

**Universidad de Sevilla**

**Facultad de Matemáticas**

**Departamento de Estadística e Investigación Operativa**

**Influencia en el modelo de curvas de crecimiento  
a través de distancias entre distribuciones**

Doctorado en Estadística e Investigación Operativa

Carlos Ferreira do Carmo de Sousa



SEVILLA - 2011



# Universidad de Sevilla

Departamento de Estadística e Investigación Operativa

## Influencia en el modelo de curvas de crecimiento a través de distancias entre distribuciones

Memoria realizada por Carlos Ferreira do Carmo de Sousa para optar al Grado de Doctor por la Universidad de Sevilla, dirigida por los Doctores D. Joaquín García de las Heras y D. Juan Muñoz Pichardo

Sevilla, noviembre de dos mil once

*Firma del doctorando,*

---

Fdo.: Carlos Ferreira do Carmo de Sousa

*Vº Bº de los Directores,*

---

Fdo.: Dr. D. Joaquín García de las Heras

---

Fdo.: Dr. D. Juan Muñoz Pichardo



# Agradecimientos

À minha família e amigos por todo o apoio moral, em particular, ao meu pai pelas muitas "viagens".

À Cláudia pela infinita paciência e dedicação. Ao Afonso dedico este trabalho.

Aos meus orientadores, Professor Doctor Joaquín Garcia de las Heras e Professor Doctor Juan Muñoz Pichardo, por todo o apoio pessoal, profissional e científico durante estes anos.

Ao Mestre Ivo Martins pela sua preciosa ajuda e disponibilidade.

Finalmente, contudo não menos importante, a todos os colegas do Instituto Superior de Engenharia da Universidade do Algarve.



# Índice general

<b>Introducción</b>	<b>III</b>
<b>1. Modelo de curvas de crecimiento</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Definición del modelo . . . . .	3
1.3. Análisis del modelo . . . . .	7
1.4. Análisis del modelo con estructura simple de Rao . . . . .	15
<b>2. Medidas de influencia propuestas en la literatura</b>	<b>19</b>
2.1. Introducción . . . . .	19
2.2. Residuos en el modelo de curvas de crecimiento . . . . .	22
2.3. Modelo de curvas de crecimiento perturbado bajo la omisión . . . . .	24
2.4. Análisis de influencia en el modelo de curvas de crecimiento . . . . .	27
2.5. Análisis de influencia en el modelo de curvas de crecimiento con estructura simple de Rao . . . . .	32
<b>3. Influencia a través de la distancia de Fréchet</b>	<b>39</b>
3.1. Introducción . . . . .	39
3.2. Análisis de influencia a través de la distancia entre distribuciones . . . . .	40
3.3. Distancia de Fréchet y su utilización en el análisis de influencia . . . . .	41
3.4. La medida de Fréchet en el modelo de curvas de crecimiento . . . . .	43
3.4.1. Influencia en combinaciones lineales de $\hat{\mathbf{B}}$ . . . . .	55
3.4.1.1. Influencia en las filas de $\hat{\mathbf{B}}$ . . . . .	60
3.4.1.2. Influencia en las columnas de $\hat{\mathbf{B}}$ . . . . .	62
3.4.1.3. Influencia en los elementos de $\hat{\mathbf{B}}$ . . . . .	64
3.5. La medida de Fréchet en el modelo de curvas de crecimiento con estructura simple de Rao . . . . .	68
3.5.1. Influencia en combinaciones lineales de $\tilde{\mathbf{B}}$ . . . . .	71
3.5.1.1. Influencia en las filas de $\tilde{\mathbf{B}}$ . . . . .	74
3.5.1.2. Influencia en las columnas de $\tilde{\mathbf{B}}$ . . . . .	75
3.5.1.3. Influencia en los elementos de $\tilde{\mathbf{B}}$ . . . . .	76
<b>4. Influencia a través de la distancia de Rao</b>	<b>79</b>
4.1. Introducción . . . . .	79
4.2. Distancia de Rao y su utilización en el análisis de influencia . . . . .	79
4.3. La medida de Rao en el modelo de curvas de crecimiento . . . . .	82
4.3.1. Influencia en combinaciones lineales de $\hat{\mathbf{B}}$ . . . . .	89

4.3.1.1.	Influencia en las filas de $\widehat{\mathbf{B}}$ . . . . .	92
4.3.1.2.	Influencia en las columnas de $\widehat{\mathbf{B}}$ . . . . .	93
4.3.1.3.	Influencia en los elementos de $\widehat{\mathbf{B}}$ . . . . .	94
4.4.	La medida de Rao en el modelo de curvas de crecimiento con estructura simple de Rao . . . . .	96
4.4.1.	Influencia en combinaciones lineales de $\widetilde{\mathbf{B}}$ . . . . .	101
4.4.1.1.	Influencia en las filas de $\widetilde{\mathbf{B}}$ . . . . .	103
4.4.1.2.	Influencia en las columnas de $\widetilde{\mathbf{B}}$ . . . . .	104
4.4.1.3.	Influencia en los elementos de $\widetilde{\mathbf{B}}$ . . . . .	104
<b>5.</b>	<b>Implementación y casos prácticos</b> . . . . .	<b>107</b>
5.1.	Introducción . . . . .	107
5.2.	Datos de Glucosa . . . . .	108
5.3.	Datos dentarios . . . . .	118
	<b>Conclusiones y futuras líneas de investigación</b> . . . . .	<b>127</b>
	<b>Bibliografía</b> . . . . .	<b>131</b>
	<b>Anexos</b> . . . . .	<b>143</b>
	<b>A. Notación</b> . . . . .	<b>145</b>
	<b>B. Algebra matricial</b> . . . . .	<b>147</b>
	<b>C. Distribuciones multivariantes</b> . . . . .	<b>151</b>

---



# Introducción

El interés del investigador, en algunas experimentaciones, se centra en analizar datos a través del tiempo para conocer la tendencia de un individuo o grupos de individuos. En este contexto se enmarca el análisis de curvas de crecimiento. El Modelo de Curvas de Crecimiento (MCC), introducido formalmente por Potthoff y Roy (1964), es un modelo multivariante generalizado, extensión del análisis de varianza multivariante, aplicable a un gran conjunto de problemas reales en muy diversas áreas científicas, especialmente útil para los estudios de experimentos con datos longitudinales y medidas repetidas. Desde la formulación del MCC, diferentes aspectos del modelo han sido sucesivamente considerados por muchos autores en diferentes trabajos (véase, por ejemplo, Kollo y von Rosen, 2005).

Por otra parte, cualquier análisis estadístico tiene como objetivo básico la obtención de conclusiones fiables a partir de los datos de las variables analizadas. Así, el papel que desempeñan las observaciones es de gran importancia para el desarrollo del estudio. No obstante, la importancia de cada observación en la construcción de un modelo es generalmente muy distinta. Un gran número de autores han presentado situaciones prácticas en las que existen observaciones experimentales que inciden considerablemente en los resultados del análisis, motivando la necesidad de identificar tales observaciones, denominadas en la literatura como observaciones influyentes u observaciones influencia, y evaluar sus efectos en el análisis estadístico que se pretende realizar. De acuerdo con la definición de Cook (Cook, 1977), una observación es considerada como influyente si su omisión de los datos da lugar a cambios sustanciales en rasgos importantes del análisis. Por otro lado, siguiendo a Cook y Weisberg (1982), puede decirse que el análisis de los datos experimentales, con objeto de encontrar estos casos relevantes, es de gran interés para las conclusiones que se obtengan de la experiencia por dos motivos:

- i) Proporciona información referida a la fiabilidad de las conclusiones y resultados obtenidos.
- ii) Puede indicar áreas del espacio muestral con efecto informativo inadecuado para una inferencia fiable y estable.

En este sentido, la identificación de estas observaciones es esencial para juzgar la robustez del modelo construido y para evitar errores de especificación en los modelos. El objetivo de esta memoria se centra en estos tópicos arriba citados: El Análisis de Influencia en Modelos de Curvas de Crecimiento. En el marco de la inferencia estadística clásica, el estudio de la influencia sobre un modelo se aborda como la evaluación del cambio resultante de la perturbación del modelo. Las distintas técnicas estadísticas propuestas para abordar el problema de la influencia se enmarcan en la aproximación genérica recogida por Cook (1977) y Cook y Weisberg (1982): para medir el efecto que tiene una observación o un conjunto de ellas sobre un aspecto de interés

del análisis estadístico, se introducen pequeñas perturbaciones (*perturbación del modelo*) y se cuantifica el cambio producido (*comparación de resultados*). En este enfoque se han propuesto un conjunto de métodos englobados en lo que genéricamente se conoce como el análisis de influencia, incluida en las técnicas de diagnóstico. Usualmente, este estudio puede ser hecho dentro de las siguientes estructuras. Estructura de verosimilitud, donde están incluidas la distancia de Cook, la distancia de Cook-Weisberg o el desplazamiento de verosimilitud. Estructura bayesiana, donde están incluidas la divergencia de Kullback-Leibler o la entropía bayesiana. Dentro de estas dos estructuras las técnicas de diagnóstico multivariantes pueden ser clasificadas, a su vez, en dos categorías: la influencia global y la influencia local.

A partir de estas dos categorías de técnicas de diagnóstico, se han propuesto medidas de influencia en prácticamente todos los modelos estadísticos. Una revisión general sobre las medidas de diagnóstico aplicadas al MCC se recoge en Pan y Fang (2002). Sin embargo, la mayoría de las técnicas de detección de observaciones influyentes están orientadas hacia la cuantificación de la influencia sobre el valor del estimador obtenido en una muestra concreta, es decir, al estudio de la influencia sobre las estimaciones de los parámetros de interés del modelo, comparando directamente los valores obtenidos en dichas estimaciones a partir del modelo postulado y del modelo perturbado. Así, tales diagnósticos de influencia tienden a evaluar algunas características aisladas de los cambios provocados por la perturbación.

Un enfoque alternativo es diseñar técnicas que permitan analizar la influencia sobre la distribución muestral de dichos estimadores. Esto es, no sólo sobre el valor del estimador obtenido en una muestra concreta, sino en toda la distribución de valores que puede tomar el estimador bajo estudio. El argumento que induce este enfoque es que, en múltiples ocasiones, la influencia afecta al comportamiento del estimador, en particular a su estructura de covarianzas y, en consecuencia, a los errores de estimación. Ello puede proporcionar medidas de influencia más globales, más completas y, por tanto, menos sensibles a posibles fluctuaciones muestrales. En este sentido, Jiménez-Gamero et al. (2002) proponen una aproximación general al problema basada en medir la distancia entre la distribución del estadístico de interés bajo ambos modelos, el modelo postulado y el modelo perturbado. Este enfoque genérico permite obtener medidas de influencia más globales en tanto que no se centra en el efecto sobre el valor de un estadístico o estimador obtenido para una muestra concreta, sino sobre la distribución de valores posibles que el estadístico o estimador puede tomar, bajo la ley de probabilidad estimada a partir de la muestra. Para una aplicación de dicha metodología a otros modelos estadísticos véase Muñoz-Pichardo (2006). Es en esta aproximación al análisis de influencia en la que se centra esta memoria.

Así, el objetivo principal de esta memoria es obtener medidas de influencia basadas en la distancia entre distribuciones, para el análisis de influencia sobre el estimador de máxima verosimilitud (EMV) de la matriz de coeficientes de regresión del MCC. Para ello se plantean los siguientes aspectos:

- Se considera como hipótesis distribucional la normalidad multivariante.
  - Se utiliza el esquema de perturbación de la omisión de las observaciones bajo estudio sobre el EMV del parámetro estructural del modelo, cuantificando la diferencia entre los resultados obtenidos para las distribuciones del modelo postulado inicialmente y del modelo perturbado.
-

- 
- Se proponen como distancias entre distribuciones: la distancia de Fréchet (Fréchet, 1957) y la distancia de Rao (Rao, 1945 y 1949).

De esta forma se obtienen diagnósticos de influencia que permiten analizar la fiabilidad de las conclusiones y resultados obtenidos del análisis estadístico y pueden indicar áreas del espacio muestral con efecto informativo inadecuado para una inferencia fiable. Tales diagnósticos complementan los propuestos en la literatura y pueden proporcionar información no incluida en la aportada por aquellos. En este sentido, los capítulos de esta memoria están estructurados de la siguiente manera.

Con la finalidad de establecer la notación básica a utilizar en este trabajo y presentar una serie de resultados utilizados en los siguientes capítulos, en el Capítulo 1, se lleva a cabo una revisión del MCC, donde se introducen algunos conceptos y resultados básicos. Se recoge, de una forma resumida, el MCC, se presentan los estimadores de mínimos cuadrados generalizados, los EMV y sus propiedades más generales. Además, se recoge el MCC con estructura simple de Rao (MCC(SCS)), uno de los modelos más utilizados en el marco de las curvas de crecimiento.

Con objeto de introducir adecuadamente los conceptos necesarios para el desarrollo de las técnicas de análisis de influencia recogidas en esta memoria y realizar una síntesis del estado actual del problema considerado, en el Capítulo 2, se comienza con una introducción a los residuos del MCC. Posteriormente, se presenta, también de forma reducida, el MCC perturbado bajo la omisión, para culminar con una revisión de algunas de las medidas de influencia global, basadas en la omisión de casos, propuestas en la literatura para identificar observaciones influyentes en el EMV de la matriz de coeficientes de regresión del MCC y del MCC(SCS).

En el Capítulo 3, se desarrolla una medida de influencia, basada en la distancia de Fréchet entre distribuciones de probabilidad, para el EMV de la matriz de coeficientes de regresión del MCC y del MCC(SCS), así como para algunas combinaciones lineales del mismo. Las medidas basadas en la distancia de Fréchet pueden ser descompuestas en dos componentes de interés: el efecto sobre la localización y el efecto sobre la dispersión.

En el Capítulo 4, se propone una medida de influencia, basada en la distancia de Rao entre distribuciones de probabilidad, para el EMV de la matriz de coeficientes de regresión del MCC y del MCC(SCS), siguiendo una estructura paralela al capítulo anterior.

Finalmente, en el Capítulo 5, para ilustrar la utilidad de los diagnósticos propuestos, se aplican las medidas de influencia presentadas en el Capítulo 2 y los planteamientos metodológicos de detección de observaciones influyentes, bajo la utilización de distancias entre distribuciones de probabilidad, desarrolladas en los Capítulos 3 y 4, a dos conjuntos de datos reales, a través de los cuales se comprueba empíricamente la adecuación de las medidas de Rao y Fréchet, comparándolas con los diagnósticos de influencia ya existentes. El análisis de influencia llevado a cabo en los dos ejemplos, se realiza desde el punto de vista individual, mediante un programa realizado en MATLAB.

Por último, se presentan algunas conclusiones, comentarios y futuras líneas de investigación.

Se completa esta memoria con tres anexos (A, B y C) en los cuales se recogen, respectivamente, alguna notación usada, una selección de definiciones y resultados correspondientes, fundamentalmente, al álgebra matricial y algunos resultados básicos de distribuciones matriciales útiles en el desarrollo de esta memoria.

---



# Capítulo 1

## Modelo de curvas de crecimiento

### 1.1. Introducción

El modelo de curvas de crecimiento (MCC) es un modelo multivariante generalizado, especialmente útil para el estudio de datos longitudinales, de medidas repetidas y el análisis de tendencias. Básicamente, se puede decir que las experiencias sobre las que se aplica esta modelización consisten en registrar una serie de medidas en sucesivos instantes de tiempo, en una o más subpoblaciones, a partir de las cuales se pretende analizar la evolución o tendencia de los individuos y las posibles diferencias en esta tendencia entre las subpoblaciones consideradas.

Tal y como viene recogido en Rao (1972), se puede considerar que el primer trabajo sobre curvas de crecimiento fue publicado por Wishart (1938). Posteriormente, Box (1950) realiza un estudio pormenorizado sobre diversos problemas en este tipo de análisis, modificando los procedimientos de Wishart y proponiendo contrastes para su estudio. En Rao (1958) se estudian algunos problemas de curvas de crecimiento, así como algunas generalizaciones de las aproximaciones dadas por Wishart, y en Rao (1959) se hace una exposición teórica general cuando se trabaja con variables normales correlacionadas aplicadas a problemas de curvas de crecimiento. Otras extensiones del artículo de Wishart fueron dadas por Leech y Healy (1959), Rao (1961) y Elston y Grizzle (1962), entre otros.

El MCC fue introducido formalmente por Potthoff y Roy (1964) y, desde su formulación, diferentes aspectos del modelo han sido sucesivamente considerados por muchos autores en diferentes trabajos. A modo de ilustración se citan algunos:

- Trabajos que abordan el estudio de diferentes contrastes de hipótesis para este modelo: Khatri (1966), Gleser y Olkin (1970), Fujikoshi (1974), Kariya (1978), Kanda (1994) y Hamid y von Rosen (2005a, 2005b).
- Trabajos que abordan diversos problemas de predicción: Lee y Geisser (1972, 1975), Reinsel (1984a, 1984b), Rao (1987), Lee (1988), Liski y Nummi (1990), Naik (1990) y Kanda (1992).
- Trabajos sobre regiones de confianza: Potthoff y Roy (1964), Rao (1965), Khatri (1966), Rao (1967), Grizzle y Allen (1969), Gleser y Olkin (1972), Fujikoshi y Nishii (1983), Srivastava y Carter (1983) y Kanda (1990a, 1994).
- Trabajos que estudian el MCC con valores perdidos: Kleinbaum (1973), Liski (1985), Srivastava (1985), von Rosen (1989) y Kanda (1994).

- Trabajos sobre estimación y momentos distribucionales: Potthoff y Roy (1964), Rao (1965, 1966), Khatri (1966), Grizzle y Allen (1969), von Rosen (1988, 1990, 1991a) y Kollo y von Rosen (2005).
- Trabajos sobre las distribuciones de los estimadores de máxima verosimilitud de los parámetros del modelo: Gleser y Olkin (1970), Kabe (1975), Fujikoshi (1985), Kenward (1986), Fujikoshi (1987), Fujikoshi y Shimizu (1989), Kollo y von Rosen (2005) y Kollo et al. (2007).
- Trabajos que abordan el análisis de residuos: von Rosen (1994), von Rosen (1995b), Hamid (2005) y Hamid y von Rosen (2006).
- Trabajos sobre técnicas de diagnosis y validación del modelo: Liski (1991), Pan y Fang (1995, 1996), Pan et al. (1997), Liu (1998), Pan et al. (1999) y Pan y Fang (2002).
- Revisiones generales de este modelo se recogen, entre otros, en los siguientes trabajos: Woolson y Leeper (1980), von Rosen (1991b) (que recopila una extensa bibliografía sobre el mismo), Nummi (1995) y Srivastava y von Rosen (1999). Un libro completo y básico sobre el MCC es Kshirsagar y Smith (1995) que recoge los resultados más relevantes sobre estimación de parámetros, contrastes de hipótesis y predicciones para valores futuros. Contribuciones más recientes sobre el modelo se recogen en Kollo y von Rosen (2005), Ohlson (2009), Srivastava et al. (2009), Ye y Wang (2009), Ohlson y von Rosen (2010), Ohlson y Srivastava (2010) y Hamid et al. (2011).

El MCC se aplica extensamente en muchas áreas de la investigación: medicina, epidemiología, biología, psicología, economía, agronomía e ingeniería, entre otras. Por ejemplo, muchos ensayos clínicos para testar nuevos medicamentos y algunos estudios de mercado se caracterizan por la repetición de la variable objetivo en diferentes unidades experimentales a lo largo del tiempo. Algunos autores (Kabe, 1987 y Tan, 1991) consideran el MCC como una generalización natural del análisis de la varianza multivariante (MANOVA). En particular, Chinchilli y Elswick (1985) llaman al MCC MANOVA generalizado o GMANOVA.

En esta memoria no se pretende realizar un estudio exhaustivo sobre las técnicas de inferencia en el MCC. El objetivo de este capítulo es presentar las herramientas necesarias para, posteriormente, poder analizar la calidad de los datos en el MCC a través del Análisis de Influencia. Para ello, se lleva a cabo una revisión del modelo, con la finalidad de establecer la notación básica y presentar una serie de resultados necesarios en los siguientes capítulos.

En la sección 1.2, se comienza con la derivación de estos modelos y se especifica la importancia de las distintas formas de la matriz de covarianzas, prestando especial atención a las dos más utilizadas: matriz de covarianzas no estructurada y matriz de covarianzas con estructura simple de Rao.

En la sección 1.3, se aborda el problema de la estimación de los modelos de curvas de crecimiento. En concreto, los procedimientos de mínimos cuadrados generalizados y de máxima verosimilitud, así como sus propiedades más generales necesarias para el desarrollo de las medidas de influencia.

La sección 1.4 se dedica a la estimación de los modelos de curvas de crecimiento con estructura simple de Rao.

---

## 1.2. Definición del modelo

Como se ha comentado anteriormente, el MCC es una técnica estadística usada frecuentemente para analizar datos longitudinales o medidas repetidas. Tales datos aparecen, por ejemplo, cuando se mide una determinada característica para cada unidad experimental en un conjunto de instantes de tiempo fijos. Así, se puede considerar la siguiente situación experimental:

- El estudio se realiza sobre una población de individuos o unidades experimentales clasificadas en un conjunto de grupos o subpoblaciones.
- Se obtiene una medida o variable en diferentes instantes de tiempo de cada unidad experimental.

Formalizando esta situación experimental, se dispone de  $r$  grupos  $G_1, \dots, G_j, \dots, G_r$ , se selecciona una muestra de cada uno de ellos y a cada uno de los individuos o unidades experimentales seleccionadas se les mide una variable de crecimiento  $Y$  en  $p$  instantes de tiempo  $t_1, \dots, t_m, \dots, t_p$ . Sea  $\underline{Y} = [Y_1 \ \dots \ Y_m \ \dots \ Y_p]^T$ , donde la variable aleatoria  $Y_m$  representa el crecimiento en el instante  $t_m$  ( $m = 1, \dots, p$ ).

En la modelización, se considera una regresión polinomial de grado  $(q - 1)$  para el vector  $\underline{Y}$  como función del tiempo  $t$ . Por lo que

$$E[\underline{Y} | G_j] = \underline{\mu}_j = \begin{bmatrix} \mu_{j1} \\ \dots \\ \mu_{jm} \\ \dots \\ \mu_{jp} \end{bmatrix} = \begin{bmatrix} \sum_{h=0}^{q-1} \beta_{hj} t_1^h \\ \dots \\ \sum_{h=0}^{q-1} \beta_{hj} t_m^h \\ \dots \\ \sum_{h=0}^{q-1} \beta_{hj} t_p^h \end{bmatrix} = \begin{bmatrix} P_j(t_1) \\ \dots \\ P_j(t_m) \\ \dots \\ P_j(t_p) \end{bmatrix},$$

siendo

$$P_j(t_m) = \sum_{h=0}^{q-1} \beta_{hj} t_m^h, \quad m = 1, 2, \dots, p, \quad p > q - 1, \quad j = 1, 2, \dots, r,$$

es decir, el valor medio de la variable crecimiento en el instante  $m$ -ésimo para el grupo  $j$ -ésimo viene dado por

$$\mu_{jm} = P_j(t_m).$$

Así, la variación en el tiempo del vector aleatorio  $\underline{Y}$  se modeliza a través de un polinomio de grado  $(q - 1)$  del tiempo. El polinomio es el mismo en cada instante, pero distinto en cada grupo, aunque se supone que se verifica la hipótesis de homocedasticidad

$$\text{Var}[\underline{Y} | G_j] = \underline{\Sigma}, \quad \forall j = 1, \dots, r,$$

es decir, las observaciones  $Y_1, \dots, Y_p$  para un mismo individuo están relacionadas a través de la matriz de varianzas-covarianzas  $\underline{\Sigma}$ , común en todos los  $r$  grupos.

Para el estudio del modelo, se selecciona, en cada grupo, una muestra aleatoria de tamaño  $n_1, \dots, n_j, \dots, n_r$ , respectivamente. El modelo muestral vendrá dado por

$$y_{m,i}^{(j)} = \sum_{h=0}^{q-1} \beta_{hj} t_m^h + \varepsilon_i = P_j(t_m) + \varepsilon_i,$$

siendo  $y_{m,i}^{(j)}$  el valor de la variable crecimiento del individuo  $i$ -ésimo,  $i = 1, \dots, n_j$ , de la muestra del grupo  $G_j$ ,  $j = 1, \dots, r$ , en el instante  $t_m$ ,  $m = 1, \dots, p$ , y  $\varepsilon_i$  su correspondiente perturbación o error.

Representando por

$$\mathbf{y}_i^{(j)} = \begin{bmatrix} y_{1,i}^{(j)} & \cdots & y_{p,i}^{(j)} \end{bmatrix}^T$$

el vector de observaciones del individuo  $i$ -ésimo de la muestra de  $G_j$ , la matriz de datos correspondiente a esta muestra es

$$\mathbf{Y}_j = \begin{bmatrix} \mathbf{y}_1^{(j)} & \cdots & \mathbf{y}_{n_j}^{(j)} \end{bmatrix},$$

$(p \times n_j)$

donde cada columna recoge la información correspondiente a cada individuo del  $j$ -ésimo grupo.

Denotando por  $n = \sum_{j=1}^r n_j$  el número total de observaciones, la matriz de datos conjunta es

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 & \cdots & \mathbf{Y}_j & \cdots & \mathbf{Y}_r \end{bmatrix}.$$

$(p \times n)$

Sea

$$\boldsymbol{\beta}_j^T = \begin{bmatrix} \beta_{0j} & \beta_{1j} & \cdots & \beta_{(q-1)j} \end{bmatrix}$$

$(1 \times q)$

el vector de los coeficientes de regresión o de la curva de crecimiento para los individuos muestrales del  $j$ -ésimo grupo y

$$\mathbf{X} = \begin{bmatrix} t_1^0 & t_1^1 & \cdots & t_1^{q-1} \\ t_2^0 & t_2^1 & \cdots & t_2^{q-1} \\ \cdots & \cdots & \cdots & \cdots \\ t_p^0 & t_p^1 & \cdots & t_p^{q-1} \end{bmatrix} \quad (1.2.1)$$

$(p \times q)$

la matriz del diseño, de forma que en cada fila se recogen las potencias correspondientes al mismo instante. Entonces

$$\mathbf{y}_i^{(j)} = \mathbf{X} \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_i^{(j)},$$

donde  $\boldsymbol{\varepsilon}_i^{(j)}$  representa el vector error o perturbación aleatoria.

Para cada muestra, se obtiene

$$\begin{aligned} \mathbf{Y}_j &= \begin{bmatrix} \mathbf{y}_1^{(j)} & \cdots & \mathbf{y}_{n_j}^{(j)} \end{bmatrix} = \begin{bmatrix} \mathbf{X} \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_1^{(j)} & \cdots & \mathbf{X} \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_{n_j}^{(j)} \end{bmatrix} = \\ &= \mathbf{X} \boldsymbol{\beta}_j \mathbf{Z}_{1n_j} + \boldsymbol{\mathcal{E}}_j, \end{aligned}$$

$(p \times n_j)$

donde por  $\mathbf{Z}_{ab}$  se representa una matriz  $(a \times b)$  constituida por unos y

$$\boldsymbol{\mathcal{E}}_j = \begin{bmatrix} \boldsymbol{\varepsilon}_1^{(j)} & \boldsymbol{\varepsilon}_2^{(j)} & \cdots & \boldsymbol{\varepsilon}_{n_j}^{(j)} \end{bmatrix}.$$



Conjuntamente, para la matriz de datos se tiene

$$\begin{aligned} \underset{(p \times n)}{\mathbf{Y}} &= [ \mathbf{Y}_1 \quad \cdots \quad \mathbf{Y}_j \quad \cdots \quad \mathbf{Y}_r ] = \\ &= [ \mathbf{X} \boldsymbol{\beta}_1 Z_{1n_1} \quad \cdots \quad \mathbf{X} \boldsymbol{\beta}_j Z_{1n_j} \quad \cdots \quad \mathbf{X} \boldsymbol{\beta}_r Z_{1n_r} ] + \boldsymbol{\mathcal{E}} = \\ &= \mathbf{X} [ \boldsymbol{\beta}_1 \quad \cdots \quad \boldsymbol{\beta}_j \quad \cdots \quad \boldsymbol{\beta}_r ] \mathbf{Z} + \boldsymbol{\mathcal{E}}, \end{aligned}$$

es decir,

$$\underset{(p \times n)}{\mathbf{Y}} = \underset{(p \times q)}{\mathbf{X}} \underset{(q \times r)}{\mathbf{B}} \underset{(r \times n)}{\mathbf{Z}} + \underset{(p \times n)}{\boldsymbol{\mathcal{E}}},$$

siendo la matriz de coeficientes de regresión

$$\underset{(q \times r)}{\mathbf{B}} = ((\beta_{hj})), \quad h = 0, \dots, q-1, \quad j = 1, \dots, r, \quad (1.2.2)$$

y

$$\mathbf{Z} = \text{diag} \left[ \begin{array}{ccc} Z_{1n_1} & \vdots & \cdots \\ & & \vdots \\ & & & Z_{1n_r} \end{array} \right] \quad (1.2.3)$$

una matriz  $(r \times n)$  diagonal por bloques, con  $Z_{1n_j}$  ( $j = 1, 2, \dots, r$ ) en los bloques diagonales y matrices nulas en los restantes bloques, y

$$\underset{(p \times n)}{\boldsymbol{\mathcal{E}}} = [ \boldsymbol{\mathcal{E}}_1 \quad \cdots \quad \boldsymbol{\mathcal{E}}_j \quad \cdots \quad \boldsymbol{\mathcal{E}}_r ] \quad (1.2.4)$$

la matriz de errores o perturbaciones aleatorias.

Para el estudio del modelo, se supone que se verifican las hipótesis clásicas sobre los vectores de errores:

$$H_1) \text{ Insesgadez, } E \left[ \boldsymbol{\varepsilon}_i^{(j)} \right] = \mathbf{0}, \quad \forall i = 1, \dots, n_j, \quad j = 1, \dots, r.$$

$$H_2) \text{ Homocedasticidad, } \text{Var} \left[ \boldsymbol{\varepsilon}_i^{(j)} \right] = \boldsymbol{\Sigma}, \quad \forall i = 1, \dots, n_j, \quad j = 1, \dots, r.$$

$$H_3) \text{ Incorrelación, } \text{Cov} \left[ \boldsymbol{\varepsilon}_i^{(j)}, \boldsymbol{\varepsilon}_s^{(l)} \right] = 0, \quad \forall (i, j) \neq (s, l).$$

Suponiendo que se cumplen las hipótesis anteriores  $H_1$ ),  $H_2$ ) y  $H_3$ ), un modelo de curvas de crecimiento viene definido por

$$\underset{(p \times n)}{\mathbf{Y}} = \underset{(p \times q)}{\mathbf{X}} \underset{(q \times r)}{\mathbf{B}} \underset{(r \times n)}{\mathbf{Z}} + \underset{(p \times n)}{\boldsymbol{\mathcal{E}}} \quad (1.2.5)$$

donde:

- $\underset{(p \times n)}{\mathbf{Y}}$  es la matriz de observaciones.
- $\underset{(q \times r)}{\mathbf{B}}$  es una matriz de parámetros desconocida o coeficientes de regresión.
- $\underset{(p \times q)}{\mathbf{X}}$  con  $q \leq p$  y  $\underset{(r \times n)}{\mathbf{Z}}$  con  $\text{rg}(\mathbf{Z}) = r \leq n$  son las matrices del diseño "inter" y "entre" individuos, respectivamente.

Supuesto cierto el modelo, se verifica que

$$E[\mathbf{Y}] = \mathbf{X}\mathbf{B}\mathbf{Z}$$

y

$$\text{Var}[\text{vec}(\mathbf{Y})] = \mathbf{I}_n \otimes \boldsymbol{\Sigma},$$

que se denotará, por simplicidad, por  $\text{Var}[\mathbf{Y}]$ , siendo  $\text{vec}(\cdot)$  y  $\otimes$  los operadores matriciales dados en las definiciones B.0.1 y B.0.5, respectivamente.

Para estudios posteriores (estimación, análisis de influencia, etc.) al modelo (1.2.5) hay que imponerle algunas hipótesis distribucionales, en particular, en esta memoria se supondrá que las columnas de la matriz de errores o perturbaciones,  $\boldsymbol{\mathcal{E}}$ , están constituidas por  $p$  variables normales multivariantes independientes con vector de media  $\mathbf{0}$  y matriz de varianzas y covarianzas común desconocida  $\boldsymbol{\Sigma}_{(p \times p)} > \mathbf{0}$ , por lo que, siguiendo la notación recogida en la sección C.2,

$$\mathbf{Y} \sim N_{p,n}(\mathbf{X}\mathbf{B}\mathbf{Z}; \boldsymbol{\Sigma}, \mathbf{I}_n).$$

Otro de los aspectos que debe de ser analizado previamente en el MCC es el estudio de la estructura de la matriz de covarianzas,  $\boldsymbol{\Sigma}$ , ya que un conocimiento adecuado de la misma, además de ser útil para la interpretación de la variación aleatoria de los datos, juega un papel fundamental en el ámbito de la diagnosis del modelo, de la detección de observaciones outliers y de la identificación de observaciones influyentes.

Aunque en la literatura se han propuesto diferentes estructuras de covarianza (Kanda, 1994 y Pan y Fang, 2002), en esta memoria se trabaja con las dos estructuras más utilizadas en el MCC:

- (1) Matriz de covarianzas no estructurada,  $\boldsymbol{\Sigma}$  es una matriz arbitraria, con la única restricción de que sea definida positiva,  $\boldsymbol{\Sigma} > \mathbf{0}$ . Este caso se utiliza en aquellas ocasiones en las cuales tanto las bases teóricas como empíricas no proporcionan información adicional sobre la estructura de la matriz de covarianzas.
- (2) Matriz de covarianzas con estructura simple de Rao (SCS), (Rao, 1967), la cual tiene como base los modelos de regresión de coeficientes aleatorios (Swamy, 1971 y Pan y Fang, 2002), una forma específica de los modelos lineales mixtos.

Para más detalles sobre estructuras de la matriz de covarianzas y su elección, puede consultarse Chinchilli y Carter (1984), Keramidias y Lee (1990), Lee (1991), Keramidias y Lee (1995) y Pan y Fang (2002).

En resumen, el Modelo de Curvas de Crecimiento (MCC) que se aborda en esta memoria viene dado por

$$\underset{(p \times n)}{\mathbf{Y}} = \underset{(p \times q)}{\mathbf{X}} \underset{(q \times r)}{\mathbf{B}} \underset{(r \times n)}{\mathbf{Z}} + \underset{(p \times n)}{\boldsymbol{\mathcal{E}}}$$

verificando las hipótesis  $H_1$ ,  $H_2$  y  $H_3$ , siendo las matrices  $\mathbf{X}$ ,  $\mathbf{B}$ ,  $\mathbf{Z}$  y  $\boldsymbol{\mathcal{E}}$  dadas en (1.2.1), (1.2.2), (1.2.3) y (1.2.4), respectivamente.

**Nota 1.2.1** Algunos autores denominan a  $\mathbf{Z}$  matriz indicadora de grupo, ya que sus elementos correspondientes, 1 y 0, indican el grupo de pertenencia de cada observación, por lo que  $\mathbf{Z}$  modela la estructura entre individuos. En este modelo,  $\mathbf{Z}$  representa la misma matriz que se utiliza en el análisis de la varianza multivariante clásico.

**Nota 1.2.2** En esta memoria, se supone que las matrices del diseño  $\mathbf{X}$  y  $\mathbf{Z}$  son de rango total, es decir,  $\text{rg}(\mathbf{X}) = q < p$  y  $\text{rg}(\mathbf{Z}) = r < n$ . En Kollo y von Rosen (2005) se aborda un estudio más general del modelo.

### 1.3. Análisis del modelo

Cuando no se dispone de información adicional, la forma más natural para el estudio del MCC es considerar que la matriz de covarianzas,  $\mathbf{\Sigma}$ , no presenta ninguna estructura, se dice en este caso que la matriz de covarianzas es no estructurada. Como se ha comentado anteriormente, la única restricción sobre la misma es que sea definida positiva,  $\mathbf{\Sigma} > \mathbf{0}$ . Al no imponerse hipótesis previas, proporciona flexibilidad al modelo, aunque aumente el número de parámetros desconocidos. Por ello, es especialmente útil cuando el número de instantes de tiempo (número de parámetros de la estructura de covarianza) es relativamente pequeño en relación al tamaño muestral y todos los individuos son medidos en los mismos instantes de tiempo.

A continuación, se recogen de forma resumida los métodos más usuales de estimación de los parámetros  $\mathbf{B}$  y  $\mathbf{\Sigma}$ : el método de los mínimos cuadrados generalizados (GLSE) (Pan y Fang, 2002 y Kariya y Kurata, 2004) y el método de estimación de máxima verosimilitud (EMV) (Pan y Fang, 2002 y Kollo y von Rosen, 2005).

Se comienza recogiendo el estimador de la matriz de coeficientes de regresión obtenido por el método GLSE y, a partir del mismo, el estimador de la matriz de covarianzas. Para tal fin, en el MCC se define la función matricial suma de cuadrados y productos cruzados debida al error

$$Q(\mathbf{B}) = (\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})^T,$$

matriz  $(p \times p)$  no negativa. Así, para el MCC, un estimador GLS  $\tilde{\mathbf{B}}$  de la matriz de coeficientes de regresión, debe satisfacer

$$\tilde{\mathbf{B}} = \arg \min_{\mathbf{B}} \{\text{tr}[Q(\mathbf{B})]\},$$

donde  $\mathbf{B}$  recorre el espacio real de matrices de orden  $(q \times r)$ .

En el siguiente teorema se recoge la solución a este problema de optimización, su demostración viene dada en Pan y Fang (2002).

**Teorema 1.3.1** Para el MCC, un GLSE de la matriz de coeficientes de regresión  $\mathbf{B}$  viene dado por

$$\tilde{\mathbf{B}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1}. \quad (1.3.1)$$

**Nota 1.3.1** Bajo las condiciones de rango total de las matrices  $\mathbf{X}$  y  $\mathbf{Z}$ , el GLSE  $\tilde{\mathbf{B}}$  es único (von Rosen, 1984 y Pan, 1988), además es una función lineal de la matriz de respuestas  $\mathbf{Y}$ .

A partir de (1.3.1), se obtiene el siguiente estimador para  $\Sigma$ ,

$$\tilde{\Sigma} = \frac{1}{n}Q(\tilde{\mathbf{B}}) = \frac{1}{n}(\mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}\mathbf{Z})^T = \frac{1}{n}\mathbf{E}\mathbf{E}^T,$$

donde  $\mathbf{E}$  es la matriz residual  $\mathbf{E} = \mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}\mathbf{Z}$ . Por abuso del lenguaje, se denomina a  $\tilde{\Sigma}$  el GLSE de  $\Sigma$ .

La matriz de suma de cuadrados y productos cruzados residual admite la siguiente descomposición

$$Q(\tilde{\mathbf{B}}) = (\mathcal{E} + \mathbf{X}(\mathbf{B} - \tilde{\mathbf{B}})\mathbf{Z})(\mathcal{E} + \mathbf{X}(\mathbf{B} - \tilde{\mathbf{B}})\mathbf{Z})^T = \mathbf{Q}_1 + \mathbf{Q}_2, \quad (1.3.2)$$

con

$$\mathbf{Q}_1 = \mathcal{E}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})\mathcal{E}^T$$

y

$$\mathbf{Q}_2 = (\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\mathcal{E}\mathbf{P}_{\mathbf{Z}^T}\mathcal{E}^T(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}),$$

siendo

$$\mathbf{P}_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$$

y

$$\mathbf{P}_{\mathbf{Z}^T} = \mathbf{Z}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}, \quad (1.3.3)$$

las matrices proyecciones generadas por  $\mathbf{X}$  y  $\mathbf{Z}^T$ , respectivamente.

Usando la descomposición (1.3.2), el GLSE de  $\Sigma$  se puede expresar por

$$\tilde{\Sigma} = \frac{1}{n}(\mathbf{S} + \mathbf{P}_{\mathbf{X}^\circ}\mathbf{Y}\mathbf{P}_{\mathbf{Z}^T}\mathbf{Y}^T\mathbf{P}_{\mathbf{X}^\circ}), \quad (1.3.4)$$

donde

$$\mathbf{S} = \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})\mathbf{Y}^T \quad (1.3.5)$$

y

$$\mathbf{P}_{\mathbf{X}^\circ} = \mathbf{X}^\circ(\mathbf{X}^{\circ T}\mathbf{X}^\circ)^{-1}\mathbf{X}^{\circ T} = \mathbf{I}_p - \mathbf{P}_{\mathbf{X}},$$

siendo  $\mathbf{X}^\circ$  una matriz cuyas columnas generan el espacio matricial ortogonal de  $\mathbf{X}$ , es decir,  $\mathcal{C}(\mathbf{X}^\circ) = \mathcal{C}(\mathbf{X})^\perp$  (véase nota B.0.1).

Además, bajo la hipótesis distribucional  $\mathbf{Y} \sim N_{p,n}(\mathbf{X}\mathbf{B}\mathbf{Z}; \Sigma, \mathbf{I}_n)$ , como el GLSE  $\tilde{\mathbf{B}}$  es una transformación lineal de la variable respuesta, teniendo en cuenta el teorema C.1.2, se obtiene el siguiente resultado.

**Teorema 1.3.2** *En el MCC, bajo la hipótesis de normalidad, la distribución del GLSE  $\tilde{\mathbf{B}}$  de la matriz de coeficientes de regresión  $\mathbf{B}$  viene dada por*

$$\tilde{\mathbf{B}} \sim N_{q,r}(\mathbf{B}; (\mathbf{X}^T\mathbf{X})^{-1}(\mathbf{X}^T\Sigma\mathbf{X})(\mathbf{X}^T\mathbf{X})^{-1}, (\mathbf{Z}\mathbf{Z}^T)^{-1}).$$

De este resultado se puede concluir que  $\tilde{\mathbf{B}}$  es un estimador insesgado de  $\mathbf{B}$ . Además, sustituyendo  $\Sigma$  por  $\tilde{\Sigma}$  se obtiene el siguiente estimador de  $\text{Var}[\tilde{\mathbf{B}}]$ ,

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}] = \frac{1}{n}(\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{S}\mathbf{X})(\mathbf{X}^T\mathbf{X})^{-1},$$

ya que, de (1.3.4) y dado que  $\mathbf{P}_{\mathbf{X}}\circ\mathbf{X} = \mathbf{0}$ , se tiene que

$$\mathbf{X}^T\tilde{\Sigma}\mathbf{X} = \frac{1}{n}\mathbf{X}^T\mathbf{S}\mathbf{X}.$$

En el MCC, considerando que  $\mathcal{E} = \mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z} \sim N_{p,n}(\mathbf{0}; \Sigma, \mathbf{I}_n)$ , la descomposición (1.3.2) y el teorema C.4.2, se obtienen los siguientes resultados:

$$\begin{aligned} \mathbf{Q}_1 &= \mathcal{E}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})\mathcal{E}^T \sim W_p(n-r, \Sigma) \Rightarrow \\ &\Rightarrow \begin{cases} \text{E}[\mathbf{Q}_1] = (n-r)\Sigma \\ \text{Var}[\mathbf{Q}_1] = (n-r)(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})\Sigma^{\otimes 2} \end{cases} \end{aligned}$$

y

$$\begin{aligned} \mathbf{Q}_2 &= (\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\mathcal{E}\mathbf{P}_{\mathbf{Z}^T}\mathcal{E}^T(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}) \sim W_p(r, (\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\Sigma(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})) \Rightarrow \\ &\Rightarrow \begin{cases} \text{E}[\mathbf{Q}_2] = r(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\Sigma(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}) \\ \text{Var}[\mathbf{Q}_2] = r(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})((\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\Sigma(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}))^{\otimes 2} \end{cases}, \end{aligned}$$

donde  $\mathbf{K}_{p^2}$  es la matriz conmutación de orden  $(p^2 \times p^2)$ , dada en la definición B.0.2, y  $\mathbf{A}^{\otimes 2} = \mathbf{A} \otimes \mathbf{A}$ . Dado que  $\mathbf{Q}_1$  y  $\mathcal{E}\mathbf{P}_{\mathbf{Z}^T}\mathcal{E}^T$  son independientes, al ser  $(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})\mathbf{P}_{\mathbf{Z}^T} = \mathbf{0}$ , las matrices  $\mathbf{Q}_1$  y  $\mathbf{Q}_2$  son mutuamente independientes. Por tanto, la distribución de  $Q(\tilde{\mathbf{B}}) = n\tilde{\Sigma}$  puede ser expresada como una suma de distribuciones de Wishart independientes, aunque de parámetros distintos. En consecuencia,

$$\text{E}[\tilde{\Sigma}] = \text{E}\left[\frac{1}{n}Q(\tilde{\mathbf{B}})\right] = \frac{1}{n}\text{E}[\mathbf{Q}_1 + \mathbf{Q}_2] = \frac{n-r}{n}\Sigma + \frac{r}{n}(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\Sigma(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}),$$

es decir, dicho estimador no es insesgado. No obstante, dado que  $r$  es fijo, el estimador  $\tilde{\Sigma}$  es asintóticamente insesgado de la componente de dispersión  $\Sigma$ . Además

$$\begin{aligned} \text{Var}[\tilde{\Sigma}] &= \text{Var}\left[\frac{1}{n}Q(\tilde{\mathbf{B}})\right] = \frac{1}{n^2}\text{Var}[\mathbf{Q}_1 + \mathbf{Q}_2] = \\ &= (\mathbf{I}_{p^2} + \mathbf{K}_{p^2})\left\{\frac{n-r}{n^2}\Sigma^{\otimes 2} + \frac{r}{n^2}((\mathbf{I}_p - \mathbf{P}_{\mathbf{X}})\Sigma(\mathbf{I}_p - \mathbf{P}_{\mathbf{X}}))^{\otimes 2}\right\}. \end{aligned}$$

A continuación, se aborda la estimación de los parámetros por el método EMV. Este método es de los más utilizados en la inferencia estadística por las propiedades asintóticas que presentan los estimadores. Existen varias aproximaciones al método, algunas de las cuales vienen recogidas en Kollo y von Rosen (2005).

En el MCC, bajo la hipótesis distribucional  $\mathbf{Y} \sim N_{p,n}(\mathbf{X}\mathbf{B}\mathbf{Z}; \mathbf{\Sigma}, \mathbf{I}_n)$  y, teniendo en cuenta la función de densidad (C.2.1), la función de verosimilitud conjunta de la muestra es

$$V(\mathbf{B}, \mathbf{\Sigma}) = (2\pi)^{-np/2} |\mathbf{\Sigma}|^{-n/2} \exp\left(-\frac{1}{2} \text{tr} [\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})^T]\right),$$

y, por tanto, el logaritmo neperiano de la función de verosimilitud vendrá dado por

$$\begin{aligned} L(\mathbf{B}, \mathbf{\Sigma}) &= \ln V(\mathbf{B}, \mathbf{\Sigma}) = \\ &= -\frac{np}{2} \ln 2\pi - \frac{n}{2} \ln \det[\mathbf{\Sigma}] - \frac{1}{2} \text{tr} [\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})^T], \end{aligned}$$

por lo que se obtendrán las siguientes ecuaciones de verosimilitud

$$\begin{cases} \frac{\partial L}{\partial \text{vec}^T(\mathbf{B})} = 0 \\ \frac{\partial L}{\partial \text{svec}^T(\mathbf{\Sigma}^{-1})} = 0 \end{cases} \Leftrightarrow \begin{cases} \mathbf{X}^T \mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})\mathbf{Z}^T = 0 \\ n\mathbf{\Sigma} - (\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\mathbf{B}\mathbf{Z})^T = 0 \end{cases}. \quad (1.3.6)$$

Este sistema ha sido resuelto por Elswick (1985) y por von Rosen (1989), obteniéndose el siguiente teorema cuya demostración puede ser vista en Srivastava y von Rosen (1999), Pan y Fang (2002) y Kollo y von Rosen (2005).

**Teorema 1.3.3** *Para el MCC, bajo la hipótesis de normalidad, si  $n > p + r$  y las matrices de diseño  $\mathbf{X}$  y  $\mathbf{Z}$  son de rango total entonces la solución  $(\hat{\mathbf{B}}, \hat{\mathbf{\Sigma}})$  del sistema de ecuaciones de verosimilitud (1.3.6) es única y viene dada por*

$$\hat{\mathbf{B}} = \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \quad (1.3.7)$$

y

$$\hat{\mathbf{\Sigma}} = \frac{1}{n} (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z})^T = \frac{1}{n} (\mathbf{S} + \hat{\mathbf{V}}\hat{\mathbf{V}}^T), \quad (1.3.8)$$

siendo

$$\hat{\mathbf{V}} = \mathbf{Y} \mathbf{P}_{\mathbf{Z}^T} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z} \quad (1.3.9)$$

y

$$\mathbf{S}_X = (\mathbf{X}^T \mathbf{S}^{-1} \mathbf{X})^{-1}, \quad (1.3.10)$$

donde  $\mathbf{S}$  está definida en (1.3.5),  $\mathbf{P}_{\mathbf{Z}^T}$  es la matriz de proyección generada por  $\mathbf{Z}^T$ , siendo  $\hat{\mathbf{B}}$  y  $\hat{\mathbf{\Sigma}}$  los estimadores de máxima verosimilitud de la matriz de coeficientes de regresión  $\mathbf{B}$  y de la componente de dispersión  $\mathbf{\Sigma}$ , respectivamente.

El teorema 1.3.3 proporciona una forma algebraica para la solución única  $(\widehat{\mathbf{B}}, \widehat{\Sigma})$  del sistema de verosimilitud (1.3.6) cuando ambas matrices explicativas  $\mathbf{X}$  y  $\mathbf{Z}$  son de rango total.

**Nota 1.3.2** Algunos comentarios respecto a los EMV.

- (i) Los EMV del MCC son función de la matriz  $\mathbf{S}$ . Por tanto, a diferencia del GLSE  $\widetilde{\mathbf{B}}$ , el EMV  $\widehat{\mathbf{B}}$  es una función no lineal de la matriz de respuestas  $\mathbf{Y}$ . No obstante,  $\widetilde{\mathbf{B}}$  se podría considerar como un caso particular de  $\widehat{\mathbf{B}}$  cuando  $\mathbf{S} = \mathbf{I}_p$ .
- (ii) Teniendo en cuenta el teorema 1.3.3, para  $n > p + r$ , como  $\mathbf{S} \sim W_p(n - r, \Sigma)$ , por la propiedad (i) del teorema C.4.2, se tiene que la matriz aleatoria  $\mathbf{S}$  es simétrica y definida positiva con probabilidad la unidad, por tanto, tiene sentido considerar  $\mathbf{S}^{-1}$ .
- (iii) Como  $\mathbf{S} \sim W_p(n - r, \Sigma)$ , por el lema C.4.1,  $E[\mathbf{S}] = (n - r)\Sigma$ , es decir, el estadístico

$$\widehat{\Sigma} = \frac{1}{n - r} \mathbf{S} \quad (1.3.11)$$

es un estimador insesgado de  $\Sigma$ .

- (iv) Considerando

$$\mathbf{X}_S^o = \mathbf{S} \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{S} \mathbf{X}^o)^{-1} \mathbf{X}^{oT}, \quad (1.3.12)$$

una matriz idempotente con  $\mathcal{C}(\mathbf{X}^o) = \mathcal{C}(\mathbf{X})^\perp$ , usando la nota B.0.2, el EMV  $\widehat{\Sigma}$  dado en (1.3.8) puede ser expresado por

$$\widehat{\Sigma} = \frac{1}{n} (\mathbf{S} + \mathbf{X}_S^o \mathbf{Y} \mathbf{P}_{\mathbf{Z}^T} \mathbf{Y}^T \mathbf{X}_S^{oT}), \quad (1.3.13)$$

que no depende de la elección de la matriz  $\mathbf{X}^o$ .

- (v) Del lema B.0.6 resulta

$$\begin{aligned} \mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} &= \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{S} \mathbf{X}^o)^{-1} \mathbf{X}^{oT} = \\ &= \mathbf{S}^{-1} \mathbf{X}_S^o = \mathbf{X}_S^{oT} \mathbf{S}^{-1} = \mathbf{X}_S^{oT} \mathbf{S}^{-1} \mathbf{X}_S^o, \end{aligned} \quad (1.3.14)$$

aplicando este resultado y la igualdad (B.0.1) a (1.3.13), se obtiene

$$\frac{1}{n} \widehat{\Sigma}^{-1} = \mathbf{S}^{-1} - \mathbf{S}^{-1} \mathbf{X}_S^o ((\mathbf{Y} \mathbf{P}_{\mathbf{Z}^T} \mathbf{Y}^T)^{-1} + \mathbf{S}^{-1} \mathbf{X}_S^o)^{-1} \mathbf{S}^{-1} \mathbf{X}_S^o$$

y dado que  $\mathbf{X}^{oT} \mathbf{X} = \mathbf{0}$ , se tiene

$$\mathbf{X}^T \widehat{\Sigma}^{-1} \mathbf{X} = n \mathbf{X}^T \mathbf{S}^{-1} \mathbf{X}.$$

- (vi) La expresión para la matriz de valores ajustados de la variable respuesta,  $\mathbf{Y}$ , viene dada en Pan y Fang (2002) y en Kollo y von Rosen (2005),

$$\widehat{\mathbf{Y}} = \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z} = \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} \mathbf{P}_{\mathbf{Z}^T}. \quad (1.3.15)$$

En el siguiente teorema se obtiene la esperanza matemática y la matriz de covarianzas del EMV  $\widehat{\mathbf{B}}$  y de la matriz de valores ajustados (véase Kollo y von Rosen, 2005).

**Teorema 1.3.4** *En el MCC, bajo la hipótesis de normalidad, sea  $\widehat{\mathbf{B}}$  dado en (1.3.7) y  $\mathbf{X}\widehat{\mathbf{B}}\mathbf{Z}$  en (1.3.15). Se verifica:*

(i)

$$\mathbb{E}[\widehat{\mathbf{B}}] = \mathbf{B} \quad (1.3.16)$$

y

$$\mathbb{E}[\mathbf{X}\widehat{\mathbf{B}}\mathbf{Z}] = \mathbf{X}\mathbf{B}\mathbf{Z},$$

o sea, el EMV  $\widehat{\mathbf{B}}$  es un estimador insesgado de la matriz de coeficientes de regresión  $\mathbf{B}$ .

(ii) Si  $n - r - (p - q) - 1 > 0$ , entonces

$$\text{Var}[\widehat{\mathbf{B}}] = c_1 (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \quad (1.3.17)$$

y

$$\text{Var}[\mathbf{X}\widehat{\mathbf{B}}\mathbf{Z}] = c_1 \mathbf{Z}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z} \otimes \mathbf{X} (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T,$$

donde

$$c_1 = \frac{n - r - 1}{n - r - (p - q) - 1} \quad (1.3.18)$$

es una constante positiva.

**Nota 1.3.3** *Como  $\text{Var}[\widehat{\mathbf{B}}]$  depende de la matriz de parámetros  $\boldsymbol{\Sigma}$  desconocida, para obtener un estimador asintóticamente insesgado y consistente de  $\text{Var}[\widehat{\mathbf{B}}]$ , basta sustituir  $\boldsymbol{\Sigma}$  en (1.3.17) por (1.3.11), es decir,*

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}] = \frac{c_1}{n - r} (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \mathbf{S}_{\mathbf{X}}, \quad (1.3.19)$$

con  $\mathbf{S}_{\mathbf{X}}$  definida en (1.3.10).

**Nota 1.3.4** *Para el cálculo de los momentos del EMV  $\widehat{\mathbf{B}}$  puede consultarse von Rosen (1991a) y Kollo y von Rosen (2005). En este último, se recogen fórmulas explícitas para los momentos de orden superior, las mismas pueden ser utilizadas para obtener aproximaciones a la distribución de  $\widehat{\mathbf{B}}$ .*

A continuación, se estudian los momentos del EMV  $\widehat{\boldsymbol{\Sigma}}$ . El cálculo de  $\mathbb{E}[\widehat{\boldsymbol{\Sigma}}]$  y el de  $\text{Var}[\widehat{\boldsymbol{\Sigma}}]$  fueron dados inicialmente por von Rosen (1991a) y vienen recogidos en el siguiente teorema, cuya demostración puede verse en Pan y Fang (2002) y Kollo y von Rosen (2005).

**Teorema 1.3.5** *Para el MCC, bajo la hipótesis de normalidad, sea  $\widehat{\boldsymbol{\Sigma}}$  el EMV de la componente de dispersión  $\boldsymbol{\Sigma}$  dado en (1.3.8). Se verifica:*

(i) Si  $n - r - (p - q) - 1 > 0$  entonces

$$\mathbb{E}[\widehat{\boldsymbol{\Sigma}}] = \boldsymbol{\Sigma} - c_2 \mathbf{X} (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T,$$

con

$$c_2 = \frac{r}{n} \frac{n - r - 2(p - q) - 1}{n - r - (p - q) - 1}.$$



(ii) Si  $n - r - (p - q) - 3 > 0$  resulta

$$\begin{aligned} \text{Var}[\widehat{\boldsymbol{\Sigma}}] &= c_3(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})[(\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T) \otimes (\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T)] + \\ &+ c_4(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})[(\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T) \otimes (\boldsymbol{\Sigma} - \mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T)] + \\ &+ c_4(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})[(\boldsymbol{\Sigma} - \mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T) \otimes (\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T)] + \\ &+ \frac{1}{n}(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})[(\boldsymbol{\Sigma} - \mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T) \otimes (\boldsymbol{\Sigma} - \mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T)] + \\ &+ c_5 \text{vec}(\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T) \text{vec}^T(\mathbf{X}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}^T), \end{aligned}$$

donde  $\mathbf{K}_{p^2}$  es la matriz conmutación de orden  $(p^2 \times p^2)$ , siendo

$$c_3 = \frac{n-r}{n^2} + \frac{2r(p-q)}{n^2(n-r-(p-q)-1)} + \frac{r}{n^2}(2s_1 + s_2 + s_3) + \frac{r^2}{n^2}s_3,$$

$$c_4 = \frac{n-(p-q)-1}{n(n-r-(p-q)-1)},$$

$$c_5 = \frac{2r(n-r-1)(n-(p-q)-1)(p-q)}{n^2(n-r-(p-q))(n-r-(p-q)-1)^2(n-r-(p-q)-3)},$$

$$s_1 = \frac{p-q}{n-r-(p-q)-1},$$

$$s_2 = \frac{(p-q)^2(n-r-(p-q)-2) + 2(p-q)}{(n-r-(p-q))(n-r-(p-q)-1)(n-r-(p-q)-3)}$$

y

$$\begin{aligned} s_3 &= \frac{p-q}{(n-r-(p-q))(n-r-(p-q)-3)} + \\ &+ \frac{(p-q)^2}{(n-r-(p-q))(n-r-(p-q)-1)(n-r-(p-q)-3)}. \end{aligned}$$

Del teorema 1.3.5 se concluye que el EMV  $\widehat{\boldsymbol{\Sigma}}$ , dado en (1.3.8), es un estimador sesgado de la matriz de covarianzas  $\boldsymbol{\Sigma}$ .

A continuación, se recogen algunos resultados relacionados con la distribución de  $\widehat{\mathbf{B}}$ , los cuales serán de gran utilidad para el desarrollo de las medidas de influencia propuestas en esta memoria. A partir de (1.3.7) se concluye que el EMV  $\widehat{\mathbf{B}}$  es una función no lineal de la matriz de respuestas,  $\mathbf{Y}$ , constituido por dos partes aleatorias:  $\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1}$  (no lineal) e  $\mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1}$ . Tal hecho hace más compleja la inferencia estadística en este modelo. En particular, es difícil obtener formas analíticas para las distribuciones de  $\widehat{\mathbf{B}}$  y  $\widehat{\boldsymbol{\Sigma}}$ , así como para sus momentos de orden superior. Diferentes autores han intentado obtener la forma exacta de la distribución de  $\widehat{\mathbf{B}}$ , entre los que se pueden citar a Gleser y Olkin (1970), Kabe (1975), Kenward (1986) y Fujikoshi (1987). Todos los resultados obtenidos dan formas complejas de difícil aplicación, por lo que se han propuesto diferentes aproximaciones. A continuación, se recogen algunas formas explícitas simples de aproximación de la función de densidad del EMV  $\widehat{\mathbf{B}}$ . Para una aproximación a la distribución del EMV  $\widehat{\boldsymbol{\Sigma}}$  de la componente de dispersión puede verse Kollo y von Rosen (2005).

Existen varias estrategias que pueden ser útiles para aproximarse a la distribución de  $\widehat{\mathbf{B}}$ . La más utilizada es la aproximación asintótica a la distribución normal, la cual se basa en la convergencia de la matriz  $\mathbf{S}$ , en concreto en que  $\frac{1}{n-r}\mathbf{S}$  converge en probabilidad a  $\boldsymbol{\Sigma}$ : la distribución de  $\widehat{\mathbf{B}}$  puede aproximarse por una distribución  $N_{q,r}(\mathbf{B}; (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$  (Kollo et al., 2007).

En el siguiente teorema, se recoge otra aproximación (véase Kollo y von Rosen, 2005). Una ventaja de este resultado es que garantiza una aproximación a una función de densidad normal matricial con la misma matriz de covarianzas que  $\widehat{\mathbf{B}}$ .

**Teorema 1.3.6** *Una aproximación para la función de densidad del estimador  $\widehat{\mathbf{B}}$ , definido en (1.3.7), viene dada por  $f_N^*(\mathbf{B}_0)$  función de densidad de una matriz aleatoria con distribución normal matricial*

$$N_{q,r}(\mathbf{B}; c_1(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1}),$$

con  $c_1$  dado en (1.3.18). El orden del error de esta aproximación es

$$|f_N^*(\mathbf{B}_0) - f_{\widehat{\mathbf{B}}}(\mathbf{B}_0)| = O(n^{-2}), \quad \forall \mathbf{B}_0.$$

Kollo y von Rosen (2005) y Kollo et al. (2007) obtienen una aproximación de la función de densidad del estimador  $\widehat{\mathbf{B}}$  más acurada que la anterior, la cual viene recogida en el siguiente teorema.

**Teorema 1.3.7** *Dada la función matricial*

$$f_{\mathbf{B}_E}(\mathbf{B}_0) = f_N(\mathbf{B}_0) \left[ 1 + \frac{1}{2}s \left( \text{tr} \left[ \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} (\mathbf{B}_0 - \mathbf{B}) \mathbf{Z} \mathbf{Z}^T (\mathbf{B}_0 - \mathbf{B})^T \right] - qr \right) \right], \quad (1.3.20)$$

con

$$s = \frac{p - q}{n - r - (p - q) - 1}, \quad (1.3.21)$$

donde  $f_N(\cdot)$  es la función de densidad de una matriz aleatoria con distribución normal matricial  $N_{q,r}(\mathbf{B}; (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$ , se verifica:

(i) *La función  $f_{\mathbf{B}_E}$  es una función de densidad si*

$$0 < 1 - \frac{1}{2}sqr < 1. \quad (1.3.22)$$

(ii) *La función  $f_{\mathbf{B}_E}$  es una aproximación de la función de densidad del estimador  $\widehat{\mathbf{B}}$ , siendo el orden de error*

$$|f_{\mathbf{B}_E}(\mathbf{B}_0) - f_{\widehat{\mathbf{B}}}(\mathbf{B}_0)| = O(n^{-2}), \quad \forall \mathbf{B}_0.$$

**Nota 1.3.5** La condición (1.3.22) es equivalente a

$$n > (p - q) \left( 1 + \frac{1}{2}qr \right) + r + 1.$$

Del teorema anterior se obtiene que, en caso de que se verifique (1.3.22),  $f_{\mathbf{B}_E}(\mathbf{B}_0)$  es una función de densidad de una distribución matricial elíptica cuyos parámetros vienen recogidos en el siguiente teorema (Kollo y von Rosen, 2005).

**Teorema 1.3.8** Si se verifica (1.3.22), la función  $f_{\mathbf{B}_E}$  dada en (1.3.20) es la función de densidad de una distribución matricial elíptica tal que:

(i)  $E[\mathbf{B}_E] = E[\widehat{\mathbf{B}}] = \mathbf{B}.$

(ii)  $\text{Var}[\mathbf{B}_E] = \text{Var}[\widehat{\mathbf{B}}] = c_1(\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1},$  con  $c_1 = s + 1$  y  $s$  definido en (1.3.21).

Por último, como se recoge en Kollo et al. (2007), esta aproximación se basa en una mixtura de distribuciones elípticas. En efecto, si se verifica la condición (1.3.22), la función de densidad  $f_{\mathbf{B}_E}$ , dada en el teorema 1.3.7, se puede expresar como mixtura de una distribución normal matricial,  $N_{q,r}(\mathbf{B}; (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}, (\mathbf{Z}\mathbf{Z}^T)^{-1})$ , con peso  $1 - sqr/2$ , y de una distribución de Kotz matricial,  $K_{q,r}(\mathbf{B}; (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}, (\mathbf{Z}\mathbf{Z}^T)^{-1})$ , con peso  $sqr/2$ , donde las dimensiones  $q$  y  $r$  vienen fijadas en el modelo (1.2.5).

## 1.4. Análisis del modelo con estructura simple de Rao

Como se ha comentado anteriormente, en la modelización no estructurada no se realiza ningún supuesto sobre las varianzas y covarianzas entre diferentes medidas. Así, paralelamente al hecho de que el número de parámetros de la covarianza crece rápidamente con el número de instantes de tiempo, este modelo tiene la desventaja de que no aprovecha las tendencias de varianzas y covarianzas en el tiempo.

Así, aunque la modelización no estructurada sea la más usada en el MCC, en determinadas ocasiones es recomendable la modelación de determinadas estructuras de covarianza en este modelo. Para una revisión de estructuras de covarianza véase Fitzmaurice et al. (2004).

Por ejemplo, cuando se pretenden explicar las variaciones provenientes de diferentes fuentes de datos (o cuando las observaciones son medidas repetidas en instantes de tiempo irregulares), se puede usar la estructura de covarianza con efectos aleatorios o coeficientes aleatorios

$$\boldsymbol{\Sigma} = \mathbf{X}\boldsymbol{\Gamma}\mathbf{X}^T + \sigma^2\mathbf{I}_p, \quad (1.4.1)$$

donde  $\boldsymbol{\Gamma} > \mathbf{0}$  y  $\sigma^2 > 0$  (Rao, 1965). Esta estructura también es analizada en Grizzle y Allen (1969), Reinsel (1982, 1985), Ware (1985), Kanda (1994), Yokoyama (1995), Nummi (1997), Pan y Fang (2002) e Ip et al. (2007).

Este es un modelo muy flexible que permite modelizar diversos tipos de comportamientos en las varianzas y covarianzas (correlaciones), entre los que se incluyen varianzas que crecen

o decrecen (como función de los tiempos de medición), así como correlaciones que pueden ser negativas o positivas. Sin embargo, no permite que la varianza sea constante si las correlaciones entre observaciones igualmente espaciadas no lo son. Además, el número de parámetros en estos modelos no está relacionado con el número de ocasiones (tiempos) en las que se realizan mediciones.

Por otro lado, cuando las lecturas en todos los tiempos tienen la misma varianza y todos los pares de medidas en la misma unidad experimental tienen la misma correlación, resulta adecuada la estructura simétrica compuesta (estructura de covarianza uniforme), definida por

$$\boldsymbol{\Sigma} = \sigma^2 [(1 - \rho)\mathbf{I}_p + \rho\mathbf{1}_p\mathbf{1}_p^T] \quad (1.4.2)$$

donde  $\sigma^2 > 0$  y  $-\frac{1}{p-1} < \rho < 1$  son parámetros desconocidos de dimensión apropiada (véase, por ejemplo, Kanda, 1990b, 1992 y 1994, Wu, 1998, Pan y Fang, 2002, Zezula, 2006 y Ye y Wang, 2009). Ello implica que el único aspecto de la covarianza entre medidas repetidas es debido a la contribución común de la unidad experimental, sin importar la distancia en el tiempo (entre pares de mediciones repetidas). Este modelo tiene solamente dos parámetros independientemente del número de instantes de tiempo. Así, es un modelo muy restringido que debe utilizarse sólo en los casos que las varianzas y correlaciones sean constantes.

Las estructuras de covarianza (1.4.1) y (1.4.2) son dos casos particulares de la estructura de covarianza simple de Rao, que, a continuación, se define.

**Definición 1.4.1** *En el modelo MCC, se dice que la componente de dispersión  $\boldsymbol{\Sigma}$  tiene una estructura de covarianza simple de Rao (SCS) si*

$$\boldsymbol{\Sigma} = \mathbf{X}\boldsymbol{\Gamma}\mathbf{X}^T + \mathbf{X}^o\boldsymbol{\Theta}\mathbf{X}^{oT},$$

donde  $\boldsymbol{\Gamma}_{(q \times q)} > \mathbf{0}$  y  $\boldsymbol{\Theta}_{((p-q) \times (p-q))} > \mathbf{0}$  son matrices desconocidas, y  $\mathbf{X}_{(p \times (p-q))}^o$  es una matriz ortogonal a la matriz  $\mathbf{X}$ , es decir,  $\mathcal{C}(\mathbf{X}^o) = \mathcal{C}(\mathbf{X})^\perp$  dado por (B.0.4).

Como alternativa en el MCC, en determinadas ocasiones es recomendable la estructura simple de Rao. Esta estructura de covarianzas fue estudiada inicialmente por Rao (1967), aunque su forma actual ha sido establecida por Geisser (1970). Resultados de interés sobre el modelo con esta estructura se encuentran en los trabajos de Grizzle y Allen (1969), Yu (1995), Kurata (1998), Pan y Fang (2002), Kariya y Kurata (2004) y Jammalamadaka et al. (2009).

**Nota 1.4.1** *En adelante, el MCC con SCS se denotará por MCC(SCS).*

A continuación, supuesta la hipótesis distribucional  $\mathbf{Y} \sim N_{p,n}(\mathbf{X}\mathbf{B}\mathbf{Z}; \boldsymbol{\Sigma}, \mathbf{I}_n)$ , se recopilan diferentes resultados sobre la estimación de los parámetros del MCC(SCS). En el siguiente teorema se recogen las expresiones analíticas de los EMV de los parámetros  $\mathbf{B}$ ,  $\boldsymbol{\Gamma}$  y  $\boldsymbol{\Theta}$ , la demostración puede verse en Pan y Fang (2002).

**Teorema 1.4.1** Para el MCC(SCS), bajo la hipótesis de normalidad, si  $n > p + r$ , los EMV de la matriz de coeficientes de regresión,  $\mathbf{B}$ , y de las componentes de dispersión,  $\mathbf{\Gamma}$  y  $\mathbf{\Theta}$ , vienen dados por

$$\tilde{\mathbf{B}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1}, \quad (1.4.3)$$

$$\tilde{\mathbf{\Gamma}} = \frac{1}{n} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{S} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (1.4.4)$$

y

$$\tilde{\mathbf{\Theta}} = \frac{1}{n} (\mathbf{X}^{oT} \mathbf{X}^o)^{-1} \mathbf{X}^{oT} \mathbf{Y} \mathbf{Y}^T \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{X}^o)^{-1},$$

respectivamente, donde  $\mathbf{S}$  está definida en (1.3.5).

Además, el EMV de  $\mathbf{\Sigma}$  es

$$\tilde{\mathbf{\Sigma}} = \mathbf{X} \tilde{\mathbf{\Gamma}} \mathbf{X}^T + \mathbf{X}^o \tilde{\mathbf{\Theta}} \mathbf{X}^{oT}. \quad (1.4.5)$$

Como se puede observar, contrariamente a lo que ocurre para el MCC sin estructura de covarianza, cuando se utiliza la estructura simple de Rao el EMV  $\tilde{\mathbf{B}}$ , dado en (1.4.3), es una transformación lineal de la variable respuesta  $\mathbf{Y}$ . Además, comparando los dos estimadores, se observa que la única diferencia reside en el hecho de que el EMV  $\tilde{\mathbf{B}}$ , definido en (1.3.7), depende de  $\mathbf{S}$ . Por otro lado, se verifica que el GLSE (1.3.1) de  $\mathbf{B}$  para el MCC coincide con el EMV de  $\mathbf{B}$  obtenido para el MCC(SCS). Kariya (1985) demuestra que la estructura SCS además de ser una condición suficiente es también necesaria para que los EMV y los GLSE de  $\mathbf{B}$  y  $\mathbf{\Sigma}$  sean idénticos, tal y como se recoge en el siguiente teorema.

**Teorema 1.4.2** Para el MCC con rango total los EMV coinciden con los GLSE si y solo si la matriz de covariancias  $\mathbf{\Sigma}$  tiene una estructura de covariancia simple de Rao.

En estos términos, en el MCC el EMV y el GLSE para  $\mathbf{B}$  coinciden si y sólo si  $\mathbf{\Sigma}$  pertenece a la clase de matrices definidas por la estructura SCS. Por tanto, en tal caso, la obtención del EMV de  $\mathbf{B}$  no depende del estimador de  $\mathbf{\Sigma}$ . Así, la estructura SCS tiene un papel especial en el MCC en el sentido de que dentro de este espacio de covarianzas el EMV de la matriz de coeficientes de regresión del MCC no depende de la elección de la estructura de covarianza, lo que facilita la inferencia estadística. En este sentido, puede ser de interés conocer las condiciones para las cuales la estructura de covarianza pertenece al espacio de la estructura SCS (véase, por ejemplo, Srivastava y von Rosen, 1999).

Bajo la hipótesis distribucional  $\mathbf{Y} \sim N_{p,n}(\mathbf{X} \mathbf{B} \mathbf{Z}; \mathbf{\Sigma}, \mathbf{I}_n)$ , dado que el EMV  $\tilde{\mathbf{B}}$  es una transformación lineal del vector de variable respuesta, se obtiene que

$$\tilde{\mathbf{B}} \sim N_{q,r}(\mathbf{B}; (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{\Sigma} \mathbf{X}) (\mathbf{X}^T \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$$

(véase el teorema 1.3.2).

Además, teniendo en cuenta la estructura SCS y dado que  $\mathbf{X}^{oT} \mathbf{X} = \mathbf{0}$ ,

$$\begin{aligned} \mathbf{\Sigma} &= \mathbf{X} \mathbf{\Gamma} \mathbf{X}^T + \mathbf{X}^o \mathbf{\Theta} \mathbf{X}^{oT} \Leftrightarrow \mathbf{X}^T \mathbf{\Sigma} \mathbf{X} = \mathbf{X}^T \mathbf{X} \mathbf{\Gamma} \mathbf{X}^T \mathbf{X} \Leftrightarrow \\ &\Leftrightarrow \mathbf{\Gamma} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{\Sigma} \mathbf{X}) (\mathbf{X}^T \mathbf{X})^{-1}, \end{aligned}$$

concluyéndose que

$$\tilde{\mathbf{B}} \sim N_{q,r}(\mathbf{B}; \mathbf{\Gamma}, (\mathbf{Z}\mathbf{Z}^T)^{-1}). \quad (1.4.6)$$

Y, por tanto, como  $\text{Var}[\tilde{\mathbf{B}}]$  depende de  $\mathbf{\Sigma}$ , considerando el EMV  $\tilde{\mathbf{\Sigma}}$  dado en (1.4.5), se obtiene el siguiente estimador

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}] = (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}. \quad (1.4.7)$$

Por otro lado, se demuestra que

$$n\tilde{\mathbf{\Gamma}} \sim W_q(n-r, \mathbf{\Gamma})$$

y

$$n\hat{\mathbf{\Theta}} \sim W_{p-q}(n, \mathbf{\Theta}).$$

Lo que indica que los EMV  $\tilde{\mathbf{B}}$  y  $\tilde{\mathbf{\Theta}}$  son estimadores insesgados de  $\mathbf{B}$  y  $\mathbf{\Theta}$ , respectivamente, y que el EMV  $\tilde{\mathbf{\Gamma}}$  es un estimador asintóticamente insesgado de  $\mathbf{\Gamma}$ . Para una demostración de estos resultados véase Pan y Fang (2002).

---

## Capítulo 2

# Medidas de influencia propuestas en la literatura

### 2.1. Introducción

Los modelos estadísticos son, de un modo general, descripciones aproximadas de fenómenos naturales más complejos. El objetivo de todo análisis estadístico es obtener conclusiones fiables a partir de los datos resultantes de una experimentación. Por tanto, la fiabilidad de las observaciones del proceso es de especial interés, ya que la eficacia del modelo estadístico y las técnicas estadísticas que se apliquen pueden verse fuertemente afectadas por una o algunas de las observaciones realizadas. Este problema ha motivado el desarrollo de métodos enfocados bien al estudio de nuevas técnicas que no se vean influenciadas excesivamente por la modelización del fenómeno natural (estadística robusta), bien al análisis de la calidad de los datos o bien al estudio de aquellas observaciones que afectan considerablemente a los resultados del análisis. En este último enfoque se han propuesto un conjunto de métodos englobados en lo que genéricamente se conoce como el análisis de influencia, incluidos en las técnicas de diagnóstico.

Las técnicas de diagnóstico tienen como objetivo detectar outliers que se desvían del modelo postulado, identificar observaciones influyentes que tengan grandes efectos en la inferencia estadística hecha a partir del modelo postulado y validar el modelo estadístico escogido. En general, no hay una relación entre outliers y observaciones influyentes, aunque un outlier puede o no considerarse una observación influyente y viceversa (véase, por ejemplo, Chatterjee y Hadi, 1988).

En general, el estudio de la influencia se ha desarrollado principalmente en el análisis de regresión lineal y a partir del trabajo de Cook (1977). De forma general, se puede decir que una observación debe considerarse como influyente si, bien de forma individual o bien conjuntamente con otras, tiene un mayor impacto que el resto de las observaciones sobre los valores de varias estimaciones y/o estadísticos. Por otro lado, los outliers son observaciones que no siguen el patrón de la mayoría de los datos (observaciones aberrantes) (Rousseeuw y van Zomeren, 1990). Se debe tener en cuenta que no siempre los outliers y/o las observaciones influyentes son necesariamente malos datos, pueden contener información más interesante y, también, pueden ser útiles para validar la elección del modelo estadístico.

La gran mayoría de las técnicas propuestas para el análisis de influencia están basadas en la aproximación genérica recogida en Cook y Weisberg (1982). Para medir el efecto que, sobre

un aspecto de interés del análisis, tiene una observación o un conjunto de ellas se introducen perturbaciones, que las afectan de alguna manera, y se cuantifica el cambio producido. Es decir, las técnicas surgen de conjugar adecuadamente perturbación del modelo y comparación de resultados. Cook y Weisberg (1982) recogen textualmente: “La idea básica en el análisis de influencia es muy simple. Introducimos pequeñas perturbaciones en la formulación del problema y entonces calculamos cuánto cambian los resultados del análisis por la perturbación.”

Cook (1987) trata de unificar el problema bajo la siguiente formulación general:

”Sea un conjunto de datos  $D$ , un modelo  $M$  postulado a priori, un resultado  $R(D, M)$  seleccionado de una síntesis de los datos y el modelo, y sea  $\omega$  un vector de perturbaciones pertenecientes a un conjunto  $\Omega$  de perturbaciones relevantes, siendo  $M(\omega)$  el modelo perturbado, de forma que  $\exists \omega_0 \in \Omega : M \approx M(\omega_0)$ . Así, el análisis de influencia consiste en la comparación de los resultados  $R(D, M)$  y  $R(D, M(\omega))$ .”

Esta formulación general ha servido como punto de partida a una gran cantidad de técnicas estadísticas dirigidas tanto al análisis de influencia, como a la estadística robusta. Cuestiones importantes en los métodos del análisis de influencia son: la elección del esquema de perturbación, el aspecto específico del análisis a evaluar y el método de medición. Las diferentes respuestas a estas cuestiones pueden originar diferentes diagnósticos.

Los métodos de identificación de observaciones influyentes se pueden clasificar en dos categorías: influencia global e influencia local. Los procedimientos usados en la influencia global se basan en la llamada metodología de omisión de casos, cuya idea principal es eliminar un conjunto de observaciones cuya influencia se pretende estudiar y elegir una métrica apropiada para medir las variaciones producidas en el modelo, en un estimador, etc. Este método es el más popular y suele ser usado en una gran variedad de áreas (Cook y Weisberg, 1982 y Chatterjee y Hadi, 1988), aunque en la práctica pueden surgir algunas interrogantes para su aplicación (Pan y Fang, 2002), por ejemplo:

- ¿Cómo decidir el número de posibles outliers y/o observaciones influencia?
- ¿Cómo se elige el subconjunto de observaciones a eliminar, una vez que el número de observaciones ha sido fijado?

La primera cuestión está relacionada con el fenómeno de enmascaramiento, unas veces son considerados outliers y/o observaciones influencia algunas observaciones que no lo son y, en otros casos, se detectan de menos. Usualmente, el problema de enmascaramiento surge cuando en la muestra existe un grupo de observaciones cuya influencia conjunta oculta el efecto individual de cada una de ellas, provocando que este efecto no sea detectado mediante el uso de técnicas que analizan una sola observación.

La segunda cuestión, tiene especial importancia para datos correlacionados, por ejemplo, para datos longitudinales, ya que usualmente en su estudio se presentan dificultades de cálculo (véase Barnett y Lewis, 1984).

Por otro lado, la influencia local mide el efecto del esquema de perturbación en el modelo ajustado usando una curvatura geométrica de alguna métrica apropiada tal como el desplazamiento de verosimilitud propuesto por Cook (1986) y la divergencia de Kullback-Leibler sugerida



por Box y Tiao (1968). De una manera general, como la influencia global puede sufrir el efecto de enmascaramiento, la influencia local puede complementar la información sobre observaciones influyentes proporcionada por la influencia global.

El análisis de observaciones influencia ha sido ampliamente estudiado en las tres últimas décadas para el modelo de regresión, sin embargo, para el modelo de curvas de crecimiento (MCC) y para el MCC con estructura simple de Rao (MCC(SCS)) las investigaciones sobre este tópico son más recientes. Entre ellas, se pueden citar:

- Pan (1994), von Rosen (1995a) y Pan y Fang (1996) que abordan la influencia global en el MCC dentro del contexto de la verosimilitud. Pan (2002) estudia el caso particular del MCC(SCS).
- La influencia local es analizada por diferentes autores, entre ellos, Pan et al. (1997) en el MCC, y Pan y Bai (2003) y Qingming y Huaixiong (2005) en el MCC(SCS).
- En el contexto bayesiano, Pan y Fung (2000) estudian la influencia global en el MCC, y Pan et al. (1999) y Peng y Yu (2000) la influencia local. Para el MCC(SCS), entre otros, Pan et al. (1994).
- Otros autores que abordan el estudio del análisis de influencia en el MCC son Kish y Chinchilli (1990), Liski (1991), Walker (1992) y Pan (1995).

En el libro Pan y Fang (2002) se puede encontrar una revisión general del análisis de influencia desde diferentes enfoques, tanto en el MCC como en el MCC(SCS).

Por último, se debe tener en cuenta que el énfasis del análisis de diagnóstico reside solamente en la detección de outliers y/o observaciones influyentes y no en la manera de como deben tratarse dichas observaciones una vez descubiertas, ya que las consideraciones finales dependen necesariamente de los contextos específicos, por lo que hacer una recomendación global es en general imposible. Como indica Hadi (1992), los diagnósticos de influencia son designados para la detección de observaciones cuya influencia sea (parezca) mayor que las restantes en un conjunto de datos. No son designados como tests de hipótesis formales, así, los valores de la influencia de una observación deberán ser comparados con los de las restantes.

El objetivo de este capítulo es hacer una revisión general de diferentes medidas de influencia propuestas en la literatura para el estimador de máxima verosimilitud (EMV) de la matriz de coeficientes de regresión en el MCC.

Dado que algunas de las medidas de influencia que se proponen en esta memoria están basadas en residuos desarrollados por von Rosen (1995b), en la sección 2.2, se realiza un breve estudio de los mismos en el MCC.

A continuación, en la sección 2.3, se plantea el estudio del MCC bajo la omisión de un conjunto de observaciones.

En la sección 2.4, se presentan algunas de las medidas de influencia propuestas en la literatura para el EMV de la matriz de coeficientes de regresión del MCC. Dichas medidas están basadas en la distancia de Cook generalizada y en el volumen del elipsoide de confianza. También, se recogen medidas de Cook para combinaciones lineales del citado estimador. Para más detalles, véase, por ejemplo, Pan y Fang (2002).

Se finaliza este capítulo con un estudio similar, al realizado anteriormente, en el MCC(SCS).

---

## 2.2. Residuos en el modelo de curvas de crecimiento

En general, el análisis de los residuos es una de las estrategias básicas en la validación de las hipótesis del modelo. A través del examen detallado de los mismos se extraen conclusiones sobre la adecuación del modelo considerado. Para los modelos lineales univariantes este tópico ha sido ampliamente analizado y la mayoría de las técnicas de diagnósticos están basadas en los residuos.

Los trabajos dedicados al estudio de los residuos en los modelos multivariantes son más limitados, en particular en el MCC. Uno de los primeros trabajos sobre ellos en el MCC fue publicado por von Rosen (1995b), en el que se interpretan, se proporcionan algunas observaciones sobre la información que se puede extraer de éstos y se establecen algunas propiedades básicas.

En el modelo lineal univariante y en el modelo lineal multivariante,  $\mathbf{Y} = \mathbf{BZ} + \boldsymbol{\varepsilon}$ , los residuos se obtienen proyectando los valores observados  $\mathbf{Y}$  sobre el complemento ortogonal del espacio generado por las columnas de la matriz de diseño  $\mathbf{Z}^T$ , es decir, sobre el espacio  $\mathcal{C}(\mathbf{Z}^T)^\perp$ . Así, la matriz de residuos es  $\mathbf{R} = \mathbf{Y}(\mathbf{I}_n - \mathbf{Z}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}) = \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})$ . Como el MCC posee una estructura bilineal, existen dos espacios de interés: el espacio columna de  $\mathbf{Z}^T$ , es decir,  $\mathcal{C}(\mathbf{Z}^T)$ , y el espacio columna de la matriz  $\mathbf{X}$  con el producto interno definido por  $\langle x, y \rangle = x^T \mathbf{S}^{-1} y$ , que se denotará por  $\mathcal{C}_{\mathbf{S}}(\mathbf{X})$ , donde  $\mathbf{S} = \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T})\mathbf{Y}^T$ . La matriz proyección sobre  $\mathcal{C}_{\mathbf{S}}(\mathbf{X})$  viene dada por  $\mathbf{X}(\mathbf{X}^T \mathbf{S}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{S}^{-1} = \mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1}$ . En consecuencia, de (1.3.15) se deduce que en el MCC la matriz de valores ajustados,  $\widehat{\mathbf{Y}} = \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z}$ , obtenida a través de la estimación de máxima verosimilitud, es la proyección ortogonal de la matriz de valores observados  $\mathbf{Y}$  sobre el espacio  $\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X})$  (definición B.0.6). Para ello basta comprobar que

$$\begin{aligned} \text{vec}(\widehat{\mathbf{Y}}) &= \text{vec}(\mathbf{X} \widehat{\mathbf{B}} \mathbf{Z}) = \text{vec}(\mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} \mathbf{P}_{\mathbf{Z}^T}) = \\ &= [\mathbf{P}_{\mathbf{Z}^T} \otimes \mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1}] \text{vec}(\mathbf{Y}). \end{aligned}$$

Así, los residuos se pueden obtener a través de la proyección de los valores observados  $\mathbf{Y}$  sobre el complemento ortogonal del espacio  $\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X})$ , es decir, sobre  $(\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X}))^\perp$ ,

$$\text{vec}(\mathbf{Y} - \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z}) = [\mathbf{I}_{np} - \mathbf{P}_{\mathbf{Z}^T} \otimes \mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1}] \text{vec}(\mathbf{Y}).$$

Para el estudio de los residuos, von Rosen (1995b) utiliza la descomposición del espacio  $(\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X}))^\perp$  en la suma ortogonal, dada en la definición B.0.7, de tres subespacios ortogonales,

$$(\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X}))^\perp = [\mathcal{C}(\mathbf{Z}^T)^\perp \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X})^\perp] \boxplus [\mathcal{C}(\mathbf{Z}^T)^\perp \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X})] \boxplus [\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_{\mathbf{S}}(\mathbf{X})^\perp].$$

A partir de ésta, propone la siguiente descomposición de los residuos,

$$\mathbf{R} = \mathbf{Y} - \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z} = \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3, \quad (2.2.1)$$

donde

$$\begin{aligned} \mathbf{R}_1 &= \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T}) - \mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} + \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z}, \\ \mathbf{R}_2 &= \mathbf{X} \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} - \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z} \quad \text{y} \\ \mathbf{R}_3 &= \mathbf{Y} \mathbf{P}_{\mathbf{Z}^T} - \mathbf{X} \widehat{\mathbf{B}} \mathbf{Z}, \end{aligned}$$

de forma que  $\mathbf{R}_1$  se obtiene a partir del subespacio  $\mathcal{C}(\mathbf{Z}^T)^\perp \otimes \mathcal{C}_S(\mathbf{X})^\perp$ ,  $\mathbf{R}_2$  del subespacio  $\mathcal{C}(\mathbf{Z}^T)^\perp \otimes \mathcal{C}_S(\mathbf{X})$  y  $\mathbf{R}_3$  del subespacio  $\mathcal{C}(\mathbf{Z}^T) \otimes \mathcal{C}_S(\mathbf{X})^\perp$ .

La descomposición (2.2.1) puede mostrar diversos aspectos útiles para validar el análisis del modelo. Por tanto, es importante examinar cada una de las tres componentes residuales para una mejor comprensión de la estructura del modelo. En particular, si se desean obtener diagnósticos basados en los residuos del modelo, se deberían investigar las componentes para determinar una información más completa de los mismos.

Como  $\mathcal{C}_S(\mathbf{X})$  es el espacio vectorial que representa la estructura "dentro de los individuos" y  $\mathcal{C}(\mathbf{Z}^T)$  representa la estructura "entre los individuos", von Rosen (1995b) propone la siguiente interpretación de los residuos:

- $\mathbf{R}_1$  representa la proyección de la diferencia entre las observaciones y la media sobre  $\mathcal{C}_S(\mathbf{X})^\perp$ .
- $\mathbf{R}_2$  representa la proyección de la diferencia entre las observaciones y la media sobre  $\mathcal{C}_S(\mathbf{X})$ .
- $\mathbf{R}_3$  representa la proyección de la diferencia entre las observaciones y la media sobre  $\mathcal{C}(\mathbf{Z}^T)$ .

Esto significa que  $\mathbf{R}_1$  y  $\mathbf{R}_3$  reflejan las hipótesis del modelo dentro de los individuos, mientras que  $\mathbf{R}_1$  y  $\mathbf{R}_2$  pueden ser utilizadas para investigar los supuestos del modelo entre los individuos.

Así, debido a la estructura bilineal del modelo, los residuos ordinarios constan de dos partes. Una parte proporciona información sobre la estructura entre los individuos, mientras que la otra proporciona información sobre la estructura dentro de los individuos. De hecho,

$$\mathbf{E} = \mathbf{R}_1 + \mathbf{R}_2 = \mathbf{Y}(\mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T}) \quad (2.2.2)$$

representa la diferencia entre las observaciones y los valores promedios en cada grupo,  $\mathbf{Y}\mathbf{P}_{\mathbf{Z}^T}$ , es decir, proporciona información sobre las desviaciones de los individuos respecto del comportamiento medio del grupo ("dentro del grupo").

Por otro lado,

$$\mathbf{R}_3 = \mathbf{Y}\mathbf{P}_{\mathbf{Z}^T} - \mathbf{X}\hat{\mathbf{B}}\mathbf{Z}$$

refleja la diferencia entre la estructura promedio observada y los valores ajustados por el modelo (estructura promedio estimada), es decir, la proporción de residuos que puede ser utilizada para comprobar si el modelo estimado se ajusta a los datos.

Además, se observa que  $\mathbf{R}_3 = \hat{\mathbf{V}}$  dado en (1.3.9) y, por tanto, un EMV de  $\Sigma$  es

$$n\hat{\Sigma} = \mathbf{S} + \mathbf{R}_3\mathbf{R}_3^T.$$

Por todo ello, von Rosen (1995b) recomienda examinar las componentes de los residuos ordinarios, obteniéndolos por separado, ya que puede ocurrir que las dos partes antes mencionadas se estén cancelando una a la otra y, por tanto, los residuos ordinarios pueden tener valores cercanos a cero, dando la impresión de que el modelo se ajusta bien a los datos.

Hamid y von Rosen (2006) aportan aspectos adicionales a la interpretación de los residuos en el MCC, desarrollando un método para establecer los residuos en el MCC extendido. Cada

uno de los residuos propuestos contiene información relevante sobre la estructura del modelo bilineal y pueden ser utilizados para validar las hipótesis del modelo y detectar observaciones influyentes y/o valores atípicos. Sin embargo, no quedan determinados los valores adecuados de tales residuos para concluir que las hipótesis subyacentes son válidas, ni para clasificar a las observaciones como influyentes u outliers. Por otro lado, se necesitan herramientas de diagnóstico basadas en estos residuos razonables para validar el modelo, así como para detectar este tipo de observaciones.

### 2.3. Modelo de curvas de crecimiento perturbado bajo la omisión

Como se ha comentado anteriormente, los procedimientos usados en la influencia global están basados en la llamada metodología de omisión de casos u observaciones (case-deletion), que es el esquema de perturbación más utilizado en el diagnóstico estadístico. Uno de los tópicos principales de la influencia global es escoger una métrica apropiada para medir la diferencia entre la inferencia resultante del conjunto de datos completo y del conjunto de datos reducido que resulta de la omisión de las observaciones cuya influencia se pretende estudiar (el modelo perturbado bajo omisión). A continuación, se hace una breve referencia a dicha aproximación y se presenta el MCC perturbado. Para más información sobre este modelo véase, por ejemplo, Pan y Fang (2002).

En el contexto de la influencia global, para medir la influencia de un subconjunto de observaciones sobre alguna característica del MCC, definido en (1.2.5), es necesario considerar el modelo anterior bajo la omisión de dichas observaciones. Para tal fin, sea  $I = \{i_1, i_2, \dots, i_k\} \subset \{1, 2, \dots, n\}$ ,  $n > p + k$ , un conjunto de subíndices que contiene los  $k$  individuos que van a ser eliminados, donde  $k$  está previamente fijado. Sin pérdida de generalidad, el conjunto indexado puede ser  $I = \{n - k + 1, n - k + 2, \dots, n\}$  de tal forma que las columnas de las matrices  $\mathbf{Y}$ ,  $\mathbf{Z}$  y  $\mathbf{E}$  (residuo definido en (2.2.2)) son reordenadas de modo que puedan obtenerse las particiones  $\mathbf{Y} = (\mathbf{Y}_{(I)} \dot{\vdash} \mathbf{Y}_I)$ , donde  $\mathbf{Y}_{(I)} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n-k})$  (observaciones no eliminadas) e  $\mathbf{Y}_I = (\mathbf{y}_{n-k+1}, \mathbf{y}_{n-k+2}, \dots, \mathbf{y}_n)$  (observaciones eliminadas a las cuales se les pretende medir su influencia),  $\mathbf{Z} = (\mathbf{Z}_{(I)} \dot{\vdash} \mathbf{Z}_I)$  y  $\mathbf{E} = (\mathbf{E}_{(I)} \dot{\vdash} \mathbf{E}_I)$ , respectivamente. En estos términos, el MCC bajo la omisión del conjunto de observaciones subindicadas por  $I$ , que se denotará por  $\text{MCC}(I)$ , vendrá dado por

$$\begin{cases} \mathbf{Y}_{(I)} = \mathbf{X} \mathbf{B} \mathbf{Z}_{(I)} + \mathcal{E}_{(I)} \\ \begin{matrix} (p \times (n-k)) & (p \times q)(q \times r) & (r \times (n-k)) & (p \times (n-k)) \end{matrix} \end{cases}, \quad (2.3.1)$$

$$\mathcal{E}_{(I)} \sim N_{p, (n-k)}(\mathbf{0}; \mathbf{\Sigma}, \mathbf{I}_{n-k})$$

donde  $\mathbf{Y}_{(I)}$ ,  $\mathbf{Z}_{(I)}$  y  $\mathcal{E}_{(I)}$  son las submatrices de  $\mathbf{Y}$ ,  $\mathbf{Z}$  y  $\mathcal{E}$ , respectivamente, obtenidas eliminando las columnas subindicadas por  $I$ . En el  $\text{MCC}(I)$ , se mantiene la misma estructura de la matriz de covarianzas que en el MCC (Pan y Fang, 2002).

Con el objetivo de estudiar la influencia de un subconjunto de observaciones  $\mathbf{Y}_I$ , del total de observaciones  $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$ , sobre el EMV  $\hat{\mathbf{B}}$  de la matriz de coeficientes de regresión  $\mathbf{B}$ , se denota por  $\hat{\mathbf{B}}_{(I)}$  el estimador de  $\mathbf{B}$ , obtenido mediante el mismo procedimiento que  $\hat{\mathbf{B}}$  pero utilizando solamente las observaciones de  $\mathbf{Y}_{(I)}$ .

Así, de acuerdo con el teorema 1.3.3, los EMV de los parámetros  $\mathbf{B}_{(I)}$  y  $\mathbf{\Sigma}_{(I)}$  para el  $\text{MCC}(I)$  son

$$\widehat{\mathbf{B}}_{(I)} = (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{Y}_{(I)} \mathbf{Z}_{(I)}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1}$$

y

$$\widehat{\boldsymbol{\Sigma}}_{(I)} = \frac{1}{n-k} \left( \mathbf{S}_{(I)} + \mathbf{X}_{\mathbf{S}_{(I)}}^o \mathbf{Y}_{(I)} \mathbf{P}_{\mathbf{Z}_{(I)}^T} \mathbf{Y}_{(I)}^T \mathbf{X}_{\mathbf{S}_{(I)}}^{oT} \right),$$

respectivamente, con

$$\mathbf{S}_{(I)} = \mathbf{Y}_{(I)} \left( \mathbf{I}_{n-k} - \mathbf{P}_{\mathbf{Z}_{(I)}^T} \right) \mathbf{Y}_{(I)}^T, \quad (2.3.2)$$

la correspondiente a  $\mathbf{S}$  para el MCC( $I$ ), y

$$\mathbf{P}_{\mathbf{Z}_{(I)}^T} = \mathbf{Z}_{(I)}^T \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \mathbf{Z}_{(I)}.$$

Además, de (1.3.16), se obtiene que

$$\mathbb{E}[\widehat{\mathbf{B}}_{(I)}] = \mathbf{B}$$

y, de (1.3.17),

$$\text{Var}[\widehat{\mathbf{B}}_{(I)}] = c_6 \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1},$$

con

$$c_6 = \frac{n-k-r-1}{n-k-r-(p-q)-1}, \quad (2.3.3)$$

y, considerando  $\widehat{\boldsymbol{\Sigma}}_{(I)} = \frac{1}{n-r-k} \mathbf{S}_{(I)}$ , un estimador consistente de  $\text{Var}[\widehat{\mathbf{B}}_{(I)}]$  vendrá dado por

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] = \frac{c_6}{n-k-r} \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1}. \quad (2.3.4)$$

Las diferencias existentes entre  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , usualmente, se miden a través de una función  $f(\widehat{\mathbf{B}}, \widehat{\mathbf{B}}_{(I)})$ , que deberá de ser capaz de cuantificar la sensibilidad del subconjunto de observaciones candidatas a ser consideradas influyentes. Estas funciones son conocidas, generalmente, como medidas de diagnóstico o medidas de influencia, la elección de las mismas dependerá de la forma de las observaciones bajo estudio y del método de estimación utilizado.

Para obtener diferentes medidas de influencia, basadas en la omisión, se necesita saber las relaciones entre los EMV de los parámetros del MCC y del MCC( $I$ ), por lo que, previamente, hay que cuantificar las relaciones entre las matrices  $\mathbf{S}$  y  $\mathbf{S}_{(I)}$  definidas en (1.3.5) y (2.3.2), respectivamente. Teniendo en cuenta el teorema C.4.2, como  $\mathbf{S}_{(I)} \sim W_p(n-r-k, \boldsymbol{\Sigma})$ , cuando  $n > r+p+k$ , el siguiente resultado proporciona una relación entre ambas matrices, así como para sus inversas.

**Lema 2.3.1** *Para el MCC*

$$\mathbf{S}_{(I)} = \mathbf{S} - \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T$$

y

$$\mathbf{S}_{(I)}^{-1} = \mathbf{S}^{-1} + \mathbf{S}^{-1} \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)^{-1} \mathbf{E}_I^T \mathbf{S}^{-1}, \quad (2.3.5)$$

donde

$$\mathbf{E}_I = \mathbf{Y}_I - \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I \quad (2.3.6)$$

y

$$\mathbf{H}_I = \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I. \quad (2.3.7)$$

**Nota 2.3.1** La existencia de  $(\mathbf{I}_k - \mathbf{H}_I)^{-1}$  está garantizada pues en el análisis de influencia sobre el MCC no tiene sentido estudiar la influencia conjunta de todas las observaciones que conforman uno de los grupos incluidos en la experimentación.

El siguiente teorema proporciona una relación entre los EMV  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , del MCC y del MCC(I), respectivamente, su demostración se puede ver en Pan y Fang (2002). Este resultado jugará un papel fundamental en el análisis de observaciones influyentes.

**Teorema 2.3.1** Sean  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$  los EMV del parámetro  $\mathbf{B}$ , basados en MCC y MCC(I), respectivamente, entonces se verifica

$$\widehat{\mathbf{B}}_{(I)} = \widehat{\mathbf{B}} - \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1}, \quad (2.3.8)$$

donde

$$\mathbf{K}_I = \mathbf{Z}_I - \mathbf{Z} \mathbf{Y}^T \mathbf{S}^{-1} \mathbf{E}_I + \mathbf{Z} \mathbf{Y}^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \quad (2.3.9)$$

y

$$\mathbf{V}_I = \mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I + \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I. \quad (2.3.10)$$

**Nota 2.3.2** Considerando la relación (1.3.14), se obtienen las expresiones alternativas

$$\mathbf{K}_I = \mathbf{Z}_I - \mathbf{Z} \mathbf{Y}^T \mathbf{X}_S^{oT} \mathbf{S}^{-1} \mathbf{X}_S^o \mathbf{E}_I \quad y \quad \mathbf{V}_I = \mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{X}_S^{oT} \mathbf{S}^{-1} \mathbf{X}_S^o \mathbf{E}_I,$$

donde  $\mathbf{X}_S^o$  está definida en (1.3.12).

A continuación se presenta la relación (2.3.8) cuando se elimina una única observación.

**Corolario 2.3.1** En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , entonces

$$\widehat{\mathbf{B}}_{(i)} = \widehat{\mathbf{B}} - \frac{1}{v_{ii}} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1}, \quad (2.3.11)$$

donde

$$\mathbf{k}_i = \mathbf{z}_i - \mathbf{Z} \mathbf{Y}^T \mathbf{S}^{-1} \mathbf{e}_i + \mathbf{Z} \mathbf{Y}^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i, \quad (2.3.12)$$

es la columna  $i$ -ésima de la matriz  $\mathbf{K} = \mathbf{Z} - \mathbf{Z} \mathbf{Y}^T \mathbf{X}_S^{oT} \mathbf{S}^{-1} \mathbf{X}_S^o \mathbf{E}$ , y

$$v_{ii} = 1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i + \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \quad (2.3.13)$$

el elemento  $(i, i)$  de la matriz  $\mathbf{V} = \mathbf{I}_n - \mathbf{P}_{\mathbf{Z}^T} - \mathbf{E}^T \mathbf{X}_S^{oT} \mathbf{S}^{-1} \mathbf{X}_S^o \mathbf{E}$ , con

$$\mathbf{e}_i = \mathbf{y}_i - \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i \quad (2.3.14)$$

la columna  $i$ -ésima del residuo  $\mathbf{E}$ , definido por (2.2.2), y

$$p_{ii} = \mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i \quad (2.3.15)$$

es el elemento  $i$ -ésimo diagonal de la matriz de proyección  $\mathbf{P}_{\mathbf{Z}^T}$ , definida por (1.3.3), donde  $\mathbf{z}_i$  es la columna  $i$ -ésima de  $\mathbf{Z}$ .

## 2.4. Análisis de influencia en el modelo de curvas de crecimiento

En este apartado se recogen diferentes medidas para el estudio de la calidad de los datos en el modelo de curvas de crecimiento. Basadas en el tópicó de omisión de casos, se recoge un resumen de algunas de las medidas propuestas en la literatura para analizar la influencia de un subconjunto de observaciones  $\mathbf{Y}_I$  en el EMV  $\widehat{\mathbf{B}}$  de la matriz de coeficientes de regresión del MCC. En particular, basada en la relación entre los EMV  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , se recoge la distancia de Cook generalizada y, a partir de ella, la medida de Cook y la medida de Welsch-Kuh.

En primer lugar, se recogen dos estadísticos importantes en el desarrollo de dichas medidas, los mismos permiten verificar si el subconjunto  $\mathbf{Y}_I$  de las observaciones eliminadas está compuesto por observaciones outliers. Para el MCC, en Pan y Fang (2002), se considera el siguiente estadístico

$$\begin{aligned} T_I &= \det[\mathbf{I}_k + \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)^{-1}] = \\ &= \frac{\det[\mathbf{V}_I]}{\det[\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I]}. \end{aligned}$$

El subconjunto  $\mathbf{Y}_I$  de las observaciones omitidas es considerado como un conjunto outlier si  $T_I$  es significativamente grande o, de forma equivalente,  $\Lambda_I = T_I^{-1}$  es suficientemente pequeño. El punto crítico, para  $T_I$ , puede ser obtenido utilizando el siguiente resultado

$$\Lambda_I \sim \Lambda(q, n - k - r - p + q, k), \quad (2.4.1)$$

es decir,  $\Lambda_I$  sigue una distribución de Wilk con parámetros  $q$ ,  $n - k - r - p + q$  y  $k$ .

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , se obtiene el estadístico

$$T_i = \Lambda_i^{-1} = 1 + \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i} = \frac{v_{ii}}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i}, \quad (2.4.2)$$

y el individuo  $i$ -ésimo,  $\mathbf{y}_i$ , será considerado una observación outlier si

$$T_i \geq 1 + \frac{qC_\alpha^*}{n - r - p} \quad (2.4.3)$$

donde  $C_\alpha^*$  es el percentil de orden  $100(1 - \alpha)\%$  de la distribución  $F_{q, n-r-p}$ .

En general, las medidas de influencia más utilizadas en la regresión están basadas en la distancia de Cook (Cook y Weisberg, 1982), la cuál está fuertemente relacionada con la función de influencia muestral propuesta en Hampel et al. (1986). Basadas en dicha distancia, a continuación, se recogen dos medidas de influencia para el EMV de la matriz de coeficientes de regresión del MCC.

Para un estimador  $\widehat{\boldsymbol{\Phi}}$ , se define la distancia de Cook generalizada

$$D_I(\mathbf{M}, \mathbf{C}) = (\widehat{\boldsymbol{\Phi}} - \widehat{\boldsymbol{\Phi}}_{(I)}) \mathbf{M} (\widehat{\boldsymbol{\Phi}} - \widehat{\boldsymbol{\Phi}}_{(I)})^T \mathbf{C}^{-1},$$

donde  $\widehat{\boldsymbol{\Phi}}_{(I)}$  es el estimador obtenido bajo la omisión de las observaciones subindicadas por  $I$ ,  $\mathbf{M}$  y  $\mathbf{C}$  son matrices constantes definidas positivas, denominadas matrices de los “pesos”.

Las diferentes elecciones de dichas matrices conducen, generalmente, a diferentes medidas que pueden ser utilizadas para evaluar la influencia del subconjunto  $\mathbf{Y}_I$  en el estimador  $\hat{\Phi}$ .

En el MCC, se obtiene la siguiente distancia de Cook generalizada

$$D_I(\mathbf{M}, \mathbf{C}) = \text{tr} \left[ (\hat{\mathbf{B}} - \hat{\mathbf{B}}_{(I)}) \mathbf{M} (\hat{\mathbf{B}} - \hat{\mathbf{B}}_{(I)})^T \mathbf{C}^{-1} \right], \quad (2.4.4)$$

donde  $\hat{\mathbf{B}}$  y  $\hat{\mathbf{B}}_{(I)}$  son los EMV de  $\mathbf{B}$  basados en MCC y MCC(I), respectivamente,  $\mathbf{M}_{(r \times r)} > \mathbf{0}$  y  $\mathbf{C}_{(q \times q)} > \mathbf{0}$  (Pan y Fang, 2002).

Además, teniendo en cuenta el teorema 2.3.1, la distancia anterior viene expresada por

$$D_I(\mathbf{M}, \mathbf{C}) = \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{M} (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \right. \\ \left. \mathbf{X} \mathbf{S}_X \mathbf{C}^{-1} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \right]. \quad (2.4.5)$$

Un valor grande de  $D_I(\mathbf{M}, \mathbf{C})$  indica que el subconjunto de observaciones  $\mathbf{Y}_I$  ejerce una gran influencia sobre  $\hat{\mathbf{B}}$ , respecto a  $\mathbf{M}$  y  $\mathbf{C}$ . Por otro lado, es obvio que la magnitud de  $D_I(\mathbf{M}, \mathbf{C})$  está determinada por el "peso" de las matrices  $\mathbf{M}$  y  $\mathbf{C}$ , que deben de ser elegidas en relación al estudio que se desee realizar.

A continuación, se recoge la distancia  $D_I(\mathbf{M}, \mathbf{C})$  para dos elecciones usuales de las matrices de los pesos, que dan origen a las medidas de influencia conocidas como medida de Cook y medida de Welsch-Kuh.

- **Medida Cook.** Considerando  $\mathbf{M} = \mathbf{Z}\mathbf{Z}^T$  y  $\mathbf{C} = qr\mathbf{S}_X$  en (2.4.5), se obtiene la medida de influencia de Cook, la cuál analiza la influencia de  $\mathbf{Y}_I$  en el centro del elipsoide de confianza de  $\mathbf{B}$ ,

$$D_I = \frac{1}{qr} \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \right] = \\ = \frac{1}{qr} \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I (\mathbf{V}_I^{-1} - \mathbf{V}_I^{-1} (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I) \mathbf{V}_I^{-1}) \right].$$

En particular, si se está interesado en medir la influencia de una única observación, la  $i$ -ésima, es decir,  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , esta medida se puede expresar por

$$D_i = \frac{1}{qrv_{ii}^2} \mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{k}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i, \quad (2.4.6)$$

donde  $v_{ii}$ ,  $\mathbf{e}_i$  y  $\mathbf{k}_i$  vienen dados en (2.3.13), (2.3.14) y (2.3.12), respectivamente.

Además, considerando

$$\Delta_i = \frac{\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{k}_i}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i},$$

se obtiene

$$D_i = \frac{1}{qr} \Lambda_i (1 - \Lambda_i) \Delta_i,$$

donde  $\Lambda_i = T_i^{-1}$ , con  $T_i$  dado en (2.4.2). Como se puede observar, la medida de Cook depende del estadístico  $T_i$ , que a su vez sirve para identificar si la observación es outlier, y de la cantidad  $\Delta_i$ , distancia generalizada en el espacio  $\mathcal{C}(\mathbf{Z})$  (Pan y Fang, 2002).



- **Medida de Welsch-Kuh.** También conocida por DFFITS (Belsley et al., 1980). Tomando  $\mathbf{M} = \mathbf{Z}\mathbf{Z}^T$  y  $\mathbf{C} = qr(\mathbf{X}^T\mathbf{S}_{(I)}^{-1}\mathbf{X})^{-1}$  en (2.4.5), la medida de influencia de Welsch-Kuh, viene dada por

$$WK_I = \frac{1}{qr} \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \left( (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)^{-1} - \mathbf{V}_I^{-1} \right) \right],$$

y mide la desviación del valor ajustado bajo omisión  $\widehat{\mathbf{Y}}_{(I)} = \mathbf{X}\widehat{\mathbf{B}}_{(I)}\mathbf{Z}$  respecto al valor ajustado  $\widehat{\mathbf{Y}} = \mathbf{X}\widehat{\mathbf{B}}\mathbf{Z}$  de todo el conjunto de datos.

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida anterior tomará la siguiente expresión

$$WK_i = \frac{1}{qr} (1 - \Lambda_i) \Delta_i = D_i T_i.$$

A continuación, se recogen otras medidas de influencia, para el MCC, las cuales pueden ser vistas, por ejemplo, en Pan y Fang (2002).

- **Medida de razón de covarianzas.** Evalúa el efecto de las observaciones  $\mathbf{Y}_I$  sobre el determinante de la matriz de covarianzas del EMV  $\widehat{\mathbf{B}}$ , viene definida por

$$CR_I = \frac{\det[\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}]]}{\det[\widehat{\text{Var}}[\widehat{\mathbf{B}}]]} = c_7^{qr} (\det[\mathbf{I}_k - \mathbf{H}_I])^{-q} T_I^{-r},$$

donde  $\widehat{\text{Var}}[\widehat{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}]$  son los estimadores de las matrices de covarianza del MCC y MCC(I) dadas en (1.3.19) y (2.3.4), respectivamente, y  $c_7 = \frac{c_6(n-r)}{c_1(n-k-r)}$ , con  $c_1$  dado por (1.3.18) y  $c_6$  por (2.3.3).

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida  $CR_I$  se puede expresar por

$$CR_i = c_8^{qr} (1 - p_{ii})^{-q} T_i^{-r},$$

con

$$c_8 = \frac{(n-r)(n-r-2)(n-r-(p-q)-1)}{(n-r-1)^2(n-r-(p-q)-2)}.$$

Generalmente, se trabaja con la medida  $|CR_I - 1|$  como un criterio para analizar la influencia de  $\mathbf{Y}_I$  sobre la covarianza de  $\widehat{\mathbf{B}}$ . Cuanto mayor sea el valor del estadístico  $|CR_I - 1|$ , mayor será la influencia del subconjunto  $\mathbf{Y}_I$  en el ajuste de la curva de crecimiento.

- **Medida de Cook-Weisberg.** Se centra en el efecto de las observaciones de  $\mathbf{Y}_I$  sobre el volumen del elipsoide de confianza de  $\mathbf{B}$ . Designando  $ACE$  y  $ACE_{(I)}$  los elipsoides asintóticos de confianza de  $\mathbf{B}$  para MCC y MCC(I), respectivamente, la medida de influencia de Cook-Weisberg viene dada por

$$CW_I = \log \left[ \frac{\text{volumen}(ACE_{(I)})}{\text{volumen}(ACE)} \right] = \frac{1}{2} \log CR_I + C_\alpha + \frac{qr}{2} \log \frac{c_1(n-k-r)}{c_6(n-r)},$$

donde  $C_\alpha = (qr/2) \log(C_\alpha^{**}/C_\alpha^*)$ , con  $C_\alpha^*$  y  $C_\alpha^{**}$ , respectivamente, los percentiles de orden  $100(1-\alpha)\%$  de las distribuciones  $GT^2(q, r, n-r-p+q)$  y  $GT^2(q, r, n-k-r-p+q)$  (distribución  $T^2$  de Hotelling generalizada dada en la definición C.5.1).

Cuando  $n$  es suficientemente grande,  $CW_I \simeq \frac{1}{2} \log CR_I + C_\alpha$ , es decir, en esta situación la medida de Cook-Weisberg es equivalente a la razón de covarianzas para analizar la influencia de  $\mathbf{Y}_I$ . La constante

$$C_\alpha + \frac{qr}{2} \log \frac{c_1(n-k-r)}{c_6(n-r)}$$

no afecta la detección de observaciones influyentes. Si  $CW_I$  es negativo (positivo) entonces el volumen del elipsoide de confianza decrece (crece) con la omisión del subconjunto  $\mathbf{Y}_I$ . Independientemente de los signos de  $CW_I$ , el subconjunto de observaciones que tenga mayor  $|CW_I|$  tendrá la mayor influencia en el ajuste de la curva de crecimiento.

- **Medida de razón de información.** Viene dada por

$$IR_I = \frac{\det[\widehat{G}_I(\mathbf{B})]}{\det[\widehat{G}(\mathbf{B})]} = \left( \frac{c_6}{c_1} \right)^{qr} CR_I^{-1},$$

donde

$$\widehat{G}(\mathbf{B}) = (n-r)(\mathbf{Z}\mathbf{Z}^T) \otimes (\mathbf{X}^T\mathbf{S}^{-1}\mathbf{X})$$

y

$$\widehat{G}_I(\mathbf{B}) = (n-k-r)(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T) \otimes (\mathbf{X}^T\mathbf{S}_{(I)}^{-1}\mathbf{X}),$$

son estimadores consistentes de la matriz de información del coeficiente de regresión  $\mathbf{B}$ , para el MCC y MCC( $I$ ), respectivamente, (Pan y Fang, 2002).

- **Medida de Andrew-Pregibon.** Mide el efecto de  $\mathbf{Y}_I$  en el EMV de la matriz de covarianza  $\boldsymbol{\Sigma}$  (Cook y Weisberg, 1982) y puede ser definida por

$$AP_I = -\frac{1}{2} \left( \log \det [\mathbf{I}_k - \mathbf{H}_I] + \log \det [\mathbf{S} + \mathbf{X}\mathbf{X}^T - \mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T] - \log \det [\mathbf{S} + \mathbf{X}\mathbf{X}^T] \right).$$

En particular, cuando  $k=1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida  $AP_I$  se reduce a

$$AP_i = -\left( \frac{1}{2} \log [1 - p_{ii}] + \log \det \left[ \mathbf{S} + \mathbf{X}\mathbf{X}^T - \frac{\mathbf{e}_i\mathbf{e}_i^T}{1 - p_{ii}} \right] - \log \det [\mathbf{S} + \mathbf{X}\mathbf{X}^T] \right).$$

En muchas situaciones, el objetivo del investigador es analizar la influencia de las observaciones  $\mathbf{Y}_I$  en un determinado subconjunto del estimador de la matriz de coeficientes de regresión  $\mathbf{B}$ . Para dicho estudio, sea  $\boldsymbol{\Psi} = \mathbf{D}\mathbf{B}\mathbf{L}$ , con  $\mathbf{D}$  y  $\mathbf{L}$  matrices de dimensiones  $(s \times q)$  y  $(r \times l)$ , respectivamente, donde  $\text{rg}(\mathbf{D}) = s \leq q$  y  $\text{rg}(\mathbf{L}) = l \leq r$ . Según Kollo y von Rosen (2005), el EMV de las  $s \times l$  combinaciones lineales de  $\mathbf{B}$  es

$$\widehat{\boldsymbol{\Psi}} = \underset{(s \times q)(q \times r)(r \times l)}{\mathbf{D} \widehat{\mathbf{B}} \mathbf{L}}. \quad (2.4.7)$$

Análogamente, para el MCC( $I$ ), el EMV de  $\boldsymbol{\Psi}$  viene dado por  $\widehat{\boldsymbol{\Psi}}_{(I)} = \mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}$ . A partir de estos resultados, las medidas de influencia recogidas anteriormente pueden ser aplicadas a las combinaciones lineales, como ilustración se recoge la medida de Cook.

Teniendo en cuenta (2.4.4) y la relación entre los EMV  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , dada por el teorema 2.3.1, la distancia de Cook generalizada para  $\widehat{\Psi}$  viene dada por

$$\begin{aligned} DL_I(\mathbf{M}, \mathbf{C}) &= \text{tr} \left[ (\widehat{\Psi} - \widehat{\Psi}_{(I)}) \mathbf{M} (\widehat{\Psi} - \widehat{\Psi}_{(I)})^T \mathbf{C}^{-1} \right] = \\ &= \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{M} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \right. \\ &\quad \left. \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{C}^{-1} \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \right], \end{aligned}$$

con matrices de pesos  $\mathbf{M}_{(l \times l)}$  y  $\mathbf{C}_{(s \times s)}$  d.p.. Tomando  $\mathbf{M} = (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1}$  y  $\mathbf{C} = sl \mathbf{D} \mathbf{S}_X \mathbf{D}^T$  la expresión general de la medida Cook para las combinaciones lineales de  $\widehat{\mathbf{B}}$  vendrá dada por

$$\begin{aligned} DL_I &= \frac{1}{sl} \text{tr} \left[ \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \mathbf{V}_I^{-1} \right. \\ &\quad \left. \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D} \mathbf{S}_X \mathbf{D}^T)^{-1} \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \right]. \end{aligned}$$

Cuando  $I = \{i\}$ ,  $k = 1$  y  $1 \leq i \leq n$ , se obtiene

$$\begin{aligned} DL_i &= \frac{1}{slv_{ii}^2} \mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{k}_i \\ &\quad \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D} \mathbf{S}_X \mathbf{D}^T)^{-1} \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i. \end{aligned} \quad (2.4.8)$$

Como aplicaciones de la medida anterior, a continuación, se presentan tres medidas resultantes de elecciones usuales de las matrices  $\mathbf{D}$  y  $\mathbf{L}$ :

- Para estudiar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $j$ -ésima fila del EMV  $\widehat{\mathbf{B}}$ , se puede elegir  $\mathbf{L} = \mathbf{I}_r$ ,  $\text{rg}(\mathbf{L}) = r$  y  $\mathbf{D} = \mathbf{d}_j^T$ ,  $\text{rg}(\mathbf{D}) = 1 \leq q$ , donde  $\mathbf{d}_j \in \mathbb{R}^q$  es un vector de dimensión  $q$  en el que la  $j$ -ésima componente es la unidad y las restantes son nulas. En estos términos, (2.4.7) se reduce a  $\widehat{\Psi} = \mathbf{d}_j^T \widehat{\mathbf{B}}$  y la medida de Cook (2.4.8) viene dada por

$$DL_{i(j\cdot)} = \frac{1}{rv_{ii}^2} \frac{\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{k}_i}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2. \quad (2.4.9)$$

- Para analizar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  en la  $m$ -ésima columna de  $\widehat{\mathbf{B}}$ , se puede escoger  $\mathbf{D} = \mathbf{I}_q$ ,  $\text{rg}(\mathbf{D}) = q$  y  $\mathbf{L} = \mathbf{h}_m$ ,  $\text{rg}(\mathbf{L}) = 1 \leq r$ , con  $\mathbf{h}_m \in \mathbb{R}^r$  un vector de dimensión  $r$  donde la  $m$ -ésima componente es la unidad y las restantes son nulas. Así, (2.4.7) se reduce a  $\widehat{\Psi} = \widehat{\mathbf{B}} \mathbf{h}_m$  y, de (2.4.8), se obtiene

$$DL_{i(m)} = \frac{1}{qv_{ii}^2} \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2. \quad (2.4.10)$$

- Si el objetivo es analizar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  sobre el  $(j, m)$ -ésimo elemento de  $\widehat{\mathbf{B}}$ ,  $1 \leq j \leq q$  y  $1 \leq m \leq r$ , basta considerar  $\mathbf{D} = \mathbf{d}_j^T$ ,  $\text{rg}(\mathbf{D}) = 1 \leq q$  y  $\mathbf{L} = \mathbf{h}_m$ ,  $\text{rg}(\mathbf{L}) = 1 \leq r$ . En este caso, (2.4.7) se reduce a  $\widehat{\Psi} = \mathbf{d}_j^T \widehat{\mathbf{B}} \mathbf{h}_m \in \mathbb{R}$  y, teniendo en cuenta  $DL_{i(j\cdot)}$  y  $DL_{i(m)}$  dadas en (2.4.9) y (2.4.10), respectivamente, de (2.4.8), se tiene

$$DL_{i(jm)} = \frac{DL_{i(j\cdot)} DL_{i(m)}}{D_i} = D_i \left[ \frac{DL_{i(j\cdot)}}{D_i} \right] \left[ \frac{DL_{i(m)}}{D_i} \right],$$

donde  $D_i$  es la medida de Cook (2.4.6) para  $\widehat{\mathbf{B}}$ . Los factores  $\frac{DL_{i(j\cdot)}}{D_i}$  y  $\frac{DL_{i(m)}}{D_i}$  representan las contribuciones de las filas y columnas de  $\widehat{\mathbf{B}}$  en la influencia sobre  $\widehat{\mathbf{B}}$ . Esta última igualdad también indica que  $DL_{i(jm)}$  es proporcional a  $DL_{i(j\cdot)}$  y a  $DL_{i(m)}$ .

## 2.5. Análisis de influencia en el modelo de curvas de crecimiento con estructura simple de Rao

Análogamente al estudio realizado para el MCC, a continuación, se presenta un resumen de medidas diagnóstico para analizar la influencia de las observaciones en el EMV  $\tilde{\mathbf{B}}$  de la matriz de coeficientes de regresión del MCC con estructura simple de Rao.

Siguiendo el mismo esquema anterior, en primer lugar, se recogen algunos resultados sobre el MCC(SCS) bajo la omisión del conjunto de observaciones subindicadas por  $I$ , el MCC( $I$ ) definido en (2.3.1) con estructura SCS, que se denotará por MCC(SCS)( $I$ ).

Del teorema 1.4.1, para el MCC(SCS)( $I$ ), se obtiene que los EMV de los parámetros  $\mathbf{B}$ ,  $\mathbf{\Gamma}$  y  $\mathbf{\Theta}$ , son, respectivamente,

$$\begin{aligned}\tilde{\mathbf{B}}_{(I)} &= (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}_{(I)} \mathbf{Z}_{(I)}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1}, \\ \tilde{\mathbf{\Gamma}}_{(I)} &= \frac{1}{n-k} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{S}_{(I)} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}\end{aligned}\quad (2.5.1)$$

y

$$\tilde{\mathbf{\Theta}}_{(I)} = \frac{1}{n-k} (\mathbf{X}^{oT} \mathbf{X}^o)^{-1} \mathbf{X}^{oT} \mathbf{Y}_{(I)} \mathbf{Y}_{(I)}^T \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{X}^o)^{-1},$$

donde  $\mathbf{S}_{(I)}$  viene dada en (2.3.2) y  $\mathcal{C}(\mathbf{X}^o) = \mathcal{C}(\mathbf{X})^\perp$ .

Además, de (1.4.5), se obtiene que

$$\tilde{\mathbf{\Sigma}}_{(I)} = \mathbf{X} \tilde{\mathbf{\Gamma}}_{(I)} \mathbf{X}^T + \mathbf{X}^o \tilde{\mathbf{\Theta}}_{(I)} \mathbf{X}^{oT}. \quad (2.5.2)$$

Teniendo en cuenta la distribución de  $\tilde{\mathbf{B}}$  para el MCC(SCS), definida por (1.4.6), para el MCC(SCS)( $I$ ) se verifica que

$$\tilde{\mathbf{B}}_{(I)} \sim N_{q,r} \left( \mathbf{B}; \mathbf{\Gamma}, \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \right).$$

Como  $\text{Var}[\tilde{\mathbf{B}}_{(I)}]$  contiene el parámetro de covarianza desconocido  $\mathbf{\Sigma}$ , se puede considerar el EMV  $\tilde{\mathbf{\Sigma}}_{(I)}$ , dado en (2.5.2), para obtener

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] = \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)}. \quad (2.5.3)$$

Utilizando el lema 2.3.1, el siguiente teorema recoge una relación importante entre los EMV de  $\mathbf{B}$ ,  $\mathbf{\Gamma}$  y  $\mathbf{\Theta}$ , para el MCC(SCS) y MCC(SCS)( $I$ ), su demostración puede ser vista en Pan y Fang (2002).

**Teorema 2.5.1** *Las relaciones entre los EMV de los parámetros  $\mathbf{B}$ ,  $\mathbf{\Gamma}$  y  $\mathbf{\Theta}$  del MCC(SCS) y del MCC(SCS)( $I$ ), respectivamente, son*

$$\tilde{\mathbf{B}}_{(I)} = \tilde{\mathbf{B}} - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1}, \quad (2.5.4)$$

$$\tilde{\mathbf{\Gamma}}_{(I)} = \frac{n}{n-k} \tilde{\mathbf{\Gamma}} - \frac{1}{n-k} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (2.5.5)$$

y

$$\tilde{\mathbf{\Theta}}_{(I)} = \frac{n}{n-k} \tilde{\mathbf{\Theta}} - \frac{1}{n-k} (\mathbf{X}^{oT} \mathbf{X}^o)^{-1} \mathbf{X}^{oT} \mathbf{Y}_I \mathbf{Y}_I^T \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{X}^o)^{-1}.$$

Además,

$$\tilde{\Sigma}_{(I)} = \frac{n}{n-k} \tilde{\Sigma} - \frac{1}{n-k} \left[ \mathbf{P}_X \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{P}_X + (\mathbf{I}_p - \mathbf{P}_X) \mathbf{Y}_I \mathbf{Y}_I^T (\mathbf{I}_p - \mathbf{P}_X) \right],$$

donde  $\mathbf{E}_I$  y  $\mathbf{H}_I$  están dadas en (2.3.6) y (2.3.7), respectivamente.

**Corolario 2.5.1** Para el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , se verifica

$$\tilde{\mathbf{B}}_{(i)} = \tilde{\mathbf{B}} - \frac{1}{1-p_{ii}} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1},$$

$$\tilde{\Gamma}_{(i)} = \frac{n}{n-1} \tilde{\Gamma} - \frac{1}{(n-1)(1-p_{ii})} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1},$$

$$\tilde{\Theta}_{(i)} = \frac{n}{n-1} \tilde{\Theta} - \frac{1}{n-1} (\mathbf{X}^{oT} \mathbf{X}^o)^{-1} \mathbf{X}^{oT} \mathbf{y}_i \mathbf{y}_i^T \mathbf{X}^o (\mathbf{X}^{oT} \mathbf{X}^o)^{-1}$$

y

$$\tilde{\Sigma}_{(i)} = \frac{n}{n-1} \tilde{\Sigma} - \frac{1}{n-1} \left[ \frac{1}{1-p_{ii}} \mathbf{P}_X \mathbf{e}_i \mathbf{e}_i^T \mathbf{P}_X + (\mathbf{I}_p - \mathbf{P}_X) \mathbf{y}_i \mathbf{y}_i^T (\mathbf{I}_p - \mathbf{P}_X) \right],$$

con  $p_{ii}$  y  $\mathbf{e}_i$  dados por (2.3.15) y (2.3.14), respectivamente.

A continuación, se recogen dos estadísticos importantes para el estudio de la influencia en el MCC(SCS) (Pan y Fang, 2002 y Pan, 2004).

Para este modelo, la distribución de

$$\Lambda_I = \det \left[ \mathbf{I}_k - \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \right] \quad (2.5.6)$$

viene dada por

$$\Lambda_I \sim \Lambda(q, n-k-r, k).$$

En el caso particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , vendrá expresado por

$$\Lambda_i = 1 - \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}}. \quad (2.5.7)$$

En el MCC(SCS), el siguiente estadístico, denominado "residuo internamente studentizado" (Pan y Fang, 2002) se utiliza para determinar si la observación  $i$ -ésima de  $\mathbf{Y}$  es outlier

$$T_i = 1 - \Lambda_i = \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}}. \quad (2.5.8)$$

Dicha observación será considerada outlier si

$$T_i \geq \frac{qC_\alpha^*}{(n-r-q) + qC_\alpha^*} \quad (2.5.9)$$

donde  $C_\alpha^*$  es el percentil de orden  $100(1-\alpha)\%$  de la distribución  $F_{q, n-r-q}$ .

Para el MCC(SCS), teniendo en cuenta el teorema 2.5.1, la distancia de Cook generalizada se puede expresar como

$$D_I(\mathbf{M}, \mathbf{C}) = \text{tr} \left[ \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{M} (\mathbf{Z}\mathbf{Z}^T)^{-1} \right. \\ \left. \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^{-1} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right]. \quad (2.5.10)$$

A partir de esta distancia, se obtiene:

- **Medida de Cook.** Considerando  $\mathbf{M} = \mathbf{Z}\mathbf{Z}^T$  y  $\mathbf{C} = \tilde{\mathbf{\Gamma}}$ , resulta

$$D_I = n \text{tr} \left[ \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{H}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \right].$$

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ ,

$$D_i = n \frac{p_{ii}}{1 - p_{ii}} T_i = n \frac{p_{ii}}{1 - p_{ii}} (1 - \Lambda_i). \quad (2.5.11)$$

Esta medida es proporcional al estadístico  $T_i$  usado para la identificación de outliers.

- **Medida de Welsch-Kuh.** Eligiendo  $\mathbf{M} = \mathbf{Z}\mathbf{Z}^T$  y  $\mathbf{C} = \tilde{\mathbf{\Gamma}}_{(I)}$ , se obtiene

$$WK_I = (n - k) \text{tr} \left[ \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S}_{(I)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{H}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \right].$$

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ ,

$$WK_i = (n - 1) \frac{p_{ii}}{1 - p_{ii}} \frac{T_i}{1 - T_i}.$$

Además, teniendo en cuenta (2.5.11), se obtiene que

$$WK_i = (n - 1) \frac{p_{ii} D_i}{np_{ii} - (1 - p_{ii}) D_i},$$

es una función creciente de  $D_i$ . Así el patrón de influencia detectado por  $WK_i$  deberá ser el mismo que el detectado por  $D_i$ . No obstante, cuando  $k > 1$  esta conclusión puede no ser válida (Pan y Fang, 2002).

**Nota 2.5.1** En el modelo lineal multivariante, el factor  $\frac{p_{ii}}{1 - p_{ii}}$  es conocido como el potencial de la  $i$ -ésima observación  $\mathbf{y}_i$ , una medida usada para detectar los casos nominados como "leverage" o "puntos de palanca" del espacio de  $\mathbf{Z}$  (véase, por ejemplo, Chatterjee y Hadi, 1988). Como se recoge en Cook y Weisberg (1982), el potencial puede ser usado para detectar configuraciones de las variables explicativas que son capaces de producir casos altamente influyentes. En el MCC, no tiene sentido esta designación dado que  $p_{ii}$  es constante en cada grupo al depender sólo de  $\mathbf{Z}$ .

**Nota 2.5.2** La cantidad  $\frac{T_i}{1 - T_i} = \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S}_{(i)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i}{1 - p_{ii}}$  se denomina "suma de cuadrados residual externamente studentizada" para el MCC(SCS) (Pan y Fang, 2002).

A continuación, se recogen otras medidas de influencia para el MCC(SCS), las cuáles pueden verse, por ejemplo, en Pan y Fang (2002).

- **Medida Razón de covarianzas.** Sean  $\widehat{\text{Var}}[\widetilde{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\widetilde{\mathbf{B}}_{(I)}]$  los estimadores de las matrices de covarianza de  $\widetilde{\mathbf{B}}$  en el MCC(SCS) y MCC(SCS)( $I$ ), dados por (1.4.7) y (2.5.3), respectivamente, entonces

$$CR_I = \frac{\det[\widehat{\text{Var}}[\widetilde{\mathbf{B}}_{(I)}]]}{\det[\widehat{\text{Var}}[\widetilde{\mathbf{B}}]]} = \left(\frac{n}{n-k}\right)^{qr} (\det[\mathbf{I}_k - \mathbf{H}_I])^{-q} \Lambda_I^r.$$

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ ,

$$CR_i = \left(\frac{n}{n-1}\right)^{qr} \frac{\Lambda_i^r}{(1-p_{ii})^q}.$$

Generalmente, se trabaja con  $|CR_I - 1|$  como un criterio para analizar la influencia de  $\mathbf{Y}_I$  sobre la covarianza de  $\widetilde{\mathbf{B}}$ .

- **Medida de Cook-Weisberg.** Considerando  $CE$  y  $CE_{(I)}$  los elipsoides de confianza de  $\mathbf{B}$ , para MCC(SCS) y MCC(SCS)( $I$ ), respectivamente, entonces

$$CW_I = \log \left[ \frac{\text{volumen}(CE_{(I)})}{\text{volumen}(CE)} \right] = \frac{1}{2} \log CR_I + C_\alpha, \quad (2.5.12)$$

donde  $C_\alpha = (qr/2) \log(C_\alpha^{**}/C_\alpha^*)$ , con  $C_\alpha^*$  y  $C_\alpha^{**}$  los percentiles de orden  $100(1-\alpha)\%$  de las distribuciones  $GT^2(q, r, n-r)$  y  $GT^2(q, r, n-k-r)$ , respectivamente.

Si  $CW_I$  es negativo (positivo) entonces el volumen del elipsoide de confianza decrece (crece) con la omisión del subconjunto  $\mathbf{Y}_I$ . El subconjunto de observaciones que tiene mayor  $|CW_I|$  tendrá la mayor influencia en el ajuste de la curva de crecimiento. La constante  $C_\alpha$  no afecta la elección de observaciones influyentes, por tanto, la relación (2.5.12) implica que la medida  $CW_I$  es equivalente a  $CR_I$  para analizar la influencia de  $\mathbf{Y}_I$ .

- **Medida de Andrews-Pregibon.** Viene dada por

$$AP_I = -\frac{1}{2} \log [\det[\mathbf{I}_k - \mathbf{H}_I] \Lambda_I].$$

En particular, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ ,

$$AP_i = -\frac{1}{2} \log [(1-p_{ii}) \Lambda_i] = -\frac{1}{2} \log [(1-p_{ii})(1-T_i)],$$

una función creciente de  $T_i$ .

Análogamente al estudio realizado para el MCC, se presenta la medida de Cook para evaluar, en el MCC(SCS), la influencia de las observaciones  $\mathbf{Y}_I$  en combinaciones lineales de  $\tilde{\mathbf{B}}$ . En este caso, los EMV vendrán dados por  $\tilde{\Psi} = \mathbf{D}\tilde{\mathbf{B}}\mathbf{L}$  para el MCC(SCS) y por  $\tilde{\Psi}_{(I)} = \mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}$  para el MCC(SCS)( $I$ ), donde  $\text{rg}(\mathbf{D}_{(s \times q)}) = s \leq q$  y  $\text{rg}(\mathbf{L}_{(r \times l)}) = l \leq r$ .

Teniendo en cuenta la relación entre los EMV  $\tilde{\mathbf{B}}$  y  $\tilde{\mathbf{B}}_{(I)}$ , dada por el teorema 2.5.1, se obtiene la distancia de Cook generalizada

$$DL_I(\mathbf{M}, \mathbf{C}) = \text{tr} \left[ \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}\mathbf{M}\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \right. \\ \left. \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \mathbf{C}^{-1} \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right].$$

La medida de Cook para  $\tilde{\Psi}$  puede ser obtenida tomando  $\mathbf{M} = (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1}$  y  $\mathbf{C} = \mathbf{D}\tilde{\Gamma}\mathbf{D}^T$ , así

$$DL_I = n \text{tr} \left[ \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \left( \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{S}\mathbf{X}) (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \right)^{-1} \right. \\ \left. \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right. \\ \left. \left( \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right)^{-1} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \right]. \quad (2.5.13)$$

Cuando  $I = \{i\}$  y  $k = 1$ ,  $1 \leq i \leq n$ , se obtiene

$$DL_i = \frac{n}{(1 - p_{ii})^2} \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \\ \mathbf{D}^T (\mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{S}\mathbf{X}) (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T)^{-1} \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i. \quad (2.5.14)$$

Como casos particulares de esta medida se pueden analizar las siguientes:

- Para analizar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $j$ -ésima fila de  $\tilde{\mathbf{B}}$ , (2.5.14) se reduce a

$$DL_{i(j \cdot)} = \frac{p_{ii}}{(1 - p_{ii})^2} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} = \\ = \frac{1}{n(1 - p_{ii})} \frac{D_i (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{T_i \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j}, \quad (2.5.15)$$

donde  $\mathbf{d}_j \in \mathbb{R}^q$  es un vector de dimensión  $q$  en el que la  $j$ -ésima componente es la unidad y las restantes son nulas.

- Cuando se pretende diagnosticar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  en la  $m$ -ésima columna de  $\tilde{\mathbf{B}}$ , (2.5.14) viene dada por

$$DL_{i(\cdot m)} = \frac{nT_i}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} = \frac{D_i (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{p_{ii} \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m}, \quad (2.5.16)$$

con  $\mathbf{h}_m \in \mathbb{R}^r$  un vector de dimensión  $r$  donde la  $m$ -ésima componente es la unidad y las restantes son nulas.



Teniendo en cuenta que

$$\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i = \begin{cases} \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i = p_{ii}, & \mathbf{h}_m = \mathbf{z}_i \\ 0 & , \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$

donde  $\mathbf{h}_m = \mathbf{z}_i$  significa que la columna de  $\tilde{\mathbf{B}}$  para la cual se pretende estudiar la influencia corresponde al grupo de la observación eliminada, de (2.5.16) viene

$$DL_{i(\cdot m)} = \begin{cases} D_i, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}. \quad (2.5.17)$$

Por tanto, para el análisis de la influencia en las columnas de  $\tilde{\mathbf{B}}$ , la medida de Cook para el MCC(SCS) asume que cada grupo tiene una curva de crecimiento diferente, de manera que un grupo no tiene influencia en el otro, una vez que  $DL_{i(\cdot m)} = 0$  cuando  $\mathbf{h}_m \neq \mathbf{z}_i$ .

- Análogamente, para estudiar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  sobre el  $(j, m)$ -ésimo elemento de  $\tilde{\mathbf{B}}$ , teniendo en cuenta  $DL_{i(j\cdot)}$  y  $DL_{i(\cdot m)}$  definidas en (2.5.15) y (2.5.16), respectivamente, de (2.5.14) se obtiene

$$DL_{i(jm)} = \frac{DL_{i(j\cdot)}DL_{i(\cdot m)}}{D_i} = D_i \left[ \frac{DL_{i(j\cdot)}}{D_i} \right] \left[ \frac{DL_{i(\cdot m)}}{D_i} \right] \quad (2.5.18)$$

donde  $D_i$  es la medida de Cook (2.5.11) para  $\tilde{\mathbf{B}}$ . Además, por (2.5.17) se tiene

$$DL_{i(jm)} = \begin{cases} DL_{i(j\cdot)}, & \mathbf{h}_m = \mathbf{z}_i \\ 0 & , \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$

o sea, usando la medida de Cook en el MCC(SCS), la influencia de la  $i$ -ésima observación es solamente en los elementos de  $\tilde{\mathbf{B}}$  que pertenecen a las columnas correspondientes a los grupos de dicha observación.



## Capítulo 3

# Influencia a través de la distancia de Fréchet

### 3.1. Introducción

En el capítulo anterior se ha presentado una amplia batería de medidas utilizadas en la detección de observaciones influyentes en el estimador de máxima verosimilitud (EMV) de la matriz de coeficientes de regresión del modelo de curvas de crecimiento (MCC) y del MCC con estructura simple de Rao (MCC(SCS)). En el Análisis de Influencia, la gran mayoría de las técnicas de detección de observaciones influyentes están orientadas hacia el estudio de la influencia sobre las estimaciones de los parámetros de interés del modelo, comparando directamente los valores obtenidos en dichas estimaciones a partir del modelo postulado y del modelo perturbado. Un enfoque alternativo es diseñar técnicas que permitan analizar la influencia sobre la distribución muestral de dichos estimadores. Esto es, no sólo sobre el valor del estimador obtenido en una muestra concreta, sino en toda la distribución de valores que puede tomar el estimador bajo estudio. Ello puede proporcionar medidas de influencia más globales, más completas y, por tanto, menos sensibles a posibles fluctuaciones muestrales. Así, Jiménez-Gamero et al. (2002) proponen una modelización genérica del problema y, en particular, utilizan la distancia de Rao para cuantificar la discrepancia entre las distribuciones muestrales. A raíz de esta modelización quedó abierta una línea de trabajo concreta que proporciona una nueva e interesante visión sobre el problema de la influencia, con la aplicación de otras distancias distribucionales. Esta aproximación al problema del diagnóstico es en la que se centra la presente memoria y, en particular, este capítulo.

Con este propósito, en este capítulo y en el siguiente se derivan medidas de influencia a través de distancias entre distribuciones de probabilidad, para la detección de observaciones influyentes en el EMV de la matriz de coeficientes de regresión del MCC y del MCC(SCS). Este es el primer paso para, posteriormente, decidir el tratamiento más adecuado que debe darse a las observaciones que se han determinado como influyentes.

Este capítulo está organizado de la siguiente manera. En la sección 3.2, se plantea la metodología para el análisis de influencia a través de distancias entre distribuciones. Esto sirve como fundamento de los principales resultados de la presente memoria.

En la sección 3.3, se expone la utilización de la distancia de Fréchet en el análisis de influencia. Donde, en particular, se evidencia el interés de la aplicación de dicha distancia una vez que se

puede interpretarse como función de una componente de localización y de una componente de dispersión.

En la sección 3.4, se obtienen medidas de influencia para el EMV de la matriz de coeficientes de regresión del MCC, así como para combinaciones lineales del mismo, basadas en la distancia de Fréchet. Se analiza su aplicabilidad, así como algunas diferencias y relaciones respecto a las medidas expuestas en la sección 2.4.

En la sección 3.5, se hace el mismo estudio que en la sección 3.4, pero para el MCC(SCS).

### 3.2. Análisis de influencia a través de la distancia entre distribuciones

En general, como se ha comentado anteriormente, el estudio de la influencia sobre un modelo se aborda como la evaluación del cambio resultante de la perturbación del modelo. Para dicha evaluación se han propuesto una gran variedad de métodos específicos para el modelo bajo estudio y para los estimadores de interés. Así, tales diagnósticos de influencia tienden a evaluar algunas características aisladas de los cambios provocados por la perturbación. Jiménez-Gamero et al. (2002) proponen una aproximación general al problema basada en medir la distancia entre la distribución del estadístico de interés bajo ambos modelos, el modelo postulado y el modelo perturbado. Dicha aproximación puede ser descrita como sigue:

”Sea  $D$  un conjunto de datos sobre los cuales se asume su adecuación a un modelo  $M$ . Sea  $R = R(D, M)$  un estadístico definido sobre los datos y sea  $F_R$  su distribución. Sea  $M^*$  una perturbación del modelo y sean  $R^*$  y  $F_{R^*}$  el estadístico y su distribución bajo el modelo perturbado. Los cambios debidos a la perturbación pueden cuantificarse a través de una distancia o divergencia  $\delta$  entre  $F_R$  y  $F_{R^*}$  :  $\delta(F_R, F_{R^*})$ .”

Este enfoque genérico permite obtener medidas de influencia más globales en tanto que no se centran en el efecto sobre el valor de un estadístico o estimador obtenido para una muestra concreta, sino sobre la distribución de valores posibles que dicho estadístico o estimador puede tomar, bajo la ley de probabilidad estimada a partir de la muestra.

En caso de que el esquema de perturbación sea la omisión de un conjunto de observaciones, la aproximación propuesta quedaría como sigue:

”Sea  $D$  un conjunto de datos sobre los cuales se asume su adecuación a un modelo  $M$ . Sea  $R = R(D, M)$  un estadístico definido sobre los datos y sea  $F_R$  su distribución. Sean  $D_{(I)}$  y  $M_{(I)}$  la perturbación del conjunto de datos y del modelo, respectivamente, provocada por la omisión de un conjunto  $I$  de observaciones, y sean  $R_{(I)}$  y  $F_{R_{(I)}}$  el estadístico y su distribución bajo el modelo perturbado. Los cambios debidos a la perturbación pueden cuantificarse a través de una distancia o divergencia  $\delta$  entre  $F_R$  y  $F_{R_{(I)}}$  :  $\delta(F_R, F_{R_{(I)}})$ .”

Son muchas las métricas entre distribuciones propuestas en la literatura. No obstante, no todas son adecuadas para el estudio de la influencia. Más aún, la adecuación de las mismas puede depender del modelo bajo estudio. Dos medidas que se han mostrado útiles en el análisis de influencia son la distancia de Rao (Rao, 1949) y la distancia de Fréchet (Fréchet, 1957). Otras métricas tales como la métrica de Kolmogorov (Kolmogorov, 1933), o la métrica de Kantorovich

o Wasserstein (Kantorovich y Rubinstein, 1958 y Vallender, 1974) pueden ser también útiles para el objetivo marcado en algunos modelos.

La utilidad de esta aproximación ha sido demostrada en situaciones concretas recogidas en los trabajos de Jiménez-Gamero et al. (2002), Muñoz-Pichardo et al. (2004) y García-Heras et al. (2006), dedicados al modelo lineal general multivariante, aunque con distintas hipótesis distribucionales y con distintas distancias entre distribuciones, en el trabajo de Muñoz-García et al. (2006) dedicado al modelo de regresión logística y en Muñoz-Pichardo et al. (2008) para el análisis de perfiles con errores aleatorios elípticamente distribuidos. Para una aplicación de dicha metodología a otros modelos estadísticos véase Muñoz-Pichardo (2006). Además, también se han utilizado para fundamentar técnicas de detección de outliers. Por ejemplo, Hadi y Nyquist (1999) utilizan la distancia de Fréchet para la detección de outliers multivariantes. Incluso la distancia de Cook puede estar incluida en esta aproximación dado que se puede obtener, bajo condiciones de normalidad, aplicando la métrica de Mahalanobis.

En este capítulo, se considera el MCC y el MCC(SCS) y se estudia el análisis de influencia en el EMV de la matriz de coeficientes de regresión y en combinaciones lineales del mismo, usando la aproximación general anteriormente mencionada para la distancia de Fréchet.

### 3.3. Distancia de Fréchet y su utilización en el análisis de influencia

Fréchet (1957) introdujo una métrica en el espacio de las distribuciones de probabilidad sobre  $\mathbb{R}$  con momentos de segundo orden finito, que más tarde fue extendida al espacio de las distribuciones multivariantes.

Para dos elementos de la clase de distribuciones con momentos de segundo orden finito,  $F_1$  y  $F_2$ , se define la métrica de Fréchet como

$$\delta(F_1, F_2) = \left( \min \left\{ E \left[ \|\mathbf{X}_1 - \mathbf{X}_2\|^2 \right] : \mathcal{L}(X_1) = F_1, \mathcal{L}(X_2) = F_2 \right\} \right)^{\frac{1}{2}},$$

donde  $\mathcal{L}(X)$  representa la distribución del vector aleatorio  $X$ .

La distribución bivalente  $H(x, y)$  para la cual se alcanza dicho mínimo es la distribución

$$H(x, y) = \min\{F_1(x), F_2(y)\}.$$

En el caso particular de que  $F_1$  y  $F_2$  pertenezcan a una familia de distribuciones cerrada con respecto a cambios de localización y escala, la distancia de Fréchet toma la forma

$$\delta(F_1, F_2) = \sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2},$$

donde  $\mu_1$ ,  $\mu_2$ ,  $\sigma_1$  y  $\sigma_2$  son, respectivamente, las medias y desviaciones típicas de  $F_1$  y  $F_2$ .

La definición anterior se puede generalizar al caso multivariante de forma directa como sigue. Sea  $\mathcal{P}(\mathbb{R}^q)$  el espacio de las distribuciones de probabilidad sobre  $\mathbb{R}^q$  con momentos de segundo orden finitos. Dadas dos distribuciones  $F_1, F_2 \in \mathcal{P}(\mathbb{R}^q)$ , la distancia de Fréchet entre  $F_1$  y  $F_2$  se define como

$$\delta(F_1, F_2) = \left( \min_{X_1, X_2} \left\{ E \left[ \|\mathbf{X}_1 - \mathbf{X}_2\|^2 \right] : \mathcal{L}(\mathbf{X}_1) = F_1, \mathcal{L}(\mathbf{X}_2) = F_2 \right\} \right)^{\frac{1}{2}}.$$

La solución obtenida para el caso univariante no es trasladable al caso multivariante, resultando en este caso complejo su cálculo.

Sin embargo, Dowson y Landau (1982) demuestran que si  $F_1$  y  $F_2$  pertenecen a una familia de distribuciones  $q$ -dimensionales cerrada con respecto a transformaciones lineales del vector aleatorio entonces la distancia de Fréchet viene dada por,

$$\begin{aligned}\delta(F_1, F_2) &= \delta[(\mu_1, \Sigma_1), (\mu_2, \Sigma_2)] = \\ &= \left( \|\mu_1 - \mu_2\|^2 + \text{tr} \left[ \Sigma_1 + \Sigma_2 - 2(\Sigma_1 \Sigma_2)^{1/2} \right] \right)^{1/2},\end{aligned}$$

donde  $\mu_1, \mu_2, \Sigma_1$  y  $\Sigma_2$  son, respectivamente, los vectores medias y las matrices de covarianzas de  $F_1$  y  $F_2$ . Como caso particular de interés de familias cerradas respecto a transformaciones lineales se encuentra la familia de distribuciones elípticas definida en la sección C.1.

La distancia de Fréchet es de gran interés ya que es invariante bajo transformaciones ortogonales de las variables asociadas y, además, es una métrica que combina las desviaciones entre los vectores media y las desviaciones entre las matrices de covarianza. Dicha métrica puede ser interpretada de la siguiente manera (Hadi y Nyquist, 1999):

- La raíz del primer término,  $\|\mu_1 - \mu_2\|^2$ , define una métrica en  $\mathbb{R}^q$ .
- La raíz del segundo término,  $\text{tr} \left[ \Sigma_1 + \Sigma_2 - 2(\Sigma_1 \Sigma_2)^{1/2} \right]$ , define una métrica sobre el espacio de todas las matrices definidas positivas de orden  $q$ .

Así, en este caso, se puede interpretar la métrica de Fréchet como una función de dos componentes: una métrica sobre la localización de las distribuciones y una métrica sobre la estructura de covarianzas de las mismas. Muñoz-Pichardo et al. (2004) denominaron a estos dos términos, la componente de localización (*LC*) y la componente de dispersión (*DC*), respectivamente, y propusieron una versión muestral para dicha métrica como un diagnóstico de influencia cuando se pretende estudiar la influencia de la observación  $i$ -ésima. Se puede generalizar la formulación de Muñoz-Pichardo et al. (2004) al caso en que se pretende estudiar la influencia de un conjunto de observaciones, es decir, al estudio de la influencia conjunta.

Sea  $Y_1, \dots, Y_n$  una muestra aleatoria de un vector  $q$ -dimensional  $Y$  cuya distribución depende de un vector de parámetros  $\theta$ . Sea  $\hat{\theta} = \hat{\theta}(Y_1, \dots, Y_n)$  un estimador de  $\theta$  y  $\hat{\theta}_{(I)}$  el estimador obtenido con la omisión de las observaciones subindicadas por  $I = \{i_1, i_2, \dots, i_k\}$ . Sean  $F(\cdot, \theta)$  y  $F_{(I)}(\cdot, \theta)$  las distribuciones de  $\hat{\theta}$  y  $\hat{\theta}_{(I)}$ , respectivamente. Teniendo en cuenta Muñoz-Pichardo et al. (2004), la versión muestral de la distancia de Fréchet para la distribución de  $\hat{\theta}$  asociada al conjunto  $I$  es

$$\delta_I(\hat{\theta}) = \delta \left( F(\cdot, \hat{\theta}), F_{(I)}(\cdot, \hat{\theta}_{(I)}) \right). \quad (3.3.1)$$

Esta expresión puede ser considerada como una medida de la influencia que el subconjunto de observaciones subindicadas por  $I$  ejerce en el estimador  $\hat{\theta} = \hat{\theta}(Y_1, \dots, Y_n)$ .

En particular, si los estimadores  $\hat{\boldsymbol{\theta}}$  y  $\hat{\boldsymbol{\theta}}_{(I)}$  siguen una distribución elíptica, por ejemplo,  $\hat{\boldsymbol{\theta}} \sim E_p(\boldsymbol{\mu}_{\hat{\boldsymbol{\theta}}}, \boldsymbol{\Phi}_{\hat{\boldsymbol{\theta}}}, g)$  y  $\hat{\boldsymbol{\theta}}_{(I)} \sim E_p(\boldsymbol{\mu}_{\hat{\boldsymbol{\theta}}_{(I)}}, \boldsymbol{\Phi}_{\hat{\boldsymbol{\theta}}_{(I)}}, g)$ , y los estimadores de los parámetros de las distribuciones son  $\hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}}$ ,  $\hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}}$ ,  $\hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}_{(I)}}$  y  $\hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}_{(I)}}$ , respectivamente, entonces de (3.3.1), para  $\text{rg}(\boldsymbol{\Phi}_{\hat{\boldsymbol{\theta}}}) = p$ , se obtiene

$$\delta_I(\hat{\boldsymbol{\theta}}) = \left[ LC_I^2(\hat{\boldsymbol{\theta}}) + DC_I^2(\hat{\boldsymbol{\theta}}) \right]^{1/2}, \quad (3.3.2)$$

con componente de localización

$$LC_I^2(\hat{\boldsymbol{\theta}}) = \left\| \hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}} - \hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}_{(I)}} \right\|^2 \quad (3.3.3)$$

y componente de dispersión

$$DC_I^2(\hat{\boldsymbol{\theta}}) = \frac{E[R^2]}{p} \text{tr} \left[ \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}} + \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}_{(I)}} - 2 \left( \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}} \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}_{(I)}} \right)^{1/2} \right], \quad (3.3.4)$$

donde  $R^2$  es una variable aleatoria no negativa usada en la representación estocástica de la variable aleatoria  $\hat{\boldsymbol{\theta}}$  (véase el teorema C.1.4).

Cuando se está interesado en medir la influencia que la observación  $i$ -ésima ejerce sobre el estimador  $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(Y_1, \dots, Y_n)$ , a partir de (3.3.1), la medida de Fréchet se puede expresar por

$$\delta_i(\hat{\boldsymbol{\theta}}) = \delta \left( F(\cdot, \hat{\boldsymbol{\theta}}), F_{(i)}(\cdot, \hat{\boldsymbol{\theta}}_{(i)}) \right) = \left[ LC_i^2(\hat{\boldsymbol{\theta}}) + DC_i^2(\hat{\boldsymbol{\theta}}) \right]^{1/2},$$

donde  $\hat{\boldsymbol{\theta}}_{(i)} = \hat{\boldsymbol{\theta}}(Y_1, \dots, Y_{i-1}, Y_{i+1}, \dots, Y_n)$  es el estimador obtenido por omisión de la  $i$ -ésima observación,  $F(\cdot, \hat{\boldsymbol{\theta}})$  y  $F_{(i)}(\cdot, \hat{\boldsymbol{\theta}}_{(i)})$  son las distribuciones de  $\hat{\boldsymbol{\theta}}$  y  $\hat{\boldsymbol{\theta}}_{(i)}$ , respectivamente. La componente de localización es

$$LC_i^2(\hat{\boldsymbol{\theta}}) = \left\| \hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}} - \hat{\boldsymbol{\mu}}_{\hat{\boldsymbol{\theta}}_{(i)}} \right\|^2$$

y la componente de dispersión

$$DC_i^2(\hat{\boldsymbol{\theta}}) = \frac{E[R^2]}{p} \text{tr} \left[ \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}} + \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}_{(i)}} - 2 \left( \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}} \hat{\boldsymbol{\Phi}}_{\hat{\boldsymbol{\theta}}_{(i)}} \right)^{1/2} \right].$$

En las dos secciones siguientes, se recoge la aplicación de la distancia de Fréchet al MCC y al MCC(SCS) bajo la hipótesis de normalidad.

### 3.4. La medida de Fréchet en el modelo de curvas de crecimiento

Siguiendo las ideas presentadas en la sección anterior, en este apartado se recoge la aplicación de la distancia de Fréchet como medida de influencia sobre el EMV de la matriz de coeficientes de regresión del MCC. En particular, se pretende estudiar la influencia de un subconjunto de observaciones en el EMV  $\hat{\mathbf{B}}$  y en combinaciones lineales de dicho estimador. El modelo perturbado usado es el MCC( $I$ ), definido en la sección 2.3.

En el contexto del análisis de influencia, para aplicar la distancia de Fréchet en el MCC, es necesario conocer las distribuciones de  $\hat{\mathbf{B}}$  y  $\hat{\mathbf{B}}_{(I)}$  (o equivalentemente de  $\text{vec}(\hat{\mathbf{B}})$  y  $\text{vec}(\hat{\mathbf{B}}_{(I)})$ ).

Este último es el EMV de  $\mathbf{B}$  bajo la omisión del conjunto de observaciones subindicadas por  $I$ , es decir, en el MCC( $I$ ). Como se recoge en el teorema (1.3.3), el EMV de  $\mathbf{B}$  en el MCC es

$$\widehat{\mathbf{B}} = \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{Y} \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1}, \quad (3.4.1)$$

con  $\mathbf{S}_X = (\mathbf{X}^T \mathbf{S}^{-1} \mathbf{X})^{-1}$ , por lo que el EMV de  $\mathbf{B}$  en el modelo perturbado MCC( $I$ ) es

$$\widehat{\mathbf{B}}_{(I)} = \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{Y}_{(I)} \mathbf{Z}_{(I)}^T \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1}. \quad (3.4.2)$$

En este sentido, no hay una forma explícita para la distribución de  $\widehat{\mathbf{B}}$ . No obstante, dicha distribución admite como aproximaciones (véase sección 1.3):

- i) La distribución normal matricial  $N_{q,r}(\mathbf{B}; c_1(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$ , con  $c_1$  dado en (1.3.18).
- ii) La mixtura de una distribución normal matricial,  $N_{q,r}(\mathbf{B}; (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$ , con peso  $1 - sqr/2$ , y de una Kotz matricial,  $K_{q,r}(\mathbf{B}; (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$ , con peso  $sqr/2$ , con  $s$  dada en (1.3.21).

Además, Kollo et al. (2007) consideran la distribución marginal de la mixtura y demuestran que dichas distribuciones continúan siendo mixturas similares, con pesos más pequeños para la distribución de Kotz. Así, cuando se reduce la dimensión de la matriz aleatoria  $\mathbf{Y}$  crece la similitud a la distribución normal. A través de este resultado se puede concluir que en casos de baja dimensión la aproximación a la distribución normal puede ser apropiada, mientras en los casos de grandes dimensiones dicha aproximación no es tan buena.

A continuación, se comparan las dos aproximaciones i) y ii) para la aplicación de la distancia de Fréchet como medida de influencia en el EMV  $\widehat{\mathbf{B}}$  del MCC.

Según lo recogido en i), la distribución de  $\text{vec}(\widehat{\mathbf{B}})$  se puede aproximar a la distribución

$$N_{qr}(\text{vec}(\mathbf{B}), c_1(\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}). \quad (3.4.3)$$

Substituyendo  $\boldsymbol{\Sigma}$  por su estimador insesgado  $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n-r} \mathbf{S}$ , se obtiene

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}] = \frac{c_1}{n-r} (\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes \mathbf{S}_X. \quad (3.4.4)$$

Por otro lado, se concluye que la distribución de  $\text{vec}(\widehat{\mathbf{B}}_{(I)})$  puede ser aproximada por la distribución

$$N_{qr}(\text{vec}(\mathbf{B}), c_6(\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}), \quad (3.4.5)$$

con  $c_6$  dada en (2.3.3). Análogamente para el MCC( $I$ ), considerando  $\widehat{\boldsymbol{\Sigma}}_{(I)} = \frac{1}{n-k-r} \mathbf{S}_{(I)}$ , se obtiene el estimador

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] = \frac{c_6}{n-k-r} \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1}. \quad (3.4.6)$$



Entonces, como la distribución normal matricial pertenece a la clase de distribuciones elípticas, de (3.3.2), la versión muestral de la distancia de Fréchet  $\delta_I(\widehat{\mathbf{B}})$ , como medida de influencia del subconjunto de observaciones  $\mathbf{Y}_I$  sobre la distribución del EMV  $\widehat{\mathbf{B}}$  del MCC, puede ser aproximada por

$$\delta_I(\widehat{\mathbf{B}}) = \left[ LC_I^2(\widehat{\mathbf{B}}) + DC_I^2(\widehat{\mathbf{B}}) \right]^{1/2}. \quad (3.4.7)$$

De (3.3.3), la componente de localización, que representa la influencia del subconjunto de observaciones  $\mathbf{Y}_I$  sobre la esperanza de la distribución de  $\widehat{\mathbf{B}}$ , viene dada por

$$LC_I^2(\widehat{\mathbf{B}}) = \left\| \text{vec}(\widehat{\mathbf{B}}) - \text{vec}(\widehat{\mathbf{B}}_{(I)}) \right\|^2 = \left\| \text{vec}(\widehat{\mathbf{B}} - \widehat{\mathbf{B}}_{(I)}) \right\|^2 = \left\| \widehat{\mathbf{B}} - \widehat{\mathbf{B}}_{(I)} \right\|^2, \quad (3.4.8)$$

donde  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$  son los EMV de  $\mathbf{B}$  para el MCC y MCC( $I$ ), dados en (3.4.1) y (3.4.2), respectivamente.

A partir de (3.3.4), la componente de dispersión, que representa la influencia de las observaciones sobre la estructura de covarianzas de dicha distribución, se puede expresar por

$$\begin{aligned} DC_I^2(\widehat{\mathbf{B}}) &= \frac{E[R^2]}{qr} \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] + \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right] = \\ &= \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] + \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right], \end{aligned} \quad (3.4.9)$$

dado que, por la nota C.1.2,  $\frac{E[R^2]}{qr} = 1$ , donde  $\widehat{\text{Var}}[\widehat{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}]$  son los estimadores de las matrices de covarianza de  $\widehat{\mathbf{B}}$  en el MCC y MCC( $I$ ), dados en (3.4.4) y (3.4.6), respectivamente.

Dicha medida permite valorar tanto el cambio en la componente de localización como en la componente de dispersión de la distribución de  $\widehat{\mathbf{B}}$  bajo la omisión de un subconjunto de observaciones.

En resumen, se ha presentado la versión muestral de la medida de Fréchet cuando el estimador  $\widehat{\mathbf{B}}$  es aproximado por la distribución normal (3.4.3).

Otra alternativa, según lo recogido en ii), es la aproximación a través de la distribución mixtura. Como se recoge en la sección 1.3, si se verifica la condición  $0 < 1 - \frac{1}{2}sqr < 1$ , la función densidad de  $\text{vec}(\widehat{\mathbf{B}})$  se puede aproximar por la función densidad de una variable aleatoria  $\text{vec}(\mathbf{B}_E)$  definida por (1.3.20). Además, teniendo en cuenta Kollo y von Rosen (2005), la distribución de  $\text{vec}(\mathbf{B}_E)$  es una distribución matricial elíptica y, según Kollo et al. (2007), es la mixtura de una distribución normal,

$$N_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}), \quad (3.4.10)$$

con peso  $1 - sqr/2$ , y de una distribución de Kotz,

$$K_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}), \quad (3.4.11)$$

con peso  $sqr/2$ .

Así, es posible enunciar el siguiente lema.

**Lema 3.4.1** Si  $0 < 1 - \frac{1}{2}sqr < 1$ , entonces

$$\text{vec}(\mathbf{B}_E) \sim E_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, g_{\mathbf{B}_E}), \quad (3.4.12)$$

con

$$g_{\mathbf{B}_E}(u) = \left(1 - \frac{sqr}{2} + \frac{s}{2}u\right) \exp\left(-\frac{u}{2}\right) \quad (3.4.13)$$

la función generadora de densidad de la distribución elíptica matricial y constante de normalización  $C_{\mathbf{B}_E} = (2\pi)^{-qr/2}$ , donde

$$s = \frac{p - q}{n - r - (p - q) - 1}. \quad (3.4.14)$$

**Demostración:** Como se recoge en el teorema 1.3.7, si  $0 < 1 - \frac{1}{2}sqr < 1$ , entonces  $\text{vec}(\mathbf{B}_E)$  posee función de densidad. Por otro lado, la distribución de  $\text{vec}(\mathbf{B}_E)$  es una mixtura de dos componentes elípticas, una distribución normal y una distribución de Kotz dadas en (3.4.10) y (3.4.11), respectivamente. En consecuencia, por los teoremas C.2.1 y C.3.1, la distribución de  $\text{vec}(\mathbf{B}_E)$  es la mixtura de las componentes

$$E_{qr}\left(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, g_N\right) \text{ y } E_{qr}\left(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, g_K\right),$$

con pesos  $1 - sqr/2$  y  $sqr/2$ , respectivamente, donde las funciones generadoras de densidad son

$$g_N(u) = \exp\left(-\frac{u}{2}\right) \quad \text{y} \quad g_K(u) = u \exp\left(-\frac{u}{2}\right),$$

y constantes de normalización

$$C_N = (2\pi)^{-qr/2} \quad \text{y} \quad C_K = \frac{1}{qr} (2\pi)^{-qr/2},$$

respectivamente.

Aplicando el teorema C.1.6, la función de densidad de  $\text{vec}(\mathbf{B}_E)$  es

$$\begin{aligned} f_{\mathbf{B}_E}(\mathbf{t}) &= (2\pi)^{-qr/2} \left| (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \right|^{-1/2} \left[ \left(1 - \frac{sqr}{2}\right) g_N(u) + \left(\frac{s}{2}\right) g_K(u) \right] = \\ &= (2\pi)^{-qr/2} \left| (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \right|^{-1/2} \left[ \left(1 - \frac{sqr}{2} + \frac{s}{2}u\right) \exp\left(-\frac{u}{2}\right) \right] \end{aligned}$$

con  $u = [\mathbf{t} - \text{vec}(\mathbf{B})]^T \left[ (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \right]^{-1} [\mathbf{t} - \text{vec}(\mathbf{B})]$ .

En consecuencia, si la función

$$g_{\mathbf{B}_E}(u) = \left(1 - \frac{sqr}{2} + \frac{s}{2}u\right) \exp\left(-\frac{u}{2}\right)$$

es una función generadora de densidad de probabilidad, quedará demostrado el enunciado.

Por las propiedades de la función gamma, se tiene que

$$I = \int_0^\infty u^{qr/2-1} g_{\mathbf{B}_E}(u) du = \left(1 - \frac{sqr}{2}\right) I_1 + \left(\frac{s}{2}\right) I_2 = 2^{qr/2} \Gamma\left(\frac{qr}{2}\right) < \infty,$$

dado que,

$$I_1 = 2^{qr/2} \int_0^\infty t^{qr/2-1} \exp(-t) dt = 2^{qr/2} \Gamma\left(\frac{qr}{2}\right) < \infty$$

e

$$I_2 = 2^{qr/2+1} \int_0^\infty t^{qr/2} \exp(-t) dt = 2^{qr/2} qr \Gamma\left(\frac{qr}{2}\right) < \infty,$$

así, por el lema C.1.1, se concluye que la función  $g_{\mathbf{B}_E}(u)$  define la función generadora de densidad de una distribución elíptica.

Por otro lado, de (C.1.1), la constante de normalización es

$$C_{B_E} = \frac{\Gamma\left(\frac{qr}{2}\right)}{2\pi^{qr/2} \int_0^\infty u^{qr-1} g_{B_E}(u^2) du} = \frac{\Gamma\left(\frac{qr}{2}\right)}{2\pi^{qr/2} \left[\left(1 - \frac{sqr}{2}\right) I_3 + \left(\frac{s}{2}\right) I_4\right]}$$

donde

$$I_3 = 2^{qr/2-1} \int_0^\infty t^{qr/2-1} \exp(-t) dt = 2^{qr/2-1} \Gamma\left(\frac{qr}{2}\right)$$

e

$$I_4 = 2^{qr/2} \int_0^\infty t^{qr/2} \exp(-t) dt = 2^{qr/2-1} qr \Gamma\left(\frac{qr}{2}\right).$$

Simplificando, se obtiene  $C_{B_E} = (2\pi)^{-qr/2}$ . ■

Por tanto, atendiendo al lema 3.4.1, la distribución de  $\text{vec}(\widehat{\mathbf{B}})$  puede ser aproximada por la distribución mixtura elíptica de  $\text{vec}(\mathbf{B}_E)$  dada por (3.4.12). Para aplicar la distancia de Fréchet, la esperanza matemática de  $\text{vec}(\widehat{\mathbf{B}})$  se obtiene directamente de (3.4.12).

Para el cálculo de la matriz de covarianzas de la distribución

$$\text{vec}(\mathbf{B}_E) \sim E_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, g_{\mathbf{B}_E}),$$

teniendo en cuenta el teorema C.1.5 es necesario calcular  $\frac{E[R^2]}{qr}$ , donde  $qr = \text{rg}((\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1})$ . Considerando la nota C.1.2 la función densidad de la variable aleatoria  $R$  es

$$f_R(u) = C_{B_E} \frac{2\pi^{qr/2}}{\Gamma\left(\frac{qr}{2}\right)} u^{qr-1} g_R(u^2), \quad u > 0,$$

con  $g_R(u^2)$  obtenida de la función  $g_{\mathbf{B}_E}(u)$ . Entonces

$$E[R^2] = \int_0^\infty u^2 f_R(u) du = \frac{2\pi^{qr/2}}{\Gamma\left(\frac{qr}{2}\right)} \left( \left(1 - \frac{sqr}{2}\right) I_4 + \left(\frac{s}{2}\right) I_5 \right),$$

donde

$$I_4 = 2^{qr/2-1} qr \Gamma\left(\frac{qr}{2}\right)$$

e

$$I_5 = 2^{qr/2+1} \int_0^\infty t^{qr/2+1} \exp(-t) dt = 2^{qr/2} qr \left(\frac{qr}{2} + 1\right) \Gamma\left(\frac{qr}{2}\right).$$

En consecuencia,

$$\frac{E[R^2]}{qr} = s + 1 = c_1. \quad (3.4.15)$$

Así, considerando la aproximación a la distribución mixtura elíptica (3.4.12), se concluye que la esperanza matemática y la matriz de covarianzas de  $\mathbf{B}_E$  son iguales a las de  $\widehat{\mathbf{B}}$ , o sea:

- $E[\mathbf{B}_E] = E[\widehat{\mathbf{B}}] = \mathbf{B}$ .
- $\text{Var}[\mathbf{B}_E] = \text{Var}[\widehat{\mathbf{B}}] = \frac{E[R^2]}{qr} (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} = c_1 (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}$ , con  $c_1 = s + 1$  y  $s$  dado en (3.4.14).

Una vez establecida la aproximación a la distribución de  $\text{vec}(\widehat{\mathbf{B}})$ , para la aplicación de la distancia de Fréchet, como medida de influencia, sobre el EMV  $\widehat{\mathbf{B}}$  del MCC, es necesario establecer la aproximación de la distribución de  $\text{vec}(\widehat{\mathbf{B}}_{(I)})$ . Además, como se recoge en la sección 3.3, las dos distribuciones han de tener las mismas funciones generadoras de densidad  $g$ .

Como consecuencia del lema 3.4.1, es posible aproximar la distribución de  $\text{vec}(\widehat{\mathbf{B}}_{(I)})$  por la distribución de

$$\text{vec}(\mathbf{B}_{E(I)}) \sim E_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, g_{\mathbf{B}_{E(I)}})$$

donde

$$g_{\mathbf{B}_{E(I)}}(u) = (2\pi)^{-qr/2} \left( 1 - \frac{s(I)qr}{2} + \frac{s(I)}{2}u \right) \exp\left(-\frac{u}{2}\right), \quad (3.4.16)$$

si

$$0 < 1 - \frac{1}{2}s(I)qr < 1,$$

con

$$s(I) = \frac{p - q}{n - k - r - (p - q) - 1}. \quad (3.4.17)$$

En este caso,

$$\text{Var}[\mathbf{B}_{E(I)}] = c_6 (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1},$$

una vez que  $\frac{E[R_{(I)}^2]}{qr} = c_6 = s(I) + 1$ , y

$$\widehat{\text{Var}}[\mathbf{B}_{E(I)}] = \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] = \frac{c_6}{n - k - r} \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \left( \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \right)^{-1}.$$

Como se puede observar, las distribuciones de  $\text{vec}(\mathbf{B}_E)$  y  $\text{vec}(\mathbf{B}_{E(I)})$  tienen diferentes funciones generadoras de densidad  $g_{\mathbf{B}_E}(u)$  y  $g_{\mathbf{B}_{E(I)}}(u)$ , dadas en (3.4.13) y (3.4.16), respectivamente. No obstante, de (3.4.17) se puede considerar

$$s(I) = s(1 + \alpha),$$

con

$$\alpha = \frac{k}{n - r - (p - q) - (1 + k)}.$$

Como la razón entre  $g_{\mathbf{B}_E}(u)$  y  $g_{\mathbf{B}_{E(I)}}(u)$  es

$$\frac{g_{\mathbf{B}_{E(I)}}(u)}{g_{\mathbf{B}_E}(u)} = 1 - \alpha \frac{s(qr - u)}{2 - s(qr - u)}$$

donde  $\alpha = O\left(\frac{1}{n}\right)$ , entonces

$$\frac{g_{\mathbf{B}_{E(I)}}(u)}{g_{\mathbf{B}_E}(u)} \approx 1.$$

En este sentido, se puede aproximar la distribución de  $\text{vec}(\mathbf{B}_{E(I)})$  y, en consecuencia, la distribución de  $\text{vec}(\widehat{\mathbf{B}}_{(I)})$ , por la distribución elíptica

$$E_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}, g_{\mathbf{B}_E}),$$

con  $g_{\mathbf{B}_E}$ , dado por (3.4.13), y usar la versión muestral de la distancia de Fréchet como aproximación para medir la influencia del subconjunto de observaciones  $\mathbf{Y}_I$  sobre el EMV  $\widehat{\mathbf{B}}$  del MCC.

Por tanto, en este caso, la citada versión muestral de la medida de Fréchet viene dada por (3.4.7), con

- $LC_I^2(\widehat{\mathbf{B}})$  igual a la componente de localización de la aproximación normal dada por (3.4.8), y
- teniendo en cuenta (3.4.15), la componente de dispersión es

$$DC_I^2(\widehat{\mathbf{B}}) = c_1 \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] + \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right]. \quad (3.4.18)$$

En resumen, se han presentado dos aproximaciones para  $\text{vec}(\widehat{\mathbf{B}})$  y las correspondientes versiones muestrales para la medida de Fréchet asociadas a  $\widehat{\mathbf{B}}$  en el MCC. Se concluye que:

- 1) Las componentes de localización de las medidas de Fréchet son las mismas.
- 2) Como las matrices de covarianzas consideradas bajo el MCC y el MCC(I) son las mismas para las dos aproximaciones, comparando las cantidades  $DC_I^2(\widehat{\mathbf{B}})$ , dadas por (3.4.9) y (3.4.18), para la aproximación a la distribución normal y para la distribución mixtura elíptica, respectivamente, se observa que estas difieren en la constante  $\frac{E[R^2]}{qr}$ .
- 3) Por lo recogido en 1) y 2), las medidas de Fréchet asociadas a las dos aproximaciones difieren en una constante asociada a la componente de dispersión  $DC_I^2(\widehat{\mathbf{B}})$ . Tal hecho no tiene consecuencias relevantes en términos del diagnóstico de influencia, ya que las conclusiones obtenidas desde el punto de vista práctico para las dos aproximaciones coinciden.

Así, en esta memoria, se presenta el desarrollo teórico de la aplicación de la distancia de Fréchet como medida de influencia cuando la distribución de  $\text{vec}(\widehat{\mathbf{B}})$  es aproximada por la distribución normal matricial (3.4.3). En los dos teoremas siguientes se presentan expresiones de las componentes  $LC_I^2(\widehat{\mathbf{B}})$  y  $DC_I^2(\widehat{\mathbf{B}})$  de la medida de Fréchet para el MCC.

**Teorema 3.4.1** *En el MCC, la componente de localización de la medida de Fréchet  $\delta_I(\widehat{\mathbf{B}})$  es*

$$LC_I^2(\widehat{\mathbf{B}}) = \text{tr} [\mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{K}_I], \quad (3.4.19)$$

con las matrices  $\mathbf{E}_I$ ,  $\mathbf{K}_I$  y  $\mathbf{V}_I$  definidas en (2.3.6), (2.3.9) y (2.3.10), respectivamente.

**Demostración:** Teniendo en cuenta la relación entre  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , presentada en el teorema 2.3.1, de (3.4.8) se obtiene

$$\begin{aligned} LC_I^2(\widehat{\mathbf{B}}) &= \left\| \widehat{\mathbf{B}} - \widehat{\mathbf{B}}_{(I)} \right\|^2 = \\ &= \left\| \widehat{\mathbf{B}} - \left( \widehat{\mathbf{B}} - \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \right) \right\|^2 = \\ &= \left\| \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \right\|^2 = \\ &= \text{tr} \left[ \left( \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \right)^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \right] = \\ &= \text{tr} [\mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{K}_I], \end{aligned}$$

la última igualdad es obtenida de la relación entre la norma y la traza de una matriz dada en (B.0.6). ■

A partir del teorema 3.4.1, se puede determinar la componente de localización de la medida de Fréchet asociada a la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.4.1** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de localización de la medida de Fréchet  $\delta_i(\widehat{\mathbf{B}})$  es*

$$LC_i^2(\widehat{\mathbf{B}}) = \frac{1}{v_{ii}^2} \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i, \quad (3.4.20)$$

con  $v_{ii}$  y  $\mathbf{k}_i$  definidos por (2.3.13) y (2.3.12), respectivamente, y  $\mathbf{e}_i$  dado en (2.3.14) es la  $i$ -ésima columna de la matriz de residuos  $\mathbf{E}$  definida en (2.2.2).

Los resultados que se recogen a continuación son útiles para el cálculo de la componente de dispersión de la medida de Fréchet para el MCC.

**Lema 3.4.2** *Sean  $\mathbf{Z}$  y  $\mathbf{Z}_{(I)}$  las matrices del diseño entre individuos del MCC y del MCC(I), respectivamente, entonces*

$$(\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} = (\mathbf{Z} \mathbf{Z}^T)^{-1} + (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1},$$

con  $\mathbf{H}_I$  dada en (2.3.7).

**Demostración:** La demostración es directa, teniendo en cuenta que  $\mathbf{Z} \mathbf{Z}^T = \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T + \mathbf{Z}_I \mathbf{Z}_I^T$ , y aplicando la igualdad (B.0.2). ■

**Lema 3.4.3** Sean  $\mathbf{S}$  y  $\mathbf{S}_{(I)}$  las matrices definidas por (1.3.5) y (2.3.2), respectivamente, para el MCC y MCC(I). Entonces

$$\left(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X}\right)^{-1} = \mathbf{S}_X - \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X.$$

**Demostración:** Usando la relación entre  $\mathbf{S}_{(I)}^{-1}$  y  $\mathbf{S}^{-1}$  definida en (2.3.5) y las igualdades (B.0.3) y (2.3.10), se obtiene

$$\begin{aligned} \left(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X}\right)^{-1} &= \left(\mathbf{X}^T \left(\mathbf{S}^{-1} + \mathbf{S}^{-1} \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)^{-1} \mathbf{E}_I^T \mathbf{S}^{-1}\right) \mathbf{X}\right)^{-1} = \\ &= \left(\mathbf{X}^T \mathbf{S}^{-1} \mathbf{X} + \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X}\right)^{-1} = \\ &= \mathbf{S}_X - \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \\ &\quad \left( (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I) + \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \right)^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X = \\ &= \mathbf{S}_X - \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X. \end{aligned}$$

■

Dados los dos lemas anteriores, en el siguiente teorema se presenta la componente de dispersión de la medida de Fréchet asociada al EMV de la matriz de coeficientes de regresión del MCC.

**Teorema 3.4.2** En el MCC, la componente de dispersión de la medida de Fréchet  $\delta_I(\widehat{\mathbf{B}})$  es

$$\begin{aligned} DC_I^2(\widehat{\mathbf{B}}) &= \frac{wc_1}{n-r} \text{tr}[\mathbf{S}_X] + \frac{c_6}{n-k-r} \left( w + \text{tr} [(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{Z}_I] \right) \\ &\quad \left( \text{tr}[\mathbf{S}_X] - \text{tr}[\mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I] \right) - \\ &\quad - 2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \text{tr} \left[ (\mathbf{S}_X^2 - \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X)^{1/2} \right] \\ &\quad \text{tr} \left[ ((\mathbf{Z} \mathbf{Z}^T)^{-2} + (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1})^{1/2} \right], \end{aligned} \tag{3.4.21}$$

con  $c_1$  y  $c_6$  dados en (1.3.18) y (2.3.3), respectivamente.

**Demostración:** Por las propiedades de la traza de matrices, se puede descomponer  $DC_I^2(\widehat{\mathbf{B}})$ , dada en (3.4.9), de la siguiente manera,

$$\begin{aligned} DC_I^2(\widehat{\mathbf{B}}) &= \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] + \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right] = \\ &= \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] \right] + \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right] - 2 \text{tr} \left[ \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right] = \\ &= DC1(\widehat{\mathbf{B}}) + DC2(\widehat{\mathbf{B}}) + DC3(\widehat{\mathbf{B}}). \end{aligned}$$

Por tanto, usando los estimadores  $\widehat{\text{Var}}[\widehat{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}]$  dados en (3.4.4) y (3.4.6), respectivamente, y la relación entre la traza y el producto de Kronecker recogida por el teorema B.0.5, se tiene

$$DC1(\widehat{\mathbf{B}}) = \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}] \right] = \frac{c_1}{n-r} \text{tr} [(\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes \mathbf{S}_X] = \frac{c_1}{n-r} \text{tr}[(\mathbf{Z} \mathbf{Z}^T)^{-1}] \text{tr}[\mathbf{S}_X],$$

$$\begin{aligned} DC2(\widehat{\mathbf{B}}) &= \text{tr} \left[ \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right] = \frac{c_6}{n-k-r} \text{tr} \left[ \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \right] = \\ &= \frac{c_6}{n-k-r} \text{tr} \left[ \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \right] \text{tr} \left[ \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \right]. \end{aligned}$$

Por otro lado, aplicando la propiedad (vii) del mismo teorema, se obtiene

$$\begin{aligned} DC3(\widehat{\mathbf{B}}) &= -2 \text{tr} \left[ \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}] \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] \right)^{1/2} \right] = \\ &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \\ &\quad \text{tr} \left[ \left( \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \otimes \mathbf{S}_{\mathbf{X}} \right) \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \right]^{1/2} = \\ &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \\ &\quad \text{tr} \left[ \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \otimes \mathbf{S}_{\mathbf{X}} \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \right]^{1/2}. \end{aligned}$$

A continuación, se calcula separadamente cada una de estas cantidades. Para el cálculo de  $DC1(\widehat{\mathbf{B}})$ , como

$$\left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} = \text{diag} \left[ \frac{1}{n_1}, \dots, \frac{1}{n_r} \right],$$

entonces

$$\text{tr} \left[ \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \right] = \sum_{j=1}^r \frac{1}{n_j} = w, \quad (3.4.22)$$

por tanto,

$$DC1(\widehat{\mathbf{B}}) = \frac{w c_1}{n-r} \text{tr} [\mathbf{S}_{\mathbf{X}}]. \quad (3.4.23)$$

Se observa que  $DC1(\widehat{\mathbf{B}})$  es una constante que no depende del subconjunto  $\mathbf{Y}_I$  de las observaciones omitidas. Para el cálculo de  $DC2(\widehat{\mathbf{B}})$ , del lema 3.4.2, se obtiene

$$\begin{aligned} \text{tr} \left[ \left( \mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T \right)^{-1} \right] &= \text{tr} \left[ \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} + \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \mathbf{Z}_I \left( \mathbf{I}_k - \mathbf{H}_I \right)^{-1} \mathbf{Z}_I^T \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \right] = \quad (3.4.24) \\ &= \text{tr} \left[ \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \right] + \text{tr} \left[ \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \mathbf{Z}_I \left( \mathbf{I}_k - \mathbf{H}_I \right)^{-1} \mathbf{Z}_I^T \left( \mathbf{Z} \mathbf{Z}^T \right)^{-1} \right] = \\ &= w + \text{tr} \left[ \left( \mathbf{I}_k - \mathbf{H}_I \right)^{-1} \mathbf{Z}_I^T \left( \mathbf{Z} \mathbf{Z}^T \right)^{-2} \mathbf{Z}_I \right], \end{aligned}$$

con  $w$  dado por (3.4.22). Por otro lado, del lema 3.4.3,

$$\text{tr} \left[ \left( \mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X} \right)^{-1} \right] = \text{tr} [\mathbf{S}_{\mathbf{X}}] - \text{tr} \left[ \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \right].$$

Así

$$\begin{aligned} DC2(\widehat{\mathbf{B}}) &= \frac{c_6}{n-k-r} \left( w + \text{tr} \left[ \left( \mathbf{I}_k - \mathbf{H}_I \right)^{-1} \mathbf{Z}_I^T \left( \mathbf{Z} \mathbf{Z}^T \right)^{-2} \mathbf{Z}_I \right] \right) \quad (3.4.25) \\ &\quad \left( \text{tr} [\mathbf{S}_{\mathbf{X}}] - \text{tr} \left[ \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \right] \right). \end{aligned}$$



Finalmente, para el cálculo de  $DC3(\widehat{\mathbf{B}})$ , como las matrices incluidas en la traza son definidas positivas, usando la propiedad (ix) del teorema B.0.5, se obtiene

$$\begin{aligned}
 DC3(\widehat{\mathbf{B}}) &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \\
 &\quad \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \right)^{1/2} \right] \text{tr} \left[ \mathbf{S}_{\mathbf{X}} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \right]^{1/2} = \\
 &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \\
 &\quad \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-2} + (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \right)^{1/2} \right] \\
 &\quad \text{tr} \left[ (\mathbf{S}_{\mathbf{X}}^2 - \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}})^{1/2} \right], \tag{3.4.26}
 \end{aligned}$$

la última igualdad es obtenida teniendo en cuenta los lemas 3.4.2 y 3.4.3. ■

A través del teorema 3.4.2, se determina la componente de dispersión de la medida de Fréchet para la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.4.2** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de dispersión de la medida de Fréchet  $\delta_i(\widehat{\mathbf{B}})$  es*

$$\begin{aligned}
 DC_i^2(\widehat{\mathbf{B}}) &= \frac{w c_1}{n-r} \text{tr} [\mathbf{S}_{\mathbf{X}}] + \\
 &\quad + \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( \text{tr} [\mathbf{S}_{\mathbf{X}}] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right) - \\
 &\quad - 2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \text{tr} \left[ \left( \mathbf{S}_{\mathbf{X}}^2 - \frac{1}{v_{ii}} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}} \right)^{1/2} \right],
 \end{aligned}$$

donde

$$c_9 = \frac{n-r-2}{n-r-(p-q)-2}, \tag{3.4.27}$$

$$w_i = \frac{p_{ii}}{(1-p_{ii})^{1/2}} + \sum_{j \neq i} \frac{1}{n_j} \tag{3.4.28}$$

y  $p_{ii}$ , dado en (2.3.15), el elemento  $i$ -ésimo diagonal de la matriz de proyección  $\mathbf{P}_{\mathbf{Z}^T}$ .

**Demostración:** Se obtiene usando las expresiones  $DC1(\widehat{\mathbf{B}})$ ,  $DC2(\widehat{\mathbf{B}})$  y  $DC3(\widehat{\mathbf{B}})$  dadas en (3.4.23), (3.4.25) y (3.4.26), respectivamente, y considerando la descomposición

$$DC_i^2(\widehat{\mathbf{B}}) = DC1_i + DC2_i + DC3_i. \tag{3.4.29}$$

Como se ha indicado en el teorema 3.4.2,  $DC1$  no depende de las observaciones omitidas, luego  $DC1_i$  no depende de la  $i$ -ésima observación, siendo

$$DC1_i = DC1 = \frac{w c_1}{n-r} \text{tr} [\mathbf{S}_{\mathbf{X}}]. \tag{3.4.30}$$

El segundo término de la descomposición (3.4.29) se obtiene teniendo en cuenta que para  $k = 1$  e  $I = \{i\}$ ,  $\mathbf{H}_I = \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i = p_{ii}$ ,  $(\mathbf{I}_k - \mathbf{H}_I)^{-1} = (1 - p_{ii})^{-1}$  y  $\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i = p_{ii}^2$ . Usando estas cantidades en (3.4.25), se obtiene

$$\begin{aligned} DC2_i &= \frac{c_9}{n-r-1} (w + \text{tr} [(1-p_{ii})^{-1} \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i]) \\ &\quad (\text{tr} [\mathbf{S}_X] - \text{tr} [v_{ii}^{-1} \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i]) = \\ &= \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( \text{tr} [\mathbf{S}_X] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right). \end{aligned} \quad (3.4.31)$$

Por último, de la expresión (3.4.26), resulta que

$$\begin{aligned} DC3_i &= -2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \\ &\quad \text{tr} \left[ \left( \mathbf{S}_X^2 - \frac{1}{v_{ii}} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \right)^{1/2} \right]. \end{aligned} \quad (3.4.32)$$

Para obtener  $w_i$  se ha de tener en cuenta que

$$(\mathbf{Z}\mathbf{Z}^T)^{-1} = \text{diag} \left[ \frac{1}{n_1}, \dots, \frac{1}{n_r} \right] \quad \text{y} \quad (\mathbf{Z}\mathbf{Z}^T)^{-2} = \text{diag} \left[ \frac{1}{n_1^2}, \dots, \frac{1}{n_r^2} \right],$$

donde  $n_j$  es el tamaño del grupo  $j$ -ésimo,  $j = 1, \dots, r$ , por lo que

$$(\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} = \text{diag} \left[ 0, \dots, 0, \frac{1}{n_j^3}, 0, \dots, 0 \right],$$

es decir, una matriz compuesta por ceros con excepción de la posición de la diagonal principal que corresponde al grupo  $j$ -ésimo del cual fue omitido la  $i$ -ésima observación, y toma el valor  $\frac{1}{n_j^3}$ , pues  $\mathbf{z}_i^T = (0, \dots, 0, 1, 0, \dots, 0)$  donde el valor 1 está en la posición correspondiente al grupo donde se ha eliminado la  $i$ -ésima observación. Por otro lado,  $p_{ii} = \frac{1}{n_j}$ , para cualquier observación  $i$  que pertenezca al  $j$ -ésimo grupo, y así

$$\begin{aligned} \frac{1}{1-p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} &= \text{diag} \left[ 0, \dots, 0, \frac{1}{(1-p_{ii})n_j^3}, 0, \dots, 0 \right] = \\ &= \text{diag} \left[ 0, \dots, 0, \frac{p_{ii}^3}{1-p_{ii}}, 0, \dots, 0 \right]. \end{aligned}$$

Como la matriz

$$(\mathbf{Z}\mathbf{Z}^T)^{-2} + \frac{1}{1-p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1}$$

es diagonal,

$$\left( (\mathbf{Z}\mathbf{Z}^T)^{-2} + \frac{1}{1-p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \right)^{1/2}$$

es la matriz diagonal con elementos diagonales la raíz cuadrada de los elementos diagonales de la matriz original. Entonces se puede hacer la simplificación

$$\begin{aligned} w_i &= \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-2} + \frac{1}{1-p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \right)^{1/2} \right] = \\ &= \frac{p_{ii}}{(1-p_{ii})^{1/2}} + \sum_{j \neq i} \frac{1}{n_j}, \end{aligned}$$

donde  $\sum_{j \neq i} \frac{1}{n_j}$  quiere decir que se suman los  $\frac{1}{n_j}$  que no corresponden al grupo de la observación  $i$ -ésima. Así, la cantidad  $w_i$  es constante en cada grupo. ■

Usando los teoremas 3.4.1 y 3.4.2, se concluye que la medida Fréchet para evaluar la influencia del subconjunto de observaciones  $\mathbf{Y}_I$  en la distribución del EMV  $\hat{\mathbf{B}}$  para el MCC es

$$\delta_I(\hat{\mathbf{B}}) = \left[ LC_I^2(\hat{\mathbf{B}}) + DC_I^2(\hat{\mathbf{B}}) \right]^{1/2},$$

donde la componente de localización  $LC_I^2(\hat{\mathbf{B}})$  viene dada por (3.4.19) y la componente de dispersión  $DC_I^2(\hat{\mathbf{B}})$  dada en (3.4.21).

Como consecuencia, de los corolarios 3.4.1 y 3.4.2, resulta que la medida de Fréchet para analizar la influencia de la observación  $i$ -ésima de  $\mathbf{Y}$  sobre la distribución del EMV  $\hat{\mathbf{B}}$  del MCC es

$$\begin{aligned} \delta_i(\hat{\mathbf{B}}) &= \left[ LC_i^2(\hat{\mathbf{B}}) + DC_i^2(\hat{\mathbf{B}}) \right]^{1/2} = \tag{3.4.33} \\ &= \left[ \frac{1}{v_{ii}^2} \mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{k}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i + \right. \\ &\quad \left. + \frac{wc_1}{n-r} \text{tr}[\mathbf{S}_{\mathbf{X}}] + \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( \text{tr}[\mathbf{S}_{\mathbf{X}}] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right) - \right. \\ &\quad \left. - 2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \text{tr} \left[ \left( \mathbf{S}_{\mathbf{X}}^2 - \frac{1}{v_{ii}} \mathbf{S}_{\mathbf{X}}^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_{\mathbf{X}} \right)^{1/2} \right] \right]^{1/2}. \end{aligned}$$

Un valor elevado de  $\delta_i$  indica que la correspondiente observación es influyente en la estimación de  $\mathbf{B}$ , una vez que su omisión provoca una considerable variación en la distribución de  $\hat{\mathbf{B}}$ . En el caso del estudio de la influencia individual, la representación gráfica de  $LC_i^2(\hat{\mathbf{B}})$  versus  $DC_i^2(\hat{\mathbf{B}})$ , para  $i = 1, \dots, n$ , puede ser muy ilustrativa para determinar el sentido de la influencia que ejerce cada observación.

### 3.4.1. Influencia en combinaciones lineales de $\hat{\mathbf{B}}$

La distancia de Fréchet puede ser extendida como medida de influencia de las observaciones parciales  $\mathbf{Y}_I$  sobre las distribuciones de los estimadores de combinaciones lineales de la matriz de coeficientes de regresión  $\mathbf{B}$ , o sea, en los estimadores de  $\Psi = \mathbf{D}\mathbf{B}\mathbf{L}$ , donde los rangos de las matrices conocidas  $\mathbf{D}_{(s \times q)}$  y  $\mathbf{L}_{(r \times l)}$  vienen dados por  $\text{rg}(\mathbf{D}) = s \leq q$  y  $\text{rg}(\mathbf{L}) = l \leq r$ , respectivamente. Teniendo en cuenta Kollo y von Rosen (2005), el EMV de  $\Psi$  es

$$\hat{\Psi} = \mathbf{D}_{(s \times q)} \hat{\mathbf{B}}_{(q \times r)} \mathbf{L}_{(r \times l)}. \tag{3.4.34}$$

En la sección anterior, para la aplicación de la distancia de Fréchet como medida de influencia, la distribución del EMV  $\text{vec}(\widehat{\mathbf{B}})$  fue aproximada por la distribución normal

$$N_{qr}(\text{vec}(\mathbf{B}), c_1(\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}),$$

teniendo en cuenta el teorema C.1.2 resulta que  $\text{vec}(\mathbf{D}\widehat{\mathbf{B}}\mathbf{L})$  puede ser aproximada por la distribución normal

$$N_{sl}(\text{vec}(\mathbf{D}\mathbf{B}\mathbf{L}), c_1 \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{D}^T),$$

donde

$$\widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}\mathbf{L}] = \frac{c_1}{n-r} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D}\mathbf{S}_X \mathbf{D}^T, \quad (3.4.35)$$

con  $c_1$  y  $\mathbf{S}_X$  dadas en (1.3.18) y (1.3.10), respectivamente.

Para el MCC( $I$ ), de la aproximación a la distribución de  $\text{vec}(\widehat{\mathbf{B}}_{(I)})$ , dada por (3.4.5), considerando el EMV  $\widehat{\boldsymbol{\Psi}}_{(I)} = \mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}$ , la distribución de  $\text{vec}(\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L})$  puede ser aproximada por

$$N_{sl}(\text{vec}(\mathbf{D}\mathbf{B}\mathbf{L}), c_6 \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D}(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{D}^T),$$

donde

$$\widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}] = \frac{c_6}{n-k-r} \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D}(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T, \quad (3.4.36)$$

con  $c_6$  dado por (2.3.3).

El cálculo de la medida de Fréchet para las combinaciones lineales de  $\widehat{\mathbf{B}}$  es análogo al de la sección anterior. En los siguientes teoremas se presentan expresiones generales de las componentes de localización y de dispersión de la medida de Fréchet asociada a las combinaciones lineales  $\widehat{\boldsymbol{\Psi}}$  para  $\mathbf{Y}_I$ , o sea, para  $\delta_I(\widehat{\boldsymbol{\Psi}}) = \delta_I(\mathbf{D}\widehat{\mathbf{B}}\mathbf{L})$ .

**Teorema 3.4.3** *En el MCC, la componente de localización de la medida de Fréchet  $\delta_I(\widehat{\boldsymbol{\Psi}})$  es*

$$LC_I^2(\widehat{\boldsymbol{\Psi}}) = \text{tr} \left[ \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \right. \\ \left. \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \right],$$

donde  $\mathbf{E}_I$ ,  $\mathbf{K}_I$  y  $\mathbf{V}_I$  están definidas en (2.3.6), (2.3.9) y (2.3.10), respectivamente.

**Demostración:** Teniendo en cuenta (3.4.8) y usando la relación (2.3.8) entre  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$ , dada por el teorema 2.3.1, se tiene

$$LC_I^2(\widehat{\boldsymbol{\Psi}}) = \left\| \mathbf{D}\widehat{\mathbf{B}}\mathbf{L} - \mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L} \right\|^2 = \left\| \mathbf{D}(\widehat{\mathbf{B}} - \widehat{\mathbf{B}}_{(I)})\mathbf{L} \right\|^2 = \\ = \left\| \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right\|^2 = \\ = \text{tr} \left[ (\mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^T \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right] = \\ = \text{tr} \left[ \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{K}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{K}_I \right],$$

la última igualdad es obtenida de la relación entre la norma y la traza de una matriz definida por (B.0.6). ■

A partir del teorema 3.4.3, se obtiene la componente de localización para la medida de Fréchet asociada a la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.4.3** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de localización de  $\delta_i(\hat{\Psi})$  es*

$$LC_i^2(\hat{\Psi}) = \frac{1}{v_{ii}^2} \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i. \quad (3.4.37)$$

En el siguiente teorema se presenta la componente de dispersión de la medida de Fréchet asociada a  $\hat{\Psi}$ .

**Teorema 3.4.4** *En el MCC, la componente de dispersión de la medida de Fréchet  $\delta_I(\hat{\Psi})$  es*

$$\begin{aligned} DC_I^2(\hat{\Psi}) &= \frac{c_1}{n-r} \text{tr} [\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L}] \text{tr} [\mathbf{D} \mathbf{S}_X \mathbf{D}^T] + \\ &+ \frac{c_6}{n-k-r} (\text{tr} [\mathbf{D} \mathbf{S}_X \mathbf{D}^T] - \text{tr} [\mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T]) \\ &(\text{tr} [\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L}] + \text{tr} [\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L}]) - \\ &- 2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \\ &\text{tr} \left[ ((\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L})^2 + \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L})^{1/2} \right] \\ &\text{tr} \left[ ((\mathbf{D} \mathbf{S}_X \mathbf{D}^T)^2 - \mathbf{D} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T)^{1/2} \right]. \end{aligned}$$

**Demostración:** Se aplica la misma descomposición usada en la sección anterior, o sea,

$$\begin{aligned} DC_I^2(\mathbf{D} \hat{\mathbf{B}} \mathbf{L}) &= \text{tr} \left[ \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}} \mathbf{L}] + \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}}_{(I)} \mathbf{L}] - 2 \left( \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}} \mathbf{L}] \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}}_{(I)} \mathbf{L}] \right)^{1/2} \right] = \\ &= \text{tr} \left[ \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}} \mathbf{L}] \right] + \text{tr} \left[ \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}}_{(I)} \mathbf{L}] \right] - 2 \text{tr} \left[ \left( \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}} \mathbf{L}] \widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}}_{(I)} \mathbf{L}] \right)^{1/2} \right] = \\ &= DC1(\hat{\Psi}) + DC2(\hat{\Psi}) + DC3(\hat{\Psi}). \end{aligned}$$

Así, usando los estimadores  $\widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}} \mathbf{L}]$  y  $\widehat{\text{Var}}[\mathbf{D} \hat{\mathbf{B}}_{(I)} \mathbf{L}]$  dados por (3.4.35) y (3.4.36), respectivamente, y las propiedades del producto de Kronecker, se tiene

$$\begin{aligned} DC1(\hat{\Psi}) &= \text{tr} \left[ \frac{c_1}{n-r} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D} \mathbf{S}_X \mathbf{D}^T \right] = \\ &= \frac{c_1}{n-r} \text{tr} [\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L}] \text{tr} [\mathbf{D} \mathbf{S}_X \mathbf{D}^T], \end{aligned} \quad (3.4.38)$$

obsérvese que, también en este caso, esta cantidad no depende del subconjunto de observaciones omitidas  $\mathbf{Y}_I$ , es decir, es una constante. Por otro lado,

$$\begin{aligned} DC2(\hat{\Psi}) &= \text{tr} \left[ \frac{c_6}{n-k-r} \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right] = \\ &= \frac{c_6}{n-k-r} \text{tr} [\mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L}] \text{tr} [\mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T] = \\ &= \frac{c_6}{n-k-r} (\text{tr} [\mathbf{D} \mathbf{S}_X \mathbf{D}^T] - \text{tr} [\mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I]) \\ &(\text{tr} [\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L}] + \text{tr} [(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I]), \end{aligned} \quad (3.4.39)$$

dado que, del lema 3.4.2,

$$\begin{aligned} \text{tr} \left[ \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \right] &= \text{tr} \left[ \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \right] + \\ &+ \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I \right] \end{aligned} \quad (3.4.40)$$

y, del lema 3.4.3,

$$\begin{aligned} \text{tr} \left[ \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right] &= \text{tr} \left[ \mathbf{D} \mathbf{S}_X \mathbf{D}^T \right] - \\ &- \text{tr} \left[ \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \right]. \end{aligned}$$

Finalmente, aplicando nuevamente los lemas anteriores,

$$\begin{aligned} DC3(\widehat{\Psi}) &= -2 \text{tr} \left[ \left( \left( \frac{c_1}{n-r} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D} \mathbf{S}_X \mathbf{D}^T \right) \right. \right. \\ &\quad \left. \left. \left( \frac{c_6}{n-k-r} \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right) \right)^{1/2} \right] = \\ &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \text{tr} \left[ \left( \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \right. \right. \\ &\quad \left. \left. \otimes \mathbf{D} \mathbf{S}_X \mathbf{D}^T \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right)^{1/2} \right] = \\ &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \text{tr} \left[ \left( \left( (\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L})^2 + \right. \right. \right. \quad (3.4.41) \\ &\quad \left. \left. \left. + \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \right) \otimes \right. \right. \\ &\quad \left. \left. \otimes \left( (\mathbf{D} \mathbf{S}_X \mathbf{D}^T)^2 - \mathbf{D} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \right) \right)^{1/2} \right] = \\ &= -2 \left( \frac{c_1 c_6}{(n-r)(n-k-r)} \right)^{1/2} \text{tr} \left[ \left( (\mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L})^2 + \right. \right. \\ &\quad \left. \left. + \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \right)^{1/2} \right] \\ &\quad \text{tr} \left[ \left( (\mathbf{D} \mathbf{S}_X \mathbf{D}^T)^2 - \mathbf{D} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \right)^{1/2} \right], \end{aligned}$$

lo que completa la demostración. ■

A partir del teorema 3.4.4 se obtiene la componente de dispersión de la medida de Fréchet asociada a la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.4.4** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de dispersión de  $\delta_i(\widehat{\Psi})$  es*

$$DC_i^2(\widehat{\Psi}) = DC1_i + DC2_i + DC3_i, \quad (3.4.42)$$

con

$$DC1_i = \frac{c_1}{n-r} \text{tr} \left[ \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \right] \text{tr} \left[ \mathbf{D} \mathbf{S}_X \mathbf{D}^T \right], \quad (3.4.43)$$

$$DC2_i = \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] + \frac{\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i}{1-p_{ii}} \right) \left( \text{tr} [\mathbf{D}\mathbf{S}_X \mathbf{D}^T] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right) \quad (3.4.44)$$

y

$$DC3_i = -2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} \left[ \text{tr} \left[ \left( (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^2 + \frac{1}{1-p_{ii}} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right)^{1/2} \right] \right. \\ \left. \text{tr} \left[ \left( (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^2 - \frac{1}{v_{ii}} \mathbf{D}\mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \right)^{1/2} \right] \right], \quad (3.4.45)$$

donde  $c_9$  dada en (3.4.27).

**Demostración:** Se utilizan las expresiones  $DC1(\widehat{\Psi})$ ,  $DC2(\widehat{\Psi})$  y  $DC3(\widehat{\Psi})$  dadas por (3.4.38), (3.4.39) y (3.4.41), respectivamente. Como se recoge en la demostración del teorema 3.4.4,  $DC1(\widehat{\Psi})$  no depende del conjunto de observaciones omitidas, por tanto  $DC1_i = DC1(\widehat{\Psi})$ , tampoco depende de la  $i$ -ésima observación.

De  $DC2(\widehat{\Psi})$ , resulta que

$$DC2_i = \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{D}\mathbf{S}_X \mathbf{D}^T] - \text{tr} [v_{ii}^{-1} \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i] \right) \\ \left( \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] + \text{tr} \left[ (1-p_{ii})^{-1} \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \right] \right) = \\ = \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{D}\mathbf{S}_X \mathbf{D}^T] - \frac{1}{v_{ii}} \text{tr} [\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i] \right) \\ \left( \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] + \frac{1}{1-p_{ii}} \text{tr} [\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i] \right) = \\ = \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] + \frac{\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i}{1-p_{ii}} \right) \\ \left( \text{tr} [\mathbf{D}\mathbf{S}_X \mathbf{D}^T] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T \mathbf{D} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right).$$

Finalmente, de forma análoga, se obtiene  $DC3_i$  a partir de  $DC3(\widehat{\Psi})$ . ■

Usando los corolarios 3.4.3 y 3.4.4, se obtiene la expresión general para la distancia de Fréchet como medida de influencia asociada a las combinaciones lineales  $\widehat{\Psi} = \mathbf{D}\widehat{\mathbf{B}}\mathbf{L}$  para la  $i$ -ésima observación de  $\mathbf{Y}$  en el MCC,

$$\delta_i(\widehat{\Psi}) = \left[ LC_i^2(\widehat{\Psi}) + DC_i^2(\widehat{\Psi}) \right]^{1/2},$$

donde  $LC_i^2(\widehat{\Psi})$  y  $DC_i^2(\widehat{\Psi})$  están dadas por (3.4.37) y (3.4.42), respectivamente. Diferentes combinaciones lineales de  $\widehat{\mathbf{B}}$  requieren la imposición de algunas restricciones a  $\mathbf{D}$  y  $\mathbf{L}$ . Tres de esas restricciones se recogen en los siguientes casos. Otras posibles elecciones de dichas matrices se recogen, por ejemplo, en Kshirsagar y Smith (1995).

**Nota 3.4.1** La expresión general de la medida de Fréchet para  $\widehat{\Psi} = \mathbf{D}\widehat{\mathbf{B}}\mathbf{L}$  incluye la medida de Fréchet para  $\widehat{\mathbf{B}}$ , sin más que considerar  $\mathbf{D} = \mathbf{I}_q$  y  $\mathbf{L} = \mathbf{I}_r$ .

### 3.4.1.1. Influencia en las filas de $\widehat{\mathbf{B}}$

Para el estudio de la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $j$ -ésima fila de  $\widehat{\mathbf{B}}$ , se puede escoger  $\mathbf{L} = \mathbf{I}_r$  y  $\mathbf{D} = \mathbf{d}_j^T$ , donde  $\mathbf{d}_j \in \mathbb{R}^q$  es un vector de dimensión  $q$  con todas sus componentes nulas salvo la  $j$ -ésima componente igual a la unidad. Así, (3.4.34) se reduce a

$$\widehat{\Psi} = \mathbf{d}_j^T{}_{(1 \times q)} \widehat{\mathbf{B}}_{(q \times r)},$$

y, de los corolarios 3.4.3 y 3.4.4, se obtiene la medida de Fréchet para evaluar la influencia de  $\mathbf{y}_i$  en la  $j$ -ésima fila de  $\widehat{\mathbf{B}}$  para el MCC,

$$\delta_{i(j)}(\widehat{\mathbf{B}}) = \left[ LC_{i(j)}^2(\widehat{\mathbf{B}}) + DC_{i(j)}^2(\widehat{\mathbf{B}}) \right]^{1/2}. \quad (3.4.46)$$

De (3.4.37), la componente de localización es

$$LC_{i(j)}^2(\widehat{\mathbf{B}}) = \frac{1}{v_{ii}^2} \mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{k}_i (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2, \quad (3.4.47)$$

dado que  $\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \in \mathbb{R}$ . Y de (3.4.42), la componente de dispersión es

$$DC_{i(j)}^2(\widehat{\mathbf{B}}) = DC1_{i(j)} + DC2_{i(j)} + DC3_{i(j)},$$

obteniéndose las siguientes expresiones para los tres sumandos.

De (3.4.43),

$$DC1_{i(j)} = \frac{c_1}{n-r} \text{tr} [(\mathbf{Z}\mathbf{Z}^T)^{-1}] \text{tr} [\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j] = \frac{c_1 w}{n-r} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j, \quad (3.4.48)$$

con  $w$  dada en (3.4.22).

De (3.4.44),

$$\begin{aligned} DC2_{i(j)} &= \frac{c_9}{n-r-1} \left( \text{tr} [(\mathbf{Z}\mathbf{Z}^T)^{-1}] + \frac{\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i}{1-p_{ii}} \right) \\ &\quad \left( \text{tr} [\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j] - \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right) = \\ &= \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j - \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right), \end{aligned} \quad (3.4.49)$$

dado que  $\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{z}_i = p_{ii}^2$ .

Y de (3.4.45),

$$\begin{aligned} DC3_{i(j)} &= -2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \\ &\quad \left( (\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j)^2 - \frac{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right)^{1/2}, \end{aligned} \quad (3.4.50)$$

con  $c_1$ ,  $c_9$  y  $w_i$  dados por (1.3.18), (3.4.27) y (3.4.28), respectivamente.



**Nota 3.4.2** La componente de dispersión  $DC_{i(j\cdot)}^2$  se puede expresar como función de la componente de localización  $LC_{i(j\cdot)}^2$ . De (3.4.47), se obtiene que

$$\frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} = \frac{v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_{i(j\cdot)}^2,$$

por lo que

$$\begin{aligned} DC_{i(j\cdot)}^2 &= \frac{c_1 w}{n-r} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j + \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j - \frac{v_{ii} LC_{i(j\cdot)}^2}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} \right) - \\ &- 2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \left( (\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j)^2 - \frac{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j v_{ii} LC_{i(j\cdot)}^2}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} \right)^{1/2}. \end{aligned}$$

**Nota 3.4.3** En el MCC, para el estudio de la influencia asociada a las filas de  $\widehat{\mathbf{B}}$ , la medida de Fréchet,  $\delta_{i(j\cdot)}$ , puede expresarse como función de la medida de Cook,  $DL_{i(j\cdot)}$ . De hecho, considerando

$$\Delta_i = \frac{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i}, \quad (3.4.51)$$

se tiene

$$\frac{v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i} = \frac{v_{ii}}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i} \frac{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i} = T_i \Delta_i^{-1},$$

donde  $T_i$ , dado en (2.4.2), es un estadístico para determinar si la  $i$ -ésima observación de  $\mathbf{Y}$  es outlier. Por tanto, de  $DL_{i(j\cdot)}$ , dada en (2.4.9), resulta

$$\frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} = \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \frac{v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i} r DL_{i(j\cdot)} = \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \frac{r T_i}{\Delta_i} DL_{i(j\cdot)}.$$

Así,

$$LC_{i(j\cdot)}^2 = \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i \frac{r T_i}{v_{ii} \Delta_i} DL_{i(j\cdot)},$$

y

$$\begin{aligned} DC_{i(j\cdot)}^2 &= \frac{c_1 w}{n-r} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j + \frac{c_9}{n-r-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( 1 - r \frac{T_i}{\Delta_i} DL_{i(j\cdot)} \right) - \\ &- 2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} w_i \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( 1 - r \frac{T_i}{\Delta_i} DL_{i(j\cdot)} \right)^{1/2}. \end{aligned}$$

### 3.4.1.2. Influencia en las columnas de $\widehat{\mathbf{B}}$

Para evaluar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $m$ -ésima columna de  $\widehat{\mathbf{B}}$ , se puede considerar  $\mathbf{D} = \mathbf{I}_q$  y  $\mathbf{L} = \mathbf{h}_m$  y, la expresión (3.4.34) se reduce a

$$\widehat{\Psi} = \widehat{\mathbf{B}}_{(q \times r)} \mathbf{h}_{m(r \times 1)},$$

donde  $\mathbf{h}_m \in \mathbb{R}^r$  es un vector de dimensión  $r$  con la  $m$ -ésima componente la unidad y las restantes nulas.

Por tanto, de los corolarios 3.4.3 y 3.4.4, se obtiene la medida de Fréchet asociada a  $\mathbf{y}_i$  sobre la  $m$ -ésima columna de  $\widehat{\mathbf{B}}$  para el MCC,

$$\delta_{i(\cdot m)}(\widehat{\mathbf{B}}) = \left[ LC_{i(\cdot m)}^2(\widehat{\mathbf{B}}) + DC_{i(\cdot m)}^2(\widehat{\mathbf{B}}) \right]^{1/2}.$$

De (3.4.37), la componente de localización es

$$LC_{i(\cdot m)}^2(\widehat{\mathbf{B}}) = \frac{1}{v_{ii}^2} (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i, \quad (3.4.52)$$

dato que  $\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \in \mathbb{R}$ .

Por otro lado, de (3.4.42), la componente de dispersión es

$$DC_{i(\cdot m)}^2(\widehat{\mathbf{B}}) = DC1_{i(\cdot m)} + DC2_{i(\cdot m)} + DC3_{i(\cdot m)}, \quad (3.4.53)$$

de forma que, de (3.4.43),

$$\begin{aligned} DC1_{i(\cdot m)} &= \frac{c_1}{n-r} \text{tr} [\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m] \text{tr} [\mathbf{S}_X] = \\ &= \frac{c_1}{n-r} \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \text{tr} [\mathbf{S}_X], \end{aligned} \quad (3.4.54)$$

de (3.4.44),

$$\begin{aligned} DC2_{i(\cdot m)} &= \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{S}_X] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right) \\ &\quad \left( \text{tr} [\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m] + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right) = \\ &= \frac{c_9}{n-r-1} \left( \text{tr} [\mathbf{S}_X] - \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} \right) \\ &\quad \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right) \end{aligned} \quad (3.4.55)$$

y, de (3.4.45),

$$\begin{aligned} DC3_{i(\cdot m)} &= -2 \left( \frac{c_1 c_9}{(n-r)(n-r-1)} \right)^{1/2} \text{tr} \left[ \left( \mathbf{S}_X^2 - \frac{1}{v_{ii}} \mathbf{S}_X^2 \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \right)^{1/2} \right] \\ &\quad \left( (\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)^{1/2}. \end{aligned} \quad (3.4.56)$$

**Nota 3.4.4** En el MCC, para el estudio de la influencia en las columnas de  $\widehat{\mathbf{B}}$  es posible hacer la siguiente descomposición. Obsérvese que

$$\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m = \begin{cases} p_{ii}, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}, \quad (3.4.57)$$

donde  $\mathbf{h}_m = \mathbf{z}_i$  representa que la columna de  $\widehat{\mathbf{B}}$  para la cual se pretende estudiar la influencia corresponde al grupo de la observación omitida. Así, dado que

$$\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m = \begin{cases} p_{ii}, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm}, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}, \quad (3.4.58)$$

se tiene

$$\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1 - p_{ii}} = \begin{cases} \frac{p_{ii}}{1 - p_{ii}}, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm}, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases} \quad (3.4.59)$$

y

$$\left( (\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1 - p_{ii}} \right)^{1/2} = \begin{cases} \sqrt{\frac{p_{ii}^2}{1 - p_{ii}}}, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm}, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}. \quad (3.4.60)$$

Por tanto, es posible calcular separadamente la componente de dispersión  $DC_{i(\cdot m)}^2$ , dada por (3.4.53), y consecuentemente la medida de Fréchet  $\delta_{i(\cdot m)}$ , cuando  $\mathbf{h}_m = \mathbf{z}_i$  y  $\mathbf{h}_m \neq \mathbf{z}_i$ . En particular, comparando las componentes  $DC1_{i(\cdot m)}$ ,  $DC2_{i(\cdot m)}$  y  $DC3_{i(\cdot m)}$  dadas en (3.4.54), (3.4.55) y (3.4.56) y las componentes  $DC1_i$ ,  $DC2_i$  y  $DC3_i$  dadas en (3.4.30), (3.4.31) y (3.4.32), respectivamente, se obtienen las relaciones

$$DC1_{i(\cdot m)} = \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m w^{-1} DC1_i,$$

$$DC2_{i(\cdot m)} = \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1 - p_{ii}} \right) \left( w + \frac{p_{ii}^2}{1 - p_{ii}} \right)^{-1} DC2_i$$

y

$$DC3_{i(\cdot m)} = \left( (\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1 - p_{ii}} \right)^{1/2} w_i^{-1} DC3_i.$$

En consecuencia, aplicando (3.4.59) y (3.4.60), se obtiene

$$\begin{aligned} DC_{i(\cdot m)}^2 &= DC1_{i(\cdot m)} + DC2_{i(\cdot m)} + DC3_{i(\cdot m)} = \\ &= \begin{cases} p_{ii} \left( w^{-1} DC1_i + \frac{1}{1 - p_{ii}} \left( w + \frac{p_{ii}^2}{1 - p_{ii}} \right)^{-1} DC2_i + \sqrt{\frac{1}{1 - p_{ii}}} w_i^{-1} DC3_i \right), & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \left( w^{-1} DC1_i + \left( w + \frac{p_{ii}^2}{1 - p_{ii}} \right)^{-1} DC2_i + w_i^{-1} DC3_i \right), & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}, \end{aligned} \quad (3.4.61)$$

con  $w$  y  $w_i$  dadas por (3.4.22) y (3.4.28), respectivamente. Además, teniendo en cuenta la componente de localización  $LC_i^2$ , dada en (3.4.20), de (3.4.52) se obtiene

$$LC_{i(\cdot m)}^2 = \frac{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_i^2. \quad (3.4.62)$$

Por tanto, es posible calcular la medida de Fréchet para las columnas de  $\widehat{\mathbf{B}}$  usando las componentes de la medida de Fréchet  $\delta_i$  dada en (3.4.33). Este hecho conlleva ventajas computacionales para el cálculo de  $\delta_{i(\cdot m)}$ .

### 3.4.1.3. Influencia en los elementos de $\widehat{\mathbf{B}}$

Para analizar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en el elemento  $(j, m)$  de  $\widehat{\mathbf{B}}$ , donde  $1 \leq j \leq q$  y  $1 \leq m \leq r$ , se considera  $\mathbf{D} = \mathbf{d}_j^T$  y  $\mathbf{L} = \mathbf{h}_m$  y, a través de la expresión (3.4.34), se obtiene

$$\widehat{\Psi} = \mathbf{d}_j^T (1 \times q) \widehat{\mathbf{B}}_{(q \times r)} \mathbf{h}_m (r \times 1) \in \mathbb{R},$$

con los vectores  $\mathbf{d}_j \in \mathbb{R}^q$  y  $\mathbf{h}_m \in \mathbb{R}^r$  definidos anteriormente.

Así, la medida de Fréchet asociada a  $\mathbf{y}_i$  sobre el elemento  $(j, m)$  del EMV  $\widehat{\mathbf{B}}$  en el MCC es

$$\delta_{i(jm)}(\widehat{\mathbf{B}}) = \left[ LC_{i(jm)}^2(\widehat{\mathbf{B}}) + DC_{i(jm)}^2(\widehat{\mathbf{B}}) \right]^{1/2},$$

donde las componentes  $LC_{i(jm)}^2$  y  $DC_{i(jm)}^2$  pueden ser obtenidas usando los corolarios 3.4.3 y 3.4.4, respectivamente.

De (3.4.37), se obtiene que la componente de localización es

$$LC_{i(jm)}^2(\widehat{\mathbf{B}}) = \frac{1}{v_{ii}^2} (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2. \quad (3.4.63)$$

Por otro lado, teniendo en cuenta  $LC_i^2$ ,  $LC_{i(j\cdot)}^2$  y  $LC_{i(\cdot m)}^2$  dados en (3.4.20), (3.4.47) y (3.4.52), respectivamente,

$$LC_{i(jm)}^2 = \frac{LC_{i(j\cdot)}^2 LC_{i(\cdot m)}^2}{LC_i^2} = LC_i^2 \left[ \frac{LC_{i(j\cdot)}^2}{LC_i^2} \right] \left[ \frac{LC_{i(\cdot m)}^2}{LC_i^2} \right] \quad (3.4.64)$$

y

$$LC_i^2 = \sum_{j=1}^q LC_{i(j\cdot)}^2 = \sum_{m=1}^r LC_{i(\cdot m)}^2 = \sum_{j=1}^q \sum_{m=1}^r LC_{i(jm)}^2.$$

**Nota 3.4.5** Los factores  $\frac{LC_{i(j\cdot)}^2}{LC_i^2}$  y  $\frac{LC_{i(\cdot m)}^2}{LC_i^2}$  representan las contribuciones de las componentes de localización, de la medida de Fréchet, para las filas y columnas de  $\widehat{\mathbf{B}}$  en la componente de localización de la influencia sobre  $\widehat{\mathbf{B}}$ . Además, la igualdad (3.4.64) también indica que  $LC_{i(jm)}^2$  es proporcional a  $LC_{i(j\cdot)}^2$  y a  $LC_{i(\cdot m)}^2$ .

De (3.4.42), la componente de dispersión para la medida de Fréchet asociada a los elementos de  $\widehat{\mathbf{B}}$  es

$$DC2_{i(jm)}^2(\widehat{\mathbf{B}}) = DC1_{i(jm)} + DC2_{i(jm)} + DC3_{i(jm)}, \quad (3.4.65)$$

con

$$\begin{aligned} DC1_{i(jm)} &= \frac{c_1}{n-r} \text{tr}[\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m] \text{tr}[\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j] = \\ &= \frac{c_1}{n-r} \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j, \end{aligned} \quad (3.4.66)$$

$$\begin{aligned} DC2_{i(jm)} &= \frac{c_9}{n-r-1} \left( \text{tr}[\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j] - \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right) \\ &\quad \left( \text{tr}[\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m] + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right) = \\ &= \frac{c_9}{n-r-1} \left( \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j - \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right) \\ &\quad \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right) \end{aligned} \quad (3.4.67)$$

y

$$\begin{aligned} DC3_{i(jm)} &= -2 \left( \frac{c_1}{n-r} \frac{c_9}{n-r-1} \right)^{1/2} \left( (\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j)^2 - \frac{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} \right)^{1/2} \\ &\quad \left( (\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)^{1/2}. \end{aligned}$$

Desarrollando (3.4.65), se obtiene

$$\begin{aligned} DC2_{i(jm)}^2 &= \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \quad (3.4.68) \\ &\quad \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right)} \right) \left( 1 + \frac{1}{1-p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \\ &= \left( \sqrt{DC1_{i(jm)}} - \sqrt{DC2_{i(jm)}} \right)^2. \end{aligned}$$

Por otro lado, teniendo en cuenta las cantidades  $DC1_i$ ,  $DC1_{i(j\cdot)}$ ,  $DC1_{i(\cdot m)}$  dadas en (3.4.30), (3.4.48) y (3.4.54), respectivamente, se concluye que

$$DC1_{i(jm)} = \frac{DC1_{i(j\cdot)} DC1_{i(\cdot m)}}{DC1_i} \quad (3.4.69)$$

y

$$DC1_i = \sum_{j=1}^q DC1_{i(j\cdot)} = \sum_{m=1}^r DC1_{i(\cdot m)} = \sum_{j=1}^q \sum_{m=1}^r DC1_{i(jm)}.$$

Análogamente,

$$DC2_{i(jm)} = \frac{DC2_{i(j\cdot)}DC2_{i(\cdot m)}}{DC2_i} \quad (3.4.70)$$

y

$$DC2_i = \sum_{j=1}^q DC2_{i(j\cdot)} = \sum_{m=1}^r DC2_{i(\cdot m)} = \sum_{j=1}^q \sum_{m=1}^r DC2_{i(jm)},$$

con  $DC2_i$ ,  $DC2_{i(j\cdot)}$  y  $DC2_{i(\cdot m)}$ , dadas en (3.4.31), (3.4.49) y (3.4.55), respectivamente.

Finalmente,

$$DC3_{i(jm)} = \frac{DC3_{i(j\cdot)}DC3_{i(\cdot m)}}{DC3_i}, \quad (3.4.71)$$

con  $DC3_i$ ,  $DC3_{i(j\cdot)}$  y  $DC3_{i(\cdot m)}$  dadas en (3.4.32), (3.4.50) y (3.4.56), respectivamente.

Por tanto, usando (3.4.64), (3.4.69), (3.4.70) y (3.4.71), también se puede expresar la medida de Fréchet asociada a los elementos de  $\hat{\mathbf{B}}$ , para la  $i$ -ésima observación de  $\mathbf{Y}$ , por

$$\begin{aligned} \delta_{i(jm)} &= \left[ LC_{i(jm)}^2 + DC_{i(jm)}^2 \right]^{1/2} = \\ &= \left[ LC_{i(jm)}^2 + DC1_{i(jm)} + DC2_{i(jm)} + DC3_{i(jm)} \right]^{1/2} = \\ &= \left[ \frac{LC_{i(j\cdot)}^2 LC_{i(\cdot m)}^2}{LC_i^2} + \frac{DC1_{i(j\cdot)}DC1_{i(\cdot m)}}{DC1_i} + \frac{DC2_{i(j\cdot)}DC2_{i(\cdot m)}}{DC2_i} + \frac{DC3_{i(j\cdot)}DC3_{i(\cdot m)}}{DC3_i} \right]^{1/2}. \end{aligned} \quad (3.4.72)$$

Otras relaciones que pueden ser útiles en algunos casos para el cálculo de estas medidas de influencia y sus componentes son las que a continuación se recogen.

En primer lugar,

$$\begin{aligned} DC_{i(jm)}^2 &= \begin{cases} \left( \sqrt{DC1_{i(jm)}} - \sqrt{DC2_{i(jm)}} \right)^2, & \mathbf{h}_m = \mathbf{z}_i \\ \left( \sqrt{DC1_{i(jm)}} - \sqrt{DC2_{i(jm)}} \right)^2, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases} = \\ &= \begin{cases} p_{ii} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \frac{1}{1-p_{ii}}} \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right)^2, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1}} \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right)^2, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}, \end{aligned} \quad (3.4.73)$$

con  $p_{mm} = \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m$ , obtenida aplicando (3.4.57), (3.4.58) y (3.4.68).

En segundo lugar, aplicando las relaciones (3.4.61), (3.4.69), (3.4.70) y (3.4.71), se tiene que

$$DC_{i(jm)}^2 = \begin{cases} p_{ii} \left( w^{-1} DC1_{i(j\cdot)} + \frac{1}{1-p_{ii}} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right)^{-1} DC2_{i(j\cdot)} + \sqrt{\frac{1}{1-p_{ii}}} w_i^{-1} DC3_{i(j\cdot)} \right), & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \left( w^{-1} DC1_{i(j\cdot)} + \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right)^{-1} DC2_{i(j\cdot)} + w_i^{-1} DC3_{i(j\cdot)} \right), & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

Finalmente, aplicando (3.4.62) en (3.4.64)

$$LC_{i(jm)}^2 = \frac{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_{i(j\cdot)}^2.$$

Por tanto, es posible calcular la medida de Fréchet para los elementos de  $\widehat{\mathbf{B}}$  usando la medida de Fréchet para las filas de  $\widehat{\mathbf{B}}$  dada en (3.4.46).

**Nota 3.4.6** La componente de dispersión,  $DC_{i(jm)}^2$ , puede ser expresada como función de la componente de localización,  $LC_{i(jm)}^2$ . En efecto, de (3.4.63), se tiene

$$\frac{1}{v_{ii}} (\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2 = \frac{v_{ii}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} LC_{i(jm)}^2,$$

por tanto, de (3.4.73), se obtiene

$$DC_{i(jm)}^2 = \begin{cases} p_{ii} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{v_{ii} LC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i)^2} \right)} \right)^2, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \left( 1 - \frac{v_{ii} LC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} \right)} \right)^2, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

**Nota 3.4.7** En el MCC, para el estudio de la influencia en los elementos de  $\widehat{\mathbf{B}}$ , la medida de Fréchet,  $\delta_{i(jm)}$ , puede ser expresada como función de la medida de Cook,  $DL_{i(jm)}$ . De (2.4.8), se obtiene

$$DL_{i(jm)} = \frac{1}{v_{ii}^2} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2 (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m}, \quad (3.4.74)$$

utilizando (3.4.58), resulta

$$LC_{i(jm)}^2 = \begin{cases} p_{ii} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j DL_{i(jm)}, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j DL_{i(jm)}, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$

y, de (3.4.73), se tiene

$$DC_{i(jm)}^2 = \begin{cases} p_{ii} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{p_{ii} v_{ii} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i)^2} \right)} \right)^2, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{c_9}{n-r-1} \left( 1 - \frac{p_{mm} v_{ii} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} \right)} \right)^2, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

### 3.5. La medida de Fréchet en el modelo de curvas de crecimiento con estructura simple de Rao

Teniendo en cuenta las ideas presentadas en la sección 3.3, se propone la utilización de la distancia de Fréchet, dada en (3.3.1), como medida de influencia de un subconjunto de observaciones  $\mathbf{Y}_I$  en la distribución del EMV  $\tilde{\mathbf{B}}$  de la matriz de coeficientes de regresión del MCC con estructura de covarianza simple de Rao, MCC(SCS), definido en la sección 1.4. Para ello, es necesario considerar el MCC(SCS) con la omisión de dichas observaciones, el modelo perturbado MCC(SCS)( $I$ ), obtenido de (2.3.1) usando la estructura simple de Rao.

Como queda recogido en la sección 1.4,

$$\text{vec}(\tilde{\mathbf{B}}) \sim N_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \mathbf{\Gamma}) \quad (3.5.1)$$

y

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}] = (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}, \quad (3.5.2)$$

con  $\tilde{\mathbf{\Gamma}}$  dado en (1.4.4). Análogamente, en la sección 2.5, se recoge que, para el MCC(SCS)( $I$ ),

$$\text{vec}(\tilde{\mathbf{B}}_{(I)}) \sim N_{qr}\left(\text{vec}(\mathbf{B}), (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \mathbf{\Gamma}\right) \quad (3.5.3)$$

y

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] = (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)}, \quad (3.5.4)$$

con  $\tilde{\mathbf{\Gamma}}_{(I)}$  dado en (2.5.1).

Así, la versión muestral de la medida de Fréchet (3.3.1) para evaluar la influencia de un subconjunto de observaciones  $\mathbf{Y}_I$  sobre el EMV  $\tilde{\mathbf{B}}$ , para el MCC(SCS), es

$$\delta_I(\tilde{\mathbf{B}}) = \left[ LC_I^2(\tilde{\mathbf{B}}) + DC_I^2(\tilde{\mathbf{B}}) \right]^{1/2},$$

de forma que la componente de localización viene dada por

$$LC_I^2(\tilde{\mathbf{B}}) = \left\| \tilde{\mathbf{B}} - \tilde{\mathbf{B}}_{(I)} \right\|^2$$

y, dado que en este caso, por la nota C.1.2,  $\frac{E[R^2]}{qr} = 1$ , la componente de dispersión es

$$DC_I^2(\tilde{\mathbf{B}}) = \text{tr} \left[ \widehat{\text{Var}}[\tilde{\mathbf{B}}] + \widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\tilde{\mathbf{B}}] \widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] \right)^{1/2} \right].$$

En los teoremas siguientes, se presentan expresiones de las componentes de localización y dispersión de la medida de Fréchet asociada al EMV  $\tilde{\mathbf{B}}$  del MCC(SCS).

**Teorema 3.5.1** *En el MCC(SCS), la componente de localización de la medida de Fréchet  $\delta_I(\tilde{\mathbf{B}})$  es*

$$LC_I^2(\tilde{\mathbf{B}}) = \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{Z}_I \right], \quad (3.5.5)$$

con  $\mathbf{E}_I$  y  $\mathbf{H}_I$  las matrices dadas en (2.3.6) y (2.3.7), respectivamente.



**Demostración:** Análoga a la realizada en el teorema 3.4.1 para el MCC, teniendo en cuenta la relación entre los EMV  $\tilde{\mathbf{B}}$  y  $\tilde{\mathbf{B}}_{(I)}$  dada por (2.5.4). ■

A partir del teorema 3.5.1, se obtiene la componente de localización de la medida de Fréchet para  $\tilde{\mathbf{B}}$  asociada a la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.5.1** En el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de localización de la medida de Fréchet  $\delta_i(\tilde{\mathbf{B}})$  es

$$LC_i^2(\tilde{\mathbf{B}}) = \frac{p_{ii}^2}{(1 - p_{ii})^2} \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i, \quad (3.5.6)$$

dado que  $\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i = p_{ii}$  y  $\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{z}_i = p_{ii}^2$ , siendo  $\mathbf{z}_i$  la columna  $i$ -ésima de  $\mathbf{Z}$  y  $\mathbf{e}_i$ , dado en (2.3.14), es la columna  $i$ -ésima de matriz de residuos  $\mathbf{E}$  definida en (2.2.2).

**Nota 3.5.1** Teniendo en cuenta la medida de Cook,  $D_i$ , para la  $i$ -ésima observación de  $\mathbf{Y}$  definida en (2.5.11),

$$LC_i^2(\tilde{\mathbf{B}}) = \left( \frac{D_i}{nT_i} \right)^2 \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i,$$

donde  $T_i$ , dado por (2.5.8), es un estadístico que se aplica para determinar si la observación  $i$ -ésima de  $\mathbf{Y}$  es outlier en el MCC(SCS).

Para obtener la componente de dispersión de la medida de Fréchet, en el MCC(SCS), es útil el siguiente resultado.

**Lema 3.5.1** Sean  $\tilde{\mathbf{\Gamma}}$  y  $\tilde{\mathbf{\Gamma}}_{(I)}$  las matrices de dispersión para el MCC(SCS) y el MCC(SCS)( $I$ ) dadas, respectivamente, por (1.4.4) y (2.5.1), entonces:

(i)

$$\text{tr}[\tilde{\mathbf{\Gamma}}_{(I)}] = \frac{1}{n - k} \left( n \text{tr}[\tilde{\mathbf{\Gamma}}] - \text{tr}[(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{E}_I] \right). \quad (3.5.7)$$

(ii)

$$\tilde{\mathbf{\Gamma}} \tilde{\mathbf{\Gamma}}_{(I)} = \frac{1}{n - k} \left( n \tilde{\mathbf{\Gamma}}^2 - \tilde{\mathbf{\Gamma}} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right). \quad (3.5.8)$$

**Demostración:** Estos resultados se obtienen a partir de la relación (2.5.5) entre  $\tilde{\mathbf{\Gamma}}$  y  $\tilde{\mathbf{\Gamma}}_{(I)}$ , presentada en el teorema 2.5.1. ■

**Teorema 3.5.2** En el MCC(SCS), la componente de dispersión de la medida de Fréchet  $\delta_I(\tilde{\mathbf{B}})$  es

$$\begin{aligned} DC_I^2(\tilde{\mathbf{B}}) = & w \text{tr}[\tilde{\mathbf{\Gamma}}] + \frac{1}{n - k} \left( w + \text{tr}[(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{Z}_I] \right) \\ & \left( n \text{tr}[\tilde{\mathbf{\Gamma}}] - \text{tr}[(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{E}_I] \right) - \\ & - \frac{2}{\sqrt{n - k}} \text{tr} \left[ ((\mathbf{Z} \mathbf{Z}^T)^{-2} + (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1})^{1/2} \right] \\ & \text{tr} \left[ \left( n \tilde{\mathbf{\Gamma}}^2 - \tilde{\mathbf{\Gamma}} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right)^{1/2} \right]. \end{aligned} \quad (3.5.9)$$

**Demostración:** Teniendo en cuenta los estimadores de las matrices de covarianza  $\widehat{\text{Var}}[\tilde{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}]$  dadas por (3.5.2) y (3.5.4), respectivamente, y aplicando las propiedades de la traza,

$$\begin{aligned}
DC_I^2(\tilde{\mathbf{B}}) &= \text{tr} \left[ \widehat{\text{Var}}[\tilde{\mathbf{B}}] + \widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] - 2 \left( \widehat{\text{Var}}[\tilde{\mathbf{B}}] \widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] \right)^{1/2} \right] = \\
&= \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right) + \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) - \right. \\
&\quad \left. - 2 \left( \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right) \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) \right)^{1/2} \right] = \\
&= \text{tr} \left[ (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right] + \text{tr} \left[ (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right] - \\
&\quad - 2 \text{tr} \left[ \left( \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right) \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) \right)^{1/2} \right] = \\
&= DC1(\tilde{\mathbf{B}}) + DC2(\tilde{\mathbf{B}}) + DC3(\tilde{\mathbf{B}}).
\end{aligned}$$

Las expresiones de estas tres componentes son obtenidas aplicando las propiedades del producto de Kronecker recogidas en el teorema B.0.5. Por un lado,

$$DC1(\tilde{\mathbf{B}}) = \text{tr} \left[ (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right] = \text{tr} \left[ (\mathbf{Z}\mathbf{Z}^T)^{-1} \right] \text{tr}[\tilde{\mathbf{\Gamma}}] = w \text{tr}[\tilde{\mathbf{\Gamma}}],$$

con  $w$  dado en (3.4.22), que no depende de  $I$ , es decir, la cantidad  $DC1(\tilde{\mathbf{B}})$  es una constante. Por otro lado, de (3.4.24) y de la expresión (3.5.7), se concluye que

$$\begin{aligned}
DC2(\tilde{\mathbf{B}}) &= \text{tr} \left[ (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right] = \text{tr} \left[ (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \right] \text{tr}[\tilde{\mathbf{\Gamma}}_{(I)}] = \\
&= \frac{1}{n-k} (w + \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{Z}_I \right]) \\
&\quad \left( n \text{tr}[\tilde{\mathbf{\Gamma}}] - \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{E}_I \right] \right).
\end{aligned}$$

Finalmente,

$$\begin{aligned}
DC3(\tilde{\mathbf{B}}) &= -2 \text{tr} \left[ \left( \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right) \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) \right)^{1/2} \right] = \\
&= -2 \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \left( \mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T \right)^{-1} \otimes \tilde{\mathbf{\Gamma}} \tilde{\mathbf{\Gamma}}_{(I)} \right)^{1/2} \right] = \\
&= -2 \text{tr} \left[ \left( \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \left( \mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T \right)^{-1} \right)^{1/2} \right) \text{tr} \left[ (\tilde{\mathbf{\Gamma}} \tilde{\mathbf{\Gamma}}_{(I)})^{1/2} \right] \right] = \\
&= -\frac{2}{\sqrt{n-k}} \text{tr} \left[ \left( (\mathbf{Z}\mathbf{Z}^T)^{-2} + (\mathbf{Z}\mathbf{Z}^T)^{-2} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \right)^{1/2} \right] \\
&\quad \text{tr} \left[ \left( n \tilde{\mathbf{\Gamma}}^2 - \tilde{\mathbf{\Gamma}} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right)^{1/2} \right],
\end{aligned}$$

donde la última igualdad se obtiene usando el lema 3.4.2 y la expresión (3.5.8). ■

A partir de este teorema, se obtiene la componente de dispersión de la medida de Fréchet para  $\tilde{\mathbf{B}}$  asociada a la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 3.5.2** En el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de dispersión de la medida de Fréchet  $\delta_i(\tilde{\mathbf{B}})$  es

$$\begin{aligned} DC_i^2(\tilde{\mathbf{B}}) &= DC1_i + DC2_i + DC3_i = \\ &= \text{wtr}[\tilde{\Gamma}] + \frac{1}{n-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( n\text{tr}[\tilde{\Gamma}] - \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}} \right) - \\ &\quad - \frac{2w_i}{\sqrt{n-1}} \text{tr} \left[ \left( n\tilde{\Gamma}^2 - \frac{1}{1-p_{ii}} \tilde{\Gamma} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right)^{1/2} \right], \end{aligned}$$

con  $w$  y  $w_i$  dados en (3.4.22) y (3.4.28), respectivamente, y  $p_{ii}^2 = \mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{z}_i$ .

De los teoremas 3.5.1 y 3.5.2, se concluye que la medida Fréchet para evaluar la influencia del subconjunto de observaciones  $\mathbf{Y}_I$  en la distribución del EMV  $\tilde{\mathbf{B}}$  para el MCC(SCS) es

$$\delta_I(\tilde{\mathbf{B}}) = \left[ LC_I^2(\tilde{\mathbf{B}}) + DC_I^2(\tilde{\mathbf{B}}) \right]^{1/2},$$

donde la componente de localización  $LC_I^2(\tilde{\mathbf{B}})$  es dada en (3.5.5) y la componente de dispersión  $DC_I^2(\tilde{\mathbf{B}})$  dada en (3.5.9).

Por otro lado, teniendo en cuenta los corolarios 3.5.1 y 3.5.2, la medida de Fréchet para estudiar la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  en la distribución del EMV  $\tilde{\mathbf{B}}$  del MCC(SCS) es

$$\begin{aligned} \delta_i(\tilde{\mathbf{B}}) &= \left[ LC_i^2(\tilde{\mathbf{B}}) + DC_i^2(\tilde{\mathbf{B}}) \right]^{1/2} = \\ &= \left[ \frac{p_{ii}^2}{(1-p_{ii})^2} \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i + \text{wtr}[\tilde{\Gamma}] + \right. \\ &\quad \left. + \frac{1}{n-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( n\text{tr}[\tilde{\Gamma}] - \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}} \right) - \right. \\ &\quad \left. - \frac{2w_i}{\sqrt{n-1}} \text{tr} \left[ \left( n\tilde{\Gamma}^2 - \frac{1}{1-p_{ii}} \tilde{\Gamma} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right)^{1/2} \right] \right]^{1/2}. \end{aligned}$$

### 3.5.1. Influencia en combinaciones lineales de $\tilde{\mathbf{B}}$

También en el MCC(SCS), la distancia de Fréchet puede ser extendida para estudiar la influencia de las observaciones  $\mathbf{Y}_I$  en la distribución de un determinado subconjunto del EMV  $\tilde{\mathbf{B}}$ . En este caso, el EMV de la expresión general para las combinaciones lineales de  $\tilde{\mathbf{B}}$  es

$$\tilde{\Psi} = \underset{(s \times q)(q \times r)(r \times l)}{\mathbf{D} \tilde{\mathbf{B}} \mathbf{L}}$$

donde, los rangos de las matrices conocidas  $\mathbf{D}$  y  $\mathbf{L}$  son  $\text{rg}(\mathbf{D}) = s \leq q$  y  $\text{rg}(\mathbf{L}) = l \leq r$ . Por otro lado, para el MCC(SCS)( $I$ ) se obtiene  $\tilde{\Psi}_{(I)} = \mathbf{D} \tilde{\mathbf{B}}_{(I)} \mathbf{L}$ . El desarrollo que se sigue es análogo al realizado en la sección 3.4.1 para el MCC.

De las distribuciones de  $\text{vec}(\tilde{\mathbf{B}})$  y  $\text{vec}(\tilde{\mathbf{B}}_{(I)})$  dadas en (3.5.1) y (3.5.3), respectivamente, y usando el teorema C.1.2, se tiene

$$\text{vec}(\mathbf{D} \tilde{\mathbf{B}} \mathbf{L}) \sim N_{sl}(\text{vec}(\mathbf{D} \tilde{\mathbf{B}} \mathbf{L}), \mathbf{L}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D} \Gamma \mathbf{D}^T),$$

$$\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}] = \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T,$$

$$\text{vec}(\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}) \sim N_{sl}\left(\text{vec}(\mathbf{D}\mathbf{B}\mathbf{L}), \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\mathbf{\Gamma}\mathbf{D}^T\right)$$

y

$$\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}] = \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\mathbf{\Gamma}}_{(I)}\mathbf{D}^T.$$

A continuación, se obtienen las expresiones generales de las componentes de localización y dispersión de la medida de Fréchet para analizar la influencia del conjunto de observaciones  $\mathbf{Y}_I$ , y de la  $i$ -ésima observación  $\mathbf{y}_i$ , sobre  $\tilde{\Psi}$ .

**Teorema 3.5.3** *En el MCC(SCS), la componente de localización para la medida de Fréchet  $\delta_I(\tilde{\Psi})$  es*

$$LC_I^2(\tilde{\Psi}) = \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I \right. \\ \left. (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I \right],$$

donde las matrices  $\mathbf{E}_I$  y  $\mathbf{H}_I$  están definidas por (2.3.6) y (2.3.7), respectivamente.

**Demostración:** Análoga a la realizada en el teorema 3.4.3, para el MCC, teniendo en cuenta la relación entre los EMV  $\tilde{\mathbf{B}}$  y  $\tilde{\mathbf{B}}_{(I)}$  dada en (2.5.4). ■

**Corolario 3.5.3** *En el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de localización de  $\delta_i(\tilde{\Psi})$  es*

$$LC_i^2(\tilde{\Psi}) = \frac{1}{(1 - p_{ii})^2} \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \quad (3.5.10) \\ \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i.$$

**Teorema 3.5.4** *En el MCC(SCS), la componente de dispersión para la medida de Fréchet  $\delta_I(\tilde{\Psi})$  es*

$$DC_I^2(\tilde{\Psi}) = DC1(\tilde{\Psi}) + DC2(\tilde{\Psi}) + DC3(\tilde{\Psi}),$$

con

$$DC1(\tilde{\Psi}) = \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] \text{tr} [\mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T],$$

$$DC2(\tilde{\Psi}) = \frac{1}{n - k} \left( \text{tr} [\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}] + \text{tr} [(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I] \right) \\ \left( n \text{tr} [\mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T] - \text{tr} [(\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I] \right)$$

y

$$DC3(\tilde{\Psi}) = -\frac{2}{\sqrt{n - k}} \\ \text{tr} \left[ \left( (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^2 + \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right)^{1/2} \right] \\ \text{tr} \left[ \left( n (\mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T)^2 - \mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T \mathbf{D} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{D}^T \right)^{1/2} \right].$$

**Demostración:** De forma análoga al teorema 3.5.2, se obtiene

$$\begin{aligned}
 DC_I^2(\tilde{\Psi}) &= \text{tr} \left[ \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}] + \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}] - 2 \left( \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}] \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}] \right)^{1/2} \right] = \\
 &= \text{tr} \left[ \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right) + \left( \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right) - \right. \\
 &\quad \left. - 2 \left( \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right) \otimes \left( \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right) \right)^{1/2} \right] = \\
 &= \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right] + \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right] - \\
 &\quad - 2\text{tr} \left[ \left( \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right) \left( \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right) \right)^{1/2} \right] = \\
 &= DC1(\tilde{\Psi}) + DC2(\tilde{\Psi}) + DC3(\tilde{\Psi}).
 \end{aligned}$$

En primer lugar,

$$DC1(\tilde{\Psi}) = \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right] = \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right] \text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T].$$

Por otro lado, considerando la igualdad (2.5.5), resulta

$$\begin{aligned}
 \text{tr}[\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T] &= \text{tr} \left[ \mathbf{D} \left( \frac{n}{n-k}\tilde{\Gamma} - \frac{1}{n-k}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} \right) \mathbf{D}^T \right] = \\
 &= \frac{1}{n-k} \left( n\text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T] - \text{tr} \left[ \mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T \right] \right),
 \end{aligned}$$

por lo que, teniendo en cuenta la expresión (3.4.40), se obtiene

$$\begin{aligned}
 DC2(\tilde{\Psi}) &= \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right] = \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \right] \text{tr}[\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T] = \\
 &= \frac{1}{n-k} \left( \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right] + \text{tr} \left[ \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right) \right. \right. \\
 &\quad \left. \left. \left( n\text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T] - \text{tr} \left[ \mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T \right] \right) \right] = \\
 &= \frac{1}{n-k} \left( \text{tr} \left[ \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right] + \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I \right] \right. \\
 &\quad \left. \left( n\text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T] - \text{tr} \left[ (\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I \right] \right) \right).
 \end{aligned}$$

Finalmente, usando la propiedad (ix) del teorema B.0.5,

$$\begin{aligned}
 DC3(\tilde{\Psi}) &= -2\text{tr} \left[ \left( \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right) \left( \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right) \right)^{1/2} \right] = \\
 &= -2\text{tr} \left[ \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right)^{1/2} \right] = \\
 &= -2\text{tr} \left[ \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \right)^{1/2} \right] \text{tr} \left[ \left( \mathbf{D}\tilde{\Gamma}\mathbf{D}^T\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T \right)^{1/2} \right] = \\
 &= -\frac{2}{\sqrt{n-k}} \\
 &\quad \text{tr} \left[ \left( \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right)^2 + \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right)^{1/2} \right] \\
 &\quad \text{tr} \left[ \left( n(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^2 - \mathbf{D}\tilde{\Gamma}\mathbf{D}^T\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T \right)^{1/2} \right].
 \end{aligned}$$

La última igualdad se obtiene teniendo en cuenta el lema 3.4.2 y la relación (2.5.5). ■

**Corolario 3.5.4** *En el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la componente de dispersión de  $\delta_i(\tilde{\Psi})$  es*

$$\begin{aligned}
DC_i^2(\tilde{\Psi}) &= \text{tr}[\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}] \text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T] + \\
&+ \frac{1}{n-1} \left( \text{tr}[\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}] + \frac{\mathbf{z}_i^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{z}_i}{1-p_{ii}} \right) \\
&\left( n\text{tr}[\mathbf{D}\tilde{\Gamma}\mathbf{D}^T] - \frac{\mathbf{e}_i^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{e}_i}{1-p_{ii}} \right) - \frac{2}{\sqrt{n-1}} \\
&\text{tr} \left( \left( \left( \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right)^2 + \frac{1}{1-p_{ii}} \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{z}_i\mathbf{z}_i^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \right)^{1/2} \right) \\
&\text{tr} \left( \left( n \left( \mathbf{D}\tilde{\Gamma}\mathbf{D}^T \right)^2 - \frac{1}{1-p_{ii}} \mathbf{D}\tilde{\Gamma}\mathbf{D}^T\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{e}_i\mathbf{e}_i^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T \right)^{1/2} \right).
\end{aligned} \tag{3.5.11}$$

Así, usando los corolarios 3.5.3 y 3.5.4, se obtiene la expresión general de la medida de Fréchet asociada a las combinaciones lineales de  $\tilde{\mathbf{B}}$  para la  $i$ -ésima observación de  $\mathbf{Y}$  en el MCC(SCS)

$$\delta_i(\tilde{\Psi}) = \left[ LC_i^2(\tilde{\Psi}) + DC_i^2(\tilde{\Psi}) \right]^{1/2}, \tag{3.5.12}$$

con  $LC_i^2(\tilde{\Psi})$  y  $DC_i^2(\tilde{\Psi})$  dados, respectivamente, por (3.5.10) y (3.5.11).

### 3.5.1.1. Influencia en las filas de $\tilde{\mathbf{B}}$

Para estudiar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en las filas de  $\tilde{\mathbf{B}}$  basta considerar  $\tilde{\Psi} = \mathbf{d}_j^T \tilde{\mathbf{B}}$ , donde  $\mathbf{d}_j$  es un vector de dimensión  $q$  con la  $j$ -ésima componente igual a la unidad y restantes nulas.

Razonando de forma análoga que en la sección 3.4.1.1, de (3.5.12), se obtiene que la medida de Fréchet asociada al MCC(SCS) para estudiar la influencia de  $\mathbf{y}_i$  sobre la  $j$ -ésima fila de  $\tilde{\mathbf{B}}$  es

$$\delta_{i(j\cdot)}(\tilde{\mathbf{B}}) = \left[ LC_{i(j\cdot)}^2(\tilde{\mathbf{B}}) + DC_{i(j\cdot)}^2(\tilde{\mathbf{B}}) \right]^{1/2},$$

donde, del corolario 3.5.3, la componente de localización es

$$LC_{i(j\cdot)}^2(\tilde{\mathbf{B}}) = \frac{p_{ii}^2}{(1-p_{ii})^2} (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2 \tag{3.5.13}$$

y, del corolario 3.5.4, la componente de dispersión

$$DC_{i(j\cdot)}^2(\tilde{\mathbf{B}}) = DC1_{i(j\cdot)} + DC2_{i(j\cdot)} + DC3_{i(j\cdot)},$$

con

$$DC1_{i(j\cdot)} = \text{tr}[(\mathbf{Z}\mathbf{Z}^T)^{-1}] \text{tr}[\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j] = w \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j,$$

$$\begin{aligned} DC2_{i(j\cdot)} &= \frac{1}{n-1} \left( \text{tr} [(\mathbf{Z}\mathbf{Z}^T)^{-1}] + \frac{1}{1-p_{ii}} \text{tr} [z_i^T (\mathbf{Z}\mathbf{Z}^T)^{-2} z_i] \right) \\ &\quad \left( n \text{tr} [\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j] - \frac{1}{1-p_{ii}} \text{tr} [(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2] \right) = \\ &= \frac{1}{n-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( n \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j - \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{1-p_{ii}} \right) \end{aligned}$$

y

$$DC3_{i(j\cdot)} = -\frac{2w_i}{\sqrt{n-1}} \left( n \left( \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \right)^2 - \frac{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{1-p_{ii}} \right)^{1/2},$$

con  $w_i$  dada por (3.4.28).

**Nota 3.5.2** También de forma análoga a la citada sección, la componente de dispersión  $DC_{i(j\cdot)}^2$  puede expresarse como función de la componente de localización  $LC_{i(j\cdot)}^2$ , o sea,

$$\begin{aligned} DC_{i(j\cdot)}^2(\tilde{\mathbf{B}}) &= w \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j + \frac{1}{n-1} \left( w + \frac{p_{ii}^2}{1-p_{ii}} \right) \left( n \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j - \frac{1-p_{ii}}{p_{ii}^2} LC_{i(j\cdot)}^2 \right) - \\ &\quad - \frac{2w_i}{\sqrt{n-1}} \left( n \left( \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \right)^2 - \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \frac{1-p_{ii}}{p_{ii}^2} LC_{i(j\cdot)}^2 \right)^{1/2}. \end{aligned}$$

**Nota 3.5.3** En el MCC(SCS), para el estudio de la influencia en las filas de  $\tilde{\mathbf{B}}$ , la medida de Fréchet puede ser expresada como función de la medida de Cook

$$\begin{aligned} \delta_{i(j\cdot)}^2 &= p_{ii} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j DL_{i(j\cdot)} + w \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j + \frac{n}{n-1} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \left( w + p_{ii} \frac{D_i}{nT_i} \right) \left( 1 - \frac{T_i}{D_i} DL_{i(j\cdot)} \right) - \\ &\quad - 2 \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \sqrt{\frac{n}{n-1}} w_i \left( 1 - \frac{T_i}{D_i} DL_{i(j\cdot)} \right)^{1/2}. \end{aligned}$$

### 3.5.1.2. Influencia en las columnas de $\tilde{\mathbf{B}}$

Razonando de forma semejante para las columnas del EMV  $\tilde{\mathbf{B}}$ , para evaluar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  basta considerar  $\tilde{\Psi} = \tilde{\mathbf{B}} \mathbf{h}_m$ , donde  $\mathbf{h}_m \in \mathbb{R}^r$  es un vector de dimensión  $r$  con la  $m$ -ésima componente unitaria y las restantes nulas.

En estos términos, de (3.5.12), la medida de Fréchet asociada al MCC(SCS) para estudiar la influencia de  $\mathbf{y}_i$  en la  $m$ -ésima columna de  $\tilde{\mathbf{B}}$  es

$$\delta_{i(m)}(\tilde{\mathbf{B}}) = \left[ LC_{i(m)}^2(\tilde{\mathbf{B}}) + DC_{i(m)}^2(\tilde{\mathbf{B}}) \right]^{1/2},$$

con componente de localización

$$LC_{i(m)}^2(\tilde{\mathbf{B}}) = \frac{1}{(1-p_{ii})^2} \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i (z_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 \quad (3.5.14)$$

y componente de dispersión

$$DC_{i(m)}^2(\tilde{\mathbf{B}}) = DC1_{i(m)} + DC2_{i(m)} + DC3_{i(m)},$$

donde

$$DC1_{i(\cdot m)} = \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \text{tr}[\tilde{\Gamma}],$$

$$DC2_{i(\cdot m)} = \frac{1}{n-1} \left( n \text{tr}[\tilde{\Gamma}] - \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}} \right) \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)$$

y

$$DC3_{i(\cdot m)} = -\frac{2}{\sqrt{n-1}} \text{tr} \left[ \left( n \tilde{\Gamma}^2 - \frac{1}{1-p_{ii}} \tilde{\Gamma} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \right)^{1/2} \right] \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \right)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)^{1/2}.$$

**Nota 3.5.4** Usando la descomposición (3.4.57), de (3.5.14), se obtiene

$$LC_{i(\cdot m)}^2 = \begin{cases} \frac{p_{ii}^2}{(1-p_{ii})^2} \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-2} \mathbf{X}^T \mathbf{e}_i, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases} = \begin{cases} LC_i^2, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$

con  $LC_i^2$  dada en (3.5.6). El hecho de que para  $\mathbf{h}_m \neq \mathbf{z}_i$  se tenga  $LC_{i(\cdot m)}^2 = 0$ , significa que, como las columnas de  $\tilde{\mathbf{B}}$  corresponden a los diferentes grupos, la componente de localización solamente contribuye para la medida de Fréchet cuando la observación eliminada pertenece al grupo correspondiente a la columna para la cual se pretende estudiar la influencia. Así, cuando  $\mathbf{h}_m \neq \mathbf{z}_i$ , la medida de Fréchet viene dada por  $\delta_{i(\cdot m)}^2 = DC_{i(\cdot m)}^2$ , la influencia de un grupo en otro es solamente en dispersión.

### 3.5.1.3. Influencia en los elementos de $\tilde{\mathbf{B}}$

Para estudiar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en los elementos del EMV  $\tilde{\mathbf{B}}$ , se considera  $\tilde{\Psi} = \mathbf{d}_j^T \tilde{\mathbf{B}} \mathbf{h}_m \in \mathbb{R}$ , con  $\mathbf{d}_j^T \in \mathbb{R}^q$  y  $\mathbf{h}_m \in \mathbb{R}^r$  definidos anteriormente.

La medida de Fréchet para evaluar la influencia de  $\mathbf{y}_i$  sobre el elemento  $(j, m)$  de  $\tilde{\mathbf{B}}$ ,  $1 \leq j \leq q$  y  $1 \leq m \leq r$ , es

$$\delta_{i(jm)}(\tilde{\mathbf{B}}) = \left[ LC_{i(jm)}^2(\tilde{\mathbf{B}}) + DC_{i(jm)}^2(\tilde{\mathbf{B}}) \right]^{1/2},$$

donde la componente de localización es

$$LC_{i(jm)}^2(\tilde{\mathbf{B}}) = \frac{1}{(1-p_{ii})^2} (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2 (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 \quad (3.5.15)$$

y la componente de dispersión es

$$DC_{i(jm)}^2(\tilde{\mathbf{B}}) = DC1_{i(jm)} + DC2_{i(jm)} + DC3_{i(jm)}, \quad (3.5.16)$$



con

$$DC1_{i(jm)} = \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j,$$

$$DC2_{i(jm)} = \frac{1}{n-1} \left( n \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j - \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{1-p_{ii}} \right) \left( \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m + \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)$$

y

$$DC3_{i(jm)} = -\frac{2}{\sqrt{n-1}} \left( n (\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j)^2 - \frac{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{1-p_{ii}} \right)^{1/2} \left( (\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 + \frac{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m (\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{1-p_{ii}} \right)^{1/2}.$$

Desarrollando (3.5.16), se obtiene

$$DC_{i(jm)}^2 = \left( \sqrt{DC1_{i(jm)}} - \sqrt{DC2_{i(jm)}} \right)^2. \quad (3.5.17)$$

Otras relaciones que se verifican en el MCC(SCS) son

$$\delta_{i(jm)} = \left[ \frac{LC_{i(j\cdot)}^2 LC_{i(\cdot m)}^2}{LC_i^2} + \frac{DC1_{i(j\cdot)} DC1_{i(\cdot m)}}{DC1_i} + \frac{DC2_{i(j\cdot)} DC2_{i(\cdot m)}}{DC2_i} + \frac{DC3_{i(j\cdot)} DC3_{i(\cdot m)}}{DC3_i} \right]^{1/2},$$

$$LC_i^2 = \sum_{j=1}^q LC_{i(j\cdot)}^2 = \sum_{m=1}^r LC_{i(\cdot m)}^2 = \sum_{j=1}^q \sum_{m=1}^r LC_{i(jm)}^2,$$

$$DC1_i = \sum_{j=1}^q DC1_{i(j\cdot)} = \sum_{m=1}^r DC1_{i(\cdot m)} = \sum_{j=1}^q \sum_{m=1}^r DC1_{i(jm)},$$

y

$$DC2_i = \sum_{j=1}^q DC2_{i(j\cdot)} = \sum_{m=1}^r DC2_{i(\cdot m)} = \sum_{j=1}^q \sum_{m=1}^r DC2_{i(jm)}.$$

**Nota 3.5.5** Usando la descomposición (3.4.57) y la expresión (3.5.15),

$$LC_{i(jm)}^2 = \begin{cases} \left( \frac{p_{ii}}{1-p_{ii}} \right)^2 (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases} = \begin{cases} LC_{i(j\cdot)}^2, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$


---

con  $LC_{i(j)}^2$ , dada en (3.5.13), y de (3.5.17), se obtiene

$$DC_{i(jm)}^2 = \begin{cases} p_{ii} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \left( 1 - \sqrt{\frac{n}{n-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{1}{n(1-p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right)} \right)^2, & \mathbf{h}_m = \mathbf{z}_i \\ p_{mm} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \left( 1 - \sqrt{\frac{n}{n-1} \left( 1 - \frac{1}{n(1-p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right)} \right)^2, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases} \quad (3.5.18)$$

El hecho de que para  $\mathbf{h}_m \neq \mathbf{z}_i$  se tenga  $LC_{i(jm)}^2 = 0$ , significa que la componente de localización solamente contribuye en la medida de Fréchet cuando la observación omitida pertenece al grupo correspondiente a la columna para la cual se pretende estudiar la influencia.

**Nota 3.5.6** Para el MCC(SCS), la medida de Fréchet,  $\delta_{i(jm)}$ , puede ser expresada como función de la medida de Cook,  $DL_{i(jm)}$ . De (2.5.14), se obtiene

$$DL_{i(jm)} = \frac{1}{(1-p_{ii})^2} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j}, \quad (3.5.19)$$

por tanto, de (3.5.15), resulta

$$LC_{i(jm)}^2 = \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m DL_{i(jm)},$$

es decir,

$$LC_{i(jm)}^2 = \begin{cases} p_{ii} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j DL_{i(jm)}, & \mathbf{h}_m = \mathbf{z}_i \\ 0 & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

Por otro lado, se observa que solamente para  $\mathbf{h}_m = \mathbf{z}_i$  es posible expresar  $DC_{i(jm)}^2$  como función de  $DL_{i(jm)}$ , dado que  $(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2 = p_{ii} \neq 0$ . En efecto, de (3.5.19), resulta

$$\frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{1-p_{ii}} = \frac{1-p_{ii}}{p_{ii}} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j DL_{i(jm)} = n \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \frac{T_i}{D_i} DL_{i(jm)}, \quad \mathbf{h}_m = \mathbf{z}_i,$$

y, de (3.5.18), viene

$$DC_{i(jm)}^2 = p_{ii} \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j \left( 1 - \sqrt{\frac{n}{n-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{T_i}{D_i} DL_{i(jm)} \right)} \right)^2, \quad \mathbf{h}_m = \mathbf{z}_i.$$

Se concluye que, para  $\mathbf{h}_m = \mathbf{z}_i$ , la medida  $\delta_{i(jm)}$  es una función creciente de  $DL_{i(jm)}$ , o sea, ambas medidas detectan las mismas observaciones influyentes en cada grupo.

## Capítulo 4

# Influencia a través de la distancia de Rao

### 4.1. Introducción

La estrategia de introducir una distancia entre diferentes poblaciones estadísticas como medida de influencia ha sido considerada en el capítulo anterior. En este capítulo, se usa la distancia de Rao (Rao, 1945 y Rao, 1949) para el análisis de observaciones influyentes en el estimador de máxima verosimilitud (EMV) de la matriz de coeficientes de regresión  $\mathbf{B}$  del modelo de curvas de crecimiento (MCC) y del MCC con estructura simple de Rao (MCC(SCS)).

El presente capítulo está organizado como sigue. En la sección 4.2, se presenta la distancia de Rao para su utilización como medida de influencia en el MCC. En particular, se evidencia el interés de la aplicación de dicha distancia por sus buenas propiedades.

En la sección 4.3, se define una expresión general de la distancia de Rao y su fórmula particular para una distribución normal matricial  $qr$ -variante. Después se propone esta distancia como medida de influencia asociada al EMV de la matriz de coeficientes de regresión del MCC, así como para combinaciones lineales del mismo. Se discute su aplicabilidad, así como algunas diferencias y relaciones respecto a las medidas expuestas en la sección 2.4 y a la medida de Fréchet.

Para finalizar, en la sección 4.4, se hace el mismo desarrollo que en la sección 4.3 para el MCC(SCS).

### 4.2. Distancia de Rao y su utilización en el análisis de influencia

Los trabajos de Rao (1945) y Rao (1949) introducen la distancia de Rao, que se presenta a continuación de una forma simplificada.

Sea  $X$  un vector que describe la población bajo estudio, de forma que su función de densidad de probabilidad pertenece a una cierta familia paramétrica

$$S = \left\{ p(x, \boldsymbol{\theta}), \boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta \subset \mathbb{R}^k \right\}$$

con respecto a alguna medida  $\sigma$ -finita  $\mu$  en alguna  $\sigma$ -álgebra de los subconjuntos del espacio muestral. Suponiendo que  $S$  verifica determinadas condiciones de regularidad (Atkinson y Mitchell, 1981) puede definirse la distancia propuesta por Rao por el procedimiento siguiente.

Sean  $(\theta_1, \theta_2, \dots, \theta_k)$  y  $(\theta_1 + d\theta_1, \theta_2 + d\theta_2, \dots, \theta_k + d\theta_k)$  dos puntos contiguos del espacio paramétrico  $\Theta$ . La diferencia relativa entre las funciones de densidad de probabilidad correspondientes a dichos parámetros es

$$\sum_{i=1}^k \frac{\partial \ln p(x, \boldsymbol{\theta})}{\partial \theta_i} d\theta_i. \quad (4.2.1)$$

La distribución de (4.2.1) recoge las consecuencias de reemplazar  $\theta_i$  por  $\theta_i + d\theta_i$ ,  $i = 1, 2, \dots, k$ . En particular, Rao (1945) considera la varianza de (4.2.1), dada por

$$ds^2(\boldsymbol{\theta}) = \sum_{i,j=1}^k g_{ij}(\boldsymbol{\theta}) d\theta_i d\theta_j, \quad (4.2.2)$$

una forma diferencial cuadrática definida positiva, donde

$$g_{ij}(\boldsymbol{\theta}) = E \left[ \frac{\partial \ln p(x, \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial \ln p(x, \boldsymbol{\theta})}{\partial \theta_j} \right], \quad i, j = 1, 2, \dots, k, \quad (4.2.3)$$

es el elemento  $(i, j)$  de la matriz de información de Fisher, la cual se asume definida positiva para  $\boldsymbol{\theta} \in \Theta$ . La cantidad (4.2.2) es una métrica del espacio de Riemann inducida por la matriz de información de Fisher, con elemento del arco  $ds = (ds^2)^{1/2}$  y tensor métrico fundamental  $g_{ij}(\boldsymbol{\theta})$ ,  $i, j = 1, 2, \dots, k$ , sobre  $\Theta$ .

Rao (1945) propone usar la forma cuadrática (4.2.2), con tensor métrico fundamental  $g_{ij}(\boldsymbol{\theta})$ , como una medida de la divergencia entre dos distribuciones asociadas a parámetros que son puntos contiguos del espacio paramétrico.

Además, Rao (1949) propone (4.2.2) como medida de distancia entre distribuciones correspondientes a dos puntos cualesquiera del espacio paramétrico, no necesariamente contiguos,  $\boldsymbol{\theta}_1$  y  $\boldsymbol{\theta}_2$  de  $\Theta$ , o sea, entre las funciones de densidad de probabilidad  $p(x, \boldsymbol{\theta}_1)$  y  $p(x, \boldsymbol{\theta}_2)$ , la distancia geodésica entre  $\boldsymbol{\theta}_1$  y  $\boldsymbol{\theta}_2$ , con respecto a la métrica inducida por (4.2.3) (usando el elemento del arco (4.2.2)). Así, la distancia geodésica sobre  $\Theta$ , basada en la matriz de información de Fisher, recibe el nombre de distancia geodésica informacional o distancia de Rao y es utilizada como distancia entre dos densidades de probabilidad de  $S$ . Las siguientes propiedades (Cuadras, 1989) hacen que el uso de la distancia de Rao sea de gran interés:

- (i) Es invariante por transformaciones admisibles de los parámetros, entendiendo por tales las transformaciones biyectivas definidas a través de funciones diferenciables con continuidad.
- (ii) Es invariante por transformaciones admisibles de las variables.
- (iii) Es la generalización natural de la distancia de Mahalanobis (Mahalanobis, 1936) para cualquier distribución paramétrica. La distancia de Mahalanobis fue introducida en el ámbito de la distribución normal multivariante.
- (iv) Posee mayor poder de separación que otras distancias. En particular, que la distancia de Matusita (Matusita, 1955), extensamente utilizada en problemas de inferencia estadística, como se refleja en la relación

$$m(F_1 F_2) \leq s(F_1 F_2),$$

donde  $F_1$  y  $F_2$  son dos distribuciones,  $m$  representa la distancia de Matusita y  $s$  la distancia de Rao.

Las propiedades presentadas en (i) y (ii) son de gran interés, pues es razonable requerir que la distancia entre dos miembros de la familia sea invariante para cualquier transformación admisible de los parámetros. También, debe de ser invariante para transformaciones admisibles del vector aleatorio  $X$  (por ejemplo, transformaciones de localización y escala). Según se recoge en Cuadras (1989), "esta propiedad de invarianza prácticamente sólo la verifica la distancia de Rao" (véase, también, Cuadras et al., 1985 y Oller y Cuadras, 1987).

Como Atkinson y Mitchell (1981) afirman, las dificultades matemáticas para calcular la distancia de Rao para una familia particular de distribuciones son considerables. No obstante, algunos autores la han calculado para algunas familias de distribuciones de interés, véase, por ejemplo, Atkinson y Mitchell (1981), Burbea (1984), Mitchell y Krzanowski (1985), Oller y Cuadras (1985), Oller (1987) y Reverter y Oller (2003).

En particular, para la familia de distribuciones normales  $q$ -variantes con matriz de covarianzas definida positiva, Atkinson y Mitchell (1981) obtienen las siguientes expresiones para la distancia de Rao:

- (i) Si las poblaciones tienen vectores de medias  $\boldsymbol{\mu}_1$  y  $\boldsymbol{\mu}_2$ , y matriz de covarianzas común  $\boldsymbol{\Sigma}$ ,

$$s(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}) = [(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)]^{1/2},$$

es decir, la distancia de Mahalanobis de estas distribuciones.

- (ii) Si las poblaciones tienen media común  $\boldsymbol{\mu}$  y matrices de covarianzas  $\boldsymbol{\Sigma}_1$  y  $\boldsymbol{\Sigma}_2$ ,

$$s(\boldsymbol{\mu}, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) = \left[ \frac{1}{2} \sum_{i=1}^q \ln^2 [\lambda_i(\boldsymbol{\Sigma}_2 \boldsymbol{\Sigma}_1^{-1})] \right]^{1/2}, \quad (4.2.4)$$

donde para cualquier matriz cuadrada  $\mathbf{A}_{(q \times q)}$ ,  $\lambda_i(\mathbf{A})$ ,  $i = 1, \dots, q$ , representan sus autovalores.

**Nota 4.2.1** *El cálculo de la distancia de Rao entre modelos normales multivariantes con medias y matrices de covarianza diferentes es un problema aún no resuelto. Se obtienen algunos resultados parciales y aproximaciones en Oller y Cuadras (1983), Burbea (1984) y Krzanowski (1996).*

Un estudio amplio sobre esta distancia y sus aplicaciones está recogido en Cuadras (1989). En relación a la aplicación de la misma al análisis de influencia, se puede citar a Jiménez-Gamero et al. (2002), en el modelo lineal multivariante, y a Muñoz-Pichardo (2006), en el modelo lineal general multivariante y en la regresión logística.

En las siguientes secciones se define una expresión general de la distancia de Rao y su fórmula particular para dos distribuciones normales matriciales  $qr$ -variantes. Después se propone esta distancia como medida de influencia asociada al EMV de la matriz de coeficientes de regresión (y en algunas combinaciones lineales del mismo) del MCC y del MCC(SCS).

### 4.3. La medida de Rao en el modelo de curvas de crecimiento

Siguiendo las ideas presentadas en la sección 3.2 del capítulo anterior, en este apartado se recoge la aplicación de la distancia de Rao como medida de influencia sobre el EMV  $\widehat{\mathbf{B}}$  de la matriz de coeficientes de regresión del MCC, bajo la hipótesis de normalidad. En particular, se pretende estudiar la influencia de un subconjunto de observaciones en el EMV  $\widehat{\mathbf{B}}$ , y en combinaciones lineales de dicho estimador. El modelo perturbado usado es la omisión de un conjunto de observaciones con un número arbitrario de casos, o sea, el MCC( $I$ ) definido por (2.3.1).

Para la aplicación de la distancia de Rao como medida de influencia en el MCC, se considera la distribución normal matricial  $N_{q,r}(\mathbf{B}; c_1(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z} \mathbf{Z}^T)^{-1})$ , presentada en la sección 1.3, como aproximación a la distribución del EMV  $\widehat{\mathbf{B}}$ . Así,

$$E[\widehat{\mathbf{B}}] = \mathbf{B}$$

y

$$\text{Var}[\widehat{\mathbf{B}}] = c_1(\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1},$$

con  $c_1$  dado en (1.3.18). Como se recoge en el capítulo anterior, un estimador consistente de  $\text{Var}[\widehat{\mathbf{B}}]$  se obtiene substituyendo  $\boldsymbol{\Sigma}$  por su estimador insesgado  $\frac{1}{n-r} \mathbf{S}$ , resultando

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}] = \frac{c_1}{n-r} (\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes \mathbf{S}_{\mathbf{X}}, \quad (4.3.1)$$

matriz simétrica de dimensión  $(qr \times qr)$ , con  $\mathbf{S}_{\mathbf{X}} = (\mathbf{X}^T \mathbf{S}^{-1} \mathbf{X})^{-1}$ .

Considerando el MCC( $I$ ) presentado en la sección 2.3, la distribución del EMV  $\widehat{\mathbf{B}}_{(I)}$  puede ser aproximada por la distribución normal matricial

$$N_{q,r}(\mathbf{B}; c_6(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1}, (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1}),$$

así

$$E[\widehat{\mathbf{B}}_{(I)}] = \mathbf{B}$$

y

$$\text{Var}[\widehat{\mathbf{B}}_{(I)}] = c_6(\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X})^{-1},$$

con  $c_6$  dado en (2.3.3). Análogamente para el MCC( $I$ ), un estimador consistente de  $\text{Var}[\widehat{\mathbf{B}}_{(I)}]$  es obtenido substituyendo  $\boldsymbol{\Sigma}$  por el estimador  $\frac{1}{n-k-r} \mathbf{S}_{(I)}$ , por tanto

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] = \frac{c_6}{n-k-r} (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1}, \quad (4.3.2)$$

matriz simétrica de dimensión  $(qr \times qr)$ .

Así, se verifica que los estimadores  $\widehat{\mathbf{B}}$  y  $\widehat{\mathbf{B}}_{(I)}$  se pueden aproximar por distribuciones normales  $qr$ -variantes con la misma matriz esperanza y que difieren en sus matrices de covarianzas. Por tanto, siguiendo la aproximación genérica al problema de la influencia recogida en la sección 3.2, como  $\widehat{\text{Var}}[\widehat{\mathbf{B}}]$  es no singular, teniendo en cuenta (4.2.4), la versión muestral del cuadrado de

la distancia de Rao para evaluar la influencia que sobre la distribución de  $\widehat{\mathbf{B}}$  ejerce el conjunto de observaciones subindicadas por  $I$ , es

$$\rho_I^2(\widehat{\mathbf{B}}) = \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}] (\widehat{\text{Var}}[\widehat{\mathbf{B}}])^{-1} \right) \right], \quad (4.3.3)$$

donde  $\lambda_i(\mathbf{A})$ ,  $i = 1, \dots, qr$ , son los autovalores de la matriz cuadrada  $\mathbf{A}_{(qr \times qr)}$ .

**Nota 4.3.1** *Por simplicidad, como medida de influencia se utilizará el cuadrado de la distancia de Rao,  $\rho_I^2$ .*

Para utilizar (4.3.3) como medida de influencia en el EMV  $\widehat{\mathbf{B}}$  del MCC, son útiles los siguientes resultados consecuencia directa de los lemas 3.4.2 y 3.4.3.

**Lema 4.3.1** *Sean  $\mathbf{Z}$  y  $\mathbf{Z}_{(I)}$  las matrices del diseño entre individuos del MCC y del MCC(I), respectivamente, entonces*

$$(\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} (\mathbf{Z} \mathbf{Z}^T) = \mathbf{I}_r + (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T,$$

donde  $\mathbf{H}_I$  es dada en (2.3.7).

**Lema 4.3.2** *Sean  $\mathbf{S}$  y  $\mathbf{S}_{(I)}$  las matrices definidas por (1.3.5) y (2.3.2), respectivamente, para el MCC y MCC(I). Entonces*

$$(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{S}_{\mathbf{X}}^{-1} = \mathbf{I}_q - \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X},$$

con  $\mathbf{E}_I$  y  $\mathbf{V}_I$  dadas en (2.3.6) y (2.3.10), respectivamente.

En el siguiente teorema se presenta la distancia de Rao como medida para estudiar la influencia de  $\mathbf{Y}_I$  en el EMV  $\widehat{\mathbf{B}}$  del MCC.

**Teorema 4.3.1** *Para el MCC, la medida de Rao asociada al subconjunto  $\mathbf{Y}_I$  sobre el EMV  $\widehat{\mathbf{B}}$  es*

$$\rho_I^2(\widehat{\mathbf{B}}) = \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 [c_{10} (1 + \lambda_j(\mathbf{L}_I)) (1 - \mu_h(\mathbf{R}_I))], \quad (4.3.4)$$

con

$$\mathbf{L}_I = (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T, \quad (4.3.5)$$

$$\mathbf{R}_I = \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \quad (4.3.6)$$

y

$$c_{10} = \frac{c_6(n-r)}{c_1(n-k-r)}, \quad (4.3.7)$$

donde  $\lambda_j(\mathbf{L}_I)$ ,  $j = 1, \dots, r$ , y  $\mu_h(\mathbf{R}_I)$ ,  $h = 1, \dots, q$ , son los autovalores de las matrices cuadradas  $\mathbf{L}_I$  y  $\mathbf{R}_I$ , respectivamente.

**Demostración:** Teniendo en cuenta los estimadores  $\widehat{\text{Var}}[\widehat{\mathbf{B}}]$  y  $\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}]$  definidos en (4.3.1) y (4.3.2), respectivamente, las propiedades (vi) y (vii) del producto de Kronencker dadas por el teorema B.0.5 y los lemas 4.3.1 y 4.3.2, se tiene que

$$\begin{aligned} & \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}](\widehat{\text{Var}}[\widehat{\mathbf{B}}])^{-1} = \\ & = \left( \frac{c_6}{n-k-r} (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \otimes (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \right) \left( \frac{c_1}{n-r} (\mathbf{Z} \mathbf{Z}^T)^{-1} \otimes \mathbf{S}_{\mathbf{X}} \right)^{-1} = \\ & = \frac{c_6(n-r)}{c_1(n-k-r)} \left( (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} (\mathbf{Z} \mathbf{Z}^T) \otimes (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{S}_{\mathbf{X}}^{-1} \right) = \\ & = c_{10} (\mathbf{I}_r + (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T) \otimes (\mathbf{I}_q - \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X}), \end{aligned}$$

matriz simétrica de dimensión  $(qr \times qr)$ . Así, de (4.3.3),

$$\begin{aligned} \rho_I^2 &= \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\widehat{\mathbf{B}}_{(I)}](\widehat{\text{Var}}[\widehat{\mathbf{B}}])^{-1} \right) \right] = \\ &= \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ c_{10} \lambda_i \left( (\mathbf{I}_r + (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T) \otimes \right. \right. \\ & \quad \left. \left. \otimes (\mathbf{I}_q - \mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X}) \right) \right] = \\ &= \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 \left[ c_{10} \left( 1 + \lambda_j \left( (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T \right) \right) \right. \\ & \quad \left. (1 - \mu_h (\mathbf{S}_{\mathbf{X}} \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X})) \right]. \end{aligned}$$

La última igualdad resulta de la aplicación de la propiedad (xii) del teorema B.0.5 y del hecho de que el producto de Kronecker es de orden  $qr$ . ■

Un valor grande de  $\rho_I^2$  indica que el correspondiente subconjunto  $\mathbf{Y}_I$  es influyente en la estimación de  $\mathbf{B}$ , dado que su omisión provoca una considerable variación en la distribución de  $\widehat{\mathbf{B}}$ .

Del teorema 4.3.1, se concluye que  $\rho_I^2$  es función de dos componentes,  $\mathbf{L}_I$  y  $\mathbf{R}_I$  dadas por (4.3.5) y (4.3.6), respectivamente. A continuación, se interpretan las cantidades  $\lambda_j(\mathbf{L}_I)$ ,  $j = 1, \dots, r$ , y  $\mu_h(\mathbf{R}_I)$ ,  $h = 1, \dots, q$ .

Se comienza por el caso más simple, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , es decir, cuando se elimina una observación del modelo. En este sentido, usando la medida de Rao para estudiar la influencia del subconjunto  $\mathbf{Y}_I$  sobre la distribución del EMV  $\widehat{\mathbf{B}}$ , se obtiene la distancia de Rao como medida de influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  sobre  $\widehat{\mathbf{B}}$  en el MCC.

**Corolario 4.3.1** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida de Rao asociada al EMV  $\widehat{\mathbf{B}}$  es*

$$\rho_i^2(\widehat{\mathbf{B}}) = \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 [c_{11} (1 + \lambda_j(\mathbf{L}_i)) (1 - \mu_h(\mathbf{R}_i))]$$

con

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T,$$



$$\mathbf{R}_i = \frac{1}{v_{ii}} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X}$$

y

$$c_{11} = \frac{c_9(n-r)}{c_1(n-r-1)}, \quad (4.3.8)$$

donde  $p_{ii}$ ,  $v_{ii}$  y  $\mathbf{e}_i$  están definidos por (2.3.15), (2.3.13) y (2.3.14), respectivamente.

En este caso, para la componente  $\mathbf{L}_i$ , se obtiene

$$\lambda_j(\mathbf{L}_i) = \lambda_j \left( \frac{1}{1-p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T \right) = \frac{1}{1-p_{ii}} \lambda_j \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T \right)$$

y, teniendo en cuenta el lema B.0.2, la matriz  $(\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T$  tiene solamente un autovalor no nulo  $p_{ii} = \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i$ , el elemento  $i$ -ésimo de la diagonal principal de la matriz de proyección  $\mathbf{P}_{\mathbf{Z}^T}$ . Por tanto,

$$\lambda_j \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T \right) = \begin{cases} p_{ii}, & j = 1, i = 1, \dots, n \\ 0, & j = 2, \dots, r, i = 1, \dots, n \end{cases},$$

es decir, todos los autovalores de la matriz  $\mathbf{L}_i$  son nulos con excepción de  $\lambda_1$ , así

$$\lambda_j(\mathbf{L}_i) = \begin{cases} \frac{p_{ii}}{1-p_{ii}}, & j = 1, i = 1, \dots, n \\ 0, & j = 2, \dots, r, i = 1, \dots, n \end{cases}. \quad (4.3.9)$$

En consecuencia,  $\lambda_j(\mathbf{L}_i)$  es constante en cada grupo.

**Nota 4.3.2** En el modelo lineal multivariante, el factor  $\frac{p_{ii}}{1-p_{ii}}$  es conocido como el potencial de la  $i$ -ésima observación  $\mathbf{y}_i$ , una medida usada para detectar los casos denominados como "leverage" del espacio de  $\mathbf{Z}$ . En el MCC no tiene sentido dicha designación dado que  $p_{ii}$  es constante en cada grupo al depender sólo de  $\mathbf{Z}$ , véase nota 2.5.1.

Análogamente para la componente  $\mathbf{R}_i$ , se obtiene

$$\mu_h(\mathbf{R}_i) = \mu_h \left( \frac{1}{v_{ii}} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i v_{ii}^{-1} \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \right) = \frac{1}{v_{ii}} \mu_h \left( \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \right),$$

donde  $\mathbf{e}_i$  es la columna  $i$ -ésima de la matriz del residuo  $\mathbf{E}$ . Del lema B.0.2, resulta que la matriz  $\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X}$  tiene solamente el autovalor  $\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i$  no nulo, así

$$\mu_h(\mathbf{R}_i) = \begin{cases} \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}}, & h = 1, i = 1, \dots, n \\ 0, & h = 2, \dots, q, i = 1, \dots, n \end{cases},$$

o sea, todos los autovalores de la matriz  $\mathbf{R}_i$  son cero con excepción de  $\mu_1$ .

Como el estadístico para determinar si la observación  $i$ -ésima es outlier en el MCC,  $T_i$  dado por (2.4.2), es

$$T_i = \Lambda_i^{-1} = 1 + \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i} = \frac{v_{ii}}{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i},$$

y dado que, por (2.3.13),

$$\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i = v_{ii} - (1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i)$$

entonces

$$\begin{aligned} \mu_1 &= \frac{\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} = \frac{v_{ii} - (1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i)}{v_{ii}} = \\ &= 1 - \frac{1 - p_{ii} - \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i}{v_{ii}} = 1 - \frac{1}{T_i} = \frac{T_i - 1}{T_i} = 1 - \Lambda_i, \end{aligned}$$

por tanto,

$$\mu_h(\mathbf{R}_i) = \begin{cases} \frac{T_i - 1}{T_i}, & h = 1, \quad i = 1, \dots, n \\ 0, & h = 2, \dots, q, \quad i = 1, \dots, n \end{cases}. \quad (4.3.10)$$

La cantidad  $\frac{T_i - 1}{T_i}$  es una función creciente de  $\mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{e}_i$  (función del residuo internamente studentizado en el modelo lineal multivariante) y es también usada como estadístico para detectar outliers en el MCC (véase Srivastava y von Rosen, 1998).

Así, la medida de Rao,  $\rho_i^2(\hat{\mathbf{B}})$ , para el MCC, puede ser interpretada como función de dos componentes:  $\mathbf{L}_i$  constante en cada grupo del modelo y  $\mathbf{R}_i$  que puede ser interpretada como una componente para la detección de outliers basada en los residuos.

Teniendo en cuenta el corolario 4.3.1 y los autovalores (4.3.9) y (4.3.10), la medida de Rao,  $\rho_i^2$ , se reduce a

$$\begin{aligned} \rho_i^2 &= \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \Lambda_i \right] + \frac{r - 1}{2} \ln^2 [c_{11} \Lambda_i] + \\ &+ \frac{q - 1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \right] + \frac{(q - 1)(r - 1)}{2} \ln^2 [c_{11}], \end{aligned}$$

con  $c_{11}$  dado por (4.3.8) y  $\Lambda_i = T_i^{-1}$ . O sea,  $\rho_i^2$  es una función creciente de  $T_i$ .

Para el caso general, la interpretación de las cantidades  $\lambda_j(\mathbf{L}_I)$ ,  $j = 1, \dots, r$ , y  $\mu_h(\mathbf{R}_I)$ ,  $h = 1, \dots, q$ , dadas por el teorema 4.3.1, es análoga al caso anterior, para  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ . En este caso, teniendo en cuenta el lema B.0.2, se obtiene que la matriz  $\mathbf{L}_I$ , dada por (4.3.5), tiene los mismos autovalores no nulos, con la misma multiplicidad, que la matriz

$$\mathbf{L}_I^* = (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I = (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{H}_I.$$

La igualdad en la multiplicidad de los autovalores no nulos de  $\mathbf{L}_I$  y  $\mathbf{L}_I^*$  se justifica por el siguiente resultado. Teniendo en cuenta el lema B.0.1, considerando las matrices  $\mathbf{A}_{(r \times k)} = (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}_I$  y  $\mathbf{C}_{(k \times r)} = (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T$ , resulta que

$$\det [\mathbf{I}_r - \mathbf{A} \mathbf{C}] = \det [\mathbf{I}_k - \mathbf{C} \mathbf{A}],$$

entonces, si  $k > r$

$$\det [\lambda \mathbf{I}_k - \mathbf{CA}] = \lambda^{k-r} \det [\lambda \mathbf{I}_r - \mathbf{AC}],$$

por tanto para  $\lambda \neq 0$ , los autovalores no nulos de  $\mathbf{L}_I^* = \mathbf{CA}$  son los mismos que de  $\mathbf{L}_I = \mathbf{AC}$ , además la multiplicidad de  $\lambda$  en  $\mathbf{L}_I^* = \mathbf{CA}$  es la misma que de  $\lambda$  en  $\mathbf{L}_I = \mathbf{AC}$ .

Por otro lado, también del lema B.0.2, se obtiene que la matriz  $\mathbf{R}_I$ , dada por (4.3.6), tiene los mismos autovalores no nulos, con la misma multiplicidad, que la matriz

$$\mathbf{R}_I^* = \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I.$$

Como de la expresión de  $\mathbf{V}_I$ , dada por (2.3.10), resulta que

$$\mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I = \mathbf{V}_I - (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I),$$

entonces

$$\mathbf{R}_I^* = \mathbf{I}_k - \mathbf{V}_I^{-1} (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I),$$

o sea, los autovalores no nulos de  $\mathbf{R}_I$  son

$$1 - \mu_h^* (\mathbf{V}_I^{-1} (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I))$$

y, por tanto,

$$1 - \mu_h(\mathbf{R}_I) = \mu_h^* (\mathbf{V}_I^{-1} (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)).$$

Además, como

$$\det [\mathbf{V}_I^{-1} (\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I)] = \frac{\det [\mathbf{I}_k - \mathbf{H}_I - \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{E}_I]}{\det [\mathbf{V}_I]} = T_I^{-1} = \Lambda_I,$$

se concluye que, en el caso general,  $\mathbf{R}_I$  puede ser considerada como una matriz basada en los residuos para detectar outlier y, así, valores grandes de  $\mu_1(\mathbf{R}_I), \mu_2(\mathbf{R}_I), \dots, \mu_q(\mathbf{R}_I)$  indican que el conjunto de observaciones indexada por  $I$  no sigue el patrón sugerido por las otras  $n - k$  observaciones.

**Nota 4.3.3** En la mayoría de los modelos estadísticos, las medidas de influencia se pueden expresar como función de una componente "leverage" y de una componente residual (véase, por ejemplo, Barrett y Ling, 1992). En el MCC, mientras la matriz  $\mathbf{R}_I$  puede ser considerada como una componente residual, la matriz  $\mathbf{L}_I$  no puede ser considerada una componente "leverage". Teniendo en cuenta la definición del MCC,  $\mathbf{Z}$  es la matriz indicador del grupo compuesta por 1 y 0, que indican el grupo del cual fue omitida una observación, entonces, la matriz  $\mathbf{P}_{\mathbf{Z}^T} = \mathbf{H} = \mathbf{Z}^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{Z}$  es constante en cada grupo y no puede ser considerada como una matriz "leverage". Para la obtención de una matriz "leverage" generalizada para el MCC se puede intentar usar el procedimiento recogido en Wei et al. (1998). Por otra parte, la matriz  $\mathbf{L}_I$ , y consecuentemente  $\mathbf{L}_I^*$ , no es constante en cada grupo, dado que cuando se elimina el conjunto  $I$  no todas las observaciones pertenecen al mismo grupo.

La siguiente descomposición, sugerida por Jiménez-Gamero et al. (2002), muestra la relación de  $\rho_I^2$  con las componentes  $\mathbf{L}_I$  y  $\mathbf{R}_I$ .

**Teorema 4.3.2** *En el MCC, sea*

$$\mathcal{L}_I = (\mathbf{1}_q \otimes \mathbf{a}), \text{ donde } \mathbf{a}^T = (\ln [1 + \lambda_1(\mathbf{L}_I)], \dots, \ln [1 + \lambda_r(\mathbf{L}_I)])$$

y

$$\mathcal{R}_I = (\mathbf{b} \otimes \mathbf{1}_r), \text{ donde } \mathbf{b}^T = (\ln [c_{10} (1 - \mu_1)(\mathbf{R}_I)], \dots, \ln [c_{10} (1 - \mu_q)(\mathbf{R}_I)]),$$

con  $c_{10}$  dada por (4.3.7), entonces

$$2\rho_I^2 = \|\mathcal{L}_I\|^2 + \|\mathcal{R}_I\|^2 + 2\|\mathcal{L}_I\| \|\mathcal{R}_I\| \cos \theta_I,$$

donde  $\|\cdot\|$  es la norma euclídea y  $\theta_I$  es el ángulo entre los vectores  $\mathcal{L}_I$  y  $\mathcal{R}_I$ , siendo  $\mathbf{1}_s = (1, \dots, 1) \in \mathbb{R}^s$ .

**Demostración:** Teniendo en cuenta (4.3.4), se obtiene

$$\begin{aligned} 2\rho_I^2 &= q \sum_{j=1}^r \ln^2 [1 + \lambda_j(\mathbf{L}_I)] + 2 \sum_{h=1}^q \sum_{j=1}^r \ln [1 + \lambda_j(\mathbf{L}_I)] \ln [c_{10} (1 - \mu_h(\mathbf{R}_I))] + \\ &+ r \sum_{h=1}^q \ln^2 [c_{10} (1 - \mu_h(\mathbf{R}_I))], \end{aligned}$$

y, dado que

$$\begin{aligned} \|\mathcal{L}_I\|^2 &= q \sum_{j=1}^r \ln^2 [1 + \lambda_j(\mathbf{L}_I)], \\ \|\mathcal{R}_I\|^2 &= r \sum_{h=1}^q \ln^2 [c_{10} (1 - \mu_h(\mathbf{R}_I))] \end{aligned}$$

y

$$\|\mathcal{L}_I\| \|\mathcal{R}_I\| \cos \theta_I = \mathcal{L}_I^T \cdot \mathcal{R}_I = \sum_{h=1}^q \sum_{j=1}^r \ln [1 + \lambda_j(\mathbf{L}_I)] \ln [c_{10} (1 - \mu_h(\mathbf{R}_I))],$$

se obtiene el resultado enunciado. ■

Así,  $\rho_I^2$  se puede expresar en función de tres componentes:  $\|\mathcal{L}_I\|^2$ ,  $\|\mathcal{R}_I\|^2$  y  $\|\mathcal{L}_I\| \|\mathcal{R}_I\| \cos \theta_I$ , que puede interpretarse como la interacción entre las dos componentes anteriores.

Teniendo en cuenta los autovalores (4.3.9) y (4.3.10), se puede aplicar el teorema 4.3.2 a la medida de Rao,  $\rho_i^2$ , asociada a la  $i$ -ésima observación de  $\mathbf{Y}$  sobre el EMV  $\hat{\mathbf{B}}$  del MCC, como se recoge a continuación.

**Nota 4.3.4** *A partir del teorema anterior, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , se tiene*

$$\mathcal{L}_i = (\mathbf{1}_q \otimes \mathbf{a}), \text{ donde } \mathbf{a}^T = \left( \ln \left[ \frac{1}{1 - p_{ii}} \right], 0, \dots, 0 \right)$$

y

$$\mathcal{R}_i = (\mathbf{b} \otimes \mathbf{1}_r), \text{ donde } \mathbf{b}^T = (\ln [c_{11} \Lambda_i], \ln [c_{11}], \dots, \ln [c_{11}]),$$

con  $c_{11}$  dado por (4.3.8),  $\Lambda_i = T_i^{-1}$  y  $\mathbf{1}_s = (1, \dots, 1) \in \mathbb{R}^s$ . En consecuencia, la medida de Rao,  $\rho_i^2$ , puede ser descompuesta por

$$2\rho_i^2 = \|\mathcal{L}_i\|^2 + \|\mathcal{R}_i\|^2 + 2\|\mathcal{L}_i\|\|\mathcal{R}_i\|\cos\theta_i,$$

donde

$$\|\mathcal{L}_i\|^2 = q \sum_{j=1}^r \ln^2 [1 + \lambda_j(\mathbf{L}_i)] = q \ln^2 \left[ \frac{1}{1 - p_{ii}} \right],$$

$$\|\mathcal{R}_i\|^2 = r \sum_{h=1}^q \ln^2 [c_{11}(1 - \mu_h(\mathbf{R}_i))] = r \ln^2 [c_{11}\Lambda_i] + (q-1)r \ln^2 [c_{11}]$$

y

$$\begin{aligned} \|\mathcal{R}_i\|\|\mathcal{L}_i\|\cos\theta_i &= \mathcal{L}_i^T \cdot \mathcal{R}_i = \sum_{h=1}^q \sum_{j=1}^r \ln [1 + \lambda_j(\mathbf{L}_i)] \ln [c_{11}(1 - \mu_h(\mathbf{R}_i))] = \\ &= \ln \left[ \frac{1}{1 - p_{ii}} \right] \ln [c_{11}\Lambda_i] + (q-1) \ln \left[ \frac{1}{1 - p_{ii}} \right] \ln [c_{11}] = \\ &= \ln \left[ \frac{1}{1 - p_{ii}} \right] (\ln [\Lambda_i] + q \ln [c_{11}]). \end{aligned}$$

#### 4.3.1. Influencia en combinaciones lineales de $\hat{\mathbf{B}}$

La distancia de Rao puede ser extendida como medida de influencia sobre las combinaciones lineales del EMV de la matriz de coeficientes de regresión  $\mathbf{B}$  del MCC. Sea  $\Psi = \mathbf{D}\mathbf{B}\mathbf{L}$ , con  $\mathbf{D}$  y  $\mathbf{L}$  matrices de dimensiones  $(s \times q)$  y  $(r \times l)$ , respectivamente, donde  $\text{rg}(\mathbf{D}) = s \leq q$  y  $\text{rg}(\mathbf{L}) = l \leq r$ . Teniendo en cuenta Kollo y von Rosen (2005), el EMV de  $\Psi$  es

$$\hat{\Psi} = \mathbf{D}_{(s \times q)} \hat{\mathbf{B}}_{(q \times r)} \mathbf{L}_{(r \times l)}. \quad (4.3.11)$$

En la sección anterior, para la aplicación de la distancia de Rao como medida de influencia, la distribución de  $\hat{\mathbf{B}}$  fue aproximada por la distribución normal

$$N_{q,r}(\mathbf{B}; c_1(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1}, (\mathbf{Z}\mathbf{Z}^T)^{-1}),$$

por el teorema C.1.2 se puede aproximar la distribución del EMV  $\hat{\Psi}_{(s \times l)} = \mathbf{D}\hat{\mathbf{B}}\mathbf{L}$  a través de la distribución normal

$$N_{s,l}(\text{vec}(\mathbf{D}\hat{\mathbf{B}}\mathbf{L}); c_1 \mathbf{D}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{D}^T, \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L}),$$

por lo tanto,

$$\widehat{\text{Var}}[\mathbf{D}\hat{\mathbf{B}}\mathbf{L}] = \frac{c_1}{n-r} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D}\mathbf{S}_X \mathbf{D}^T, \quad (4.3.12)$$

matriz simétrica de dimensión  $(sl \times sl)$ .

Por otro lado, considerando las combinaciones lineales de  $\hat{\mathbf{B}}$  en el MCC(I), la distribución del EMV  $\hat{\Psi}_{(I)} = \mathbf{D}\hat{\mathbf{B}}_{(I)}\mathbf{L}$  puede ser aproximada por la distribución normal

$$N_{s,l}(\mathbf{D}\hat{\mathbf{B}}\mathbf{L}; c_6 \mathbf{D}(\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{D}^T, \mathbf{L}^T (\mathbf{Z}_{(I)} \mathbf{Z}_{(I)}^T)^{-1} \mathbf{L})$$

y

$$\widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}] = \frac{c_6}{n-k-r} \mathbf{L}^T (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T, \quad (4.3.13)$$

matriz simétrica de dimensión  $(sl \times sl)$ .

Por tanto, se verifica que la distribución de los EMV  $\widehat{\Psi}$  y  $\widehat{\Psi}_{(I)}$  pueden ser aproximadas por distribuciones normales  $sl$ -variantes con esperanza común y matrices de covarianza diferentes. Teniendo en cuenta (4.2.4), como  $\widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}\mathbf{L}]$  es no singular, la versión muestral del cuadrado de la distancia de Rao, para evaluar la influencia que sobre la distribución de  $\widehat{\Psi} = \mathbf{D}\widehat{\mathbf{B}}\mathbf{L}$  ejerce el conjunto de observaciones subindicadas por  $I$ , es

$$\rho_I^2(\widehat{\Psi}) = \frac{1}{2} \sum_{i=1}^{sl} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}] (\widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}\mathbf{L}])^{-1} \right) \right], \quad (4.3.14)$$

donde  $\lambda_i(\mathbf{A})$ ,  $i = 1, \dots, sl$ , representan los autovalores de la matriz cuadrada  $\mathbf{A}_{(sl \times sl)}$ .

Usando la expresión anterior, la medida de Rao para estudiar la influencia del subconjunto de observaciones  $\mathbf{Y}_I$  en combinaciones lineales del EMV  $\widehat{\mathbf{B}}$  para el MCC es presentada en el siguiente teorema.

**Teorema 4.3.3** *En el MCC, la medida de Rao asociada a  $\mathbf{Y}_I$  sobre  $\widehat{\Psi}$  es*

$$\rho_I^2(\widehat{\Psi}) = \frac{1}{2} \sum_{a=1}^l \sum_{b=1}^s \ln^2 [c_{10} (1 + \lambda_a(\mathbf{L}_I)) (1 - \mu_b(\mathbf{R}_I))],$$

siendo

$$\mathbf{L}_I = \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1}$$

y

$$\mathbf{R}_I = \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1},$$

con  $c_{10}$  dada por (4.3.7), donde  $\lambda_a(\mathbf{L}_I)$ ,  $a = 1, \dots, l$ , y  $\mu_b(\mathbf{R}_I)$ ,  $b = 1, \dots, s$ , representan los autovalores de las matrices cuadradas  $\mathbf{R}_I$  y  $\mathbf{L}_I$ , respectivamente.

**Demostración:** De los lemas 4.3.1 y 4.3.2, resulta

$$\begin{aligned} \left( \mathbf{L}^T (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \right) \left( \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \right)^{-1} &= \mathbf{I}_l + \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \\ &\quad \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \end{aligned}$$

y

$$\begin{aligned} \left( \mathbf{D} (\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right) (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1} &= \mathbf{I}_s - \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \\ &\quad \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1}. \end{aligned}$$

Por tanto, teniendo en cuenta los estimadores de las matrices de covarianza, (4.3.12) y (4.3.13), y el teorema B.0.5, se obtiene

$$\begin{aligned}
& \widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}] \left( \widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}\mathbf{L}] \right)^{-1} = \\
& = \left( \frac{c_6}{n-k-r} \mathbf{L}^T (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} \otimes \mathbf{D}(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T \right) \\
& \quad \left( \frac{c_1}{n-r} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} \otimes \mathbf{D}\mathbf{S}_X \mathbf{D}^T \right)^{-1} = \\
& = c_{10} \left( \mathbf{L}^T (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \otimes \mathbf{D}(\mathbf{X}^T \mathbf{S}_{(I)}^{-1} \mathbf{X})^{-1} \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1} \right) = \\
& = c_{10} \left( \mathbf{I}_l + \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \right) \otimes \\
& \quad \otimes \left( \mathbf{I}_s - \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1} \right),
\end{aligned}$$

matriz simétrica de dimensión  $(sl \times sl)$ . Así, de (4.3.14),

$$\begin{aligned}
\rho_I^2(\widehat{\Psi}) &= \frac{1}{2} \sum_{i=1}^{sl} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}_{(I)}\mathbf{L}] \left( \widehat{\text{Var}}[\mathbf{D}\widehat{\mathbf{B}}\mathbf{L}] \right)^{-1} \right) \right] = \\
&= \frac{1}{2} \sum_{i=1}^{sl} \ln^2 \left[ c_{10} \lambda_i \left( \mathbf{I}_s - \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1} \right) \otimes \right. \\
& \quad \left. \otimes \left( \mathbf{I}_l + \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \right) \right] = \\
&= \frac{1}{2} \sum_{a=1}^l \sum_{b=1}^s \ln^2 \left[ c_{10} \left( 1 - \mu_b \left( \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{E}_I \mathbf{V}_I^{-1} \mathbf{E}_I^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1} \right) \right) \right. \\
& \quad \left. \left( 1 + \lambda_a \left( \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1} \right) \right) \right].
\end{aligned}$$

Como el producto de Kronecker es de orden  $sl$ , la demostración queda concluida. ■

La medida de Rao, para estudiar la influencia del subconjunto  $\mathbf{Y}_I$  sobre las combinaciones lineales,  $\widehat{\Psi} = \mathbf{D}\widehat{\mathbf{B}}\mathbf{L}$ , se puede particularizar como medida de influencia de la  $i$ -ésima observación de  $\mathbf{Y}$ .

**Corolario 4.3.2** *En el MCC, cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ ,*

$$\rho_i^2(\widehat{\Psi}) = \frac{1}{2} \sum_{a=1}^l \sum_{b=1}^s \ln^2 [c_{11} (1 + \lambda_a(\mathbf{L}_i)) (1 - \mu_b(\mathbf{R}_i))] \quad (4.3.15)$$

con

$$\begin{aligned}
\mathbf{L}_i &= \frac{1}{1 - p_{ii}} \mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1}, \\
\mathbf{R}_i &= \frac{1}{v_{ii}} \mathbf{D}\mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X} \mathbf{S}_X \mathbf{D}^T (\mathbf{D}\mathbf{S}_X \mathbf{D}^T)^{-1}
\end{aligned}$$

y  $c_{11}$  dado por (4.3.8).

Además, por el lema B.0.2, se concluye que las matrices  $\mathbf{L}_i$  y  $\mathbf{R}_i$  implicadas en la medida de Rao (4.3.15) tienen solamente un autovalor no nulo

El corolario 4.3.2 proporciona la distancia de Rao, para el MCC, como medida de influencia cuando se pretende analizar el efecto de la  $i$ -ésima observación de  $\mathbf{Y}$  en un determinado subconjunto de  $\widehat{\mathbf{B}}$ , basta para ello sustituir  $\mathbf{D}$  y  $\mathbf{L}$  por matrices adecuadas. A continuación, se presenta la distancia de Rao como medida de influencia para tres posibles elecciones de las matrices  $\mathbf{D}$  y  $\mathbf{L}$ .

#### 4.3.1.1. Influencia en las filas de $\widehat{\mathbf{B}}$

Para estudiar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $j$ -ésima fila de  $\widehat{\mathbf{B}}$ , se puede escoger  $\mathbf{L} = \mathbf{I}_r$ ,  $\text{rg}(\mathbf{L}) = r$  y  $\mathbf{D} = \mathbf{d}_j^T$ ,  $\text{rg}(\mathbf{D}) = 1 \leq q$ , donde  $\mathbf{d}_j$  es un vector de dimensión  $q$  con la  $j$ -ésima componente la unidad y las restantes cero y, de (4.3.11), se tiene  $\widehat{\Psi}_{(1 \times r)} = \mathbf{d}_j^T \widehat{\mathbf{B}}_{(q \times r)}$ . Por tanto, del corolario 4.3.2, se obtiene

$$\mathbf{R}_i = \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \in \mathbb{R}$$

y

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T.$$

Teniendo en cuenta (4.3.12) y (4.3.13), se observa que

$$\widehat{\text{Var}}[\mathbf{d}_j^T \widehat{\mathbf{B}}_{(i)}] (\widehat{\text{Var}}[\mathbf{d}_j^T \widehat{\mathbf{B}}])^{-1}$$

es una matriz simétrica de dimensión  $(r \times r)$ , concluyéndose que la medida de Rao para estudiar la influencia de  $\mathbf{y}_i$  sobre la  $j$ -ésima fila de  $\widehat{\mathbf{B}}$  en el MCC es

$$\rho_{i(j \cdot)}^2(\widehat{\mathbf{B}}) = \frac{1}{2} \sum_{a=1}^r \ln^2 [c_{11} (1 - \mathbf{R}_i) (1 + \lambda_a(\mathbf{L}_i))]. \quad (4.3.16)$$

Para el desarrollo de la medida de Rao para las filas de  $\widehat{\mathbf{B}}$ , se observa que para  $\rho_{i(j \cdot)}^2$ , la componente  $\mathbf{L}_i$  es la misma que para la medida de Rao,  $\rho_i^2$ , por tanto, usando (4.3.9), de (4.3.16) se obtiene

$$\begin{aligned} \rho_{i(j \cdot)}^2 &= \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \left( 1 - \frac{1}{v_{ii}} \left( \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right) \right) \right] + \\ &+ \frac{r-1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{1}{v_{ii}} \left( \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right) \right]. \end{aligned} \quad (4.3.17)$$



**Nota 4.3.5** En el MCC, para el estudio de la influencia en las filas de  $\widehat{\mathbf{B}}$ , la medida de Rao,  $\rho_{i(j)}^2$ , puede ser expresada como función de la componente de localización,  $LC_{i(j)}^2$ , de la medida de Fréchet  $\delta_{i(j)}$ . En efecto, de (3.4.47), se obtiene

$$\frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} = \frac{v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_{i(j)}^2,$$

por tanto, de (4.3.17),

$$\begin{aligned} \rho_{i(j)}^2 &= \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \left( 1 - \frac{v_{ii}}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_{i(j)}^2 \right) \right] + \\ &+ \frac{r-1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{v_{ii}}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-2} \mathbf{k}_i} LC_{i(j)}^2 \right) \right]. \end{aligned}$$

**Nota 4.3.6** Teniendo en cuenta que

$$\frac{v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i} = T_i \Delta_i^{-1},$$

(véase nota 3.4.3), de la expresión de  $DL_{i(j)}$ , dada en (2.4.9), se concluye que

$$\frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} = \frac{r v_{ii}}{\mathbf{k}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{k}_i} DL_{i(j)} = r \frac{T_i}{\Delta_i} DL_{i(j)}.$$

De este modo,

$$\rho_{i(j)}^2 = \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \left( 1 - \frac{r T_i}{\Delta_i} DL_{i(j)} \right) \right] + \frac{r-1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{r T_i}{\Delta_i} DL_{i(j)} \right) \right].$$

Es decir, en el MCC, la medida de Rao,  $\rho_{i(j)}^2$ , asociada a la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  sobre las filas de  $\widehat{\mathbf{B}}$ , puede ser expresada como función de la medida de Cook,  $DL_{i(j)}$ , del estadístico  $T_i$ , una medida que permite identificar outliers, y de la cantidad  $\Delta_i$ , dada en (3.4.51), la cual es una distancia generalizada en el espacio  $\mathcal{C}(\mathbf{Z})$ .

#### 4.3.1.2. Influencia en las columnas de $\widehat{\mathbf{B}}$

Para evaluar la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en la  $m$ -ésima columna del EMV  $\widehat{\mathbf{B}}$  del MCC, se puede escoger  $\mathbf{D} = \mathbf{I}_q$ ,  $\text{rg}(\mathbf{D}) = q$  y  $\mathbf{L} = \mathbf{h}_m$ ,  $\text{rg}(\mathbf{L}) = 1 \leq r$ , con  $\mathbf{h}_m \in \mathbb{R}^r$  un vector de dimensión  $r$  donde la  $m$ -ésima componente es la unidad y las restantes son cero, y (4.3.11) se reduce a  $\widehat{\Psi}_{(q \times 1)} = \widehat{\mathbf{B}}_{(q \times r)} \mathbf{h}_m (r \times 1)$ . Por tanto, del corolario 4.3.2, se obtiene

$$\mathbf{R}_i = \frac{1}{v_{ii}} \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i \mathbf{e}_i^T \mathbf{S}^{-1} \mathbf{X}$$

y

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m} \in \mathbb{R}.$$

Como, por (4.3.12) y (4.3.13),

$$\widehat{\text{Var}}[\widehat{\mathbf{B}}_{(i)} \mathbf{h}_m] (\widehat{\text{Var}}[\widehat{\mathbf{B}} \mathbf{h}_m])^{-1}$$

es una matriz simétrica de dimensión  $(q \times q)$ , se obtiene la medida de Rao para estudiar la influencia de  $\mathbf{y}_i$  en la  $m$ -ésima columna de  $\widehat{\mathbf{B}}$  para el MCC

$$\rho_{i(\cdot m)}^2(\widehat{\mathbf{B}}) = \frac{1}{2} \sum_{b=1}^q \ln^2 [c_{11} (1 + \mathbf{L}_i) (1 - \mu_b(\mathbf{R}_i))]. \quad (4.3.18)$$

Como  $\rho_{i(\cdot m)}^2$  tiene la misma componente  $\mathbf{R}_i$  que  $\rho_i^2$ , usando (4.3.10) y (4.3.18),

$$\begin{aligned} \rho_{i(\cdot m)}^2 &= \frac{1}{2} \ln^2 \left[ c_{11} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \Lambda_i \right] \\ &+ \frac{q-1}{2} \ln^2 \left[ c_{11} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \right], \end{aligned} \quad (4.3.19)$$

con  $\Lambda_i = T_i^{-1}$ .

Por otro lado, como de la descomposición (3.4.57) se obtiene

$$\frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m} = \begin{cases} \frac{p_{ii}}{1 - p_{ii}}, & \mathbf{h}_m = \mathbf{z}_i \\ 0, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}, \quad (4.3.20)$$

la medida de Rao,  $\rho_{i(\cdot m)}^2$ , se reduce a

$$\rho_{i(\cdot m)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \Lambda_i \right] + \frac{q-1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 [c_{11} \Lambda_i] + \frac{q-1}{2} \ln^2 [c_{11}] & , \mathbf{h}_m \neq \mathbf{z}_i \end{cases}. \quad (4.3.21)$$

**Nota 4.3.7** Se observa que las dos cantidades de  $\rho_{i(\cdot m)}^2$  dadas en (4.3.21) no dependen de  $m$  (columnas o grupos) solamente de  $i$  (observaciones). O sea, usando la medida de Rao para el estudio de la influencia en las columnas de  $\widehat{\mathbf{B}}$ , se concluye que la influencia de una observación en los grupos no correspondientes a dicha observación es siempre la misma, es decir, cada grupo de observaciones ejerce la misma influencia en cada uno de los restantes grupos. En este sentido, cuando  $r > 2$ , la simplificación (4.3.21) tiene ventajas computacionales en relación a (4.3.19).

#### 4.3.1.3. Influencia en los elementos de $\widehat{\mathbf{B}}$

Se considera ahora la influencia de la  $i$ -ésima observación  $\mathbf{y}_i$  de  $\mathbf{Y}$  en el  $(j, m)$ -ésimo elemento del EMV de la matriz de coeficientes de regresión  $\mathbf{B}$ , con  $1 \leq j \leq q$  y  $1 \leq m \leq r$ . En este caso,  $\mathbf{D} = \mathbf{d}_j^T$ ,  $\text{rg}(\mathbf{D}) = 1 \leq q$ , y  $\mathbf{L} = \mathbf{h}_m$ ,  $\text{rg}(\mathbf{L}) = 1 \leq r$ , y la expresión (4.3.11) se reduce a  $\widehat{\Psi} = \mathbf{d}_j^T (\mathbf{1} \times q) \widehat{\mathbf{B}}_{(q \times r)} \mathbf{h}_m (r \times 1) \in \mathbb{R}$ . Del corolario 4.3.2, se obtiene

$$\mathbf{R}_i = \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j^T} \in \mathbb{R}$$

y

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \in \mathbb{R},$$

donde la componente  $\mathbf{R}_i$  coincide con la de  $\rho_{i(j\cdot)}^2$  y  $\mathbf{L}_i$  con la de  $\rho_{i(\cdot m)}^2$ .

Como, de (4.3.12) y (4.3.13), se observa que

$$\widehat{\text{Var}}[\mathbf{d}_j^T \widehat{\mathbf{B}}_{(i)} \mathbf{h}_m] (\widehat{\text{Var}}[\mathbf{d}_j^T \widehat{\mathbf{B}} \mathbf{h}_m])^{-1} \in \mathbb{R},$$

del corolario 4.3.2, se obtiene la medida de Rao asociada a  $\mathbf{y}_i$  sobre el  $(j, m)$ -ésimo elemento de  $\widehat{\mathbf{B}}$  para el MCC

$$\begin{aligned} \rho_{i(jm)}^2(\widehat{\mathbf{B}}) &= \frac{1}{2} \ln^2 [c_{11} (1 + \mathbf{L}_i) (1 - \mathbf{R}_i)] = \\ &= \frac{1}{2} \ln^2 \left[ c_{11} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right]. \end{aligned} \quad (4.3.22)$$

Por otro lado, usando la descomposición (4.3.20) se obtiene la siguiente simplificación

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} \right) \right], & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}. \quad (4.3.23)$$

**Nota 4.3.8** Se observa que las dos cantidades en la medida  $\rho_{i(jm)}^2$  dada por (4.3.23) no dependen de  $m$  (columnas o grupos), solamente de  $i$  (observaciones) y  $j$  (filas). O sea, la influencia de una observación en los elementos de  $\widehat{\mathbf{B}}$  que pertenezcan a columnas no correspondientes al grupo de dicha observación es siempre la misma. Así, la expresión (4.3.23) tiene ventajas computacionales respecto a la expresión (4.3.22) para el cálculo de  $\rho_{i(jm)}^2$ , cuando  $r > 2$ .

**Nota 4.3.9** La medida de Rao,  $\rho_{i(jm)}^2$ , para los elementos de  $\widehat{\mathbf{B}}$  se puede expresar como función de las componentes de la medida de Fréchet,  $\delta_{i(jm)}$ . En relación a la componente de localización,  $LC_{i(jm)}^2$ , de (3.4.63),

$$\frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{v_{ii}} = \frac{v_{ii}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} LC_{i(jm)}^2,$$

entonces de (4.3.23), se tiene

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1 - p_{ii}} \right) \left( 1 - \frac{v_{ii} LC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i)^2} \right) \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{v_{ii} LC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j (\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} \right) \right], & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

Para la componente de dispersión,  $DC_{i(jm)}^2$ , efectivamente, de (3.4.68), se obtiene

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ \frac{n-r}{c_1} \left( \sqrt{\frac{c_1}{n-r}} + \sqrt{\frac{DC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m}} \right)^2 \right], & DC_{2i(jm)} > DC_{1i(jm)} \\ \frac{1}{2} \ln^2 \left[ \frac{n-r}{c_1} \left( \sqrt{\frac{c_1}{n-r}} - \sqrt{\frac{DC_{i(jm)}^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j \mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m}} \right)^2 \right], & DC_{2i(jm)} < DC_{1i(jm)} \end{cases},$$

donde  $DC_{1i(jm)}$  y  $DC_{2i(jm)}$ , definidas en (3.4.66) y (3.4.67), respectivamente, son dos de las componentes de  $DC_{i(jm)}^2$ . Por tanto, la medida  $\rho_{i(jm)}^2$  es una función creciente de  $DC_{i(jm)}^2$ .

**Nota 4.3.10** La medida de Rao,  $\rho_{i(jm)}^2$ , se puede expresar como función de  $DL_{i(jm)}$ , la medida de Cook para los elementos de  $\widehat{\mathbf{B}}$ . Como de (3.4.74), se puede considerar

$$\frac{1}{v_{ii}} \frac{(\mathbf{d}_j^T \mathbf{S}_X \mathbf{X}^T \mathbf{S}^{-1} \mathbf{e}_i)^2}{\mathbf{d}_j^T \mathbf{S}_X \mathbf{d}_j} = \begin{cases} \frac{v_{ii} p_{ii} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i)^2}, & \mathbf{h}_m = \mathbf{z}_i \\ \frac{v_{ii} p_{mm} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}, & \mathbf{h}_m \neq \mathbf{z}_i \end{cases},$$

de (4.3.23), se obtiene

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ c_{11} \left( \frac{1}{1-p_{ii}} \right) \left( 1 - \frac{v_{ii} p_{ii} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i)^2} \right) \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 \left[ c_{11} \left( 1 - \frac{v_{ii} p_{mm} DL_{i(jm)}}{(\mathbf{k}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2} \right) \right], & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

#### 4.4. La medida de Rao en el modelo de curvas de crecimiento con estructura simple de Rao

En este apartado se recoge la aplicación de la distancia de Rao como medida de influencia sobre el EMV  $\widetilde{\mathbf{B}}$  de la matriz de coeficientes de regresión del MCC(SCS), bajo la hipótesis de normalidad.

De las secciones 1.4 y 2.5, si las columnas de las matrices del error  $\mathcal{E}$  y  $\mathcal{E}_{(I)}$ , para el MCC(SCS) y MCC(SCS)( $I$ ), respectivamente, satisfacen  $\mathcal{E}_j \sim N_p(0, \boldsymbol{\Sigma})$ ,  $j = 1, \dots, n$ , con  $\boldsymbol{\Sigma}$  no singular,  $\text{vec}(\widetilde{\mathbf{B}})$  y  $\text{vec}(\widetilde{\mathbf{B}}_{(I)})$  se distribuyen según distribuciones normales  $qr$ -variantes que difieren solamente en sus varianzas. En efecto,

$$\text{vec}(\widetilde{\mathbf{B}}) \sim N_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \boldsymbol{\Gamma})$$

y

$$\widehat{\text{Var}}[\widetilde{\mathbf{B}}] = (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \widetilde{\boldsymbol{\Gamma}},$$

matriz simétrica de dimensión  $(qr \times qr)$ , con  $\widetilde{\boldsymbol{\Gamma}}$  dado por (1.4.4).

Análogamente,

$$\text{vec}(\widetilde{\mathbf{B}}_{(I)}) \sim N_{qr}(\text{vec}(\mathbf{B}), (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \boldsymbol{\Gamma})$$

y

$$\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] = (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)},$$

matriz simétrica de dimensión  $(qr \times qr)$ , con  $\tilde{\mathbf{\Gamma}}_{(I)}$  dado por (2.5.1).

Entonces los estimadores  $\tilde{\mathbf{B}}$  y  $\tilde{\mathbf{B}}_{(I)}$  siguen distribuciones normales  $qr$ -variantes con matrices de valores esperados iguales y matrices de covarianzas diferentes. Así, siguiendo el razonamiento de la sección anterior, como  $\widehat{\text{Var}}[\tilde{\mathbf{B}}]$  es no singular, teniendo en cuenta (4.2.4), la versión muestral del cuadrado de la distancia de Rao para evaluar la influencia que sobre la distribución de  $\tilde{\mathbf{B}}$  ejerce el conjunto de observaciones subindicadas por  $I$ , en el MCC(SCS), es

$$\rho_I^2(\tilde{\mathbf{B}}) = \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}] (\widehat{\text{Var}}[\tilde{\mathbf{B}}])^{-1} \right) \right], \quad (4.4.1)$$

donde  $\lambda_i(\mathbf{A})$ ,  $i = 1, \dots, qr$ , son los autovalores de la matriz cuadrada  $\mathbf{A}_{(qr \times qr)}$ .

El siguiente lema es útil para el desarrollo de (4.4.1) como medida de influencia en el EMV  $\tilde{\mathbf{B}}$ .

**Lema 4.4.1** Sean  $\tilde{\mathbf{\Gamma}}$  y  $\tilde{\mathbf{\Gamma}}_{(I)}$  los EMV de la componente de dispersión  $\mathbf{\Gamma}$  del MCC(SCS) y del MCC(SCS)( $I$ ), respectivamente, entonces

$$\tilde{\mathbf{\Gamma}}_{(I)}\tilde{\mathbf{\Gamma}}^{-1} = \frac{n}{n-k} (\mathbf{I}_q - (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{S}\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{X})).$$

**Demostración:** Usando la relación entre  $\tilde{\mathbf{\Gamma}}$  y  $\tilde{\mathbf{\Gamma}}_{(I)}$ , dada en (2.5.5), se obtiene

$$\begin{aligned} \tilde{\mathbf{\Gamma}}_{(I)}\tilde{\mathbf{\Gamma}}^{-1} &= \left( \frac{n}{n-k}\tilde{\mathbf{\Gamma}} - \frac{1}{n-k}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} \right) \tilde{\mathbf{\Gamma}}^{-1} = \\ &= \frac{n}{n-k}\tilde{\mathbf{\Gamma}}\tilde{\mathbf{\Gamma}}^{-1} - \frac{1}{n-k}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\tilde{\mathbf{\Gamma}}^{-1} = \\ &= \frac{n}{n-k}\mathbf{I}_q - \frac{1}{n-k}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}n(\mathbf{X}^T\mathbf{X}) \\ &\quad (\mathbf{X}^T\mathbf{S}\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{X}) = \\ &= \frac{n}{n-k} (\mathbf{I}_q - (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{S}\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{X})), \end{aligned}$$

con  $\mathbf{E}_I$  y  $\mathbf{H}_I$  dadas en (2.3.6) y (2.3.7), respectivamente. ■

**Teorema 4.4.1** Para el MCC(SCS), la medida de Rao asociada al subconjunto  $\mathbf{Y}_I$  sobre el EMV  $\tilde{\mathbf{B}}$  es

$$\rho_I^2(\tilde{\mathbf{B}}) = \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 \left[ \frac{n}{n-k} (1 + \lambda_j(\mathbf{L}_I) (1 - \mu_h(\mathbf{R}_I))) \right],$$

con

$$\mathbf{L}_I = (\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T$$

y

$$\mathbf{R}_I = \mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{S}\mathbf{X})^{-1},$$

donde  $\lambda_j(\mathbf{L}_I)$ ,  $j = 1, \dots, r$ , y  $\mu_h(\mathbf{R}_I)$ ,  $h = 1, \dots, q$ , son los autovalores de las matrices cuadradas  $\mathbf{L}_I$  y  $\mathbf{R}_I$ , respectivamente.

**Demostración:** Teniendo en cuenta las propiedades (vi) y (vii) del producto de Kronecker, dadas por el teorema B.0.5, y los lemas 4.3.1 y 4.4.1 se tiene

$$\begin{aligned}
\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}](\widehat{\text{Var}}[\tilde{\mathbf{B}}])^{-1} &= \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) \left( (\mathbf{Z}\mathbf{Z}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}} \right)^{-1} = \\
&= \left( (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} \otimes \tilde{\mathbf{\Gamma}}_{(I)} \right) \left( (\mathbf{Z}\mathbf{Z}^T) \otimes \tilde{\mathbf{\Gamma}}^{-1} \right) = \\
&= (\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1} (\mathbf{Z}\mathbf{Z}^T) \otimes \tilde{\mathbf{\Gamma}}_{(I)} \tilde{\mathbf{\Gamma}}^{-1} = \\
&= \frac{n}{n-k} (\mathbf{I}_r + (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T) \otimes \\
&\quad \otimes (\mathbf{I}_q - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X})),
\end{aligned}$$

matriz simétrica de dimensión  $(qr \times qr)$ . Así, de (4.4.1), se obtiene

$$\begin{aligned}
\rho_I^2 &= \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ \lambda_i (\widehat{\text{Var}}[\tilde{\mathbf{B}}_{(I)}](\widehat{\text{Var}}[\tilde{\mathbf{B}}])^{-1}) \right] = \\
&= \frac{1}{2} \sum_{i=1}^{qr} \ln^2 \left[ \frac{n}{n-k} \lambda_i \left( (\mathbf{I}_r + (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T) \otimes \right. \right. \\
&\quad \left. \left. \otimes (\mathbf{I}_q - (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X})) \right) \right] = \\
&= \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 \left[ \frac{n}{n-k} (1 + \lambda_j ((\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T)) \right. \\
&\quad \left. (1 - \mu_h ((\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X}))) \right] = \\
&= \frac{1}{2} \sum_{h=1}^q \sum_{j=1}^r \ln^2 \left[ \frac{n}{n-k} (1 + \lambda_j ((\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{Z}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{Z}_I^T)) \right. \\
&\quad \left. (1 - \mu_h (\mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1} \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1})) \right].
\end{aligned}$$

Usando la propiedad (xii) del teorema B.0.5, teniendo en cuenta que el producto de Kronecker es de orden  $qr$  y el lema B.0.3 se concluye la demostración. ■

También para el MCC(SCS), se puede interpretar la medida de Rao,  $\rho_I^2$ , como función de las componentes  $\mathbf{L}_I$  y  $\mathbf{R}_I$ . Como se puede observar, la componente  $\mathbf{L}_I$  es idéntica a la obtenida en el MCC. Por otro lado, del lema B.0.2, se concluye que la componente  $\mathbf{R}_I$  tiene los mismos autovalores no nulos, con la misma multiplicidad, que

$$\mathbf{R}_I^* = \mathbf{E}_I^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E}_I (\mathbf{I}_k - \mathbf{H}_I)^{-1}. \quad (4.4.2)$$

De la expresión (2.5.6) se obtiene que  $\Lambda_I = \det[\mathbf{I}_k - \mathbf{R}_I^*]$ , por lo que la matriz  $\mathbf{R}_I^*$ , basada en los residuos, puede ser usada en la detección de outliers.

**Nota 4.4.1** Por razonamientos análogos a los usados para el MCC en la sección 4.3, también para el MCC(SCS) se puede descomponer la medida de Rao de la siguiente manera

$$2\rho_I^2 = \|\mathcal{L}_I\|^2 + \|\mathcal{R}_I\|^2 + 2\|\mathcal{L}_I\| \|\mathcal{R}_I\| \cos \theta_I,$$

donde

$$\mathcal{L}_I = (\mathbf{1}_q \otimes \mathbf{a}), \text{ con } \mathbf{a}^T = (\ln[1 + \lambda_1(\mathbf{L}_I)], \dots, \ln[1 + \lambda_r(\mathbf{L}_I)]),$$

$$\mathcal{R}_I = (\mathbf{b} \otimes \mathbf{1}_r), \text{ con } \mathbf{b}^T = \left( \ln \left[ \frac{n}{n-k} (1 - \mu_1(\mathbf{R}_I)) \right], \dots, \ln \left[ \frac{n}{n-k} (1 - \mu_q(\mathbf{R}_I)) \right] \right),$$

$\|\cdot\|$  es la norma euclídea y  $\theta_I$  es el ángulo entre los vectores  $\mathcal{L}_I$  y  $\mathcal{R}_I$ , siendo  $\mathbf{1}_s = (1, \dots, 1) \in \mathbb{R}^s$ .

Usando el teorema 4.4.1, se obtiene la distancia de Rao como medida de la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  sobre la distribución del EMV  $\tilde{\mathbf{B}}$ .

**Corolario 4.4.1** Para el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida de Rao asociada al EMV  $\tilde{\mathbf{B}}$  es

$$\rho_i^2(\tilde{\mathbf{B}}) = \frac{1}{2} \sum_{j=1}^r \sum_{h=1}^q \ln^2 \left[ \frac{n}{n-1} (1 + \lambda_j(\mathbf{L}_i)) (1 - \mu_h(\mathbf{R}_i)) \right],$$

con

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T$$

y

$$\mathbf{R}_i = \frac{1}{1 - p_{ii}} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1},$$

donde  $p_{ii}$  y  $\mathbf{e}_i$  están definidos por (2.3.15) y (2.3.14), respectivamente.

Para el desarrollo de la medida  $\rho_i^2(\tilde{\mathbf{B}})$ , se observa que  $\mathbf{L}_i$  es la misma que en el MCC, así de (4.3.9), se tiene que

$$\lambda_j(\mathbf{L}_i) = \begin{cases} \frac{p_{ii}}{1-p_{ii}}, & j = 1, \quad i = 1, \dots, n \\ 0, & j = 2, \dots, r, \quad i = 1, \dots, n \end{cases}. \quad (4.4.3)$$

Por otro lado, teniendo en cuenta la relación entre los autovalores de las matrices  $\mathbf{R}_I$  y  $\mathbf{R}_I^*$  dada en (4.4.2), la matriz  $\mathbf{R}_i$  tiene solamente el autovalor no nulo  $\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i$ , o sea,

$$\mu_h(\mathbf{R}_i) = \begin{cases} \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i}{1-p_{ii}}, & h = 1, \quad i = 1, \dots, n \\ 0, & h = 2, \dots, q, \quad i = 1, \dots, n \end{cases},$$

y como, en el MCC(SCS), un estadístico para determinar si la observación  $i$ -ésima de  $\mathbf{Y}$  es outlier (véase la sección 2.5), es

$$T_i = \frac{\mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i}{1 - p_{ii}},$$

el residuo internamente studentizado para el MCC(SCS), se concluye que

$$\mu_h(\mathbf{R}_i) = \begin{cases} T_i, & h = 1, \quad i = 1, \dots, n \\ 0, & h = 2, \dots, q, \quad i = 1, \dots, n \end{cases}. \quad (4.4.4)$$

Así, usando los autovalores (4.4.3) y (4.4.4) en el corolario 4.4.1, la medida de Rao, asociada a la observación  $\mathbf{y}_i$  sobre el EMV  $\tilde{\mathbf{B}}$  del MCC(SCS), se reduce a

$$\rho_i^2(\tilde{\mathbf{B}}) = \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( \frac{1}{1-p_{ii}} \right) \Lambda_i \right] + \frac{r-1}{2} \ln^2 \left[ \frac{n}{n-1} \Lambda_i \right] + \\ + \frac{q-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( \frac{1}{1-p_{ii}} \right) \right] + \frac{(q-1)(r-1)}{2} \ln^2 \left[ \frac{n}{n-1} \right],$$

donde,  $\Lambda_i = 1 - T_i$ . En consecuencia, la medida de Rao,  $\rho_i^2$ , es una función creciente de  $T_i$ .

**Nota 4.4.2** En el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , teniendo en cuenta los autovalores (4.4.3) y (4.4.4), la medida de Rao puede decomponerse de la siguiente manera

$$2\rho_i^2 = \|\mathcal{L}_i\|^2 + \|\mathcal{R}_i\|^2 + 2\|\mathcal{L}_i\| \|\mathcal{R}_i\| \cos \theta_i,$$

con

$$\mathcal{L}_i = (\mathbf{1}_q \otimes \mathbf{a}), \text{ donde } \mathbf{a}^T = \left( \ln \left[ \frac{1}{1-p_{ii}} \right], 0, \dots, 0 \right)$$

y

$$\mathcal{R}_i = (\mathbf{b} \otimes \mathbf{1}_r), \text{ donde } \mathbf{b}^T = \left( \ln \left[ \frac{n}{n-1} \Lambda_i \right], \ln \left[ \frac{n}{n-1} \right], \dots, \ln \left[ \frac{n}{n-1} \right] \right),$$

donde  $\Lambda_i = 1 - T_i$  y  $\mathbf{1}_s = (1, \dots, 1) \in \mathbb{R}^s$ . En este sentido, se obtiene

$$\|\mathcal{L}_i\|^2 = q \sum_{j=1}^r \ln^2 [1 + \lambda_j(\mathbf{L}_i)] = q \ln^2 \left[ \frac{1}{1-p_{ii}} \right],$$

$$\|\mathcal{R}_i\|^2 = r \sum_{h=1}^q \ln^2 \left[ \frac{n}{n-1} (1 - \mu_h(\mathbf{R}_i)) \right] = r \ln^2 \left[ \frac{n}{n-1} \Lambda_i \right] + (q-1)r \ln^2 \left[ \frac{n}{n-1} \right]$$

y

$$\|\mathcal{R}_i\| \|\mathcal{L}_i\| \cos \theta = \mathcal{L}_i^T \cdot \mathcal{R}_i = \sum_{j=1}^r \sum_{h=1}^q \ln [1 + \lambda_j(\mathbf{L}_i)] \ln \left[ \frac{n}{n-1} (1 - \mu_h(\mathbf{R}_i)) \right] = \\ = \ln \left[ \frac{1}{1-p_{ii}} \right] \left( \ln [\Lambda_i] + q \ln \left[ \frac{n}{n-1} \right] \right).$$

**Nota 4.4.3** En el MCC(SCS), es posible expresar  $\rho_i^2(\tilde{\mathbf{B}})$  como función de las medidas presentadas en la sección 2.5. Por ejemplo, de la medida de Cook,  $D_i$ , para la  $i$ -ésima observación de  $\mathbf{Y}$ , dada en (2.5.11), se obtiene

$$\Lambda_i = 1 - \frac{1-p_{ii}}{np_{ii}} D_i,$$

entonces

$$\rho_i^2(\tilde{\mathbf{B}}) = \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{1-p_{ii}}{np_{ii}} D_i \right) \right] + \frac{q-1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1-p_{ii}} \right] + \\ + \frac{r-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{1-p_{ii}}{np_{ii}} D_i \right) \right] + \frac{(q-1)(r-1)}{2} \ln^2 \left[ \frac{n}{n-1} \right].$$

Es decir, la medida de Rao puede ser considerada como una función creciente de  $p_{ii}$  (constante en cada grupo) y  $D_i$ .



#### 4.4.1. Influencia en combinaciones lineales de $\tilde{\mathbf{B}}$

La extensión de distancia de Rao como medida de influencia en combinaciones lineales del EMV de la matriz de coeficientes de regresión del MCC(SCS) es similar al realizado para el MCC. En este caso, se consideran los EMV  $\tilde{\Psi} = \mathbf{D}_{(s \times q)} \tilde{\mathbf{B}}_{(q \times r)} \mathbf{L}_{(r \times l)}$  y  $\tilde{\Psi}_{(I)} = \mathbf{D}_{(s \times q)} \tilde{\mathbf{B}}_{(I)(q \times r)} \mathbf{L}_{(r \times l)}$ , donde  $\text{rg}(\mathbf{D}) = s \leq q$  y  $\text{rg}(\mathbf{L}) = l \leq r$ .

Teniendo en cuenta las distribuciones de  $\text{vec}(\tilde{\mathbf{B}})$  y  $\text{vec}(\tilde{\mathbf{B}}_{(I)})$ , usando el teorema C.1.2, se obtiene

$$\text{vec}(\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}) \sim N_{sl}(\text{vec}(\mathbf{D}\mathbf{B}\mathbf{L}); \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\mathbf{\Gamma}\mathbf{D}^T)$$

y

$$\text{vec}(\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}) \sim N_{sl}\left(\text{vec}(\mathbf{D}\mathbf{B}\mathbf{L}), \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\mathbf{\Gamma}\mathbf{D}^T\right),$$

resultando

$$\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}] = \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T \quad (4.4.5)$$

y

$$\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}] = \mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\mathbf{\Gamma}}_{(I)}\mathbf{D}^T, \quad (4.4.6)$$

matrices simétricas de dimensión  $(sl \times sl)$ .

Entonces, de (4.2.4), como  $\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}]$  es no singular, la versión muestral del cuadrado de la distancia de Rao, como medida de influencia del subconjunto de observaciones  $\mathbf{Y}_I$  en las combinaciones lineales de  $\tilde{\mathbf{B}}$ , es

$$\rho_I^2(\tilde{\Psi}) = \frac{1}{2} \sum_{i=1}^{sl} \ln^2 \left[ \lambda_i \left( \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}] \left( \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}] \right)^{-1} \right) \right], \quad (4.4.7)$$

donde  $\lambda_i(\mathbf{A})$ ,  $i = 1, \dots, sl$ , representan los autovalores de la matriz cuadrada  $\mathbf{A}_{(sl \times sl)}$ .

A partir de la expresión general (4.4.7) se enuncia el siguiente teorema.

**Teorema 4.4.2** *Para el MCC(SCS), la medida de Rao asociada a  $\mathbf{Y}_I$  sobre  $\tilde{\Psi}$  es*

$$\rho_I^2(\tilde{\Psi}) = \frac{1}{2} \sum_{a=1}^l \sum_{b=1}^s \ln^2 \left[ \frac{n}{n-k} (1 + \lambda_a(\mathbf{L}_I)) (1 - \mu_b(\mathbf{R}_I)) \right],$$

con

$$\mathbf{L}_I = \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1}$$

y

$$\mathbf{R}_I = \frac{1}{n} \mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T(\mathbf{D}\tilde{\mathbf{\Gamma}}\mathbf{D}^T)^{-1},$$

donde  $\lambda_a(\mathbf{L}_I)$ ,  $a = 1, \dots, l$ , y  $\mu_b(\mathbf{R}_I)$ ,  $b = 1, \dots, s$ , son los autovalores de las matrices cuadradas  $\mathbf{L}_I$  y  $\mathbf{R}_I$ , respectivamente.

**Demostración:** Del lema 3.4.2, se obtiene

$$\begin{aligned} (\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L})(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1} &= \mathbf{I}_l + \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1} \\ &\quad \mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1}, \end{aligned}$$

y, de 2.5.5,

$$\begin{aligned} (\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T)(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1} &= \frac{n}{n-k}(\mathbf{I}_s - \frac{1}{n}\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1} \\ &\quad \mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1}). \end{aligned}$$

Usando los estimadores de las matrices de covarianza, (4.4.5) y (4.4.6), y el teorema B.0.5, se obtiene que

$$\begin{aligned} \widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}_{(I)}\mathbf{L}](\widehat{\text{Var}}[\mathbf{D}\tilde{\mathbf{B}}\mathbf{L}])^{-1} &= \\ &= (\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T)(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1} = \\ &= (\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L} \otimes \mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T)((\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1} \otimes (\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1}) = \\ &= (\mathbf{L}^T(\mathbf{Z}_{(I)}\mathbf{Z}_{(I)}^T)^{-1}\mathbf{L})(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1} \otimes (\mathbf{D}\tilde{\Gamma}_{(I)}\mathbf{D}^T)(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1} = \\ &= \frac{n}{n-k}(\mathbf{I}_l + \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{Z}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{Z}_I^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L}(\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1}\mathbf{L})^{-1}) \otimes \\ &\quad \otimes (\mathbf{I}_s - \frac{1}{n}\mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}_I(\mathbf{I}_k - \mathbf{H}_I)^{-1}\mathbf{E}_I^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{D}^T(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1}), \end{aligned}$$

matriz simétrica de dimensión  $(sl \times sl)$ , de (4.4.7), teniendo en cuenta que el producto de Kronecker es de orden  $sl$ , se concluye la demostración. ■

El corolario siguiente permite considerar la distancia de Rao como medida de influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  en un determinado subconjunto del EMV  $\tilde{\mathbf{B}}$ .

**Corolario 4.4.2** Para el MCC(SCS), cuando  $k = 1$  e  $I = \{i\}$ ,  $1 \leq i \leq n$ , la medida de Rao asociada a  $\tilde{\Psi}$  es

$$\rho_i^2(\tilde{\Psi}) = \frac{1}{2} \sum_{a=1}^l \sum_{b=1}^s \ln^2 \left[ \frac{n}{n-1} (1 + \lambda_a(\mathbf{L}_i)) (1 - \mu_b(\mathbf{R}_i)) \right],$$

con

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} \mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L} (\mathbf{L}^T(\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{L})^{-1}$$

y

$$\mathbf{R}_i = \frac{1}{n(1 - p_{ii})} \mathbf{D}(\mathbf{X}^T\mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} \mathbf{D}^T(\mathbf{D}\tilde{\Gamma}\mathbf{D}^T)^{-1}.$$

A continuación, se presentan las medidas de Rao resultantes de tres posibles elecciones de las matrices  $\mathbf{D}$  y  $\mathbf{L}$ .

4.4.1.1. Influencia en las filas de  $\tilde{\mathbf{B}}$

Para estudiar la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  en la  $j$ -ésima fila de  $\tilde{\mathbf{B}}$ , en el MCC(SCS), se tiene  $\tilde{\Psi} = \mathbf{d}_j^T \tilde{\mathbf{B}}$ . Entonces, del corolario 4.4.2, la medida de Rao en este caso es

$$\rho_{i(j)}^2(\tilde{\mathbf{B}}) = \frac{1}{2} \sum_{a=1}^r \ln^2 \left[ \frac{n}{n-1} (1 - \mathbf{R}_i) (1 + \lambda_a(\mathbf{L}_i)) \right], \quad (4.4.8)$$

con

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{z}_i \mathbf{z}_i^T$$

y

$$\mathbf{R}_i = \frac{1}{n(1 - p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \in \mathbb{R}.$$

Por otro lado, teniendo en cuenta que la componente  $\mathbf{L}_i$  es la misma que en la medida de Rao  $\rho_i^2(\tilde{\mathbf{B}})$ , usando (4.4.3), se obtiene

$$\begin{aligned} \rho_{i(j)}^2 &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1 - p_{ii}} \left( 1 - \frac{1}{n(1 - p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right) \right] + \\ &+ \frac{r-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{1}{n(1 - p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right) \right]. \end{aligned}$$

**Nota 4.4.4** En el MCC(SCS), la medida  $\rho_{i(j)}^2$  puede expresarse como función de  $LC_{i(j)}^2$ , la componente de localización de la medida de Fréchet para las filas de  $\tilde{\mathbf{B}}$ . En efecto, de (3.5.13),

$$\frac{1}{1 - p_{ii}} (\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2 = \frac{1 - p_{ii}}{p_{ii}^2} LC_{i(j)}^2,$$

por tanto

$$\begin{aligned} \rho_{i(j)}^2 &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1 - p_{ii}} \left( 1 - \frac{1}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \frac{1 - p_{ii}}{np_{ii}^2} LC_{i(j)}^2 \right) \right] + \\ &+ \frac{r-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{1}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \frac{1 - p_{ii}}{np_{ii}^2} LC_{i(j)}^2 \right) \right]. \end{aligned}$$

**Nota 4.4.5** En el MCC(SCS), se puede expresar la medida de Rao  $\rho_{i(j)}^2$  como función de  $DL_{i(j)}$ , la medida de Cook para las filas de  $\tilde{\mathbf{B}}$ . Dado que, por (2.5.15), se tiene

$$\frac{1}{1 - p_{ii}} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} = \frac{1 - p_{ii}}{p_{ii}} DL_{i(j)} = \frac{nT_i}{D_i} DL_{i(j)}, \quad (4.4.9)$$

en consecuencia

$$\rho_{i(j)}^2 = \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1 - p_{ii}} \left( 1 - \frac{T_i}{D_i} DL_{i(j)} \right) \right] + \frac{r-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{T_i}{D_i} DL_{i(j)} \right) \right],$$

siendo  $D_i$  la medida de Cook dada en (2.5.11) y  $T_i$  dado en (2.5.8).

#### 4.4.1.2. Influencia en las columnas de $\tilde{\mathbf{B}}$

Para evaluar la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  en la  $m$ -ésima columna de  $\tilde{\mathbf{B}}$ , en el MCC(SCS), basta considerar  $\tilde{\Psi} = \tilde{\mathbf{B}}\mathbf{h}_m$ . Entonces, del corolario 4.4.2, la medida de Rao en este caso es

$$\rho_{i(\cdot m)}^2(\tilde{\mathbf{B}}) = \frac{1}{2} \sum_{b=1}^q \ln^2 \left[ \frac{n}{n-1} (1 + \mathbf{L}_i) (1 - \mu_b(\mathbf{R}_i)) \right], \quad (4.4.10)$$

con

$$\mathbf{L}_i = \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \in \mathbb{R}$$

y

$$\mathbf{R}_i = \frac{1}{1 - p_{ii}} \mathbf{X}^T \mathbf{e}_i \mathbf{e}_i^T \mathbf{X} (\mathbf{X}^T \mathbf{S} \mathbf{X})^{-1}.$$

Teniendo en cuenta que la componente  $\mathbf{R}_i$  es la misma que en la medida de Rao  $\rho_i^2(\tilde{\mathbf{B}})$ , usando (4.4.4), se obtiene que

$$\begin{aligned} \rho_{i(\cdot m)}^2 &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \Lambda_i \right] + \\ &+ \frac{q-1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \right]. \end{aligned}$$

Por otro lado, considerando la descomposición (4.3.20),

$$\rho_{i(\cdot m)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1-p_{ii}} \Lambda_i \right] + \frac{q-1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1-p_{ii}} \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \Lambda_i \right] + \frac{q-1}{2} \ln^2 \left[ \frac{n}{n-1} \right], & \mathbf{h}_m \neq \mathbf{z}_i \end{cases}.$$

En resumen, las cantidades de la medida de Rao  $\rho_{i(\cdot m)}^2$  no dependen de  $m$ . En este sentido, se tienen las mismas ventajas computacionales que las mencionadas en la nota 4.3.7, para el MCC.

#### 4.4.1.3. Influencia en los elementos de $\tilde{\mathbf{B}}$

En relación al estudio de la influencia de la  $i$ -ésima observación de  $\mathbf{Y}$  en el  $(j, m)$ -ésimo elemento de  $\tilde{\mathbf{B}}$ , para el MCC(SCS), se considera  $\tilde{\Psi} = \mathbf{d}_j^T \tilde{\mathbf{B}} \mathbf{h}_m \in \mathbb{R}$ . Teniendo en cuenta la medida de Rao asociada a las filas y columnas de  $\tilde{\mathbf{B}}$  dadas por (4.4.8) y (4.4.10), respectivamente, del corolario 4.4.2, se obtiene

$$\begin{aligned} \rho_{i(jm)}^2 &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} (1 + \mathbf{L}_i) (1 - \mathbf{R}_i) \right] = \\ &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 + \frac{1}{1 - p_{ii}} \frac{(\mathbf{z}_i^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m)^2}{\mathbf{h}_m^T (\mathbf{Z}\mathbf{Z}^T)^{-1} \mathbf{h}_m} \right) \right. \\ &\quad \left. \left( 1 - \frac{1}{n(1 - p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right) \right] \end{aligned} \quad (4.4.11)$$

la medida de Rao asociada a la observación  $\mathbf{y}_i$  sobre el  $(j, m)$ -ésimo elemento de  $\tilde{\mathbf{B}}$ .

Por otro lado, considerando la descomposición (4.3.20),

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \frac{1}{1-p_{ii}} \left( 1 - \frac{1}{n(1-p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right) \right], & \mathbf{h}_m = \mathbf{z}_i \\ \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{1}{n(1-p_{ii})} \frac{(\mathbf{d}_j^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{e}_i)^2}{\mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j} \right) \right] & , \mathbf{h}_m \neq \mathbf{z}_i \end{cases} .$$

Se observa, como en el MCC, que las cantidades de la medida de Rao  $\rho_{i(jm)}^2$  no dependen de  $m$ , es decir, se tienen las mismas ventajas computacionales que las mencionadas en la nota 4.3.8.

**Nota 4.4.6** En el MCC(SCS), se puede expresar  $\rho_{i(jm)}^2$  como función de  $DC_{i(jm)}^2$ , la componente de dispersión de la medida de Fréchet  $\delta_{i(jm)}$ . En efecto, de (3.5.17), se obtiene

$$\rho_{i(jm)}^2 = \begin{cases} \frac{1}{2} \ln^2 \left[ \left( 1 + \sqrt{\frac{DC_{i(jm)}^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j}} \right)^2 \right], & DC_{2i(jm)} > DC_{1i(jm)} \\ \frac{1}{2} \ln^2 \left[ \left( 1 - \sqrt{\frac{DC_{i(jm)}^2}{\mathbf{h}_m^T (\mathbf{Z} \mathbf{Z}^T)^{-1} \mathbf{h}_m \mathbf{d}_j^T \tilde{\Gamma} \mathbf{d}_j}} \right)^2 \right], & DC_{2i(jm)} < DC_{1i(jm)} \end{cases} .$$

**Nota 4.4.7** También es posible establecer una relación entre la medida de Rao  $\rho_{i(jm)}^2$  y las medidas de Cook. De (4.4.9) y (2.5.16), usando (4.4.11), se obtiene

$$\begin{aligned} \rho_{i(jm)}^2 &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 + \frac{p_{ii}}{(1-p_{ii})D_i} DL_{i(\cdot m)} \right) \left( 1 - \frac{1-p_{ii}}{np_{ii}} DL_{i(j\cdot)} \right) \right] = \\ &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 + \frac{1}{nT_i} DL_{i(\cdot m)} \right) \left( 1 - \frac{T_i}{D_i} DL_{i(j\cdot)} \right) \right] = \\ &= \frac{1}{2} \ln^2 \left[ \frac{n}{n-1} \left( 1 - \frac{T_i}{D_i} DL_{i(j\cdot)} + \frac{1}{nT_i} DL_{i(\cdot m)} - \frac{1}{n} DL_{i(jm)} \right) \right], \end{aligned}$$

donde las medidas de Cook  $D_i$ ,  $DL_{i(j\cdot)}$ ,  $DL_{i(\cdot m)}$  y  $DL_{i(jm)}$ , para el MCC(SCS), están definidas en (2.5.11), (2.5.15), (2.5.16) y (2.5.18), respectivamente.



## Capítulo 5

# Implementación y casos prácticos

### 5.1. Introducción

Para ilustrar esta memoria, en este capítulo se aplican, a algunos conjuntos de datos reales, las medidas de influencia presentadas en el capítulo 2 y los planteamientos metodológicos de detección de observaciones influyentes, bajo la utilización de distancias entre distribuciones de probabilidad, desarrolladas en los capítulos 3 y 4. El análisis de influencia, llevado a cabo en los ejemplos, se realiza desde el enfoque de la influencia individual.

El primer conjunto corresponde a los datos recogidos por Zerbe (1979) en un trabajo sobre la asociación entre la hiperglucemia y la hiperinsulinemia, a través de diversas medidas repetidas de fosfato inorgánico plasmático. Este conjunto de datos ya ha sido utilizado en la bibliografía para el MCC, en particular para el análisis de influencia en este campo. En Pan y Fang (2002) es utilizado en el ámbito del análisis de influencia en el EMV de la matriz de coeficientes de regresión. Esto permite la comparación del análisis de influencia realizado mediante la distancia entre distribuciones con las técnicas utilizadas por otros autores.

En el segundo ejemplo, se estudia la identificación de observaciones influyentes en el EMV de la matriz de coeficientes de regresión del MCC(SCS). Para ello, se utiliza un conjunto de datos incluido en un artículo clásico del MANOVA generalizado de Potthoff y Roy (1964). Los datos corresponden a una medida dental (la distancia, en milímetros, del centro de la glándula pituitaria a la fisura pterigomaxilar) a un conjunto de niños y niñas a distintas edades. Estos datos han sido utilizados en el estudio de técnicas de detección de observaciones influyentes en diversos trabajos, pudiéndose citar, entre otros, el trabajo de Pan (2002).

Con estos ejemplos sólo se pretende ilustrar el uso y la aplicación de las técnicas de análisis de influencia propuestas a través de la aplicación de las distancias de Fréchet y de Rao, por lo que no se pretende realizar un análisis detallado de los datos y de los modelos aplicados.

En general, en el análisis de influencia, algunos autores proponen para las medidas de influencia puntos de corte a partir de los cuales se deben considerar las observaciones influyentes. Sin embargo, otros autores proponen medir la influencia de todas las observaciones y considerar como influyentes aquellas que obtengan valores considerablemente mayores en relación con el resto. Ello puede realizarse a través de una representación gráfica de los valores del diagnóstico considerado frente al índice de casos (*index plots*). Ambas estrategias están cargadas con cierto

nivel de subjetividad. No obstante, posiblemente, la segunda esté más en la línea de la propia definición de observación influencia, por lo que en esta memoria se ha optado por ella.

El *software* básico utilizado para desarrollar e implementar las medidas de influencia propuestas en esta memoria ha sido MATLAB, aunque fácilmente pueden ser trasladadas para su implementación en R.

## 5.2. Datos de Glucosa

Zerbe (1979) analiza un conjunto de datos obtenidos en un estudio sobre la asociación entre la hiperglucemia y la hiperinsulinemia, a través de diversas medidas repetidas de fosfato inorgánico plasmático. El estudio se realizó en el Área de Investigación Clínica Pediátrica del Medical Center de la Universidad de Colorado. En la experiencia, se administran diversos tests de tolerancia a la glucosa a dos grupos ( $r = 2$ ) de pacientes: grupo control, con  $n_1 = 13$  individuos, y grupo de pacientes afectados de obesidad, con  $n_2 = 20$  individuos. Las observaciones son mediciones de fosfato inorgánico plasmático (mg/dl) obtenidas a partir de muestras de sangre, para cada uno de los  $n = 33$  pacientes, tomadas en  $p = 8$  instantes sucesivos de tiempo (0 h., 0,5 h., 1 h., 1,5 h., 2 h., 3 h., 4 h. y 5 h.) después de ser administrada la dosis estándar de glucosa oral.

Este conjunto de datos ha sido posteriormente analizado por Reinsel (1984b), Chi y Reinsel (1989), Keramidas y Lee (1995) y, en el ámbito del análisis de influencia, por Pan y Fang (2002).

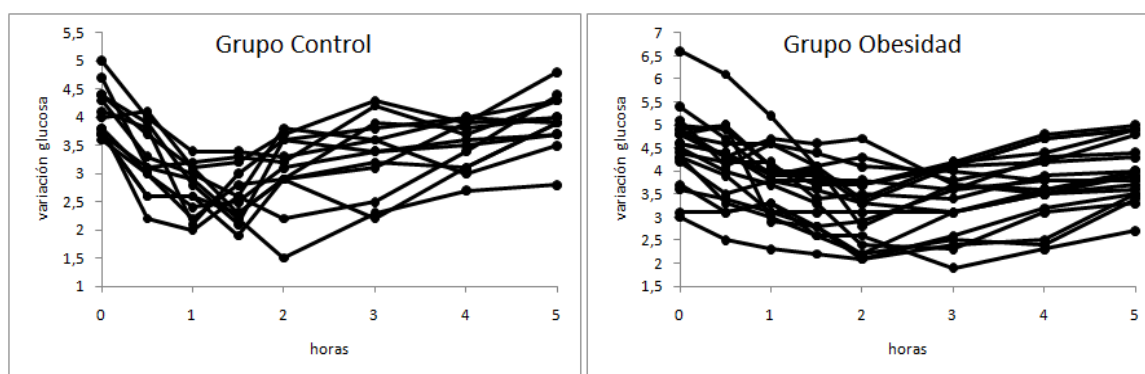


Figura 5.2.1: Curvas de crecimiento de los niveles de fosfato inorgánico plasmático (mg/dl) en individuos del Grupo Control y del Grupo Obesidad

Para la aplicación de las medidas de influencia es necesario el ajuste de los datos a través del MCC, en este sentido la figura 5.2.1 ilustra las curvas de crecimiento de los niveles de fosfato inorgánico plasmático en los dos grupos. A partir de ella, parece razonable considerar un polinomio de grado 2 o superior para modelizar el crecimiento medio de la variable objetivo. Pan y Fang (2002) consideran el ajuste de un MCC con matriz de covarianzas no estructurada, a través de un polinomio de grado 2, obteniéndose como coeficientes de determinación  $R^2 = 0,6564$ , para el grupo control, y  $R^2 = 0,9423$ , para el grupo obesidad. Si se utiliza un



polinomio de grado 3, los coeficientes de determinación son  $R^2 = 0,903$  y  $R^2 = 0,971$ , respectivamente. Por ello, en esta memoria se opta por un modelo basado en el polinomio de grado 3.

En consecuencia, el modelo de curvas de crecimiento considerado en la modelización de esta experiencia tiene las siguientes dimensiones: dos grupos ( $r = 2$ ), el grupo control ( $n_1 = 13$  pacientes) y el grupo obesidad ( $n_2 = 20$  pacientes); ocho medidas repetidas o instantes de tiempo considerados ( $p = 8$ ); polinomio de tercer grado ( $q = 4$ ). Así, las matrices de diseño son

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0,5 & 1 & 1,5 & 2 & 3 & 4 & 5 \\ 0 & 0,25 & 1 & 2,25 & 4 & 9 & 16 & 25 \\ 0 & 0,125 & 1 & 3,375 & 8 & 27 & 64 & 125 \end{bmatrix}^T \quad y \quad \mathbf{Z} = \begin{bmatrix} \mathbf{1}_{13}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{20}^T \end{bmatrix}.$$

Al considerar un MCC con matriz de covarianzas no estructurada, la única restricción sobre la matriz de varianzas y covarianzas es que  $\boldsymbol{\Sigma} > \mathbf{0}$ . En este sentido se obtienen los EMV ( $\hat{\mathbf{B}}$ ) y GLSE ( $\tilde{\mathbf{B}}$ ) para la matriz de coeficientes de regresión  $\mathbf{B}$  del MCC,

$$\hat{\mathbf{B}} = \begin{bmatrix} 4,0106 & 4,5038 \\ -1,7660 & -1,0475 \\ 0,6491 & 0,2920 \\ -0,0608 & -0,0204 \end{bmatrix} \quad y \quad \tilde{\mathbf{B}} = \begin{bmatrix} 4,0040 & 4,5924 \\ -1,7702 & -1,1939 \\ 0,7370 & 0,3278 \\ -0,0777 & -0,0222 \end{bmatrix}.$$

La figura 5.2.2 ilustra el ajuste de las curvas obtenidas de los EMV y GLSE, respectivamente. Para el grupo control se concluye que existe una gran diferencia entre las dos curvas, mientras que en el grupo obesidad dicha diferencia es menos significativa. Considerando los EMV, los polinómios ajustados para el modelo son:  $y = 4,0106 - 1,7660t + 0,6491t^2 - 0,0608t^3$  (grupo control) e  $y = 4,5038 - 1,0475t + 0,2920t^2 - 0,0204t^3$  (grupo obesidad), donde  $t$  es la variable tiempo.

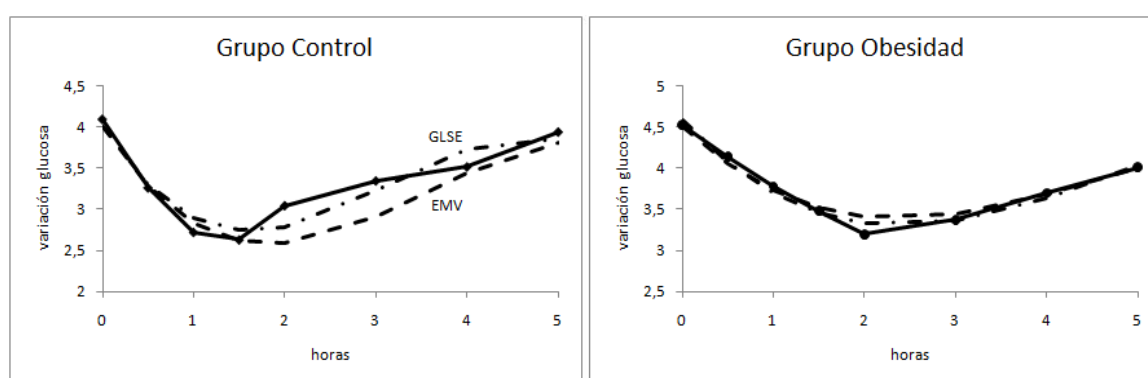


Figura 5.2.2: Modelo de Curvas de Crecimiento ajustado al nivel de fosfato inorgánico plasmático a través de los estimadores MV ( $\hat{\mathbf{B}}$ ) y GLS ( $\tilde{\mathbf{B}}$ ) (línea continua: variación media muestral; línea discontinua: EMV; línea "punto-rama": GLSE)

Una vez realizado el ajuste del modelo, a continuación, se desarrolla el estudio de la diagnóstico del mismo, especialmente el análisis de influencia. En primer lugar se aplica el criterio de detección de outliers (2.4.3), a través del estadístico  $T_i$  dado en (2.4.2).

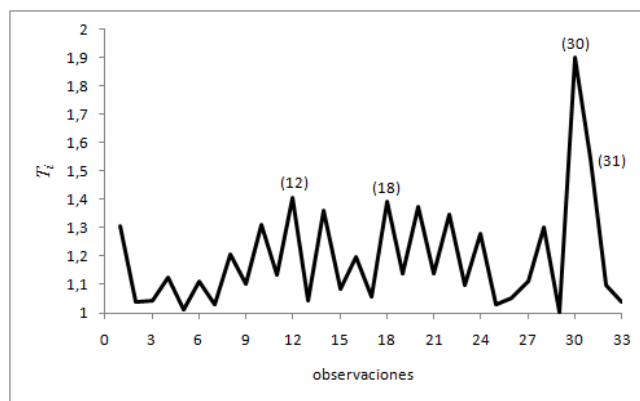


Figura 5.2.3: Medida  $T_i$  para detección de outliers

La figura 5.2.3 muestra que los valores mayores de  $T_i$  ( $1 \leq i \leq 33$ ) se alcanzan en el caso 12 ( $T_{12} = 1,4050$ ), para el grupo control, y en los casos 18 ( $T_{18} = 1,3908$ ), 30 ( $T_{30} = 1,9003$ ) y 31 ( $T_{31} = 1,5444$ ), para el grupo obesidad. Si se utiliza el criterio (2.4.3), para  $\alpha = 0,05$ ,

$$1 + \frac{qC_{\alpha}^*}{n - r - p} = 1,4863,$$

solamente las observaciones 30 y 31 (grupo obesidad) pueden ser consideradas outliers.

Identificadas las observaciones outliers, a continuación, se presentan los resultados de la aplicación de las medidas de influencia existentes en la literatura sobre  $\hat{\mathbf{B}}$ . En la figura 5.2.4 se representan diagnósticos de influencia propuestos, recogidos en el Capítulo 2 (donde se señalan aquellas observaciones con valor del diagnóstico más elevado).

A partir de estos gráficos se pueden señalar diversos aspectos:

- Las medidas de Cook,  $D_i$ , y de Welsch-Kuh,  $WK_i$  proporcionan resultados similares, determinando a los individuos 1, 10 (grupo control) y 24, 30 (grupo obesidad) como las observaciones más influyentes.
- Las medidas de Andrew-Pregibon,  $AP_i$ , y de Cook-Weisberg,  $|CW_i|$ , detectan los individuos 5 (grupo control) y 30 (grupo obesidad) como las observaciones más influyentes.
- La medida de razón de información,  $IR_i$ , detecta los individuos 30 y 31 (grupo obesidad) como las observaciones más influyentes.
- La medida de razón de covarianzas,  $|CR_i - 1|$ , es la que proporciona menor información sobre las observaciones influyentes en este caso, al no existir observaciones que se separen marcadamente de las restantes respecto de dicha medida.

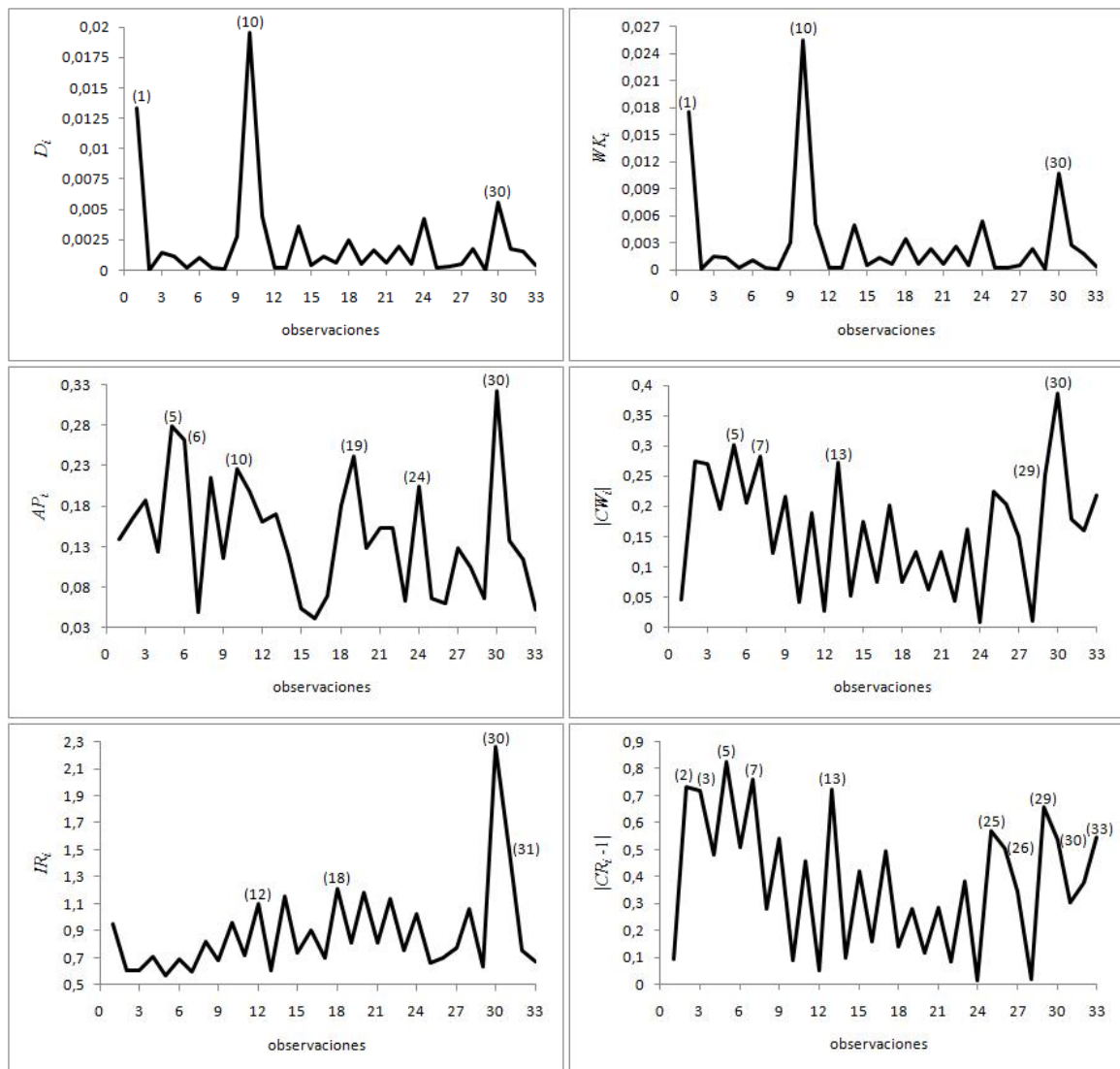


Figura 5.2.4: Medidas de influencia propuestas en la literatura

- El estudio conjunto de las medidas anteriores conduce a poder afirmar que las observaciones más influyentes corresponden a los individuos 1, 5 y 10 (grupo control) y 30 (grupo obesidad).
- De un modo general, se puede observar que las medidas de influencia propuestas en la literatura proporcionan información diferente sobre la influencia, dependiendo del estadístico o estimador sobre el que se centra el análisis. Así, este ejemplo es útil para ilustrar que el uso de una única medida de influencia proporciona información incompleta sobre las posibles observaciones influyentes.

A continuación, se presenta la aplicación de las medidas de influencia desarrolladas en esta memoria a este conjunto de datos.

La figura 5.2.5 representa la medida de Rao,  $\rho_i$ , pudiéndose observar su semejanza con la gráfica 5.2.3 del estadístico  $T_i$ . De hecho, como se ha comentado en el desarrollo teórico, dicha medida es función creciente de  $T_i$ . En conclusión, la medida  $\rho_i$  determina como casos más influyentes los casos outliers detectados por  $T_i$ .

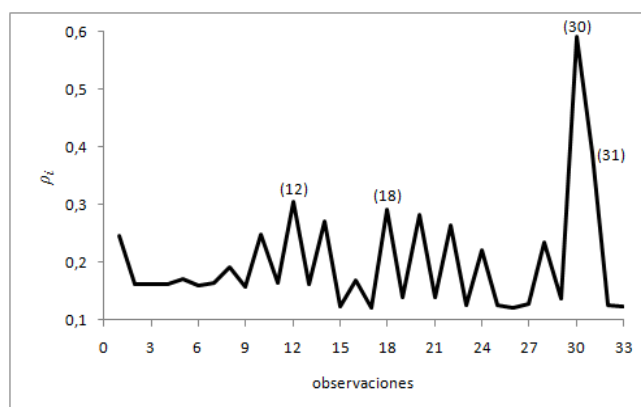


Figura 5.2.5: Medida de Rao

La figura 5.2.6 recoge la medida de Fréchet,  $\delta_i$ , a partir de la cual se puede determinar los individuos 1, 9 y 10 (grupo control) y los individuos 18, 22, 24 y 30 (grupo obesidad) como las observaciones más influyentes. La figura también recoge el gráfico de dispersión de las dos componentes de la medida de Fréchet, es decir, de  $LC_i^2$  (componente de localización) versus  $DC_i^2$  (componente de dispersión), para  $i = 1, \dots, n$ , que puede ser muy ilustrativo para determinar el sentido de la influencia que ejerce cada observación.

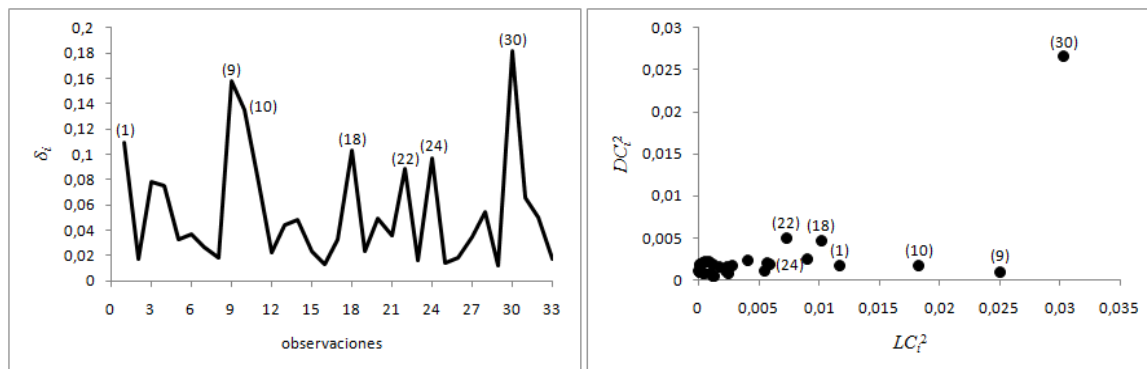


Figura 5.2.6: Medida de Fréchet

El análisis detallado de esta medida y de sus componentes en este conjunto de datos, además de permitir confirmar la idea de que es una medida de influencia que proporciona información más completa que las anteriormente analizadas, permite señalar diversos aspectos que se recogen a continuación:

- El individuo 30 (grupo obesidad) es el caso más influyente, según la medida de Fréchet. Este caso, detectado como outlier, también es determinado como influyente por las medidas  $AP_i$  de Andrew-Pregibon,  $CW_i$  de Cook-Weisberg y la razón de información,  $IR_i$ . Las medidas  $D_i$  de Cook y  $WK_i$  de Welsch-Kuh la detectan, pero de forma menos relevante. La concordancia de las distintas medidas se debe a que este caso influye en la componente de localización de la distribución muestral del estimador  $\hat{\mathbf{B}}$  y en la componente de dispersión o estructura de covarianzas de la misma, como se puede ver en el plot de dispersión de ambas componentes.
- El caso 9 (grupo control) se detecta como el segundo caso más influyente a través de la medida de Fréchet, debido especialmente a su impacto sobre la componente de localización de la distribución muestral del estimador  $\hat{\mathbf{B}}$ . Sin embargo, esta observación no es detectada por las restantes medidas de influencia.
- Los individuos 1 y 10, ambos del grupo control, también son detectados por la medida de Fréchet por su influencia sobre la componente de localización. Ambos casos son detectados por las medidas  $D_i$  de Cook y  $WK_i$  de Welsch-Kuh, aunque no por las restantes medidas.
- En el grupo obesidad, además del caso 30, se detectan como moderadamente influyentes a través de la medida de Fréchet las observaciones 18, 22 y 24. El primer de ellos presenta valores de las medidas  $T_i$  y  $\rho_i$  elevados, así como valores moderados de las medidas de Cook y Welsch-Kuh. También los casos 22 y 24 son levemente resaltados por estas medidas.

Una vez presentadas las medidas de influencia sobre el EMV de la matriz de coeficientes de regresión, a continuación, se recoge un resumen del estudio de la influencia en las filas, columnas y elementos de dicho estimador. Teniendo en cuenta la definición de  $\mathbf{B}$ , las filas de  $\hat{\mathbf{B}}$  contienen los estimadores de los coeficientes de los polinomios ajustados de los dos grupos, mientras que las columnas contienen los estimadores de los coeficientes de regresión asociados al grupo control y al grupo obesidad (cada columna de  $\hat{\mathbf{B}}$  corresponde a un grupo). Los diagnósticos de influencia utilizados son las medidas de Cook, de Fréchet y de Rao. Se han obtenido dichas medidas y las siguientes gráficas recogen los resultados más relevantes.

Como ilustración se presenta el estudio realizado en los coeficientes del término de 2º grado de las curvas ajustadas, o sea, para la fila tercera de  $\hat{\mathbf{B}}$ , cuyos resultados se recogen en la figura 5.2.7. De una manera general se pueden realizar las siguientes consideraciones:

- Como se ha comentado en la sección 3.5.1.1, la medida de Fréchet,  $\delta_{i(j\cdot)}$ , se puede expresar como función de la distancia de Cook,  $DL_{i(j\cdot)}$ , por lo que mantienen un cierto nivel de coincidencia los resultados obtenidos a través de ambas medidas. Sin embargo, no hay coincidencia completa (véanse las observaciones 30 y 31). Por otra parte, la medida de Fréchet tiene la ventaja de permitir obtener información sobre la influencia de las

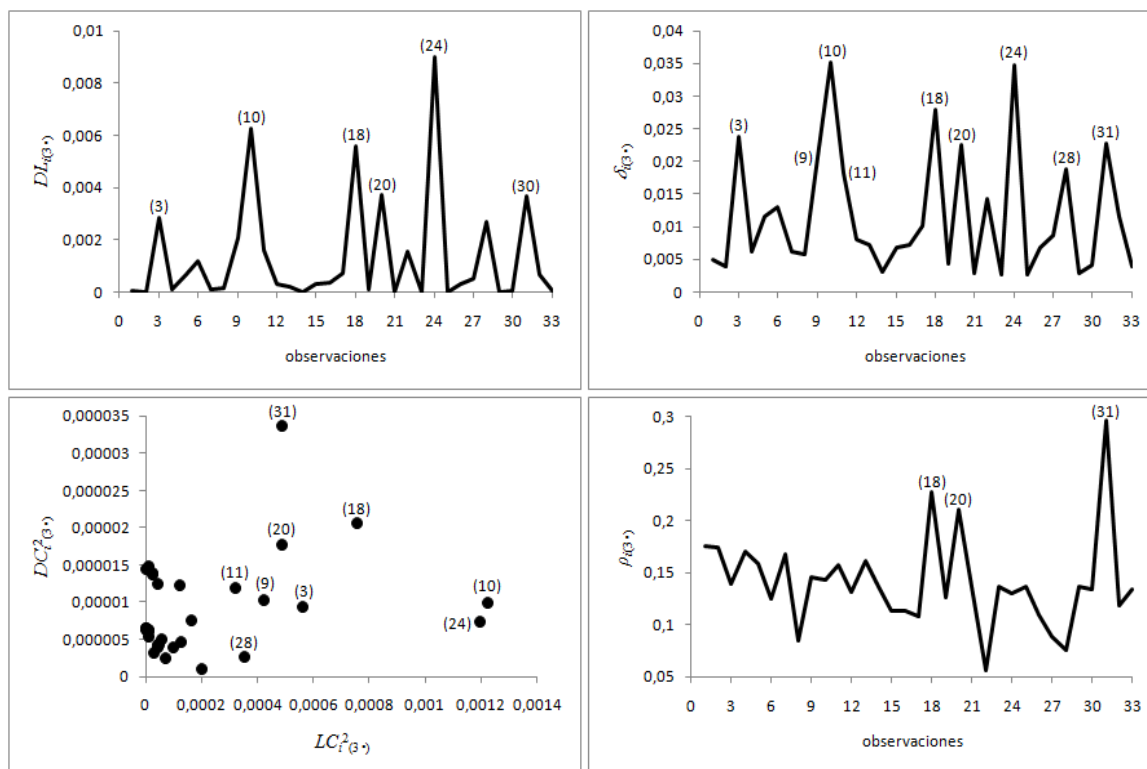


Figura 5.2.7: Medidas de influencia sobre la fila tercera de  $\hat{\mathbf{B}}$

observaciones en la localización y en la dispersión de la distribución del parámetro bajo estudio. Como ilustración de este comentario se puede observar que los casos 10 y 18 presentan similares niveles de influencia según las distancias de Cook y de Fréchet, aunque, a través del estudio de las componentes de localización y dispersión, se puede afirmar que el caso 18 ejerce la influencia en ambas componentes, mientras que el caso 10 la ejerce fundamentalmente en la localización.

- La medida de Fréchet y la de Rao detectan observaciones diferentes como más influyentes para las filas de  $\hat{\mathbf{B}}$ . Mientras la medida de Fréchet detecta observaciones influyentes en los dos grupos para todas las filas de  $\hat{\mathbf{B}}$ , la medida de Rao, de una manera general, solamente presenta relevancia en la influencia de las observaciones del 2º grupo. En particular, la medida de Rao presenta cierta concordancia en los resultados con la componente de dispersión de la medida de Fréchet.

En las figuras 5.2.8 y 5.2.9 se recogen los resultados correspondientes al estudio de la influencia sobre las columnas de  $\hat{\mathbf{B}}$ . A partir de este análisis, sin entrar en detalles específicos de las distintas observaciones, se pueden realizar diversas consideraciones:

- Se observa nuevamente la falta de coincidencia en los resultados y conclusiones que se pueden extraer con las medidas de Cook ( $DL_{i(m)}$ ), de Fréchet ( $\delta_{i(m)}$ ) y de Rao ( $\rho_{i(m)}$ ), además de la mayor información proporcionada por la distancia de Fréchet.
- Los casos más influyentes sobre el estimador de la matriz de coeficientes de regresión no son necesariamente los más influyentes sobre el estimador de cualquier subconjunto de parámetros de la misma, en particular, sobre sus columnas. Este aspecto puede quedar ilustrado por la figura 5.2.6, en la que se puede ver que el individuo 9, un paciente del grupo control, es la observación más influyente de dicho grupo para el ajuste del modelo de curvas de crecimiento del nivel de fosfato inorgánico plasmático, no es una observación influyente sobre la 2ª columna del estimador, como se ilustra en la figura 5.2.9.

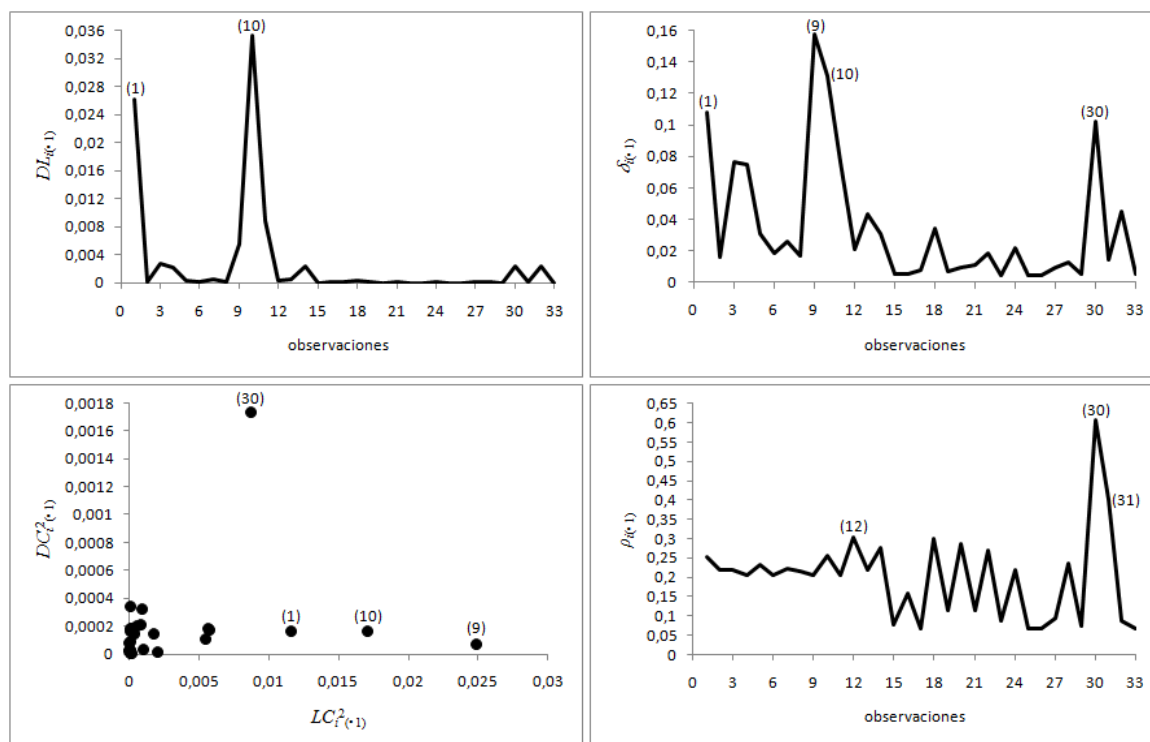


Figura 5.2.8: Medidas de influencia sobre la columna primera de  $\hat{\mathbf{B}}$

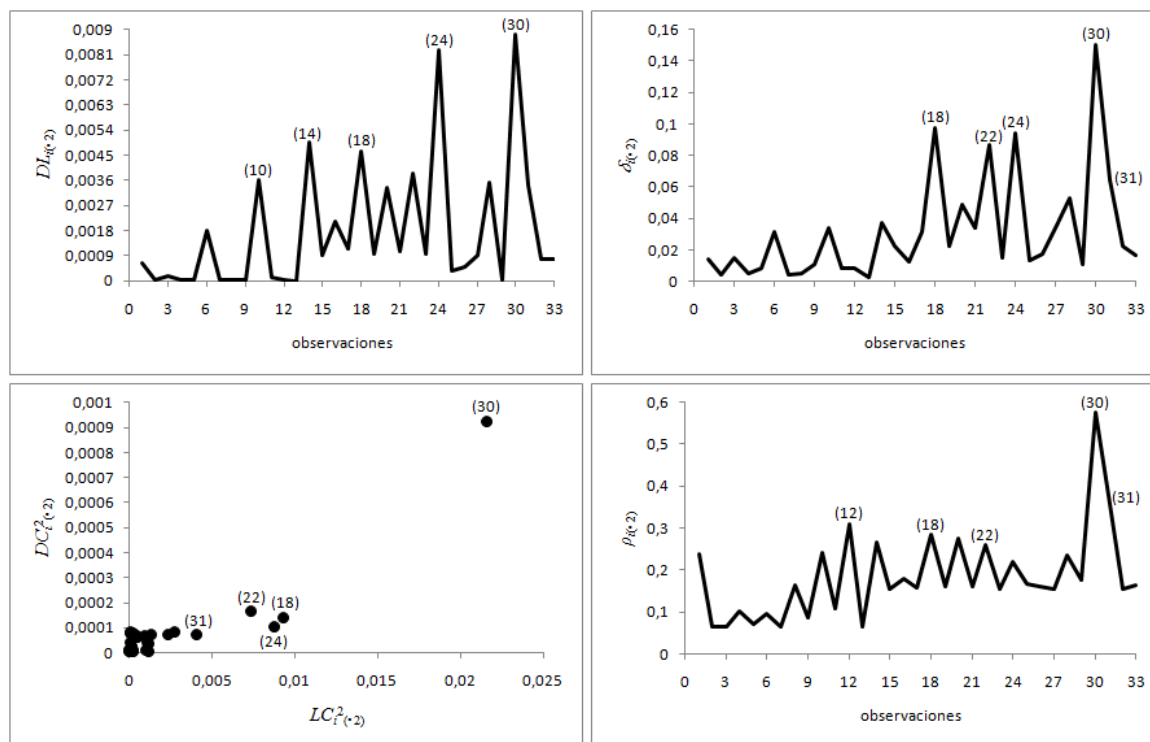


Figura 5.2.9: Medidas de influencia sobre la columna segunda de  $\widehat{\mathbf{B}}$

El estudio de la influencia de las observaciones en los elementos  $(j, m)$  de  $\widehat{\mathbf{B}}$  es ilustrado en las figuras 5.2.10 y 5.2.11, a partir de las cuales se pueden corroborar algunas de las conclusiones generales anteriormente expresadas, independientemente de las conclusiones concretas que sobre los individuos se pueden extraer a partir de un análisis detallado de las mismas.

Como conclusión, con este ejemplo se muestra la validez de las medidas propuestas en esta memoria para detectar observaciones altamente influyentes en el Modelo de Curvas de Crecimiento. Los casos determinados como influyentes son básicamente los mismos que en los estudios realizados mediante otras técnicas, aunque con nuevas aportaciones y algunos matices diferenciales propios. Por tanto, se puede afirmar que las medidas de Fréchet y de Rao constituyen una buena alternativa y un buen complemento a las medidas propuestas en la literatura para el MCC.



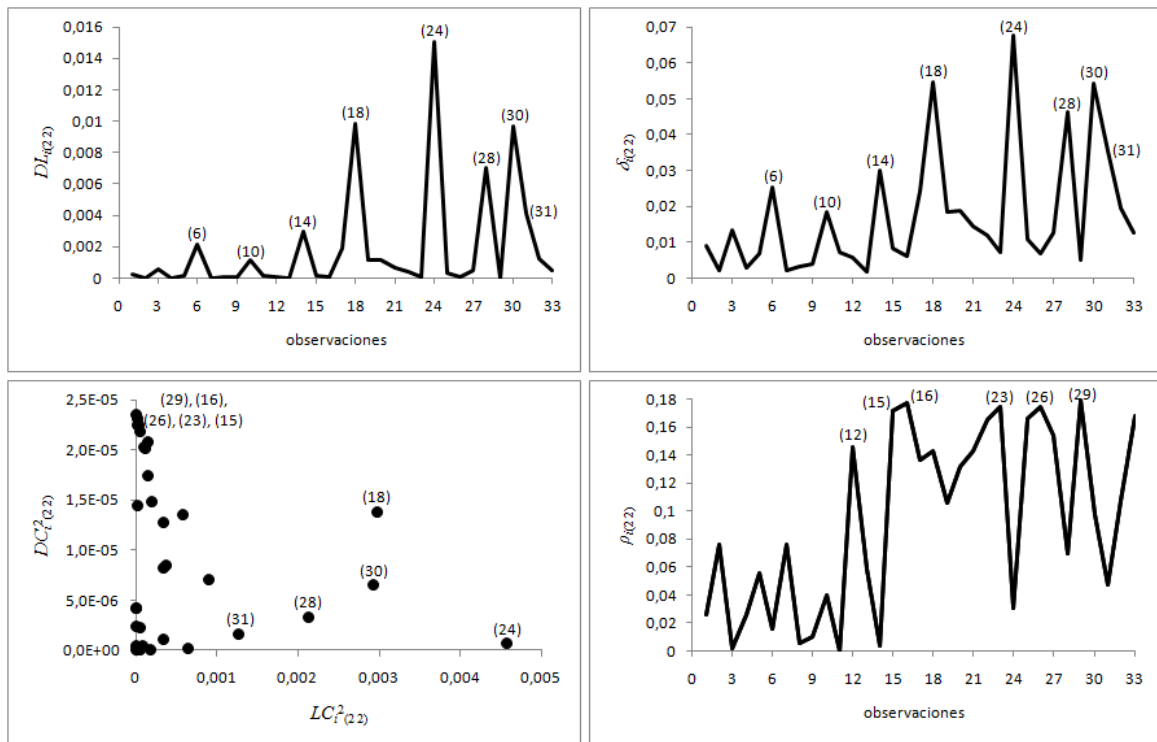


Figura 5.2.10: Medidas de influencia sobre el elemento  $\hat{b}_{22}$

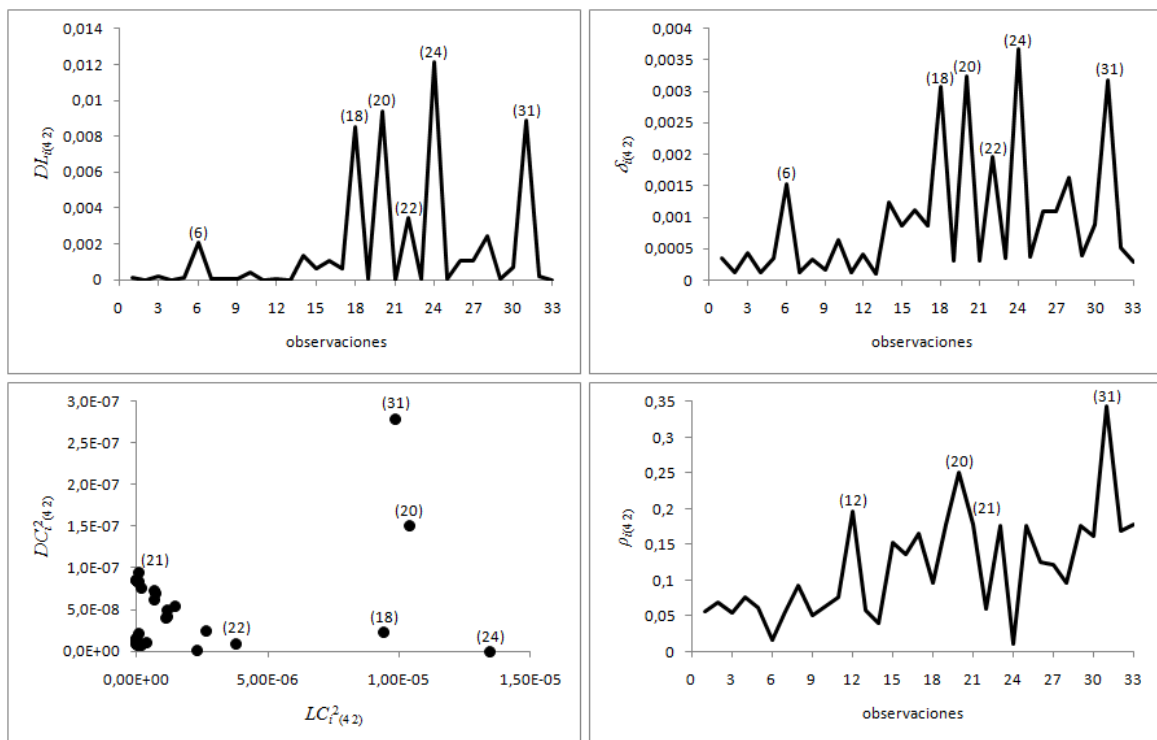


Figura 5.2.11: Medidas de influencia sobre el elemento  $\hat{b}_{42}$

### 5.3. Datos dentarios

Potthoff y Roy (1964) recogen un conjunto de datos resultantes de un estudio de ortodoncia sobre la relación entre una medida dental y la edad de un conjunto de niños de ambos sexos ( $r = 2$ ). La variable objetivo es la distancia, en milímetros, del centro de la glándula pituitaria a la fisura pterigomaxilar, en  $n_1 = 11$  niñas y  $n_2 = 16$  niños a las edades de 8, 10, 12 y 14 años ( $p = 4$ ).

Este conjunto de datos ha sido analizado en numerosos artículos sobre el MCC, entre otros, en Ohlson y von Rosen (2010) y Hamid et al. (2011) y, en particular, en el ámbito del diagnóstico de influencia para el MCC(SCS) por Pan (2002) y Pan y Fang (2002).

A continuación, se aplican los planteamientos metodológicos desarrollados anteriormente para la identificación de observaciones influyentes en el EMV de la matriz de coeficientes de regresión del MCC(SCS) y en algunas de sus combinaciones lineales.

La figura 5.3.1 ilustra las curvas de crecimiento dental en los dos grupos, Pan y Fang (2002) sugieren la aplicación del ajuste lineal (polinomio de grado 1,  $q = 2$ ) a tales curvas en ambos grupos.

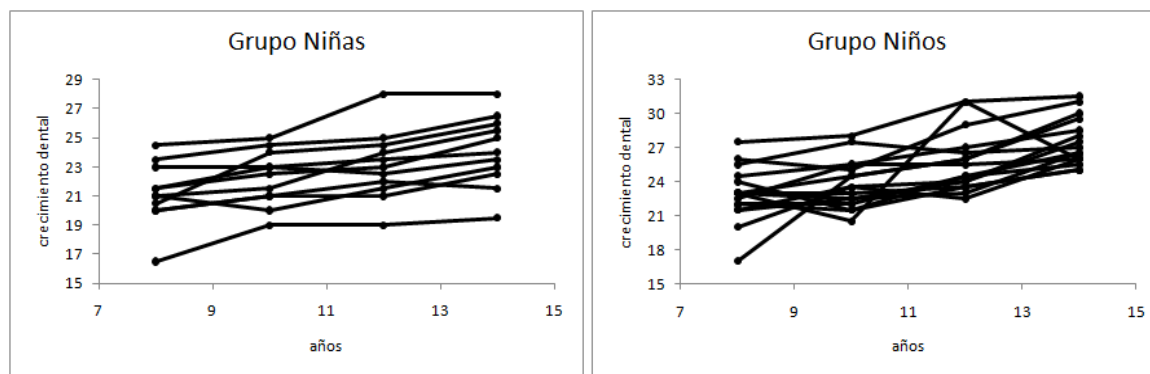


Figura 5.3.1: Curvas de crecimiento dental

Las matrices de diseño vienen dadas por

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 8 & 10 & 12 & 14 \end{bmatrix}^T \quad \text{y} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{1}_{11}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{16}^T \end{bmatrix}.$$

Por otro lado, Pan (2002) afirma que la estructura SCS, dada en la sección 1.4, es una estructura de covarianza apropiada para este conjunto de datos. En este sentido, el EMV de la

matriz de coeficientes de regresión  $\mathbf{B}$  del MCC(SCS) es

$$\tilde{\mathbf{B}} = \begin{bmatrix} 17,3727 & 16,3406 \\ 0,4795 & 0,7844 \end{bmatrix}.$$

Teniendo en cuenta las columnas de  $\tilde{\mathbf{B}}$ , las rectas  $y = 0,4795t + 17,3727$  e  $y = 0,7844t + 16,3406$ , donde  $t$  es la variable tiempo, son usadas para el ajuste de las curvas de crecimiento del grupo de las niñas y del grupo de los niños, respectivamente. El ajuste de las rectas obtenidas del EMV a la variación media del crecimiento de las observaciones de los dos grupos se presenta en la figura 5.3.2, con coeficientes de determinación  $R^2 = 0,9988$  (grupo niñas) y  $R^2 = 0,9818$  (grupo niños).

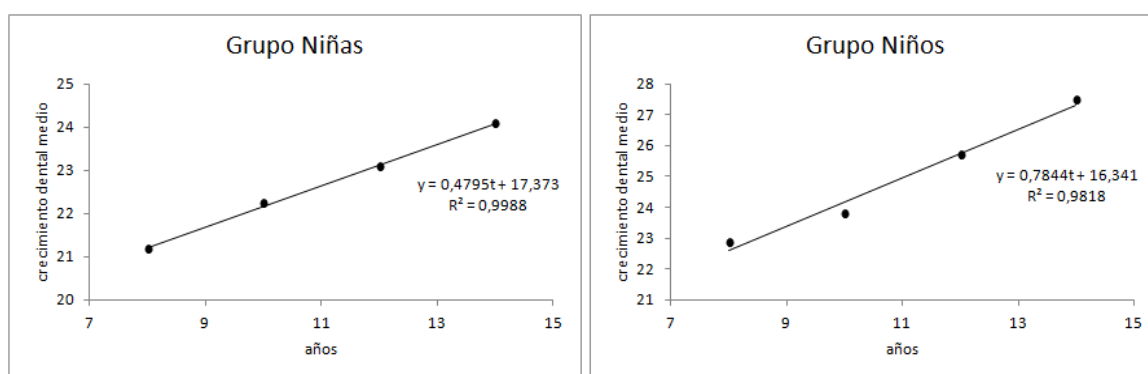


Figura 5.3.2: Modelo de Curvas de Crecimiento ajustado a la media de los dos grupos

Ajustado el modelo, a continuación, se realiza el análisis de diagnóstico de las diferentes observaciones en  $\tilde{\mathbf{B}}$ . Para la identificación de las observaciones candidatas a outliers, se usa el estadístico  $T_i$  dado en (2.5.8) a través del criterio (2.5.9), presentado en la sección 2.5. Para  $\alpha = 0,05$ , se obtiene

$$\frac{qC_{\alpha}^{r*}}{(n - r - q) + qC_{\alpha}^{r*}} = 0,2293,$$

por lo que las observaciones 21 ( $T_{21} = 0,2367$ ) y 24 ( $T_{24} = 0,5119$ ) del grupo de los niños pueden ser consideradas outliers. Estas conclusiones se ilustran en el gráfico de  $T_i$  ( $1 \leq i \leq 27$ ), recogido en la figura 5.3.3, a través del cual se observa que también conviene prestar atención a las observaciones 10 ( $T_{10} = 0,2011$ ), 11 ( $T_{11} = 0,1679$ ), del grupo de las niñas, y 15 ( $T_{15} = 0,1791$ ).

En relación a las medidas de influencia propuestas en la literatura, cuyos gráficos se presentan en la figura 5.3.3, se pueden realizar las siguientes consideraciones:

- Se observa una semejanza entre los gráficos de las medidas de Cook,  $D_i$ , de Welsch-Kuh,  $WK_i$ , y de Andrew-Pregibon,  $AP_i$ , o sea, dichas medidas identifican las mismas observaciones como las más influyentes sobre  $\tilde{\mathbf{B}}$ : los individuos 10, 11 (grupo de las niñas), 15, 21 y 24 (grupo de los niños). Esta conclusión concuerda con la detección de outliers, de hecho, dichas medidas son todas funciones crecientes de  $T_i$  (véase sección 2.5).

- Existe analogía entre la forma de los gráficos de la medida de Cook-Weisberg,  $|CW_i|$ , y de la medida de razón de covarianzas,  $|CR_i - 1|$ . No obstante, esta última medida detecta los individuos 1, 5, 7 (grupo de las niñas), 14 y 24 (grupo de los niños) como las observaciones más influyentes del conjunto de datos, mientras  $|CW_i|$  proporciona una mayor relevancia al individuo 24 como observación más influyente.
- El estudio conjunto de las medidas anteriores conduce a poder afirmar que las observaciones más influyentes corresponden a los individuos 1, 5, 7, 10, 11 (grupo de las niñas), 14, 15, 21 y 24 (grupo de los niños), siendo el individuo 24 el más influyente sobre  $\hat{B}$  para este conjunto de datos.

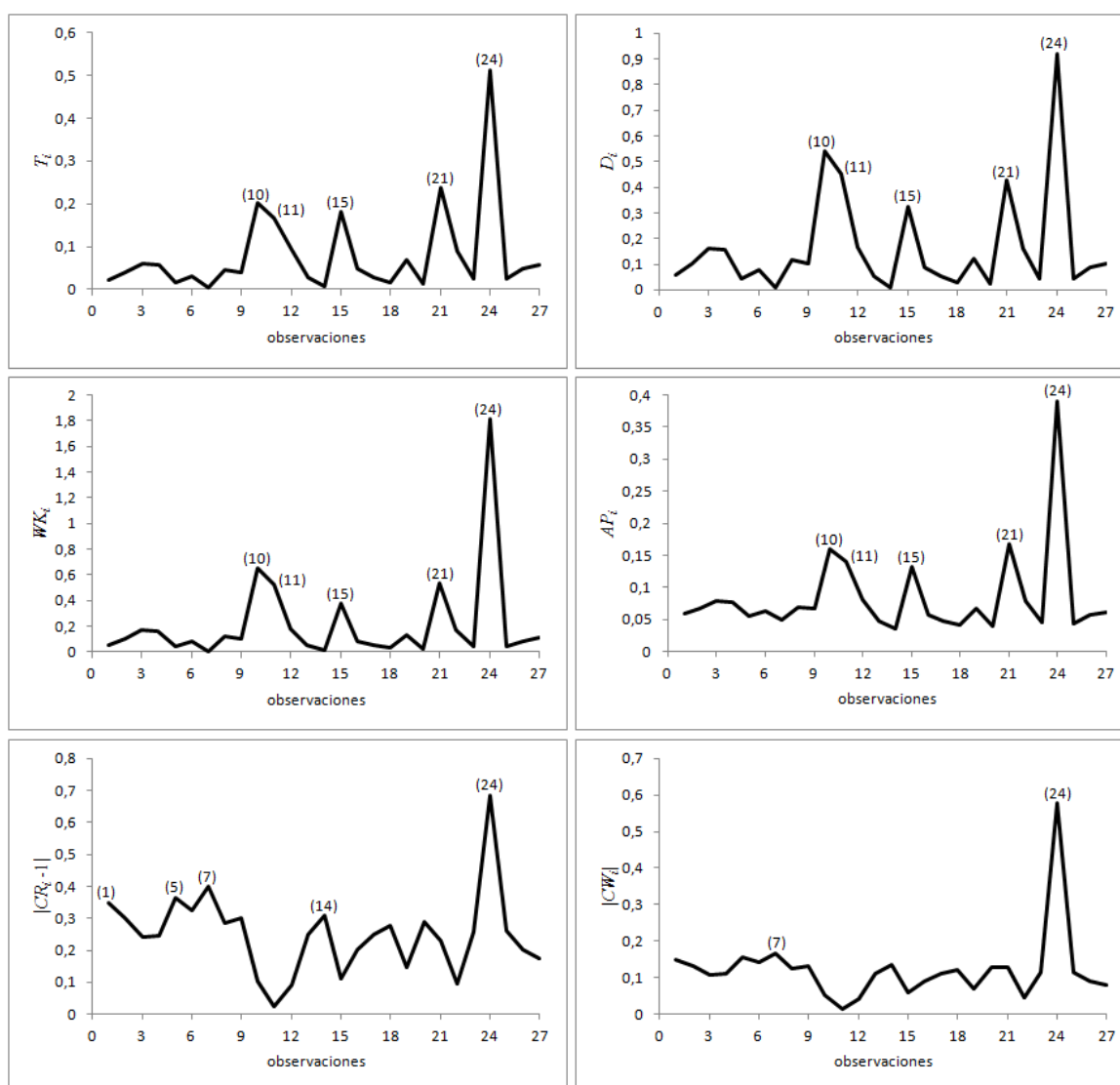


Figura 5.3.3: Medida de detección de outliers y medidas de influencia propuestas en la literatura

Respecto a la aplicación de las medidas de influencia desarrolladas en esta memoria a este conjunto de datos, la figura 5.3.4 presenta el gráfico de la medida de Rao,  $\rho_i$ , y la figura 5.3.5 presenta la medida de Fréchet,  $\delta_i$ , y el gráfico de dispersión de  $LC_i^2$  versus  $DL_i^2$ ,  $i = 1, \dots, 27$ , las dos componentes de dicha medida. Este último gráfico ilustra la influencia de las diferentes observaciones en localización y dispersión en la distribución de  $\tilde{\mathbf{B}}$ .

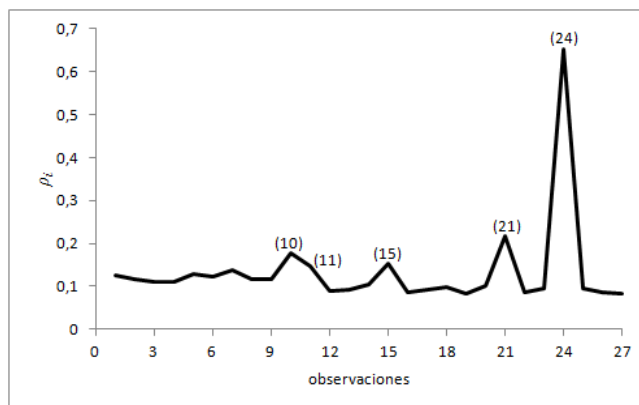


Figura 5.3.4: Medida de Rao

Analizando dichas figuras y comparándolas con las medidas propuestas en la literatura (figura 5.3.3), se puede concluir lo siguiente:

- La medida de Rao,  $\rho_i$ , determina las mismas observaciones influyentes, en el mismo orden de magnitud, que  $D_i$ ,  $WK_i$  y  $AP_i$ , por tanto, la detección de observaciones influyentes por la medida  $\rho_i$  está de acuerdo con la detección de outliers a través del estadístico  $T_i$ . De hecho, también para el MCC(SCS), la medida  $\rho_i$  es una función creciente de  $T_i$  (véase sección 4.4).

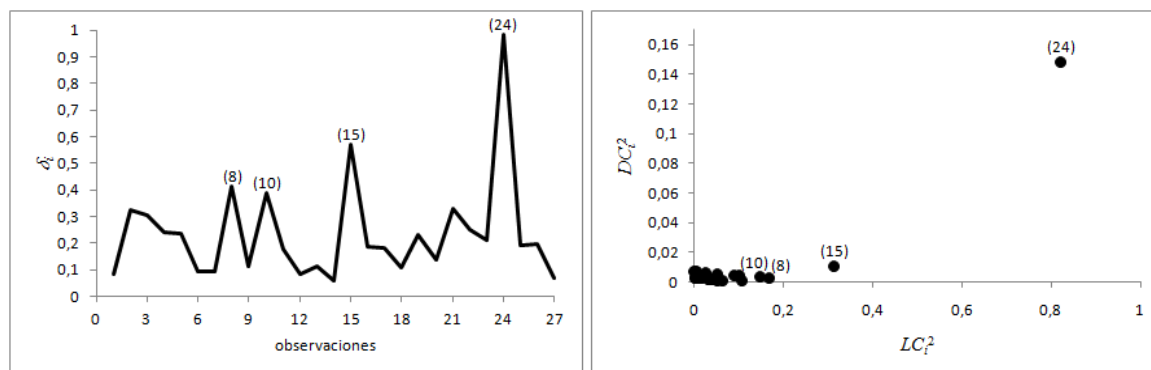


Figura 5.3.5: Medida de Fréchet

- La medida Fréchet,  $\delta_i$ , también detecta el caso 24 (grupo de los niños) como la observación más influyente en el ajuste del crecimiento del conjunto de datos dentarios. Por otro lado, se concluye que la observación 24 influye en la distribución de  $\tilde{\mathbf{B}}$  tanto en su localización como en su dispersión.
- La medida de Fréchet, además de detectar, de una manera general, las mismas observaciones influyentes que las restantes medidas (con más o menos relevancia), detecta el individuo 8 como la observación más influyente del grupo de las niñas sobre  $\tilde{\mathbf{B}}$ , observándose que dicha influencia es más grande en la localización de la distribución de  $\tilde{\mathbf{B}}$ .
- También en el MCC(SCS), se puede obtener la conclusión de que la medida de Fréchet es la que proporciona información más completa en la identificación de observaciones influyentes sobre  $\tilde{\mathbf{B}}$ .

A continuación, sin entrar en grandes detalles específicos de las distintas observaciones  $\mathbf{y}_i$ , ( $1 \leq i \leq 27$ ), se estudia la influencia en las filas, columnas y elementos de  $\tilde{\mathbf{B}}$ . Sus filas corresponden a las intersecciones en el origen y en las pendientes del tiempo en el análisis del crecimiento, respectivamente, para los dos grupos, mientras que las columnas corresponden a los coeficientes de las rectas ajustadas a los dos grupos.

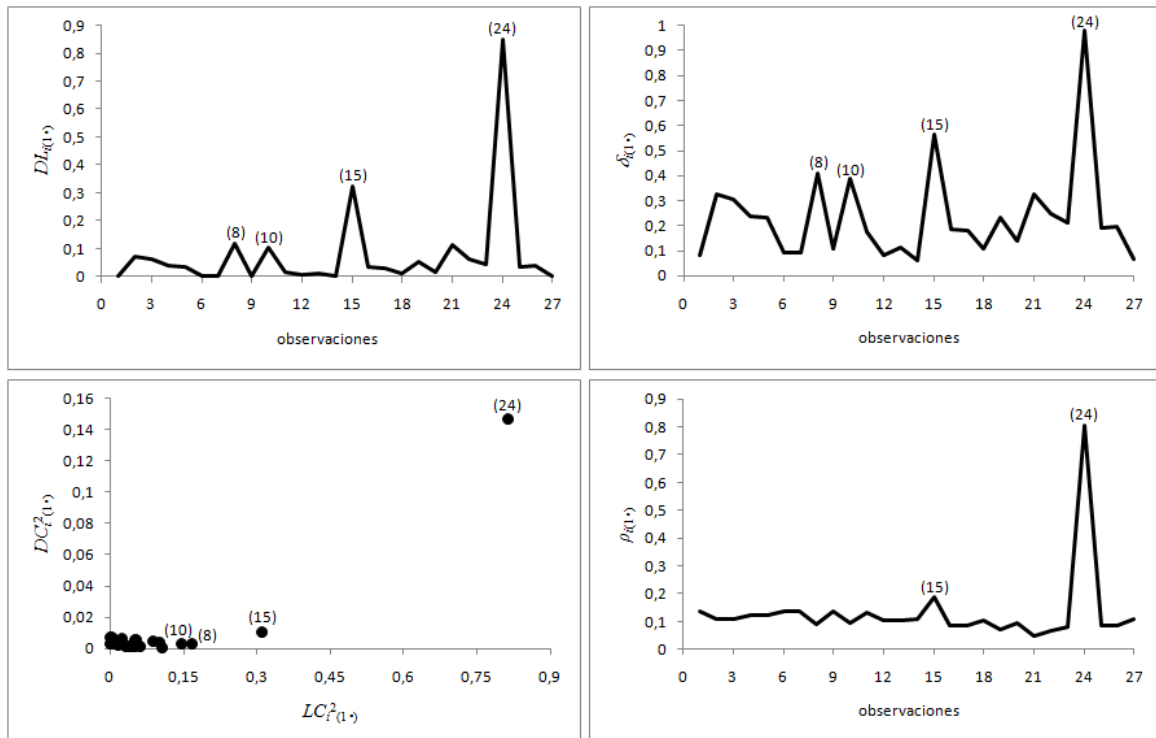


Figura 5.3.6: Medidas de influencia sobre la fila primera de  $\tilde{\mathbf{B}}$

En relación a las filas de  $\tilde{\mathbf{B}}$ , de las figuras 5.3.6 y 5.3.7 sobresalen las siguientes consideraciones:

- Las medidas de Cook,  $DL_{i(j\cdot)}$ , y Fréchet,  $\delta_{i(j\cdot)}$ , detectan las mismas observaciones influyentes (véase la nota 3.5.3), en la que  $\delta_{i(j\cdot)}$  se puede expresar como función de  $DL_{i(j\cdot)}$ . Pero, la medida de Fréchet tiene la ventaja de dar información sobre la influencia de cada observación en la localización y la dispersión.
- Hay una gran diferencia entre los gráficos de las medidas  $\delta_{i(j\cdot)}$  y  $\rho_{i(j\cdot)}$ . La medida de Fréchet detecta observaciones influyentes en la distribución de las dos filas de  $\tilde{\mathbf{B}}$  para los dos grupos, mientras que, de una manera general, la medida de Rao detecta básicamente la observación 24 como influyente en las dos filas. La medida de Rao presenta cierta concordancia en los resultados con la componente de dispersión de la medida de Fréchet.

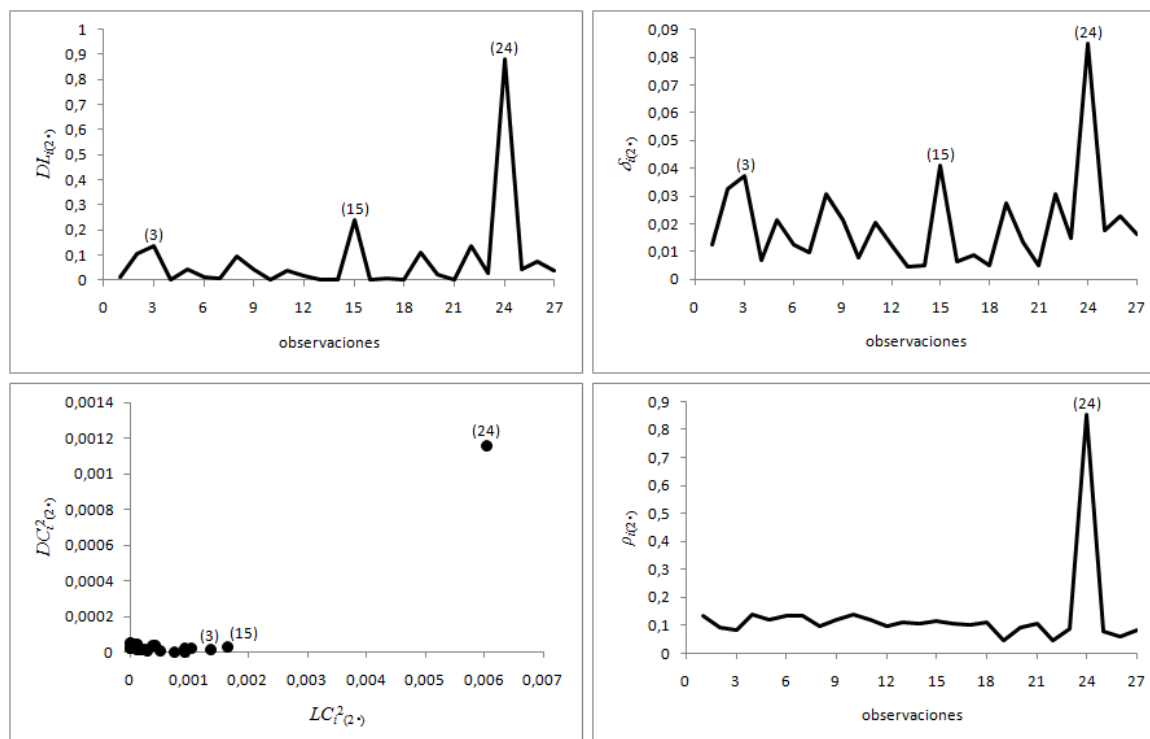


Figura 5.3.7: Medidas de influencia sobre la fila segunda de  $\tilde{\mathbf{B}}$

Para el análisis de la influencia en las columnas de  $\tilde{\mathbf{B}}$ , las figuras 5.3.8 y 5.3.9 recogen los resultados obtenidos para las medidas de Cook,  $DL_{i(\cdot m)}$ , Fréchet,  $\delta_{i(\cdot m)}$ , y de Rao,  $\rho_{i(\cdot m)}$ . Se puede concluir lo siguiente:

- De una manera general, no hay concordancia en los resultados y conclusiones que se pueden extraer de las medidas de Cook, Fréchet y Rao con respecto a la detección de observaciones influyentes en el ajuste lineal del crecimiento de los dos grupos.

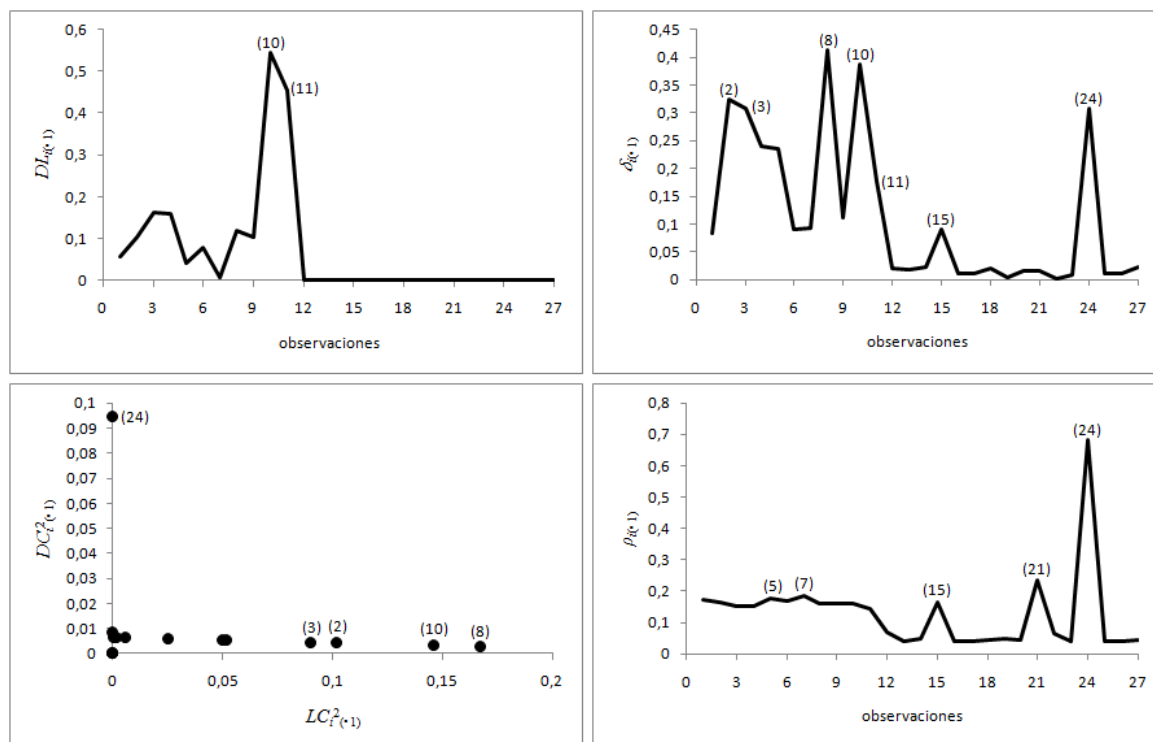


Figura 5.3.8: Medidas de influencia sobre la columna primera de  $\tilde{\mathbf{B}}$

- Contrariamente a las medidas de Fréchet y Rao, la medida de Cook para el MCC(SCS) asume que cada grupo tiene una curva de crecimiento diferente, de manera que las observaciones de un grupo no tienen influencia en la variación del otro grupo.
- Para la medida de Fréchet, la influencia de una observación en la estimación de la ecuación de un grupo al que dicha observación no pertenece sólo se presenta en la componente de dispersión (véase nota 3.5.4), es el caso de la observación 24 en el ajuste lineal del grupo de las niñas. La influencia de las observaciones 8 y 10, que pertenecen al grupo de las niñas, en el ajuste del dicho grupo se presenta en ambas componentes, localización y dispersión.
- La medida de Rao, de una manera general, detecta observaciones influyentes en concordancia con los outliers (como se recoge en la sección 4.4.1.2,  $\rho_{i(\cdot m)}$  es una función creciente de  $T_i$ ).
- Usando las medidas de Fréchet y de Cook, se concluye que el individuo que ejerce mayor influencia en el modelo ajustado no es necesariamente influyente (o el más influyente) en la estimación de un subconjunto de los parámetros.



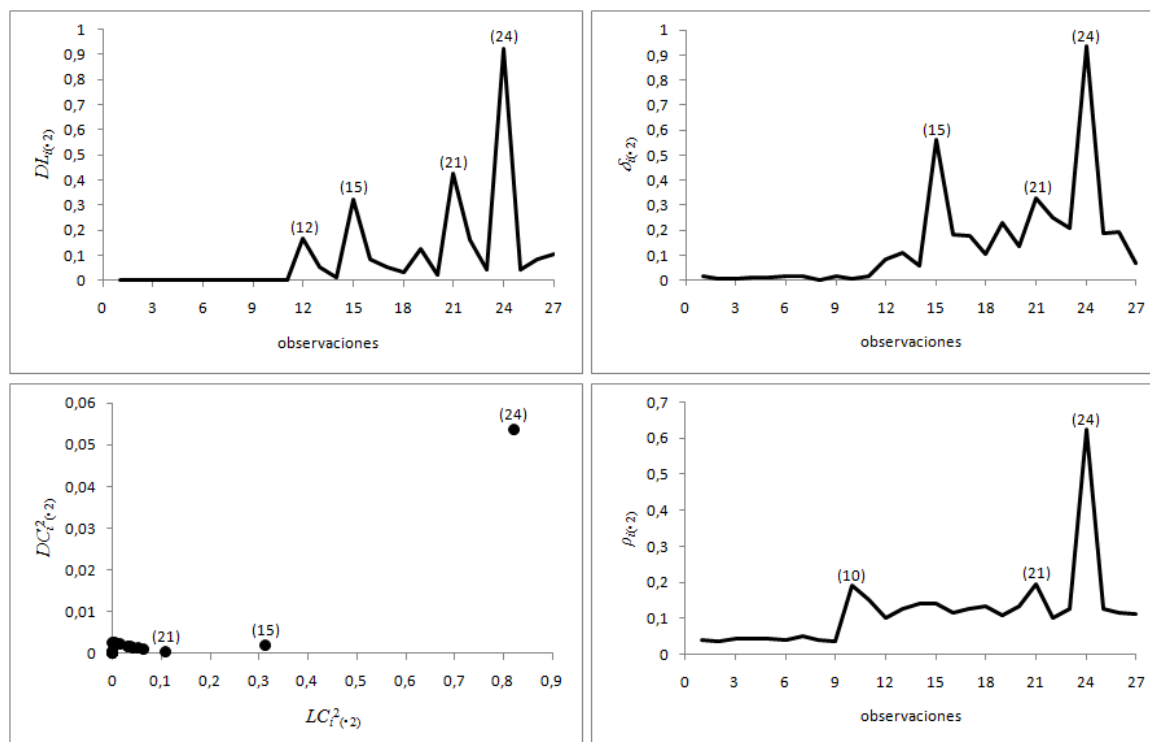


Figura 5.3.9: Medidas de influencia sobre la columna segunda de  $\tilde{\mathbf{B}}$

El estudio de la influencia de las observaciones en cada elemento  $(j, m)$  de  $\tilde{\mathbf{B}}$  corrobora algunas de las conclusiones generales anteriormente expresadas, además puede servir como un complemento al análisis previamente realizado. Como ilustración en las figuras 5.3.10 y 5.3.11 se presentan los resultados de la medida de Cook,  $DL_{i(jm)}$ , de Fréchet,  $\delta_{i(jm)}$ , y de Rao  $\rho_{i(jm)}$  para dos de los elementos de  $\tilde{\mathbf{B}}$  (las pendientes de las curvas ajustadas a los dos grupos).

Como conclusión, al igual que en el caso anterior, con este ejemplo se muestra la validez de las medidas propuestas en esta memoria para identificar observaciones altamente influyentes en el Modelo de Curvas de Crecimiento con estructura simple de Rao. Por tanto, se puede afirmar que las medidas de Fréchet y de Rao constituyen una buena alternativa y un buen complemento a las medidas propuestas en la literatura para el MCC(SCS).

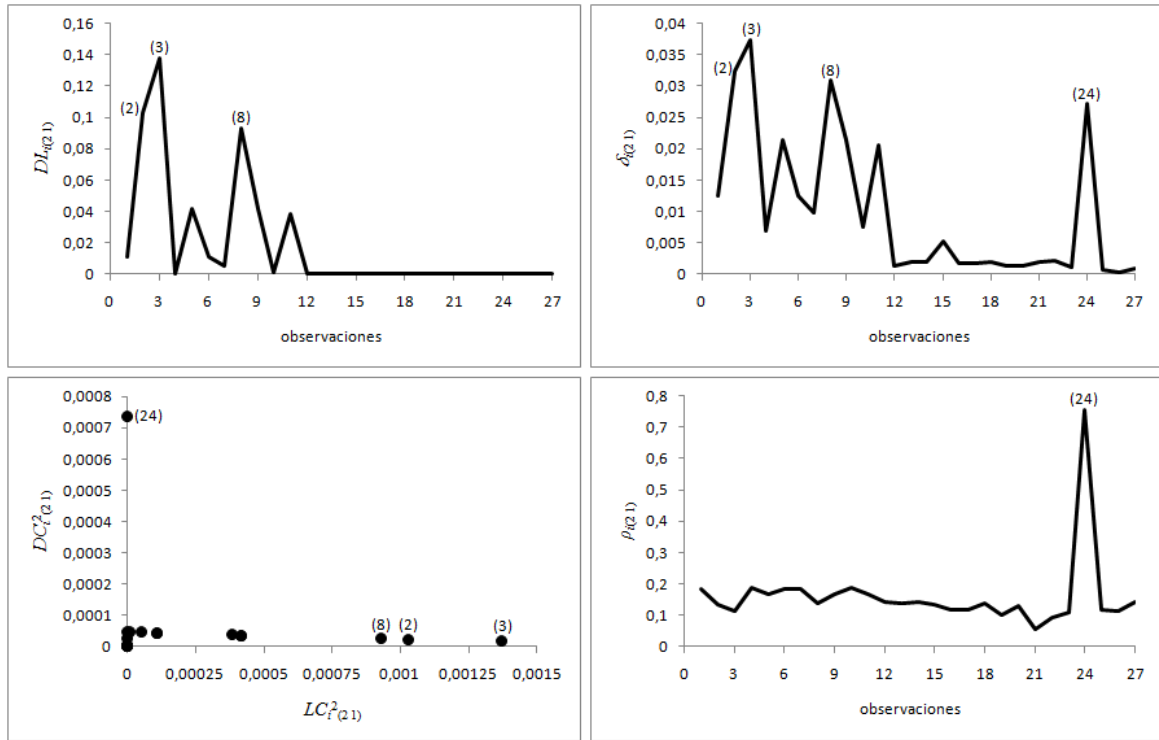


Figura 5.3.10: Medidas de influencia sobre el elemento  $\tilde{b}_{21}$

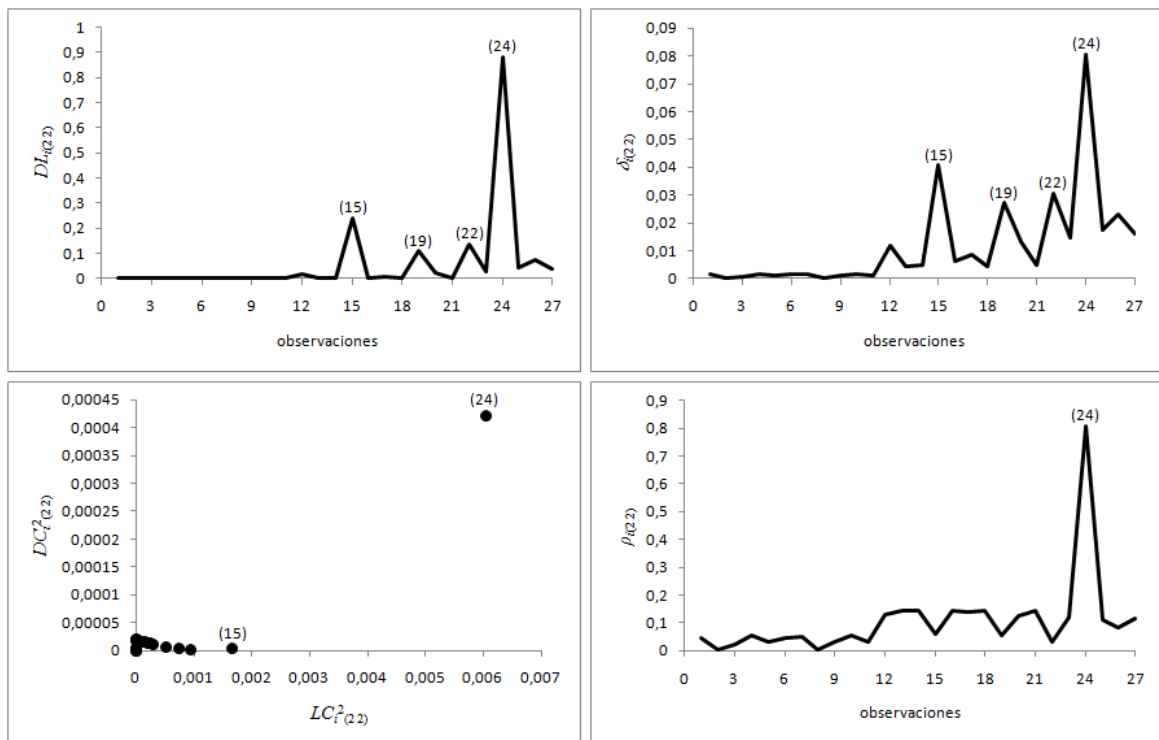


Figura 5.3.11: Medidas de influencia sobre el elemento  $\tilde{b}_{22}$

# Conclusiones y futuras líneas de investigación

El objetivo planteado para esta memoria fue el estudio del análisis de influencia en el modelo de curvas de crecimiento (MCC) a través de la aplicación de las distancias entre distribuciones de probabilidad. Tradicionalmente, el análisis de influencia sobre un estadístico se ha centrado en la comparación de los valores obtenidos tras una perturbación del modelo y/o del conjunto de datos. Esto es, sólo sobre el valor del estadístico obtenido en una muestra concreta. La aproximación al problema que se propone en esta memoria se centra globalmente en la distribución de valores que puede tomar el estadístico bajo estudio. Ello proporciona medidas de influencia más globales y, por tanto, menos sensibles a posibles fluctuaciones muestrales. En este sentido, se proponen dos nuevas medidas de influencia alternativas a las medidas existentes, basadas en dos distancias entre distribuciones de probabilidad clásicas en Estadística, la distancia de Fréchet (Fréchet, 1957) y la distancia de Rao (Rao, 1949), para el estimador de máxima verosimilitud (EMV) de la matriz de coeficientes de regresión del MCC y del MCC bajo la estructura simple de Rao (MCC(SCS)). Para ello, se usan las técnicas de análisis de influencia global, considerando como modelo perturbado uno de los más comúnmente utilizados en la literatura estadística del análisis de influencia, la omisión de un conjunto de observaciones que se ha denotado por  $MCC(I)$ .

Las principales aportaciones y conclusiones de este trabajo se resumen a continuación.

En el MCC, bajo la hipótesis de normalidad de las observaciones, el EMV de la matriz de coeficientes de regresión  $\mathbf{B}$  es una función no lineal de la matriz de respuestas, que no se distribuye según una distribución normal multivariante. No obstante, su distribución puede ser aproximada bien por una distribución normal matricial o bien por la mixtura de dos distribuciones elípticas matriciales, una distribución normal y una distribución de Kotz. Por otro lado, el EMV de  $\mathbf{B}$  en el MCC(SCS) es una función lineal de las respuestas que se distribuye según una normal matricial. Sobre estas distribuciones y las correspondientes en el modelo perturbado se aplican las distancias de Fréchet y Rao, obteniéndose medidas de influencia que se suman y complementan a los diagnósticos de influencia ya propuestos. Estas medidas proporcionan información relevante sobre aspectos de interés no contemplados en las existentes, tales como las componentes de localización y dispersión de la distancia de Fréchet.

La metodología antes citada se ha aplicado en el MCC(SCS), dado que el modelo posee una estructura de covarianzas que engloba un amplio número de estructuras de gran interés práctico, tales como la estructura de efectos aleatorios y la estructura uniforme (véase sección 1.4). De forma similar al modelo general, se obtienen medidas de influencia que complementan a las ya propuestas en la literatura.

En líneas generales, este trabajo ha ido especialmente dirigido a la obtención, mediante las distancias de Fréchet y Rao, de medidas de influencia sobre el EMV de la matriz de coeficientes de regresión del MCC y del MCC(SCS), que complementaran y fueran alternativas a las medidas existentes. La principal conclusión que se puede obtener de este trabajo es que, como se ha puesto de relieve, es posible la utilización de las distancias de Fréchet y Rao como medidas de influencia en el MCC y en el MCC(SCS). Además, dichas medidas son útiles para detectar observaciones influyentes y para interpretar diferentes aspectos de los efectos de dichas observaciones. La medida de Fréchet tiene la gran ventaja de permitir la interpretación de los efectos de las observaciones en las componentes de localización y dispersión de la distribución de los estimadores bajo estudio. Por otra parte, la medida de Rao es función creciente de los residuos, por lo que está directamente relacionada con las técnicas de detección de outliers.

A partir de los comentarios anteriores y el número elevado de medidas de influencia propuestas previamente en la literatura, cuestiones que pueden plantearse son: "¿qué medida de influencia debe ser usada en cada momento?", "¿cuál es la mejor medida de influencia?". Resulta complejo encontrar respuestas para estas cuestiones, dado que en un modelo estadístico hay múltiples aspectos que deben ser analizados, estimados o contrastados, y cada diagnóstico de influencia centra su interés sobre uno de estos aspectos, proporcionando, por tanto, información complementaria al resto de los diagnósticos.

Partiendo de los resultados incluidos en la memoria y las conclusiones antes recogidas surgen algunas futuras líneas de investigación que pueden extender y ampliar el trabajo realizado en esta tesis.

Una de ellas es el análisis de influencia sobre modelos con otras estructuras de covarianza. Ante los resultados obtenidos, parece razonable desarrollar la aplicación de las distancias de Fréchet y Rao como medidas de influencia a otras estructuras de covarianza usadas en el MCC. Por ejemplo, la estructura esférica (Pan y Fang, 2002) y la estructura autoregressiva de 1ª orden o serial (Lee, 1988 y Kanda, 1994).

Otra línea de investigación abierta es explorar la posibilidad de desarrollar las medidas de influencia basadas en la distancia de Fréchet y Rao en el modelo de curvas de crecimiento extendido

$$\mathbf{Y} = \sum_{i=1}^m \mathbf{X}_i \mathbf{B}_i \mathbf{Z}_i + \mathcal{E},$$

que se configura como suma de modelos de curvas de crecimiento, ampliamente desarrollado en Kollo y von Rosen (2005). Este modelo es también conocido como modelo de suma de perfiles (Verbyla y Venables, 1988).

También se pueden ampliar los resultados de esta memoria con el estudio de la relación entre las medidas propuestas y los residuos del modelo definidos por von Rosen (1995b). Para los modelos univariantes, el análisis de los residuos ha sido ampliamente abordado y la mayoría de las técnicas de diagnóstico están basadas en los mismos. Los trabajos dedicados al estudio de los residuos en los modelos multivariantes son más limitados, en particular en el MCC. von Rosen (1995b) concluye que, debido a la estructura bilineal de este modelo, los residuos ordinarios se pueden descomponer en tres componentes que dan información sobre las estructuras

”entre” y ”dentro” de los individuos. Los diagnósticos para la detección de outliers y observaciones influyentes en el MCC y en el MCC(SCS) propuestos en la actualidad dependen de la componente de los residuos que proporciona información sobre la estructura ”entre” individuos, es decir, los residuos definidos por (2.2.2). No obstante, puede ser útil analizar las relaciones de estas medidas con las otras componentes de los residuos, no sólo desde el punto de vista del análisis de influencia, sino en la validación de las hipótesis subyacentes en el modelo.

Finalmente se podría abordar el problema del enmascaramiento, que se produce cuando la muestra incluye un grupo de observaciones tales que su influencia conjunta enmascara el efecto individual de cada una de ellas, provocando que éste no sea detectado mediante el uso de los estadísticos que analizan la influencia individual de cada observación. Aunque en la presente memoria se han incluido las medidas de influencia conjunta basadas en las distancias de Fréchet y Rao, puede resultar de interés profundizar en el estudio de estas medidas con objeto de evitar el efecto enmascaramiento. Como se recoge en Pan y Fang (1995) y Pan y Bai (2003), los problemas del enmascaramiento en el MCC pueden ser parcialmente tratados usando la influencia local o siguiendo los planteamientos de Rousseeuw y van Zomeren (1990) y de Pan et al. (2000). No obstante, el problema del enmascaramiento no ha sido aún satisfactoriamente resuelto para el MCC. Por ello, parece conveniente intentar encontrar una estrategia para estudiar este problema basada en las medidas propuestas en esta memoria. Para ello, se pueden seguir los planteamientos de Peña y Yohai (1995), un método usado para el modelo de regresión lineal que permite seleccionar eficientemente grupos de observaciones influyentes que pasan desapercibidos al emplear estadísticos de influencia individual y con un coste computacional muy inferior al de otros métodos propuestos en la literatura. Otra alternativa, es el enfoque de Lawrence (1995), usado para la distancia de Cook como medida de influencia en el modelo de regresión.

---



# Bibliografía

- Abadir, K. M. y Magnus, J. R. (2005). *Matrix Algebra*. Cambridge University Press, Cambridge.
- Anderson, T. W. y Fang, K. T. (1990). Theory and applications of elliptically contoured and related distributions. *Technical Report, 24*. Department of Statistics, Standford University, Standford.
- Atkinson, C. y Mitchell, A. F. S. (1981). Rao's distance measure. *Sankhyá A*, **43**(3): 345–365.
- Barnett, V. y Lewis, T. (1984). *Outliers in Statistical Data*. John Wiley and Sons, New York.
- Barrett, B. E. y Ling, R. F. (1992). General classes of influence measures for multivariate regression. *Journal of the American Statistical Association*, **87**(417): 184–191.
- Belsley, D. A., Kuh, E. y Welsch, R. E. (1980). *Regression Diagnostic. Identifying Influential Data and Sources of Collinearity*. John Wiley and Sons, New York.
- Bilodeau, M. y Brenner, D. (1999). *Theory of Multivariate Statistics*. Springer-Verlag, New York.
- Box, G. E. P. (1950). Problems in the analysis of growth and wear curves. *Biometrics*, **6**(4): 362–389.
- Box, G. E. P. y Tiao, G. C. (1968). A bayesian approach to some outlier problems. *Biometrika*, **55**(1): 119–129.
- Burbea, J. (1984). Informative geometry of probability spaces. *Technical Reports, 84-52*. Center for Multivariate Analysis, University of Pittsburgh, Pittsburgh.
- Chatterjee, S. y Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression*. John Wiley and Sons, New York.
- Chi, E. M. y Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association*, **84**(406): 452–459.
- Chinchilli, V. M. y Carter, W. H. (1984). A likelihood ratio test for a patterned covariance matrix in a multivariate growth-curve model. *Biometrics*, **40**(1): 151–156.
- Chinchilli, V. M. y Elswick, R. K. (1985). A mixture of the MANOVA and GMANOVA models. *Communications in Statistics - Theory and Methods*, **14**(12): 3075–3089.
- Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics*, **19**(1): 15–18.

- Cook, R. D. (1986). Assesment of local influence (with discussion). *Journal of the Royal Statistical Society B*, **48**(2): 133–169.
- Cook, R. D. (1987). Influence assessment. *Journal of Applied Statistics*, **14**(2): 117–131.
- Cook, R. D. y Weisberg, S. (1982). *Residual and Influence in Regression*. Chapman and Hall, New York.
- Cuadras, C. M. (1989). Distancias estadísticas. *Estadística Española*, **30**(119): 295–378.
- Cuadras, C. M., Oller, J. M., Arcas, A. y Rios, M. (1985). Métodos geométricos de la estadística. *Questiío*, **9**(4): 219–250.
- Dowson, D. C. y Landau, B. V. (1982). The Fréchet distance between multivariate normal distributions. *Journal of Multivariate Analysis*, **12**(3): 450–455.
- Elston, R. C. y Grizzle, J. E. (1962). Estimation of time-response curves and their confidence bands. *Biometrics*, **18**(2): 148–159.
- Elswick, R. K. (1985). *The missing data problem as applied to the extended version of the GMANOVA model*. PhD thesis. Virginia Commonwealth University, Richmond, Virginia.
- Fang, K. T., Kotz, S. y Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.
- Fitzmaurice, G. M., Laird, N. M. y Ware, J. H. (2004). *Applied Longitudinal Analysis*. John Wiley and Sons, New Jersey.
- Fréchet, M. (1957). Sur la distance de deux lois de probabilité. *C. R. Acad. Sci. Paris*, **244**: 689–692.
- Fujikoshi, Y. (1974). Asymptotic expansions of the non-null distributions of three statistics in GMANOVA. *Annals of the Institute of Statistical Mathematics*, **26**(1): 289–297.
- Fujikoshi, Y. (1985). An error bound for an asymptotic expansion of the distribution function of an estimate in a multivariate linear model. *The Annals of Statistics*, **13**: 827–831.
- Fujikoshi, Y. (1987). Error bounds for asymptotic expansions of the distribution of the MLE in a GMANOVA model. *Annals of the Institute of Statistical Mathematics A*, **39**(1): 153–161.
- Fujikoshi, Y. y Nishii, R. (1983). Asymptotic comparison of three confidence regions in analysis of growth curves. *Journal of the Japan Statistical Society*, **13**: 151–155.
- Fujikoshi, Y. y Shimizu, R. (1989). Asymptotic expansions of some mixtures of the multivariate normal distribution and their error bounds. *The Annals of Statistics*, **17**(3): 1124–1132.
- García-Heras, J., Muñoz-García, J., Muñoz-Pichardo, J. M. y Pardo, L. (2006). Influence measures based on Cressie-Read divergence measures in multivariate linear model. *Communications in Statistics - Theory and Methods*, **35**(11): 2055–2073.
- Geisser, S. (1970). Bayesian analysis of growth curves. *Sankhyā A*, **32**: 53–64.
-



- Gleser, L. J. y Olkin, I. (1970). Linear models in multivariate analysis. *En* R. Bose, I. Chakravarti, P. Mahalanobis, C. Rao y K. Smith (eds), *Essays in Probability and Statistics*. University of North Carolina Press, Chapel Hill, North Carolina. pp. 267–292.
- Gleser, L. J. y Olkin, I. (1972). Estimation for a regression model with an unknown covariance matrix. *En* L. M. Cam, J. Neyman y E. L. Scott (eds), *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*. Vol. 1. Univ. of Calif. Press, Berkeley y Los Angeles. pp. 541–568.
- Grizzle, J. E. y Allen, D. M. (1969). Analysis of growth and dose response curves. *Biometrics*, **25**(2): 357–381.
- Gupta, A. K. y Nagar, D. (2000). *Matrix Variate Distributions*. Chapman and Hall/CRC, Boca Raton.
- Gupta, A. K. y Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, Dordrecht.
- Gupta, A. K. y Varga, T. (1994). A new class of matrix variate elliptically contoured distributions. *Journal of the Italian Statistical Society*, **3**(2): 255–270.
- Hadi, A. S. (1992). A new measure of overall potential influence in linear regression. *Computational Statistics and Data Analysis*, **14**(1): 1–27.
- Hadi, A. S. y Nyquist, H. (1999). Fréchet distance as a tool for diagnosing multivariate data. *Linear Algebra and its Applications*, **289**(1/3): 183–201.
- Hamid, J. S. (2005). *New residuals in multivariate bilinear models*. PhD thesis. Department of Biometry and Engineering, Swedish University of Agricultural Sciences.
- Hamid, J. S., Beyene, J. y von Rosen, D. (2011). A novel trace test for the mean parameters in a multivariate growth curve model. *Journal of Multivariate Analysis*, **102**(2): 238–251.
- Hamid, J. S. y von Rosen, D. (2005a). An approximate critical point for a test in the growth curve model: A conditional approach. *Research Report, 2005:04*. Centre of Biostochastics, Swedish University of Agriculture Sciences, Uppsala.
- Hamid, J. S. y von Rosen, D. (2005b). Hypothesis testing via the residuals in two GMANOVA models. *Research Report, 2005:05*. Centre of Biostochastics, Swedish University of Agriculture Sciences, Uppsala.
- Hamid, J. S. y von Rosen, D. (2006). Residuals in the Extended Growth Curve Model. *Scandinavian Journal of Statistics*, **33**(1): 121–138.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. y Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley and Sons, New York.
- Ip, W. C., Wu, M. X., Wang, S. G. y Wong, H. (2007). Estimation for parameters of interest in random effects growth curve models. *Journal of Multivariate Analysis*, **98**(2): 317–327.
- Jammalamadaka, S. R., Siddiqi, M. A. y Ghosh, K. (2009). Analysis of microtubule dynamics using growth curve models. *Journal of Applied Statistics*, **36**(6): 621–631.
-

- Jiménez-Gamero, M. D., Muñoz-Pichardo, J. M., Muñoz-García, J. y Pascual-Acosta, A. (2002). Rao distance as a measure of influence in the multivariate linear model. *Journal of Applied Statistics*, **29**(6): 841–854.
- Kabe, D. G. (1975). Some results for the GMANOVA model. *Communication in Statistics - Theory and Methods*, **4**(9): 813–820.
- Kabe, D. G. (1987). GMANOVA model ANOVA and ANCOVA theory. *Journal of Industrial Mathematics Society*, **37**: 37–51.
- Kanda, T. (1990a). Confidence region of growth curve model with covariance structures. *Technical Report*, 276. Statistical Research Group, Hiroshima University.
- Kanda, T. (1990b). Growth curve model with a uniform covariance structure. *Technical Report* 278. Statistical Research Group, Hiroshima University.
- Kanda, T. (1992). MSE's of prediction in growth curve model with covariance structures. *Annals of the Institute of Statistical Mathematics*, **44**(3): 519–528.
- Kanda, T. (1994). Growth curve model with covariance structures. *Hiroshima Mathematical Journal*, **24**(1): 135–176.
- Kantorovich, L. y Rubinstein, G. (1958). On the space of completely additive functions. *Vestnik Leningrad Univ., Ser. Mat. Mekh. i Astron.*, **13**(7): 52–59.
- Kariya, T. (1978). The General MANOVA problem. *The Annals of Statistics*, **6**(1): 200–214.
- Kariya, T. (1985). *Testing in the Multivariate General Linear Model*. Vol. 22 of *Economic Research Series*. Kinokuniya Printing Company, Tokyo.
- Kariya, T. y Kurata, H. (2004). *Generalized Least Squares*. Wiley Series in Probability and Statistics. John Wiley and Sons, Chichester.
- Kenward, M. G. (1986). The distribution of a generalized least squares estimator with covariance adjustment. *Journal of Multivariate Analysis*, **20**(2): 244–250.
- Keramidas, E. M. y Lee, J. C. (1990). Forecasting technological substitutions with concurrent short time series. *Journal of the American Statistical Association*, **85**(411): 625–632.
- Keramidas, E. M. y Lee, J. C. (1995). Selection of a covariance structure for growth curves. *Biometrical Journal*, **37**(7): 783–797.
- Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curves. *Annals of the Institute of Statistical Mathematics*, **18**(1): 75–86.
- Kish, C. W. y Chinchilli, V. M. (1990). Diagnostics for identifying influential cases in GMANOVA. *Communications in Statistics - Theory and Methods*, **19**: 2683–2704.
- Kleinbaum, D. G. (1973). A generalization of the growth curve model which allows missing data. *Journal of multivariate analysis*, **3**: 117–124.
- Kollo, T., Roos, A. y von Rosen, D. (2007). Approximation of the distribution of the location parameter in the growth curve model. *Scandinavian journal of statistics*, **34**(3): 499–510.
-

- Kollo, T. y Roos, A. (2005). On Kotz-type elliptical distributions. *En* J. Fan y G. Li (eds), *Contemporary multivariate analysis and experimental designs, in celebration of professor Kai-Tai Fang's 65th birthday*. World Scientific, New Jersey. pp. 159–170.
- Kollo, T. y von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*. Springer, Dordrecht.
- Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Instituto Italiano degli Attuari*, **4**: 83–91.
- Kotz, S. (1975). Multivariate distributions at a cross-road. *En* G. P. Patil, S. Kotz y J. K. Ord (eds), *Statistical Distributions in Scientific Work*. Vol. 1. D. Reidel Publishing Company, Dordrecht. pp. 247–270.
- Krazanowski, W. J. (1996). Rao's distance between normal populations that have common principal components. *Biometrics*, **52**(4): 1467–1471.
- Kshirsagar, A. M. y Smith, W. (1995). *Growth Curves*. Marcel Dekker, New York.
- Kurata, H. (1998). A generalization of Rao's covariance structure with applications to several linear models. *Journal of Multivariate Analysis*, **67**(2): 297–305.
- Lawrence, A. J. (1995). Deletion influence and masking in regression. *Journal of the Royal Statistical Society B*, **57**(1): 181–189.
- Lee, J. C. (1988). Prediction and estimation of growth curve with special covariance structures. *Journal of the American Statistical Association*, **83**(402): 432–440.
- Lee, J. C. (1991). Tests and model selection for the general growth curve model. *Biometrics*, **47**(1): 147–159.
- Lee, J. C. y Geisser, S. (1972). Growth curve prediction. *Sankhyā A*, **34**: 393–412.
- Lee, J. C. y Geisser, S. (1975). Applications of growth curve prediction. *Sankhyā A*, **37**: 239–256.
- Leech, F. B. y Healy, J. R. M. (1959). The analysis of experiments on growth rate. *Biometrics*, **15**(1): 98–106.
- Liski, J. C. (1985). Estimation from incomplete data in growth curves models. *Communications in Statistics - Simulation and Computation*, **14**(1): 13–27.
- Liski, J. C. (1991). Detecting influential measurements in a growth curves model. *Biometrics*, **47**(2): 659–668.
- Liski, J. C. y Nummi, T. (1990). Prediction in growth curve models using the EM algorithm. *Computational Statistics and Data Analysis*, **10**(2): 99–108.
- Liu, A. (1998). Detecting influential covariates in a growth curve model. *Computational Statistics and Data Analysis*, **28**(1): 105–113.
- Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proc. Nat. Inst. Sci. India*, **2**(1): 49–55.
-

- Matusita, K. (1955). Decision rules based on the distance for problems of fit, two samples and estimation. *Annals of Mathematical Statistics*, **26**(4): 631–640.
- Mitchell, A. F. S. y Krzanowski, J. (1985). The Mahalanobis distance and elliptic distributions. *Biometrika*, **72**(2): 464–467.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley and Sons, New York.
- Muñoz-García, J., Muñoz-Pichardo, J. M. y Pardo, L. (2006). Cressie and Read power-divergences as influence measures for logistic regression models. *Computational Statistics and Data Analysis*, **50**(11): 3199–3221.
- Muñoz-Pichardo, J. M. (2006). Análisis de influencia a través de distancias entre distribuciones. *Proyecto de investigación*. Departamento de Estadística e Investigación Operativa, Universidad de Sevilla, Sevilla.
- Muñoz-Pichardo, J. M., Enguix-González, A., Muñoz-García, J. y Pascual-Acosta, A. (2004). The Fréchet metric as a measure of influence in multivariate linear models with random errors elliptically distributed. *Computational Statistics and Data Analysis*, **46**(3): 469–491.
- Muñoz-Pichardo, J. M., Moreno-Rebollo, J. L., Enguix-González, A. y Pascual-Acosta, A. (2008). Influence measures on profiles analysis with elliptical data through Fréchet's metric. *Metrika*, **68**(1): 111–127.
- Nadarajah, S. (2003). The Kotz-type distribution with applications. *Statistics*, **37**(4): 341–358.
- Naik, D. (1990). Prediction intervals for growth curves. *Journal of Applied Statistics*, **17**(2): 245–254.
- Nummi, T. (1995). *Estimation and prediction in the growth curve models with applications*. PhD thesis. Department of Mathematical Science, University of Tampere, Finland.
- Nummi, T. (1997). Estimation in a random effects growth curve model. *Journal of Applied Statistics*, **24**(2): 157–168.
- Ohlson, M. (2009). *Studies in estimation of patterned covariance matrices*. PhD thesis. Department of Mathematics, Linköping University, Sweden.
- Ohlson, M. y Srivastava, M. S. (2010). Profile analysis for a growth curve model. *Journal of the Japan Statistical Society*, **40**(1): 1–21.
- Ohlson, M. y von Rosen, D. (2010). Explicit estimators of parameters in the growth curve Model with linearly structured covariance matrices. *Journal of Multivariate Analysis*, **101**(5): 1284–1295.
- Oller, J. M. (1987). Information metric for extreme value and logistic probability distributions. *Sankhyā A*, **49**(1): 17–23.
- Oller, J. M. y Cuadras, C. M. (1983). Sobre una distancia definida para la distribución normal multivariante. *Actas XIII Jornadas de Est. I.O.*. Vol. II, secc.III. Valladolid. pp. 168–173.
-

- Oller, J. M. y Cuadras, C. M. (1985). Rao's distance for negative multinomial distributions. *Sankhyá A*, **47**(1): 75–83.
- Oller, J. M. y Cuadras, C. M. (1987). Sobre ciertas condiciones que deben verificar las distancias en espacios probabilísticos. *Actas XV reunión SEIO*. Vol. 2. Universidad de Oviedo. pp. 503–509.
- Pan, J. X. (1988). Generalized least square estimator of regression parameters and Gauss-Markov theorem in growth curve model. *Journal of Mathematical Statistics and Applied Probability*, **2**: 169–185.
- Pan, J. X. (1994). Detecting influential observation in growth curve model with arbitrary covariance structure. *Research Report, 51*. Department of Mathematics, Hong Kong Baptist College, Hong Kong.
- Pan, J. X. (1995). *Multivariate statistical diagnostics with application to the growth curve model*. PhD thesis. Department of Mathematics, Hong Kong Baptist University.
- Pan, J. X. (2002). Influential observations identification in the growth curve model with Rao's simple covariance structure. *Communications in Statistics - Theory and Methods*, **31**(5): 813–831.
- Pan, J. X. (2004). Discordant outlier detection in the growth curve model with Rao's simple covariance structure. *Statistics and Probability Letters*, **69**(2): 135–142.
- Pan, J. X., Fang, K. T. y Liski, E. P. (1994). Bayesian local influence in the growth curve model with Rao's simple covariance structure. *Journal of Multivariate Analysis*, **58**(1): 55–81.
- Pan, J. X., Fang, K. T. y von Rosen, D. (1997). Local influence assessment in the growth curve model with unstructured covariance. *Journal of Statistical Planning and Inference*, **62**(2): 263–278.
- Pan, J. X., Fang, K. T. y von Rosen, D. (1999). Bayesian local influence in the growth curve model with unstructured covariance. *Biometrical Journal*, **41**(6): 641–658.
- Pan, J. X., Fung, W. y Fang, K. T. (2000). Multiple outlier detection in multivariate data using projection pursuit techniques. *Journal of Statistical Inference and Planning*, **83**(1): 153–167.
- Pan, J. X. y Bai, P. (2003). Local influence analysis in the growth curve model with Rao's simple covariance structure. *Journal of Applied Statistics*, **30**(7): 771–781.
- Pan, J. X. y Fang, K. (2002). *Growth Curve Models and Statistical Diagnostics*. Springer-Verlag, New York.
- Pan, J. X. y Fang, K. T. (1995). Multiple outlier detection in growth curve model with unstructured covariance matrix. *Annals of the Institute of Statistical Mathematics*, **47**(1): 137–153.
- Pan, J. X. y Fang, K. T. (1996). Influential observation in the growth curve model with unstructured covariance matrix. *Computational Statistics and Data Analysis*, **22**(1): 71–87.
-

- Pan, J. X. y Fung, W. K. (2000). Bayesian influence assessment in the growth curve model with unstructured covariance. *Annals of the Institute of Statistical Mathematics*, **52**(4): 737–752.
- Peña, D. y Yohai, V. J. (1995). The detection of influential subsets in linear regression by using an influence matrix. *Journal of the Royal Statistical Society B*, **57**(1): 145–156.
- Peng, B. y Yu, F. (2000). Bayesian local influence assessments in a growth curve model with general covariance structure. *Acta Mathematica Scientia B*, **20**(4): 563–570.
- Potthoff, R. F. y Roy, S. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**(3/4): 313–326.
- Qingming, Z. y Huaixiong, Z. (2005). Local influence assessment in a multivariate t-model with Rao's simple structure. *Acta Mathematica Scientia B*, **25**(1): 179–192.
- Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, **37**: 81–91.
- Rao, C. R. (1949). On the distance between two populations. *Sankhyā*, **9**: 246–248.
- Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrics*, **14**(1): 1–17.
- Rao, C. R. (1959). Some problems involving linear hypothesis in multivariate analysis. *Biometrika*, **46**(1/2): 49–58.
- Rao, C. R. (1961). Some observations on multivariate statistical methods in anthropological research. *Bulletin of the International Statistical Institute*, **38**: 99–109.
- Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, **52**(3/4): 447–458.
- Rao, C. R. (1966). Covariance adjustment and related problems in multivariate analysis. En P. Krishnaiah (ed.), *Multivariate Analysis I*. Academic Press, New York. pp. 87–103.
- Rao, C. R. (1967). Least square theory using an estimated dispersion matrix and its applications to measurement of signals. En L. M. Lecam y J. Neyman (eds), *5th Berkeley Symposium on Mathematical, Statistics and Probability*. University of California Press, Berkeley. pp. 355–372.
- Rao, C. R. (1972). Recent trends of research work in multivariate analysis. *Biometrics*, **28**(1): 3–22.
- Rao, C. R. (1987). Prediction of future observations in growth curve models. With discussion and a reply by the author. *Statistical Science*, **2**(4): 434–471.
- Reinsel, G. (1982). Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure. *Journal of the American Statistical Association*, **77**(377): 190–195.
- Reinsel, G. (1984a). Effects of the estimation of covariance matrix parameters in the generalized multivariate linear model. *Communications in statistics - Theory and methods*, **13**(5): 639–650.
-

- Reinsel, G. (1984b). Estimation and prediction in a multivariate random effects generalized linear model. *Journal of the American Statistical Association*, **79**(386): 406–414.
- Reinsel, G. (1985). Mean squared error properties of empirical Bayes estimators in a multivariate random effects general linear model. *Journal of the American Statistical Association*, **80**(391): 642–650.
- Rencher, A. C. (2002). *Methods of Multivariate Analysis, 2nd edition*. Wiley Series in Probability and Statistics, Chichester.
- Reverter, F. y Oller, J. M. (2003). Computing the Rao distance for Gamma distributions. *Journal of Computational and Applied Mathematics*, **157**(1): 155–167.
- Rousseeuw, P. J. y van Zomeren, B. C. (1990). Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, **85**(411): 633–651.
- Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*. John Wiley and Son, New York.
- Seber, G. A. F. (1984). *Multivariate Observations*. John Wiley, New York.
- Seber, G. A. F. (2008). *A Matrix Handbook for Statisticians*. John Wiley, New York.
- Srivastava, M. S. (1985). Multivariate data with missing observations. *Communications in statistics - Theory and methods*, **14**(4): 775–792.
- Srivastava, M. S., von Rosen, T. y von Rosen, D. (2009). Estimation and testing in general multivariate linear models with Kronecker product covariance structure. *Sankhyā A*, **71**(2): 137–163.
- Srivastava, M. S. y von Rosen, D. (1998). Outliers in multivariate regression models. *Journal of Multivariate Analysis*, **65**(2): 195–208.
- Srivastava, M. S. y von Rosen, D. (1999). Growth curve models. En S. Ghosh (ed.), *Multivariate analysis, design of experiments, and survey sampling*. Vol. 159. Marcel Dekker, New York. pp. 547–578.
- Srivastava, M. y Carter, E. M. (1983). *An Introduction to Applied Multivariate Statistics*. North-Holland, New York.
- Swamy, P. A. V. B. (1971). *Statistical Inference in Random Coefficient Regression Models*. Vol. 55 of *Lecture Notes in Operational Research and of Mathematical Systems*. Springer-Verlag, Berlin y New York.
- Tan, M. (1991). Improved estimators for the GMANOVA problem with application to Monte Carlo simulation. *Journal of Multivariate Analysis*, **38**(2): 262–274.
- Timm, N. H. (2002). *Applied Multivariate Analysis*. Springer-Verlag, New York.
- Vallender, S. (1974). Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability and its Applications*, **18**(4): 784–786.
- Verbyla, A. P. y Venables, W. (1988). An extension of the growth curve model. *Biometrika*, **75**(1): 129–138.
-

- von Rosen, D. (1984). Maximum likelihood estimates in multivariate linear normal models with special references to the growth curve model. *Research Report, 135*. Dept. of Math. Statist., University of Stockholm, Stockholm.
- von Rosen, D. (1988). Moments for the inverted Wishart distribution. *Scandinavian Journal of Statistics*, **15**: 97–109.
- von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models. *Journal of Multivariate Analysis*, **31**(2): 187–200.
- von Rosen, D. (1990). Moments for a multivariate linear normal model with application to growth curve model. *Journal of Multivariate Analysis*, **35**(2): 243–259.
- von Rosen, D. (1991a). Moments of maximum likelihood estimators in the growth curve model. *Statistics*, **22**(1): 111–131.
- von Rosen, D. (1991b). The growth curve model: A review. *Communication in Statistics - Theory and Methods*, **20**(9): 2791–2822.
- von Rosen, D. (1994). A study of residuals in the growth curve model. *Research Report, 1994:36*. Department of Mathematics, Uppsala University, Uppsala.
- von Rosen, D. (1995a). Influential observations in multivariate linear models. *Scandinavian Journal of Statistics*, **22**: 207–222.
- von Rosen, D. (1995b). Residual in the growth curve model. *Annals of the Institute of Statistical Mathematics*, **47**(1): 129–136.
- Walker, J. H. (1992). *Influence measures for the GMANOVA model*. PhD thesis. The Graduate School of Cornell University.
- Ware, J. H. (1985). Linear models for the analysis of longitudinal studies. *The American Statistician*, **39**(2): 95–101.
- Wei, B. C., Hu, Y. Q. y Fung, W. K. (1998). Generalized leverage and its applications. *Scandinavian Journal of Statistics*, **25**(1): 25–37.
- Wishart, J. (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, **20A**(1/2): 32–52.
- Wishart, J. (1938). Growth-rate determinations in nutrition studies with the bacon pig, and their analysis. *Biometrika*, **30**(1/2): 16–28.
- Woolson, R. F. y Leeper, J. D. (1980). Growth curve analysis of complete and incomplete longitudinal data. *Communications in Statistics - Theory and Methods*, **9**(14): 1491–1513.
- Wu, Q. G. (1998). Existence conditions of the uniformly minimum risk unbiased estimators in extended growth curve models. *Journal of Statistical Planning and Inference*, **69**(2): 101–114.
- Ye, R. y Wang, S. (2009). Estimating parameters in extended growth curve models with special covariance structures. *Journal of Statistical Planning and Inference*, **139**(8): 2746–2756.
-



- 
- Yokoyama, T. (1995). Statistical inference on some mixed MANOVA-GMANOVA models with random effects. *Hiroshima Mathematical Journal*, **25**(3): 441–474.
- Yu, F. (1995). Outlier score test in a growth curve model. *Systems Science and Mathematical Sciences*, **8**(2): 157–165.
- Zerbe, G. O. (1979). Randomization analysis of the completely randomized design extended to growth and response curves. *Journal of the American Statistical Association*, **74**(365): 215–221.
- Zezula, I. (2006). Special variance structures in the growth curve model. *Journal of multivariate analysis*, **97**(3): 606–618.
-



# **ANEXOS**



# Anexo A

## Notación

### Funciones

$\Gamma(z)$  – Función gamma

$\Gamma_m(a)$  – Función gamma multivariante

### Notación matricial

$\mathbf{1}_p$  – Vector ( $p \times 1$ ) compuesto por unos

$\mathbf{A}_{(m \times n)}$  – Matriz de orden ( $m \times n$ )

$\mathbf{A} = ((a_{ij}))$  – Matriz con elementos  $a_{ij}$

$\mathbf{A}^T$  – Transpuesta de la matriz  $\mathbf{A}$

$\mathbf{A}^o$  – Matriz ortogonal a  $\mathbf{A}$

$\text{diag}(\mathbf{A})$  – Matriz diagonal  $\mathbf{A}$

$\mathbf{I}_p$  – Matriz identidad de orden  $p$

$\mathbf{E}_{ij}(p, q)$  – Matriz de orden ( $p \times q$ ) con los  $(i, j)$ -ésimos elementos 1 y los restantes 0

$\mathbf{K}_{pq}$  – Matriz de conmutación (vec-permutación) de orden ( $pq \times pq$ ), denotada para  $p = q$  por  $\mathbf{K}_{p^2}$

$[\mathbf{A} \dot{\vdash} \mathbf{B}]$  – Matriz particionada en bloques

$\mathbf{A} > \mathbf{0}$  – Matriz  $\mathbf{A}$  definida positiva (d.p)

### Funciones y operadores matriciales

$\text{rg}(\mathbf{A})$  – Rango de una matriz  $\mathbf{A}$

$\text{tr}[\mathbf{A}]$  – Traza de una matriz cuadrada  $\mathbf{A}$

$\det[\mathbf{A}]$  – Determinante de una matriz cuadrada  $\mathbf{A}$

$\|\mathbf{A}\|$  – Norma de Frobenius (o euclídea) de una matriz  $\mathbf{A}$

$\mathbf{P}_A$  – Matriz de proyección de  $\mathbf{A}$ , i.e.,  $\mathbf{P}_A = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$

$\text{vec}(\mathbf{A})$  – Operador vec de una matriz  $\mathbf{A}$

$\text{svec}(\mathbf{A})$  – Operador svec de una matriz  $\mathbf{A}$

$\mathbf{A} \otimes \mathbf{B}$  – Producto de Kronecker de las matrices  $\mathbf{A}$  y  $\mathbf{B}$

$\text{Var}[\mathbf{X}]$  – Matriz de covarianzas de la matriz aleatoria  $\mathbf{X}$ , i.e.,  $\text{Var}[\mathbf{X}] = \text{Var}[\text{vec}(\mathbf{X})]$

### Espacios vectoriales

$\mathcal{C}(\mathbf{A})$  – Espacio columna de  $\mathbf{A}$ , el espacio generado por las columnas de  $\mathbf{A}$

$\mathcal{C}(\mathbf{A}^\circ) = \mathcal{C}(\mathbf{A})^\perp$  – Espacio ortogonal a  $\mathcal{C}(\mathbf{A})$ , las columnas de  $\mathbf{A}^\circ$  generan el espacio  $\mathcal{C}(\mathbf{A})^\perp$

$\mathbf{V}_1 \boxplus \mathbf{V}_2$  – Suma ortogonal de los subespacios vectoriales  $\mathbf{V}_1$  y  $\mathbf{V}_2$

$\mathbf{V} \otimes \mathbf{W}$  – Producto tensorial de los espacios vectoriales  $\mathbf{V}$  y  $\mathbf{W}$

### Distribuciones de probabilidad

$E_{p,n}(\mathbf{M}, \mathbf{V}, \mathbf{W}, \phi)$  – Distribución elíptica matricial

$N_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W})$  – Distribución normal matricial

$K_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W})$  – Distribución de Kotz matricial

$W_p(n, \Sigma)$  – Distribución de Wishart

$GT^2(p, v_h, v_e)$  – Distribución  $T^2$  de Hotelling generalizada

## Anexo B

# Algebra matricial

En este anexo se recogen diferentes definiciones y resultados correspondientes, fundamentalmente, al álgebra matricial. Para demostraciones y otros detalles véase, por ejemplo, Muirhead (1982), Searle (1982), Seber (1984), Bilodeau y Brenner (1999), Rencher (2002), Timm (2002), Abadir y Magnus (2005), Kollo y von Rosen (2005) y Seber (2008).

**Lema B.0.1** *Dadas las matrices  $\mathbf{A}_{(m \times n)}$  y  $\mathbf{B}_{(n \times m)}$  entonces*

$$\det [\mathbf{I}_m + \mathbf{AB}] = \det [\mathbf{I}_n + \mathbf{BA}].$$

**Lema B.0.2** *Para las matrices  $\mathbf{A}_{(m \times n)}$  y  $\mathbf{B}_{(n \times m)}$ , los autovalores no nulos de  $\mathbf{AB}$  y  $\mathbf{BA}$  son idénticos.*

**Lema B.0.3** *Si  $\mathbf{B} = \mathbf{CAC}^{-1}$ , donde  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  son matrices  $(m \times m)$ , entonces  $\mathbf{A}$  y  $\mathbf{B}$  tienen los mismos autovalores.*

**Lema B.0.4** *Sean  $\mathbf{A}_{(p \times p)}$  y  $\mathbf{B}_{(q \times q)}$  matrices no singulares, considerando las matrices  $\mathbf{C}_{(p \times q)}$  y  $\mathbf{D}_{(q \times p)}$ , entonces:*

(i)

$$(\mathbf{A} + \mathbf{CBD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{CB}(\mathbf{B} + \mathbf{BDA}^{-1}\mathbf{CB})^{-1}\mathbf{BDA}^{-1}. \quad (\text{B.0.1})$$

(ii) *Si  $\mathbf{B} = \mathbf{I}_q$  en (i), resulta*

$$(\mathbf{A} + \mathbf{CD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{I}_q + \mathbf{DA}^{-1}\mathbf{C})^{-1}\mathbf{DA}^{-1}. \quad (\text{B.0.2})$$

(ii) *Cuando,  $\mathbf{A} = \mathbf{I}_p$  y  $\mathbf{B} = \mathbf{I}_q$ , en (i), se obtiene*

$$(\mathbf{I}_p + \mathbf{CD})^{-1} = \mathbf{I}_p - \mathbf{C}(\mathbf{I}_q + \mathbf{DC})^{-1}\mathbf{D}.$$

**Lema B.0.5** *Sea*

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

donde  $\mathbf{Q}$  puede ser rectangular. Cuando  $\mathbf{A}$  es cuadrada y no singular, la matriz  $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$  se denomina complemento Schur de  $\mathbf{A}$  en  $\mathbf{Q}$ , además

$$(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})^{-1}\mathbf{BD}^{-1}. \quad (\text{B.0.3})$$

**Teorema B.0.1** *Sean  $\mathcal{C}(\mathbf{A})$  y  $\mathcal{C}(\mathbf{B})$  los espacios columna de  $\mathbf{A}$  y  $\mathbf{B}$ , respectivamente. Entonces  $\mathcal{C}(\mathbf{A}) \subseteq \mathcal{C}(\mathbf{B})$  si y solo si  $\mathbf{A} = \mathbf{BQ}$  para alguna matriz  $\mathbf{Q}$ .*

**Teorema B.0.2** *Sea  $\mathbf{A}_{(p \times q)}$  una matriz arbitraria de rango total, la matriz  $\mathbf{P}_\mathbf{A} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ , asociada a la proyección ortogonal sobre el espacio  $\mathcal{C}(\mathbf{A})$ , es simétrica, idempotente y tiene rango igual al rango de  $\mathbf{A}$ .*

**Nota B.0.1** *Dada  $\mathbf{A}_{(p \times q)}$  una matriz arbitraria con  $\text{rg}(\mathbf{A}) = q < p$ , existen infinitas matrices  $\mathbf{A}_{(p \times (p-q))}^*$  tales que*

$$\mathcal{C}(\mathbf{A}^*) = \mathcal{C}(\mathbf{A})^\perp,$$

el espacio matricial ortogonal a  $\mathcal{C}(\mathbf{A})$ . En particular, la matriz  $\mathbf{I} - \mathbf{P}_\mathbf{A}$ , de rango  $p - q$ , puede representarse por

$$\mathbf{I} - \mathbf{P}_\mathbf{A} = \begin{bmatrix} \mathbf{A}^\circ : \mathbf{Q} \end{bmatrix},$$

donde  $\mathbf{Q}_{((p-q) \times q)}$  es una matriz arbitraria y  $\mathbf{A}_{(p \times (p-q))}^\circ$  es tal que

$$\mathcal{C}(\mathbf{A}^\circ) = \mathcal{C}(\mathbf{A})^\perp = \{ \mathbf{A}^\circ : \mathbf{A}^\circ(p \times (p-q)), r(\mathbf{A}^\circ) = p - q, \mathbf{A}^T\mathbf{A}^\circ = \mathbf{0} \}, \quad (\text{B.0.4})$$

o sea,  $\mathbf{A}^\circ$  es una matriz cuyas columnas generan  $\mathcal{C}(\mathbf{A})^\perp$ . Es decir, la matriz  $\mathbf{A}^\circ$  puede ser obtenida con  $p - q$  columnas linealmente independientes de  $\mathbf{I} - \mathbf{P}_\mathbf{A}$ ,

$$\mathcal{C}(\mathbf{I} - \mathbf{P}_\mathbf{A}) = \mathcal{C}(\mathbf{A}^\circ) = \mathcal{C}(\mathbf{A})^\perp.$$

**Lema B.0.6** *Sea  $\mathbf{S}_{(p \times p)} > \mathbf{0}$  y  $\mathbf{A}_{(p \times q)}$  una matriz arbitraria con  $\text{rg}(\mathbf{A}) = q$ , entonces*

$$\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{A}(\mathbf{A}^T\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{S}^{-1} + \mathbf{A}^\circ(\mathbf{A}^{\circ T}\mathbf{S}\mathbf{A}^\circ)^{-1}\mathbf{A}^{\circ T} \quad (\text{B.0.5})$$

donde  $\mathcal{C}(\mathbf{A}^\circ) = \mathcal{C}(\mathbf{A})^\perp$ .

Del lema anterior, se puede concluir que, mientras (B.0.5) depende de la matriz  $\mathbf{A}_{(p \times (p-q))}^\circ$  que satisface  $\mathbf{A}^T\mathbf{A}^\circ = \mathbf{0}$ , ésta no depende de la elección específica de  $\mathbf{A}^\circ$  en  $\mathcal{C}(\mathbf{A})^\perp$ , una vez que  $\mathbf{A}^\circ(\mathbf{A}^{\circ T}\mathbf{S}\mathbf{A}^\circ)^{-1}\mathbf{A}^{\circ T}$  no depende de una elección específica de  $\mathbf{A}^\circ$ , por ser ésta arbitraria. Para más detalles véase, por ejemplo, Searle (1982).

**Nota B.0.2** *De (B.0.5), resulta  $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T + \mathbf{A}^\circ(\mathbf{A}^{\circ T}\mathbf{A}^\circ)^{-1}\mathbf{A}^{\circ T} = \mathbf{I}_p$ , es decir,  $\mathbf{P}_\mathbf{A} + \mathbf{P}_{\mathbf{A}^\circ} = \mathbf{I}_p$ .*



**Definición B.0.1** Sea  $\mathbf{A}_{(m \times n)} = ((a_{ij}))$ , se representa por  $\text{vec}(\mathbf{A})$  el vector

$$\text{vec}(\mathbf{A}) = (a_{11}, \dots, a_{n1}, a_{12}, \dots, a_{n2}, a_{m1}, \dots, a_{mn})^T.$$

**Definición B.0.2** Sea  $\mathbf{A}_{(m \times n)}$ , se llama matriz de conmutación a la matriz única  $\mathbf{K}_{mn}$  de orden  $(mn \times mn)$  tal que

$$\mathbf{K}_{mn} \text{vec}(\mathbf{A}) = \text{vec}(\mathbf{A}^T),$$

$$\text{donde } \mathbf{K}_{mn} = \sum_{i=1}^m \sum_{j=1}^n \mathbf{E}_{ij}(m, n) \otimes \mathbf{E}_{ij}^T(m, n).$$

**Definición B.0.3** Sea  $\mathbf{A}_{(n \times n)} = ((a_{ij}))_{1 \leq i, j \leq n}$  una matriz simétrica, se representa por  $\text{svec}(\mathbf{A})$  el vector

$$\text{svec}(\mathbf{A}) = (a_{11}, \dots, a_{n1}, a_{22}, \dots, a_{n2}, \dots, a_{nn})^T \in \mathbb{R}^{n^*},$$

donde  $n^* = \frac{n(n+1)}{2}$ . El vector  $\text{svec}(\mathbf{A})$  es también denotado por  $\text{vech}(\mathbf{A})$ .

**Teorema B.0.3** Sea  $\mathbf{A} = ((a_{ij}))$  una matriz de orden  $(p \times q)$ , entonces

$$\text{tr}[\mathbf{A}^T \mathbf{A}] = \text{tr}[\mathbf{A} \mathbf{A}^T] = \sum_{i=1}^p \sum_{j=1}^q a_{ij}^2.$$

Además,  $\text{tr}[\mathbf{A}^T \mathbf{A}] = 0$  si y solo si  $\mathbf{A} = \mathbf{0}$ .

**Definición B.0.4** La norma de Frobenius (o euclídea),  $\|\mathbf{A}\|$ , de una matriz  $\mathbf{A} = ((a_{ij}))$  de orden  $(p \times q)$ , se define por

$$\|\mathbf{A}\| = \sqrt{\sum_{i=1}^p \sum_{j=1}^q a_{ij}^2}. \quad (\text{B.0.6})$$

**Teorema B.0.4** Sea  $\mathbf{A}$  una matriz de orden  $(p \times q)$  y  $\mathbf{a} = \text{vec}(\mathbf{A})$ , entonces  $\|\mathbf{A}\| = \|\mathbf{a}\|$ .

**Definición B.0.5** Sean  $\mathbf{A}_{(n \times m)}$  y  $\mathbf{B}_{(p \times q)}$ , se define el producto de Kronecker de  $\mathbf{A}$  por  $\mathbf{B}$  como la matriz por bloques de orden  $(np \times mq)$  tal que

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11} \mathbf{B} & \dots & a_{1m} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{n1} \mathbf{B} & \dots & a_{nm} \mathbf{B} \end{bmatrix}.$$

En general, para cualquiera matriz  $\mathbf{A}$ ,  $\mathbf{A}^{\otimes k} = \underbrace{\mathbf{A} \otimes \mathbf{A} \otimes \dots \otimes \mathbf{A}}_{k \text{ veces}} y \mathbf{I}^{\otimes 0} = \mathbf{I}$ .

**Teorema B.0.5** *Son válidas las siguientes propiedades para el producto de Kronecker:*

- (i)  $\alpha \otimes \mathbf{A} = \alpha \mathbf{A} = \mathbf{A} \otimes \alpha, \forall \alpha \in \mathbb{R}.$
- (ii)  $(\alpha \mathbf{A}) \otimes (\beta \mathbf{B}) = \alpha\beta (\mathbf{A} \otimes \mathbf{B}), \forall \alpha, \beta \in \mathbb{R}.$
- (iii)  $(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T.$
- (iv)  $\text{rg}(\mathbf{A} \otimes \mathbf{B}) = \text{rg}(\mathbf{A}) \text{rg}(\mathbf{B})$
- (v) Sea  $\mathbf{A}_{(m \times n)}$  y  $\mathbf{B}_{(p \times q)}$ , entonces  $\mathbf{K}_{pm}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{qn}$  y  $\mathbf{K}_{pm}(\mathbf{A} \otimes \mathbf{B})\mathbf{K}_{nq} = \mathbf{B} \otimes \mathbf{A}.$
- (vi) Si  $\mathbf{A}$  y  $\mathbf{B}$  son no singulares, entonces también  $\mathbf{A} \otimes \mathbf{B}$  es no singular. Además,  $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}.$
- (vii) Sean  $\mathbf{A}_{(m \times n)}$ ,  $\mathbf{B}_{(p \times q)}$ ,  $\mathbf{C}_{(n \times r)}$  y  $\mathbf{D}_{(q \times s)}$ , entonces  $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}.$
- (viii) Sean  $\mathbf{A}_{(n \times n)}$  y  $\mathbf{B}_{(q \times q)}$  entonces:
  - a)  $\text{tr}[\mathbf{A} \otimes \mathbf{B}] = \text{tr}[\mathbf{A}] \text{tr}[\mathbf{B}].$
  - b)  $\det[\mathbf{A} \otimes \mathbf{B}] = (\det[\mathbf{A}])^q (\det[\mathbf{B}])^n.$
- (ix) Si  $\mathbf{A} > \mathbf{0}$  y  $\mathbf{B} > \mathbf{0} \Rightarrow \exists \mathbf{A}^{1/2}$  y  $\mathbf{B}^{1/2}$  y  $\mathbf{A} \otimes \mathbf{B} = (\mathbf{A}^{1/2} \otimes \mathbf{B}^{1/2})(\mathbf{A}^{1/2} \otimes \mathbf{B}^{1/2})$ , o sea,  $(\mathbf{A} \otimes \mathbf{B})^{1/2} = \mathbf{A}^{1/2} \otimes \mathbf{B}^{1/2}.$
- (x) Si  $\mathbf{A} > \mathbf{0}$  y  $\mathbf{B} > \mathbf{0}$  entonces  $\mathbf{A} \otimes \mathbf{B} > \mathbf{0}.$
- (xi)  $\|\mathbf{A} \otimes \mathbf{B}\| = \|\mathbf{A}\| \|\mathbf{B}\|$ , donde  $\|\cdot\|$  es la norma de Frobenius.
- (xii) Sean las matrices  $\mathbf{A}_{(m \times m)}$  y  $\mathbf{B}_{(n \times n)}$  con autovalores  $\lambda_1, \dots, \lambda_m$ , y  $\mu_1, \dots, \mu_n$ , respectivamente. Los  $mn$  autovalores de  $\mathbf{A} \otimes \mathbf{B}$  son  $\lambda_i \mu_j$  ( $i = 1, \dots, m; j = 1, \dots, n$ ).

**Lema B.0.7** *Suponiendo que las siguientes matrices son de dimensiones apropiadas, entonces:*

- (i)  $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}).$
- (ii)  $\text{tr}[\mathbf{ABCD}] = \text{vec}^T(\mathbf{D})(\mathbf{A} \otimes \mathbf{C}^T) \text{vec}(\mathbf{B}^T).$

**Definición B.0.6** *Sean  $V$  y  $W$  dos espacios vectoriales de dimensiones  $m$  y  $n$  respectivamente. Sean  $\{v_1, \dots, v_m\}$  y  $\{w_1, \dots, w_n\}$  bases de los espacios  $V$  y  $W$ , respectivamente. Se denomina producto tensorial de  $V$  y  $W$ , y se nota por  $V \otimes W$ , al espacio vectorial generado por los vectores  $\{v_i \otimes w_j : i = 1, \dots, m; j = 1, \dots, n\}$ , donde  $v_i \otimes w_j$  representa el vector  $mn$ -dimensional producto de Kronecker de los vectores  $v_i$  y  $w_j$ .*

**Definición B.0.7** *Sean  $V_1$  y  $V_2$  dos subespacios de un espacio vectorial  $V$ , ortogonales entre sí ( $V_1 \subseteq V_2^\perp$  ó  $V_2 \subseteq V_1^\perp$ ). Se define la suma ortogonal de  $V_1$  y  $V_2$ , y se denota por  $V_1 \boxplus V_2$ , al subespacio vectorial definido por*

$$V_1 \boxplus V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}.$$

## Anexo C

# Distribuciones multivariantes

En este anexo se recogen algunos resultados básicos de distribuciones multivariantes útiles en el desarrollo de esta memoria. Las demostraciones de estos resultados, así como un estudio más detallado, están recogidas en Anderson y Fang (1990), Fang et al. (1990), Gupta y Varga (1993), Gupta y Varga (1994), Gupta y Nagar (2000) y Kollo y von Rosen (2005).

### C.1. Distribución elíptica

La familia de distribuciones elípticas puede ser considerada como una extensión de la normal multivariante, de gran interés práctico en tanto que conserva muchas de las propiedades de esta última, tales como la simetría o constituir una familia cerrada ante transformaciones lineales. Por otra parte, muchos de los resultados y propiedades de la distribución normal multivariante pueden ser fácilmente trasladados, de forma directa o con leves modificaciones, a la familia elíptica. Así, además del interés que tiene para la modelización de poblaciones multivariantes, representa una herramienta relevante en técnicas estadísticas robustas en el análisis multivariante.

La familia de distribuciones elípticas es una clase general de distribuciones cuyos contornos de densidad constante tienen forma elíptica, como la familia normal, con colas más o menos pesadas.

Existen diversas formas de introducir la clase de distribuciones matriciales elípticas, con el objetivo de describir una matriz aleatoria  $\mathbf{U}_{(p \times n)}$ , cuyas columnas pueden ser consideradas como observaciones de una población  $p$ -dimensional.

**Definición C.1.1** Una matriz aleatoria  $\mathbf{U}_{(p \times n)}$  se dice que tiene distribución matricial elíptica si admite la función característica  $\varphi_{\mathbf{U}}(\mathbf{T}) = \exp(\text{tr}[i\mathbf{T}^T \mathbf{M}])\phi(\text{tr}[\mathbf{T}^T \mathbf{V} \mathbf{T} \mathbf{W}])$ , con  $\mathbf{T}_{(p \times n)}$ ,  $\mathbf{M}_{(p \times p)}$ ,  $\mathbf{V}_{(p \times p)} \geq \mathbf{0}$ ,  $\mathbf{W}_{(n \times n)} \geq \mathbf{0}$  y  $\phi : [0, \infty) \rightarrow \mathbb{R}$  la función característica generadora. Se denota por  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$ .

Si  $n = 1$  se dice que  $\mathbf{U}$  sigue una distribución elíptica multivariante, y se representa por  $E_p(\mathbf{m}, \mathbf{V}, \phi)$ , siendo su función característica  $\varphi_{\mathbf{U}}(\mathbf{t}) = \exp(it^T \mathbf{m})\phi(\mathbf{t}^T \mathbf{V} \mathbf{t})$ .

El siguiente teorema establece la relación entre la distribución elíptica matricial y la distribución elíptica multivariante. Permite generalizar los resultados de las distribuciones elípticas multivariantes al caso matricial. Más precisamente, si  $\mathbf{U}_{(p \times n)}$  tiene una distribución matricial elíptica entonces es posible aplicar los resultados conocidos de los vectores a  $\text{vec}(\mathbf{U})$ .

**Teorema C.1.1** Sea  $\mathbf{U}_{(p \times n)}$  una matriz aleatoria, entonces  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$  si y solo si  $\text{vec}(\mathbf{U}) \sim E_{pn}(\text{vec}(\mathbf{M}), \mathbf{W} \otimes \mathbf{V}, \phi)$ .

**Nota C.1.1** Las matrices  $\mathbf{M}_{(p \times n)}$ ,  $\mathbf{V}_{(p \times p)}$  y  $\mathbf{W}_{(n \times n)}$  son los parámetros de la distribución, donde  $\mathbf{M} \in \mathbb{R}^{p \times n}$  puede ser considerado el parámetro matricial de localización y  $\mathbf{W} \otimes \mathbf{V} \in \mathbb{R}^{pn \times pn}$  el parámetro matricial de escala.

**Teorema C.1.2** Sea  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$  y sean  $\mathbf{A}_{(q \times p)}$ ,  $\mathbf{B}_{(m \times n)}$  y  $\mathbf{C}_{(q \times m)}$  matrices constantes, entonces

$$\mathbf{A}\mathbf{U}\mathbf{B}^T + \mathbf{C} \sim E_{q,m}(\mathbf{A}\mathbf{M}\mathbf{B}^T + \mathbf{C}; \mathbf{A}\mathbf{V}\mathbf{A}^T, \mathbf{B}\mathbf{W}\mathbf{B}^T, \phi).$$

Para distribuciones elípticas, la existencia de la función densidad de probabilidad (f.d.p.) no está siempre garantizada. No obstante, la existencia de la f.d.p. está garantizada si  $\mathbf{U}_{(p \times n)}$  es una matriz aleatoria cuya distribución conjunta es absolutamente continua (Gupta y Varga, 1994).

**Teorema C.1.3** Sea  $\mathbf{U}_{(p \times n)}$  una matriz aleatoria cuya distribución conjunta es absolutamente continua. Entonces  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$  si y solo si la f.d.p. de  $\mathbf{U}$  puede ser expresada por

$$f_{\mathbf{U}}(\mathbf{U}) = C_{np} |\mathbf{V}|^{-\frac{n}{2}} |\mathbf{W}|^{-\frac{p}{2}} g \left( \text{tr} \left[ \mathbf{V}^{-1} (\mathbf{U} - \mathbf{M}) \mathbf{W}^{-1} (\mathbf{U} - \mathbf{M})^T \right] \right),$$

donde la función  $g : \mathbb{R} \rightarrow [0, \infty)$ , denominada función generadora de densidad, es tal que

$$\int_0^{\infty} u^{np/2-1} g(u) du < \infty$$

y  $C_{np}$  es la constante de normalización

$$C_{np} = \frac{\Gamma(np/2)}{2\pi^{np/2} \int_0^{\infty} u^{np-1} g(u^2) du}, \quad (\text{C.1.1})$$

con  $\Gamma : [0, \infty[ \rightarrow \mathbb{R}$  la función gamma. En este caso, se denota  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, g)$ .

**Lema C.1.1** Toda función  $g : \mathbb{R} \rightarrow [0, \infty)$  tal que

$$\int_0^{\infty} u^{np/2-1} g(u) du < \infty$$

es función generadora de una distribución elíptica, con constante de normalización  $C_{np}$  dada por (C.1.1).

**Teorema C.1.4** Sea  $\mathbf{U}_{(p \times n)}$  una matriz aleatoria y  $\mathbf{M}_{(p \times n)}$ ,  $\mathbf{V}_{(p \times p)} \geq \mathbf{0}$ ,  $\mathbf{W}_{(n \times n)} \geq \mathbf{0}$  matrices constantes con  $\text{rg}(\mathbf{V}) = q$  y  $\text{rg}(\mathbf{W}) = m$ . Entonces  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$  si y solo si se puede representar estocásticamente ( $\stackrel{d}{=}$ ) por

$$\mathbf{U} \stackrel{d}{=} \mathbf{M} + R\boldsymbol{\tau}\boldsymbol{\Delta}\boldsymbol{\gamma}^T, \quad (\text{C.1.2})$$

donde  $R$  es una variable aleatoria no negativa independiente de  $\boldsymbol{\Delta}_{(q \times m)}$ ,  $\text{vec}(\boldsymbol{\Delta}^T)$  está uniformemente distribuida en la esfera unidad de  $\mathbb{R}^{qm}$ ,  $\mathbf{V} = \boldsymbol{\tau}\boldsymbol{\tau}^T$  con  $\boldsymbol{\tau}$  una matriz arbitraria de dimensiones  $(p \times q)$  y  $\mathbf{W} = \boldsymbol{\gamma}\boldsymbol{\gamma}^T$  con  $\boldsymbol{\gamma}$  una matriz arbitraria de dimensiones  $(n \times m)$ .

**Nota C.1.2** Si  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, g)$  entonces la f.d.p. de la variable aleatoria  $R$  incluida en su representación estocástica (C.1.2), es

$$f_R(r) = C_{np} \frac{2\pi^{np/2}}{\Gamma(np/2)} r^{np-1} g(r^2), \quad r > 0.$$

**Teorema C.1.5** Sea  $\mathbf{U} \sim E_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W}, \phi)$ .

(i) Si  $\mathbf{U}$  tiene momentos de primera orden finitos, entonces  $E[\mathbf{U}] = \mathbf{M}$ .

(ii) Si  $\mathbf{U}$  tiene momentos de segunda orden finitos, entonces

$$\text{Var}[\mathbf{U}] = \text{Var}[\text{vec}(\mathbf{U})] = -2\phi'(0)(\mathbf{W} \otimes \mathbf{V}) = \frac{E[R^2]}{np}(\mathbf{W} \otimes \mathbf{V}).$$

**Teorema C.1.6** Sea  $X$  una mezcla de  $k$  distribuciones  $E_p(\mu, \mathbf{V}, g_j)$  con función de densidad de probabilidad  $f_j$ , entonces la f.d.p. de  $X$  se puede expresar por

$$f_X(x) = |\mathbf{V}|^{-1/2} \sum_{j=1}^k \pi_j C_j g_j [(x - \mu)^T \mathbf{V}^{-1}(x - \mu)]$$

donde  $C_j$  son las constantes de normalización de cada distribución componente.

## C.2. Distribución normal matricial

La distribución normal matricial generaliza la distribución normal multivariante en un contexto matricial sus propiedades se exponen ampliamente en Kollo y von Rosen (2005). La distribución normal matricial puede ser definida por su función densidad de probabilidad.

**Definición C.2.1** Se dice que la matriz aleatoria  $\mathbf{U}_{(p \times n)}$  sigue una distribución normal matricial cuando admite la siguiente función de densidad

$$f_U(\mathbf{U}) = (2\pi)^{-\frac{np}{2}} |\mathbf{V}|^{-\frac{n}{2}} |\mathbf{W}|^{-\frac{p}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\mathbf{V}^{-1}(\mathbf{U} - \mathbf{M})\mathbf{W}^{-1}(\mathbf{U} - \mathbf{M})^T\right]\right), \quad (\text{C.2.1})$$

donde  $\mathbf{M}_{(p \times n)}$ ,  $\mathbf{V}_{(p \times p)} > \mathbf{0}$  y  $\mathbf{W}_{(n \times n)} > \mathbf{0}$ . En este caso, se denota  $\mathbf{U} \sim N_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W})$ .

**Teorema C.2.1** La distribución normal matricial es un caso particular de la distribución elíptica matricial con función característica generadora

$$\phi(u) = \exp\left(-\frac{u}{2}\right),$$

función generadora de densidad

$$g(u) = \exp\left(-\frac{u}{2}\right)$$

y constante de normalización

$$C_{np} = (2\pi)^{-\frac{np}{2}}.$$

**Corolario C.2.1** Sea  $\mathbf{U}_{(p \times n)}$  una matriz aleatoria con distribución normal matricial con parámetros  $\mathbf{M}$ ,  $\mathbf{V}$  y  $\mathbf{W}$ ,  $\mathbf{U} \sim N_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W})$ . Entonces:

- (i)  $E[\mathbf{U}] = \mathbf{M}$ .
- (ii)  $\text{Var}[\mathbf{U}] = \mathbf{W} \otimes \mathbf{V}$ .

**Nota C.2.1** Si  $\mathbf{m}_i$  representa la  $i$ -ésima columna de  $\mathbf{M}$ , entonces si  $\mathbf{W} = \mathbf{I}_n$ , las columnas de  $\mathbf{U}$  están independientemente distribuidas según  $N_p(\mathbf{m}_i, \mathbf{V})$ , para  $i = 1, \dots, n$ . Además, si  $\mathbf{V} = \mathbf{I}_p$  todos los elementos de  $\mathbf{U}$  son mutuamente independientes.

### C.3. Distribución de Kotz matricial

La familia de distribuciones tipo-Kotz fue introducida por Kotz (1975) como una generalización de la distribución normal multivariante. Nadarajah (2003) recoge una revisión de esta familia de distribuciones, con referencias a diferentes aplicaciones. Véase Kollo y Roos (2005) para una aplicación al MCC.

**Definición C.3.1** Se dice que una matriz aleatoria  $\mathbf{U}_{(p \times n)}$  sigue la distribución de Kotz matricial con parámetros  $\mathbf{M}$ ,  $\mathbf{V}$  y  $\mathbf{W}$ , y se denota por  $\mathbf{U} \sim K_{p,n}(\mathbf{M}; \mathbf{V}, \mathbf{W})$ , si su función densidad de probabilidad es

$$f_{\mathbf{U}}(\mathbf{U}) = \frac{1}{np} (2\pi)^{-\frac{np}{2}} |\mathbf{V}|^{-\frac{p}{2}} |\mathbf{W}|^{-\frac{n}{2}} g \left( \text{tr} \left[ \mathbf{V}^{-1} (\mathbf{U} - \mathbf{M}) \mathbf{W}^{-1} (\mathbf{U} - \mathbf{M})^T \right] \right),$$

donde  $g(u) = u \exp(-u/2)$ ,  $\mathbf{M}_{(p \times n)}$ ,  $\mathbf{V}_{(p \times p)} > \mathbf{0}$  y  $\mathbf{W}_{(n \times n)} > \mathbf{0}$ .

**Teorema C.3.1** La distribución de Kotz matricial es un caso particular de la distribución matricial elíptica con función característica generadora

$$\phi(u) = \left( 1 - \frac{u}{np} \right) \exp \left( -\frac{u}{2} \right),$$

función generadora de densidad

$$g(u) = u \exp \left( -\frac{u}{2} \right)$$

y constante de normalización

$$C_{np} = \frac{1}{np} (2\pi)^{-\frac{np}{2}}.$$

**Corolario C.3.1** Sea  $\mathbf{U}_{(p \times n)}$  una matriz aleatoria con distribución de Kotz matricial con parámetros  $\mathbf{M}$ ,  $\mathbf{V}$  y  $\mathbf{W}$ . Entonces:

- (i)  $E[\mathbf{U}] = \mathbf{M}$ .
- (ii)  $\text{Var}[\mathbf{U}] = \left( 1 + \frac{2}{np} \right) (\mathbf{W} \otimes \mathbf{V})$ .

## C.4. Distribución de Wishart

La distribución matricial conocida por distribución de Wishart (Wishart, 1928) está asociada a la distribución normal matricial, generaliza la distribución  $\chi^2$  y juega un papel importante en inferencia multivariante, en particular, en el análisis de covarianza y en la detección de outliers y observaciones influyentes.

**Definición C.4.1** Sea  $\mathbf{U}_{(p \times n)} \sim N_{p,n}(\mathbf{0}; \mathbf{\Sigma}, \mathbf{I}_n)$ , entonces la matriz  $\mathbf{W}_{(p \times p)} = \mathbf{U}\mathbf{U}^T$  tiene distribución de Wishart centrada con  $n$  grados de libertad y matriz asociada  $\mathbf{\Sigma}_{(p \times p)} > \mathbf{0}$ , denotándose por  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$ .

**Teorema C.4.1** Sea  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$  con  $n \geq p$ , la función densidad de probabilidad de  $\mathbf{W}$  es

$$f_{\mathbf{W}}(\mathbf{W}) = 2^{-pn/2} \left[ \Gamma_p \left( \frac{n}{2} \right) \right]^{-1} |\mathbf{\Sigma}|^{-n/2} |\mathbf{W}|^{(n-p-1)/2} \exp \left( -\frac{1}{2} \text{tr} [\mathbf{\Sigma}^{-1} \mathbf{W}] \right), \quad \mathbf{W} > \mathbf{0},$$

donde  $\Gamma_p(\cdot)$  denota la función gamma multivariante

$$\Gamma_p(a) = \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma \left( a - \frac{1}{2}(i-1) \right)$$

y  $\Gamma(\cdot)$  es la función gamma.

**Lema C.4.1** Sea  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$  entonces:

- (i)  $E[\mathbf{W}] = n\mathbf{\Sigma}$ .
- (ii)  $\text{Var}[\mathbf{W}] = n(\mathbf{I}_{p^2} + \mathbf{K}_{p^2})(\mathbf{\Sigma} \otimes \mathbf{\Sigma})$ , donde  $\mathbf{K}_{p^2}$  es la matriz de conmutación.

A continuación, se presentan algunas propiedades sobre la distribución de Wishart, para más detalles véase, por ejemplo, Kollo y von Rosen (2005).

**Teorema C.4.2** Se verifican las siguientes propiedades para la distribución de Wishart:

- (i) Si  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$ , con  $n \geq p$  y  $\mathbf{\Sigma} > \mathbf{0}$ , entonces  $P(\mathbf{W} > \mathbf{0}) = 1$ , es decir, con probabilidad 1 tiene sentido hablar de  $\mathbf{W}^{-1}$ .
- (ii) Suponiendo que  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$  y que  $\mathbf{C}_{(q \times p)}$  con  $r(\mathbf{C}) = q$  entonces la matriz no singular  $\mathbf{C}\mathbf{W}\mathbf{C}^T \sim W_q(n, \mathbf{C}\mathbf{\Sigma}\mathbf{C}^T)$ .
- (iii) Suponiendo que  $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$  y que  $\mathbf{C}_{(p \times q)}$  con  $r(\mathbf{C}) = q$ . Entonces  $(\mathbf{C}^T \mathbf{W}^{-1} \mathbf{C})^{-1} \sim W_q(n - p + q, (\mathbf{C}^T \mathbf{\Sigma}^{-1} \mathbf{C})^{-1})$ .
- (iv) Sea  $\mathbf{X}_{(p \times n)} \sim N_{p,n}(\mathbf{M}; \mathbf{\Sigma}, \mathbf{I}_n)$  y  $\mathbf{P}_{(n \times n)}$  una matriz simétrica. Entonces  $\mathbf{X}\mathbf{P}\mathbf{X}^T$  tiene distribución de Wishart si y solo si  $\mathbf{P}$  es una matriz idempotente.
- (v) Sea  $\mathbf{X}_{(p \times n)} \sim N_{p,n}(\mathbf{M}; \mathbf{\Sigma}, \mathbf{I}_n)$  y  $\mathbf{P}_{(n \times n)}$  una matriz simétrica idempotente, tal que,  $\mathbf{M}\mathbf{P} = \mathbf{0}$ . Entonces  $\mathbf{X}\mathbf{P}\mathbf{X}^T \sim W_p(\text{rg}(\mathbf{P}), \mathbf{\Sigma})$ .
- (vi) Las variables Wishart  $\mathbf{X}^T \mathbf{A}\mathbf{X}$  y  $\mathbf{X}^T \mathbf{B}\mathbf{X}$  son independientes si y solo si  $\mathbf{A}\mathbf{B} = \mathbf{0}$ .
- (vii) Es reproductiva con respecto a los grados de libertad.

### C.5. Distribución $T^2$ de Hotelling generalizada

Las distribuciones basadas en la traza y los autovalores de matrices aleatorias definidas positivas son muy utilizadas en modelos multivariantes. Un caso particular de estas distribuciones es la distribución  $T^2$  de Hotelling generalizada.

**Definición C.5.1** Sean  $\mathbf{H} \sim W_p(v_h, \Sigma)$  y  $\mathbf{E} \sim W_p(v_e, \Sigma)$  independientes con  $v_e \geq p$ , entonces se define el estadístico traza de Lawley-Hotelling o  $T_g^2$  de Hotelling generalizado por

$$T_g^2 = v_e \text{tr} [\mathbf{H} \mathbf{E}^{-1}] = v_e \text{tr} [\mathbf{E}^{-1/2} \mathbf{H} \mathbf{E}^{-1/2}].$$

La distribución de  $T_g^2$  se denota por  $GT^2(p, v_h, v_e)$  y se denomina distribución  $T^2$  de Hotelling generalizada o distribución de Lawley-Hotelling con parámetros  $p, v_h$  y  $v_e$ .