conduce a la función de densidad de  $y_i$ :

$$f(y_i) = \frac{1}{\beta [\frac{1}{2}(n-q-i+1), \frac{q}{2}]} e^{-y_i [\frac{1}{2}(n-q-i+1)-1]} \sum_{u=0}^{\frac{q}{2}-1} \binom{\frac{q}{2}-1}{u} (-1)^u [e^{-y_i}]^u e^{-y_i}$$

Si suponemos conocida la función de densidad de la variable  $W$  la expresión (67) puede considerarse como suma de dos variables aleatorias  $-\log w^{(N)}$ ,  $y_{N+1}$  cuyas funciones de densidad son de la forma  $s^k e^{-sx}$  y  $e^{-bx}$ ; por tanto es aplicable a estas funciones el proceso señalado en el teorema 4. para obtener la función de densidad de la variable suma. Operando en este sentido y realizando la consiguiente transformación posterior en (66) Schatzoff llega a la función de densidad  $f(w)^{(N+1)}$  resultando ser ésta de la misma estructura que la expuesta en (66).

La función de densidad de la variable  $W$  ha sido obtenida para el caso en que el parámetro  $q$  es par, hecho que da lugar a que las series particulares de la forma (65) que en este caso intervienen sean finitas. Cuando  $q$  es impar y  $p$  es par, el teorema 2. permite afirmar que la función de densidad de la variable  $W$  sigue siendo de la forma (66) aunque es necesario realizar en ella el correspondiente intercambio de parámetros. Cuando  $p$  y  $q$  son impares no puede aplicarse el proceso que aquí se ha expuesto.

#### b) Procedimiento de Pillai-Gupta.

Pillai y Gupta han realizado una notable mejora del método de Schatzoff en el sentido que disminuyen el número de reiteraciones en la aplicación del teorema 4. Partiendo del resultado del teorema 3. cuando  $p = 1$  y 2 la función de densidad de  $W$  es la de las variables de las distribuciones  $\beta[\frac{n-q}{2}, \frac{q}{2}]$  y cuadrado de  $\beta(n-q-1, q)$  respectivamente.

Si  $p=3$

$$W = X_1 \cdot X_2 \cdot X_3 = z_1 \cdot X_3$$

Las transformaciones  $y_3 = -\lg x_3$  y  $y_1 = \lg z_1$  permiten obtener

$$-\lg W = y_1 + y_3$$

Puesto que las funciones de densidad de  $y_1$  y  $y_3$  son de la forma expuesta en el teorema 4, su aplicación conduce a la obtención de la función de densidad de  $-\lg W$  y a través de la consiguiente transformación a  $f(w) : w \in \mathcal{L}(3, q, n)$

De la misma forma si  $p=4$

$$W = X_1 \cdot X_2 \cdot X_3 \cdot X_4 = z_1 \cdot z_2$$

utilizando  $y_1 = -\lg z_1$  ;  $y_2 = -\lg z_2$

$$-\lg W = y_1 + y_2$$

Una única aplicación del teorema 4, permite llegar a  $f(w) : w \in \mathcal{L}(4, p, n)$

Cuando  $p > 4$  puede llegarse a expresiones análogas pero con más sumandos

Para  $p=5$  
$$-\lg W = y_1 + y_2 + y_3$$

"  $p=6$  
$$-\lg W = y_1 + y_2 + y_3$$

En cualquier caso el número de sumandos de estas expresiones es siempre menor que el que resulta de emplear el procedimiento de Schatzoff, lo cual supone una disminución en el número de reiteraciones en el resultado del teorema 4, para obtener  $f(w)$

Pillai y Gupta han obtenido en forma explícita la función de densidad de la variable  $W \in \mathcal{L}(p, q, n) : p=3,4,5,6$ , independientemente

de que el parámetro  $q$  sea par o impar. Estos autores han completado una tabulación realizada con anterioridad por Schatzoff empleando para valores de  $p$  elevados la aproximación de Rao expuesta en el apartado anterior.

#### 4. 5. 1. 2. Criterios de Pillai

Otros criterios obtenidos a partir de matrices con distribuciones de Wishart independientes en los que están fundamentados algunos contrastes de hipótesis relacionados con el modelo canónico son los debidos a Pillai. Como en el caso del criterio  $\Lambda$  de Wilks el llevar a la práctica tales contrastes requiere conocer las distribuciones teóricas de los criterios.

En este apartado expondremos las definiciones de los cuatro criterios debidos a este autor así como sus distribuciones aproximadas. Supondremos que las matrices A y B que intervienen en sus definiciones poseen distribuciones independientes:  $A \in W_p [\Sigma, n-q]$ ;  $B \in W_p [\Sigma, q]$

##### Criterio I.

La variable aleatoria  $v_1$  del primer criterio de Pillai se define por la relación <sup>(1)</sup>

$$v_1 = \text{tr } B[A+B]^{-1} \quad (71)$$

Si se considera el determinante

$$|B(A+B)^{-1} - \lambda I| = 0 \quad (72)$$

ya que las matrices A y B son definidas positivas y simétricas las raíces de (72) son positivas y si se expresan ordenadas en forma decreciente en una matriz diagonal constituyen la matriz

---

(1) Pillai, K. C. [ 37 ] pág. 495.

diagonalizada de  $B(A+B)^{-1}$ . Puesto que la diagonalización conserva la traza, otra forma de expresar el criterio de Pillai, es como sigue:

$$v_1 = \sum_{i=1}^p \lambda_i^2 \quad (73)$$

La distribución de este criterio ha sido estudiada por su autor habiendo llegado al siguiente resultado: la distribución de la variable  $v_1$  se aproxima a la distribución  $\beta$  de primera clase siendo su función de densidad

$$k^* v_1^{p(2u+p+1)/2-1} \left(1 - \frac{v_1}{p}\right)^{p(2k+p+1)/2-1} \quad (74)$$

$$: k^* = \left[ p^{p(2u+p+1)/2} \beta \left[ \frac{p}{2}(2k+p+1); \frac{p}{2}(2u+p+1) \right] \right]^{-1}$$

$$u = \frac{1}{2}(q-p-1) ; \quad k = \frac{1}{2}(n-p-q-2)$$

En la práctica esta aproximación es buena si cuando  $p=2$  ;  $m+k \gg 30$ . Para valores de  $p > 2$  , las condiciones para utilizar tal aproximación requieren que por cada aumento de  $p$  en una unidad  $m+k$  lo haga en 10.

En una publicación posterior <sup>(1)</sup> Pillai estudia las distribuciones de ciertas funciones de las raíces de la ecuación (72). Teniendo en cuenta que el criterio  $\Lambda$  de Wilks es la potencia  $p$ -ésima de la media geométrica de tales raíces y que su primer criterio es  $p$  veces la media aritmética establece tres nuevos criterios basados en la media armónica y otros tipos de funciones de las raíces anteriores. A estos criterios los denominaremos II-III y IV de Pillai. Los expondremos a continuación:

---

(1) Pillai, K. C. [38] págs. 117-119.

### Criterio II.

La variable aleatoria  $v_2$  del segundo criterio de Pillai se define por la relación

$$v_2 = \frac{p}{\sum_{i=1}^p \frac{1}{1-\lambda_i}} \quad (75)$$

en donde  $\lambda_i: i=1 \dots p$  son las raíces de la ecuación (72).

Pillai ha demostrado que la distribución de la variable  $v_2$  se aproxima a la que tiene por función de densidad

$$f(v_2) = \frac{v_2^{pk} (1-v_2)^{p(2u+p+1)/2-1}}{\beta [pk+1, \frac{1}{2} p(2u+p+1)]} \quad (76)$$

en donde,  $k = \frac{1}{2}(n-p-q-2)$  ;  $u = \frac{1}{2}(q-p-1)$

La condición para utilizar esta aproximación es análoga a la expuesta en el criterio I haciendo cumplir a  $k+p$  los requisitos exigidos a  $m+k$ .

### Criterio III.

La variable aleatoria  $v_3$  del tercer criterio de Pillai queda definida según la siguiente relación

$$v_3 = \frac{p}{\sum_{i=1}^p \frac{1}{\lambda_i}} \quad (77)$$

en donde,  $\lambda_i: i=1 \dots p$  son las raíces de la ecuación (72).

También su autor ha estudiado la distribución aproximada de  $v_3$  llegando al siguiente resultado:

$$f(\sigma_3) = \frac{\sigma_3^{p\omega} (1-\sigma_3)^{p(2k+p+1)/2-1}}{\beta [p\omega+1, p/2(2k+p+1)]} \quad 0 < \sigma_3 < 1 \quad (78)$$

La condición para aplicar esta aproximación requiere que  $m+p$  cumpla los requisitos exigidos a  $m+k$  en el criterio I.

Criterio IV.

Se define la variable aleatoria  $v_4$  de este criterio como:

$$v_4 = \frac{p}{\sum 1/\theta_i} \quad (79)$$

en donde,  $\theta_i : i = 1, 2, \dots, p$  son las raíces de la ecuación

$$|BA^{-1} - \theta I| = 0 \quad (80)$$

La distribución de la variable  $v_4$  puede obtenerse de la anterior teniendo en cuenta que  $v_4 = \frac{\sigma_3}{1-\sigma_3}$ ; esta resulta ser

$$f(v_4) = \frac{\sigma_3^{p\omega}}{\beta [p\omega+1, p/2(2k+p+1)]} (1+\sigma_4)^{-[p(2\omega+2k+p+1)/2+1]} \quad 0 < v_4 < \infty \quad (81)$$

con idénticas condiciones para su utilización que las exigidas en el criterio III.

Las funciones de densidad de las distribuciones a que se aproximan las variables de los cuatro criterios están tabuladas por lo cual no presenta ningún inconveniente su utilización en las pruebas de hipótesis basadas en estadísticos con tales distribuciones. En apartados posteriores incluiremos la utilización de estas distribuciones en el contraste de hipótesis relativas al modelo canónico

#### 4.5.1.3. Criterio de Hotelling.

Si A y B son matrices con las características señaladas en el apartado 4.4.1. a) se define la variable aleatoria del criterio de Hotelling a través de la siguiente expresión:

$$t = \text{tr } BA^{-1} \quad (82)$$

En notación abreviada diremos que  $t \in \tau_0^2$

De forma análoga a los criterios anteriores esta variable puede expresarse en términos de las raíces características de una ecuación determinante: en este caso

$$|BA^{-1} - \theta I| \quad (83)$$

La variable de este criterio en función de tales raíces resulta ser:

$$t = \sum_{i=1}^p \theta_i$$

pudiendo expresarse también <sup>(1)</sup>

$$t = \sum_{i=1}^p \frac{\lambda_i^0}{1 + \lambda_i}$$

en donde,  $\lambda_i : i=1, 2, \dots, p$ . son las raíces de la ecuación (82).

La distribución de la variable t definida en (82) ha sido estudiada por varios autores llegando a la función de densidad exacta para ciertos casos particulares. En relación a las aproximaciones una que resulta de gran utilidad es la siguiente: <sup>(2)</sup>

$$\text{tr } BA^{-1} \rightarrow \chi^2(pq)$$

(1) Véase 4.4.1. b).

(2) Véase Morrison, D. [33], pág. 198.

#### 4.5.1.4. Criterio de Roy

Se define la variable aleatoria del criterio de Roy como la mayor raíz de la ecuación:

$$| \beta - \lambda (A + \beta) | = 0 \quad (84)$$

en donde, A y B son matrices con las características señaladas en el apartado 4.4.1. a)

La distribución de la mayor raíz de la ecuación (84) ha sido obtenida por Roy a partir de la distribución conjunta de las raíces de esta ecuación. Con posterioridad D. N. Nanda ha obtenido la distribución de la mayor, menor y cualquier raíz intermedia de (84). En un primer trabajo <sup>(1)</sup> halló las distribuciones de tales raíces para valores de  $p = 2, 3, 4$ , y 5, en base a considerar el último producto de (41) como un determinante de Vandermonde

$$\prod_{i > j} (\lambda_i - \lambda_j) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_p \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{p-1} & \lambda_2^{p-1} & \dots & \lambda_p^{p-1} \end{vmatrix} \quad (85)$$

A través de sucesivas transformaciones llegó a la distribución de la mayor raíz característica de (84). En una segunda publicación <sup>(2)</sup> este autor presenta soluciones asintóticas de la distribución anterior.

Hecks <sup>(3)</sup> ha tabulado esta distribución facilitando así su aplicación en determinados contrastes de hipótesis.

(1) Nanda, D. N. [ 34 ] pág. 48-57.

(2) Nanda, D. N. [ 35 ] pág. 340-350.

(3) Hecks, D. L. [ 18 ] pág. 625-642.

Las distribuciones de todos los criterios incluidos en el apartado 4.5.1. se han estudiado para el caso que suele conocerse en la literatura estadística como "null case". Esta denominación está fundamentada en el hecho de haber supuesto que la matriz  $B$  sigue la distribución de Wishart centrada. Hasta aquí al hablar de tal distribución nos hemos referido siempre a este caso es decir, a la distribución de la matriz  $B = E^T E$  en que cada columna de  $E$  poseen distribuciones normales multivariantes independientes de vector de medias nulo.

La distribución de estos criterios también puede derivarse cuando la matriz  $B$  ha sido obtenida como producto de dos matrices de variables aleatorias de la forma  $E^T E$  pero cuyas columnas no son de media nula. En este caso suele decirse que la distribución de  $B$  es Wishart no centrada. Si se obtienen las distribuciones de los criterios anteriores utilizando una matriz  $B$  de las características anteriores suele decirse que tales distribuciones se refieren al "non null case".

La distribución de los criterios expuestos en 4.5.1. con las características señaladas en el párrafo anterior ha sido obtenida sólo en alguno de ellos y plantea serias dificultades. En los tests de hipótesis que utilizaremos en el modelo canónico únicamente nos referiremos a la distribución de los criterios como "null case". El motivo de este comentario ha sido únicamente el poner de manifiesto algunas de las restricciones supuestas en la obtención de las distribuciones de los criterios anteriores y la de señalar un campo de trabajo cuya teoría no ha alcanzado su desarrollo completo.

#### 4. 5. 2. Contrastes de hipótesis relativos a las correlaciones canónicas

Suele ser frecuente en los contrastes de hipótesis acerca de las correlaciones canónicas establecer las siguientes etapas :

- 1ª. - Contraste de la hipótesis de que todas las correlaciones canónicas poblacionales son nulas.
- 2ª. - Si la hipótesis anterior es rechazada se llevará a cabo el contraste de la hipótesis de que las últimas  $p-1$  correlaciones canónicas poblacionales son nulas. Se continua el proceso en la medida que van siendo rechazadas las hipótesis de la forma anterior hasta el contraste de la hipótesis de que la última correlación canónica es nula.

Nos referiremos en primer lugar al contraste de la hipótesis del apartado 1ª, y a continuación en forma genérica al resto.

##### 4. 5. 2. 1. Contraste de la hipótesis de que todas las correlaciones canónicas poblacionales son nulas.

Cuando  $X$  e  $Y$  son dos vectores de variables aleatorias con distribución conjunta normal las siguientes hipótesis son equivalentes:

- a) Los vectores  $X$  e  $Y$  son independientes.
- b) Los coeficientes de los modelos de regresión lineal multivariante de  $X$  sobre  $Y$  e  $Y$  sobre  $X$  son nulos.
- c) Las correlaciones canónicas poblacionales entre  $X$  e  $Y$  son nulas.

Para realizar el contraste de cualquiera de las hipótesis anteriores tendremos en cuenta alguno de los resultados expuestos en 4.4.1. tal como:

Las matrices  $Q$  y  $D_{11.2}$  tienen las distribuciones:

$$\begin{aligned} Q &= D_{12} D_{22}^{-1} D_{21} \in W_p [\Sigma_{11}, q] \\ D_{11.2} &= D_{11} - D_{12} D_{22}^{-1} D_{21} \in W_p [\Sigma_{11}, n-q] \end{aligned}$$

y además tales matrices son independientes cuando las particiones de la matriz  $D$  se han obtenido con muestras de tamaño  $N = n+1$  procedentes de dos vectores de variables aleatorias independientes con distribución conjunta normal. Cuando ésto es así, el estadístico

$$\frac{|D_{11.2}|}{|D_{11.2} + Q|} \in \mathcal{L}(p, q, n) \quad (86)$$

Puesto que las correlaciones canónicas son soluciones del determinante

$$|D_{12} D_{22}^{-1} D_{21} - c^2 D_{11}| = 0 \quad (87)$$

teniendo en cuenta las expresiones de  $Q$  y  $D_{11.2}$  en términos de las particiones de  $D$  (86) puede expresarse:

$$\prod_{i=1}^p (1 - c_i^2) \in \mathcal{L}(p, q, n) \quad (88)$$

Procediendo de forma similar puede llegarse a otros estadísticos que sean función de las correlaciones canónicas muestrales cuyas distribuciones en el muestreo cuando se cumple la hipótesis del contraste <sup>(1)</sup> sean las de los criterios expuestos en 4.5.1.1.

El estadístico:

$$\text{tr } Q [Q + D_{11.2}]^{-1} = \text{tr } D_{12} D_{22}^{-1} D_{21} D_{11}^{-1} = \sum_{i=1}^p c_i^2 \quad (89)$$

(1) En este caso el cumplimiento de la hipótesis del contraste va incluido en suponer que  $Q$  y  $D_{11.2}$  siguen distribuciones de Wishart independientes.

se distribuye en el muestreo según el primer criterio de Pillai

Igualmente

$$\frac{p}{\sum_{i=1}^p \frac{1}{1-c_i^2}} \quad (90)$$

se distribuye según el segundo criterio de Pillai

$$\frac{p}{\sum_{i=1}^p \frac{1}{c_i^2}} \quad (91)$$

se distribuye según el tercer criterio de Pillai

$$\frac{p}{\sum_{i=1}^p \frac{1-c_i^2}{c_i^2}} \quad (92)$$

se distribuye según el cuarto criterio de Pillai

$$\sum_{i=1}^p \frac{c_i^2}{1+c_i^2} \quad (93)$$

se distribuye según el criterio de Hotelling

Cualquiera de los estadísticos (88), (89), (90), (91), (92) y (93) puede utilizarse para el contraste de la hipótesis de que las correlaciones canónicas poblacionales entre los vectores  $X$  e  $Y$  son nulas. El procedimiento a seguir sería el que exponemos a continuación:

A partir de una muestra de tamaño  $N \gg p+q+1$  de los dos vectores de variables se calculan las correlaciones canónicas muestrales y en base a ellas el valor que toman los estadísticos (88)...(93). Tales valores se comparan con el cuantil  $1-\alpha$  de las distribuciones de probabilidades que siguen en el muestreo el estadístico utilizado. Si el valor obtenido en el estadístico supera al valor de la distribución se rechazará al nivel de sig que las muestras obtenidas proceden de vectores con correlaciones canónicas poblacionales nulas. La cuestión así planteada es simple y de apli-

cación inmediata cuando las distribuciones teóricas están tabuladas. Una cuestión a tener en cuenta en este tipo de contrastes es la siguiente: ¿Cuál de los tests expuestos se empleará para el contraste de la hipótesis de que las correlaciones canónicas entre dos vectores son nulas? Enjuiciaremos esta interrogante bajo el siguiente comentario de Pillai<sup>(1)</sup> "es conveniente que los estadísticos a emplear estén basados en funciones de las raíces muestrales por ser estas últimas invariantes ante ciertas transformaciones lineales; otra cuestión a tener en cuenta es la disposición o no de grandes muestras que puede afectar a las aproximaciones que se hagan en la distribución del estadístico. Además hay que considerar propiedades tales como insesgadez, admisibilidad, etc.

Hemos dejado para el final de este apartado el contraste de la hipótesis de independencia entre dos vectores de variables basado en el criterio de Roy. El test para tal contraste está construido en base al principio de unión-intersección. Lo expondremos a continuación:

Consideremos dos combinaciones lineales de los vectores de referencia.

$$\begin{aligned} \mu &= a^T x \\ \nu &= b^T y \end{aligned} : \Sigma_{xy} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (94)$$

El coeficiente de correlación entre las combinaciones lineales anteriores es función de a y b

$$\lambda(a,b) = \frac{a^T \Sigma_{12} b}{\sqrt{a^T \Sigma_{11} a \quad b^T \Sigma_{22} b}} \quad (95)$$

La hipótesis de que  $\Sigma_{12}=0$  es equivalente a que todos los coeficientes de correlación entre las infinitas combinaciones lineales de la forma (94) son nulas.

(1) Pillai, K. C. [33] pág. 117 -121.

La estimación del coeficiente de correlación (95) viene dada por la expresión:

$$r(a,b) = \frac{a^T S_{12} b}{\sqrt{a^T S_{11} a \quad b^T S_{22} b}} \quad (96)$$

Aceptaremos la hipótesis de independencia si  $r^2(a,b) \leq r_p^2$  siendo  $r_p^2$  el cuantil  $1 - \alpha$  de la distribución teórica del coeficiente de correlación  $r^2$  cuando la hipótesis del contraste es cierta, para un  $\alpha$  prefijado. La región de aceptación para la hipótesis anterior es la intersección:

$$\bigcap_{a,b} [r^2(a,b)] \leq r_p^2 \quad (97)$$

siendo equivalente a

$$\max_{a,b} r^2(a,b)$$

El máximo valor de  $r^2(a,b)$  se obtiene cuando  $r$  es la mayor correlación canónica muestral entre  $X$  e  $Y$  con lo cual queda justificado la utilización de  $C^2$  como estadístico para el contraste de la hipótesis de independencia entre dos vectores.

#### 4.5.2.2. Contraste de la hipótesis de que parte de las correlaciones canónicas poblacionales son nulas.

Cuando la hipótesis de que todas las correlaciones canónicas poblacionales son nulas ha sido rechazada puede ensayarse la hipótesis de que las últimas  $p-r$  correlaciones canónicas son nulas.

En la definición del modelo canónico muestral se puso de manifiesto cómo la matriz  $D_{12}$  es de rango  $p \leq q$  aunque ésta se haya obtenido con muestras procedentes de dos vec-

tores de variables aleatorias con matriz  $\Sigma_{12}$  de rango  $r < p$ . Nos corresponde en este apartado fundamentar el contraste de la hipótesis de que dichas muestras proceden de vectores con sólo  $r$  correlaciones canónicas poblacionales no nulas.

El procedimiento para obtener los tests que nos permitan realizar el contraste de esta hipótesis es similar al expuesto en 4.5.2.2. Nos centraremos aquí en los tests basados en el criterio  $\Lambda$  de Wilks.

El estadístico<sup>(1)</sup>

$$\sum_{i=r+1}^p (1-c_i^2) \in \mathcal{L}(p-r, q-r, n-r) \quad (98)$$

lo cual permite llevar a cabo el anterior contraste. Además la distribución de este estadístico se aproxima a  $\chi^2$  según la siguiente relación :  $-(n-\frac{1}{2}(p+q+1)) \log \prod_{i=r+1}^p (1-c_i^2) \rightarrow \chi^2_{(p-r) \times (q-r)}$

Lawley ha estudiado una serie de aproximaciones para la distribución del estadístico (98) siendo un caso particular de éstas la obtenida por Bartlett. Citaremos las dos siguientes:

13. - Cuando X e Y son dos vectores aleatorios<sup>(2)</sup>

$$-\left[ n-r-\frac{1}{2}(p+q-1) + \sum_{i=1}^r \frac{1}{\lambda_i^2} \right] \log \prod_{i=r+1}^p (1-c_i^2) \rightarrow \chi^2_{[(p-r) \times (q-r)]} \quad (99)$$

Si las primeras correlaciones canónicas poblacionales son "1".  $\sum_{i=1}^r \frac{1}{\lambda_i^2} = r$  con lo cual la expresión (99) coincide con la de Bartlett.

(1) Véase Kshirsagar, A. [28], pág.326.

(2) Lawley, D.N. [29], pág. 128-136.

2ª. - Si  $Y$  es un vector no aleatorio la aproximación obtenida por Lawley es en este caso (1)

$$- \left[ n - \frac{1}{2}(p+q+1) + \sum_{i=1}^r \frac{1}{\beta_i^2} \right] \lg \prod_{i=1}^p (1-c_i^2) \rightarrow \chi^2 [(p-v) \times (q-v)] \quad (100)$$

en donde  $\beta_i$  son los elementos  $b_{ii}$  correspondientes a la matriz de los coeficientes de regresión de  $X$  sobre  $Y$ . Cuando los elementos de  $b_{ii}$  son de elevado valor  $\sum \frac{1}{\beta_i^2}$  puede considerarse despreciable. Si en lugar de  $\beta_i$  se sustituyen los valores estimados;  $\hat{\beta}_i = \frac{c_i^2}{1-c_i^2}$  con lo cual:

$$\left[ n - \frac{1}{2}(p+q+1) + \sum_{i=1}^r \frac{c_i^2}{1-c_i^2} \right] \lg \prod_{i=1}^p (1-c_i^2) \rightarrow \chi^2 [(p-v) \times (q-v)] \quad (101)$$

#### 4.5.3. Contrastes de hipótesis relativas a los coeficientes del modelo de regresión lineal multivariante de $X$ sobre $Y$

En 4.5.2.1. hemos señalado que las hipótesis de nulidad de coeficientes de correlación canónica poblacionales entre  $X$  e  $Y$  y nulidad de los coeficientes de regresión en el modelo de regresión lineal multivariante de  $X$  sobre  $Y$  son equivalentes, por tanto, para contrastar la hipótesis  $B = [0]$  (2) nos referiremos a 4.5.2.1.

Los contrastes acerca de los coeficientes de regresión que aquí estudiaremos serán los siguientes:

1º. - Si  $B = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_a & \beta_b \end{bmatrix}$  la matriz  $\beta_a$  es una matriz predeterminada  $\beta_c^*$

2º. - Como caso particular del anterior la matriz  $\beta_a = [0]$

(1) Véase Lawley, D.N. [ 30 ] pág. 66.

(2) Denominaremos  $B$  a la matriz de coeficientes del modelo de regresión lineal multivariante de  $X$  sobre  $Y$ . Tal matriz en términos de las particiones de la matriz de covarianzas adopta la forma  $B = \Sigma_{12} \Sigma_{22}^{-1}$ . Véase 2.1.

4.5.3.1. Contraste de la hipótesis de que la matriz  $B_a$  es una matriz prefijada  $B_a^*$

Diseñaremos la prueba estadística para el contraste de esta hipótesis mediante el procedimiento de razón de verosimilitudes <sup>(1)</sup>. Tal procedimiento consiste en obtener el estadístico para realizar el contraste como cociente entre el máximo de la función de verosimilitud de la muestra y el máximo de la función de verosimilitud de la muestra cuando se cumple la hipótesis del contraste. Para el caso que aquí nos ocupa dichos máximos resultan ser: <sup>(2)</sup>

$$Max.L = \frac{|D_{11.2}|^{1/2 N} e^{-1/2 PN}}{(2\pi)^{-1/2 PN}} ; D_{11.2} = D_{11} - D_{12} D_{22}^{-1} D_{21} \quad (102)$$

y

$$Max.L^* = \frac{|D_{11.2} + (\hat{\beta}_a - \beta_a^*)^T D_{aa.b} (\hat{\beta}_a - \beta_a^*)|^{1/2 N} e^{-1/2 PN}}{(2\pi)^{-1/2 PN}} \quad (103)$$

en donde  $D_{22} = \begin{bmatrix} D_{aa} & D_{ab} \\ D_{ba} & D_{bb} \end{bmatrix}$  y  $\hat{\beta}_a = [D_{11a} - D_{11b} D_{bb}^{-1} D_{11a}]^{-1} D_{11a.b}$  ;  $D_{11.2} = [D_{11a} | D_{11b}]$

$$: D_{aa.b} = D_{aa} - D_{ab} D_{bb}^{-1} D_{ba}$$

Obtendremos el estadístico para contrastar la hipótesis  $B_a = B_a^*$  a partir del cociente entre (102) y (103). Si denominamos  $U$  a tal cociente

$$\lambda = U^{\frac{3}{N}} = \frac{|D_{11.2}|}{|D_{11.2} + (\hat{\beta}_a - \beta_a^*)^T D_{aa.b} (\hat{\beta}_a - \beta_a^*)|} = \frac{|D_{11.2}|}{|D_{11.2}^*|} \quad (104)$$

Para llevar a cabo el contraste de la hipótesis anterior es necesario conocer la distribución en el muestreo del estadístico (104). Puesto que las matrices  $D_{11.2}$  y  $D_{11.2}^*$  poseen distribuciones Wishart independientes: <sup>(3)</sup>

(1) Veáse Fz. Trocóniz, A. [ 12 ] Cap. 35.

(2) Veáse Anderson, T.W. [ 2 ] pag. 188

(3) Veáse Anderson, T.W. [ 2 ] pág. 191

$$D_{11.2} \in W_p [\Sigma_{11}, n-q]$$

$$D_{11.2}^* \in W_p [\Sigma_{11}, q_2]$$

el estadístico  $\lambda$  tiene en el muestreo la distribución . . .

$$\lambda \in \mathcal{L} (p, q_1, n)$$

Fijado este resultado el contraste se llevará a cabo como sigue: a partir de una muestra de tamaño  $N$  se evaluará el estadístico (104) si el valor resultante supera al cuantil  $1-\alpha$  de la distribución  $\mathcal{L}$  para un nivel de significación determinado  $\alpha$  se rechazará a tal nivel la hipótesis de que la muestra procede de dos vectores de variables aleatorias en que los coeficientes  $B_a$  de su modelo lineal asociado son la matriz  $B_a^*$ .

4.5.3.2. Contraste de la hipótesis de que la matriz de coeficientes de regresión  $B_a \equiv [0]$

Como ya indicábamos en la introducción este contraste constituye un caso particular del anterior en que la matriz  $B_a^*$  prefijada es una matriz de ceros. Con él se trata de averiguar qué componentes del vector  $\beta$  y puede considerarse que no tienen ninguna contribución en la "explicación" de la matriz de covarianzas del vector  $X$ . El estadístico a utilizar es análogo al anterior si bien la matriz  $D_{11.2}^*$  se reduce en este caso concretamente a

$$D_{11.2}^* = D_{11.2} + (\hat{B}_a D_{aa.b} \hat{B}_a)$$

Daremos una justificación intuitiva de tal estadístico: el numerador contiene el determinante de la matriz de errores de regresión cuando no se ha efectuado ninguna restricción en los coeficientes. Es decir la diferencia entre las matrices  $D_{11}$  y  $D_{12} D_{22}^{-1} D_{21}$ . El denominador su

pone la diferencia entre  $D_{11}$  y la reconstrucción de esta matriz por la regresión de  $X$  sobre  $y_{q_1}, y_{q_2}, \dots, y_{q_r}$ ; es decir  $D_{11} - D_{11}^* D_{11}^{-1} D_{11}$ . Si los componentes  $y_1, y_2, \dots, y_{q_r}$  son independientes de  $X$ , esta matriz muestral tendrá unos elementos próximos a los de  $D_{12} D_{11}^{-1} D_{21}$  y por tanto una evaluación del estadístico conducirá a un resultado próximo a la unidad. Una prueba estadística más concreta se realizará comparando tal cociente con el valor del criterio  $\mathcal{L}(p, q, n)$  para un determinado nivel de significación. Si el cociente supera al valor en la distribución  $\mathcal{L}$  se rechazará a tal nivel que los coeficientes de regresión  $B_a$  son nulos.

Haremos un último comentario en torno a los coeficientes del modelo de regresión lineal multivariante de  $Y$  sobre  $X$ . Dificultades de tipo práctico son las que impiden el diseño de pruebas para el contraste de hipótesis relativas a tales coeficientes. Estas dificultades son las mismas que en 4.4. nos impidió obtener la distribución en el muestreo de los vectores canónicos  $H$  y pueden sintetizarse en que el numerador y denominador del estadístico (104) no siguen distribuciones Wishart. No tenemos referencia de que se haya estudiado la distribución en el muestreo del estadístico que resulta de dividir los máximos de la función de verosimilitud en el caso que aquí hemos mencionado.

CAPITULO V: Los modelos muestrales de los modelos ca-  
nónicos parciales.

### 5. 1. Introducción

Si el modelo muestral en el análisis canónico clásico ha tenido un tratamiento adecuado por lo menos en el "null case" y con distribuciones normales no podemos decir lo mismo de los modelos muestrales asociados a los incluidos en el análisis canónico parcial. La tónica seguida por sus autores ha sido la de exponer el modelo matemático y si acaso referirse a algún test de hipótesis en relación a los parámetros poblacionales.

Creemos que han sido dos las causas del estado actual que presenta la estadística de la inferencia en estos modelos; su reciente aparición dentro del análisis multivariante y las dificultades de estudio que este tipo de desarrollo lleva consigo. En relación a la primera diremos que modelos canónicos tales como los  $G_1$  y  $G_2$  expuestos en el capítulo II - datan de fechas todavía muy recientes por lo que es presumible que se esté todavía trabajando en las distribuciones en el muestreo de los estimadores de sus parámetros. En cuanto a la segunda causa es bien sabido los problemas de tipo matemático que se plantean en el desarrollo de toda teoría muestral dentro del análisis multivariante, en este caso tales problemas se ven acrecentados al referirse a vectores de variables aleatorias obtenidos de los originales a través de sucesivas transformaciones.

Daremos por descontado que todos los desarrollos efectuados en este tipo de modelos se han realizado bajo la hipótesis de normalidad en la distribución de probabilidades conjunta de los vectores de variables aleatorias que intervienen. Esta hipótesis goza de singular importancia en la interpretación de los resultados obtenidos en los modelos parciales en términos de las matrices de covarianzas de las distribuciones condicionadas.

Distinguiremos en este capítulo dos partes bien diferenciadas: La primera de ellas estará dedicada al estudio del modelo muestral del análisis canónico parcial de Rao, R. B. Este autor en la exposición de tal modelo únicamente hace referencia al procedimiento de obtención de estimadores de máxima verosimilitud para las correlaciones canónicas parciales. Posteriormente Timm y Carlson presentan un test de hipótesis para el contraste de que las correlaciones canónicas parciales poblacionales son nulas. Completaremos en este trabajo una serie de cuestiones relacionadas con el modelo muestral del análisis canónico que fue expuesto por Rao; expondremos el procedimiento de llegar a estimadores de máxima verosimilitud de los vectores canónicos parciales así como las distribuciones en el muestreo de los estimadores de máxima verosimilitud de las correlaciones canónicas parciales y de los vectores canónicos obtenidas con un conjunto de restricciones análogas a las que se utilizaron en el 4.4.2 y 4.4.3; también expondremos un conjunto de pruebas de hipótesis relativas a las correlaciones canónicas parciales; éstas estarán basadas en los criterios introducidos en el capítulo IV.

La segunda parte la dedicaremos a exponer los resultados muestrales a que se ha llegado en el resto de modelos canónicos parciales.

Aunque este capítulo está dedicado a la inferencia en los modelos parciales no queremos terminar este trabajo sin hacer una referencia al modelo muestral del análisis canónico generalizado. Citaremos un comentario de Paul Horst a tal efecto "ningún intento ha sido hecho en obtener test de significación para es

tos modelos; es probable que tales derivaciones serían de bastante dificultad". Aunque ha transcurrido cierto tiempo desde que Horst presentó sus modelos canónicos generalizados no hemos encontrado en la literatura sobre el tema ninguna referencia que trate la cuestión aquí considerada.

Puesto que toda la problemática de los modelos canónicos generalizados gira en torno a la interdependencia entre más de dos vectores de variables aleatorias exponenciales, presentaremos los tests de hipótesis para llevar a cabo el contraste de tal independencia.

## 5.2. El modelo muestral del análisis canónico parcial de Rao.

### 5.2.1. Consideraciones preliminares

Sean  $(X^T, Y^T, Z^T)$  vectores de variables aleatorias definidas conjuntamente con distribución  $N(\omega, \Sigma)$  de dimensiones  $p \times q \times r$ :  $p \leq q$  y consideremos una muestra de tamaño  $N = n+1 \gg p+q+r$ . Dicha muestra puede ser expresada en la siguiente matriz:

$$W = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \dots & x_{np} \\ y_{11} & y_{21} & \dots & y_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1q} & y_{2q} & \dots & y_{nq} \\ z_{11} & z_{21} & \dots & z_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1r} & z_{2r} & \dots & z_{nr} \end{bmatrix} = \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} \quad (1)$$

Teniendo en cuenta las matrices introducidas en el apartado 4.1. la matriz de covarianzas muestral entre  $(X^T, Y^T, Z^T)$  da lugar a:

$$S = \frac{1}{N} W \left[ I - \frac{E_{NN}}{N} \right] W^T \quad (2)$$

Denominaremos  $D$  a la matriz

$$D = W \left[ I - \frac{E_{NN}}{N} \right] W^T \quad (3)$$

$$S = \frac{D}{N} \quad (4)$$

La expresión particionada de la matriz  $W$  permite obtener las particiones de la matriz definida en (3)

$$D = \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix} \left[ I - \frac{E_{NN}}{N} \right] \begin{bmatrix} \bar{X}^T & \bar{Y}^T & \bar{Z}^T \end{bmatrix} \quad (5)$$

$$D = \begin{bmatrix} \bar{X} \left[ I - \frac{E_{NN}}{N} \right] \bar{X}^T & \bar{X} \left[ I - \frac{E_{NN}}{N} \right] \bar{Y}^T & \bar{X} \left[ I - \frac{E_{NN}}{N} \right] \bar{Z}^T \\ \bar{Y} \left[ I - \frac{E_{NN}}{N} \right] \bar{X}^T & \bar{Y} \left[ I - \frac{E_{NN}}{N} \right] \bar{Y}^T & \bar{Y} \left[ I - \frac{E_{NN}}{N} \right] \bar{Z}^T \\ \bar{Z} \left[ I - \frac{E_{NN}}{N} \right] \bar{X}^T & \bar{Z} \left[ I - \frac{E_{NN}}{N} \right] \bar{Y}^T & \bar{Z} \left[ I - \frac{E_{NN}}{N} \right] \bar{Z}^T \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix} \quad (6)$$

y por (4):

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} = \begin{bmatrix} D_{11/N} & D_{12/N} & D_{13/N} \\ D_{21/N} & D_{22/N} & D_{23/N} \\ D_{31/N} & D_{32/N} & D_{33/N} \end{bmatrix} \quad (7)$$

### 5. 2. 2. Definición de las correlaciones y vectores canónicos parciales muestrales

Definiremos tales correlaciones y vectores de forma análoga a como se hizo en el modelo poblacional, es decir, sustituyendo las particiones de la matriz de covarianza poblacional por análogas particiones obtenidas a partir de la matriz D.

En base al comentario anterior denominaremos vectores canónicos parciales muestrales a los vectores

$$\begin{aligned} u &= G^T e_x \\ v &= H^T e_y. \end{aligned} \quad (8)$$

en donde G y H son las soluciones de los sistemas

$$\begin{bmatrix} D_{12.3} & D_{22.3}^{-1} & D_{21.3} - c^2 & D_{11.3} \end{bmatrix} G = 0 \quad (9)$$

$$\begin{bmatrix} D_{21.3} & D_{11.3}^{-1} & D_{12.3} - c^2 & D_{22.3} \end{bmatrix} H = 0 \quad (10)$$

$$; \quad D_{ij.3} = D_{ij} - D_{i3} D_{33}^{-1} D_{3j} \quad ; \quad i, j = 1, 2.$$

y los vectores de variables  $e_x, e_y$ :

$$\begin{aligned} e_x &= x - \Sigma_{13} \Sigma_{33}^{-1} z. \\ e_y &= y - \Sigma_{23} \Sigma_{33}^{-1} z. \end{aligned} \quad (11)$$

Igualmente se denominan correlaciones canónicas parciales muestrales a las raíces cuadradas con signo positivo ordenadas en forma decreciente de las soluciones comunes a las ecuaciones determinantes:

$$\begin{aligned} | D_{12.3} \quad D_{22.3}^{-1} \quad D_{21.3} - c^2 D_{11.3} | &= 0 \\ | D_{21.3} \quad D_{11.3}^{-1} \quad D_{12.3} - c^2 D_{22.3} | &= 0 \end{aligned} \quad (12)$$

### 5. 2. 3. Estimadores de máxima verosimilitud de las correlaciones y vectores canónicos parciales.

La obtención de estimadores de máxima verosimilitud para las correlaciones y vectores canónicos parciales está basada en una serie de resultados que vamos a exponer a continuación:

- 1º. - La matriz  $NS = D$  constituye un estimador de máxima verosimilitud de la matriz  $N\Sigma$  <sup>(1)</sup>
- 2º. - Si se considera la matriz de covarianzas entre los vectores de variables  $(e_x, e_y)$

$$\Sigma_{.3} = \begin{bmatrix} \Sigma_{11.3} & \Sigma_{12.3} \\ \Sigma_{21.3} & \Sigma_{22.3} \end{bmatrix} \quad (13)$$

el estimador de máxima verosimilitud de  $N\Sigma$  es la matriz: <sup>(2)</sup>  $D_{.3} = \begin{bmatrix} D_{11.3} & D_{12.3} \\ D_{21.3} & D_{22.3} \end{bmatrix}$

(1) Véase nota (1) de la página 139.

(2) Véase Anderson, T.W., [2] pág. 181.

3º.- La transformación de las variables aleatorias contenidas en la matriz  $D$  a las variables contenidas en las matrices  $G$ ,  $H$  y  $C$  según las siguientes ecuaciones:

$$\begin{aligned} G^T D_{11.3} G &= I_p \\ H^T D_{21.3} G &= C^* \\ H^T D_{22.3} H &= I_q \end{aligned} ; C = \begin{bmatrix} C^* & 10 \\ \hline 0 & \dots & 0 \\ & & c_p & \vdots & 0 \end{bmatrix}_P$$

es una transformación donde coinciden el número de ecuaciones y de nuevas variables <sup>(1)</sup>

Los resultados contenidos en los puntos 2º y 3º permiten la aplicación a este caso concreto del teorema fundamental de 4.3. - IV. Por cumplirse las hipótesis de tal teorema  $G$ ,  $H$  y  $C$  son las matrices que contienen los estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos parciales.

#### 5. 2. 4. Distribuciones en el muestreo de las correlaciones y vectores de coeficientes canónicos.

##### 5. 2. 4. 1. Introducción

Utilizaremos en este apartado un conjunto de resultados que ya fueron presentados en 4.4.1. que nos servirán como base para la obtención de las distribuciones que aquí nos ocupan. En particular nos referiremos a los siguientes:

Si  $A^*$  y  $B^*$  son dos matrices de variables aleatorias con distribuciones independientes

$$\begin{aligned} A^* &\in W_p [\Sigma, m] \\ B^* &\in W_p [\Sigma, n] \end{aligned}$$

la distribución de las raíces características y vectores asocia-

(1) La afirmación anterior es cierta si en la matriz  $H_2 : H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}_P$  se exige a sus elementos las mismas restricciones que en el apartado 4.3. Cap. IV.

dos al sistema :

$$[A^* - \phi [A^* + B^*]] Y = 0$$

tienen las siguientes funciones de densidad:

$$f(F) = \frac{\pi^{k/2} \prod_{i=1}^p \Gamma[\frac{1}{2}(u+u+1-i)]}{\prod_{i=1}^p \Gamma[\frac{1}{2}(u+1-i)] \Gamma[\frac{1}{2}(u+1-i)] \Gamma[\frac{1}{2}(p+1-i)]} \prod_{i=1}^p t_i^{\frac{1}{2}(u-p-1)} \prod_{i=1}^p (1-t_i)^{\frac{1}{2}(u-p-1)} \prod_{i < j} (t_i - t_j) \quad (14)$$

en donde  $F = \begin{bmatrix} t_1 & & & 0 \\ & t_2 & & \\ & & \dots & \\ 0 & & & t_p \end{bmatrix}$  y :  $1 > t_1 > t_2 \dots > t_p > 0$

$$f(Y) = \frac{\pi \Gamma[(p+1-i)] \sum_{i=1}^p \frac{(u+u)}{2} \pi^{-\frac{1}{2} p^2} |Y^{-1}|^{-(u+u-p)} e^{-\frac{1}{2} tr \Sigma^{-1} Y^{-1} Y^{-1}}}{\int_{\mathcal{J}} \Gamma[(m+n-2)p] \Gamma[\frac{1}{2}(u+n+1-i)]} \quad (15)$$

La obtención de las distribuciones en el muestreo de las correlaciones y vectores canónicos parciales requiere demostrar que las matrices :

$$\begin{matrix} D_{11.3} - D_{12.3} D_{22.3}^{-1} D_{21.3} \\ D_{12.3} D_{22.3}^{-1} D_{21.3} \end{matrix}$$

poseen distribuciones Wishart independientes con parámetro común.  $\Sigma_{11}$ . La demostración de este resultado exige una serie de hipótesis bastante restrictivas y es para el caso en que se cumplan tales hipótesis para el que obtendremos las distribuciones ya citadas. Igualmente para llegar a tal resultado es necesario el estudio de las distribuciones en el muestreo de algunas matrices de variables aleatorias que van a ser expuestas en el apartado siguiente.

5.2.4.2. Distribución en el muestreo de ciertas matrices de variables aleatorias

Consideremos la matriz D obtenida con muestras con las

características señaladas en 5.2.1.

1º. - La matriz  $D \stackrel{(1)}{\in} W_{p+q+r} [\Sigma, n]$

2º. - La matriz  $\stackrel{(2)}{}$

$$\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} - \begin{bmatrix} D_{13} \\ D_{23} \end{bmatrix} D_{33}^{-1} \begin{bmatrix} D_{31} & D_{32} \end{bmatrix} \in W_{p+q} [\Sigma_{03}, n-r] \quad (16)$$

cuando  $D_{13}$   $D_{23}$   $D_{33}^{-1}$  se han obtenido a partir de una matriz  $Z$  de elementos no aleatorios.

3º. - Si  $D$  es una matriz obtenida a partir de una muestra de tamaño  $N$  procedente de una distribución normal

$$N_{p+q} \left[ \mu, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right]$$

Cuando a)  $N \geq p+q+1$

b) y es un vector no aleatorio

c)  $\Sigma_{12} = 0$

Las matrices  $D_{11,2}$  y  $D_{12} D_{22}^{-1} D_{21}$  siguen distribuciones Wishart independientes

$$\begin{aligned} D_{12} D_{22}^{-1} D_{21} &\in W_p [\Sigma_{11}, q] \\ D_{11,2} &\in W_p [\Sigma_{11}, n-q] \end{aligned} \quad (17)$$

Operando en (16) se obtiene

- 
- (1) La obtención de la distribución en el muestreo de la matriz  $D$  cuando ésta se ha construido con muestras de tamaño  $N$  de un colectivo  $N_{p+q+r}(\mu, \Sigma)$  puede verse en Anderson, T.W., [2] cap. 7.
- (2) Véase Khsirsagar, A. [28] pág. 109.

$$\begin{bmatrix} D_{11.3} & D_{12.3} \\ D_{21.3} & D_{22.3} \end{bmatrix} \in W [\Sigma_{.3}, n-r]$$

La aplicación del resultado del apartado 3 a esta matriz nos permite realizar la siguiente afirmación:

$$D_{11.3} - D_{12.3} D_{22.3}^{-1} D_{21.3} \in W_p [\Sigma_{11.3}, n-r-q] \quad (18)$$

$$D_{12.3} D_{22.3}^{-1} D_{21.3} \in W_p [\Sigma_{11.3}, q] \quad (19)$$

y además son independientes cuando se cumplen las siguientes hipótesis

- a)  $n-r \geq p+q+1$ .
- b)  $e_y$  es un vector no aleatorio
- c)  $\Sigma_{12.3} = 0$

#### 5.2.4.3. Distribución en el muestreo de las correlaciones canónicas parciales muestrales.

Las distribuciones de las matrices estudiadas en el apartado anterior nos va a permitir llegar a la distribución en el muestreo de las correlaciones canónicas parciales muestrales cuando éstas se han obtenido con muestras procedentes de vectores  $e_x$  y  $e_y$  independientes.

En efecto, considerando la ecuación determinante

$$| D_{12.3} D_{22.3}^{-1} D_{21.3} - c^2 D_{11.3} | = 0$$

y los resultados obtenidos en (18) y (19) es inmediato obtener la función de densidad de tal distribución sin más que realizar los cambios convenientes en los parámetros de la función (14). Esta resulta ser:

$$f(c_1^2 \dots c_p^2) = \frac{\pi^{1/2} \prod_{i=1}^p \Gamma[\frac{1}{2}(n+q-v+1-i)]}{\prod_{i=1}^p \Gamma[\frac{1}{2}(n-v+1-i)] \Gamma[\frac{1}{2}(q+1-i)] \Gamma[\frac{1}{2}(p+1-i)]} \prod_{i=1}^p c_i^{2k(q-p-i)} \prod_{i=1}^p (1-c_i^2)^{\frac{1}{2}(n-v-p-i)} \prod_{i=1}^p (c_i^2 - c_{i+1}^2) \quad (20)$$

#### 5.2.4.4. Distribución en el muestreo de los vectores de coeficientes canónicos parciales.

Por idéntico motivo que en el modelo canónico de dos vectores de variables aleatorias únicamente obtendremos aquí los vectores canónicos contenidos en la matriz  $G$  que satisface la ecuación

$$[D_{12.3} D_{22.3}^{-1} D_{21.3} - c^2 D_{11.3}] G = 0$$

De nuevo los resultados obtenidos en (18) y (19) permiten llegar a la función de densidad de tales variables a partir de la función (15) sin más que realizar la adaptación necesaria en sus parámetros. Procediendo de esta forma llegamos a la siguiente función

$$f(G) = \frac{\pi \Gamma[(p+1-c)]}{2^{(n-v-2)q} \Gamma[\frac{1}{2}(n-v+1-i)]} \sum_{i=1,3}^{\frac{-(n-v)}{3}} \pi^{-\frac{1}{2} p^2} |G^{-1}|^{-(n-v+p)} e^{-\frac{1}{2} tr \sum_{i=1,3} G^{-1} G^{-1}} \quad (21)$$

### 5.2.5. Pruebas de hipótesis en el modelo canónico parcial

El conjunto de resultados a que se ha llegado en el apartado 4.5. nos va a permitir fundamentar una serie de contrastes de hipótesis en torno a los parámetros poblacionales del modelo.

Un resultado básico en el que nos apoyaremos en todas las cuestiones que aquí vamos a tratar es el que hemos puesto de manifiesto en las expresiones (18) y (19). Los comentarios anteriores nos permiten la construcción de una serie de estadísticos que vamos a exponer a continuación y que utilizaremos en los contrastes de hipótesis acerca de las correlaciones canónicas.

#### 5.2.5.1. Contrastes de hipótesis relativos a las correlaciones canónicas parciales

##### 5.2.5.1.1. Contraste de la hipótesis de que todas las correlaciones canónicas parciales son nulas

Los resultados obtenidos (19) y (20) permiten fundamentar los tests de hipótesis que utilizaremos en el contraste que aquí nos ocupa. Puesto que las matrices  $D_{11.3}$  y  $D_{12.3}$   $D_{22.3}^{-1}$   $D_{21.3}$  poseen distribuciones Wishart independientes en base a ellas pueden construirse una serie de estadísticos cuyas distribuciones en el muestreo son los criterios expuestos en 4.5.2.1.

Ya que las matrices anteriores únicamente siguen las distribuciones citadas cuando se cumple la hipótesis del con-

traste (correlaciones canónicas parciales poblacionales nulas) es inmediato llegar a los tests estadísticos para realizar la prueba mencionada.

Entre la variedad de estadísticos que pueden formarse <sup>(1)</sup> nos referiremos aquí a los que siguen en el muestreo distribuciones de uso más frecuente como son:

$$\prod_{i=1}^p (1 - c_i^2) \in \Lambda(p, q, n-r) \quad (20)$$

$$\sum_{i=1}^p \frac{c_i^2}{1+c_i^2} \in \tau_0^2 \quad (21)$$

Cualquiera de estos dos estadísticos puede utilizarse para el contraste de la hipótesis del enunciado. Procederemos como sigue: Si evaluados a partir de una muestra (20) y (21) superan a los valores en las distribuciones respectivas a un determinado nivel de significación  $\alpha$  se rechazará a tal nivel la hipótesis de que las muestras proceden de vectores de variables aleatorias con correlaciones canónicas parciales poblacionales nulas.

El contraste puede realizarse también en base a las aproximaciones estudiadas del criterio  $\Lambda_i$  sin más que realizar la conveniente modificación en sus parámetros. Procediendo de esta forma se llega al siguiente resultado:

$$- \left[ (n-r) - \frac{p+q+1}{2} \right] \text{Ln.} \prod_{i=1}^p (1 - c_i^2) \rightarrow \chi^2(p \cdot q) \quad (22)$$

(1) También en base a las raíces de la ecuación  $|D_{11.3} D_{22.3} D_{33.3} - c^2 D_{11.3}| = 0$  pueden construirse estadísticos que sigan en el muestro los criterios de Pillai y Roy, y a partir de ellos realizar el contraste que aquí nos ocupa.

5. 2. 5. 1. 2. Contraste de la hipótesis de que parte de las correlaciones canónicas parciales son nulas

Cuando ha sido rechazada la hipótesis anterior puede ensayarse la siguiente: "las últimas  $p-s$  correlaciones canónicas parciales poblacionales son nulas".

Una simple corrección de los parámetros en la distribución del estadístico (98) permite llegar a la del que será utilizado en el contraste de esta hipótesis; tal es :

$$\sum_{i=s+1}^p (1-c_i^2) \in \mathcal{L}(p-r, q-r, n-r-s) \quad (23)$$

que puede aproximarse a la distribución  $\chi^2$  según la expresión

$$- \left[ n-r - \frac{1}{2}(p+q+1) \right] \ln \prod_{i=s+1}^p (1-c_i^2) \rightarrow \chi^2 \left[ (p-r)(q-r) \right] \quad (24)$$

De la misma forma podrían obtenerse para este caso concreto las dos aproximaciones debidas a Lawley <sup>(1)</sup>.

5. 3. El modelo muestral en el resto de modelos parciales

Nos referiremos aquí al modelo muestral de los modelos canónicos biparcial  $G_1$  y  $G_2$ . Un tratamiento paralelo al realizado en el modelo de Rao podría efectuarse previa prueba de que las matrices de variables aleatorias que intervienen en la definición de las correlaciones canónicas muestrales de este tipo de modelos siguen distribuciones de Wishart independientes.

---

(1) Véase las expresiones (99) y (100) de este capítulo.

Unicamente expondremos aquí una serie de tests de hipótesis que para sus respectivos modelos han elaborado Timm y Carlson y Sik -Yum -Lee. Todos ellos están referidos a las correlaciones canónicas entre vectores de variables resultantes de eliminar las influencias lineales supuestas.

1º Tests de hipótesis acerca de las correlaciones canónicas biparciales.

Para el contraste de la hipótesis de que las correlaciones canónicas  $X_1$  y  $X_2$  después de eliminados los efectos lineales supuestos en  $G$  -III son nulas, Timm y Carlson<sup>(1)</sup> consideran el siguiente estadístico<sup>(2)</sup>

$$\left[ (N - \max(n_3, n_4) - 1) - \frac{n_1 + n_2 + 1}{2} \right] \ln \prod_{i=1}^p (1 - c_i^2)^{(b.p.)} \quad (25)$$

mostrando que se aproxima a la distribución  $\chi^2_{(n_1, n_2)}$

2º Tests de hipótesis acerca de las correlaciones canónicas  $G_1$  y  $G_2$ .

Para el contraste de las hipótesis que aquí hemos mencionado Sik - Yum - Lee<sup>(2)</sup> propone los siguientes estadísticos:

$$\left[ N - n_3 - \max(n_4, n_3) - \frac{n_1 + n_2 + 1}{2} \right] \ln \prod_{i=1}^p (1 - c_i^2)^{(G_1)} \quad (26)$$

$$\left[ N - n_3 - \max(n_4, n_3) - \frac{n_1 + n_2 + 1}{2} \right] \ln \prod_{i=1}^p (1 - c_i^2)^{(G_2)} \quad (27)$$

(1) Timm y Carlson [48] pág. 165-196

(2) Véase la notación utilizada en 2. 2. 3., 2. 2. 4. y 2. 2. 5.

(3) Sik-Yum-Lee [47], pág. 417-421.

Cuya distribución en el muestreo se aproxima a  $\chi^2(n_1 \times n_2)$

#### 5. 4. Referencia al modelo muestral de los modelos canónicos generalizados

A falta de un desarrollo específico de los modelos muestrales del análisis canónico generalizado expondremos aquí un aspecto relacionado con esta cuestión tal como es el contraste de independencia entre un conjunto de más de dos vectores de variables aleatorias.

Wilks<sup>(1)</sup> ha obtenido el estadístico para el contraste de esta hipótesis mediante el procedimiento de razón de verosimilitudes. Expondremos una síntesis de tal procedimiento: Si  $(X_1, X_2, \dots, X_k)$  son vectores de variables aleatorias de dimensiones  $n_1, n_2, \dots, n_k$  con distribución conjunta normal de parámetros  $(m, \Sigma)$  la hipótesis de que  $(X_1, X_2, \dots, X_k)$  son independientes es equivalente a la de que las particiones de  $\Sigma$ ;  $\Sigma_{ij} = [0] \forall i \neq j$ . Si a partir de  $N$  observaciones de los vectores  $X_1, X_2, \dots, X_k$  se construye la función de verosimilitud de la muestra el cociente de los máximos de tal función cuando no se hace ninguna restricción y cuando se incluye en ella el cumplimiento de la hipótesis del contraste resulta ser: (2)

$$\lambda = \frac{|D|^{1/2 N}}{\prod_{i=1}^k |D_{ii}|^{1/2 N}} \quad (28)$$

siendo  $D_{ii}$  las particiones diagonales de la matriz  $D$  definida en 4. 1. (6).

Para llevar a cabo el contraste de la hipótesis que aquí nos ocupa es necesario conocer la distribución en el muestreo del es-

(1) Wilks, S.S. [52] pág. 309-326.

(2) Anderson, T.W. [2], pág. 230-233.

estadístico (28). Entre las aproximaciones dadas a tal distribución citaremos la siguiente:

$$\frac{n}{k} \ln \lambda^{2/n} \rightarrow \chi^2(t) \quad : N=n+1 \quad (29)$$

cuando la hipótesis del contraste es cierta<sup>(1)</sup>; siendo

$$\begin{aligned} \kappa^{-1} &= 1 - \frac{1}{2t} (2p_1 + 3p_2) \\ &: t = \frac{1}{2} p_2 \quad \text{y} \quad p_s = \left( \sum_{i=1}^k n_i^s \right) - \sum_{i=1}^k n_i^s \quad s=1,2. \end{aligned}$$

Si a partir de una muestra se evalúa el estadístico (28) la comparación con el valor en la distribución  $\chi^2(t)$  a un determinado nivel de significación  $p$  permite establecer la regla de decisión de la prueba.

Si  $\frac{n}{k} \ln \lambda^{2/n} > \chi^2_p(t)$  se rechazará la hipótesis de que las muestras proceden de vectores de variables aleatorias independientes

---

(1) Véase Morrison, D., [ 33 ] pág. 212.

C O N C L U S I O N E S . -

---

## CONCLUSIONES

Dentro del capítulo dedicado a las conclusiones de esta memoria distinguiremos dos tipos: Las que constituyen resultados de carácter general dentro del análisis canónico y las que contienen nuestras aportaciones al análisis canónico.

### A) CONCLUSIONES DE CARACTER GENERAL

PRIMERA: La mayor parte de los métodos del análisis multivariante pueden agruparse en dos grandes bloques: Análisis externo e interno. En las técnicas que constituyen el análisis externo se estudia la relación entre un vector de variables  $X$  (puede ser o no aleatorio) y otro vector externo  $Y$  (ídem), mientras que las incluidas en el análisis interno tienen por objeto analizar la relación y variabilidad entre los componentes de un único vector  $X$  (modelos de componentes principales y factorial).

SEGUNDA: El análisis canónico es la técnica por excelencia del análisis externo y constituye un marco genérico en el que quedan incluidos otros tipos de métodos multivariantes como son:

- Análisis de varianza y covarianza multivariantes. En este caso el vector externo  $Y$  está formado por componentes que son variables "dummy".

- Análisis discriminante multivariante. A este método se llega cuando se supone que el vector  $X$  es el que tiene componentes "dummy". En este caso los componentes del vector canónico  $V$  constituyen las funciones discriminantes.

- Tablas de contingencia. También pueden formularse como caso particular del análisis canónico en donde tanto el vector  $X$  como  $Y$  tienen componentes "dummy".

TERCERA: El análisis canónico introducido por Hotelling constituye la generalización natural del modelo de regresión múltiple extendido este al caso en que el regresando lo constituya un vector de variables aleatorias. Las primeras variables canónicas son las combinaciones lineales de los vectores de referencia  $X$ ,  $Y$  que tienen correlación máxima; las  $s$  segundas variables canónicas son las combinaciones lineales de los vectores de referencia incorreladas con las primeras y que tienen correlación máxima; las  $r$ -ésimas variables canónicas ( $r = \text{rango de } \Sigma_{12}$  en donde  $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ ) es la matriz de covarianzas conjunta de los vectores  $X$  e  $Y$ ), son las combinaciones lineales de los vectores de referencia incorreladas con las previamente obtenidas que presentan máxima correlación.

B) CONCLUSIONES QUE CONSTITUYEN NUESTRAS APORTACIONES AL ANALISIS CANONICO.

PRIMERA: Si  $u_1$  y  $v_1$  son las primeras variables canónicas las regresiones de  $u_1$  sobre  $Y$  y de  $v_1$  sobre  $X$  en donde  $X$  e  $Y$  son los vectores aleatorios de referencia, tienen unos coeficientes de regresión que son proporcionales a la primera pareja de vectores de coeficientes canónicos  $(m_1, l_1)$  siendo el factor de proporcionalidad en ambas regresiones la primera correlación canónica  $\lambda_1$ . Además el cuadrado de los errores de regresión en ambas regresiones es  $1 - \lambda_1^2$ . Análoga propiedad cumplen las segundas ...  $r$ -ésimas variables canónicas.

SEGUNDA: Los vectores canónicos son entes matemáticos que pueden no tener significación en términos de variables concretas. Los métodos de rotación permiten hallar vectores transformados de

los canónicos que conserven sus principales propiedades y cuya identificación, debido al contraste creado entre los elementos de los nuevos vectores de coeficientes, sea más fácil. Cuando se transforman ortogonalmente y con la misma matriz de transformación, los vectores canónicos se conserva la incorrelación intra-grupos así como la traza de la partición inter-grupos de la matriz de covarianzas conjunta. Fijadas estas características de la transformación ortogonal común para su cálculo no existe inconveniente en utilizar los procedimientos de rotación ortogonal del análisis factorial.

TERCERA: El modelo canónico parcial más general de los expuestos en la literatura estadística (Modelo  $G_2$ -biparcial de Sik-Yum-Lee) puede generalizarse bajo una estructura de dependencias a la que hemos denominado "extensión lineal del modelo de Rao". Según esta generalización el modelo canónico parcial L(3) es el modelo de Rao, L(5) es el modelo  $G_2$ -biparcial de Sik-Yum-Lee y L(7) modelo que hemos desarrollado en esta memoria incluye como casos particulares a los anteriores.

CUARTA: Igualmente el modelo  $G_2$ -biparcial puede generalizarse bajo otra estructura de dependencias a la que hemos denominado "extensión circular del modelo de Rao". En base a esta generalización el modelo C(3) coincide con el de Rao; C(5) coincide con el  $G_2$ -biparcial y C(7) modelo que hemos desarrollado en este trabajo incluye como casos particulares a los anteriores y al modelo L(7).

QUINTA: Las extensiones lineal y circular del modelo de Rao pueden sintetizarse en forma genérica en los modelos L(2n+1) y C(2n+1) que aquí han sido expuestos. Tales modelos coinciden para n= 1 y 2;

si  $n > 2$  los modelos canónicos que se originan tienen estructuras de dependencia asociadas en donde la del modelo  $L(2n+1)$  siempre está incluida en la del modelo  $C(2n+1)$ .

SEXTA : El análisis canónico de Hotelling no puede generalizarse a más de dos vectores aleatorios de referencia sin pérdida de alguna de las propiedades de los vectores canónicos como es la incorrelación entre componentes de distinto ordinal al considerarlos intergrupos.

SEPTIMA : Los modelos canónicos generalizados admiten una nueva extensión como es la consideración en tales modelos de vectores de variables aleatorias en los que se han eliminado las influencias lineales de otros. A estos modelos presentados en esta memoria los hemos denominado modelos canónicos parciales generalizados. Con la anotación ya introducida hemos expuesto los siguientes:  $L(3)-C(3)$  generalizado, parte y biparcial generalizado,  $G_1$  generalizado,  $L(5)-C(5)$  generalizado. También pueden extenderse las estructuras de los grafos de los modelos  $L(7) \dots L(2n+1)$  y  $C(7) \dots C(2n+1)$  sin más que considerar bloques de vectores aleatorios.

OCTAVA : Cuando las correlaciones y vectores de coeficientes canónicos parciales muestrales se obtienen como soluciones de los sistemas:

$$\begin{aligned} [D_{12.3} D_{22.3}^{-1} D_{31.3} - c^2 D_{11.3}] G &= 0 & : G^T D_{11.3} G &= I_p \\ [D_{21.3} D_{11.3}^{-1} D_{12.3} - c^2 D_{22.3}] H &= 0 & : H^T D_{22.3} H &= I_q \end{aligned}$$

las matrices de variables aleatorias  $C$ ,  $G$  y  $H$  constituyen estimadores de máxima verosimilitud de las correlaciones y vectores de coeficientes canónicos poblacionales  $\mathcal{L}$ ,  $L$  y  $M$ .

NOVENA: Las distribuciones en el muestreo de los estimadores de máxima verosimilitud de las matrices  $\Lambda$  y  $L$  del modelo canónico parcial de Rao pueden obtenerse en forma similar a las de las distribuciones de estos estimadores en el modelo canónico de Hotelling en base a que las matrices de variables aleatorias  $D_{11.3} - D_{12.3} D_{22.3}^{-1} D_{21.3}$  y  $D_{22.3} - D_{21.3} D_{11.3}^{-1} D_{12.3}$  que intervienen en la definición de las correlaciones y vectores canónicos parciales muestrales poseen distribuciones de Wishart independientes cuando tales matrices se han obtenido con muestras de vectores que previa eliminación de las influencias lineales de otro son independientes y con distribución conjunta normal. El resultado anterior nos ha permitido obtener las dos conclusiones siguientes:

DECIMA: La función de densidad de la distribución en el muestreo de las correlaciones canónicas parciales muestrales cuando éstas se obtienen en vectores con las características señaladas en la conclusión anterior es la misma que función de densidad de la distribución en el muestreo de tales correlaciones en el modelo de Hotelling. Únicamente hay que sustituir en esta última el parámetro  $N$  por  $N-r$  en donde  $N$  es el tamaño de la muestra y  $r$  el número de componentes del vector aleatorio cuya influencia lineal se elimina.

UNDECIMA: La función de densidad de la distribución en el muestreo de la matriz de vectores de coeficientes canónicos parciales  $G$  también puede obtenerse a partir de la función de densidad de las variables contenidas en idéntica matriz en el modelo de Hotelling sin más que sustituir en esta última el parámetro  $N$  por  $N-r$  y la matriz  $\Sigma_{11}$  por  $\Sigma_{11.3}$

BIBLIOGRAFIA

CONSULTADA. -

- [1] AAKER, D. A. (1971). *Multivariate Analysis in Marketing Theory and Application*. Wadsworth Publishing Company Inc. California.
- [2] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons Inc. New York.
- [3] AZNAR, A. (1975). *El análisis factorial en la Econometría. Una aplicación a las provincias españolas*. Tesis Doctoral.
- [4] BARTTLET, M. S. (1939). A note on test of significance in multivariate analysis. *Proc. Camb. Phil. Soc.* 35 pg. 180.
- [5] BARTTLET, M. S. (1948). A note on the statistical estimation of supply and demand relations from time series. *Econométriķa*. 16. pág. 323-329.
- [6] BERTIER, P. et BOUROCHE, J. M. (1975). *Analyse des données multidimensionnelles*, P. U. F. Paris.
- [7] CLIFF, N. and KRUSS, D. (1976). Interpretation of canonical analysis: Rotated versus unrotated solutions. *Psychometrika*. 41 pag. 35-42.
- [8] COOLEY, W. and LOHNES, P. (1971) *Multivariate Data Analysis*. John Wiley and Sons, Inc. New York.

- [9] CRAMER, H. , (1970). Métodos Matemáticos de Estadística. Aguilar. Madrid.
- [10] DEEMER, W. L. and OLKIN, I. (1951) The jacobians of certain matrix transformations useful in multivariate analysis. *Biometrika* 38, pág. 345-367.
- [11] DRHYMES, P. J. (1970). *Econometrics statistical foundations and applications*. Harper Row.
- [12] Fz. TROCONIZ, A. (1975). *Probabilidades, estadística, procesos aleatorios*. Grafor. Bilbao.
- [13] Fz. TROCONIZ, A. (1978). *Matemáticas Generales I*. Grafor. Bilbao.
- [14] GIRSHICK, M. A. (1939). On the sampling theory of roots of determinantal equations. *Annals of Mathematical Statistics*, 10, pág. 203-224.
- [15] GOOD, I. J. (1969). Some applications of the singular decomposition of a matrix. *Technometrics*. 11. pag. 823-831.
- [16] GREEN, P. E. (1978). *Mathematical Tools for applied Multivariate analysis*. Academic Press.
- [17] HARMAN, H. (1976). *Modern Factor Analysis*. University of Chicago Press. Chicago.

- [18] HECKS, D. L., (1960). Charts of some upper percentage points of the distribution of the largest characteristic root. *The Annals of Mathematical Statistics*, 31. pag. 625-642.
- [19] HORST, P. (1961) Relations among  $m$  sets of measures. *Psychometrika*. 26. pag. 129-149.
- [20] HORST, P. (1961) Generalized canonical correlations and their applications to experimental data. *Journal of clinical Psychology* (monograph supplement.) 14- pag. 331-347.
- [21] HORST, P. (1965) *Factor Analysis of Data matrices*. Holt Rinehart and Winston Inc. New York.
- [22] HOTELLING, H. (1935) The most predictable criterion. *Journal of Educational Psychology*, 26. pag. 139-142.
- [23] HOTELLING, H. (1936) Relation between two sets of variates. *Biometrika*, 28. pag. 321-377.
- [24] JÖRESKOG, K. G., KLOVAN, J.E. and REYMENT R. A. (1976) *Geological Factor Analysis*. Elsevier Scientific Publishing Company. Amsterdam.

- [25] JOSHINSTON, J. (1975) *Métodos de Econometría*. Vicens Vives Barcelona.
- [26] KENDALL, M. G. and STUART, A. (1968) *The advanced theory of Statistics*. Vol-3. Charles Griffin and Company London.
- [27] KHATRI, G. C. (1976) A note on Multiple and Canonical correlation for a singular covariance matrix. *Psychometrika*, 41. pág. 465-470.
- [28] KSHIRSAGAR, A. M. (1972) *Multivariate Analysis*. Marcel Dekker, Inc. New York.
- [29] LAWLEY, D. N. (1956) Tests of significance for the latent roots of covariance and correlations matrices. *Biometrika*, 43. pág. 128-136.
- [30] LAWLEY, D. N. (1959) Tests of significance in canonical analysis. *Biometrika* 46. pág. 59-66.
- [31] LAWLEY, D. N. and MAXWELL, A. E. (1971) *Factor Analysis as Statistical method*. Butterworths and Co. London.
- [32] LEBART, L. et FENELON, J. P. (1971) *Statistique et Informatique appliquée*. Dunod. Paris.

- [33] MORRISON, D. (1967) *Multivariate Statistical Methods*. McGraw-Hill, New York.
- [34] NANDA, D. N., (1947) Distribution of a root of a determinantal equation. *The Annals of Mathematical Statistics*. 18, pag. 47-48.
- [35] NANDA, D. N., (1948) Limiting Distribution of a root of a determinantal equation. *The Annals of Mathematical Statistics*, 19, pag. 340-350.
- [36] OLKIN, I, and RUBIN, H. (1964) Multivariate Beta distributions and independence properties of Wishart distribution. *The Annals of Mathematical Statistics*, 35 pag. 261-267.
- [37] PILLAI, K. C. (1953) On the distribution of the sum of the roots of a determinantal equations. *The Annals of Mathematical Statistics*. 24 pag. 495
- [38] PILLAI, K. C. (1955) Some new test criteria in multivariate analysis. *The Annals of Mathematical Statistics*. 26 pag. 117-121.
- [39] PILLAI, K. C. and GUPTA, A. (1969) On the exact distribution of Wilks criterion. *Biometrika* 56. pag. 109-118.

- [40] RAO, C. R. (1951) An asymptotic expansion of the distribution of Wilk's criterion. Bull Inst. Inter. Stats. 33 pág. 177.
- [41] RAO, C.R. and MITRA, S. K. (1971) Generalized inverse of matrices and its applications. John Wiley and Sons. Inc. New York.
- [42] RAO, R. B. , (1969) Partial canonical correlations. Trabajos de Estadística e Investigación Operativa. 20. pag. 211-219.
- [43] ROY, S. N. (1958) Test of multiple independence and the associated confidence bound. The Annals of Mathematical Statistics, 29, pág. 491-502.
- [44] ROY, S. N. (1957) Some aspects of Multivariate Analysis. John Wiley and sons. New York.
- [45] SCHATZOFF, M. (1966) Exact distributions of Wilks like likelihood ratio criterion. Biometrika. 53. pag. 347-358.
- [46] SEAL, H. (1966) Multivariate Statistical Analysis for Biologist. Methuen. London.
- [47] SIK-YUM-LEE (1978) Generalizations of the Partial, Part and Bipartial canonical correlation analysis. Psychometrika. 43. pág. 427-430.

- [48] TIMM, N. H. and CARLSON, J. E. (1977) Part and Bipartial canonical correlation analysis. *Psychometrika*. 41. pág. 149-1956.
- [49] TINTNER, G., (1963) *Econometrics*. John Wiley and Sons. Inc. New York.
- [50] VAN DE GEER, J. P. (1971) *Introduction to Multivariate Analysis for the Social Sciencies*. W. H. Freeman and Company. San Francisco.
- [51] WAUGH, F. V. (1942) Regresions between sets of variables. *Econometrika*. 10. pág. 290-310.
- [52] WILKS, S. S. (1935) On the independence of  $k$  sets of normally distributed statistical variables. *Econometrika*. 3, . pág. 309-326.
- [53] YONG-SIN-LEE (1971) Distribution of the canonical correlations and asyntotic expansions for distributions of certain independence test statistics. *The Annals of Mathematical Statistics*, 42. pág. 526-537.

