

UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA
INSTITUTO DE MATEMÁTICA APLICADA DEL LITORAL

**Estimación Robusta para el Modelo Lineal
Multivariado basado en τ -escala, con Rango Completo
y Rango Reducido**

Tesis presentada en el Departamento de Matemáticas de la Universidad Nacional del Litoral
como requisito para optar del título de Doctor en Matemática.

Lcda. Ananda Paola Rivas Marquina

Directora: Dra. Liliana Forzani

Codirectora: Dra. Pamela Llop

Jurados:

Adrover, Jorge

Bergesio, Andrea

Rodriguez, Daniela

Valdora, Marina

Szretter, Maria Eugenia

Departamento de Matemática

2024

Índice general

1	DEFINICIÓN DEL τ-FUNCIONAL Y PROPIEDADES	11
1.1	Motivación del τ -estimador para el MLM	11
1.2	Definición del τ -funcional de escala	12
1.3	τ -funcional de escala basado en una muestra	15
1.4	Existencia del τ -funcional de estimación para el MLM	16
1.5	Ecuaciones que definen el τ -estimador basado en una muestra	17
1.6	Unicidad del τ -funcional de estimación para el MLM y consistencia del mismo basado en una muestra	18
2	DISTRIBUCIÓN ASINTÓTICA (CASO: RANGO COMPLETO)	20
2.1	El sistema de ecuaciones (1.22) en forma matricial	20
2.1.1	Consistencia de $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \mathbf{\Sigma}_n)$	21
2.1.2	Distribución asintótica de $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \mathbf{\Sigma}_0^*, s_n(\mathbf{B}_0, \mathbf{\Sigma}_0^*))$	22
2.1.3	Convergencia de derivadas	24
2.2	Normalidad asintótica del τ -estimador (rango completo)	25
3	DISTRIBUCIÓN ASINTÓTICA (CASO: RANGO REDUCIDO)	29
3.1	Preliminares	29
3.2	Definición, existencia y unicidad del τ -funcional de estimación para el MLMRR en los contextos poblacional y muestral	31
3.3	Consistencia del τ -estimador de $(\mathbf{Q}_0, \mathbf{R}_0, \mathbf{\Sigma}_0^*)$	34
3.4	Ecuaciones de τ -estimaciones	34
3.5	Normalidad asintótica del τ -estimador (rango reducido)	36
4	ESTUDIOS DE SIMULACIÓN	38
4.1	Métodos	38
4.2	Modelo Matricial	39
4.3	Medidas de error	39
4.4	Estudio de simulación 1	40
4.4.1	Resultados	41
4.5	Estudio de simulación 2	62
4.5.1	Resultados	63
5	APLICACIÓN A DATOS REALES	74
5.1	Medida de error	74
5.2	Datos bioquímicos	75
5.3	Factores determinantes de niveles de retinol y beta-caroteno en plasma	84
6	CONCLUSIONES	91

	Bibliografía	92
A	APÉNDICE DEL CAPÍTULO 1	95
A.1	Demostración del Lema 1.1.1	95
A.2	Demostración del Lema 1.2.1	96
A.3	Demostración del Lema 1.2.2	97
A.4	Demostración del Lema 1.2.3	98
A.5	Demostración del Teorema 2	99
B	APÉNDICE DEL CAPÍTULO 2	111
B.1	Demostración del Lema 2.1.1	111
B.2	Demostración del Teorema 5	112
B.3	Demostración del Lema 2.1.2	120
B.4	Demostración del Teorema 6	122
B.5	Demostración del Lema 2.1.3	130
B.6	Demostración del Teorema 8	132
C	APÉNDICE DEL CAPÍTULO 3	162
C.1	Demostración de la Observación 3.2.4	162
C.2	Demostración del Teorema 14	162
C.3	Demostración del Teorema 15	163
C.4	Demostración del Teorema 16	164
	Notación	168

Resumen

La mayoría de los procedimientos estadísticos clásicos para estimar los parámetros del modelo de regresión lineal multivariado están desarrollados para modelos cuyas hipótesis son rígidas, tales como errores normales y errores igualmente distribuidos, entre otras. Si bien bajo estas hipótesis se deducen procedimientos óptimos, estos métodos son muy sensibles al incumplimiento de las mismas, por ejemplo, ante la presencia en la muestra de observaciones atípicas.

Los procedimientos estadísticos robustos tienen como objetivo permitir inferencias válidas aún cuando el modelo no se cumple exactamente o hay una proporción de datos que no siguen el modelo. Estos métodos son capaces de detectar y manejar datos “anómalos” y se esfuerzan por ofrecer estimaciones eficientes. Es decir, buscan proporcionar estimaciones casi tan precisas como las clásicas cuando los datos siguen el modelo central y, en presencia de datos atípicos, comportarse de manera similar a los métodos clásicos aplicados a datos no contaminados.

En este contexto, esta tesis estudia estimadores robustos para los parámetros del modelo lineal multivariado cuando la matriz de coeficientes de regresión es de rango reducido. Se utilizan estimadores basados en τ -escala para permitir la estimación simultánea de los coeficientes de regresión y la matriz de covarianza de los errores. Es importante resaltar que las bases de los estimadores robustos de tipo τ -escala ya están definidas y derivan de trabajos previos, como los de (García Ben et al., 2006), (Lopuhaa, 1991) y (Bergesio et al., 2021). Aunque estos autores no se centran explícitamente en el modelo lineal multivariado con la matriz de coeficientes de regresión de rango reducido, sus hallazgos y contribuciones son fundamentales para el análisis realizado en este trabajo.

Los objetivos de esta tesis son proponer estimadores robustos de tipo τ -escala para estos parámetros, analizar su consistencia bajo condiciones generales y derivar ecuaciones de estimación que se asemejan a las de máxima verosimilitud. Además, se estudia el comportamiento asintótico de estos estimadores en dos casos: cuando el coeficiente de regresión tiene rango completo y cuando tiene rango reducido. Esta investigación representa un avance con respecto al estudio previo de (García Ben et al., 2006), ya que, mientras que el trabajo citado se centra exclusivamente en el comportamiento asintótico de la matriz de coeficientes de regresión, este estudio examina de manera conjunta tanto la matriz de coeficientes de regresión, en sus escenarios de rango completo y rango reducido, como la matriz de covarianza de los errores, la cual se considera desconocida.

En el contexto detallado anteriormente, se presentan simulaciones diseñadas para ilustrar las relaciones, distinciones y beneficios de la metodología propuesta en comparación con los métodos existentes. Las simulaciones se llevan a cabo bajo diversas condiciones, replicando escenarios típicos de aplicación que contemplan la presencia de datos atípicos y variaciones en la estructura subyacente de los datos. A través de estos estudios, se busca resaltar no solo la efectividad del τ -estimador, sino también su posicionamiento dentro del marco de

estimadores robustos y su relevancia en la estimación de parámetros en modelos lineales multivariados.

Finalmente, se realiza la validación de la nueva metodología de estimación utilizando datos reales, lo que permite comparar su desempeño con las técnicas existentes, resaltando así sus ventajas y contribuciones al análisis de datos.

Agradecimientos

En primer lugar, deseo expresar mi más sincero agradecimiento a mis profesoras Liliana Forzani y Pamela Llop por su dedicación y paciencia durante el desarrollo de este trabajo. Su orientación académica, así como su disposición para aclarar dudas y compartir su experiencia, han sido esenciales para la culminación de esta tesis.

Extiendo mi gratitud al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y a la Universidad Nacional del Litoral por el apoyo financiero e institucional brindado, que ha sido fundamental para llevar a cabo esta investigación como parte de mi formación doctoral.

También quiero agradecer especialmente a la profesora Andrea Bergesio, cuyo acompañamiento y asesoramiento resultaron imprescindibles para la correcta ejecución de este estudio.

A mi familia y amigos, por su cariño incondicional, su comprensión y por brindarme el ánimo necesario para continuar en los momentos más desafiantes.

Introducción

El objetivo de estudiar métodos de regresión es crear un modelo matemático que pueda explicar la dependencia entre una variable respuesta y una o más variables independientes. Teniendo en cuenta esto, una situación frecuente en el análisis estadístico es la del modelo de regresión lineal multivariado (MLM, por sus siglas en inglés) dado por

$$\mathbf{y} = \mathbf{B}_0^t \mathbf{x} + \mathbf{u}, \quad (1)$$

donde $\mathbf{y} \in \mathbb{R}^q$, $\mathbf{x} \in \mathbb{R}^p$ y $\mathbf{B}_0 \in \mathbb{R}^{p \times q}$, con $\text{rango}(\mathbf{B}_0) = d \leq \min\{p, q\}$. Se denominará M_0 a la distribución de \mathbf{x} y F_0 a la distribución de $\mathbf{u} \in \mathbb{R}^q$, que se asume independiente de \mathbf{x} , con una matriz de covarianza Σ_0 definida positiva y media $\mathbf{0}$. La distribución conjunta de (\mathbf{x}, \mathbf{y}) se denotará por P . Si el modelo incluye un intercepto, el último elemento del vector \mathbf{x} será igual a 1. A partir de ahora, se supondrá que el espacio Θ de los posibles parámetros para el MLM (1) es

$$\Theta = \{\boldsymbol{\theta} = (\mathbf{B}, \Sigma) \in \mathbb{R}^{p \times q} \times S_+^q\},$$

donde S_+^q denota la clase de todas las matrices simétricas definidas positivas de tamaño $q \times q$. El método más utilizado para la obtención de estimadores de los parámetros del modelo es el método de máxima verosimilitud (Knight, 1999). Cuando una muestra aleatoria $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n \in \mathbb{R}^p \times \mathbb{R}^q$ de pares que satisfacen el modelo (1) con $n \leq p + q$, tiene errores $\mathbf{u}_i \sim \mathcal{N}_q(\mathbf{0}, \Sigma_0)$, y el rango de \mathbf{B}_0 es completo, i.e. $\text{rango}(\mathbf{B}_0) = \min\{p, q\}$, los estimadores de máxima verosimilitud (MLE, por sus siglas en inglés) de \mathbf{B}_0 y Σ_0 están dados por,

$$\hat{\mathbf{B}} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y}, \quad (2)$$

y

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i(\hat{\mathbf{B}}) \mathbf{u}_i^t(\hat{\mathbf{B}}), \quad (3)$$

donde \mathbf{X} es la matriz $n \times p$ cuya i -ésima fila es \mathbf{x}_i^t , \mathbf{Y} es la matriz $n \times q$ cuya i -ésima fila es \mathbf{y}_i^t y $\mathbf{u}_i(\hat{\mathbf{B}}) = \mathbf{y}_i - \hat{\mathbf{B}}^t \mathbf{x}_i$. Se debe recordar que en la obtención del MLE de (\mathbf{B}_0, Σ_0) se buscan $\hat{\mathbf{B}}$ y $\hat{\Sigma}$ tales que

$$(\hat{\mathbf{B}}, \hat{\Sigma}) = \underset{\mathbf{B}, \Sigma}{\text{argmax}} L(\mathbf{B}, \Sigma),$$

donde la función de log-verosimilitud es,

$$\begin{aligned} L(\mathbf{B}, \Sigma) &= \ln \left(\prod_{i=1}^n \frac{1}{(2\pi)^{q/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} \mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B}) \right\} \right) \\ &= \ln \left((2\pi)^{-nq/2} |\Sigma|^{-n/2} \right) - \frac{1}{2} \sum_{i=1}^n \mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B}) \\ &= -\frac{n}{2} \ln (2\pi)^q - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \sum_{i=1}^n \mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B}). \end{aligned}$$

Equivalentemente, se buscan $\hat{\mathbf{B}}$ y $\hat{\Sigma}$ tales que

$$(\hat{\mathbf{B}}, \hat{\Sigma}) = \underset{\mathbf{B}, \Sigma}{\operatorname{argmin}} \mathcal{L}(\mathbf{B}, \Sigma),$$

donde $\mathcal{L}(\mathbf{B}, \Sigma) = -\frac{2}{n}L(\mathbf{B}, \Sigma) - \ln(2\pi)^q$.

En el Capítulo 1 se demostrará que cuando $\mathbf{u}_i \sim \mathcal{N}_q(\mathbf{0}, \Sigma_0)$, el MLE de (\mathbf{B}_0, Σ_0) también satisface el problema de minimización dado por,

$$\begin{aligned} (\hat{\mathbf{B}}, \hat{\Sigma}) &= \underset{\mathbf{B}, \Sigma}{\operatorname{argmin}} |\Sigma| \\ \text{sujeto a} & \\ \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B}) &= q, \end{aligned} \tag{4}$$

donde $\sqrt{\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B})}$ es la distancia de Mahalanobis para cada $i = 1, \dots, n$. Dado que el MLE en este caso se obtiene a partir del supuesto de normalidad de los errores \mathbf{u}_i para $i = 1, \dots, n$, puede resultar muy sensible a valores atípicos y puede ser muy ineficiente incluso para pequeñas desviaciones de la normalidad (Carroll and Ruppert, 1985). Luego, se busca un estimador alternativo que presente las siguientes propiedades:

1. Alto punto de ruptura, es decir, que sea capaz de tolerar una alta proporción de observaciones atípicas sin afectar significativamente su desempeño. Se debe recordar que el punto de ruptura se refiere a la proporción de observaciones atípicas que un estimador puede manejar antes de que su influencia cause un impacto desmedido en la estimación, llevándola a resultados incorrectos o inestables.
2. Alta eficiencia de estimación, que se refiere a la capacidad de un estimador para proporcionar estimaciones precisas y consistentes, lo que implica que su varianza asintótica es baja en comparación con otros estimadores.

Puede observarse que el problema (4) involucra la escala cuadrática de la distancia de Mahalanobis en la restricción. En consecuencia, algunos autores han introducido variantes que reemplazan la escala cuadrática por otras con buenas propiedades de robustez. (Croux, 1994) demostró que el M-estimador de escala puede combinar alto punto de ruptura con alta eficiencia en condiciones de normalidad. Sin embargo, (Hössjer, 1992) había probado que el estimador de regresión basado en la M-escala no puede satisfacer ambas propiedades. Este fue el problema que resolvieron (Yohai and Zamar, 1988) proponiendo un τ -estimador en regresión univariada mediante la minimización de una τ -escala de los residuos. En (García Ben et al., 2006) se presenta una extensión de este τ -estimador para el modelo lineal multivariado, reemplazando la restricción del problema (4) por una restricción basada en la τ -escala robusta. Esta propuesta demostró que combina ambas propiedades.

El objetivo en esta tesis es analizar el comportamiento asintótico conjunto del τ -estimador (denotado por $(\tilde{\mathbf{B}}, \tilde{\Sigma})$) presentado por (García Ben et al., 2006) tanto para el caso de \mathbf{B}_0 con rango completo como de rango reducido. Este análisis se realizará a partir de los resultados

descritos en (García Ben et al., 2006), que están enfocados en \mathbf{B}_0 de rango completo y Σ_0 conocida, en (Bergesio et al., 2021) que analizan la consistencia para el caso de \mathbf{B}_0 con rango reducido y en (Bura et al., 2018), que analizan un M-estimador de regresión con rango reducido.

La tesis está organizada de la siguiente manera:

En el Capítulo 1 se introducen las definiciones y condiciones necesarias para definir el τ -funcional de estimación, se presentan los resultados de existencia, unicidad y consistencia para el MLM, y se derivan las ecuaciones de estimación del τ -estimador propuesto.

En el Capítulo 2, se obtiene la distribución asintótica conjunta de $(\tilde{\mathbf{B}}, \tilde{\Sigma})$ en el contexto de rango completo.

En el Capítulo 3, se obtiene la distribución asintótica conjunta de $(\tilde{\mathbf{B}}, \tilde{\Sigma})$ para el caso de \mathbf{B}_0 rango reducido.

El Capítulo 4 está destinado a proporcionar un análisis detallado del comportamiento de los nuevos τ -estimadores a través de un estudio de simulación. En este estudio, se llevará a cabo una comparación con otros estimadores robustos y estimadores de rango reducido. Además, estos estimadores se someterán a un análisis comparativo bajo la influencia de la contaminación.

Por último, en el Capítulo 5 se presentarán algunos ejemplos con datos reales con el propósito de comparar el estimador en estudio con otros estimadores robustos y estimadores de rango reducido.

Sólo las demostraciones principales serán incluidas en cada capítulo, mientras que las demás serán referidas a los Apéndices.

1 Definición del τ -funcional y propiedades

En este capítulo se presentan las definiciones y resultados esenciales que serán utilizados en el resto de la tesis. En particular, en la Sección 1.1 se define el τ -funcional de estimación. El τ -funcional de escala para el MLM (1) que da lugar a dicho funcional de estimación se define en la Sección 1.2. En la Sección 1.3 se expone el τ -funcional de escala basado en una muestra. Posteriormente, en la Sección 1.4 se expone un resultado de existencia del τ -funcional de estimación y en la Sección 1.5 se presenta el sistema de ecuaciones que define al τ -estimador basado en una muestra. Finalmente la unicidad del τ -funcional de estimación y la consistencia del mismo basado en una muestra se expone en la Sección 1.6.

1.1 Motivación del τ -estimador para el MLM

Antes de dar la definición del τ -funcional de estimación para el MLM (1) y tal como se ha mencionado en la introducción, en esta sección se demuestra que el MLE de $(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$ se obtiene como la solución de un problema de minimización con una restricción de igualdad. Más precisamente, en el Lema 1.1.1, se demuestra que el MLE para el parámetro $(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$, bajo normalidad de los errores, satisface el problema de optimización planteado en (4).

Antes se introducen algunas definiciones necesarias para tal fin.

Definición 1.1.1. Sea $\mathbf{V} \in \mathbb{R}^{q \times q}$ una matriz simétrica y definida positiva. La **norma de Mahalanobis** de $\mathbf{u} \in \mathbb{R}^q$ con respecto a \mathbf{V} , $d_M : \mathbb{R}^q \rightarrow \mathbb{R}$, está definida como,

$$d_M(\mathbf{u}, \mathbf{V}) = (\mathbf{u}^t \mathbf{V}^{-1} \mathbf{u})^{1/2}.$$

Observación 1.1.1. Sea (\mathbf{x}, \mathbf{y}) un vector aleatorio que satisface el MLM (1). La norma de Mahalanobis de $\mathbf{u}(\mathbf{B}) = \mathbf{y} - \mathbf{B}^t \mathbf{x}$ con respecto a $\boldsymbol{\Sigma}$ se denota como,

$$\begin{aligned} d(\mathbf{B}, \boldsymbol{\Sigma}) &:= d_M(\mathbf{u}(\mathbf{B}), \boldsymbol{\Sigma}) \\ &= ((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}. \end{aligned} \quad (1.1)$$

Definición 1.1.2. Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, una muestra aleatoria que satisface el MLM (1). Se definen los **errores muestrales** $\mathbf{u}_i(\mathbf{B})$, para cualquier matriz $\mathbf{B} \in \mathbb{R}^{p \times q}$, como

$$\mathbf{u}_i(\mathbf{B}) := \mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i, \quad (1.2)$$

y en este caso la distancia establecida en la Definición 1.1.1 resulta

$$d_i(\mathbf{B}, \boldsymbol{\Sigma}) = (\mathbf{u}_i(\mathbf{B})^t \boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}))^{1/2}. \quad (1.3)$$

Por simplicidad de notación en algunas ocasiones se denotará $\mathbf{u}_i := \mathbf{u}_i(\mathbf{B}_0)$.

Lema 1.1.1. Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, una muestra aleatoria que satisfice el MLM (1) con errores normales $\mathbf{u}_i(B_0)$. El par (\mathbf{B}, Σ) que es solución del problema de minimización

$$\begin{aligned} & \underset{\mathbf{B}, \Sigma}{\operatorname{argmin}} |\Sigma| \\ & \text{sujeto a} \\ & \frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, \Sigma) = q, \end{aligned} \tag{1.4}$$

coincide con el MLE $(\hat{\mathbf{B}}, \hat{\Sigma})$ dado en (2) y (3).

La demostración de este Lema es presentada en el Apéndice A.1.

Puede observarse que el problema planteado en (1.4) involucra en la restricción una escala definida como la escala cuadrática de las distancias de Mahalanobis. En (García Ben et al., 2006) se propone cambiar la escala cuadrática por un τ -funcional de escala que será definido más adelante en el presente capítulo. A continuación se presenta la solución al problema de minimización propuesto por los autores como un funcional de estimación. En lo que sigue, se denotará por $\mathcal{P}(E)$ al espacio de distribuciones sobre un espacio medible E . Antes se introduce una definición.

Definición 1.1.3. Para cada $Q \in \mathcal{P}(\mathbb{R}^{p+q})$ y $(\mathbf{B}, \Sigma) \in \Theta$ se define $H_{Q, \mathbf{B}, \Sigma}$ como la distribución de $d(\mathbf{B}, \Sigma)$ (ver Observación 1.1.1) cuando (\mathbf{x}, \mathbf{y}) tiene distribución Q .

Definición 1.1.4. Un τ -funcional de estimación para el MLM (1), está dado por el valor $(\tilde{\mathbf{B}}, \tilde{\Sigma}) \in \Theta$ que resuelve el problema de minimización

$$\begin{aligned} & (\tilde{\mathbf{B}}, \tilde{\Sigma}) = \underset{\mathbf{B}, \Sigma}{\operatorname{argmin}} |\Sigma| \\ & \text{sujeto a} \\ & \tau^2(H_{P, \mathbf{B}, \Sigma}) = \kappa_2, \end{aligned} \tag{1.5}$$

donde el τ -funcional de escala $\tau^2(H_{P, \mathbf{B}, \Sigma})$ y κ_2 , serán definidos en la siguiente sección.

1.2 Definición del τ -funcional de escala

Para definir el τ -funcional de escala, es necesario introducir el M-estimador de escala, las funciones a partir de las cuales se definen y sus propiedades.

Definición 1.2.1. (Yohai, 1987; Maronna et al., 2006) Una ρ -función acotada (o simplemente ρ -función) es una función $\rho : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ que satisface:

- A1) $\rho(0) = 0$;
- A2) $\rho(-u) = \rho(u)$;
- A3) si $0 \leq u \leq v$, entonces $\rho(u) \leq \rho(v)$;

A4) ρ pertenece a la clase de funciones C_2 ;

A5) $0 < a = \sup_u \rho(u) < \infty$;

A6) si $\rho(u) < a$ y $0 \leq u < v$, entonces $\rho(u) < \rho(v)$.

Asimismo, para asegurar la consistencia de Fisher de un τ -estimador, se impondrá otra condición a esta ρ -función.

A7) ρ satisface que $2\rho(v) - \psi(v)v > 0$ para $v > 0$, donde $\psi = \rho'$ y ρ' es la derivada de ρ .

Definición 1.2.2. Dada una ρ -función y un valor $\kappa > 0$, un **M-funcional de escala** para una distribución H , está definido por el valor $s(H)$ que satisface

$$E_H \left(\rho \left(\frac{|v|}{s(H)} \right) \right) = \kappa, \quad (1.6)$$

cuando v tiene distribución H .

Definición 1.2.3. Dado el par (\mathbf{x}, \mathbf{y}) que satisface el MLM (1), $(\mathbf{B}, \Sigma) \in \Theta$, una ρ -función y un valor $\kappa > 0$, un M-funcional de escala queda definido tomando $v = d(\mathbf{B}, \Sigma)$ y $H = H_{P, \mathbf{B}, \Sigma}$ en (1.6). Así, un **M-funcional de escala para el MLM (1)**, es el valor $s(H_{P, \mathbf{B}, \Sigma})$ que satisface

$$E_{H_{P, \mathbf{B}, \Sigma}} \left(\rho \left(\frac{d(\mathbf{B}, \Sigma)}{s(H_{P, \mathbf{B}, \Sigma})} \right) \right) = \kappa. \quad (1.7)$$

En lo que sigue, se expondrán las definiciones y condiciones suficientes para que el problema (1.7) tenga solución única.

Definición 1.2.4. Sean $\mathbf{a} \in \mathbb{R}^p$, $\mathbf{b} \in \mathbb{R}^q$, $c \in \mathbb{R}$ con \mathbf{a} o \mathbf{b} no nulos, se define como **hiperplano** $H_{(\mathbf{a}, \mathbf{b}, c)}$ a

$$H_{(\mathbf{a}, \mathbf{b}, c)} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{p+q} : \mathbf{a}^t \mathbf{x} + \mathbf{b}^t \mathbf{y} = c\}. \quad (1.8)$$

Además, se define la siguiente propiedad para $\epsilon > 0$.

(\mathbf{S}_ϵ) La probabilidad P tiene la propiedad (\mathbf{S}_ϵ) si para todo $H_{(\mathbf{a}, \mathbf{b}, c)}$, se tiene que $P(H_{(\mathbf{a}, \mathbf{b}, c)}) < \epsilon$. (1.9)

Lema 1.2.1. Si P satisface la propiedad (\mathbf{S}_ϵ) entonces $P(\{(\mathbf{x}, \mathbf{y}) : d(\mathbf{B}, \Sigma) = 0\}) < \epsilon$.

La demostración de este Lema es presentada en el Apéndice A.2.

La propiedad (\mathbf{S}_ϵ) se utilizará en la demostración de existencia y unicidad del M-funcional de escala para MLM, enunciada a continuación.

Lema 1.2.2. Sea (\mathbf{x}, \mathbf{y}) un vector que sigue el modelo MLM (1). Sea ρ una función que verifica A1)-A6) y P que satisface (S_ϵ) para algún $0 < \epsilon \leq 1 - \frac{\kappa}{a}$, con a definido en A5) y $0 < \kappa < (1 - \epsilon)a$. Entonces para cada $(\mathbf{B}, \Sigma) \in \Theta$ la ecuación (1.7) tiene solución y además es única.

La demostración de este Lema es presentada en el Apéndice A.3.

A continuación se define el τ -funcional de escala.

Definición 1.2.5. Dadas dos ρ -funciones ρ_1 y ρ_2 y $\kappa_1 > 0$, el τ -funcional de escala está definido por

$$\tau^2(H) = s^2(H)E_H \left(\rho_2 \left(\frac{|v|}{s(H)} \right) \right), \quad (1.10)$$

donde v tiene distribución H y $s(H)$ es el M -funcional de escala definido implícitamente por la ecuación

$$E_H \left(\rho_1 \left(\frac{|v|}{s(H)} \right) \right) = \kappa_1.$$

Definición 1.2.6. Sea (\mathbf{x}, \mathbf{y}) un vector que satisface el MLM (1) y sea $(\mathbf{B}, \Sigma) \in \Theta$. Entonces, el τ -funcional de escala para el MLM, se define como

$$\tau^2(H_{P,\mathbf{B},\Sigma}) = s^2(H_{P,\mathbf{B},\Sigma})E_{H_{P,\mathbf{B},\Sigma}} \left(\rho_2 \left(\frac{d(\mathbf{B}, \Sigma)}{s(H_{P,\mathbf{B},\Sigma})} \right) \right),$$

donde $H_{P,\mathbf{B},\Sigma}$ es la distribución de $d(\mathbf{B}, \Sigma)$ cuando (\mathbf{x}, \mathbf{y}) tiene distribución P , ρ_2 es una ρ -función y $s(H_{P,\mathbf{B},\Sigma})$ está definido mediante la ecuación

$$E_{H_{P,\mathbf{B},\Sigma}} \left(\rho_1 \left(\frac{d(\mathbf{B}, \Sigma)}{s(H_{P,\mathbf{B},\Sigma})} \right) \right) = \kappa_1,$$

donde ρ_1 es una ρ -función y $\kappa_1 > 0$.

Si se cumplen las hipótesis del Lema 1.2.2 para $\rho = \rho_1$ y $\kappa = \kappa_1$, se establece que el funcional de escala $s(H_{P,\mathbf{B},\Sigma})$ existe para el MLM con $(\mathbf{B}, \Sigma) \in \Theta$. Esto garantiza que $\tau^2(H_{P,\mathbf{B},\Sigma})$ también existe para el MLM con $(\mathbf{B}, \Sigma) \in \Theta$.

Los funcionales de escala de tipo M y τ anteriormente definidos cumplen la siguiente propiedad.

Lema 1.2.3. Para todo $\lambda > 0$, se verifica que

$$\begin{aligned} s^2(H_{P,\mathbf{B},\lambda\Sigma}) &= \frac{1}{\lambda} s^2(H_{P,\mathbf{B},\Sigma}), \\ \tau^2(H_{P,\mathbf{B},\lambda\Sigma}) &= \frac{1}{\lambda} \tau^2(H_{P,\mathbf{B},\Sigma}). \end{aligned}$$

La demostración de este Lema es presentada en el Apéndice A.4.

1.3 τ -funcional de escala basado en una muestra

Para representar el τ -funcional de estimación basado en una muestra aleatoria, las definiciones de la sección anterior serán dadas a partir de la medida de probabilidad empírica P_n que se define, para cualquier subconjunto A medible de \mathbb{R}^{p+q} , como

$$P_n(A) = \frac{1}{n} \sum_{i=1}^n I_A(\mathbf{x}_i, \mathbf{y}_i),$$

donde I_A es la función indicadora de A . Aquí $H_{P_n, \mathbf{B}, \Sigma}$ denotará la distribución de $d(\mathbf{B}, \Sigma)$ cuando (\mathbf{x}, \mathbf{y}) tiene distribución P_n .

Definición 1.3.1. Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, una muestra aleatoria que satisfice el MLM (1) y sea $(\mathbf{B}, \Sigma) \in \Theta$. Se define una **M-escala** como el valor de $s(H_{P_n, \mathbf{B}, \Sigma})$ que es solución de

$$\frac{1}{n} \sum_{i=1}^n \rho \left(\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})} \right) = \frac{1}{n} \sum_{i=1}^n \rho \left(\frac{(\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1} \mathbf{u}_i(\mathbf{B}))^{1/2}}{s(H_{P_n, \mathbf{B}, \Sigma})} \right) = \kappa, \quad (1.11)$$

donde κ es una constante y ρ una ρ -función. Por simplicidad de notación en algunas ocasiones se denotará $s_n(\mathbf{B}, \Sigma) := s(H_{P_n, \mathbf{B}, \Sigma})$.

En lo que sigue se presentan las definiciones y condiciones necesarias y suficientes para la existencia de una solución al problema (1.11). Estas son similares a las dadas por (Yohai, 1987) y a las presentadas para el Lema 1.2.2.

Primero se define en el caso muestral el término h_n (García Ben et al., 2006) como el máximo número de observaciones $(\mathbf{x}_i, \mathbf{y}_i)$ de la muestra que están en un hiperplano, es decir,

$$h_n = \max_{\|\mathbf{a}\| + \|\mathbf{b}\| > 0} \#\{i : \mathbf{a}^t \mathbf{x}_i + \mathbf{b}^t \mathbf{y}_i = c\}. \quad (1.12)$$

Observación 1.3.1. Si P_n satisface que $\frac{h_n}{n} < 1 - \frac{\kappa}{a}$ entonces P_n cumple la propiedad (S_ϵ) para todo ϵ tal que $\frac{h_n}{n} < \epsilon < 1 - \frac{\kappa}{a}$. Luego, P_n está en las hipótesis del Lema 1.2.1 y por lo tanto existe única solución para (1.7) que, en el caso de P_n , es (1.11) probando así la existencia y unicidad de la M-escala en este caso.

A partir de la existencia de la M-escala, se define la τ -escala para una muestra de observaciones de la siguiente manera.

Definición 1.3.2. Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, una muestra aleatoria que sigue el MLM (1). La **τ -escala** está definida por

$$\tau^2(H_{P_n, \mathbf{B}, \Sigma}) = s^2(H_{P_n, \mathbf{B}, \Sigma}) \frac{1}{n} \sum_{i=1}^n \rho_2 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})} \right), \quad (1.13)$$

donde ρ_2 es una ρ -función y $s(H_{P_n, \mathbf{B}, \Sigma})$ es la M-escala que está definida mediante la ecuación

$$\frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})} \right) \right) = \kappa_1,$$

ρ_1 es una ρ -función y κ_1 una constante. Por simplicidad de notación en algunas ocasiones se denotará $\tau_n(\mathbf{B}, \Sigma) := \tau(H_{P_n, \mathbf{B}, \Sigma})$.

La idea intuitiva que motivó la definición de esta escala es la siguiente (Ver (Zamar, 1994) para un caso particular): Si la función ρ_2 es aproximadamente cuadrática cerca de cero y los valores $\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})}$ son relativamente pequeños, entonces $s^2(H_{P_n, \mathbf{B}, \Sigma})\rho_2\left(\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})}\right) \approx d_i^2(\mathbf{B}, \Sigma)$ y la τ -escala no será muy diferente de la escala cuadrática de las distancias de Mahalanobis presentada en la restricción del problema (4). Por otro lado, si los valores $\frac{d_i(\mathbf{B}, \Sigma)}{s(H_{P_n, \mathbf{B}, \Sigma})}$ son grandes, entonces la influencia de esos puntos será reducida por las propiedades de la función ρ_2 . Por lo tanto, los estimadores que se definan con la τ -escala serán eficientes cuando los datos sean aproximadamente normales y resistentes a la presencia de outliers, siempre que las funciones ρ_1 y ρ_2 sean escogidas convenientemente.

Como la τ -escala para una muestra es igual al τ -funcional de escala evaluado en P_n , por el Lema 1.2.3, s_n y τ_n cumplen las propiedades de dicho lema.

1.4 Existencia del τ -funcional de estimación para el MLM

En esta sección se introducen las condiciones suficientes para la existencia del τ -funcional de estimación, tanto en el caso poblacional (1.5), como el muestral que será definido más adelante.

En lo que sigue, ρ_1 y ρ_2 serán ρ -funciones que satisfacen A1)-A7). Asumiendo \tilde{P} como la distribución de (\mathbf{x}, \mathbf{y}) , para $i = 1, 2$ se define

$$\kappa_i := E_{H_{\tilde{P}, \mathbf{B}_0, \Sigma_0}}(\rho_i(d(\mathbf{B}_0, \Sigma_0))). \quad (1.14)$$

Teorema 1. (Bergesio et al., 2021, Teorema 2) Sea (\mathbf{x}, \mathbf{y}) un vector aleatorio que cumple con el MLM (1) con P satisfaciendo la propiedad (S_ϵ) para algún $0 < \epsilon \leq 1 - \frac{\kappa_1}{a_1}$, con $0 < a_1 = \sup_u \rho_1(u) < \infty$ y $0 < \kappa_1 < (1 - \epsilon)a_1$. Entonces, existe al menos un valor de (\mathbf{B}, Σ) que cumple con las condiciones impuestas por el funcional $(\tilde{\mathbf{B}}, \tilde{\Sigma})$ en (1.5).

A partir de la Definición 1.1.4 del τ -funcional de estimación se define la versión muestral del mismo como sigue.

Definición 1.4.1. El τ -funcional de estimación evaluado en la medida empírica para el MLM (1), está definido como el valor $(\mathbf{B}_n, \Sigma_n) \in \Theta$ que resuelve el problema de minimización

$$\begin{aligned} (\mathbf{B}_n, \Sigma_n) &= \underset{(\mathbf{B}, \Sigma) \in \Theta}{\operatorname{argmin}} |\Sigma| \\ \text{sujeto a} & \end{aligned} \quad (1.15)$$

$$\tau^2(H_{P_n, \mathbf{B}, \Sigma}) = \kappa_2.$$

De aquí en adelante cuando no haya ambigüedad, se referirá a un τ -estimador como un τ -funcional de estimación evaluado en la medida empírica.

(García Ben et al., 2006) establecieron la existencia de al menos una solución $(\mathbf{B}_n, \Sigma_n) \in \Theta$ para el problema (1.15) siempre que las ρ -funciones ρ_1 y ρ_2 cumplan con las condiciones A1)–A6) y $h_n/n < 1 - \kappa_1/a_1$.

1.5 Ecuaciones que definen el τ -estimador basado en una muestra

Resolver el problema de minimización con una restricción (1.15) conduce a resolver un sistema de ecuaciones. Para esto, en (Lopuhaa, 1991) se utilizan multiplicadores de Lagrange. En esta tesis, se optará por introducir un sistema de ecuaciones que permitirá hallar el mínimo de la función objetivo y se agregará al sistema la ecuación impuesta por la restricción.

Antes de plantear el sistema de ecuaciones, se introduce la notación necesaria para tal fin.

Sean,

$$\psi_i(v) := \rho'_i(v), \quad (1.16)$$

para $i = 1, 2$, con ρ_1 , y ρ_2 , ρ -funciones que cumplen las condiciones A1)–A7) y sean, además,

$$w_n^*(v) := \frac{\psi_n^*(v)}{v}, \quad (1.17)$$

donde,

$$\psi_n^*(v) = \psi_{n, \mathbf{B}, \Sigma}^*(v) := C_n \psi_1(v) + D_n \psi_2(v), \quad (1.18)$$

con

$$C_n = C_n(\mathbf{B}, \Sigma) := \frac{1}{n} \sum_{i=1}^n [2\rho_2(d_i^*(\mathbf{B}, \Sigma)) - \psi_2(d_i^*(\mathbf{B}, \Sigma))d_i^*(\mathbf{B}, \Sigma)], \quad (1.19)$$

$$D_n = D_n(\mathbf{B}, \Sigma) := \frac{1}{n} \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma))d_i^*(\mathbf{B}, \Sigma), \quad (1.20)$$

$$d_i^*(\mathbf{B}, \Sigma) := \frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)}. \quad (1.21)$$

Observación 1.5.1. Por la definición de la M -escala dada en la ecuación que sigue a (1.11), como $0 < \kappa_1 < a_1 = \sup_u \rho_1(u) < \infty$, entonces existe al menos un $d_j^*(\mathbf{B}, \Sigma)$ tal que $0 < \rho_1(d_i^*(\mathbf{B}, \Sigma)) < a_1$ y por lo tanto $\psi_1(d_i^*(\mathbf{B}, \Sigma)) > 0$ y $d_i^*(\mathbf{B}, \Sigma) > 0$ con lo cual $D_n > 0$.

Observación 1.5.2. A partir de la observación anterior, donde se prueba que existe un j tal que $0 < d_j^*(\mathbf{B}, \Sigma)$, resulta que $\sum_i \rho_2(d_i^*(\mathbf{B}, \Sigma)) > 0$.

Finalmente, utilizando esta notación se tiene el teorema que define el sistema de ecuaciones resultante.

Teorema 2. *Se supone que P_n satisface las condiciones de la Observación 1.3.1. Cuando las funciones ρ_1 y ρ_2 verifican las condiciones A1)-A7) y existe una solución $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \in \Theta$ para el problema (1.15), entonces dicha solución satisface las ecuaciones*

$$\left\{ \begin{array}{l} \sum_{i=1}^n w_n^*(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t = \mathbf{0}, \\ \sum_{i=1}^n \left\{ q w_n^*(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B}) - s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_n^*(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \boldsymbol{\Sigma} \right\} = \mathbf{0}, \\ \sum_{i=1}^n \left\{ s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \rho_2(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) - \kappa_2 \right\} = 0, \end{array} \right. \quad (1.22)$$

donde $s_n(\mathbf{B}, \boldsymbol{\Sigma})$ es la M -escala definida en (1.11) con $\rho = \rho_1$ y $\kappa = \kappa_1$.

La demostración de este teorema se encuentra en el Apéndice A.5.

De acuerdo con la primera ecuación de (1.22), la j -ésima columna de \mathbf{B}_n es el estimador de mínimos cuadrados ponderado correspondiente a la regresión univariada cuya variable dependiente es la j -ésima componente de \mathbf{y} , el vector de variables predictoras es \mathbf{x} , y la i -observación tiene peso dado por $w_n^*(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))$. Además, de acuerdo con la segunda ecuación de (1.22), $\boldsymbol{\Sigma}_n$ es proporcional a la covarianza muestral ponderada de residuos con los mismos pesos. Por último, la tercera ecuación de (1.22), representa la restricción que se le imponía el problema de optimización inicial.

En la siguiente sección se estudiará la consistencia del τ -estimador para el MLM para el caso en que los errores en (1) tengan una distribución elíptica, puesto que tener esta propiedad será de gran ayuda para el estudio de la distribución asintótica.

1.6 Unicidad del τ -funcional de estimación para el MLM y consistencia del mismo basado en una muestra

Para demostrar la convergencia de $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ hacia una solución única $(\tilde{\mathbf{B}}, \tilde{\boldsymbol{\Sigma}})$, es esencial establecer la unicidad de $(\tilde{\mathbf{B}}, \tilde{\boldsymbol{\Sigma}})$ dada en (1.5).

Observación 1.6.1. *Por simplicidad, de ahora en adelante dadas las ρ -funciones ρ_1 y ρ_2 y los valores κ_1 y κ_2 , se denotará,*

$$\boldsymbol{\Sigma}_0^* := r \boldsymbol{\Sigma}_0, \quad (1.23)$$

donde $r = \frac{\tau^2(H_{P, \mathbf{B}_0, \boldsymbol{\Sigma}_0})}{\kappa_2}$. Por la propiedad del M -funcional de escala presentada en el Lema 1.2.3 se tiene,

$$s^2(H_{P, \mathbf{B}_0, \boldsymbol{\Sigma}_0^*}) = \frac{\kappa_2}{E_{H_{P, \mathbf{B}_0, \boldsymbol{\Sigma}_0}} \left(\rho_2 \left(\frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{k_0} \right) \right)}, \quad (1.24)$$

donde

$$k_0 := s(H_{P, \mathbf{B}_0, \boldsymbol{\Sigma}_0}). \quad (1.25)$$

Notar que si $P = \tilde{P}$ en (1.14) se tiene que $r = 1$.

En lo que resta de este capítulo y en el siguiente, se asumirá la siguiente condición:

A8) Sean ρ_1 y ρ_2 ρ -funciones acotadas y sea f_0^* una función no creciente y con al menos un punto de decrecimiento en el intervalo donde ambas funciones ρ_1 y ρ_2 son estrictamente crecientes. La distribución F_0 de \mathbf{u} tiene densidad $f_0(\mathbf{u})$ dada por

$$f_0(\mathbf{u}) = \frac{f_0^*(\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1} \mathbf{u})}{|\boldsymbol{\Sigma}_0|^{1/2}}. \quad (1.26)$$

Bajo esta hipótesis P tiene la propiedad (\mathbf{S}_ϵ) para todo $\epsilon > 0$.

Teorema 3. (García Ben et al., 2006, Lema A.12) Sea (\mathbf{x}, \mathbf{y}) el vector aleatorio, con distribución P , que sigue el MLM (1) con parámetros \mathbf{B}_0 y $\boldsymbol{\Sigma}_0$ donde el error \mathbf{u} tiene una distribución que satisface A8). Se supone también que los ρ_k , $k = 1, 2$ satisfacen A1)-A7). Entonces el problema:

$$\begin{aligned} & \underset{\mathbf{B}, \boldsymbol{\Sigma}}{\operatorname{argmin}} |\boldsymbol{\Sigma}| \\ & \text{sujeto a} \end{aligned} \quad (1.27)$$

$$\tau^2(H_{P, \mathbf{B}, \boldsymbol{\Sigma}}) = \kappa_2,$$

tiene solución única, dada por $(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$.

En otras palabras, se observa que el Teorema 3 muestra que bajo ciertos supuestos, la función de estimación τ del MLM está bien definida en el sentido de que devuelve un solo valor en el espacio de parámetros. También alcanza un mínimo en los valores verdaderos de los parámetros, a excepción de la constante multiplicativa de la matriz de covarianza.

En (García Ben et al., 2006) se estableció la consistencia del τ -estimador de rango completo, lo que garantiza la consistencia para el τ -estimador $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ definido en (1.15). A continuación, se enuncia dicho teorema.

Teorema 4. (García Ben et al., 2006, Teorema 4) Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Sea $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ el τ -estimador definido en (1.15). Bajo las hipótesis del Teorema 3 se tiene que,

$$\lim_{n \rightarrow \infty} \mathbf{B}_n = \mathbf{B}_0 \text{ c.s.}$$

$$\lim_{n \rightarrow \infty} \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}_0^* \text{ c.s.}$$

2 Distribución Asintótica (Caso: Rango Completo)

En el presente capítulo se prueba la distribución asintótica del τ -estimador del MLM (1) (\mathbf{B}_n, Σ_n) cuando $\text{rango}(\mathbf{B}_0) = \min\{p, q\}$. La distribución asintótica de \mathbf{B}_n para el caso particular de Σ_0 conocida fue estudiada en (García Ben et al., 2006).

Para obtener este resultado, se utiliza el estimador (\mathbf{B}_n, Σ_n) definido como una solución del sistema de ecuaciones (1.22). Para comenzar, dicho sistema se presenta de manera matricial y luego se estudian la distribución del estimador.

En este capítulo, además de las condiciones A1)-A8) se presentan los siguientes supuestos que serán necesarios para probar algunos resultados:

A9) \mathbf{x} es aleatorio, con distribución M_0 , tiene momento de orden 4 finito, tiene $E(\mathbf{x}) = \mathbf{0}$ y matriz de covarianza $\Sigma_{\mathbf{x}}$ no singular;

A10) Para $k = 1, 2$, existe m_k tal que $\rho_k(v)$ es constante para $|v| > m_k$.

A continuación, se enuncian los resultados que son necesarios para la prueba de la distribución asintótica.

2.1 El sistema de ecuaciones (1.22) en forma matricial

Para comenzar se introduce la notación necesaria para el objetivo propuesto.

Definición 2.1.1. Cuando las funciones ρ_1 y ρ_2 son diferenciables se define, para cada $i = 1, \dots, n$,

$$\mathbf{A}_i(\mathbf{B}, \Sigma, t) := \begin{pmatrix} \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, t)) \\ \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, t)) \\ \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, t)) \\ \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, t)) \\ \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, t) \end{pmatrix}, \quad (2.1)$$

con

$$\begin{cases} \Psi_{i,k,\mathbf{B}}(\mathbf{B}, \Sigma, t) = \frac{\psi_k\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right)}{\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right)} \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t, \\ \Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, t) = q \frac{\psi_k\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right)}{\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right)} \mathbf{u}_i(\mathbf{B}) \mathbf{u}_i(\mathbf{B})^t - t^2 \psi_k\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right) \left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right) \Sigma, \\ \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, t) = t^2 \rho_2\left(\frac{d_i(\mathbf{B}, \Sigma)}{t}\right) - \kappa_2, \end{cases} \quad (2.2)$$

para $k = 1, 2$ con $\mathbf{u}_i(\mathbf{B})$, $d_i(\mathbf{B}, \Sigma)$, $\psi_1(\cdot)$ y $\psi_2(\cdot)$ definidos en (1.2), (1.16) y (1.3), respectivamente. Además se define la matriz

$$\mathbf{M}_{(C_n, D_n)}(\mathbf{B}, \Sigma) := \begin{pmatrix} C_n(\mathbf{B}, \Sigma)\mathbf{I}_{pq} & \mathbf{0} & D_n(\mathbf{B}, \Sigma)\mathbf{I}_{pq} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_n(\mathbf{B}, \Sigma)\mathbf{I}_{q(q+1)/2} & \mathbf{0} & D_n(\mathbf{B}, \Sigma)\mathbf{I}_{q(q+1)/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}. \quad (2.3)$$

donde $C_n(\mathbf{B}, \Sigma)$ y $D_n(\mathbf{B}, \Sigma)$ fueron definidas en (1.20) y (1.19), respectivamente.

A partir de estas definiciones se tiene el siguiente resultado.

Lema 2.1.1. *El sistema (1.22) puede ser escrito como*

$$\mathbf{M}_{(C_n, D_n)}(\mathbf{B}, \Sigma) \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)) = \mathbf{0}. \quad (2.4)$$

La demostración de este resultado se muestra en el Apéndice B.1.

Para abordar el análisis de la distribución asintótica del τ -estimador (\mathbf{B}_n, Σ_n) se utiliza el sistema planteado en (2.4). Para ello, en las siguientes secciones se prueban resultados asintóticos para cada componente del mismo.

2.1.1 Consistencia de $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n)$

Para probar la consistencia de $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n)$ se estudia el comportamiento asintótico de $C_n(\mathbf{B}_n, \Sigma_n)$ y de $D_n(\mathbf{B}_n, \Sigma_n)$. Para ello se definen,

$$C(\mathbf{B}, \Sigma) = E_{H_{P, \mathbf{B}, \Sigma}} \left(2\rho_2 \left(\frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)} \right) - \psi_2 \left(\frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)} \right) \frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)} \right)$$

$$D(\mathbf{B}, \Sigma) = E_{H_{P, \mathbf{B}, \Sigma}} \left(\psi_1 \left(\frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)} \right) \frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)} \right).$$

Observación 2.1.1. *Usando las definiciones anteriores y los resultados del Lema 1.2.3, se obtiene que*

$$C(\mathbf{B}_0, \Sigma_0^*) = E_{H_{P, \mathbf{B}_0, \Sigma_0}} \left(2\rho_2 \left(\frac{d(\mathbf{B}_0, \Sigma_0)}{k_0} \right) - \psi_2 \left(\frac{d(\mathbf{B}_0, \Sigma_0)}{k_0} \right) \frac{d(\mathbf{B}_0, \Sigma_0)}{k_0} \right) = C(\mathbf{B}_0, \Sigma_0), \quad (2.5)$$

$$D(\mathbf{B}_0, \Sigma_0^*) = E_{H_{P, \mathbf{B}_0, \Sigma_0}} \left(\psi_1 \left(\frac{d(\mathbf{B}_0, \Sigma_0)}{k_0} \right) \frac{d(\mathbf{B}_0, \Sigma_0)}{k_0} \right) = D(\mathbf{B}_0, \Sigma_0), \quad (2.6)$$

con k_0 definido en (1.25). Además, la matriz $\mathbf{M}_{(C, D)}(\mathbf{B}_0, \Sigma_0)$ queda definida de la siguiente manera

$$\mathbf{M}_{(C, D)}(\mathbf{B}_0, \Sigma_0) := \begin{pmatrix} C(\mathbf{B}_0, \Sigma_0)\mathbf{I}_{pq} & \mathbf{0} & D(\mathbf{B}_0, \Sigma_0)\mathbf{I}_{pq} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C(\mathbf{B}_0, \Sigma_0)\mathbf{I}_{q(q+1)/2} & \mathbf{0} & D(\mathbf{B}_0, \Sigma_0)\mathbf{I}_{q(q+1)/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}, \quad (2.7)$$

donde $C(\mathbf{B}_0, \Sigma_0)$ y $D(\mathbf{B}_0, \Sigma_0)$ están dadas por (2.5) y (2.6), respectivamente.

Teorema 5. (*Lopuhaa, 1991; García Ben et al., 2006*) Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Se asume que ρ_1 satisface A1)-A6) y A10), ρ_2 satisface A1)-A7) y A10), \mathbf{u}_i definidos en (1.2) satisfacen A8), y \mathbf{x}_i satisfacen A9). Si (\mathbf{B}_n, Σ_n) es una sucesión consistente (c.s.) de $(\mathbf{B}_0, \Sigma_0^*)$ entonces $C_n(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} C(\mathbf{B}_0, \Sigma_0)$ y $D_n(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} D(\mathbf{B}_0, \Sigma_0)$.

La demostración de este resultado está en el Apéndice B.2.

Corolario 2.1.1. Se asumen los supuestos del Teorema 5. Si (\mathbf{B}_n, Σ_n) es una sucesión consistente (c.s.) de $(\mathbf{B}_0, \Sigma_0^*)$ entonces $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} \mathbf{M}_{(C, D)}(\mathbf{B}_0, \Sigma_0)$.

2.1.2 Distribución asintótica de $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*))$

Para probar la consistencia de $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*))$ se observa que

$$\begin{aligned} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) &= \sum_{i=1}^n [\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) - \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \\ &\quad + \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)), \end{aligned} \quad (2.8)$$

por lo que los dos términos de la derecha de esta igualdad se estudian por separado. Para ello se definen

$$d_0 := d(\mathbf{B}_0, \Sigma_0) = ((\mathbf{y} - \mathbf{B}_0^t \mathbf{x})^t \Sigma_0^{-1} (\mathbf{y} - \mathbf{B}_0^t \mathbf{x}))^{1/2}, \quad (2.9)$$

$$d_0^* := \frac{d_0}{k_0}, \quad (2.10)$$

y, para $i = 1, \dots, n$,

$$d_{(i)0}^*(\mathbf{B}_0, \Sigma_0) := \frac{d_i(\mathbf{B}_0, \Sigma_0)}{k_0}, \quad (2.11)$$

donde k_0 fue definido en (1.25).

El primer resultado corresponde a la convergencia del primer término de (2.8).

Lema 2.1.2. Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Se asume que ρ_1 satisface A1)-A6) y A10), ρ_2 satisface A1)-A7) y A10), \mathbf{u}_i satisfacen A8), \mathbf{x}_i satisfacen A9). Entonces,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n [\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) - \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \xrightarrow{p} \mathbf{0}.$$

La demostración de este resultado está en el Apéndice B.3. A continuación se analizará la distribución asintótica para $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))$.

Teorema 6. Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Se supone que ρ_1 satisface A1)-A6) y A10), ρ_2 satisface A1)-A7) y A10), \mathbf{u}_i satisfacen A8) y \mathbf{x}_i satisfacen A9). Entonces,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \text{Var}(\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)))),$$

donde la varianza en la expresión límite se toma respecto de la distribución de (\mathbf{x}, \mathbf{y}) bajo el MLM (1), y

$$\text{Var}(\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))) := \begin{pmatrix} L_{11} & \mathbf{0} & L_{13} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_{22} & \mathbf{0} & L_{24} & L_{25} \\ L_{13}^t & \mathbf{0} & L_{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_{24}^t & \mathbf{0} & L_{44} & L_{45} \\ \mathbf{0} & L_{25}^t & \mathbf{0} & L_{45}^t & L_{55} \end{pmatrix}, \quad (2.12)$$

con

$$L_{11} = E[\mathbf{x}_i \mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0) \right],$$

$$L_{13} = E[\mathbf{x}_i \mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_1(d_{(i)0}^*) \psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}} \right) \mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0) \right],$$

$$\begin{aligned} L_{22} = & q^2 E \left[\left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] + k_0^4 E[\psi_1^2(d_{(i)0}^*) d_{(i)0}^{*2}] \\ & \times \text{vech}(\Sigma_0) \text{vech}^t(\Sigma_0) - q k_0^2 E[\psi_1^2(d_{(i)0}^*) \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0))] \text{vech}^t(\Sigma_0) \\ & - q k_0^2 \text{vech}(\Sigma_0) E[\psi_1^2(d_{(i)0}^*) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0))], \end{aligned}$$

$$\begin{aligned} L_{24} = & q^2 E \left[\frac{\psi_1(d_{(i)0}^*) \psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}} \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] + k_0^4 E[\psi_1(d_{(i)0}^*) \psi_2(d_{(i)0}^*) d_{(i)0}^{*2}] \\ & \times \text{vech}(\Sigma_0) \text{vech}^t(\Sigma_0) - q k_0^2 E[\psi_1(d_{(i)0}^*) \psi_2(d_{(i)0}^*) \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0))] \text{vech}^t(\Sigma_0) \\ & - q k_0^2 \text{vech}(\Sigma_0) E[\psi_1(d_{(i)0}^*) \psi_2(d_{(i)0}^*) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0))], \end{aligned}$$

$$\begin{aligned} L_{25} = & q E \left[\left\{ \frac{k_0^2}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] \\ & - k_0^2 E \left[\left\{ \frac{k_0^2}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \psi_1(d_{(i)0}^*) d_{(i)0}^* \right] \text{vech}(\Sigma_0), \end{aligned}$$

$$L_{33} = E[\mathbf{x}_i \mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0) \right],$$

$$\begin{aligned}
 L_{44} &= q^2 E \left[\left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] + k_0^4 E \left[\psi_2^2(d_{(i)0}^*) d_{(i)0}^{*2} \right] \\
 &\quad \times \text{vech}(\boldsymbol{\Sigma}_0) \text{vech}^t(\boldsymbol{\Sigma}_0) - q k_0^2 E \left[\psi_2^2(d_{(i)0}^*) \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] \text{vech}^t(\boldsymbol{\Sigma}_0) \\
 &\quad - q k_0^2 \text{vech}(\boldsymbol{\Sigma}_0) E \left[\psi_2^2(d_{(i)0}^*) \text{vech}^t(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right], \\
 L_{45} &= q E \left[\left\{ \frac{k_0^2}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] \\
 &\quad - k_0^2 E \left[\left\{ \frac{k_0^2}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \psi_2(d_{(i)0}^*) d_{(i)0}^* \right] \text{vech}(\boldsymbol{\Sigma}_0), \\
 L_{55} &= \frac{k_0^4}{r^2} E \left[\rho_2^2(d_{(i)0}^*) \right] - \frac{2}{r} \kappa_2 k_0^2 E \left[\rho_2(d_{(i)0}^*) \right] + \kappa_2^2,
 \end{aligned}$$

con $d_{(i)0}^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0) = d_{(i)0}^*$.

La demostración de este resultado se encuentra en el Apéndice B.4.

A partir del Lema 2.1.2 se prueba entonces que $\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) \right)$ es asintóticamente equivalente a $\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) \right)$ y por tanto convergen a la misma distribución (la dada en el Teorema 6). Como consecuencia de estos resultados se obtiene entonces el siguiente teorema, que corresponde a la distribución asintótica de $\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$.

Teorema 7. *Bajo los supuestos del Lema 2.1.2 se tiene que,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \text{Var}(\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))), \quad (2.13)$$

con $\text{Var}(\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))$ definida en (2.12).

2.1.3 Convergencia de derivadas

A continuación, se presenta la deducción de las derivadas de los elementos que conforman a $\mathbf{A}_i(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$ para $i = 1, \dots, n$. Para ello, se denota a la derivada de la función $\sum_{i=1}^n \mathbf{A}_i(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$ como

$$(F_{1,n}(\mathbf{B}, \boldsymbol{\Sigma}), F_{2,n}(\mathbf{B}, \boldsymbol{\Sigma})), \quad (2.14)$$

donde la primera columna $F_{1,n}(\mathbf{B}, \boldsymbol{\Sigma})$ corresponde a la de la derivada de $\sum_{i=1}^n \mathbf{A}_i(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$ con respecto a $\text{vec}^t(\mathbf{B}^t)$ y la segunda columna, $F_{2,n}(\mathbf{B}, \boldsymbol{\Sigma})$, corresponde a la derivada con

respecto a $\text{vech}^t(\Sigma)$, es decir,

$$(F_{1,n}(\mathbf{B}, \Sigma), F_{2,n}(\mathbf{B}, \Sigma)) := \sum_{i=1}^n \begin{pmatrix} \frac{\partial \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial [\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))]}{\partial \text{vech}^t(\Sigma)} \\ \frac{\partial \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vech}^t(\Sigma)} \\ \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vech}^t(\Sigma)} \\ \frac{\partial \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))}{\partial \text{vech}^t(\Sigma)} \\ \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))}{\partial \text{vech}^t(\Sigma)} \end{pmatrix}. \quad (2.15)$$

La expresión matricial anterior es de tamaño $(2pq + q(q + 1) + 1) \times (pq + q(q + 1)/2)$.

Estas derivadas serán parte de la matriz de covarianza de la distribución asintótica del τ -estimador, para la cual también es necesario introducir la siguiente notación.

Definición 2.1.2. *A partir de dos sucesiones $\{(\mathbf{B}_n, \Sigma_n)\}_{n \geq 1}$, $\{(\mathbf{B}_n^*, \Sigma_n^*)\}_{n \geq 1}$ contenidas en Θ se define la sucesión de matrices*

$$\mathbf{V}_n := \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) (F_{1,n}(\mathbf{B}_n^*, \Sigma_n^*), F_{2,n}(\mathbf{B}_n^*, \Sigma_n^*)) \quad (2.16)$$

de tamaño $(pq + q(q + 1)/2 + 1) \times (pq + q(q + 1)/2)$, con $(F_{1,n}(\mathbf{B}_n^*, \Sigma_n^*), F_{2,n}(\mathbf{B}_n^*, \Sigma_n^*))$ y $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) = \mathbf{M}_{C_n(\mathbf{B}_n, \Sigma_n), D_n(\mathbf{B}_n, \Sigma_n)}(\mathbf{B}_n, \Sigma_n)$ que fueron definidos en (2.15) y (2.3), respectivamente.

Para la convergencia en distribución del τ -estimador es necesario el siguiente resultado.

Lema 2.1.3. *Asumiendo que las columnas de \mathbf{V}_n son linealmente independientes (con probabilidad uno) la inversa de Moore-Penrose, denotada por \mathbf{V}_n^\dagger , de la matriz \mathbf{V}_n definida en (2.16) es la definida en (B.39) y verifica que $\mathbf{V}_n^\dagger \mathbf{V}_n = \mathbf{I}_{(pq+q(q+1)/2)}$.*

La demostración de este resultado está en el Apéndice B.5.

Teorema 8. *Sean dos sucesiones $\{(\mathbf{B}_n, \Sigma_n)\}_{n \geq 1}$, $\{(\mathbf{B}_n^*, \Sigma_n^*)\}_{n \geq 1}$ contenidas en Θ que convergen casi seguramente a $(\mathbf{B}_0, \Sigma_0^*)$, con $r \neq 0$. Entonces la sucesión $n\mathbf{V}_n^\dagger$, con \mathbf{V}_n^\dagger definida en el Lema 2.1.3, converge casi seguramente a \mathbf{V}^\dagger definida en (B.67).*

La demostración de este resultado está en el Apéndice B.6.

2.2 Normalidad asintótica del τ -estimador (rango completo)

El siguiente Teorema establece la distribución asintótica del τ -estimador del MLM (1) (\mathbf{B}_n, Σ_n) cuando \mathbf{B}_0 tiene rango completo. En lo que sigue de este capítulo y en los siguientes,

para simplificar la notación tanto como sea posible, se concatenarán los vectores columna sin transponerlos; es decir, se escribirá (\mathbf{a}, \mathbf{b}) en lugar de $(\mathbf{a}^t, \mathbf{b}^t)^t$, para cualesquiera vectores \mathbf{a} y \mathbf{b} .

Teorema 9. *Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Se asume que ρ_1 satisface A1)-A6) y A10), ρ_2 satisface A1)-A7) y A10), \mathbf{u}_i satisfacen A8) y \mathbf{x}_i satisfacen A9), para $i \geq 1$. Si (\mathbf{B}_n, Σ_n) es una sucesión que satisface el sistema de ecuaciones (2.4) y es fuertemente consistente para $(\mathbf{B}_0, \Sigma_0^*)$, con $r \neq 0$, entonces*

$$\sqrt{n}(\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{W}),$$

con

$$\mathbf{W} = \mathbf{V}^\dagger \mathbf{M}_{(C,D)}(\mathbf{B}_0, \Sigma_0) \text{Var} [\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \mathbf{M}_{(C,D)}^t(\mathbf{B}_0, \Sigma_0) (\mathbf{V}^\dagger)^t, \quad (2.17)$$

donde $\mathbf{A}_i(\cdot)$, \mathbf{V}^\dagger , $\mathbf{M}_{(C,D)}(\mathbf{B}_0, \Sigma_0)$ y $\text{Var} [\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))]$ fueron definidas en (2.1), (B.67), (2.7) y (2.12), respectivamente.

Demostración del Teorema 9. Sea (\mathbf{B}_n, Σ_n) una sucesión consistente (c.s.) de $(\mathbf{B}_0, \Sigma_0^*)$ que satisface el sistema de ecuaciones (2.4). Aplicando el Teorema de Valor Medio en los puntos (\mathbf{B}_n, Σ_n) y $(\mathbf{B}_0, \Sigma_0^*)$ a la función $\sum_{i=1}^n \mathbf{A}_i(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))$ que es continuamente diferenciable por ser un vector compuesto por funciones continuamente diferenciables, se tiene que

$$\begin{aligned} \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_n, \Sigma_n, s_n(\mathbf{B}_n, \Sigma_n)) - \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) &= (F_{1,n}(\mathbf{B}_n^*, \Sigma_n^*), F_{2,n}(\mathbf{B}_n^*, \Sigma_n^*)) \\ &\times (\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*)), \end{aligned} \quad (2.18)$$

donde, $(F_{1,n}, F_{2,n})$ es una matriz cuyas columnas son las derivadas de $\sum_{i=1}^n \mathbf{A}_i(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))$ con respecto a $\text{vec}(\mathbf{B}^t)$ y $\text{vech}(\Sigma)$ con tamaño $(2pq + q(q+1) + 1) \times (pq + q(q+1)/2)$, y $(\mathbf{B}_n^*, \Sigma_n^*)$ es un valor que se encuentra entre (\mathbf{B}_n, Σ_n) y $(\mathbf{B}_0, \Sigma_0^*)$. Multiplicando la matriz $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n)$ por izquierda en ambos miembros de (2.18) se tiene que

$$\begin{aligned} \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_n, \Sigma_n, s_n(\mathbf{B}_n, \Sigma_n)) - \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) \\ = \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) (F_{1,n}(\mathbf{B}_n^*, \Sigma_n^*), F_{2,n}(\mathbf{B}_n^*, \Sigma_n^*)) (\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*)). \end{aligned} \quad (2.19)$$

Dado que (\mathbf{B}_n, Σ_n) cumple con el sistema de ecuaciones (2.4) se tiene que

$$\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_n, \Sigma_n, s_n(\mathbf{B}_n, \Sigma_n)) = \mathbf{0}.$$

Sustituyendo en (2.19) resulta,

$$\begin{aligned} -\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) \sum_{i=1}^n \mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*)) &= \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \Sigma_n) (F_{1,n}(\mathbf{B}_n^*, \Sigma_n^*), F_{2,n}(\mathbf{B}_n^*, \Sigma_n^*)) \\ &\times (\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*)) \\ &= \mathbf{V}_n (\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*)), \end{aligned} \quad (2.20)$$

con \mathbf{V}_n definida en (2.16). Premultiplicando a ambos lados de la ecuación (2.20) por $\sqrt{n}\mathbf{V}_n^\dagger$ con la inversa de Moore-Penrose definida en el Lema 2.1.3, se obtiene

$$-n\mathbf{V}_n^\dagger\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) = \sqrt{n}(\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}_0^*)).$$

Ahora, para los términos del lado izquierdo de esta igualdad se tienen los siguientes resultados:

- (a) Por el Teorema 8, la sucesión $n\mathbf{V}_n^\dagger$ converge casi seguramente a \mathbf{V}^\dagger definida en (B.67).
- (b) Por el Corolario 2.1.1, la sucesión de matrices $\mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ converge en probabilidad a $\mathbf{M}_{(C, D)}(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$ definida en (2.7).
- (c) Por el Teorema 7, $-\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$ converge en distribución a una normal con media $\mathbf{0}$ y matriz de covarianza (2.12).

Por lo tanto, la distribución asintótica de $(\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}_0^*))$ que es equivalente a escribir

$$\sqrt{n}\begin{pmatrix} \text{vec}(\mathbf{B}_n^t) - \text{vec}(\mathbf{B}_0^t) \\ \text{vech}(\boldsymbol{\Sigma}_n) - \text{vech}(\boldsymbol{\Sigma}_0^*) \end{pmatrix}$$

se obtiene a partir de (a), (b), (c) y el Teorema de Slutsky, obteniendo así que dicha distribución es normal con media

$$\mathbf{V}^\dagger\mathbf{M}_{(C, D)}(\mathbf{B}_0, \boldsymbol{\Sigma}_0)\mathbf{0} = \mathbf{0},$$

y matriz de covarianza \mathbf{W} definida en (2.17). □

Habiendo examinado previamente el Teorema 9, es posible modificar ligeramente las condiciones y aún obtener el mismo resultado. Bajo la condición A8), la propiedad (\mathbf{S}_ϵ) es verdadera. Esto implica que existe al menos una solución para (1.15), es decir, la sucesión de τ -estimadores existe. Además, según el Teorema 4, esta sucesión de τ -estimadores es fuertemente consistente. También, conforme al Teorema 2, esta sucesión satisface el sistema (2.4). Por lo tanto, tiene una distribución asintótica establecida en el Teorema 9. De este modo, el resultado del Teorema 9 puede ser obtenido bajo diferentes condiciones, y se expresa de la siguiente manera,

Teorema 10. *Sea $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Se asume que ρ_1 satisface A1)-A6) y A10), ρ_2 satisface A1)-A7) y A10), \mathbf{u}_i satisfacen A8) y \mathbf{x}_i satisfacen A9), para $i \geq 1$. Si $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ es una sucesión que satisface (1.15), con $r \neq 0$, entonces*

$$\sqrt{n}(\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}_0^*)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{W}),$$

con

$$\mathbf{W} = \mathbf{V}^\dagger\mathbf{M}_{(C, D)}(\mathbf{B}_0, \boldsymbol{\Sigma}_0)\text{Var}[\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))]\mathbf{M}_{(C, D)}^t(\mathbf{B}_0, \boldsymbol{\Sigma}_0)(\mathbf{V}^\dagger)^t, \quad (2.21)$$

donde $\mathbf{A}_i(\cdot)$, \mathbf{V}^\dagger , $\mathbf{M}_{(C, D)}(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$ y $\text{Var}[\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))]$ fueron definidas en (2.1), (B.67), (2.7) y (2.12), respectivamente.

El conocimiento de la distribución asintótica de los τ -estimadores tiene implicaciones significativas en la inferencia estadística. Al establecer la distribución asintótica, es posible deducir regiones de confianza y diseñar tests de hipótesis robustos. Estas regiones de confianza cuantifican la incertidumbre en las estimaciones de los parámetros, proporcionando intervalos donde los verdaderos valores probablemente se encuentren. Los tests de hipótesis basados en la distribución asintótica de los τ -estimadores evalúan la validez de suposiciones sobre los parámetros del modelo, facilitando la toma de decisiones basadas en datos.

De esta manera, los resultados obtenidos no sólo proporcionan una base teórica sólida para los τ -estimadores, sino que también abre la puerta a aplicaciones prácticas cruciales en la estadística inferencial, fortaleciendo la capacidad de los investigadores para realizar análisis precisos y fiables.

3 Distribución Asintótica (Caso: Rango Reducido)

3.1 Preliminares

En el ámbito del análisis de datos multivariados, los modelos lineales desempeñan un papel fundamental al permitir la comprensión de las relaciones entre múltiples variables predictoras y una variable respuesta. Sin embargo, cuando se trabaja con un gran número de variables, los modelos de rango completo pueden presentar diversas limitaciones, como la multicolinealidad y la alta dimensionalidad, que pueden conducir a estimaciones inestables y dificultades en la interpretación de los coeficientes.

Para abordar estas problemáticas, se ha desarrollado el concepto de modelos de rango reducido. Estos modelos parten del supuesto de que la matriz de coeficientes de regresión tiene un rango menor que el número total de variables predictoras y de respuesta. Esto implica que las relaciones entre las variables pueden ser representadas por un menor número de combinaciones lineales, lo que resulta en una reducción tanto en la dimensionalidad del problema como en el número de parámetros a estimar. Como bien indican [Velu and Reinsel \(2013\)](#), estos modelos ofrecen una alternativa poderosa para el análisis de datos multivariados al reducir la complejidad del problema y mejorar la interpretación de las relaciones entre las variables.

Los modelos de rango reducido se basan en la identificación de un subespacio de menor dimensión que captura la mayor parte de la variabilidad presente en las variables de respuesta, dado el conjunto de variables predictoras. Al imponer una estructura de rango reducido en la matriz de coeficientes, estos modelos no solo mitigan los efectos de la multicolinealidad, sino que también promueven una mayor estabilidad en las estimaciones y una mejor capacidad de generalización a nuevos datos.

La teoría asintótica para estimadores de rango reducido en modelos lineales multivariados se ha estudiado ampliamente bajo normalidad de los errores cuando el MLE es utilizado ([Bura and Cook, 2003](#); [Velu and Reinsel, 2013](#); [Cook et al., 2015](#); [Bura et al., 2018](#); [Zou et al., 2022](#)). El objetivo de este capítulo es definir y analizar el τ -estimador, cuando \mathbf{B}_0 tiene rango reducido. Para este objetivo se usará la notación y las derivadas obtenidas en el Capítulo 2, además de las ideas y resultados presentados por ([Bergesio et al., 2021](#); [Bura et al., 2018](#)).

En este capítulo se considera un modelo lineal multivariado de rango reducido (MLMRR), cuya formulación es análoga a la presentada en (1), con la particularidad de que la matriz, \mathbf{B}_0 , posee un rango $d < \min(p, q)$. Así, en adelante, se asumirá que el espacio Θ de los posibles parámetros para el MLMRR está definido como sigue:

$$\Theta_d = \{\boldsymbol{\theta} = (\mathbf{B}^t, \boldsymbol{\Sigma}) \in \mathbb{R}_d^{q \times p} \times S_+^q\}, \quad (3.1)$$

donde S_+^q representa la clase de todas las matrices simétricas definidas positivas de dimensión $q \times q$, y $\mathbb{R}_d^{q \times p}$ corresponde al conjunto de matrices $q \times p$ de rango d .

Según [Velu and Reinsel \(2013\)](#), la reducción en el rango de \mathbf{B}_0 tiene dos implicaciones claves:

1. Dado que $d < q$, la matriz \mathbf{B}_0^t tiene d filas linealmente independientes, lo que implica la existencia de $q - d$ restricciones lineales en la matriz de coeficientes \mathbf{B}_0^t de la forma $l_j^t \mathbf{B}_0^t = \mathbf{0}$, donde $j = 1, \dots, q - d$. Estas restricciones permiten modelar las combinaciones lineales $l_j^t \mathbf{y}$ sin utilizar al vector de variables predictoras \mathbf{x} , dependiendo únicamente de la distribución de \mathbf{u} .
2. La matriz \mathbf{B}_0^t puede expresarse como $\mathbf{B}_0^t = \mathbf{S}\mathbf{T}$, donde \mathbf{S} y \mathbf{T} son matrices $q \times d$ y $d \times p$, respectivamente, ambas de rango d . Las d columnas de \mathbf{S} pueden considerarse como una base para el espacio columna de \mathbf{B}_0^t , mientras que las d filas de \mathbf{T} constituyen una base para el espacio fila de \mathbf{B}_0^t . Por lo tanto, el modelo $\mathbf{y} = \mathbf{B}_0^t \mathbf{x} + \mathbf{u}$ puede reescribirse como $\mathbf{y} = \mathbf{S}(\mathbf{T}\mathbf{x}) + \mathbf{u}$, donde $\mathbf{T}\mathbf{x}$ tiene d componentes que corresponden a d combinaciones lineales del vector de predictores \mathbf{x} , las cuales pueden ser suficientes para modelar las variables respuestas.

Teniendo en cuenta las ideas presentadas por [Velu and Reinsel \(2013\)](#), es necesario exhibir de manera formal una forma de particionar la matriz \mathbf{B}_0^t , cuyo rango es $d < \min\{p, q\}$. Siguiendo a [Bura et al., \(2018\)](#), se introducen algunas definiciones y notaciones necesarias para tal fin.

Definición 3.1.1. *Se define la función*

$$g : \text{vec}(\mathbb{R}_d^{q \times d}) \times \text{vec}(\mathbb{R}_d^{d \times p}) \times \text{vech}(S_+^q) \rightarrow \text{vec}(\mathbb{R}_d^{q \times p}) \times \text{vech}(S_+^q) \\ (\text{vec}(\mathbf{S}), \text{vec}(\mathbf{T}), \text{vech}(\boldsymbol{\Sigma})) \mapsto (\text{vec}(\mathbf{S}\mathbf{T}), \text{vech}(\boldsymbol{\Sigma})).$$

Si $\text{vec}(\mathbf{S})$ y $\text{vec}(\mathbf{T})$ fuesen los únicos que determinaran a $\text{vec}(\mathbf{S}\mathbf{T})$, se procedería a realizar un procedimiento similar al propuesto en el [Capítulo 2](#). Sin embargo $\text{vec}(\mathbf{S}\mathbf{T}) = \text{vec}(\mathbf{S}\mathbf{A}\mathbf{A}^{-1}\mathbf{T})$ para cualquier matriz \mathbf{A} de tamaño $d \times d$ invertible. A continuación usando [\(Bura et al., 2018\)](#), se construye una parametrización de \mathbf{B}_0^t , que permite la unicidad de la descomposición. Vale la pena observar que aunque \mathbf{B}_0^t puede descomponerse de infinitas maneras, \mathbf{B}_0^t es siempre único.

Para ello, sin pérdida de generalidad, se asume de aquí en adelante que

$$\mathbf{B}_0^t \in \mathbb{R}_{first,d}^{q \times p} = \left\{ \mathbf{C} \in \mathbb{R}_d^{q \times p} : \text{las primeras } d \text{ filas de } \mathbf{C} \text{ son linealmente independientes} \right\}.$$

A nivel poblacional, esta representación se puede garantizar reordenando las coordenadas del vector respuesta \mathbf{y} .

En consecuencia, \mathbf{B}_0^t puede particionarse como

$$\mathbf{B}_0^t = \begin{pmatrix} \mathbf{R}_0 \\ \mathbf{Q}_0 \mathbf{R}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix} \mathbf{R}_0, \quad (3.2)$$

con $\mathbf{Q}_0 \in \mathbb{R}^{(q-d) \times d}$ y $\mathbf{R}_0 \in \mathbb{R}_d^{d \times p}$. Entonces tomando en cuenta esta manera de reescribir la matriz \mathbf{B}_0^t del MLMRR se presenta la siguiente definición.

Definición 3.1.2. Se definen las siguientes funciones

$$h : \text{vec}(\mathbb{R}^{(q-d) \times d}) \times \text{vec}(\mathbb{R}_d^{d \times p}) \times \text{vech}(S_+^q) \rightarrow \text{vec}(\mathbb{R}_{first,d}^{q \times p}) \times \text{vech}(S_+^q)$$

$$(\text{vec}(\mathbf{Q}), \text{vec}(\mathbf{R}), \text{vech}(\mathbf{\Sigma})) \mapsto \left(\text{vec} \begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}, \text{vech}(\mathbf{\Sigma}) \right),$$

y

$$h_M : \mathbb{R}^{(q-d) \times d} \times \mathbb{R}_d^{d \times p} \times S_+^q \rightarrow \mathbb{R}_{first,d}^{q \times p} \times S_+^q$$

$$(\mathbf{Q}, \mathbf{R}, \mathbf{\Sigma}) \mapsto \left(\begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}, \mathbf{\Sigma} \right).$$

Dado que $(\mathbf{B}_0^t, \mathbf{\Sigma}_0) \in \mathbb{R}_{first,d}^{q \times p} \times S_+^q$ es el parámetro del MLMRR, se denota como $(\mathbf{Q}_0, \mathbf{R}_0, \mathbf{\Sigma}_0) \in \mathbb{R}^{(q-d) \times d} \times \mathbb{R}_d^{d \times p} \times S_+^q$, el único punto que cumple con $h_M(\mathbf{Q}_0, \mathbf{R}_0, \mathbf{\Sigma}_0) = (\mathbf{B}_0^t, \mathbf{\Sigma}_0)$.

Lema 3.1.1. Las funciones h y h_M son biyectivas, bicontinuas y dos veces diferenciables.

La demostración de este Lema se encuentra en la Proposición 3.2 de [Bura et al. \(2018\)](#).

3.2 Definición, existencia y unicidad del τ -funcional de estimación para el MLMRR en los contextos poblacional y muestral

En esta sección se define el τ -estimador para el modelo lineal multivariado de rango reducido. Además, se cita el teorema que demuestra su existencia y se proporciona una explicación detallada de su formulación y aplicación dentro del contexto del modelo propuesto.

Definición 3.2.1. Un τ -funcional de estimación para el MLMRR está dado por el valor $(\tilde{\mathbf{B}}, \tilde{\mathbf{\Sigma}}) \in \Theta_d$ que resuelve el problema de minimización

$$(\tilde{\mathbf{B}}, \tilde{\mathbf{\Sigma}}) = \underset{(\mathbf{B}, \mathbf{\Sigma}) \in \Theta_d}{\text{argmin}} |\mathbf{\Sigma}|$$

sujeto a

$$\tau^2(H_{P, \mathbf{B}, \mathbf{\Sigma}}) = \kappa_2. \tag{3.3}$$

Observación 3.2.1. La biyectividad de la función h_M permite formular una versión equivalente a la Definición 3.2.1 en el espacio de parámetros definido por $\Theta_d^* = \mathbb{R}^{(q-d) \times d} \times \mathbb{R}_d^{d \times p} \times S_+^q$. Un τ -funcional de estimación para el MLMRR está dado por el valor $(\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \tilde{\mathbf{\Sigma}}) \in \Theta_d^*$ que resuelve el problema

$$(\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \tilde{\mathbf{\Sigma}}) = \underset{(\mathbf{Q}, \mathbf{R}, \mathbf{\Sigma}) \in \Theta_d^*}{\text{argmin}} |\mathbf{\Sigma}|$$

sujeto a

$$\tau^2(H_{P, \mathbf{Q}, \mathbf{R}, \mathbf{\Sigma}}) = \kappa_2. \tag{3.4}$$

En Szretter (2017) y Bergesio et al. (2021) se muestra la existencia del τ -funcional de estimación para el modelo Principal Fitted Components (PFC) (Ver Teorema 6.1 y Teorema 2, respectivamente) que puede verse como un modelo lineal multivariado de rango completo o rango reducido.

En el contexto de la presente tesis, el resultado se puede enunciar a través del siguiente teorema.

Teorema 11. *Sea (\mathbf{x}, \mathbf{y}) que cumple con el MLMRR y sea P su distribución. Si ρ_1 satisface A1)-A6), ρ_2 satisface A1)-A7) y P satisface (\mathcal{S}_ϵ) (Ver (1.9)), existe al menos un valor de $(\mathbf{B}, \Sigma) \in \Theta_d$ que cumple con las condiciones impuestas por el funcional $(\tilde{\mathbf{B}}, \tilde{\Sigma})$.*

La demostración se presenta en Szretter (2017) y Bergesio et al. (2021) y se basa esencialmente en el trabajo de (Lopuhaa, 1991) que establece la existencia de un τ -funcional en el contexto de estimar la posición y la matriz de covarianza para el modelo multivariado.

La atención se centra ahora en la versión muestral del τ -funcional de estimación. Esta definición permitirá un análisis de su existencia en el contexto de los datos muestrales, después de haber establecido su contraparte poblacional. Se presenta a continuación la definición del τ -funcional de estimación muestral para el modelo de rango reducido.

Definición 3.2.2. *El τ -funcional de estimación evaluado en la medida empírica para el MLMRR está definido como el valor $(\mathbf{B}_n, \Sigma_n) \in \Theta_d$ que resuelve el problema de minimización*

$$(\mathbf{B}_n, \Sigma_n) = \underset{(\mathbf{B}, \Sigma) \in \Theta_d}{\operatorname{argmin}} |\Sigma|$$

sujeto a (3.5)

$$\tau^2(H_{P_n, \mathbf{B}, \Sigma}) = \kappa_2.$$

De aquí en adelante cuando no haya ambigüedad, se referirá a un τ -estimador como un τ -funcional de estimación evaluado en la medida empírica.

Observación 3.2.2. *Nuevamente, se adapta la definición a los parámetros en Θ_d^* . Para ello, el τ -funcional de estimación evaluado en la medida empírica para el MLMRR está definido como el valor $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) \in \Theta_d^*$ que resuelve el problema de minimización*

$$(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) = \underset{(\mathbf{Q}, \mathbf{R}, \Sigma) \in \Theta_d^*}{\operatorname{argmin}} |\Sigma|$$

sujeto a (3.6)

$$\tau^2(H_{P_n, \mathbf{Q}, \mathbf{R}, \Sigma}) = \kappa_2.$$

Observación 3.2.3. *La situación de tener una muestra representa un caso especial del Teorema 11, lo que implica que la existencia de al menos una solución del problema (3.6) puede ser inferida directamente de dicho teorema. Si se considera $h_n < n\epsilon < n \left(1 - \frac{\kappa_1}{a_1}\right)$, donde h_n está definido en (1.12) lo cual garantiza que P_n está en las hipótesis del Teorema 11, entonces se asegura que existe al menos una solución del problema que define a $(\mathbf{B}_n, \Sigma_n) \in \Theta_d$ para P_n . Por lo tanto, una condición suficiente para la existencia del estimador (\mathbf{B}_n, Σ_n) de*

los parámetros del MLMRR es que las observaciones de la muestra $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, para $i = 1, \dots, n$, se encuentren en posición general.

Después de haber establecido la existencia de al menos una solución para el problema asociado al τ -estimador, se procede a verificar la unicidad de dicha solución. A continuación, se presenta el siguiente teorema que da condiciones suficientes para la unicidad del τ -funcional de estimación.

A continuación se presenta el caso particular de errores que cumplen A8) del Teorema 6.2 de Szretter (2017).

Teorema 12. (Szretter, 2017, Teorema 6.2) Sea (\mathbf{x}, \mathbf{y}) el vector aleatorio, con distribución P , que sigue el MLMRR con parámetros \mathbf{B}_0 y Σ_0 donde el error \mathbf{u} tiene una distribución que satisface A8). Se supone también que los ρ_k , $k = 1, 2$ satisfacen A1)-A7). Entonces el problema:

$$\begin{aligned} & \underset{(\mathbf{B}, \Sigma) \in \Theta_d}{\operatorname{argmin}} |\Sigma| \\ & \text{sujeto a} \\ & \tau^2(H_{P, \mathbf{B}, \Sigma}) = \kappa_2, \end{aligned}$$

tiene solución única, dada por $(\mathbf{B}_0, \Sigma_0^*)$.

La existencia y unicidad se han demostrado en términos de (\mathbf{B}, Σ) . A continuación, se explicará cómo estos resultados pueden emplearse para establecer la existencia y unicidad en términos de $(\mathbf{Q}, \mathbf{R}, \Sigma)$.

Observación 3.2.4. Se conoce por el Teorema 11, que existe al menos un valor $(\tilde{\mathbf{B}}^t, \tilde{\Sigma}) \in \mathbb{R}_{first,d}^{q \times p} \times \mathcal{S}_+^q$ tal que resuelve (3.3), entonces utilizando la biyectividad de la función $h_M(\cdot)$ (Ver Lema 3.1.1) existe al menos un elemento $(\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \tilde{\Sigma})$ en el dominio de dicha función, tal que $h_M(\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \tilde{\Sigma}) = (\tilde{\mathbf{B}}^t, \tilde{\Sigma})$ y por lo tanto resuelve el Problema (3.6). Ahora bien, la manera de encontrar los parámetros $(\tilde{\mathbf{Q}}, \tilde{\mathbf{R}}, \tilde{\Sigma})$ es a través de la inversa de la función $h_M(\cdot)$, que se define como

$$\begin{aligned} h_M^{-1} : \mathbb{R}_{first,d}^{q \times p} \times \mathcal{S}_+^q & \rightarrow \mathbb{R}^{(q-d) \times d} \times \mathbb{R}_d^{d \times p} \times \mathcal{S}_+^q \\ \left(\begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix}, \Sigma \right) & \mapsto (\mathbf{F}\mathbf{E}^t(\mathbf{E}\mathbf{E}^t)^{-1}, \mathbf{E}, \Sigma), \end{aligned}$$

ya que $\begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix} \in \mathbb{R}_{first,d}^{q \times p}$ con $\mathbf{E} \in \mathbb{R}_d^{d \times p}$, entonces existe $\mathbf{G} \in \mathbb{R}^{(q-d) \times d}$ tal que $\begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_d \\ \mathbf{G} \end{pmatrix} \mathbf{E}$. Por otro lado, por el Teorema 12, el problema (3.3) tiene única solución llamada $(\mathbf{B}_0, \Sigma_0^*)$, entonces a través de la función $h_M^{-1}(\cdot)$ existen únicos valores $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*) \in \Theta_d^*$ tales que $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*) = h_M^{-1}(\mathbf{B}_0^t, \Sigma_0^*)$, y por tanto es la única solución para el problema (3.6). En consecuencia, hasta el momento se tiene que existe una solución del problema (3.6) en $\mathbb{R}^{(q-d) \times d} \times \mathbb{R}_d^{d \times p} \times \mathcal{S}_+^q$ y además es única. En el caso de la muestra, se presenta una situación similar, donde por cada $(\mathbf{B}_n^t, \Sigma_n) \in \mathbb{R}_{first,d}^{q \times p} \times \mathcal{S}_+^q$ que resuelve el problema (3.5) hay un único $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) \in \Theta_d^*$ tal que resuelve el problema (3.6), $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) = h_M^{-1}(\mathbf{B}_n^t, \Sigma_n)$, donde $\mathbf{B}_n^t = \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_n \end{pmatrix} \mathbf{R}_n$.

La demostración de esta Observación se encuentra en el Apéndice C.1.

En las secciones siguientes de este capítulo, se profundizará en el análisis de la consistencia y la distribución asintótica de los τ -estimadores. Estos análisis son fundamentales para comprender el comportamiento y las propiedades de los estimadores en el contexto del MLMRR.

3.3 Consistencia del τ -estimador de $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$

Se analiza a continuación la consistencia de los τ -estimadores en el contexto del MLMRR, basándose en el Teorema 5 de Bergesio et al. (2021). Este teorema establece la consistencia del estimador (\mathbf{B}_n, Σ_n) como se describe a continuación,

Teorema 13. (Bergesio et al., 2021, Teorema 5) *Sea (\mathbf{x}, \mathbf{y}) el vector aleatorio con distribución P , que sigue el MLMRR con (\mathbf{B}_0, Σ_0) unívocamente definido y donde el error \mathbf{u} tiene una distribución que satisface A8). Se supone también que los ρ_k , $k = 1, 2$ satisfacen A1)-A7). Sea P_n una medida empírica obtenida desde la muestra $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$ de observaciones independientes e idénticamente distribuidas de P , entonces $\lim_{n \rightarrow \infty} (\mathbf{B}_n, \Sigma_n) = (\mathbf{B}_0, \Sigma_0^*)$ a.s. donde $(\mathbf{B}_n, \Sigma_n) \in \Theta_d^*$ es el τ -funcional de estimación evaluado en la medida empírica, definido en la Definición 3.2.2.*

Considerando lo expuesto anteriormente y recordando que $\mathbf{B}_0^t \in \mathbb{R}_{first,d}^{q \times p}$, al aprovechar la definición y las propiedades de la función $h_M(\cdot)$, se logra establecer la consistencia del MLMRR en función de la partición matricial correspondiente, tal como se detalla a continuación

Teorema 14. *Sea (\mathbf{x}, \mathbf{y}) el vector aleatorio con distribución P , que sigue el MLMRR con $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0) \in \Theta_d^*$ unívocamente definido, con $\mathbf{B}_0^t = \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix} \mathbf{R}_0$ y donde el error \mathbf{u} tiene una distribución que satisface A8). Se supone también que los ρ_k , $k = 1, 2$ satisfacen A1)-A7). Sea P_n una medida empírica obtenida desde la muestra $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$ de observaciones independientes e idénticamente distribuidas de P , entonces $\lim_{n \rightarrow \infty} (\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) = (\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$ a.s. donde $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) = h_M^{-1}(\mathbf{B}_n, \Sigma_n)$, y (\mathbf{B}_n, Σ_n) es el τ -funcional de estimación evaluado en la medida empírica, definido en la Definición 3.2.2.*

La demostración de este resultado se encuentra en el Apéndice C.2.

3.4 Ecuaciones de τ -estimaciones

Con la existencia confirmada, a continuación se presenta, de forma análoga al sistema (1.22), un sistema de ecuaciones que permite encontrar al τ -estimador para el modelo MLMRR. A diferencia del sistema de ecuaciones (1.22) expresado en términos de \mathbf{B} y Σ , el sistema de ecuaciones de las τ -estimaciones para el MLMRR ahora está expresado en términos de \mathbf{Q} , \mathbf{R} , y Σ . Por la Observación 3.2.3, se sabe que una condición suficiente para que el estimador

$(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ de los parámetros del MLMRR exista, es que las observaciones de la muestra $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$ estén en posición general. Para facilitar la notación de lo que sigue, se definen h_1 y h_2 por

$$h_M(\mathbf{Q}, \mathbf{R}, \Sigma) = \left(\begin{pmatrix} \mathbf{R} \\ \mathbf{Q}\mathbf{R} \end{pmatrix}, \Sigma \right) = (h_1(\mathbf{Q}, \mathbf{R}), h_2(\Sigma)). \quad (3.7)$$

Teorema 15. *Si las funciones ρ_1 y ρ_2 verifican las condiciones A1)-A7) y P_n satisface $\frac{h_n}{n} < 1 - \frac{\kappa_1}{a_1}$, existe al menos una solución $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) \in \Theta_d^*$ para el problema (3.6). Además, toda solución al problema (3.6) satisface las ecuaciones*

$$\left\{ \begin{array}{l} (\mathbf{R} \otimes (\mathbf{0}_{(q-d) \times d} \quad \mathbf{I}_{(q-d)}) \Sigma^{-1}) \sum_{i=1}^n \{\omega_n^*(d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) \text{vec}(\mathbf{u}_i(h_1(\mathbf{Q}, \mathbf{R})) \mathbf{x}_i^t)\} = 0, \\ (\mathbf{I}_p \otimes ((\mathbf{I}_d \quad \mathbf{Q}^t) \Sigma^{-1})) \sum_{i=1}^n \{\omega_n^*(d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) \text{vec}(\mathbf{u}_i(h_1(\mathbf{Q}, \mathbf{R})) \mathbf{x}_i^t)\} = 0, \\ \sum_{i=1}^n \{q\omega_n^*(d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) \text{vech}(\mathbf{u}_i(h_1(\mathbf{Q}, \mathbf{R})) \mathbf{u}_i^t(h_1(\mathbf{Q}, \mathbf{R}))) - s_n^2(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \psi_n^*(d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \text{vech}(\Sigma)\} = 0, \\ \sum_{i=1}^n \{s_n^2(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \rho_2(d_i^*(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) - \kappa_2\} = 0, \end{array} \right. \quad (3.8)$$

donde $\omega_n^*(\cdot)$, $\psi_n^*(\cdot)$ y $d_i^*(\cdot)$ están definidos en (1.17), (1.18) y (1.21), respectivamente; además $s_n(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))$ está definido en (1.11), con $\rho = \rho_1$ y $\kappa = \kappa_1$.

La demostración de este teorema se encuentra en el Apéndice C.3.

A continuación, se define una notación que se utilizará para presentar el sistema de ecuaciones de manera matricial para luego estudiar la distribución asintótica.

Definición 3.4.1. *Se define $\mathbf{G}_{(\mathbf{Q}, \mathbf{R}, \Sigma)}$ como la matriz*

$$\mathbf{G}_{(\mathbf{Q}, \mathbf{R}, \Sigma)} := \begin{pmatrix} (\mathbf{R} \otimes (\mathbf{0}_{(q-d) \times d} \quad \mathbf{I}_{(q-d)}) \Sigma^{-1}) & \mathbf{0} & \mathbf{0} \\ \mathbf{I}_p \otimes ((\mathbf{I}_d \quad \mathbf{Q}^t) \Sigma^{-1}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{q(q+1)/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix}, \quad (3.9)$$

de tamaño $[(q-d)d + dp + q(q+1)/2 + 1] \times [pq + q(q+1)/2 + 1]$, y que depende de los parámetros $(\mathbf{Q}, \mathbf{R}, \Sigma)$.

Lema 3.4.1. *Al considerar la matriz definida anteriormente, se tiene que el sistema (3.8) para el caso reducido se puede escribir de la siguiente forma matricial*

$$\mathbf{G}_{(\mathbf{Q}, \mathbf{R}, \Sigma)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}, \mathbf{R}, \Sigma), s_n(h_M(\mathbf{Q}, \mathbf{R}, \Sigma))) = \mathbf{0}, \quad (3.10)$$

con $\mathbf{A}_i(\cdot)$ y $\mathbf{M}_{(C_n, D_n)}(\cdot)$ definidos en (2.1) y (2.3), respectivamente.

Este lema es una consecuencia inmediata del Lema 2.1.1.

3.5 Normalidad asintótica del τ -estimador (rango reducido)

Como el problema de optimización depende ahora de \mathbf{Q} , \mathbf{R} , y Σ , entonces se requiere analizar la convergencia asintótica de los estimadores de estos tres parámetros, como se enuncia a continuación.

Teorema 16. *Sea $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, con $i \geq 1$ una sucesión de vectores aleatorios que satisface el MLMRR con parámetros $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0) \in \Theta_d^*$. Para cada n , sea $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) \in \Theta_d^*$ una solución del sistema de ecuaciones (3.8), o equivalentemente (3.10). Se supone que la sucesión $\{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)\}_{n \geq 1}$ converge a.s. a $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$. Si se verifican las condiciones A1)-A10), y $r \neq 0$, entonces $n^{1/2}(\text{vec}(\mathbf{Q}_n - \mathbf{Q}_0), \text{vec}(\mathbf{R}_n - \mathbf{R}_0), \text{vech}(\Sigma_n - \Sigma_0^*))$ converge a una distribución normal con media $\mathbf{0}$, y matriz de covarianza*

$$\mathbf{W}_R \text{Var}[\mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)))](\mathbf{W}_R)^t, \quad (3.11)$$

con

$$\mathbf{W}_R := (\mathbf{V}_R)^\dagger \mathbf{G}_{(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)} \mathbf{M}_{(C,D)}(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)), \quad (3.12)$$

y

$$\mathbf{V}_R := \mathbf{G}_{(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)} \mathbf{V} J_R(\mathbf{Q}_0, \mathbf{R}_0), \quad (3.13)$$

donde las matrices \mathbf{V} , $\mathbf{G}_{(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)}$ y $\mathbf{M}_{(C,D)}(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))$ fueron definidas en (B.66), (3.10), y (2.3), respectivamente. Además,

$$J_R(\mathbf{Q}_0, \mathbf{R}_0) := \begin{pmatrix} \mathbf{R}_0^t \otimes \begin{pmatrix} \mathbf{0}_{d \times (q-d)} \\ \mathbf{I}_{(q-d)} \end{pmatrix} & \mathbf{I}_p \otimes \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{q(q+1)/2} \end{pmatrix}. \quad (3.14)$$

La demostración de este teorema se encuentra en el Apéndice C.4.

Usando el Teorema anterior, se obtiene la convergencia asintótica

Teorema 17. *Sea $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, con $i \geq 1$ una sucesión de vectores aleatorios que satisface el MLMRR con parámetros Σ_0 y $\mathbf{B}_0 \in \mathbb{R}_{first,d}^{p \times q}$. Se supone que ρ_1 satisface A1)-A6) y A8)-A10), y ρ_2 satisface A1)-A10). Si (\mathbf{B}_n, Σ_n) es una sucesión de τ -estimadores, entonces $n^{1/2}(\text{vec}(\mathbf{B}_n^t - \mathbf{B}_0^t), \text{vech}(\Sigma_n - \Sigma_0^*))$ converge a una distribución normal con media $\mathbf{0}$, y matriz de covarianza*

$$J_R(\mathbf{Q}_0, \mathbf{R}_0) \mathbf{W}_R \text{Var}[\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)))](J_R(\mathbf{Q}_0, \mathbf{R}_0) \mathbf{W}_R)^t \quad (3.15)$$

donde $\mathbf{A}_i(\cdot)$, $J_R(\mathbf{Q}_0, \mathbf{R}_0)$ y \mathbf{W}_R fueron definidas en (2.1), (3.14) y (3.12), respectivamente.

Demostración del Teorema 17. Como la función $h(\cdot)$ definida en (3.1.2) es una función continua, dos veces diferenciable, por el Teorema 16, se tiene que $n^{1/2}(\text{vec}(\mathbf{Q}_n - \mathbf{Q}_0), \text{vec}(\mathbf{R}_n -$

\mathbf{R}_0), $vech(\boldsymbol{\Sigma}_n - \boldsymbol{\Sigma}_0^*)$) converge a una distribución normal con media cero y matriz de covarianza dada en (3.11). Entonces aplicando el método delta con $h(\cdot)$ resulta $n^{1/2}(h(\text{vec}(\mathbf{Q}_n), \text{vec}(\mathbf{R}_n), \text{vech}(\boldsymbol{\Sigma}_n)) - h(\text{vec}(\mathbf{Q}_0), \text{vec}(\mathbf{R}_0), \text{vech}(\boldsymbol{\Sigma}_0^*)))$ converge a una distribución normal con media cero y matriz de covarianza dada por

$$J_R(\mathbf{Q}_0, \mathbf{R}_0) \mathbf{W}_R \text{Var}[\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, k_0)] (\mathbf{W}_R)^t (J_R(\mathbf{Q}_0, \mathbf{R}_0))^t,$$

con

$$\begin{aligned}
 J_R(\mathbf{Q}_0, \mathbf{R}_0) &= \left(\begin{array}{ccc} \frac{\partial \text{vec}(h(\mathbf{Q}, \mathbf{R}, \boldsymbol{\Sigma}))}{\partial \text{vec}^t(\mathbf{Q})} & \frac{\partial \text{vec}(h(\mathbf{Q}, \mathbf{R}, \boldsymbol{\Sigma}))}{\partial \text{vec}^t(\mathbf{R})} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{q(q+1)/2} \end{array} \right) \Bigg|_{(\mathbf{Q}_0, \mathbf{R}_0, \boldsymbol{\Sigma}_0^*)} \\
 &= \left(\begin{array}{ccc} \mathbf{R}_0^t \otimes \begin{pmatrix} \mathbf{0}_{d \times (q-d)} \\ \mathbf{I}_{(q-d)} \end{pmatrix} & \mathbf{I}_p \otimes \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{q(q+1)/2} \end{array} \right).
 \end{aligned}$$

□

4 Estudios de Simulación

En este capítulo se exponen dos estudios de simulación destinados a mostrar el desempeño del τ -estimador y a contrastar su comportamiento con otros estimadores propuestos para el modelo de regresión multivariado, específicamente con otros métodos de estimación existentes que resultan robustos ante la presencia de datos que se apartan del modelo en el contexto de rango reducido. A continuación, se detallan los estimadores que serán objeto de comparación.

4.1 Métodos

El τ -estimador para el MLMRR (1) con matriz de coeficientes $\mathbf{B}_0^t \in \mathbb{R}_{first,d}^{q \times p}$, se compara con cuatro metodologías de estimación diferentes: el estimador de rango reducido clásico obtenido por máxima verosimilitud suponiendo que el término del error en el MLMRR tiene distribución normal (RRR), (Velu and Reinsel, 2013). El estimador robusto de rango completo (MM) presentado por Tatsuoka and Tyler (2000), dos estimadores robustos de rango reducido R4 analizados en She and Chen (2017) y R5 estudiado en Zhao et al. (2017). Además del estimador R5, se considera una variante modificada que se denotará como R5_d, la cual se deriva de R5 utilizando el rango que se obtiene mediante validación cruzada de 5-submuestras. Este estimador se desarrolló como parte del trabajo de esta tesis para abordar la limitación de R5, que necesita recibir el rango de la matriz \mathbf{B}_0 como parámetro de entrada. Para el τ -estimador se utiliza la implementación proporcionada por el paquete R: tauPFC (Ver (Bergesio et al., 2023)), el cual provee la posibilidad de encontrar el τ -estimador facilitándole la información del rango, denotado como Tau y la obtención del τ -estimador usando la estimación del rango de la matriz de coeficientes de regresión a través de validación cruzada de 5-submuestras para el modelo PFC (Bergesio et al., 2021), este estimador se denota como Tau_d.

Para obtener los τ -estimadores es necesario definir los valores de ciertas constantes relacionadas con las funciones ρ_1 y ρ_2 que se elijan para el cálculo de los estimadores. Dichas constantes, κ_i , m_i , $i = 1, 2$ definidas en ecuación (1.14) y condición A10) respectivamente, intervienen en la definición del punto de ruptura y la eficiencia asintótica del estimador. El paquete tauPFC permite obtener los valores de estas constantes para el máximo punto de ruptura 0.5 y la eficiencia deseada. Por otro lado, en el algoritmo que provee dicho paquete, se utiliza un estimador inicial para (\mathbf{B}_0, Σ_0) robusto con alto punto de ruptura y alta eficiencia, para el cual también es necesario definir las constantes κ_i , m_i , $i = 1, 2$.

Los niveles de eficiencia considerados se dividen en tres casos: el primero, con una eficiencia del 85 % tanto para el cálculo de los estimadores iniciales que requiere el algoritmo como para el τ -estimador; el segundo, con una eficiencia del 90 % para ambos estimadores; y finalmente, el tercer caso, con una eficiencia del 90 % para definir las constantes de los τ -estimadores y del 85 % para las estimaciones iniciales. Dado que los resultados muestran errores similares

en todos los casos, se presentan en las tablas los valores correspondientes al tercer caso.

4.2 Modelo Matricial

Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, vectores aleatorios independientes e idénticamente distribuidos que satisfacen el MLMRR con errores \mathbf{u}_i . El MLMRR se puede expresar de manera matricial como sigue,

$$\mathbf{Y} = \mathbf{X}\mathbf{B}_0 + \mathbf{U}, \quad (4.1)$$

donde \mathbf{X} es la matriz $n \times p$ cuya i -ésima fila es \mathbf{x}_i^t , \mathbf{Y} es la matriz $n \times q$ cuya i -ésima fila es \mathbf{y}_i^t , \mathbf{U} es la matriz $n \times q$ cuya i -ésima fila es $\mathbf{u}_i^t(\mathbf{B}_0) = \mathbf{y}_i^t - \mathbf{x}_i^t\mathbf{B}_0$ y $\mathbf{B}_0 = \mathbf{R}_0^t \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix}^t$ con $\mathbf{R}_0^t \in \mathbb{R}^{p \times d}$, $\begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix}^t \in \mathbb{R}^{d \times q}$, es decir $(\mathbf{Q}_0, \mathbf{R}_0, \boldsymbol{\Sigma}_0) \in \Theta_d^*$ definido en la Observación 3.2.1.

4.3 Medidas de error

En un estudio de simulación típico para evaluar estimadores, se repite el experimento de generar muestras que siguen un modelo, estimar los parámetros y comparar las características del estimador $(\hat{\mathbf{B}}, \hat{\boldsymbol{\Sigma}})$ con los valores verdaderos $(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$. Se denomina N a la cantidad de veces que se realiza este procedimiento. De esta manera, se obtiene un error en cada una de las N repeticiones del experimento, y luego estos errores se resumen en una única medida de error.

Se emplean dos medidas de error para evaluar el desempeño de los modelos: el *Error Cuadrático Medio recortado al 10%* y el *Error Cuadrático Medio Ponderado recortado al 10%*. Estas métricas se calculan utilizando la norma de Frobenius $\|\cdot\|_F$ para medir las distancias entre matrices. Cuando estas medidas se utilizan para analizar el comportamiento de proyecciones del tipo $\mathbf{X}\mathbf{B}$, se utiliza una nueva matriz de datos que se denomina \mathbf{X}_{test} generada sin contaminación bajo el escenario del experimento considerado. Esta estrategia garantiza una evaluación robusta de la capacidad predictiva de los estimadores en el contexto del MLMRR.

Error Cuadrático Medio (MSE por su sigla en inglés): Para cada una de las N repeticiones queda definida la cantidad MSE dada por

$$\frac{1}{qn_{test}} \|\mathbf{X}_{test}\mathbf{B}_0 - \mathbf{X}_{test}\hat{\mathbf{B}}\|_F^2, \quad (4.2)$$

donde n_{test} representa el número de filas que posee la matriz \mathbf{X}_{test} .

Error Cuadrático Medio recortado al 10%: De los N valores (4.2), se considera la media recortada al 10%. Este error se denota por $\text{Err}(\hat{\mathbf{B}})$.

Error Cuadrático Medio Ponderado: Para cada una de las N repeticiones se define el Error cuadrático medio ponderado como la cantidad:

$$\frac{1}{qn_{\text{test}}}\text{traza}\left((\mathbf{X}_{\text{test}}\mathbf{B}_0 - \mathbf{X}_{\text{test}}\widehat{\mathbf{B}})\boldsymbol{\Sigma}_0^{-1}(\mathbf{X}_{\text{test}}\mathbf{B}_0 - \mathbf{X}_{\text{test}}\widehat{\mathbf{B}})^t\right), \quad (4.3)$$

donde n_{test} representa el número de filas que posee la matriz \mathbf{X}_{test} y $\boldsymbol{\Sigma}_0$ la matriz de covarianza del error, descrita en cada escenario.

Error Cuadrático Medio Ponderado recortado al 10%: De los N valores (4.3), se considera la media recortada al 10%. Se denota por $\text{Err}(\widehat{\mathbf{B}}, \boldsymbol{\Sigma}_0)$.

4.4 Estudio de simulación 1

Los escenarios considerados en este primer estudio de simulación son los analizados por [She and Chen \(2017\)](#) y se describen a continuación.

Escenario I: Para este escenario en el modelo (4.1) se consideró $n = 100$, $p = 12$, $q = 8$ y rango de \mathbf{B}_0 igual a $d = 3$.

- Las matrices \mathbf{R}_0 y \mathbf{Q}_0 que forman la matriz de coeficientes \mathbf{B}_0 poseen elementos generados a partir muestras i.i.d. de una distribución $\mathcal{N}(0, 1)$.
- La matriz \mathbf{X} es construida generando n filas i.i.d. con distribución M_0 dada por $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_x)$, donde $\boldsymbol{\Sigma}_x$ es una matriz $p \times p$ con elementos 1 en la diagonal, y 0.5 fuera de ella.
- Las filas de la matriz de errores \mathbf{U} se generan mediante el siguiente procedimiento.

Primero, se obtiene una matriz \mathbf{D} de tamaño $n \times q$ donde cada una de la n filas se genera a partir de una distribución normal multivariada $\mathcal{N}(\mathbf{0}, \mathbf{I}_q)$. Posteriormente, se calcula el valor

$$\sigma = \frac{3}{4} \frac{[\text{SVD}(\mathbf{X}\mathbf{B}_0)]_d}{\sqrt{\sum_{i=1}^n \sum_{j=1}^q \mathbf{D}_{ij}^2}}, \quad (4.4)$$

donde $[\text{SVD}(\mathbf{X}\mathbf{B}_0)]_d$ es el d -ésimo valor singular (ordenados en forma decreciente) de la matriz $\mathbf{X}\mathbf{B}_0$, y \mathbf{D}_{ij} son los elementos de la matriz \mathbf{D} . Finalmente, se obtienen los valores de la matriz \mathbf{U} multiplicando cada elemento de \mathbf{D} por σ .

Escenario II: Este escenario es similar al *Escenario I*, pero en este caso, las filas de \mathbf{D} se generan a partir de una distribución $\mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})$, donde $\boldsymbol{\Lambda}$ es una matriz $q \times q$ con elementos 1 en la diagonal, y 0.5 fuera de ella. Así el *Escenario II* se diferencia del *Escenario I* únicamente por la matriz de covarianza de los errores.

Como consecuencia de cada escenario de simulación, se tomará la matriz $\boldsymbol{\Sigma}_0$ aproximada por $\sigma^2\boldsymbol{\Sigma}$ donde $\boldsymbol{\Sigma}$ es la matriz de covarianza de los vectores filas de la matriz \mathbf{D} , descrita en

cada escenario, ($\Sigma = \mathbf{I}_q$ en el Escenario I y $\Sigma = \mathbf{\Lambda}$ en el Escenario II) y σ está definido por (4.4). Además se observa que el denominador de (4.4) es la norma Frobenius de la matriz \mathbf{D} .

En este estudio de simulación, se busca comparar el desempeño de los estimadores propuestos para el MLMRR cuando los datos no se ajustan al modelo supuesto como verdadero. Para esto, se introducen dos tipos de contaminación que modifican una proporción de los datos generados según los modelos descritos en los *Escenarios I y II*. Además, una característica deseable de los estimadores robustos es su buen comportamiento cuando la muestra sigue el modelo establecido. Por esta razón, los resultados también incluyen los errores cuando los datos no están contaminados.

Contaminación I: Siguiendo la propuesta de [She and Chen \(2017\)](#), los outliers son introducidos por medio de una matriz \mathbf{C} donde sus primeras $O\%$ filas son no nulas, cada fila contiene al vector \mathbf{v} que en la coordenada j tiene a la desviación estándar de la columna j de $\mathbf{X}\mathbf{B}_0$, multiplicado por un valor elegido al azar en el conjunto $\{1, -1\}$ y también multiplicado por el escalar $\alpha \in \{2, 4\}$. Las restantes filas de \mathbf{C} son nulas. Se considera $O\% \in \{0\%, 5\%, 10\%, 15\%, 20\%\}$.

Para hacer el problema aún más desafiante, se reemplazan todos los elementos de las dos primeras filas de \mathbf{X} por 10, y se denota esta nueva matriz como \mathbf{X}_{10} . Esto produce valores atípicos con alto nivel de apalancamiento. El valor $O\% = 0\%$ indica que la matriz \mathbf{C} no está contaminada. Sin embargo, como \mathbf{Y} se genera calculando $\mathbf{Y} = \mathbf{X}_{10}\mathbf{B}_0 + \mathbf{U} + \mathbf{C}$ los datos presentan la contaminación que produce la matriz \mathbf{X}_{10} .

Contaminación II: La contaminación se produce utilizando el mismo método que se emplea en el caso de la *Contaminación I*, con la excepción de que en la matriz \mathbf{X} se reemplaza el $O\%$ de las primeras filas no nulas por el valor 10.

4.4.1 Resultados

Para los escenarios anteriores se realizaron $N = 200$ réplicas con $n_{test} = 100$. En cada escenario y para cada método se presentan diagramas de cajas de los errores cuadráticos medios (**MSE**) (4.2) que se obtienen en cada una de las N repeticiones. En las tablas que se presentan en esta sección, se incluyen los Errores Cuadráticos Medios recortados al 10%, los Errores Cuadráticos Medios Ponderados recortados al 10%, calculados siguiendo la metodología de ([She and Chen, 2017](#)). Además, se presentan los errores estándar recortados al 10% asociados a los errores cuadráticos medios (MSE), calculados también según ([She and Chen, 2017](#)), reportados en las ecuaciones (4.2) y (4.3). Estos errores estándar, son denotados como Estándar $\text{Err}(\hat{\mathbf{B}})$ y Estándar $\text{Err}(\hat{\mathbf{B}}, \Sigma_0)$, respectivamente. En la columna Rango se indica el rango promedio obtenido en las N repeticiones con los métodos **R4**, **R5_d** y **Tau_d**. En estos métodos el rango d es estimado por el algoritmo. Para los restantes estimadores, se reporta el rango que se ingresa como parámetro del ajuste. El método **MM** no hace reducción de la dimensión, entonces el rango es directamente el valor de $\min\{p, q\} = 8$.

En las Tablas 1 y 2, y en las Figuras 1 y 2, se muestran los errores utilizados para estudiar cuánto se alejan los estimadores robustos del óptimo en escenarios libres de contaminación.

Esto implica que los datos no sólo están exentos de la contaminación proporcionada por la matriz \mathbf{C} , sino también de cualquier alteración en las primeras dos filas de \mathbf{X} . Como es de esperar, bajo estas condiciones ideales, se observa que el estimador clásico RRR exhibe el mejor rendimiento en ambas tablas, siendo comparativamente superior a las demás propuestas. Para ambos escenarios sin contaminación, se observa un pobre desempeño del estimador MM que no tiene en cuenta la reducción del rango.

Las Tablas 3 a 10 y las Figuras 3 a 18, resumen los resultados para los escenarios I y II, y contaminación del tipo I y II, con $\alpha = 2$ y $\alpha = 4$. En lo que sigue se realiza un análisis detallado de cada escenario y nivel de contaminación, pero como resultado general puede observarse que, en presencia de contaminación, el estimador de máxima verosimilitud (RRR) no presenta un buen comportamiento. Debido a estos resultados y para una mejor visualización de los gráficos, este estimador no se ha incluido en los mismos (aunque sí en las tablas), al igual que los estimadores R5 y R5_d. En el caso de la contaminación II, solo se incluyó a R5, mientras que R5_d no se incluyó debido a que mostró resultados similares a R5 (Ver Figuras 11-18).

En contexto del escenario I, con contaminación I y $\alpha = 2$ (Tabla 3 y Figuras 3-4) los estimadores R5 y el estimador R5_d presentan los peores comportamientos cuando son comparados con otras propuestas robustas. Por otro lado, R4 presenta el mejor comportamiento y en el caso del 20 % de contaminación, se observa que el rendimiento de Tau es muy similar al obtenido para R4. También se ve en la contaminación del 20 % que para una replicación R4 obtiene un MSE de un orden de magnitud mayor a los restantes, lo cual desfavorece su performance. Para $\alpha = 4$ (Tabla 4 y Figuras 5-6) el comportamiento de los estimadores R4, Tau y Tau_d, es similar para las contaminaciones del 0 al 15 %, siendo siempre mejor el desempeño del R4. Para el caso del 20 % de contaminación, el R4 presenta un incremento notable en los errores, mientras que los errores de Tau y Tau_d se mantienen pequeños. Nuevamente, al aumentar la contaminación como puede verse en la Figura 6, la distribución del MSE del R4 presenta creciente cantidad de outliers cuyos valores también aumentan.

En el escenario II con contaminación I (Tablas 5-6 y Figuras 7-10) se observa que R5 y R5_d, como en el Escenario I con la misma contaminación, son los estimadores que mayor error y variabilidad presentan entre los estimadores robustos propuestos. Por otro lado, Tau y Tau_d se vuelven mejores competidores del R4, y consiguen mejores resultados para los mayores porcentajes de contaminación que se analizan. Nuevamente aparecen outliers en los errores del R4 en los casos de 15 % y 20 % de contaminación.

Cuando se usa la contaminación II para el escenario I (Tablas 7-8 y Figuras 11-14), los estimadores R5 y R5_d presentan mejores resultados que para el escenario de contaminación anterior, aunque su comportamiento se ve afectado con el aumento del nivel de contaminación. Por otro lado, el estimador R4 es el que presenta mejores resultados en este contexto, mostrando errores comparables con los τ -estimadores pero mostrando un comportamiento levemente mejor para niveles de contaminación más altos. Finalmente, el estimador MM presenta errores más elevados que el resto de los competidores.

En el caso del escenario II con contaminación II, para $\alpha = 2$ y $\alpha = 4$ (Tablas 9-10 y Figuras 15-18), los estimadores Tau y Tau_d exhiben errores cuadráticos medios más bajos que el resto

de los competidores presentando además, como también puede notarse en los escenarios anteriores, una gran estabilidad a medida que el nivel de contaminación aumenta.

Tabla 1. Resultados de simulación del Escenario I sin contaminación.

Estimadores	$\text{Err}(\hat{\mathbf{B}})$	Estándar $\text{Err}(\hat{\mathbf{B}})$	$\text{Err}(\hat{\mathbf{B}}, \Sigma_0)$	Estándar $\text{Err}(\hat{\mathbf{B}}, \Sigma_0)$	Rango
RRR	0.2156	0.1264	0.0750	0.0165	3.00
MM	0.4297	0.2564	0.1476	0.0260	8.00
R4	0.2251	0.1378	0.0772	0.0195	3.16
R5	0.2218	0.1310	0.0769	0.0166	3.00
$R5_d$	0.2289	0.1531	0.0756	0.0155	3.00
Tau	0.2739	0.1665	0.0948	0.0222	3.00
Tau_d	0.2777	0.1714	0.0950	0.0225	3.00

Figura 1 – Distribución de MSE sin contaminar para el Escenario I.

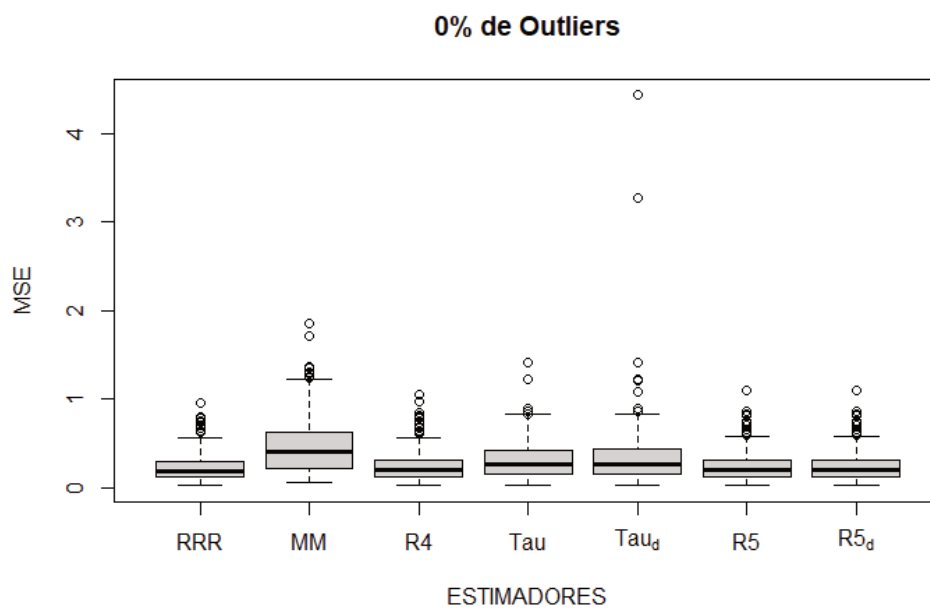


Tabla 2. Resultados de simulación del Escenario II sin contaminación.

Estimadores	$Err(\hat{\mathbf{B}})$	Estándar Err($\hat{\mathbf{B}}$)	$Err(\hat{\mathbf{B}}, \Sigma_0)$	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
RRR	0.1617	0.1015	0.0733	0.0167	3.00
MM	0.4159	0.2546	0.1470	0.0268	8.00
R4	0.2329	0.1534	0.0916	0.0225	3.11
R5	0.2192	0.1397	0.0968	0.0238	3.00
R5 _d	0.2198	0.1523	0.0969	0.0241	3.00
Tau	0.2087	0.1369	0.0947	0.0232	3.00
Tau _d	0.2086	0.1362	0.0947	0.0235	3.00

Figura 2 – Distribución de MSE sin contaminar para el Escenario II.

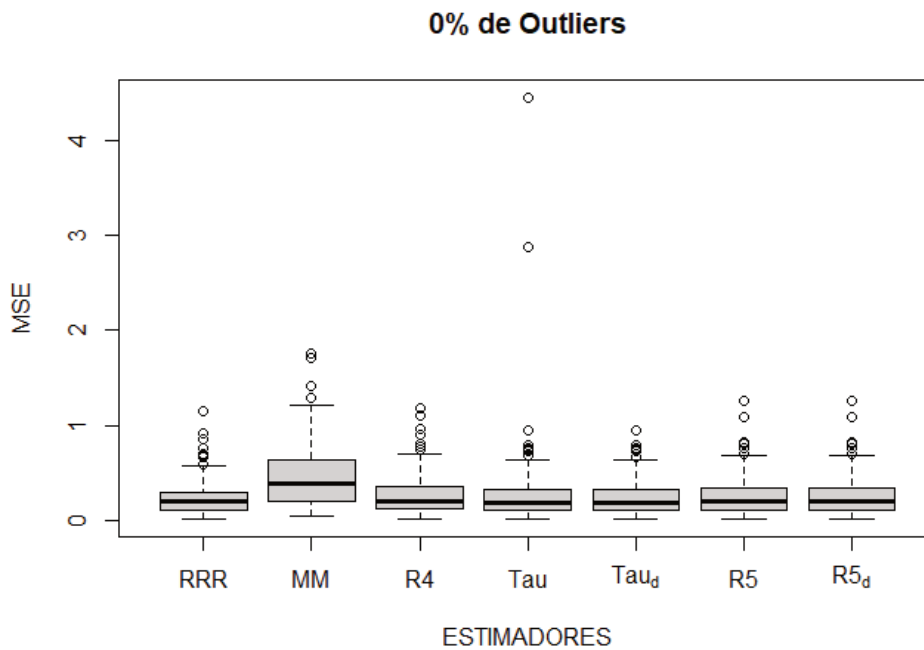


Tabla 3. Resultados de simulación del Escenario I con contaminación I y $\alpha = 2$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
0 %	RRR	51.4735	45.7170	20.2645	19.5846	2.05
	MM	0.5484	0.3588	0.1785	0.0364	8.00
	R4	0.3035	0.1980	0.0990	0.0237	3.00
	R5	4.4187	3.0101	1.4900	0.5670	3.00
	R5 _d	4.4179	2.6927	1.4281	0.5079	4.94
	Tau	0.3924	0.2565	0.1284	0.0350	3.00
	Tau _d	0.3965	0.2612	0.1287	0.0355	3.00
5 %	RRR	46.2563	42.9088	18.5399	18.9936	1.91
	MM	0.5589	0.3639	0.1825	0.0369	8.00
	R4	0.3113	0.2038	0.1015	0.0242	3.01
	R5	4.8810	3.3268	1.7000	0.7518	3.00
	R5 _d	5.0925	3.4143	1.6298	0.6725	4.81
	Tau	0.3954	0.2595	0.1296	0.0343	3.00
	Tau _d	0.4062	0.2731	0.1306	0.0352	3.00
10 %	RRR	49.9783	45.3147	19.9959	19.4862	1.32
	MM	0.5813	0.3701	0.1921	0.0398	8.00
	R4	0.3272	0.2130	0.1072	0.0256	3.00
	R5	5.4429	3.8314	1.8883	0.9301	3.00
	R5 _d	5.9823	4.0109	1.9279	0.8371	4.79
	Tau	0.4020	0.2583	0.1313	0.0326	3.00
	Tau _d	0.4108	0.2614	0.1322	0.0342	3.00
15 %	RRR	55.5951	49.7756	21.8839	20.9081	1.18
	MM	0.6082	0.3880	0.2007	0.0422	8.00
	R4	0.3480	0.2257	0.1146	0.0291	3.03
	R5	6.9620	5.2826	2.4089	1.3801	3.00
	R5 _d	8.1206	5.8288	2.5982	1.3654	4.66
	Tau	0.4097	0.2607	0.1347	0.0334	3.00
	Tau _d	0.4269	0.2812	0.1365	0.0366	3.00
20 %	RRR	58.4286	52.2743	23.1537	22.1003	1.10
	MM	0.6474	0.4169	0.2131	0.0455	8.00
	R4	0.4215	0.2976	0.1302	0.0360	2.97
	R5	10.1101	8.7684	3.5477	2.7901	3.00
	R5 _d	12.2868	10.8691	3.8815	2.5212	4.41
	Tau	0.4267	0.2759	0.1403	0.0346	3.00
	Tau _d	0.5305	0.4098	0.1866	0.1818	2.930

Figura 3 – Distribución de MSE para el Escenario I con contaminación I (5% – 10%) y $\alpha = 2$.

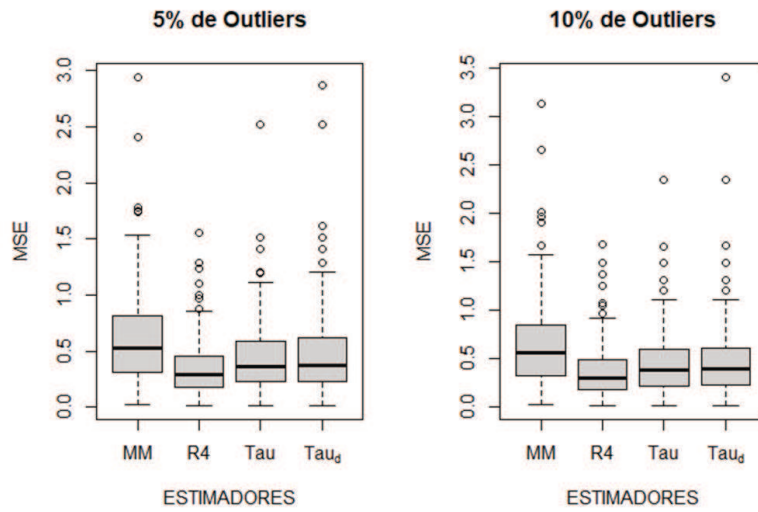


Figura 4 – Distribución de MSE para el Escenario I con contaminación I (15% – 20%) y $\alpha = 2$.

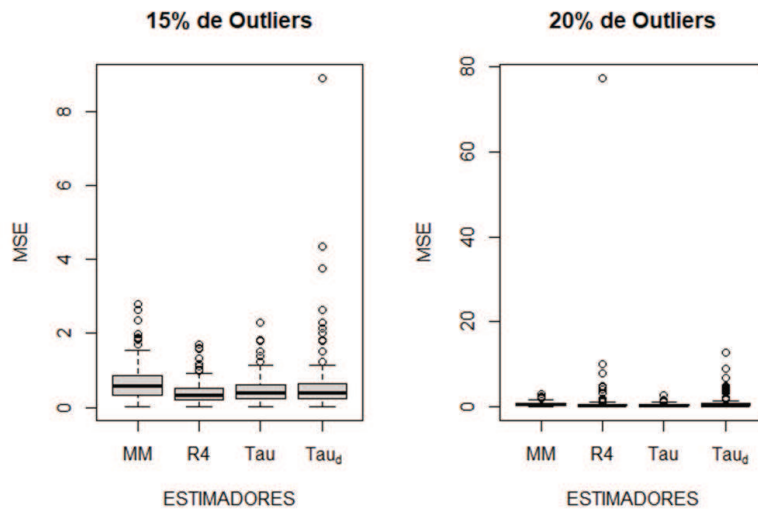


Tabla 4. Resultados de simulación del Escenario I con contaminación I y $\alpha = 4$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
0 %	RRR	13.3286	11.6588	4.4537	4.3186	2.10
	MM	0.5097	0.3273	0.1499	0.0243	8.00
	R4	0.2408	0.1561	0.0708	0.0169	3.00
	R5	1.3853	0.9083	0.4292	0.1638	3.00
	R5 _d	1.4256	0.7930	0.4823	0.1703	4.81
	Tau	0.3194	0.2055	0.0940	0.0229	3.00
	Tau _d	0.3220	0.2075	0.0944	0.0237	3.00
5 %	RRR	20.7278	18.9521	6.9431	7.1322	1.08
	MM	0.5280	0.3376	0.1552	0.0251	8.00
	R4	0.2512	0.1557	0.0750	0.0177	3.00
	R5	1.2535	0.8379	0.3897	0.1851	3.00
	R5 _d	1.4664	0.9177	0.4946	0.2195	4.77
	Tau	0.3266	0.2025	0.0971	0.0236	3.00
	Tau _d	0.3299	0.2033	0.0980	0.0242	3.00
10 %	RRR	28.7548	25.9513	10.0128	10.5007	1.08
	MM	0.5514	0.3545	0.1631	0.0274	8.00
	R4	0.2708	0.1760	0.0803	0.0198	2.98
	R5	1.6964	1.0882	0.5230	0.2429	3.00
	R5 _d	2.0899	1.2921	0.7133	0.3181	4.72
	Tau	0.3370	0.2115	0.1011	0.0251	3.00
	Tau _d	0.3542	0.2313	0.1013	0.0262	3.00
15 %	RRR	40.5510	36.5203	14.4711	15.9426	1.01
	MM	0.5810	0.3719	0.1711	0.0298	8.00
	R4	0.3255	0.2373	0.0895	0.0260	2.94
	R5	2.2389	1.4788	0.7034	0.4343	3.00
	R5 _d	2.6924	1.8659	0.9057	0.5230	4.54
	Tau	0.3445	0.2181	0.1036	0.0261	3.00
	Tau _d	0.3876	0.2719	0.1077	0.0301	3.00
20 %	RRR	52.3554	46.6334	18.7685	19.6109	1.01
	MM	0.6124	0.3862	0.1807	0.0329	8.00
	R4	2.2766	6.0726	0.6987	1.9983	2.75
	R5	3.8979	3.3630	1.1954	1.0639	3.00
	R5 _d	4.5622	3.2901	1.5964	1.1033	4.50
	Tau	0.3577	0.2208	0.1077	0.0271	3.00
	Tau _d	0.5375	0.4663	0.1733	0.1988	2.88

Figura 5 – Distribución de MSE para el Escenario I con contaminación I (5% – 10%) y $\alpha = 4$.

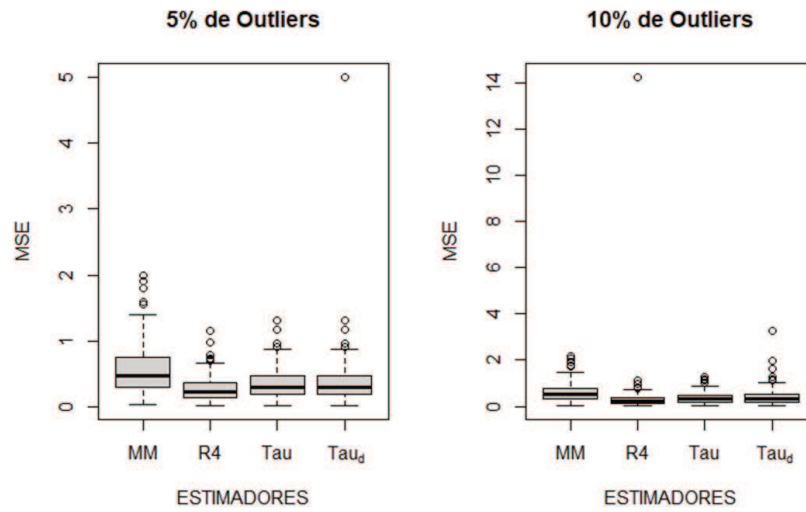


Figura 6 – Distribución de MSE para el Escenario I con contaminación I (15% – 20%) y $\alpha = 4$.

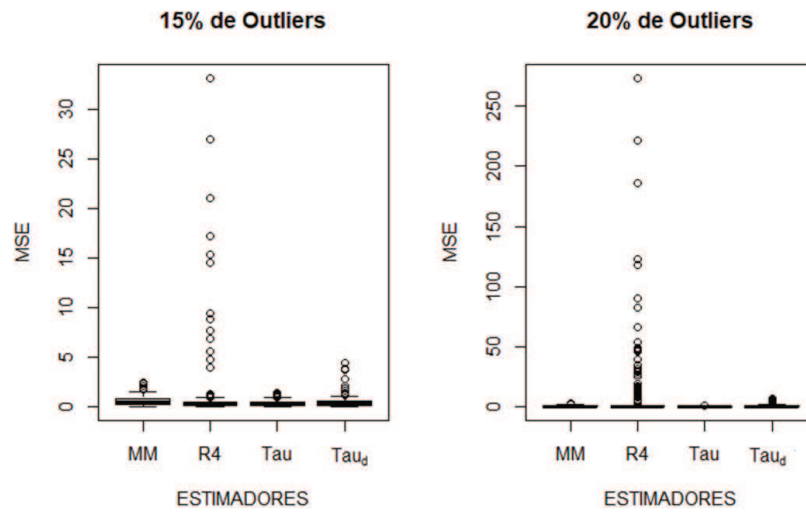


Tabla 5. Resultados de simulación del Escenario II con contaminación I y $\alpha = 2$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
0 %	RRR	13.2710	12.6121	9.0283	9.9289	2.08
	MM	0.4302	0.2726	0.1492	0.0272	8.00
	R4	0.2094	0.1382	0.0898	0.0212	3.00
	R5	1.2193	0.8243	0.7071	0.3147	3.00
	R5 _d	1.2391	0.7076	0.7546	0.3369	4.93
	Tau	0.2161	0.1386	0.0966	0.0206	3.00
	Tau _d	0.2169	0.1394	0.0968	0.0211	3.00
5 %	RRR	13.8040	13.3848	9.3036	10.1059	1.87
	MM	0.4455	0.2860	0.1528	0.0261	8.00
	R4	0.2222	0.1526	0.0938	0.0220	3.00
	R5	1.4163	0.9664	0.8325	0.4253	3.00
	R5 _d	1.4979	0.9067	0.9389	0.4977	4.80
	Tau	0.2194	0.1424	0.0985	0.0208	3.00
	Tau _d	0.2209	0.1444	0.0991	0.0209	3.00
10 %	RRR	15.9741	15.0865	10.7585	11.3799	1.44
	MM	0.4655	0.2971	0.1607	0.0294	8.00
	R4	0.2388	0.1666	0.1007	0.0245	3.02
	R5	1.6292	1.1470	0.9549	0.4840	3.00
	R5 _d	1.7928	1.1513	1.1199	0.6036	4.78
	Tau	0.2301	0.1557	0.1022	0.0242	3.00
	Tau _d	0.2292	0.1542	0.1020	0.0244	3.00
15 %	RRR	19.3379	17.6435	13.1157	13.2025	1.16
	MM	0.4981	0.3216	0.1686	0.0265	8.00
	R4	0.2574	0.1809	0.1078	0.0261	3.02
	R5	2.1384	1.5603	1.2326	0.7017	3.00
	R5 _d	2.4083	1.6352	1.5743	0.9796	4.68
	Tau	0.2349	0.1563	0.1037	0.0229	3.00
	Tau _d	0.2349	0.1567	0.1036	0.0233	3.00
20 %	RRR	22.7457	20.5895	15.2833	15.5061	1.10
	MM	0.5463	0.3579	0.1817	0.0326	8.00
	R4	0.3309	0.2436	0.1223	0.0360	3.01
	R5	3.4490	2.8635	2.0356	1.5919	3.00
	R5 _d	3.3896	2.4894	2.2589	1.5740	4.42
	Tau	0.2517	0.1713	0.1092	0.0240	3.00
	Tau _d	0.2641	0.1858	0.1104	0.0261	3.00

Figura 7 – Distribución de MSE para el Escenario II con contaminación I (5% – 10%) y $\alpha = 2$.

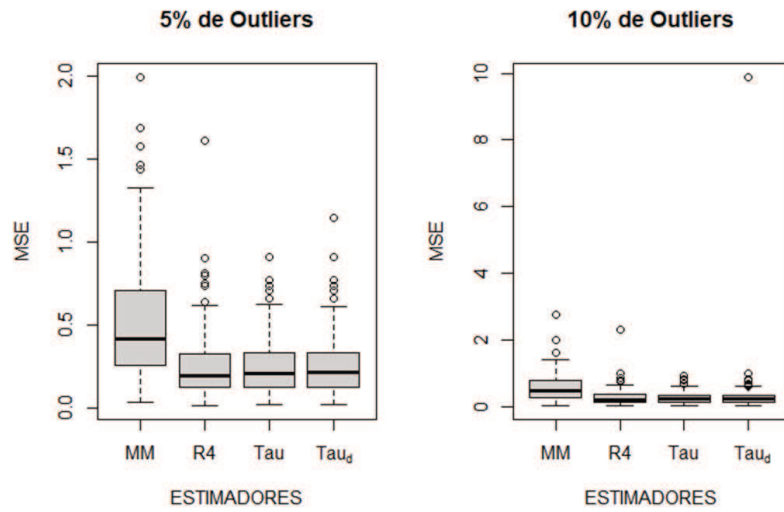


Figura 8 – Distribución de MSE para el Escenario II con contaminación I (15% – 20%) y $\alpha = 2$.

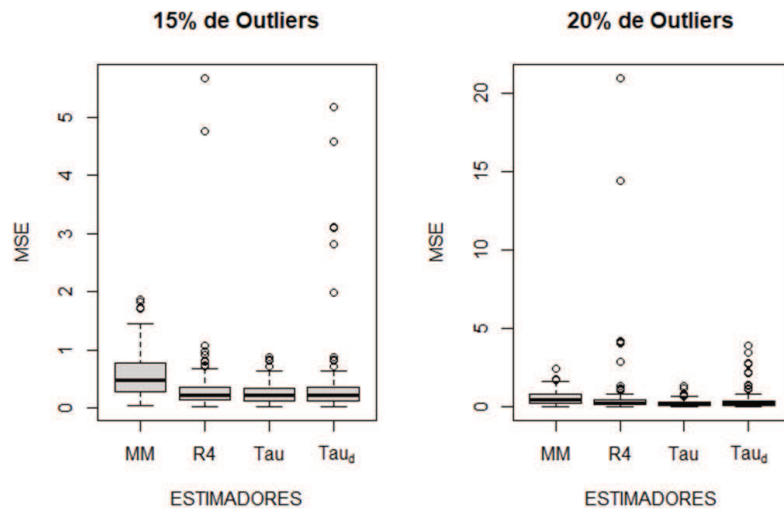


Tabla 6. Resultados de simulación del Escenario II con contaminación I y $\alpha = 4$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
0 %	RRR	12.4392	11.4065	9.7440	10.2706	2.11
	MM	0.4401	0.3315	0.1507	0.0251	8.00
	R4	0.2124	0.1597	0.0909	0.0241	3.00
	R5	1.2659	0.9440	0.7513	0.3154	3.00
	R5 _d	1.6147	1.004	0.7205	0.3218	4.80
	Tau	0.2142	0.1574	0.0962	0.0251	3.00
	Tau _d	0.2143	0.1577	0.0959	0.0250	3.00
5 %	RRR	18.2494	17.6228	14.5277	15.5410	1.40
	MM	0.4550	0.3476	0.1543	0.0265	8.00
	R4	0.2209	0.1747	0.0933	0.0246	3.00
	R5	1.1582	0.9242	0.6864	0.3578	3.00
	R5 _d	1.7033	1.045	0.7701	0.3667	4.60
	Tau	0.2159	0.1593	0.0963	0.0235	3.00
	Tau _d	0.2158	0.1601	0.0958	0.0232	3.00
10 %	RRR	27.0153	24.0243	20.3230	20.1388	1.08
	MM	0.4709	0.3523	0.1624	0.0279	8.00
	R4	0.2402	0.1923	0.1001	0.0250	3.00
	R5	1.4910	1.1480	0.8731	0.4249	3.00
	R5 _d	2.085	1.5213	1.0056	0.5966	4.60
	Tau	0.2210	0.1689	0.0980	0.0242	3.00
	Tau _d	0.2271	0.1751	0.0985	0.0249	3.00
15 %	RRR	37.2504	34.5223	28.7597	30.1951	1.01
	MM	0.4977	0.3746	0.1704	0.0303	8.00
	R4	0.3075	0.2754	0.1130	0.0365	2.95
	R5	1.9367	1.5434	1.1260	0.6429	3.00
	R5 _d	3.7025	2.7498	1.8744	1.3195	4.70
	Tau	0.2306	0.1776	0.1009	0.0238	3.00
	Tau _d	0.2332	0.1810	0.1015	0.0243	3.00
20 %	RRR	48.1254	42.2796	36.5055	36.3531	1.00
	MM	0.5246	0.3976	0.1823	0.0312	8.00
	R4	1.3117	3.1135	0.7088	1.9243	2.83
	R5	3.3309	3.1210	2.0527	1.8553	3.00
	R5 _d	5.5494	4.6433	2.8235	2.2107	4.55
	Tau	0.2446	0.1874	0.1059	0.0280	3.00
	Tau _d	0.2737	0.2102	0.1107	0.0325	3.00

Figura 9 – Distribución de MSE para el Escenario II con contaminación I (5% – 10%) y $\alpha = 4$.

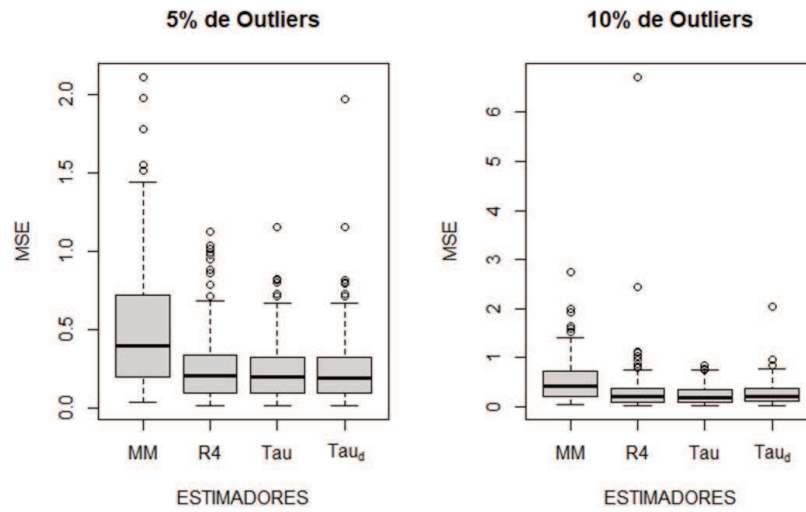


Figura 10 – Distribución de MSE para el Escenario II con contaminación I (15% – 20%) y $\alpha = 4$.

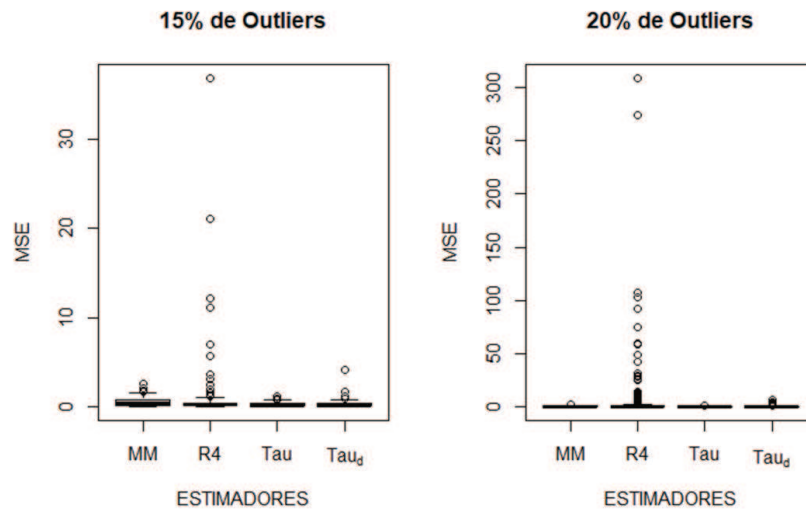


Tabla 7. Resultados de simulación del Escenario I con contaminación II y $\alpha = 2$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
5 %	RRR	0.4750	0.2894	0.1726	0.0932	2.60
	MM	0.4465	0.2631	0.1542	0.0280	8.00
	R4	0.2182	0.1258	0.0762	0.0171	3.00
	R5	0.2824	0.1714	0.0980	0.0246	3.00
	R5 _d	0.2792	0.1703	0.0972	0.0222	3.00
	Tau	0.2766	0.1650	0.0961	0.0221	3.00
	Tau _d	0.2860	0.1768	0.0972	0.0231	3.00
10 %	RRR	0.8414	0.5329	0.3277	0.2326	1.96
	MM	0.4683	0.2787	0.1613	0.0290	8.00
	R4	0.2315	0.1361	0.0801	0.0172	3.00
	R5	0.3059	0.1834	0.1063	0.0277	3.00
	R5 _d	0.3199	0.1971	0.1104	0.0303	3.00
	Tau	0.2859	0.1730	0.0990	0.0222	3.00
	Tau _d	0.3052	0.1880	0.1012	0.0246	3.00
15 %	RRR	1.4267	0.9089	0.5748	0.4380	1.48
	MM	0.4969	0.2904	0.1727	0.0299	8.00
	R4	0.2480	0.1478	0.0864	0.0201	3.00
	R5	0.4056	0.2630	0.1408	0.0475	3.00
	R5 _d	0.4048	0.2491	0.1419	0.0444	3.00
	Tau	0.2967	0.1841	0.1024	0.0249	3.00
	Tau _d	0.3192	0.2085	0.1050	0.0275	3.00
20 %	RRR	2.2312	1.5344	0.8712	0.6484	1.24
	MM	0.5276	0.3103	0.1830	0.0302	8.00
	R4	0.2655	0.1606	0.0924	0.0217	3.00
	R5	0.4848	0.3301	0.1684	0.0575	3.00
	R5 _d	0.4711	0.2974	0.1666	0.0597	3.00
	Tau	0.3057	0.1882	0.1061	0.0247	3.00
	Tau _d	0.4300	0.4064	0.1552	0.1766	2.92

Figura 11 – Distribución de MSE para el Escenario I con contaminación II (5% – 10%) y $\alpha = 2$.

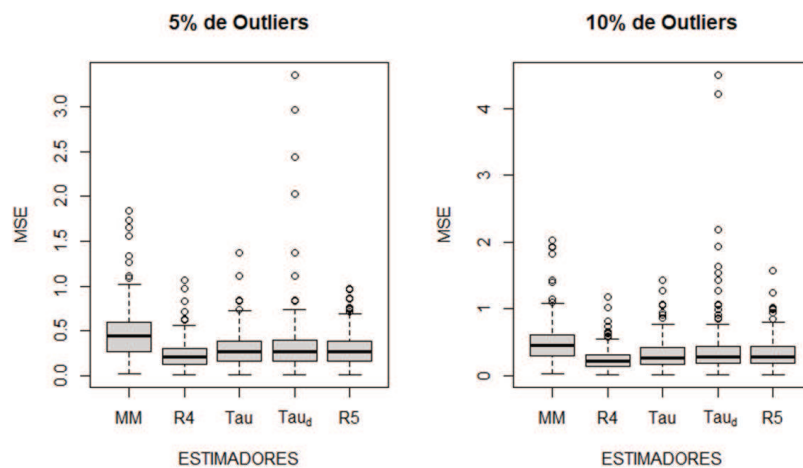


Figura 12 – Distribución de MSE para el Escenario I con contaminación II (15% – 20%) y $\alpha = 2$.

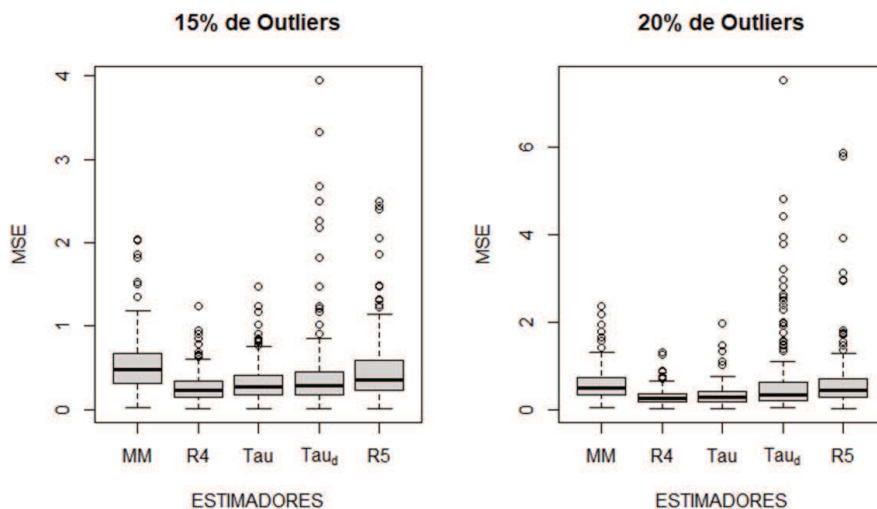


Tabla 8. Resultados de simulación del Escenario I con contaminación II y $\alpha = 4$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
5 %	RRR	1.0286	0.7263	0.3561	0.2625	1.68
	MM	0.4863	0.3272	0.1535	0.0285	8.00
	R4	0.2304	0.1548	0.0744	0.0196	3.00
	R5	0.2977	0.2081	0.0957	0.0252	3.00
	R5 _d	0.2807	0.1829	0.0954	0.0252	3.00
	Tau	0.3020	0.2037	0.0965	0.0249	3.00
	Tau _d	0.3089	0.2132	0.0980	0.0263	3.00
10 %	RRR	2.4384	1.7729	0.8883	0.7395	1.15
	MM	0.5017	0.3332	0.1600	0.0283	8.00
	R4	0.2431	0.1591	0.0789	0.0196	3.00
	R5	0.3353	0.2305	0.1071	0.0315	3.00
	R5 _d	0.3162	0.2146	0.1072	0.0296	3.00
	Tau	0.3070	0.2051	0.0984	0.0231	3.00
	Tau _d	0.3174	0.2218	0.0997	0.0246	3.00
15 %	RRR	4.9032	3.7180	1.9359	1.7642	1.03
	MM	0.5269	0.3480	0.1689	0.0329	8.00
	R4	0.2561	0.1675	0.0827	0.0211	3.00
	R5	0.4591	0.3329	0.1461	0.0502	3.00
	R5 _d	0.4180	0.2843	0.1399	0.0426	3.00
	Tau	0.3145	0.2050	0.1018	0.0260	3.00
	Tau _d	0.3514	0.2397	0.1053	0.0307	3.00
20 %	RRR	7.4258	5.4030	2.9218	2.5038	1.05
	MM	0.5618	0.3782	0.1776	0.0343	8.00
	R4	0.2790	0.1842	0.0888	0.0223	2.99
	R5	0.5470	0.3749	0.1779	0.0690	3.00
	R5 _d	0.5313	0.3922	0.1779	0.0692	3.00
	Tau	0.3260	0.2135	0.1046	0.0269	3.00
	Tau _d	0.4379	0.3651	0.1411	0.1470	2.94

Figura 13 – Distribución de MSE para el Escenario I con contaminación II (5% – 10%) y $\alpha = 4$.

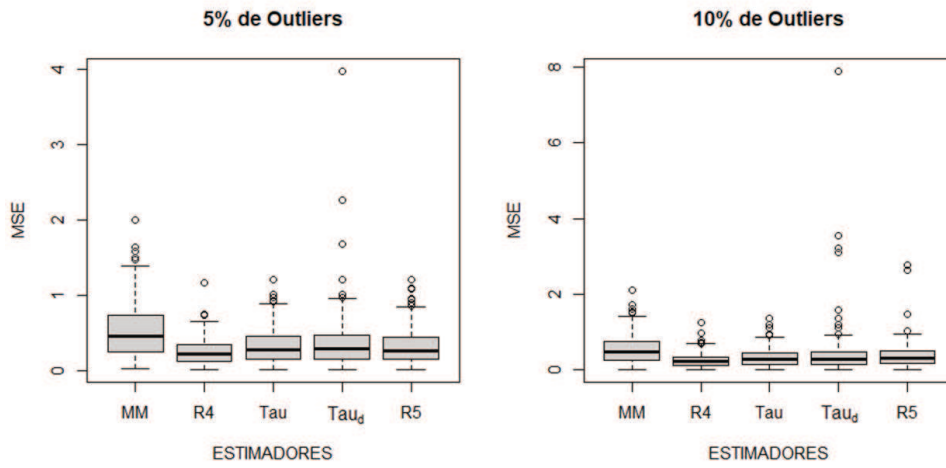


Figura 14 – Distribución de MSE para el Escenario I con contaminación II (15% – 20%) y $\alpha = 4$.

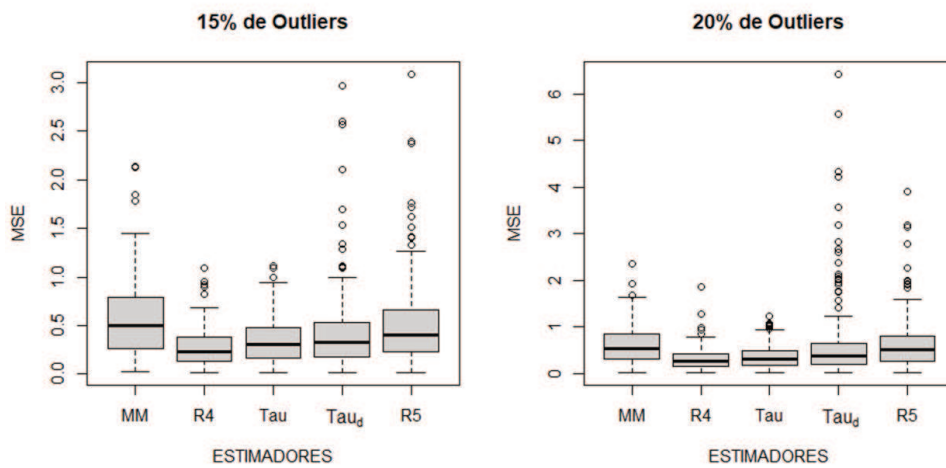


Tabla 9. Resultados de simulación del Escenario II con contaminación II y $\alpha = 2$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
5 %	RRR	0.3883	0.2455	0.2380	0.1599	2.56
	MM	0.4289	0.2616	0.1529	0.0272	8.00
	R4	0.2162	0.1353	0.0934	0.0223	3.02
	R5	0.2759	0.1797	0.1311	0.0386	3.00
	R5 _d	0.2760	0.1793	0.1277	0.0359	3.00
	Tau	0.2078	0.1284	0.0959	0.0219	3.00
	Tau _d	0.2088	0.1299	0.0965	0.0221	3.00
10 %	RRR	0.7745	0.5171	0.5163	0.4115	1.94
	MM	0.4466	0.2689	0.1592	0.0285	8.00
	R4	0.2284	0.1454	0.0985	0.0229	3.02
	R5	0.3042	0.1955	0.1481	0.0419	3.00
	R5 _d	0.3106	0.2021	0.1499	0.0443	3.00
	Tau	0.2112	0.1338	0.0978	0.0222	3.00
	Tau _d	0.2148	0.1390	0.0987	0.0235	3.00
15 %	RRR	1.3027	0.8412	0.9693	0.8594	1.48
	MM	0.4768	0.2913	0.1707	0.0296	8.00
	R4	0.2487	0.1600	0.1084	0.0264	3.02
	R5	0.4147	0.2804	0.2142	0.0781	3.00
	R5 _d	0.4059	0.2721	0.2056	0.0719	3.00
	Tau	0.2172	0.1346	0.1023	0.0244	3.00
	Tau _d	0.2185	0.1377	0.1026	0.0245	3.00
20 %	RRR	2.1246	1.4950	1.4995	1.2396	1.25
	MM	0.5054	0.3125	0.1818	0.0295	8.00
	R4	0.2748	0.1875	0.1157	0.0309	3.04
	R5	0.5000	0.3625	0.2656	0.1075	3.00
	R5 _d	0.5199	0.3746	0.2631	0.1121	3.00
	Tau	0.2298	0.1498	0.1063	0.0258	3.00
	Tau _d	0.2587	0.1868	0.1100	0.0300	3.00

Figura 15 – Distribución de MSE para el Escenario II con contaminación II (5% – 10%) y $\alpha = 2$.

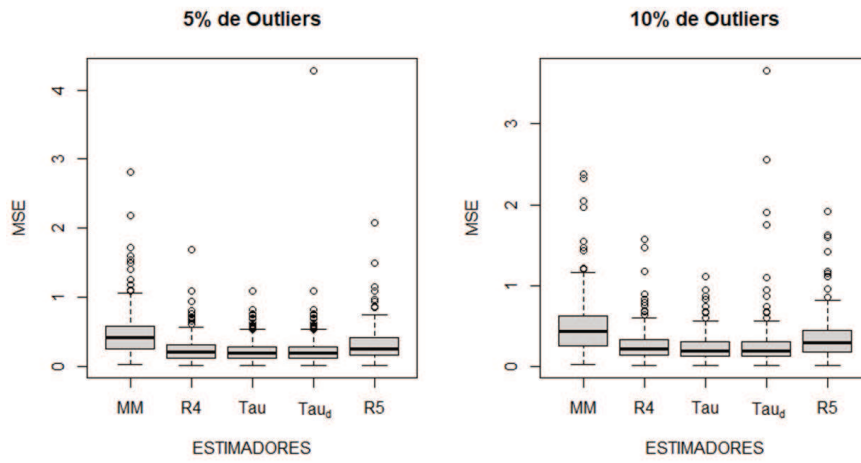


Figura 16 – Distribución de MSE para el Escenario II con contaminación II (15% – 20%) y $\alpha = 2$.

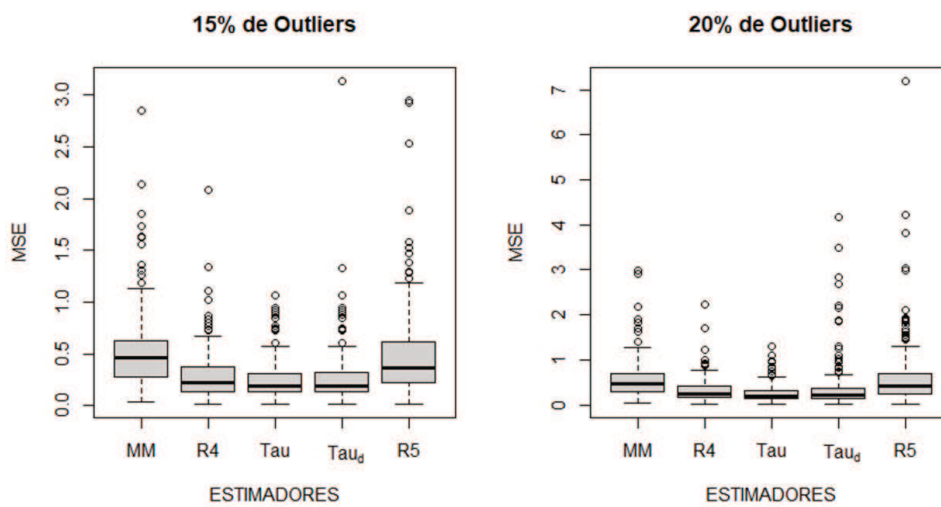


Tabla 10. Resultados de simulación del Escenario II con contaminación II y $\alpha = 4$.

O %	Estimadores	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}, \Sigma_0$)	Estándar Err($\hat{\mathbf{B}}, \Sigma_0$)	Rango
5 %	RRR	0.8543	0.5703	0.5742	0.5222	1.71
	MM	0.4262	0.2697	0.1547	0.0269	8.00
	R4	0.2187	0.1504	0.0945	0.0245	3.00
	R5	0.2714	0.1865	0.1276	0.0351	3.00
	R5 _d	0.2835	0.1889	0.1343	0.0377	3.00
	Tau	0.2167	0.1532	0.0984	0.0234	3.00
	Tau _d	0.2194	0.1569	0.0991	0.0239	3.00
10 %	RRR	2.3247	1.7264	1.6125	1.4951	1.16
	MM	0.4486	0.2817	0.1653	0.0267	8.00
	R4	0.2315	0.1553	0.1024	0.0266	3.01
	R5	0.3159	0.2128	0.1530	0.0435	3.00
	R5 _d	0.3097	0.1988	0.1511	0.0415	3.00
	Tau	0.2169	0.1492	0.1008	0.0223	3.00
	Tau _d	0.2194	0.1533	0.1014	0.0228	3.00
15 %	RRR	4.0775	3.1678	2.8329	2.5660	1.05
	MM	0.4784	0.3143	0.1726	0.0289	8.00
	R4	0.2500	0.1769	0.1095	0.0299	3.02
	R5	0.4053	0.2868	0.2067	0.0701	3.00
	R5 _d	0.4188	0.2846	0.2134	0.0745	3.00
	Tau	0.2236	0.1606	0.1043	0.0253	3.00
	Tau _d	0.2311	0.1727	0.1054	0.0266	3.00
20 %	RRR	6.7090	5.8144	4.6768	4.3882	1.05
	MM	0.4993	0.3229	0.1852	0.0312	8.00
	R4	0.2695	0.1887	0.1167	0.0303	3.04
	R5	0.4913	0.3531	0.2597	0.1002	3.00
	R5 _d	0.4883	0.3460	0.2588	0.0999	3.00
	Tau	0.2356	0.1684	0.1083	0.0252	3.00
	Tau _d	0.2563	0.1963	0.1113	0.0282	3.00

Figura 17 – Distribución de MSE para el Escenario II con contaminación II (5% – 10%) y $\alpha = 4$.

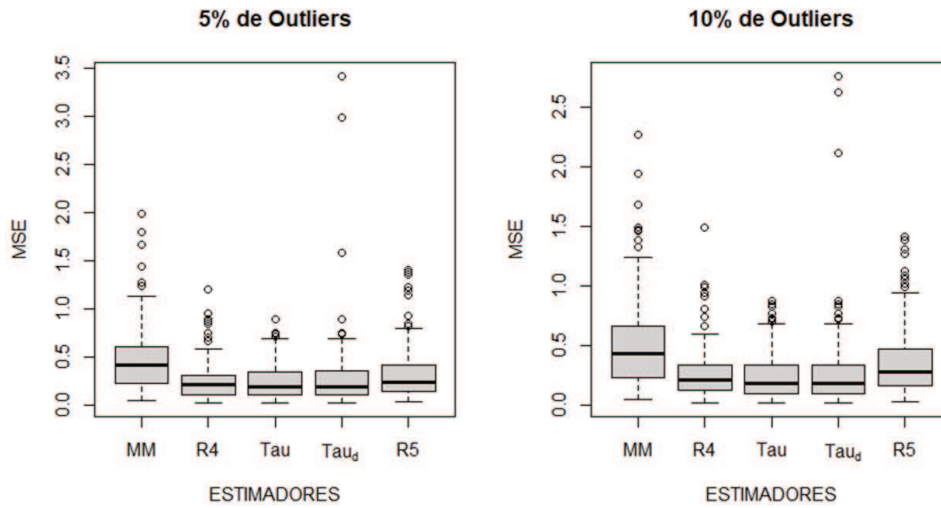
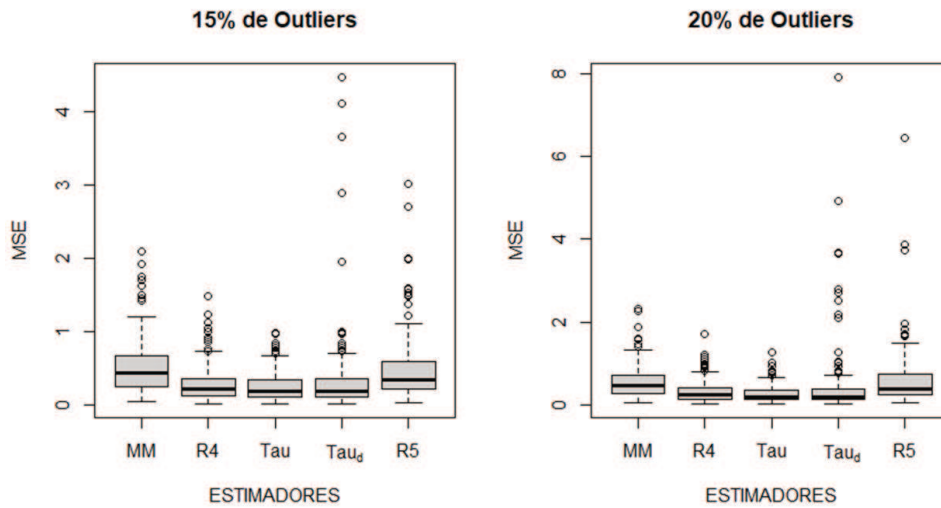


Figura 18 – Distribución de MSE para el Escenario II con contaminación II (15% – 20%) y $\alpha = 4$.



4.5 Estudio de simulación 2

En esta sección se lleva a cabo un estudio de simulación utilizando el escenario propuesto por Zhao et al. (2017). El propósito de este estudio de simulación es analizar y comparar el desempeño de alguno de los métodos descritos en la Sección 4.1, proponiendo distintos valores para el rango de la matriz \mathbf{B}_0 en el modelo (4.1).

Se estudiaron los tamaños muestrales $n = 50, 100$, $p = q = 6$ y para todos los casos analizados se consideró el rango de \mathbf{B}_0 igual a $d = 2$.

- Las matrices \mathbf{R}_0 y \mathbf{Q}_0 que forman la matriz de coeficientes \mathbf{B}_0 son dadas por $\mathbf{R}_0^t = \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 0.4 & 0.8 & 1.2 & 1.6 & 2 \end{pmatrix}^t$ y la matriz $\begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_0 \end{pmatrix}^t$ es generada aleatoriamente usando el siguiente proceso: Se crea una matriz de tamaño $q \times d$ cuyos elementos se generan aleatoriamente en forma independiente siguiendo una distribución normal estándar. Posteriormente, esta matriz se somete a una descomposición en valores singulares, esta técnica permite descomponerla en tres componentes principales: una matriz de vectores singulares izquierdos, una matriz diagonal de valores singulares, y la transposición conjugada de una matriz de vectores singulares derechos. En este proceso, se seleccionan los dos primeros vectores singulares izquierdos, representados por la matriz U , y los dos primeros vectores singulares derechos, representados por la matriz V . A continuación, se realiza la multiplicación de matrices U por la transpuesta de V (UV^t), lo que resulta en una aproximación de rango reducido de la matriz original, conservando las características más significativas de la misma.
- La matriz \mathbf{X} es construida a partir de una distribución gaussiana multivariada con $\text{Cov}(\mathbf{X}_{ij}, \mathbf{X}_{i'j'}) = \rho^{|j-j'|}$. Para este estudio se consideraron $\rho = 0.1, 0.4, 0.7$.
- Los elementos de la matriz \mathbf{U} , correspondiente al término del error en el modelo (4.1), contienen observaciones provenientes de las siguientes distribuciones univariadas, generadas de manera independiente e idénticamente distribuidas:
 - *Distribución Normal*: Los valores de \mathbf{U} se obtienen a partir de una distribución normal con media igual a 0 y desviación estándar igual a σ . Los parámetros considerados para σ son 0.5, 1 y 2.
 - *Distribución t de Student*: Los elementos de la matriz \mathbf{U} se generan a partir de una distribución t de Student con 3 grados de libertad, a la cual se le multiplica un valor σ . Los valores de σ considerados son 0.5, 1 y 2.
 - *Distribución Log-normal*: Los valores de \mathbf{U} provienen de una distribución log-normal ajustada para garantizar una desviación estándar igual a σ . Los valores de σ considerados son 0.5, 1 y 2.

Este estudio se realizó sin y con contaminación, utilizando el siguiente esquema para contaminar los datos.

Contaminación: La contaminación analizada es similar a la *Contaminación I* del Estudio de Simulación 1 presente en la Sección 4.4, es decir, siguiendo la propuesta de She and Chen (2017) donde los valores atípicos se introducen mediante una matriz \mathbf{C} cuyos elementos son nulos, excepto las primeras $O\%$ de sus filas. La matriz \mathbf{C} se compone de valores que son α -veces la desviación estándar de cada columna de \mathbf{XB}_0 , multiplicada por un valor aleatorio seleccionado en el conjunto $\{-1, 1\}$. Dado que en este estudio se desea comparar el comportamiento de los métodos para diferentes rangos y diferentes distribuciones en los errores, únicamente se reportó el caso de contaminación para $O\% = 10\%$ y $\alpha = 2$. Con el fin de aumentar el desafío del problema, se reemplazan todos los elementos de las dos primeras filas de la matriz de diseño \mathbf{X} por el valor 10, dando lugar a valores atípicos con un alto nivel de apalancamiento. Esta matriz modificada se denota como \mathbf{X}_{10} . Posteriormente, la respuesta \mathbf{Y} se genera utilizando la expresión $\mathbf{Y} = \mathbf{X}_{10}\mathbf{B}_0 + \mathbf{U} + \mathbf{C}$.

4.5.1 Resultados

Tal como lo descrito en la Sección 4.4.1, para cada uno de los escenarios anteriores se realizaron $N = 100$ réplicas y se reportan los errores *Errores Cuadráticos Medios recortados* detallados en la Sección 4.3 y sus errores estándar. En este caso los errores fueron calculados utilizando $n_{\text{test}} = 50$.

Dado que este estudio de simulación está destinado a mostrar el comportamiento de cada estimador en función del rango, de los métodos descritos en la Sección 4.1 se consideraron aquellos que se obtienen a partir de algoritmos que utilizan el rango que reciben como parámetro de entrada: RRR, R5 y Tau. Por lo tanto, los estimadores Tau_d , R5_d , que utilizan el rango que es estimado, fueron excluidos de este estudio como así también el estimador MM ya que por su naturaleza es un estimador de rango completo. Si bien el estimador R4 elige el rango, el mismo sí fue considerado en este estudio. Aunque los resultados obtenidos no se reportan en las tablas, más adelante en esta misma sección se comenta sobre su desempeño. Si bien el verdadero rango (o sea, el rango de \mathbf{B}_0) es 2, en los distintos ajustes cada algoritmo fue corrido fijando el valor de $d \in \{1, 2, 3, 4, 5, 6\}$.

Los resultados sin contaminación se reportan en las Tablas 11-19 para los tres tipos de distribuciones de errores, donde se consideró, por simplicidad de exposición, $n = 50$ y $d = 1, 2, 3, 6$, ya que los resultados para $n = 100$ dan errores más pequeños pero con la misma tendencia. También para facilitar la lectura de las tablas se excluyeron los casos $d = 4, 5$, ya que los resultados presentaban un comportamiento similar a los reportados. En el contexto de la distribución normal (Tablas 11-13), se observa que los métodos RRR, R5 y Tau son comparables, con una leve ventaja para el RRR, como es de esperar ya que es el estimador óptimo cuando los errores tienen distribución normal. Para errores con distribución t de Student (Tablas 14-16) y con distribución lognormal (Tablas 17-19) los estimadores R5 y Tau tienen un mejor comportamiento cuando es comparado con el RRR. En todos los casos estudiados sin contaminación, el estimador R5 presenta resultados levemente más favorables que el estimador Tau.

Los resultados con contaminación se encuentran en las Tablas 20-28 para los tres tipos de

distribuciones de errores, donde se consideró nuevamente $n = 50$ y $d = 1, 2, 3, 6$. En todos los casos puede observarse que el estimador **Tau** reporta resultados considerablemente mejores que el resto de estimadores. Más precisamente, en este caso contaminado puede observarse que el estimador robusto **R5** muestra un mal desempeño, comparable con el estimador clásico **RRR**.

Tanto en el caso sin contaminar como en el contaminado, puede observarse que para todos los estimadores considerados los errores cuadráticos medios recortados más bajos se obtuvieron para el rango 2 que es el rango verdadero del modelo. Cuando se subestima el rango, lo errores son mayores que cuando se sobreestima. Este comportamiento se acentúa más para el caso contaminado.

Si bien los errores para el método **R4** no fueron reportados en las tablas para diferentes rangos pues, como se ha mencionado anteriormente este método estima el rango, el mismo se ha tenido en cuenta en las simulaciones mostrando errores competitivos con el resto de los estimadores para escenarios sin contaminación y empeorando considerablemente su desempeño para casos contaminados.

Finalmente, cuando se comparan los métodos sin y con contaminación, puede observarse que en el caso normal y lognormal para $\sigma = 0.5, 1$, el estimador **Tau** mostró resultados similares en ambos casos, esto es, sin y con contaminación (Tablas 11-12, 20-21, y Tablas 17-18, 26-27, respectivamente). Lo mismo ocurre con la distribución t de Student con $\sigma = 0.5$ (Tablas 14 y 23). Para el caso contaminado de la normal y lognormal con $\sigma = 2$ (Tablas 22 y 28) y t de Student con $\sigma = 1, 2$ (Tablas 24 y 25), el método **Tau** mostró peores resultados comparados con los obtenidos sin contaminar (Tablas 13, 19, 15, 16, respectivamente). Sin embargo, los errores para el resto de los estimadores aumentaron en un orden de magnitud o más resultando, en todos los casos, en errores muchos mayores que aquellos reportados por el **Tau** estimador.

Tabla 11. Resultados de simulación para error normal con $n = 50$, $\sigma = 0.5$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.3439	0.0670	0.4564	0.0926	0.4770	0.1008
	$d = 2$	0.0096	0.0030	0.0095	0.0030	0.0095	0.0030
	$d = 3$	0.0147	0.0043	0.0147	0.0043	0.0147	0.0043
	$d = 6$	0.0178	0.0049	0.0178	0.0049	0.0178	0.0049
R5	$d = 1$	0.3558	0.0724	0.4743	0.0992	0.4866	0.1042
	$d = 2$	0.0105	0.0034	0.0104	0.0034	0.0103	0.0034
	$d = 3$	0.0161	0.0048	0.0160	0.0047	0.0159	0.0047
	$d = 6$	0.0192	0.0052	0.0192	0.0053	0.0192	0.0053
Tau	$d = 1$	0.3667	0.0697	0.4842	0.0977	0.5037	0.1085
	$d = 2$	0.0137	0.0057	0.0136	0.0056	0.0135	0.0056
	$d = 3$	0.0200	0.0065	0.0198	0.0064	0.0198	0.0064
	$d = 6$	0.0233	0.0065	0.0233	0.0065	0.0233	0.0065

Tabla 12. Resultados de simulación para error normal con $n = 50$, $\sigma = 1$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.3611	0.0696	0.4739	0.0953	0.4944	0.1019
	$d = 2$	0.0391	0.0129	0.0386	0.0126	0.0384	0.0125
	$d = 3$	0.0593	0.0174	0.0589	0.0174	0.0588	0.0174
	$d = 6$	0.0711	0.0198	0.0711	0.0198	0.0711	0.0198
R5	$d = 1$	0.3671	0.0701	0.4833	0.0980	0.5002	0.1044
	$d = 2$	0.0428	0.0144	0.0424	0.0142	0.0423	0.0143
	$d = 3$	0.0648	0.0190	0.0643	0.0189	0.0640	0.0188
	$d = 6$	0.0770	0.0210	0.0770	0.0210	0.0770	0.0210
Tau	$d = 1$	0.3924	0.0745	0.5108	0.1014	0.5305	0.1153
	$d = 2$	0.0560	0.0232	0.0552	0.0228	0.0546	0.0226
	$d = 3$	0.0802	0.0252	0.0794	0.0253	0.0795	0.0259
	$d = 6$	0.0933	0.0260	0.0933	0.0260	0.0933	0.0260

Tabla 13. Resultados de simulación para error normal con $n = 50$, $\sigma = 2$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.4304	0.0872	0.5415	0.1102	0.5627	0.1155
	$d = 2$	0.1683	0.0617	0.1640	0.0583	0.1618	0.0558
	$d = 3$	0.2416	0.0717	0.2392	0.0708	0.2383	0.0711
	$d = 6$	0.2843	0.0791	0.2843	0.0791	0.2843	0.0791
R5	$d = 1$	0.4394	0.0836	0.5526	0.1106	0.5990	0.2306
	$d = 2$	0.1855	0.0715	0.1860	0.0787	0.2109	0.0987
	$d = 3$	0.2621	0.0762	0.2599	0.0761	0.2616	0.0768
	$d = 6$	0.3079	0.0840	0.3080	0.0840	0.3080	0.0840
Tau	$d = 1$	0.5002	0.1267	0.6196	0.1450	0.6375	0.1526
	$d = 2$	0.2518	0.1192	0.2458	0.1236	0.2381	0.1111
	$d = 3$	0.3299	0.1058	0.3279	0.1111	0.3196	0.1004
	$d = 6$	0.3732	0.1040	0.3732	0.1040	0.3732	0.1040

Tabla 14. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 0.5$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.3554	0.0657	0.4687	0.0893	0.4904	0.0981
	$d = 2$	0.0276	0.0210	0.0271	0.0183	0.0276	0.0214
	$d = 3$	0.0482	0.0682	0.0481	0.0683	0.0480	0.0683
	$d = 6$	0.0556	0.0687	0.0556	0.0687	0.0556	0.0687
R5	$d = 1$	0.3559	0.0669	0.4755	0.0941	0.5013	0.1025
	$d = 2$	0.0164	0.0062	0.0162	0.0061	0.0162	0.0061
	$d = 3$	0.0270	0.0089	0.0269	0.0089	0.0270	0.0090
	$d = 6$	0.0323	0.0102	0.0323	0.0102	0.0323	0.0102
Tau	$d = 1$	0.3918	0.1157	0.5127	0.1468	0.5246	0.1318
	$d = 2$	0.0249	0.0097	0.0247	0.0094	0.0247	0.0097
	$d = 3$	0.0361	0.0130	0.0362	0.0126	0.0364	0.0125
	$d = 6$	0.0439	0.0147	0.0439	0.0147	0.0439	0.0147

Tabla 15. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 1$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.4241	0.2177	0.5342	0.1995	0.5547	0.1931
	$d = 2$	0.1376	0.2982	0.1369	0.3063	0.1299	0.2401
	$d = 3$	0.1945	0.2737	0.1935	0.2737	0.1932	0.2737
	$d = 6$	0.2225	0.2748	0.2225	0.2748	0.2225	0.2748
R5	$d = 1$	0.3804	0.0660	0.4982	0.0902	0.5258	0.1011
	$d = 2$	0.0674	0.0262	0.0666	0.0257	0.0688	0.0275
	$d = 3$	0.1087	0.0358	0.1083	0.0356	0.1088	0.0364
	$d = 6$	0.1290	0.0408	0.1291	0.0409	0.1290	0.0408
Tau	$d = 1$	0.4370	0.1180	0.5610	0.1637	0.5698	0.1478
	$d = 2$	0.1029	0.0412	0.1018	0.0406	0.1014	0.0415
	$d = 3$	0.1462	0.0529	0.1444	0.0512	0.1443	0.0508
	$d = 6$	0.1755	0.0587	0.1755	0.0587	0.1755	0.0587

Tabla 16. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 2$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.7493	1.0925	0.8336	1.0964	0.8540	1.1372
	$d = 2$	0.6444	1.1258	0.6204	1.1446	0.6014	1.1053
	$d = 3$	0.7976	1.1000	0.7901	1.1004	0.7863	1.1000
	$d = 6$	0.8898	1.0994	0.8898	1.0994	0.8898	1.0994
R5	$d = 1$	0.5125	0.1072	0.6305	0.1258	0.7692	0.3626
	$d = 2$	0.3228	0.1527	0.3436	0.1780	0.4648	0.2780
	$d = 3$	0.4461	0.1552	0.4460	0.1587	0.4546	0.1663
	$d = 6$	0.5161	0.1633	0.5160	0.1632	0.5159	0.1632
Tau	$d = 1$	0.6328	0.1848	0.7636	0.2391	0.7525	0.2019
	$d = 2$	0.4854	0.2246	0.4628	0.2184	0.4611	0.2139
	$d = 3$	0.6043	0.2251	0.5965	0.2225	0.5893	0.2163
	$d = 6$	0.7018	0.2347	0.7018	0.2347	0.7018	0.2347

Tabla 17. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 0.5$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.3438	0.0669	0.4565	0.0928	0.4772	0.1011
	$d = 2$	0.0097	0.0033	0.0096	0.0033	0.0096	0.0034
	$d = 3$	0.0151	0.0046	0.0150	0.0046	0.0150	0.0046
	$d = 6$	0.0180	0.0054	0.0180	0.0054	0.0180	0.0054
R5	$d = 1$	0.3568	0.0748	0.4758	0.1015	0.4873	0.1047
	$d = 2$	0.0076	0.0028	0.0076	0.0027	0.0075	0.0027
	$d = 3$	0.0120	0.0037	0.0120	0.0037	0.0120	0.0038
	$d = 6$	0.0143	0.0042	0.0143	0.0042	0.0143	0.0042
Tau	$d = 1$	0.3673	0.0691	0.4830	0.0934	0.5028	0.1105
	$d = 2$	0.0114	0.0049	0.0112	0.0048	0.0113	0.0049
	$d = 3$	0.0162	0.0058	0.0163	0.0059	0.0163	0.0059
	$d = 6$	0.0198	0.0068	0.0198	0.0068	0.0198	0.0068

Tabla 18. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 1$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.3610	0.0696	0.4743	0.0957	0.4950	0.1025
	$d = 2$	0.0400	0.0189	0.0393	0.0178	0.0392	0.018
	$d = 3$	0.0619	0.0259	0.0618	0.0260	0.0617	0.0260
	$d = 6$	0.0728	0.0286	0.0728	0.0286	0.0728	0.0286
R5	$d = 1$	0.3607	0.0723	0.4783	0.0994	0.4927	0.1036
	$d = 2$	0.0188	0.0075	0.0186	0.0074	0.0186	0.0074
	$d = 3$	0.0306	0.0104	0.0305	0.0104	0.0304	0.0103
	$d = 6$	0.0365	0.0122	0.0365	0.0122	0.0365	0.0122
Tau	$d = 1$	0.3903	0.0814	0.5158	0.1088	0.5282	0.1172
	$d = 2$	0.0317	0.0146	0.0312	0.0143	0.0312	0.0142
	$d = 3$	0.0451	0.0175	0.0452	0.0173	0.0452	0.0173
	$d = 6$	0.0560	0.0201	0.0560	0.0201	0.0560	0.0201

Tabla 19. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 2$, sin contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	0.4606	0.2851	0.5650	0.2556	0.5844	0.2591
	$d = 2$	0.1975	0.2859	0.1927	0.2924	0.1911	0.2936
	$d = 3$	0.2650	0.2541	0.2649	0.2544	0.2643	0.2539
	$d = 6$	0.3018	0.2591	0.3018	0.2591	0.3018	0.2591
R5	$d = 1$	0.3711	0.0718	0.4892	0.0998	0.5059	0.1037
	$d = 2$	0.0354	0.0162	0.0354	0.0163	0.0357	0.0166
	$d = 3$	0.0605	0.0245	0.0605	0.0247	0.0602	0.0244
	$d = 6$	0.0721	0.0287	0.0721	0.0287	0.0721	0.0287
Tau	$d = 1$	0.4410	0.1311	0.5668	0.1787	0.5764	0.1749
	$d = 2$	0.0750	0.0355	0.0749	0.0354	0.0750	0.0358
	$d = 3$	0.1083	0.0459	0.1086	0.0471	0.1093	0.0470
	$d = 6$	0.1329	0.0522	0.1331	0.0524	0.1331	0.0524

Tabla 20. Resultados de simulación para error normal con $n = 50$, $\sigma = 0.5$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.0902	0.3327	1.6123	0.4888	2.3356	0.7247
	$d = 2$	0.7123	0.1975	1.1383	0.3448	1.9020	0.6368
	$d = 3$	0.7199	0.2113	1.1570	0.3642	1.9132	0.6535
	$d = 6$	0.7253	0.2104	1.1624	0.3635	1.9188	0.6530
R5	$d = 1$	0.9379	0.3160	1.4826	0.4965	2.2728	0.7905
	$d = 2$	0.6904	0.2335	1.1348	0.3916	1.9455	0.7220
	$d = 3$	0.6894	0.2154	1.1260	0.3800	1.8775	0.6888
	$d = 6$	0.7284	0.2128	1.1658	0.3692	1.9176	0.6703
Tau	$d = 1$	0.3713	0.0832	0.4896	0.1120	0.5046	0.1100
	$d = 2$	0.0143	0.0049	0.0141	0.0048	0.0140	0.0048
	$d = 3$	0.0214	0.0065	0.0210	0.0063	0.0209	0.0062
	$d = 6$	0.0258	0.0072	0.0256	0.0069	0.0255	0.0069

Tabla 21. Resultados de simulación para error normal con $n = 50$, $\sigma = 1$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.1435	0.3414	1.6619	0.5023	2.3691	0.7443
	$d = 2$	0.7432	0.2074	1.1695	0.3569	1.9395	0.6489
	$d = 3$	0.7669	0.2191	1.2050	0.3741	1.9627	0.6669
	$d = 6$	0.7855	0.2172	1.2242	0.3725	1.9832	0.6655
R5	$d = 1$	1.0967	0.3297	1.6005	0.5145	2.3654	0.8150
	$d = 2$	0.7719	0.2494	1.2340	0.4241	2.0378	0.7449
	$d = 3$	0.7784	0.2453	1.2148	0.4020	1.9952	0.7186
	$d = 6$	0.7953	0.2194	1.2318	0.3772	1.9863	0.6817
Tau	$d = 1$	0.4060	0.1065	0.5179	0.1248	0.5300	0.1130
	$d = 2$	0.0615	0.0229	0.0588	0.0215	0.0582	0.0208
	$d = 3$	0.0878	0.0278	0.0861	0.0265	0.0851	0.0258
	$d = 6$	0.1048	0.0304	0.1039	0.0296	0.1030	0.0280

Tabla 22. Resultados de simulación para error normal con $n = 50$, $\sigma = 2$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.2600	0.3628	1.7881	0.5399	2.4756	0.8056
	$d = 2$	0.8771	0.2494	1.3090	0.4069	2.0946	0.7101
	$d = 3$	0.9673	0.2553	1.4114	0.4140	2.1783	0.7196
	$d = 6$	1.0255	0.2532	1.4712	0.4142	2.2425	0.7194
R5	$d = 1$	1.2686	0.3823	1.8066	0.5673	2.5737	0.8955
	$d = 2$	0.9556	0.3201	1.4160	0.4668	2.2937	0.8182
	$d = 3$	1.0182	0.2670	1.4742	0.4347	2.2802	0.7667
	$d = 6$	1.0519	0.2559	1.4978	0.4198	2.2682	0.7389
Tau	$d = 1$	0.9439	0.4436	1.0436	0.6010	0.8391	0.5631
	$d = 2$	0.6317	0.4017	0.6488	0.5937	0.4493	0.5546
	$d = 3$	0.7165	0.3872	0.7305	0.5552	0.5388	0.5118
	$d = 6$	0.7941	0.4075	0.8184	0.5884	0.6212	0.5369

Tabla 23. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 0.5$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.1266	0.3406	1.6413	0.4682	2.3421	0.6473
	$d = 2$	0.7295	0.2171	1.1450	0.3443	1.8883	0.5994
	$d = 3$	0.7388	0.2202	1.1661	0.3524	1.9026	0.6085
	$d = 6$	0.7519	0.2187	1.1799	0.3527	1.9163	0.6088
R5	$d = 1$	1.0063	0.2955	1.5130	0.4450	2.2765	0.7164
	$d = 2$	0.7141	0.2456	1.1104	0.3669	1.9121	0.6501
	$d = 3$	0.7018	0.2049	1.1205	0.3588	1.8410	0.6294
	$d = 6$	0.7306	0.1938	1.1568	0.3333	1.8855	0.5960
Tau	$d = 1$	0.4149	0.1334	0.5375	0.1745	0.5379	0.1705
	$d = 2$	0.0320	0.0132	0.0315	0.0128	0.0315	0.0129
	$d = 3$	0.0455	0.0176	0.0453	0.0178	0.0461	0.0192
	$d = 6$	0.0540	0.0193	0.0542	0.0197	0.0552	0.0218

Tabla 24. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 1$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.2534	0.5044	1.7713	0.6257	2.4477	0.8087
	$d = 2$	0.8432	0.4199	1.2656	0.5231	2.0161	0.7595
	$d = 3$	0.9003	0.4287	1.3306	0.5397	2.0746	0.7806
	$d = 6$	0.9392	0.4268	1.3725	0.5396	2.1189	0.7810
R5	$d = 1$	1.1858	0.3340	1.6836	0.4959	2.4496	0.7580
	$d = 2$	0.7986	0.2499	1.2368	0.3845	2.0600	0.6897
	$d = 3$	0.8285	0.2529	1.2756	0.4010	2.0183	0.6754
	$d = 6$	0.8501	0.2235	1.2838	0.3690	2.0249	0.6360
Tau	$d = 1$	0.5445	0.2606	0.6601	0.3029	0.6129	0.2194
	$d = 2$	0.1937	0.1949	0.1674	0.2033	0.1306	0.0580
	$d = 3$	0.2433	0.2008	0.2180	0.2062	0.1839	0.0757
	$d = 6$	0.2784	0.2028	0.2531	0.2057	0.2192	0.0849

Tabla 25. Resultados de simulación para error t de Student con $n = 50$, $\sigma = 2$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.5706	1.4223	2.1200	1.5438	2.8027	1.7562
	$d = 2$	1.3378	1.4009	1.7714	1.5073	2.5883	1.7349
	$d = 3$	1.5683	1.4205	2.0123	1.5301	2.7939	1.7671
	$d = 6$	1.6906	1.4218	2.1438	1.5330	2.9278	1.7710
R5	$d = 1$	1.4051	0.3810	1.9769	0.5798	2.8292	0.8929
	$d = 2$	1.1056	0.3725	1.5912	0.5498	2.5097	0.8816
	$d = 3$	1.2350	0.3810	1.6940	0.5403	2.5193	0.8539
	$d = 6$	1.3203	0.3909	1.7815	0.5480	2.5594	0.8527
Tau	$d = 1$	1.2713	0.4452	1.4175	0.6290	1.2547	0.7318
	$d = 2$	1.0683	0.5042	1.0798	0.6240	0.9659	0.7713
	$d = 3$	1.1932	0.5000	1.2037	0.5912	1.1258	0.7359
	$d = 6$	1.3290	0.5028	1.3370	0.6226	1.2690	0.7750

Tabla 26. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 0.5$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.0895	0.3342	1.6139	0.4937	2.3349	0.7271
	$d = 2$	0.7115	0.1967	1.1376	0.3439	1.9020	0.6347
	$d = 3$	0.7188	0.2098	1.1555	0.3624	1.9112	0.6512
	$d = 6$	0.7241	0.2091	1.1608	0.3619	1.9167	0.6509
R5	$d = 1$	0.9205	0.3064	1.4665	0.4672	2.2871	0.8015
	$d = 2$	0.6957	0.2387	1.1288	0.4026	1.9305	0.7300
	$d = 3$	0.6857	0.2115	1.1180	0.3832	1.8549	0.6848
	$d = 6$	0.7260	0.2131	1.1614	0.3687	1.9120	0.6673
Tau	$d = 1$	0.3766	0.0847	0.4924	0.1094	0.5076	0.1111
	$d = 2$	0.0134	0.0051	0.0132	0.0051	0.0131	0.0051
	$d = 3$	0.0197	0.0066	0.0196	0.0064	0.0194	0.0065
	$d = 6$	0.0237	0.0076	0.0237	0.0075	0.0238	0.0076

Tabla 27. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 1$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.1464	0.3507	1.6638	0.5123	2.3682	0.7515
	$d = 2$	0.7409	0.2047	1.1664	0.3523	1.9375	0.6459
	$d = 3$	0.7639	0.2128	1.2013	0.3658	1.9572	0.6576
	$d = 6$	0.7816	0.2115	1.2190	0.3651	1.9762	0.6565
R5	$d = 1$	1.0860	0.3376	1.5878	0.5221	2.3321	0.8275
	$d = 2$	0.7348	0.2397	1.1844	0.4071	1.9753	0.7463
	$d = 3$	0.7338	0.2346	1.1802	0.4025	1.9185	0.7141
	$d = 6$	0.7655	0.2175	1.2039	0.3740	1.9580	0.6753
Tau	$d = 1$	0.4046	0.0953	0.5177	0.1216	0.5429	0.1895
	$d = 2$	0.0446	0.0194	0.0435	0.0184	0.0432	0.0184
	$d = 3$	0.0644	0.0230	0.0663	0.0267	0.0663	0.0266
	$d = 6$	0.0775	0.0254	0.0790	0.0280	0.0800	0.0289

Tabla 28. Resultados de simulación para error lognormal con $n = 50$, $\sigma = 2$, con contaminación.

Estimadores	Rango	$\rho=0.1$		$\rho=0.4$		$\rho=0.7$	
		Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)	Err($\hat{\mathbf{B}}$)	Estándar Err($\hat{\mathbf{B}}$)
RRR	$d = 1$	1.2867	0.3881	1.8117	0.5497	2.4934	0.7980
	$d = 2$	0.8957	0.3335	1.3214	0.4383	2.1079	0.6916
	$d = 3$	0.9757	0.3160	1.4129	0.4266	2.1748	0.6927
	$d = 6$	1.0215	0.3190	1.4636	0.4277	2.2296	0.6914
R5	$d = 1$	1.2189	0.3912	1.7401	0.5787	2.5359	0.9131
	$d = 2$	0.7981	0.2621	1.2552	0.4447	2.1368	0.8102
	$d = 3$	0.8260	0.2621	1.2773	0.4239	2.0649	0.7432
	$d = 6$	0.8395	0.2402	1.2905	0.4061	2.0675	0.7227
Tau	$d = 1$	0.4818	0.1840	0.5842	0.1741	0.5888	0.1565
	$d = 2$	0.1244	0.0966	0.1158	0.0595	0.1156	0.0626
	$d = 3$	0.1786	0.1048	0.1715	0.0789	0.1706	0.0738
	$d = 6$	0.2120	0.1080	0.2091	0.0842	0.2076	0.0804

5 Aplicación a datos reales

En el presente capítulo se evalúa el desempeño del τ -estimador (Tau y Tau_d) cuando es aplicado a datos reales y se lo compara con los estimadores analizados en el Capítulo 4 (Sección 4.1). En cada caso se incluyen las citas bibliográficas de las cuales han sido extraídos los datos y que proporcionan contexto para la interpretación de los resultados obtenidos.

5.1 Medida de error

Dado que los valores reales de los parámetros ($\mathbf{B}_0, \mathbf{\Sigma}_0$) son desconocidos, se recurre a la validación simple para evaluar el rendimiento del modelo. El conjunto de datos se divide de manera aleatoria en dos subconjuntos: el 70 % de las observaciones se utiliza para entrenar el modelo, permitiendo la estimación de los parámetros, mientras que el 30 % restante se destina a la fase de validación, donde se calculan los errores de predicción. Se repite el proceso N veces, y los errores son resumidos en una única métrica de error que refleja el rendimiento general del modelo.

Para evaluar el rendimiento del modelo en cada iteración, se calcula el *Error Cuadrático Medio* (MSE), que mide la discrepancia al cuadrado entre los valores observados y los valores predichos por el modelo en el conjunto de validación de esa iteración. Matemáticamente, el MSE para cada iteración se expresa como:

$$\text{MSE} = \frac{1}{n_{\text{test}}} \|\mathbf{Y}_{\text{test}} - \mathbf{X}_{\text{test}} \hat{\mathbf{B}}\|_F^2, \quad (5.1)$$

donde \mathbf{X}_{test} representa la matriz de valores de las características en el conjunto de validación, n_{test} es el número de observaciones en \mathbf{X}_{test} , y \mathbf{Y}_{test} denota la matriz de valores observados en el conjunto de validación.

A partir de los N valores de MSE obtenidos en (5.1), se calcula la media recortada al 10 %, excluyendo los valores más extremos para evitar el efecto de los posibles valores atípicos. Este promedio recortado se denomina **Err**.

Adicionalmente, se utiliza la técnica de validación cruzada “dejar uno fuera” (LOOCV, por sus siglas en inglés) para generar gráficos de valores observados frente a valores predichos. En LOOCV, el conjunto de datos se divide en n particiones, donde n es el número total de observaciones en el conjunto. El modelo se entrena n veces, utilizando en cada iteración $n - 1$ observaciones para el entrenamiento y la restante para la validación. Este proceso se repite para cada dato del conjunto, asegurando que todas las observaciones contribuyan tanto al entrenamiento como a la evaluación. Los gráficos resultantes permiten evaluar la capacidad predictiva del modelo al comparar visualmente las predicciones con los datos reales, proporcionando una visión más completa sobre el ajuste del modelo a los datos.

Por último, se estimaron los coeficientes del modelo utilizando el método bootstrap, generando múltiples muestras mediante remuestreo a partir de los datos originales para construir una

distribución empírica de los estimadores. A partir de esta distribución, se calcularon los errores estándar (SE), que reflejan la variabilidad de los coeficientes estimados. Los resultados, incluyendo los coeficientes estimados y sus errores estándar, se presentan organizados en una tabla para cada uno de los conjuntos de datos analizados, facilitando la comparación y el análisis de las estimaciones.

5.2 Datos bioquímicos

En este análisis se utilizan datos bioquímicos provenientes de un estudio realizado por [Fairfield Smith et al. \(1962\)](#), que comprenden mediciones de diversas características en muestras de orina tomadas de 17 hombres, clasificados en dos grupos según su peso: sobrepeso y bajo peso. Cada sujeto proporcionó dos muestras de orina por la mañana, con la excepción de uno que contribuyó solo con una muestra, resultando en un total de 33 muestras disponibles. Las muestras son tratadas como independientes, sin tener en cuenta las posibles correlaciones dentro de las dos determinaciones de cada sujeto. Estos datos fueron analizados en el contexto de regresión lineal multivariada de rango reducido por [Velu and Reinsel \(2013\)](#) (Secciones 1.5 y 2.7).

El objetivo principal del estudio es examinar cómo factores personales, como el peso corporal, y variables relacionadas con el análisis de orina, como el volumen de orina y la gravedad específica, se asocian con las concentraciones de cinco componentes bioquímicos en la orina. Se busca identificar y analizar las asociaciones entre estas variables predictoras y los niveles de pigmento de creatinina, fosfato, fósforo, creatinina y colina en las muestras de orina. Las cinco variables respuesta consideradas en el estudio son

- $\mathbf{y}_{(:,1)}$ pigmento de creatinina
- $\mathbf{y}_{(:,2)}$ fosfato (mg/mL)
- $\mathbf{y}_{(:,3)}$ fósforo (mg/mL)
- $\mathbf{y}_{(:,4)}$ creatinina (mg/mL)
- $\mathbf{y}_{(:,5)}$ colina ($\mu\text{g/mL}$)

y las tres variables predictoras consideradas son

- $\mathbf{x}_{(:,1)}$ peso (lbs/100)
- $\mathbf{x}_{(:,2)}$ volumen (mL/100)
- $\mathbf{x}_{(:,3)}$ 100 (gravedad específica - 1).

El conjunto de datos multivariados para este ejemplo se presenta en la siguiente tabla:

Tabla 29. Datos bioquímicos sobre muestras de orina de pacientes de un estudio realizado por (Fairfield Smith et al., 1962).

$\mathbf{Y}_{(:,1)}$	$\mathbf{Y}_{(:,2)}$	$\mathbf{Y}_{(:,3)}$	$\mathbf{Y}_{(:,4)}$	$\mathbf{Y}_{(:,5)}$	$\mathbf{X}_{(:,1)}$	$\mathbf{X}_{(:,2)}$	$\mathbf{X}_{(:,3)}$
17.60	1.50	1.50	1.88	7.50	0.98	2.05	2.40
13.40	1.65	1.32	2.24	7.10	0.98	1.60	3.20
20.30	0.90	0.89	1.28	2.30	1.15	4.80	1.70
22.30	1.75	1.50	2.24	4.00	1.28	2.30	3.00
18.50	1.20	1.03	1.84	2.00	1.28	2.15	2.70
12.10	1.90	1.87	2.40	16.80	1.22	2.15	2.50
10.10	2.30	2.08	2.68	0.90	1.22	1.90	2.80
14.70	2.35	2.55	3.00	2.00	1.30	1.75	2.40
14.40	2.50	2.38	2.84	3.80	1.30	1.55	2.70
18.10	1.50	1.20	2.60	14.50	1.35	2.20	3.10
16.90	1.40	1.15	1.72	8.00	1.35	3.05	3.20
23.70	1.65	1.58	1.60	4.90	1.35	2.75	2.00
18.00	1.60	1.68	2.00	3.60	1.35	2.10	2.30
14.80	2.45	2.15	3.12	12.00	1.40	1.70	3.10
15.60	1.65	1.42	2.56	5.20	1.40	2.35	2.80
16.20	1.65	1.62	2.04	10.20	1.41	1.85	2.10
17.50	1.05	1.56	1.48	9.60	1.41	2.65	1.50
14.10	2.70	2.77	2.56	6.90	1.62	3.05	2.60
22.50	0.85	1.65	1.20	3.50	1.62	4.30	1.60
17.00	0.70	0.97	1.24	1.90	2.05	3.50	1.80
12.50	0.80	0.80	0.64	0.70	2.05	4.75	1.00
21.50	1.80	1.77	2.60	8.30	2.30	1.95	3.30
13.00	2.20	1.85	3.84	13.00	2.30	1.60	3.50
13.00	3.55	3.18	3.48	18.30	2.15	2.40	3.30
12.00	3.65	2.40	3.00	14.50	2.15	2.70	3.40
22.80	0.55	1.00	1.14	3.30	2.30	4.75	1.60
18.40	1.05	1.17	1.36	4.90	2.30	4.90	2.80
8.70	4.25	3.62	3.84	19.50	2.62	1.15	2.50
9.40	3.85	3.36	5.12	1.30	2.62	0.97	2.80
15.00	2.45	2.38	2.40	20.00	2.55	3.25	2.70
12.90	1.70	1.74	2.48	1.00	2.55	3.10	2.30
12.10	1.80	2.00	2.24	5.00	2.70	2.45	2.50
11.50	2.25	2.25	3.12	5.10	2.70	2.20	3.40

El presente estudio se inicia con la selección del rango adecuado para el modelo lineal multivariado (1). De acuerdo con la Tabla 2.2 del Velu and Reinsel (2013), las pruebas de razón de verosimilitud (LR) sugieren que el rango del modelo podría ser 1 o 2. Para corroborar estas afirmaciones, se empleó la función `rank_trace` del paquete `rrr` (un paquete dentro del entorno de R), que estima que el rango adecuado para el RRR es 2. Además, los resultados para el estimador R4 indicaron un rango estimado de 2, mientras que \mathbf{Tau}_d mostró un rango estimado de 1.

En esta dirección, tomando en cuenta lo anterior, se compararán los distintos estimadores propuestos considerando tanto $d = 2$ como $d = 1$. Este enfoque permitirá verificar cómo varían las estimaciones de los coeficientes de regresión bajo diferentes supuestos de rango, lo que proporcionará una comprensión más completa del modelo y asegurará que la selección del rango no afecte negativamente la validez de los resultados.

Una vez estimado el rango, se procede a identificar y evaluar los valores atípicos en los datos mediante la distancia de Mahalanobis de los residuos (1.1), siguiendo un enfoque similar al descrito en [García Ben et al. \(2006\)](#). Se considerarán como valores atípicos aquellas observaciones cuyo valor de la distancia de Mahalanobis d_i supere el umbral crítico $\sqrt{\chi_{5,0.99}^2}$, donde $\chi_{5,0.99}^2$ es el percentil 0.99 de la distribución χ^2 con 5 grados de libertad, es decir, si $d_i > \sqrt{\chi_{5,0.99}^2}$ para $i = 1, \dots, 33$. En este análisis, el estimador **RRR** identificó un total de 0 valores atípicos, mientras que el τ -estimador reveló 5 valores atípicos para un rango $d = 2$.

En lo que sigue, para comparar el desempeño de los estimadores propuestos, se llevará a cabo un análisis de validación simple descrito en la Sección 5.1, con $N = 200$, considerando dos escenarios: uno que incluya todos los datos y otro que excluya los valores atípicos identificados, con el objetivo de evaluar la influencia de estos en el rendimiento del modelo. Se empleará el MSE recortado al 10% (**Err**) para evaluar la validación. Adicionalmente, se aplicará LOOCV de manera exploratoria para generar gráficos de valores observados versus predichos.

Tabla 30. Errores (**Err**) para los datos bioquímicos obtenidos a partir de validación simple con toda la muestra ($n = 33$) y diferentes rangos.

Estimadores	$d = 1$	$d = 2$
RRR	45.51	49.10
R5	53.27	51.61
Tau	49.40	51.60

Al analizar los resultados de la Tabla 30, se observa que el estimador **RRR** presenta el menor error para $d = 1$ con un valor de 45.51, seguido por el estimador **Tau**, cuyo error es 49.40. Por otro lado, el estimador **R5** tiene el mayor error en este rango, alcanzando 53.27.

Para $d = 2$, el estimador **RRR** sigue mostrando un buen desempeño con un error de 49.10. Sin embargo, el estimador **Tau** se mantiene competitivo, con un error de 51.60, mientras que el estimador **R5** muestra una ligera mejora al reducir su error a 51.61, aunque sigue siendo el de mayor error en términos generales.

En conclusión, aunque el estimador **RRR** tiene el mejor desempeño general, el estimador **Tau** se posiciona como una alternativa robusta y competitiva en ambas configuraciones, destacando especialmente en el rango $d = 1$.

El estimador Tau_d y el estimador **R4** no se han incluido en la Tabla 30 debido a que ambos estiman el rango, lo que los diferencia de los demás métodos evaluados. El Tau_d presentó un

error de 49.55, ligeramente superior al de Tau, con un rango promedio estimado cercano a 1. Por su parte, el estimador R4 estimó un rango promedio de 1.52 y obtuvo un error de 57.30.

Figura 19 – Distribución del MSE para los datos bioquímicos obtenidos a partir de validación simple con toda la muestra ($n = 33$) y $d = 1$.

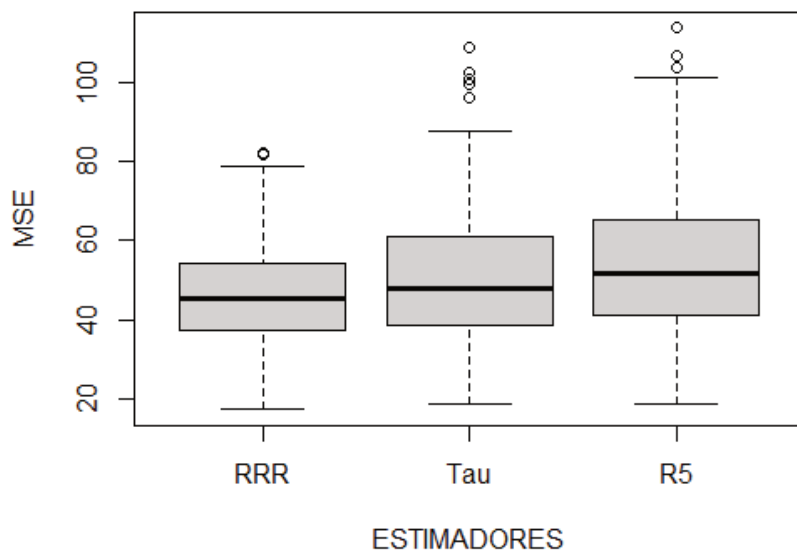
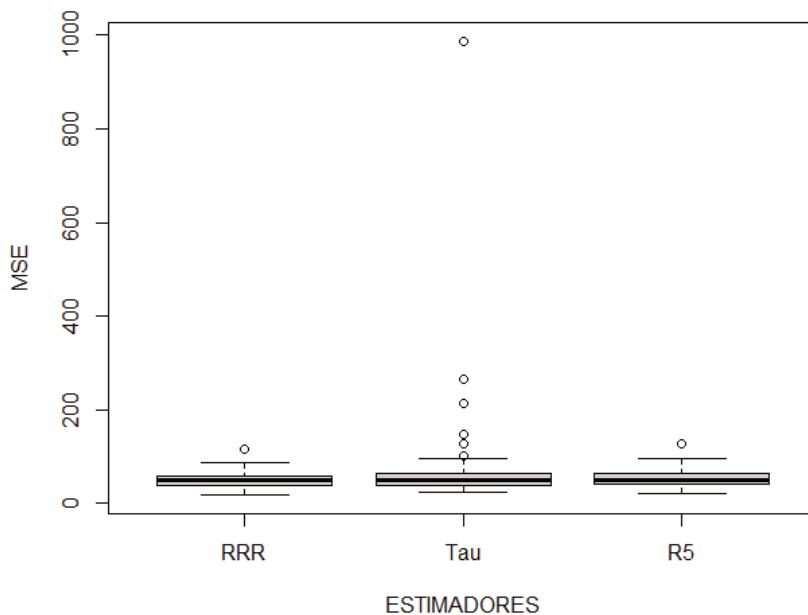


Figura 20 – Distribución del MSE para los datos bioquímicos obtenidos a partir de validación simple con toda la muestra ($n = 33$) y $d = 2$.



En las Figuras 19 y 20, se observa que los diagramas de caja de los estimadores presentan distribuciones bastante similares entre sí. En la Figura 19, el estimador RRR muestra un error más pequeño al del estimador Tau. Además, se pueden observar pocos valores atípicos.

En el caso $d = 2$, como se muestra en la Tabla 30 y en la Figura 20, el estimador RRR presenta el menor error, seguido muy de cerca por el estimador Tau, que se mantiene competitivo en su desempeño. Sin embargo, en el diagrama de cajas se observa que el Tau-estimador presenta varios valores atípicos.

Tabla 31. Errores (**Err**) para los datos bioquímicos obtenidos a partir de validación simple retirando los 5 valores atípicos y con diferentes rangos.

Estimadores	$d = 1$	$d = 2$
RRR	38.72	41.30
R5	40.22	43.53
Tau	45.19	50.16

Realizando el análisis bajo el MLM (1) para $d = 1$ y $d = 2$ y eliminando los 5 valores atípicos identificados, los resultados presentados en la Tabla 31 muestran una disminución general en los errores comparados con la Tabla 30.

Para los estimadores con rango 2, el estimador RRR presenta el menor error, con un valor de 41.30, lo que sugiere un mejor ajuste en ausencia de valores atípicos. Sin embargo, el estimador Tau, aunque con un error mayor (50.16), sigue siendo competitivo. Por otro lado, el estimador R5 presenta un error intermedio de 43.53, posicionándose como una alternativa menos favorable en esta configuración.

En el análisis con rango 1, el estimador RRR también obtiene el menor error (38.72), seguido por el estimador R5 con un error de 40.22. Aunque el estimador Tau presenta un mayor error (45.19) en esta configuración, estos resultados reflejan la estabilidad del estimador Tau en contextos con y sin valores atípicos, reforzando su aplicabilidad en datos con estructuras complejas.

Cabe destacar que el estimador Tau_d , que estima el rango automáticamente, no está incluido en la Tabla 31, pero su resultado (45.04) indica un error cercano al reportado por Tau, lo que lo posiciona como una alternativa viable en escenarios donde el rango no es conocido a priori. El análisis realizado para el estimador R4 arrojó un rango promedio estimado igual a 1.69 y un valor de **Err** igual a 49.62

Este análisis destaca que los modelos de rango 1 ofrecen un ajuste más preciso cuando se eliminan valores atípicos, y que los estimadores cercanos a este rango también se comportan de manera competitiva.

Por último, se llevó a cabo el análisis de LOOCV descrito en la Sección 5.1. El mismo se realizó utilizando todos los datos sin descontar los datos atípicos y se fijó un rango de $d = 1$ para evaluar el comportamiento de los estimadores. Dado que los resultados obtenidos para

el caso de $d = 2$ se asemejan a los obtenidos en el caso $d = 1$, se ha decidido no incluir los gráficos correspondientes.

Figura 21 – Valores predichos versus valores observados usando LOOCV, para la respuesta $y_{(\cdot,2)}$, utilizando toda la muestra ($n = 33$) y rango $d = 1$.

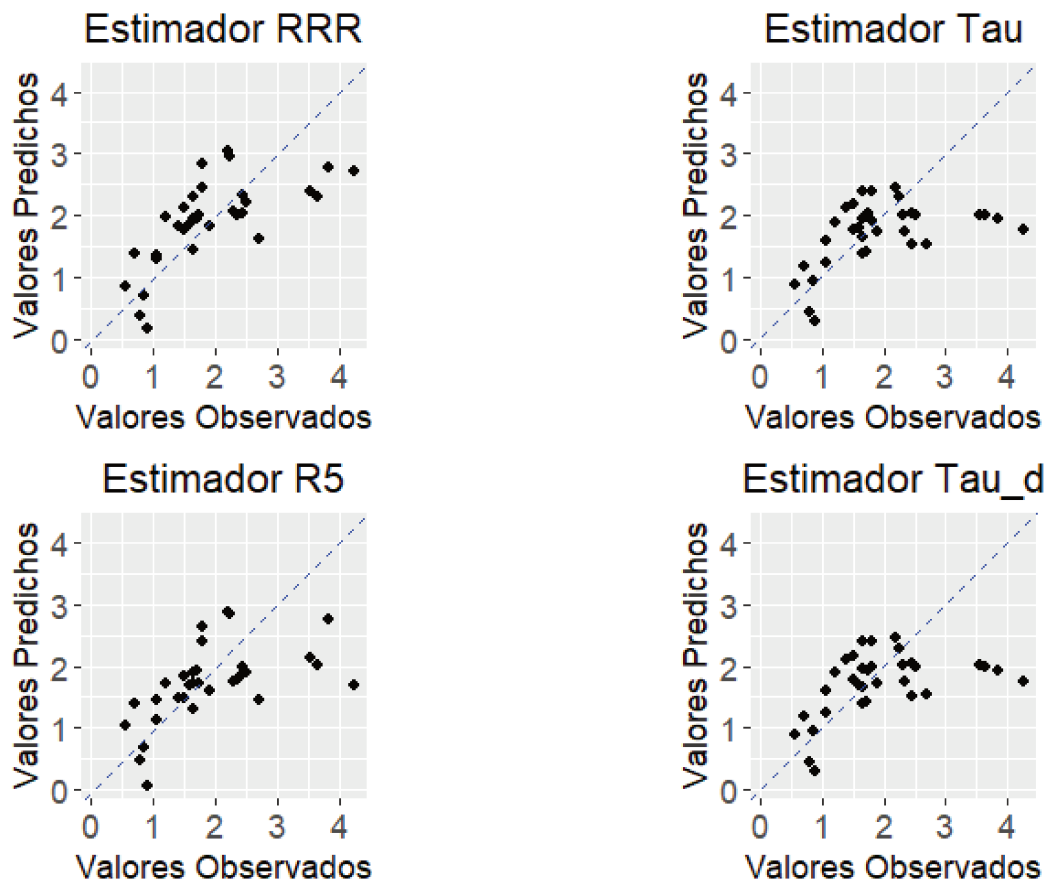


Figura 22 – Valores predichos versus valores observados usando LOOCV, para la respuesta $y_{(:,3)}$, utilizando toda la muestra ($n = 33$) y rango $d = 1$.

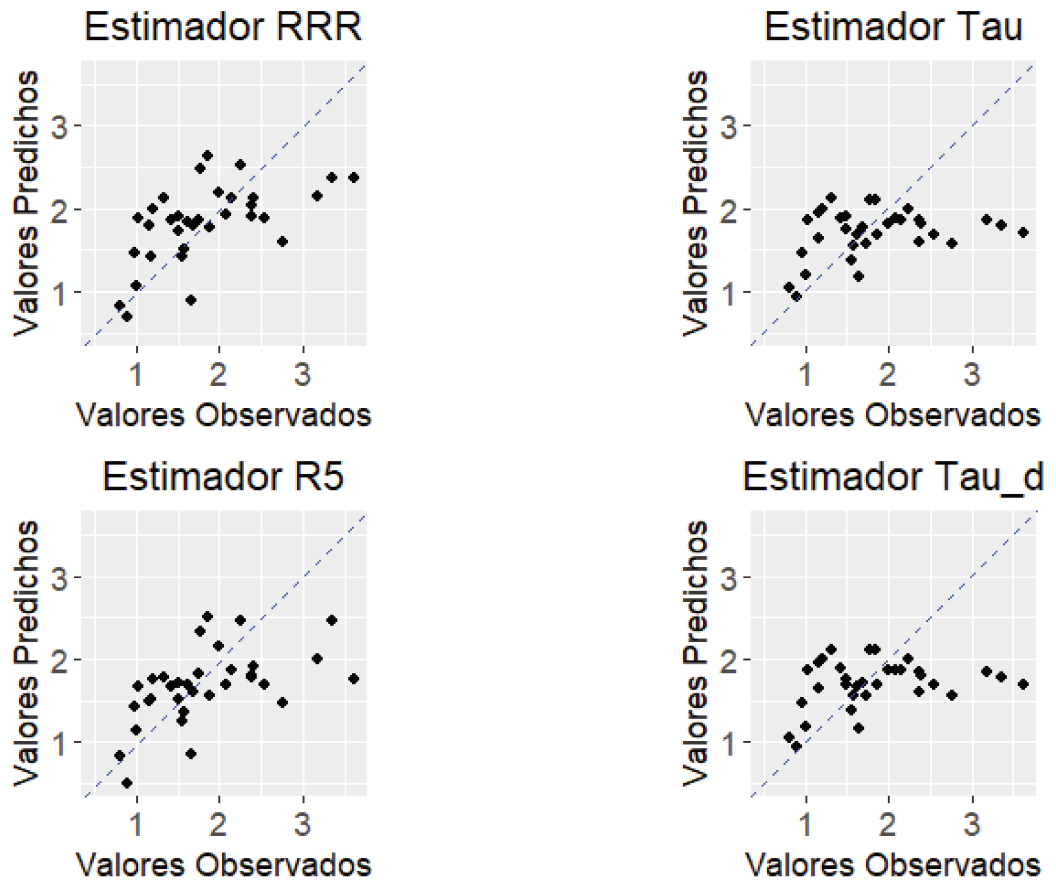
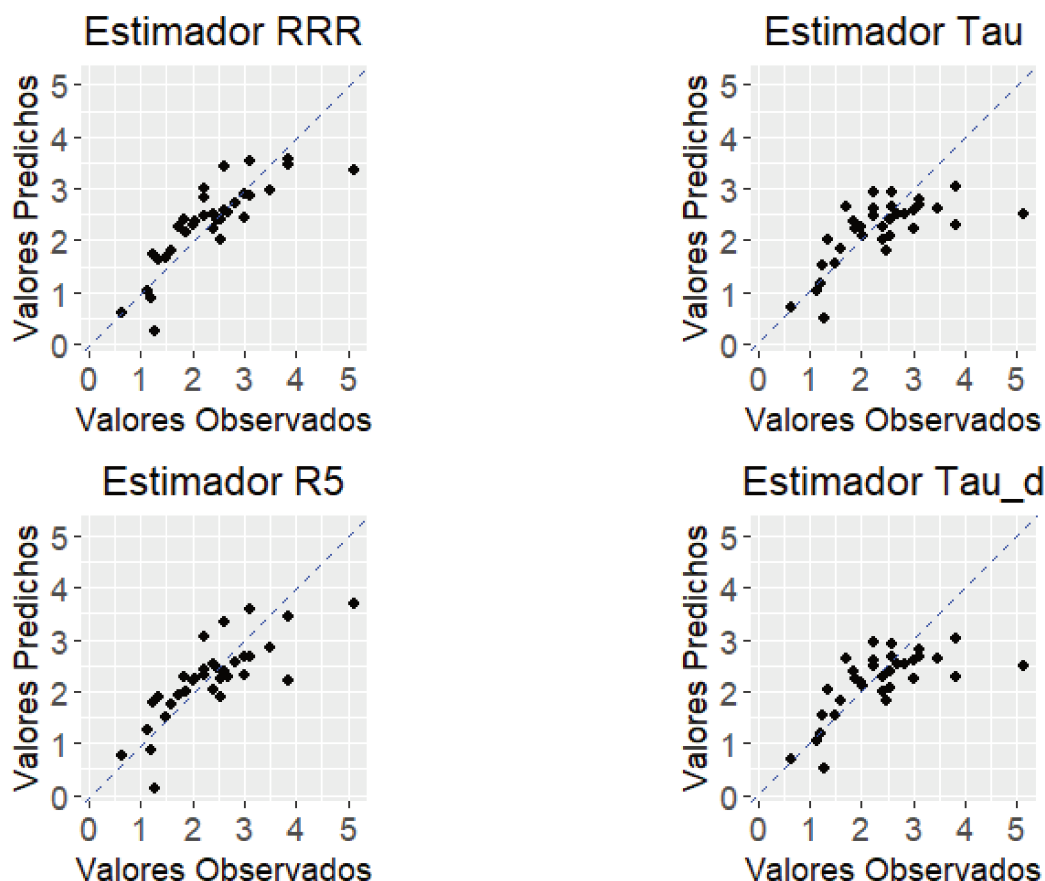


Figura 23 – Valores predichos versus valores observados usando LOOCV, para la respuesta $y_{(\cdot,4)}$, utilizando toda la muestra ($n = 33$) y rango $d = 1$.



Las Figuras 21-23 revelaron un buen ajuste para las variables respuesta $y_{(\cdot,2)}$, $y_{(\cdot,3)}$ y $y_{(\cdot,4)}$ ya que los puntos se acercan bastante a la línea diagonal. En contraste, $y_{(\cdot,1)}$ no exhibe un ajuste tan preciso, mientras que $y_{(\cdot,5)}$ muestra una relación notablemente más débil y opuesta. Por consiguiente, se decidió omitir la presentación de sus gráficos en este análisis. Entre los estimadores evaluados, Tau y Tau_d muestran un comportamiento similar.

La Tabla 32 muestra los coeficientes de regresión junto con sus errores estándar (SE) calculados a partir de un procedimiento bootstrap con 1000 muestras, tal como fue descrito en la Sección 5.1. Todos los estimadores muestran un comportamiento similar, con coeficientes que varían ligeramente pero no presentan patrones notables de dispersión, sólo pequeñas desviaciones en los coeficientes estimados dependiendo del rango utilizado o el valor de d , pero sin diferencias importantes en los errores estándar. Cuando los estimadores RRR, Tau, y R5 utilizan el mismo rango, sus estimaciones y errores estándar tienden a comportarse de manera más similar.

Tabla 32. Coeficientes de regresión y errores estándar para las estimaciones utilizando diferentes estimadores para diferentes rangos, utilizando toda la muestra ($n = 33$).

	MLE (Rango completo)		RRR ($d = 2$)		RRR ($d = 1$)		Tau (Rango completo)		Tau ($d = 2$)		Tau ($d = 1$)		Tau _d (Rango estimado:1)		R5 ($d = 2$)		R5 ($d = 1$)	
	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE
Y1																		
X1	-2.83	0.88	-2.66	0.75	-1.42	0.8	-2.69	1.5	-2.47	1.42	-1.22	1.47	-1.28	1.48	-2.89	0.71	-2.16	1.18
X2	1.96	0.43	2.02	0.42	1.38	0.57	2.03	0.86	1.82	0.95	1.1	0.91	1.13	0.92	2.02	0.38	1.51	0.85
X3	0.12	1.09	0.16	1.15	-1.3	0.75	-0.31	1.63	-0.62	1.61	-1.29	1.53	-1.26	1.56	-0.3	1	-0.67	1.31
Y2																		
X1	0.58	0.21	0.55	0.21	0.43	0.21	0.41	0.42	0.4	0.4	0.27	0.39	0.28	0.39	0.58	0.25	0.51	0.34
X2	-0.47	0.14	-0.48	0.13	-0.42	0.13	-0.43	0.22	-0.38	0.22	-0.28	0.22	-0.29	0.22	-0.44	0.15	-0.35	0.24
X3	0.29	0.25	0.28	0.22	0.41	0.19	0.31	0.35	0.38	0.32	0.42	0.34	0.41	0.34	0.25	0.22	0.27	0.36
Y3																		
X1	0.56	0.15	0.54	0.15	0.3	0.16	0.48	0.29	0.45	0.29	0.22	0.27	0.22	0.28	0.54	0.17	0.39	0.25
X2	-0.41	0.1	-0.42	0.09	-0.29	0.1	-0.39	0.17	-0.33	0.18	-0.2	0.16	-0.2	0.16	-0.39	0.11	-0.27	0.17
X3	-0.04	0.17	-0.03	0.17	0.27	0.12	-0.03	0.29	0.07	0.28	0.23	0.25	0.22	0.25	-0.01	0.19	0.15	0.25
Y4																		
X1	0.6	0.18	0.6	0.19	0.5	0.24	0.49	0.34	0.48	0.36	0.32	0.44	0.33	0.44	0.66	0.23	0.59	0.38
X2	-0.57	0.12	-0.56	0.12	-0.5	0.15	-0.49	0.21	-0.46	0.22	-0.34	0.24	-0.35	0.24	-0.51	0.15	-0.41	0.28
X3	0.36	0.15	0.37	0.15	0.48	0.21	0.45	0.26	0.48	0.27	0.49	0.36	0.49	0.36	0.39	0.18	0.32	0.42
Y5																		
X1	1.34	1.94	0.78	1.68	1.28	0.72	0.8	3.47	0.64	2.86	0.72	1.73	0.73	1.78	0.73	2.09	1.25	1.97
X2	-0.52	1.22	-0.91	1.14	-1.31	0.52	-0.56	1.8	-0.75	1.59	-0.85	1.08	-0.86	1.12	-0.24	1.28	-0.72	1.32
X3	3.06	1.77	2.58	1.72	1.5	0.98	2.66	2.63	2.33	2.2	1.6	1.54	1.6	1.55	3.15	1.75	1.8	1.94

5.3 Factores determinantes de niveles de retinol y beta-caroteno en plasma

Los datos utilizados en este ejemplo fueron extraídos de https://lib.stat.cmu.edu/datasets/Plasma_Retinol y analizados anteriormente en Nierenberg et al. (1989) y García Ben et al. (2006). La información contenida en estos datos fue recopilada con el objetivo de investigar el impacto de la ingesta dietética y las concentraciones plasmáticas de retinol, beta-caroteno y otros carotenoides en el riesgo de desarrollar ciertos tipos de cáncer. El estudio de Nierenberg et al. (1989) examinó los factores determinantes de las concentraciones plasmáticas de estos micronutrientes, analizando las asociaciones entre la concentración plasmática de beta-caroteno y retinol con once variables predictoras que se detallan a continuación. Esta información fue recopilada de 315 personas sometidas a procedimientos quirúrgicos electivos durante tres años para biopsiar o extirpar lesiones, que se determinaron como no cancerosas, en el pulmón, colon, mama, piel, ovario o útero.

Siguiendo la metodología propuesta por García Ben et al. (2006), se reexaminó la información utilizando un conjunto de 14 predictores, en lugar de los 12 originales, para analizar su influencia en los niveles plasmáticos de beta-caroteno y retinol. Algunas de las variables originales presentaban dos o tres categorías de respuesta. Se busca entonces comprender cómo estas variables influyen en las concentraciones plasmáticas de beta-caroteno y retinol, y cómo pueden utilizarse para predecirlas de manera adecuada.

Las variables de respuesta consideradas fueron

- $\mathbf{y}_{(.,1)}$ *betaplasma*: Beta-caroteno plasmático (ng/ml)
- $\mathbf{y}_{(.,2)}$ *retplasma*: Retinol plasmático (ng/ml)

y las variables predictoras se describen a continuación

- $\mathbf{x}_{(.,1)}$ Edad (años)
- $\mathbf{x}_{(.,2)}$ Masculino (1 = Masculino, 0 = Femenino)
- *smokstat*: Esta variable describe el estado de tabaquismo, donde 1 = nunca ha fumado, 2 = ex fumador y 3 = fumador actual.
 - $\mathbf{x}_{(.,3)}$ Ser fumador actual (1 si es fumador actual, 0 de lo contrario)
 - $\mathbf{x}_{(.,4)}$ Ser exfumador (1 si es ex fumador, 0 de lo contrario)
- $\mathbf{x}_{(.,5)}$ *quetelet*: Índice de masa corporal (peso/(altura²))
- *vituse*: Esta variable describe el uso de vitaminas, donde 1 = “Sí, con frecuencia”, 2 = “Sí, no con frecuencia” y 3 = “No”.
 - $\mathbf{x}_{(.,6)}$ Uso frecuente de vitaminas (1 si usa vitaminas con frecuencia, 0 de lo contrario)

- $\mathbf{x}_{(\cdot,7)}$ No usa vitaminas con frecuencia (1 si no usa vitaminas con frecuencia, 0 de lo contrario)
- $\mathbf{x}_{(\cdot,8)}$ *calorías*: Número de calorías consumidas por día
- $\mathbf{x}_{(\cdot,9)}$ *grasa*: Gramos de grasa consumidos por día
- $\mathbf{x}_{(\cdot,10)}$ *fibra*: Gramos de fibra consumidos por día
- $\mathbf{x}_{(\cdot,11)}$ *alcohol*: Número de bebidas alcohólicas consumidas por semana
- $\mathbf{x}_{(\cdot,12)}$ *colesterol*: Colesterol consumido (mg/día)
- $\mathbf{x}_{(\cdot,13)}$ *betadiet*: Beta-caroteno dietético consumido (mcg por día)
- $\mathbf{x}_{(\cdot,14)}$ *retdiet*: Retinol dietético consumido (mcg/día)

Para obtener estimadores con estos datos el primer paso consiste en estimar el rango de la reducción. Utilizando toda la muestra, se empleó la función `rank_trace` del paquete `rrr` en R, que estimó que el rango adecuado para el RRR es 1. Además, se calcularon los estimadores Tau_d y R4, los cuales obtuvieron un rango estimado de 1 en ambos casos.

Utilizando entonces el rango estimado igual a 1, considerando el enfoque propuesto por [García Ben et al. \(2006\)](#), una vez calculados los estimadores RRR y Tau, se identificaron los valores atípicos, considerando como valores atípicos aquellas observaciones cuyo valor de la distancia de Mahalanobis d_i supere el umbral crítico $\sqrt{\chi_{2,0.99}^2}$, donde $\chi_{2,0.99}^2$ es el percentil 0.99 de la distribución χ^2 con 2 grados de libertad, es decir, si $d_i > \sqrt{\chi_{2,0.99}^2}$ para $i = 1, \dots, 315$. La Figura 24 muestra un diagrama de cajas comparativo para estas distancias del estimador RRR y del τ -estimador. En este análisis, se observa que, en promedio, los residuos del τ -estimador son mayores que los del RRR. En efecto, el diagrama de cajas de la Figura 24 revela que el τ -estimador identifica una mayor cantidad de valores atípicos en comparación con el RRR. Específicamente, el RRR detecta 13 valores atípicos, mientras que el τ -estimador detecta 23. En contraste, en el estudio de [García Ben et al. \(2006\)](#), con un rango completo de 2, el τ -estimador detectó 27 valores atípicos. Esta diferencia en el número de valores atípicos puede atribuirse al uso del modelo de rango 1 y al programa `tauPFC`, utilizado para estimar el τ -estimador en este trabajo, implementa las funciones ρ_1 y ρ_2 , de la familia ρ -casi-óptima propuesta por [Muler and Yohai \(2003\)](#), mientras que [García Ben et al. \(2006\)](#) utilizó la familia bicuadrada de Tukey para su estimación del τ -estimador.

Para una evaluación del rendimiento de los estimadores, se llevó a cabo un análisis de validación simple con $N = 200$ iteraciones, tal como se describió en la Sección 5.1.

Los resultados obtenidos con toda la muestra ($n = 315$) y $d = 1$ muestran que el estimador Tau presenta el error más bajo, con un valor de 76305.21, lo que evidencia un buen desempeño en la predicción de los niveles de retinol y beta-caroteno en plasma. Por otro lado, el estimador R4 alcanzó un error de 91093.53, considerablemente superior al error de 77927.44 obtenido por el estimador RRR. Finalmente, el estimador Tau_d obtuvo un error de 76378.55, cercano al de Tau, con un rango estimado próximo a 1. Esto puede apreciarse en la Figura 25.

Figura 24 – Diagrama de caja de las distancia de Mahalanobis para los residuos calculados con los métodos RRR y Tau, con $d = 1$.

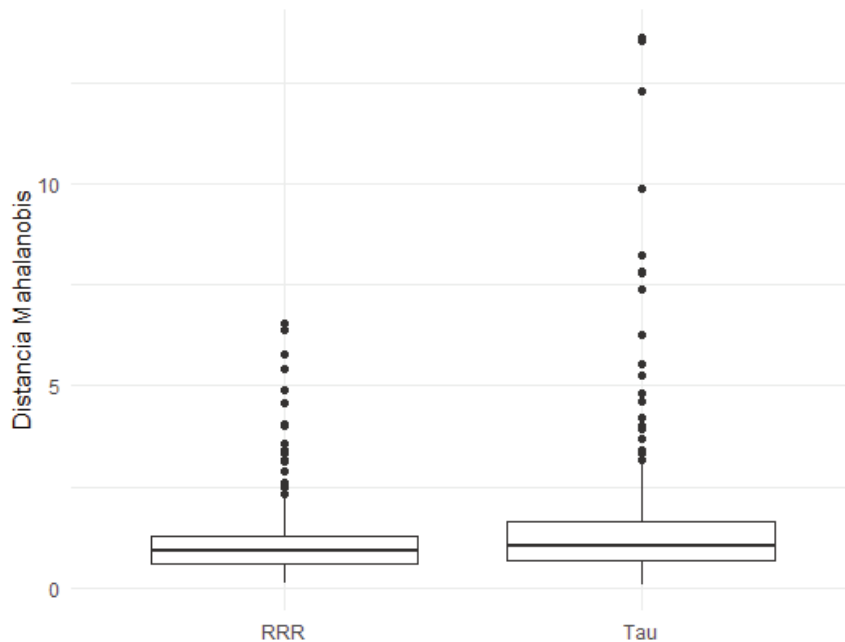
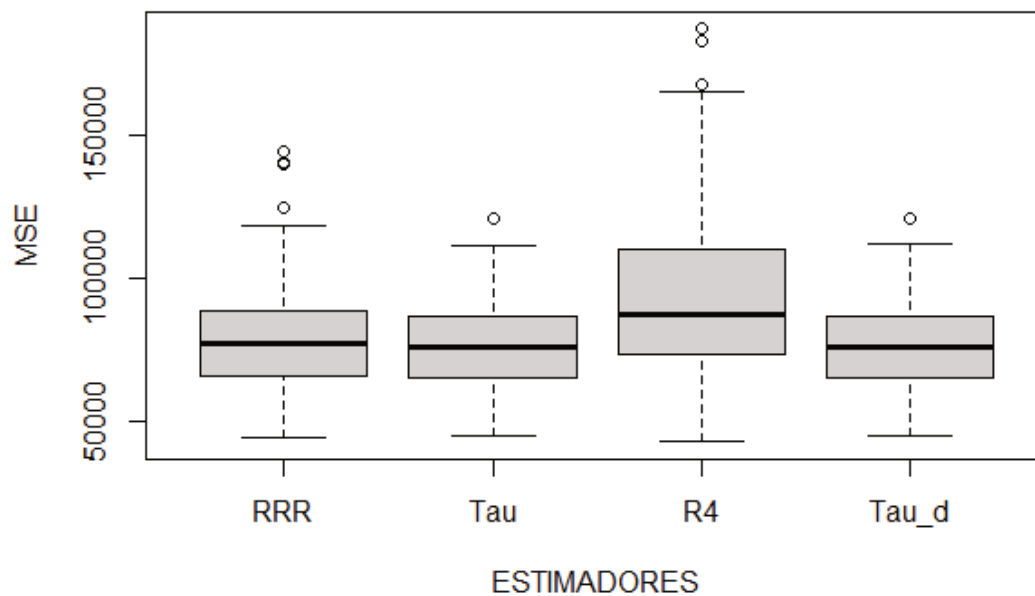


Figura 25 – Distribución del MSE para los datos de retinol y beta-caroteno ($n = 315$), con $d = 1$, excepto para R4 y Tau_d , que estiman el rango.



Para la validación simple (ver Sección 5.1) con $N = 200$, excluyendo los 23 valores atípicos identificados previamente, los resultados muestran que el estimador **Tau** presenta un desempeño muy cercano al estimador **RRR**, con un error (**Err**) de 41885.27, solo ligeramente superior al de **RRR**, que es 40873.01. Por otro lado, el estimador **R4** tiene un error de 52726.18, considerablemente mayor. Además, el estimador \mathbf{Tau}_d obtuvo un **Err** de 41914.57. Es importante señalar que tanto **R4** como \mathbf{Tau}_d estiman el rango, siendo el promedio de **R4** exactamente 1, mientras que el promedio de \mathbf{Tau}_d es cercano a 1. El estimador **R5** fue excluido del análisis debido a que presentó un error considerablemente mayor en comparación con los otros estimadores. Aunque una posible explicación podría estar relacionada con la presencia de variables dicotómicas en este conjunto de datos, no se cuenta con evidencia suficiente para confirmar esta hipótesis. Su desempeño inconsistente en este caso justifica su no inclusión en los resultados reportados.

En la Tabla 33 se presentan los coeficientes estimados de regresión para los estimadores **Tau**, \mathbf{Tau}_d , **RRR** y **R4**, junto con sus respectivos errores estándar (SE), calculados mediante el método bootstrap con 1000 muestras, como se detalla en la Sección 5.1, utilizando la muestra completa ($n = 315$). Se incluyeron únicamente aquellas variables que resultaron estadísticamente significativas a un nivel del 0.05 para al menos alguna de las dos variables respuestas, ajustadas mediante el método de mínimos cuadrados ordinarios. Para determinar la significancia estadística, se realizó un análisis basado en las pruebas de hipótesis sobre los coeficientes de las variables, evaluando si los valores p asociados eran menores a 0.05. En caso de que una variable fuera significativa en al menos una de las ecuaciones, esta se incluyó en el conjunto final de variables seleccionadas, independientemente de su desempeño en la otra ecuación. A su vez, con el objetivo de evaluar el comportamiento y la robustez de cada estimador, se emplearon $d = 1$ y $d = 2$, permitiendo una comparación detallada entre los resultados obtenidos. Este análisis permite ver las diferencias y similitudes entre los estimadores en la muestra original.

Para la respuesta BETAPLASMA, los estimadores **Tau** con $d = 1$ son los que presentan los SE más bajos, indicando mayor precisión en sus estimaciones, mientras que los estimadores **MLE** y **RRR** tienden a tener SE más altos en comparación con los demás. Aunque el estimador **R4** también muestra SE bajos, sus coeficientes estimados difieren considerablemente de los otros estimadores. Con la excepción del estimador **R4**, los coeficientes estimados con el resto de los estimadores son similares cuando los errores estándar son pequeños. Sin embargo, para los coeficientes con SE más elevados, las estimaciones tienden a diferir entre sí, se observan diferencias más marcadas, especialmente en las variables $\mathbf{x}_{(2)}$, $\mathbf{x}_{(6)}$ y $\mathbf{x}_{(10)}$, al comparar los estimadores **MLE**, **RRR**, **Tau** (con $d = 1$ y $d = 2$) y \mathbf{Tau}_d . Por otro lado, en RETPLASMA, las estimaciones de los coeficientes difieren considerablemente entre todos los estimadores, sin un patrón claro observable. Este comportamiento podría explicar los grandes errores **Err** obtenidos para todos los estimadores.

En la Tabla 34, se comparan los resultados de los estimadores con y sin la presencia de valores atípicos. En este análisis, los valores atípicos fueron eliminados únicamente de los estimadores **MLE** y **RRR**, mientras que **Tau** y \mathbf{Tau}_d utilizaron la muestra completa. A pesar de considerar todas las observaciones ($n = 315$), los resultados de los estimadores **Tau** y

Tau_d son comparables a los obtenidos con RRR cuando el rango utilizado fue 1, especialmente después de la eliminación de los 23 valores atípicos.

Sin la presencia de valores atípicos, los Tau estimadores funcionan adecuadamente en general para BETAPLASMA, con excepción de la variable $\mathbf{x}_{(6)}$. En este caso, los coeficientes estimados mediante Tau presentan SE más grandes, y sus valores difieren notablemente de los obtenidos con los estimadores MLE y RRR. Además, las estimaciones de los coeficientes tienden a ser más consistentes entre los diferentes métodos dependiendo del rango utilizado, ya sea completo o reducido. Esto sugiere que el rango influye en la consistencia de los resultados obtenidos por los estimadores.

Tabla 33. Coeficientes de regresión y errores estándar para las estimaciones utilizando diferentes estimadores para diferentes rangos, con toda la muestra ($n = 315$).

	MLE (Rango completo)		RRR ($d = 1$)		Tau (Rango completo)		Tau ($d = 1$)		Tau _d (Rango estimado:1)		R4 (Rango estimado:1)	
	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE
BETAPLASMA												
$\mathbf{x}_{(,1)}$	0.76	0.71	0.74	1	0.85	0.57	0.68	0.77	0.7	0.76	1.68	0.2
$\mathbf{x}_{(,2)}$	-34.28	27.13	-28.7	39.28	-37.2	18.94	-36.96	18.51	-36.87	18.62	0.42	0.91
$\mathbf{x}_{(,5)}$	-5.88	1.13	-4.96	1.94	-3.35	0.92	-3.11	1.01	-3.16	1	1.9	0.48
$\mathbf{x}_{(,6)}$	79.58	21.15	69.59	28.39	18.22	19.64	18.9	19.02	18.88	19.06	3.26	5.1
$\mathbf{x}_{(,10)}$	6.09	3.3	5.3	3.83	3.84	1.67	3.79	1.94	3.8	1.9	-0.37	1.23
$\mathbf{x}_{(,13)}$	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	≈0	≈0
RETPLASMA												
$\mathbf{x}_{(,1)}$	2.58	0.85	0.57	0.87	2.93	1.13	0.37	0.7	0.64	1.13	5.94	0.62
$\mathbf{x}_{(,2)}$	93.42	53.41	14.9	37.33	14.83	74.18	2.14	31.34	5.24	41.39	1.54	3.16
$\mathbf{x}_{(,5)}$	1.04	1.95	-0.26	2.12	0.43	2.14	0.13	1.99	0.18	2.02	6.81	1.32
$\mathbf{x}_{(,6)}$	30.65	28.81	13.85	31.67	0.15	40.84	1.56	15.44	1.64	20	11.08	16.42
$\mathbf{x}_{(,10)}$	-3.79	3.25	-0.26	2.61	-5.23	4.09	-0.93	2.7	-1.47	3.36	-1.54	3.32
$\mathbf{x}_{(,13)}$	≈ 0	0.01	≈ 0	0.01	≈ 0	0.01	≈ 0	0.01	≈ 0	0.01	≈ 0	0.01

Tabla 34. Coeficientes de regresión y errores estándar estimados con diferentes métodos para distintos rangos: comparación entre estimaciones sin 23 valores atípicos (MLE y RRR) y con toda la muestra ($n = 315$) (Tau y Tau_d)

	MLE (Rango completo)		RRR ($d = 1$)		Tau (Rango completo)		Tau ($d = 1$)		Tau _d (Rango estimado:1)	
	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE	Estimación	SE
BETAPLASMA										
$\mathbf{x}_{(\cdot,1)}$	0.92	0.39	0.82	0.49	0.85	0.57	0.68	0.77	0.7	0.76
$\mathbf{x}_{(\cdot,2)}$	-34.8	15.08	-35.56	15.83	-37.2	18.94	-36.96	18.51	-36.87	18.62
$\mathbf{x}_{(\cdot,5)}$	-3.85	0.75	-3.82	0.79	-3.35	0.92	-3.11	1.01	-3.16	1
$\mathbf{x}_{(\cdot,6)}$	32.69	12.32	31	12.67	18.22	19.64	18.9	19.02	18.88	19.06
$\mathbf{x}_{(\cdot,10)}$	3.15	1.34	3.27	1.47	3.84	1.67	3.79	1.94	3.8	1.9
$\mathbf{x}_{(\cdot,13)}$	0.01	≈ 0	0.01	≈ 0	0.01	0.01	0.01	0.01	0.01	0.01
RETPLASMA										
$\mathbf{x}_{(\cdot,1)}$	2.53	0.87	0.14	0.38	2.93	1.13	0.37	0.7	0.64	1.13
$\mathbf{x}_{(\cdot,2)}$	33.4	45.67	0.67	14.88	14.83	74.18	2.14	31.34	5.24	41.39
$\mathbf{x}_{(\cdot,5)}$	0.77	1.88	-0.04	1.43	0.43	2.14	0.13	1.99	0.18	2.02
$\mathbf{x}_{(\cdot,6)}$	20.62	28.04	1.58	12.26	0.15	40.84	1.56	15.44	1.64	20
$\mathbf{x}_{(\cdot,10)}$	-5.05	3.47	-0.18	1.32	-5.23	4.09	-0.93	2.7	-1.47	3.36
$\mathbf{x}_{(\cdot,13)}$	≈ 0	0.01	≈ 0	≈ 0	≈ 0	0.01	≈ 0	0.01	≈ 0	0.01

6 Conclusiones

En esta tesis, se ha extendido el uso de los τ -estimadores, una clase de estimadores robustos previamente desarrollada, al ámbito de los modelos lineales multivariados (MLM) de rango reducido. Esta aplicación novedosa permite la estimación conjunta de los coeficientes de regresión y de la matriz de covarianza de los errores en modelos de coeficientes de rango reducido. En datos simulados, estos estimadores demostraron ser altamente eficientes, mostrando una robustez significativa frente a datos atípicos. Además, cuando se aplicaron a datos no contaminados, los τ -estimadores exhibieron un comportamiento comparable al de los métodos clásicos, lo que valida su versatilidad.

El análisis asintótico de estos estimadores en dos escenarios —con coeficientes de regresión de rango completo y rango reducido— mostró que los τ -estimadores mantienen propiedades deseables, como la normalidad asintótica. Esto refuerza su aplicabilidad para la estimación precisa de parámetros en los modelos lineales multivariados.

Este trabajo se distingue de investigaciones anteriores al explorar el comportamiento de los τ -estimadores en modelos de rango reducido con matriz de covarianza desconocida, lo que ofrece un marco más realista y adaptable a una amplia gama de situaciones prácticas.

En resumen, los resultados obtenidos confirman tanto la eficacia como la robustez de los τ -estimadores en el contexto de los modelos lineales multivariados de rango reducido. Estos hallazgos representan una contribución al avance del conocimiento en la estimación robusta de parámetros y abren nuevas vías para futuras investigaciones.

Bibliografía

- Bergesio, A., Szretter, M., and Yohai, V. J. (2021). A robust proposal of estimation for the sufficient dimension reduction problem. TEST, 30. páginas 4, 9, 16, 29, 32, 34, 38
- Bergesio, A., Szretter, M., and Yohai, V. J. (2023). tauPFC: Computes Robust Estimators for the PFC Model. R package version 0.0.1. páginas 38
- Billingsley, P. (1968). Convergence of Probability Measures. Wiley Series in Probability and Mathematical Statistics. Wiley. páginas 121
- Bura, E. and Cook, R. D. (2003). Rank estimation in reduced-rank regression. Journal of Multivariate Analysis, 87(1):159–176. páginas 29
- Bura, E., Duarte, S., Forzani, L., Smucler, E., and Sued, M. (2018). Asymptotic theory for maximum likelihood estimates in reduced-rank multivariate generalized linear models. Statistics, 52(5):1005–1024. páginas 9, 29, 30, 31, 163
- Carroll, R. J. and Ruppert, D. (1985). Transformations in regression: A robust analysis. Technometrics, 27(1):1–12. páginas 8
- Cook, R. D., Forzani, L., and Zhang, X. (2015). Envelopes and reduced-rank regression. Biometrika, 102(2):439–456. páginas 29
- Croux, C. (1994). Efficient high-breakdown m-estimators of scale. Statistics & Probability Letters, 19(5):371–379. páginas 8
- Fairfield Smith, H., Gnanadesikan, R., and Hughes, J. B. (1962). Multivariate analysis of variance (manova). Biometrics, 18:22. páginas 75, 76
- Fasano, M. (2009). Robust estimation in nonlinear regression. PhD thesis, Universidad de La Plata. páginas 112, 115, 121
- García Ben, M., Martínez, E., and Yohai, V. J. (2006). Robust estimation for the multivariate linear model based on τ -scale. Journal of Multivariate Analysis, 97:1600–1622. páginas 4, 8, 9, 12, 15, 17, 19, 20, 22, 77, 84, 85
- Hössjer, O. (1992). On the optimality of s-estimators. Statistics & Probability Letters, 14:413–419. páginas 8
- Knight, K. (1999). Mathematical Statistics. Chapman & Hall/CRC Texts in Statistical Science. CRC Press. páginas 7
- Lopuhaa, H. (1991). Multivariate τ -estimators for location and scatter. Canad. J. Statist., 19:307–321. páginas 4, 17, 22, 32
- Lütkepohl, H. (1996). Handbook of Matrices. Graduate texts in mathematics. Springer. páginas 132

- Magnus, J. R. and Neudecker, H. (1980). The elimination matrix: Some lemmas and applications. SIAM J. Matrix Anal. Appl., 1(4):422–449. páginas 109
- Magnus, J. R. and Neudecker, H. (1999). Matrix Differential Calculus with Applications in Statistics and Econometrics. John Wiley, second edition. páginas 100, 134
- Maronna, R. A., Martin, R. D., and Yohai, V. J. (2006). Robust Statistics: Theory and Methods. John Wiley. páginas 12
- Maruskin, J. (2012). Essential Linear Algebra. Cayenne Canyon Press. páginas 131, 160, 164
- Muler, N. and Yohai, V. J. (2003). Robust estimates for arch processes. Journal of Time Series Analysis, 23. páginas 85
- Nierenberg, D., Stukel, T., Baron, J., Dain, B., and Greenberg, E. R. (1989). Determinants of plasma levels of beta-carotene and retinol. skin cancer prevention study group. American journal of epidemiology, 130 3:511–21. páginas 84
- Petersen, K. B. and Pedersen, M. S. (2012). The matrix cookbook. Version 20121115. páginas 100, 131
- She, Y. and Chen, K. (2017). Robust reduced-rank regression. Biometrika, 104(3):633–647. páginas 38, 40, 41, 63
- Szretter, M. (2017). Estimación robusta para un modelo de reducción de la dimensión. PhD thesis, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. páginas 32, 33
- Tatsuoka, K. and Tyler, D. (2000). The uniqueness of s and m-functionals under non-elliptical distributions. Ann. Statist., 28. páginas 38
- Velu, R. and Reinsel, G. (2013). Multivariate Reduced-Rank Regression: Theory and Applications. Lecture Notes in Statistics. Springer New York. páginas 29, 30, 38, 75, 76
- Yohai, V. J. (1974). Robust estimation in the linear model. The Annals of Statistics, 2(3):562–567. páginas 113
- Yohai, V. J. (1985). High breakdown point and high efficiency robust estimates for regression. Report 66, Department of Statistics, University of Washington. páginas 118, 121
- Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. The Annals of Statistics, 15(2):642–656. páginas 12, 15
- Yohai, V. J. and Zamar, R. (1986). High breakdown point estimates of regression by means of the minimization of an efficient scale. Report 84, Department of Statistics, University of Washington. páginas 115, 120

-
- Yohai, V. J. and Zamar, R. H. (1988). High breakdown-point estimates of regression by means of the minimization of an efficient scale. Journal of the American Statistical Association, 83(402):406–413. páginas 8
- Zamar, R. (1994). Estimación robusta. Estadística Española, 36(137):327 – 387. páginas 16
- Zhao, W., Lian, H., and Ma, S. (2017). Robust reduced-rank modeling via rank regression. Journal of Statistical Planning and Inference, 180:1–12. páginas 38, 62
- Zou, C., Ke, Y., and Zhang, W. (2022). Estimation of low rank high-dimensional multivariate linear models for multi-response data. Journal of the American Statistical Association, 117(538):693–703. páginas 29

A Apéndice del Capítulo 1

A.1 Demostración del Lema 1.1.1

Demostración del Lema. Antes de comenzar formalmente la demostración, se observan algunas propiedades: para $k > 0$

- i. $\ln |k\Sigma| = \ln k^q |\Sigma| = q \ln k + \ln |\Sigma|$,
- ii. $\frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, k\Sigma) = \frac{1}{n} \sum_{i=1}^n \mathbf{u}_i^t (k\Sigma)^{-1} \mathbf{u}_i = \frac{1}{k} \left(\frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, \Sigma) \right)$.

Para $c > 0$, se denotará $(\hat{\mathbf{B}}_c, \hat{\Sigma}_c)$ a la solución del siguiente problema

$$\min_{\mathbf{B}, \Sigma} |\Sigma| \text{ sujeto a } \frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, \Sigma) = c. \quad (\text{A.1})$$

Evaluando una constante veces la función de verosimilitud y llamándola nuevamente \mathcal{L} , en $(\hat{\mathbf{B}}_q, \hat{\Sigma}_q)$, se obtiene

$$\begin{aligned} \mathcal{L}(\hat{\mathbf{B}}_q, \hat{\Sigma}_q) &= \ln |\hat{\Sigma}_q| + \frac{1}{n} \sum_{i=1}^n d_i^2(\hat{\mathbf{B}}_q, \hat{\Sigma}_q) && \text{Definición 1.1.1} \\ &= \ln |\hat{\Sigma}_q| + q && \text{porque cumple la restricción de (1.4).} \end{aligned}$$

Por otra parte, para $c > 0$ se observa que el valor $\left(\hat{\mathbf{B}}_c, \frac{c}{q} \hat{\Sigma}_c \right)$ cumple que

$$\frac{1}{n} \sum_{i=1}^n d_i^2\left(\hat{\mathbf{B}}_c, \frac{c}{q} \hat{\Sigma}_c\right) = \frac{q}{c} \left(\frac{1}{n} \sum_{i=1}^n d_i^2(\hat{\mathbf{B}}_c, \hat{\Sigma}_c) \right) = q,$$

es decir, $\left(\hat{\mathbf{B}}_c, \frac{c}{q} \hat{\Sigma}_c \right)$ también satisface la restricción del problema (1.4) pero como, $\hat{\Sigma}_q$

minimiza el determinante, entonces $\mathcal{L}(\hat{\mathbf{B}}_q, \hat{\Sigma}_q) \leq \ln \left(\left| \frac{c}{q} \hat{\Sigma}_c \right| \right) + q$, usando la propiedad i., para todo $c > 0$ resulta que,

$$\begin{aligned} \mathcal{L}(\hat{\mathbf{B}}_q, \hat{\Sigma}_q) &\leq q \ln c - q \ln q + q + \ln |\hat{\Sigma}_c|, \\ &\leq c + \ln(|\hat{\Sigma}_c|) = \mathcal{L}(\hat{\mathbf{B}}_c, \hat{\Sigma}_c), \end{aligned} \quad (\text{A.2})$$

ya que $r(q) = q \ln c - q \ln q + q$ alcanza su máximo cuando $q = c$. De manera general, sean $(\mathbf{B}, \Sigma) \in \Theta$ cualesquiera, que satisfacen el modelo (1) con \mathbf{u} el vector de errores que tiene una distribución normal multivariada y además, $\frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, \Sigma) > 0$ con probabilidad uno (por ser \mathbf{u} normal), si se denota

$$c_1 := \frac{1}{n} \sum_{i=1}^n d_i^2(\mathbf{B}, \Sigma) > 0, \quad (\text{A.3})$$

y usando la definición (A.1) para c_1 se obtiene que,

$$\mathcal{L}(\widehat{\mathbf{B}}_{c_1}, \widehat{\boldsymbol{\Sigma}}_{c_1}) \leq \mathcal{L}(\mathbf{B}, \boldsymbol{\Sigma}), \quad (\text{A.4})$$

ya que $(\widehat{\mathbf{B}}_{c_1}, \widehat{\boldsymbol{\Sigma}}_{c_1})$ minimiza la función \mathcal{L} . Sin embargo, usando (A.2) para $c = c_1$ se tiene, por (A.4) que

$$\mathcal{L}(\widehat{\mathbf{B}}_q, \widehat{\boldsymbol{\Sigma}}_q) \leq \mathcal{L}(\widehat{\mathbf{B}}_{c_1}, \widehat{\boldsymbol{\Sigma}}_{c_1}) \leq \mathcal{L}(\mathbf{B}, \boldsymbol{\Sigma}). \quad (\text{A.5})$$

Esto indica que $(\widehat{\mathbf{B}}_q, \widehat{\boldsymbol{\Sigma}}_q)$ es el estimador MLE.

Recíprocamente, sea $(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}})$ el estimador MLE, se debe probar que es solución del problema (1.4). Lo primero a demostrar es que dicho estimador satisface la restricción del problema (1.4). Denotando $\widehat{\mathbf{u}}_i = \mathbf{u}_i(\widehat{\mathbf{B}})$ se tiene que,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n d_i^2(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}}) &= \frac{1}{n} \sum_{i=1}^n (\widehat{\mathbf{u}}_i^t \widehat{\boldsymbol{\Sigma}}^{-1} \widehat{\mathbf{u}}_i) \\ &= \frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{u}}_i^t \left(\frac{1}{n} \sum_{i=1}^n \widehat{\mathbf{u}}_i \widehat{\mathbf{u}}_i^t \right)^{-1} \widehat{\mathbf{u}}_i \quad \text{usando (3)} \\ &= \sum_{i=1}^n \text{tr} \left(\widehat{\mathbf{u}}_i^t \left(\sum_{i=1}^n \widehat{\mathbf{u}}_i \widehat{\mathbf{u}}_i^t \right)^{-1} \widehat{\mathbf{u}}_i \right) \\ &= \text{tr} \left(\sum_{i=1}^n \widehat{\mathbf{u}}_i \widehat{\mathbf{u}}_i^t \left(\sum_{i=1}^n \widehat{\mathbf{u}}_i \widehat{\mathbf{u}}_i^t \right)^{-1} \right) = q. \end{aligned}$$

Como $(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}})$ es MLE, se cumple $\mathcal{L}(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}}) \leq \mathcal{L}(\mathbf{B}, \boldsymbol{\Sigma})$ para cada $(\mathbf{B}, \boldsymbol{\Sigma}) \in \boldsymbol{\Theta}$. Sea $(\mathbf{B}_q, \boldsymbol{\Sigma}_q)$ tal que cumple la restricción (1.4), entonces

$$\begin{aligned} \mathcal{L}(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}}) &= \ln |\widehat{\boldsymbol{\Sigma}}| + q \\ &\leq \mathcal{L}(\mathbf{B}_q, \boldsymbol{\Sigma}_q) = \ln |\boldsymbol{\Sigma}_q| + q, \end{aligned}$$

así, $\ln |\widehat{\boldsymbol{\Sigma}}| \leq \ln |\boldsymbol{\Sigma}_q|$ y como la función logaritmo (\ln) es una función creciente entonces $|\widehat{\boldsymbol{\Sigma}}| \leq |\boldsymbol{\Sigma}_q|$, obteniendo que $(\widehat{\mathbf{B}}, \widehat{\boldsymbol{\Sigma}})$ es solución del problema (1.4). \square

A.2 Demostración del Lema 1.2.1

Demostración del Lema. Sea P una probabilidad que satisface la propiedad (\mathcal{S}_ϵ) , se quiere probar que $P(\{(\mathbf{x}, \mathbf{y}) : d(\mathbf{B}, \boldsymbol{\Sigma}) = 0\}) < \epsilon$. Se conoce que $\boldsymbol{\Sigma} \in \mathcal{S}_q^+$, por tanto el conjunto de valores (\mathbf{x}, \mathbf{y}) que satisfacen la igualdad $d(\mathbf{B}, \boldsymbol{\Sigma}) = 0$, son los mismos que satisfacen la

igualdad $\mathbf{y} - \mathbf{B}^t \mathbf{x} = \mathbf{0}$. Además, tomando la matriz $\mathbf{B}^t = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_q \end{pmatrix}$ con $\mathbf{b}_j^t \in \mathbb{R}^p$ con $j = 1, \dots, q$,

y considerando a $\mathbf{e}_i^t \in \mathbb{R}^q$ vectores unitarios con $i = 1, \dots, q$, por la propiedad (\mathcal{S}_ϵ) , los hiperplanos $H_{(-\mathbf{b}_1, \mathbf{e}_1, 0)}, \dots, H_{(-\mathbf{b}_q, \mathbf{e}_q, 0)}$ verifican

$$\begin{aligned} P(H_{(-\mathbf{b}_1, \mathbf{e}_1, 0)}) &< \epsilon, \\ &\vdots \\ P(H_{(-\mathbf{b}_q, \mathbf{e}_q, 0)}) &< \epsilon. \end{aligned}$$

Luego, utilizando dichas desigualdades, se tiene que

$$\begin{aligned} P(\{(\mathbf{x}, \mathbf{y}) : d(\mathbf{B}, \boldsymbol{\Sigma}) = 0\}) &= P(\{(\mathbf{x}, \mathbf{y}) : \mathbf{y} - \mathbf{B}^t \mathbf{x} = \mathbf{0}\}) \\ &= P(H_{(-\mathbf{b}_1, \mathbf{e}_1, 0)} \cap H_{(-\mathbf{b}_2, \mathbf{e}_2, 0)} \cap \dots \cap H_{(-\mathbf{b}_q, \mathbf{e}_q, 0)}) \\ &< \epsilon, \end{aligned}$$

obteniendo la desigualdad deseada. □

A.3 Demostración del Lema 1.2.2

Demostración del Lema. Sea $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$ fijo y sea $S_{(\mathbf{B}, \boldsymbol{\Sigma})} : \mathbb{R}_{>0} \rightarrow \mathbb{R}$ definida como

$$S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s) = E_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \left[\rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s} \right) \right].$$

En primer lugar, $S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s)$ es una función continua. Esta característica se deriva del Teorema de Convergencia Dominada y se fundamenta en la estructura de la función como una composición de funciones, todas ellas continuas en sus respectivos dominios. Luego, empleando el Teorema de Convergencia Dominada y las propiedades (A1)-(A5) de la ρ -función, resulta

$$\begin{aligned} \lim_{s \rightarrow +\infty} S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s) &= \lim_{s \rightarrow +\infty} E_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \left[\rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s} \right) \right] = 0, \\ \lim_{s \rightarrow 0^+} S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s) &= \lim_{s \rightarrow 0^+} \int \rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s} \right) d_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \\ &= \lim_{s \rightarrow 0^+} \int_{\{d(\mathbf{B}, \boldsymbol{\Sigma}) > 0\}} \rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s} \right) d_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \\ &= aP(d(\mathbf{B}, \boldsymbol{\Sigma}) > 0) \\ &= a[1 - P(d(\mathbf{B}, \boldsymbol{\Sigma}) = 0)]. \end{aligned}$$

Ya que P satisface (S_ϵ) para algún $\epsilon \leq 1 - \frac{\kappa}{a}$, entonces para $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$ se tiene que $a[1 - P(d(\mathbf{B}, \boldsymbol{\Sigma}) = 0)] > \kappa$.

Como $S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s)$ es continua, para $\kappa > 0$ existe $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}})$ que satisface (1.7), es decir, existe un M-funcional de escala positivo $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}})$ para MLM, tal que $S_{(\mathbf{B}, \boldsymbol{\Sigma})}(s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}})) = \kappa$.

Para demostrar la unicidad de (1.7), se procede por contradicción. Se asume que existen s_1 y s_2 con $s_1 < s_2$ tales que

$$E_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \left[\rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2} \right) \right] = E_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \left[\rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_1} \right) \right] = \kappa, \quad (\text{A.6})$$

es decir,

$$E_{H_{P, \mathbf{B}, \boldsymbol{\Sigma}}} \left[\rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2} \right) - \rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_1} \right) \right] = 0,$$

y ya ρ es una función creciente, en casi todo punto se tiene que,

$$\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2}\right) \leq \rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_1}\right),$$

entonces

$$\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2}\right) = \rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_1}\right),$$

para $d(\mathbf{B}, \boldsymbol{\Sigma}) > 0$ con $s_1 < s_2$. Usando A6) $\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2}\right) = a$, esto implica que

$$\begin{aligned} \kappa &= E_{H_{P,\mathbf{B},\boldsymbol{\Sigma}}} \left[\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2}\right) \right] \\ &= \int_{\{d(\mathbf{B}, \boldsymbol{\Sigma}) > 0\}} \rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s_2}\right) d_{H_{P,\mathbf{B},\boldsymbol{\Sigma}}} \\ &= a[1 - P(d(\mathbf{B}, \boldsymbol{\Sigma}) = 0)]. \end{aligned}$$

Como por hipótesis P satisface (S_ϵ) y utilizando el Lema 1.2.1, se tiene que $P[d(\mathbf{B}, \boldsymbol{\Sigma}) = 0] < 1 - \frac{\kappa}{a}$ y entonces se produce una contradicción, que vino de suponer (A.6). Teniendo en cuenta este resultado, la solución de la ecuación (1.7) es única. \square

A.4 Demostración del Lema 1.2.3

Demostración del Lema. Sea $\lambda > 0$, por definición de $s(H_{P,\mathbf{B},\boldsymbol{\Sigma}})$ dada en (1.7), se tiene que

$$\begin{aligned} \kappa &= E_{H_{P,\mathbf{B},\boldsymbol{\Sigma}}} \left[\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(H_{P,\mathbf{B},\boldsymbol{\Sigma}})}\right) \right] \\ &= E_{H_{P,\mathbf{B},\boldsymbol{\Sigma}}} \left[\rho\left(\frac{\sqrt{\lambda}(\mathbf{u}^t(\mathbf{B})(\lambda\boldsymbol{\Sigma})^{-1}\mathbf{u}(\mathbf{B}))^{1/2}}{s(H_{P,\mathbf{B},\boldsymbol{\Sigma}})}\right) \right] && \text{por (1.1)} \\ &= E_{H_{P,\mathbf{B},\lambda\boldsymbol{\Sigma}}} \left[\rho\left(\frac{\sqrt{\lambda}d(\mathbf{B}, \lambda\boldsymbol{\Sigma})}{s(H_{P,\mathbf{B},\boldsymbol{\Sigma}})}\right) \right], \end{aligned}$$

por lo tanto, para $\kappa > 0$ se tiene que $s(H_{P,\mathbf{B},\lambda\boldsymbol{\Sigma}})$ existe y

$$s(H_{P,\mathbf{B},\lambda\boldsymbol{\Sigma}}) = \frac{1}{\sqrt{\lambda}}s(H_{P,\mathbf{B},\boldsymbol{\Sigma}}), \tag{A.7}$$

obteniendo el primer resultado deseado. Para la demostración de la segunda igualdad del Lema, se usa la definición de τ -funcional de escala dada en (1.10) para $H_{P,\mathbf{B},\lambda\Sigma}$,

$$\begin{aligned}
 \tau^2(H_{P,\mathbf{B},\lambda\Sigma}) &= s^2(H_{P,\mathbf{B},\lambda\Sigma}) E_{H_{P,\mathbf{B},\lambda\Sigma}} \left(\rho_2 \left(\frac{d(\mathbf{B}, \lambda\Sigma)}{s(H_{P,\mathbf{B},\lambda\Sigma})} \right) \right) \\
 &= \frac{1}{\lambda} s^2(H_{P,\mathbf{B},\Sigma}) E_{H_{P,\mathbf{B},\lambda\Sigma}} \left(\rho_2 \left(\frac{d(\mathbf{B}, \lambda\Sigma)}{\frac{1}{\sqrt{\lambda}} s(H_{P,\mathbf{B},\Sigma})} \right) \right) && \text{por (A.7),} \\
 &= \frac{1}{\lambda} s^2(H_{P,\mathbf{B},\Sigma}) E_P \left(\rho_2 \left(\frac{(\mathbf{u}^t(\mathbf{B})(\Sigma)^{-1} \mathbf{u}(\mathbf{B}))^{1/2}}{s(H_{P,\mathbf{B},\Sigma})} \right) \right) && \text{por (1.1),} \\
 &= \frac{1}{\lambda} s^2(H_{P,\mathbf{B},\Sigma}) E_{H_{P,\mathbf{B},\Sigma}} \left(\rho_2 \left(\frac{d(\mathbf{B}, \Sigma)}{s(H_{P,\mathbf{B},\Sigma})} \right) \right) \\
 &= \frac{1}{\lambda} \tau^2(H_{P,\mathbf{B},\Sigma}) && \text{usando (1.10) para } H_{P,\mathbf{B},\Sigma}.
 \end{aligned}$$

□

A.5 Demostración del Teorema 2

A continuación se establecen algunos resultados básicos que serán útiles para probar el Teorema 2.

Lema A.5.1. *Las derivadas de $d_i(\mathbf{B}, \Sigma)$ con respecto a $\text{vec}^t(\mathbf{B}^t)$, $\text{vech}^t(\Sigma)$ son,*

$$\frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} = - \frac{1}{d_i(\mathbf{B}, \Sigma)} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}), \quad (\text{A.8})$$

$$\frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)} = - \frac{1}{2d_i(\mathbf{B}, \Sigma)} [(\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})] \mathbf{D}_q, \quad (\text{A.9})$$

donde \mathbf{D}_q es la matriz de duplicación de tamaño $q^2 \times q(q+1)/2$, que conecta el operador vec y el operador vech como sigue $\mathbf{D}_q \text{vech}(\mathbf{A}) = \text{vec}(\mathbf{A})$, para cualquier matriz $\mathbf{A} \in S^q$.

Demostración del Lema A.5.1.

$$\begin{aligned}
 \frac{\partial d_i(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} &= \frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))^{1/2}}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &= \frac{1}{2((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))^{1/2}} \frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &= \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \left[(\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i)^t \boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i) \right] \\
 &= \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}^t \left[-2\boldsymbol{\Sigma}^{-1} (\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i) \mathbf{x}_i^t \right] \quad \text{resultado de (Petersen and Pedersen, 2012)} \\
 &= \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}^t \left[-2\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \right] \\
 &= \frac{1}{d_i(\mathbf{B}, \boldsymbol{\Sigma})} \left[-(\mathbf{x}_i \otimes \mathbf{I}_q) (\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B})) \right]^t \quad \text{por vectorización de matrices} \\
 &= - \frac{(\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})(\mathbf{x}_i^t \otimes \mathbf{I}_q)}{d_i(\mathbf{B}, \boldsymbol{\Sigma})}.
 \end{aligned}$$

Usando definición de vectorización de matrices y propiedades del producto de Kronecker resulta

$$\begin{aligned}
 (\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})(\mathbf{x}_i^t \otimes \mathbf{I}_q) &= \text{vec}^t(\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t) \\
 &= (\mathbf{x}_i \otimes \boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}))^t \\
 &= \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1}),
 \end{aligned}$$

en consecuencia

$$\frac{\partial d_i(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} = - \frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})}{d_i(\mathbf{B}, \boldsymbol{\Sigma})}.$$

Por otro lado, para obtener la derivada de $d_i(\mathbf{B}, \boldsymbol{\Sigma})$ con respecto a $\text{vech}^t(\boldsymbol{\Sigma})$ se usarán resultados presentes en (Petersen and Pedersen, 2012) y (Magnus and Neudecker, 1999). Así,

$$\begin{aligned}
 \frac{\partial d_i(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} &= \frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))^{1/2}}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\
 &= \frac{1}{2((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))^{1/2}} \frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))}{\partial \text{vech}^t(\boldsymbol{\Sigma})},
 \end{aligned}$$

luego, por la regla de la cadena se tiene que,

$$\frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} = \frac{\partial ((\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})\mathbf{u}_i(\mathbf{B}))}{\partial \text{vec}^t(\boldsymbol{\Sigma})} \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})}.$$

En consecuencia,

$$\begin{aligned}
 \frac{\partial d_i(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} &= \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}^t(-\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1}) \mathbf{D}_q \\
 &= \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \left[-(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \text{vec}(\mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B})) \right]^t \mathbf{D}_q \\
 &= - \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_q \\
 &= - \frac{1}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} [(\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B})\boldsymbol{\Sigma}^{-1})] \mathbf{D}_q.
 \end{aligned}$$

□

En lo que sigue, se asumirá que se cumplen las hipótesis del Teorema 2.

Lema A.5.2. *Las derivadas de la M-escala $s_n(\mathbf{B}, \Sigma)$ (para $\rho = \rho_1$ y $\kappa = \kappa_1$) con respecto a $vec^t(\mathbf{B}^t)$ y $vech^t(\Sigma)$ son,*

$$\frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)} = - \frac{s_n(\mathbf{B}, \Sigma) \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{1}{d_i(\mathbf{B}, \Sigma)} (\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}))}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_i(\mathbf{B}, \Sigma)}, \quad (\text{A.10})$$

$$\frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} = - \frac{s_n(\mathbf{B}, \Sigma) \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{1}{2d_i(\mathbf{B}, \Sigma)} (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \mathbf{D}_q}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_i(\mathbf{B}, \Sigma)}. \quad (\text{A.11})$$

Demostración del Lema A.5.2. Para obtener la derivada de $s_n(\mathbf{B}, \Sigma)$ con respecto a $vec^t(\mathbf{B}^t)$, a partir de la definición de la M-escala dada en (1.11), se obtiene

$$\frac{\partial \kappa_1}{\partial vec^t(\mathbf{B}^t)} = \frac{\partial}{\partial vec^t(\mathbf{B}^t)} \left(\frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \right).$$

Como κ_1 es una constante, se tiene,

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial vec^t(\mathbf{B}^t)} \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \right) \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{\partial}{\partial vec^t(\mathbf{B}^t)} \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) && \text{por regla de la cadena} \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left[\frac{\frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)} s_n(\mathbf{B}, \Sigma) - d_i(\mathbf{B}, \Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)}}{s_n^2(\mathbf{B}, \Sigma)} \right] && \text{por regla del cociente} \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left[\frac{\left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) s_n(\mathbf{B}, \Sigma) - d_i(\mathbf{B}, \Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)}}{s_n^2(\mathbf{B}, \Sigma)} \right] && \text{por (A.8)} \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left[\left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) \frac{1}{s_n(\mathbf{B}, \Sigma)} - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)} \right]. \end{aligned}$$

Se despeja la derivada de $s_n(\mathbf{B}, \Sigma)$ con respecto a \mathbf{B}^t , obteniendo

$$\begin{aligned} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vect}(\mathbf{B}^t)} &= \frac{\frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) \frac{1}{s_n(\mathbf{B}, \Sigma)}}{\frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)}} \\ &= \frac{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right)}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)}}. \end{aligned}$$

Por otro lado, para la derivada de $s_n(\mathbf{B}, \Sigma)$ con respecto a $\text{vech}^t(\Sigma)$ se utiliza un procedimiento similar al anterior, y se obtiene

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\partial}{\partial \text{vech}^t(\Sigma)} \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \right) \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{\partial}{\partial \text{vech}^t(\Sigma)} \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) && \text{por regla de la cadena} \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left[\frac{\frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)} s_n(\mathbf{B}, \Sigma) - d_i(\mathbf{B}, \Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)}}{s_n^2(\mathbf{B}, \Sigma)} \right] && \text{por regla del cociente} \\ &= \frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left[\left(-\frac{(\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{2d_i(\mathbf{B}, \Sigma)} \mathbf{D}_q \right) \frac{1}{s_n(\mathbf{B}, \Sigma)} \right. \\ &\quad \left. - \frac{d_i(\mathbf{B}, \Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)}}{s_n^2(\mathbf{B}, \Sigma)} \right] && \text{por (A.9).} \end{aligned}$$

Despejando la derivada de $s_n(\mathbf{B}, \Sigma)$ respecto a $\text{vech}^t(\Sigma)$ resulta,

$$\begin{aligned} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)} &= \frac{\frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(-\frac{(\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{2d_i(\mathbf{B}, \Sigma)} \mathbf{D}_q \right) \frac{1}{s_n(\mathbf{B}, \Sigma)}}{\frac{1}{n} \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)}} \\ &= -\frac{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(\frac{(\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{2d_i(\mathbf{B}, \Sigma)} \mathbf{D}_q \right)}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)}}. \end{aligned}$$

□

Lema A.5.3. Las derivadas de $d_i^*(\mathbf{B}, \Sigma)$ con respecto a $\text{vec}^t(\mathbf{B}^t)$ y $\text{vech}^t(\Sigma)$ son,

$$\frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} = \frac{1}{s_n(\mathbf{B}, \Sigma)} \left[-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} + d_i(\mathbf{B}, \Sigma) \frac{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{\mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B})\Sigma^{-1})}{d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right], \quad (\text{A.12})$$

$$\begin{aligned} \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)} &= \frac{1}{2s_n(\mathbf{B}, \Sigma)} \left[-\frac{(\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B}))}{d_i(\mathbf{B}, \Sigma)} + d_i(\mathbf{B}, \Sigma) \frac{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B})}{d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right] \\ &\quad \times (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q. \end{aligned} \quad (\text{A.13})$$

Demostración del Lema A.5.3. Para obtener la derivada de $d_i^*(\mathbf{B}, \Sigma)$ con respecto a $\text{vec}^t(\mathbf{B}^t)$ se emplea la regla del cociente,

$$\begin{aligned} \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} &= \frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{s_n(\mathbf{B}, \Sigma)} - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \\ &= \left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) \frac{1}{s_n(\mathbf{B}, \Sigma)} - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \quad \text{por (A.8)} \\ &= -\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)s_n(\mathbf{B}, \Sigma)} - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \\ &\quad \times \left(\frac{s_n(\mathbf{B}, \Sigma) \sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(-\frac{\mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B})\Sigma^{-1})}{d_j(\mathbf{B}, \Sigma)} \right)}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right) \quad \text{por (A.10)} \\ &= -\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B})\Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)s_n(\mathbf{B}, \Sigma)} + \frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \left(\frac{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(\frac{\mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B})\Sigma^{-1})}{d_j(\mathbf{B}, \Sigma)} \right)}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right), \end{aligned} \quad (\text{A.14})$$

así, la primera afirmación queda probada. Para la derivada de $d_i^*(\mathbf{B}, \Sigma)$ con respecto a

$vech^t(\Sigma)$ se usará un procedimiento similar al anterior,

$$\begin{aligned}
 \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} &= \frac{\partial d_i(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} \frac{1}{s_n(\mathbf{B}, \Sigma)} - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} \\
 &= \left(-\frac{1}{2d_i(\mathbf{B}, \Sigma)} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q \right) \frac{1}{s_n(\mathbf{B}, \Sigma)} \\
 &\quad - \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} \quad \text{por (A.9)} \\
 & \quad \quad \quad (A.15) \\
 &= \left(-\frac{1}{2d_i(\mathbf{B}, \Sigma)} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q \right) \frac{1}{s_n(\mathbf{B}, \Sigma)} \\
 &\quad + \frac{d_i(\mathbf{B}, \Sigma)}{s_n^2(\mathbf{B}, \Sigma)} \left(\frac{s_n(\mathbf{B}, \Sigma) \sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{(\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q}{2d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right),
 \end{aligned}$$

donde la última igualdad se obtiene por (A.11). Luego, simplificando y extrayendo factor común, la igualdad anterior se reduce a,

$$\begin{aligned}
 \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vech^t(\Sigma)} &= -\frac{(\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q}{2s_n(\mathbf{B}, \Sigma) d_i(\mathbf{B}, \Sigma)} + \frac{d_i(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \\
 &\quad \times \left(\frac{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \left(\frac{(\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q}{2d_j(\mathbf{B}, \Sigma)} \right)}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right) \\
 &= \frac{1}{2s_n(\mathbf{B}, \Sigma)} \left[-\frac{(\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q}{d_i(\mathbf{B}, \Sigma)} + d_i(\mathbf{B}, \Sigma) \left(\frac{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) \frac{(\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B})) (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q}{d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1 \left(\frac{d_j(\mathbf{B}, \Sigma)}{s_n(\mathbf{B}, \Sigma)} \right) d_j(\mathbf{B}, \Sigma)} \right) \right] \\
 &\quad \times (\Sigma^{-1} \otimes \Sigma^{-1}) \mathbf{D}_q,
 \end{aligned}$$

lo cual completa la prueba. \square

A continuación se presenta un problema de minimización equivalente al que define al τ -estimador. Para ello se define la función $g(\mathbf{B}, \Sigma) = |\Sigma| \tau^{2q} (H_{P_n, \mathbf{B}, \Sigma})$.

Lema A.5.4. *El problema presentado en (1.15) es equivalente a minimizar $g(\mathbf{B}, \Sigma)$ con $(\mathbf{B}, \Sigma) \in \Theta$ que satisface la restricción de (1.15).*

Demostración del Lema A.5.4. Primero, se verificará que para cualquier valor real λ resulta,

$$g(\mathbf{B}, \lambda \Sigma) = g(\mathbf{B}, \Sigma). \quad (A.16)$$

En efecto,

$$\begin{aligned}
 g(\mathbf{B}, \lambda \boldsymbol{\Sigma}) &= |\lambda \boldsymbol{\Sigma}| \tau^{2q}(H_{P_n, \mathbf{B}, \lambda \boldsymbol{\Sigma}}) && \text{definición de } g \\
 &= \lambda^q |\boldsymbol{\Sigma}| \frac{1}{\lambda^q} \tau^{2q}(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}}) \\
 &= |\boldsymbol{\Sigma}| \tau^{2q}(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}}) \\
 &= g(\mathbf{B}, \boldsymbol{\Sigma}). && \text{definición de } g
 \end{aligned}$$

Se quiere probar que el τ -estimador del problema (1.15) también minimiza $g(\mathbf{B}, \boldsymbol{\Sigma})$. Ya que $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ verifica el problema (1.15) resulta

$$\begin{aligned}
 g(\mathbf{B}_n, \boldsymbol{\Sigma}_n) &= |\boldsymbol{\Sigma}_n| \tau^{2q}(H_{P_n, \mathbf{B}_n, \boldsymbol{\Sigma}_n}) \\
 &= |\boldsymbol{\Sigma}_n| \kappa_2^q.
 \end{aligned} \tag{A.17}$$

Se toma ahora cualquier $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$, existe un $c \in \mathbb{R}$ tal que

$$\tau^2(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}}) = c \kappa_2.$$

Dado que,

$$\kappa_2 = \frac{1}{c} \tau^2(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}}) = \tau^2(H_{P_n, \mathbf{B}, c\boldsymbol{\Sigma}}), \tag{A.18}$$

$(\mathbf{B}, c\boldsymbol{\Sigma})$ satisface la restricción impuesta en (1.15) entonces,

$$\begin{aligned}
 g(\mathbf{B}_n, \boldsymbol{\Sigma}_n) &= |\boldsymbol{\Sigma}_n| \kappa_2^q && \text{por (A.17)} \\
 &\leq |c\boldsymbol{\Sigma}| \tau^{2q}(H_{P_n, \mathbf{B}, c\boldsymbol{\Sigma}}) && \text{por (A.18)} \\
 &= g(\mathbf{B}, c\boldsymbol{\Sigma}) \\
 &= g(\mathbf{B}, \boldsymbol{\Sigma}), && \text{por (A.16)}
 \end{aligned}$$

en consecuencia, $(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$ minimiza la función g para todo $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$. Sea $(\mathbf{B}_{min}, \boldsymbol{\Sigma}_{min})$ un valor que minimiza la función $g(\cdot)$ y satisface la restricción presente en (1.15), entonces para cualquier $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$ que también satisfaga dicha restricción se tiene,

$$\begin{aligned}
 g(\mathbf{B}_{min}, \boldsymbol{\Sigma}_{min}) &= |\boldsymbol{\Sigma}_{min}| \tau^{2q}(H_{P_n, \mathbf{B}_{min}, \boldsymbol{\Sigma}_{min}}) \\
 &= |\boldsymbol{\Sigma}_{min}| \kappa_2^q \\
 &\leq g(\mathbf{B}, \boldsymbol{\Sigma}) \\
 &= |\boldsymbol{\Sigma}| \tau^{2q}(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}}) \\
 &= |\boldsymbol{\Sigma}| \kappa_2^q,
 \end{aligned}$$

de manera que, $|\boldsymbol{\Sigma}_{min}| \leq |\boldsymbol{\Sigma}|$. En consecuencia, $(\mathbf{B}_{min}, \boldsymbol{\Sigma}_{min})$ minimiza el problema (1.15). \square

Demostración del Teorema 2. En esta demostración cuando no exista ambigüedad, se denotará $d_i := d_i(\mathbf{B}, \boldsymbol{\Sigma})$, $d_i^* := d_i^*(\mathbf{B}, \boldsymbol{\Sigma})$ y $s := s_n(\mathbf{B}, \boldsymbol{\Sigma})$. Por otro lado, se usará el resultado obtenido en el Lema A.5.4 y por tanto se minimiza la función $g(\mathbf{B}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}| \tau^{2q}(H_{P_n, \mathbf{B}, \boldsymbol{\Sigma}})$

que es equivalente a minimizar $\log g(\mathbf{B}, \Sigma)$,

$$\begin{aligned} \log(g(\mathbf{B}, \Sigma)) &= \log \left(|\Sigma| \tau^{2q} (d_1(\mathbf{B}, \Sigma), \dots, d_n(\mathbf{B}, \Sigma)) \right) \\ &= \log |\Sigma| + q \log \left(s^2 \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma)) \right) \quad \text{de (1.13) y (1.21)} \\ &= \log |\Sigma| + q \left[\log s^2 + \log \left(\frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma)) \right) \right]. \end{aligned}$$

A partir de esta ecuación se tiene que,

$$\frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} = q \left[\frac{\partial \log s^2}{\partial \text{vect}(\mathbf{B}^t)} + \frac{\partial \log \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} \right]. \quad (\text{A.19})$$

Por un lado,

$$\frac{\partial \log s^2}{\partial \text{vect}(\mathbf{B}^t)} = \frac{2 \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma)) \left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right)}{\sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma)) d_i(\mathbf{B}, \Sigma)}. \quad \text{por (A.10)} \quad (\text{A.20})$$

Por otro lado,

$$\frac{\partial \log \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} = \frac{1}{\sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \Sigma))} \left[\sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial (d_i^*(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} \right],$$

luego por el Lema A.5.3, se sigue que,

$$\begin{aligned} \frac{\partial \log \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} &= \frac{\sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}, \Sigma))}{s \sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \Sigma))} \left[\left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) \right. \\ &\quad \left. + d_i(\mathbf{B}, \Sigma) \left(\frac{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \Sigma)) \frac{\mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B}) \Sigma^{-1})}{d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \Sigma)) d_j(\mathbf{B}, \Sigma)} \right) \right], \quad (\text{A.21}) \end{aligned}$$

sustituyendo (A.20) y (A.21) en (A.19) se obtiene,

$$\begin{aligned} \frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vect}(\mathbf{B}^t)} &= q \left[\frac{2 \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma)) \left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right)}{\sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma)) d_i(\mathbf{B}, \Sigma)} + \frac{\sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}, \Sigma))}{s \sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \Sigma))} \right. \\ &\quad \left. \times \left[\left(-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1})}{d_i(\mathbf{B}, \Sigma)} \right) + d_i(\mathbf{B}, \Sigma) \left(\frac{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \Sigma)) \frac{\mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B}) \Sigma^{-1})}{d_j(\mathbf{B}, \Sigma)}}{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \Sigma)) d_j(\mathbf{B}, \Sigma)} \right) \right] \right]. \end{aligned}$$

Igualando a cero y multiplicando por $\frac{1}{q} \sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \Sigma)) \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \Sigma)) d_i(\mathbf{B}, \Sigma)$, la ecuación se reduce a,

$$\begin{aligned} \mathbf{0} &= -2 \sum_{i=1}^n \left[\frac{\psi_1(d_i^*)}{d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \right] \sum_{j=1}^n \rho_2(d_j^*) - \sum_{i=1}^n \left[\frac{\psi_2(d_i^*)}{s d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \right] \sum_{j=1}^n \psi_1(d_j^*) d_j \\ &\quad + \sum_{i=1}^n \psi_2(d_i^*) d_i^* \sum_{j=1}^n \left[\frac{\psi_1(d_j^*)}{d_j} \mathbf{x}_j^t \otimes (\mathbf{u}_j^t(\mathbf{B}) \Sigma^{-1}) \right] \\ &= - \sum_{i=1}^n \left[\frac{\psi_1(d_i^*)}{d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \right] \left\{ 2 \sum_{j=1}^n \rho_2(d_j^*) - \sum_{j=1}^n \psi_2(d_j^*) d_j^* \right\} \\ &\quad - \sum_{i=1}^n \left[\frac{\psi_2(d_i^*)}{d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \right] \sum_{j=1}^n \psi_1(d_j^*) d_j^*. \end{aligned}$$

Por las definiciones de C_n y D_n dadas en (1.19) y (1.20), respectivamente, la ecuación resulta,

$$\begin{aligned} \mathbf{0} &= \sum_{i=1}^n \left(C_n \frac{\psi_1(d_i^*)}{d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) + D_n \frac{\psi_2(d_i^*)}{d_i} \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \right) \\ &= \sum_{i=1}^n \left[C_n \frac{\psi_1(d_i^*)}{d_i} + D_n \frac{\psi_2(d_i^*)}{d_i} \right] \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \\ &= \frac{1}{s} \sum_{i=1}^n w_n^*(d_i^*) \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \Sigma^{-1}) \quad \text{por (1.17)} \\ &= \frac{1}{s} \sum_{i=1}^n w_n^*(d_i^*) \text{vec}^t(\Sigma^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t) \\ &= \frac{1}{s} \sum_{i=1}^n w_n^*(d_i^*) \text{vec}^t(\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t) (\mathbf{I}_p \otimes \Sigma^{-1}). \quad (\text{A.22}) \end{aligned}$$

Postmultiplicando a ambos lados de la igualdad por $(\mathbf{I}_p \otimes \Sigma)$, se obtiene

$$\mathbf{0} = \sum_{i=1}^n w_n^*(d_i^*) \text{vec}^t(\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t),$$

que es equivalente a,

$$\mathbf{0} = \sum_{i=1}^n w_n^*(d_i^*) \mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t,$$

consiguiendo la primera ecuación del sistema (1.22). Para terminar, resta calcular la derivada de $\log g(\mathbf{B}, \Sigma)$ con respecto a $\text{vech}^t(\Sigma)$, es decir,

$$\frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vech}^t(\Sigma)} = \frac{\partial \log |\Sigma|}{\partial \text{vech}^t(\Sigma)} + q \frac{\partial \log \left(s^2 \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \Sigma)) \right)}{\partial \text{vech}^t(\Sigma)}. \quad (\text{A.23})$$

Por un lado,

$$\frac{\partial \log |\Sigma|}{\partial \text{vech}^t(\Sigma)} = \text{vec}^t(\Sigma^{-1}) \mathbf{D}_q. \quad (\text{A.24})$$

Por otro lado,

$$\begin{aligned}
 \frac{\partial \log \left(s^2 \frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right)}{\partial \text{vech}^t(\boldsymbol{\Sigma})} &= \frac{\partial \log s^2}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + \frac{\partial \log \left(\frac{1}{n} \sum_{i=1}^n \rho_2(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right)}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\
 &= \frac{2}{s} \frac{\partial s}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + \frac{\sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{\partial (d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{\partial \text{vech}^t(\boldsymbol{\Sigma})}}{\sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \boldsymbol{\Sigma}))} \\
 &= \frac{2}{s} \left[\frac{s \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \left(-\frac{(\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \right)}{\sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i(\mathbf{B}, \boldsymbol{\Sigma})} \mathbf{D}_q \right. \\
 &\quad + \frac{\sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{\sum_{k=1}^n \rho_2(d_k^*(\mathbf{B}, \boldsymbol{\Sigma}))} \left[\frac{1}{2s} \left(-\frac{(\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B}))}{d_i(\mathbf{B}, \boldsymbol{\Sigma})} + d_i(\mathbf{B}, \boldsymbol{\Sigma}) \right) \right. \\
 &\quad \left. \left. \times \frac{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{(\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B}))}{d_j(\mathbf{B}, \boldsymbol{\Sigma})}}{\sum_{j=1}^n \psi_1(d_j^*(\mathbf{B}, \boldsymbol{\Sigma})) d_j(\mathbf{B}, \boldsymbol{\Sigma})} \right) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{D}_q \right], \tag{A.25}
 \end{aligned}$$

la última igualdad se debe a (A.11) y (A.13). Sustituyendo (A.25) y (A.24) en (A.23) e igualando a cero, resulta,

$$\begin{aligned}
 \mathbf{0} = \text{vec}^t(\boldsymbol{\Sigma}^{-1}) \mathbf{D}_q + q &\left\{ \frac{2}{s} \left[\frac{s \sum_{i=1}^n \psi_1(d_i^*) \left(-\frac{1}{2d_i} (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1}) \otimes (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1}) \right)}{\sum_{i=1}^n \psi_1(d_i^*) d_i} \right] \right. \\
 &\left. + \frac{\sum_{i=1}^n \psi_2(d_i^*)}{\sum_{k=1}^n \rho_2(d_k^*)} \left[\frac{1}{2s} \left(-\frac{1}{d_i} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) + d_i \frac{\sum_{j=1}^n \psi_1(d_j^*) \frac{1}{d_j} (\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B}))}{\sum_{j=1}^n \psi_1(d_j^*) d_j} \right) (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \right] \right\} \mathbf{D}_q,
 \end{aligned}$$

multiplicando por derecha por $s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) \mathbf{L}_q(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \mathbf{D}_q$ y usando la propiedad

$\mathbf{D}_q \mathbf{L}_q(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \mathbf{D}_q = (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \mathbf{D}_q$ (Ver (Magnus and Neudecker, 1980)) se tiene,

$$\begin{aligned}
 \mathbf{0} &= \text{vec}^t(\boldsymbol{\Sigma}^{-1}) s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \mathbf{D}_q + q \left[s \sum_{i=1}^n \psi_1(d_i^*) \left(-\frac{1}{d_i} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \right) \right. \\
 &\quad \times \sum_{k=1}^n \rho_2(d_k^*) \mathbf{D}_q + \sum_{i=1}^n \psi_2(d_i^*) \left(-\frac{1}{2d_i} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \right) \sum_{j=1}^n \psi_1(d_j^*) d_j \\
 &\quad \left. + \sum_{i=1}^n \psi_2(d_i^*) d_i \left(\sum_{j=1}^n \psi_1(d_j^*) \left(\frac{1}{2d_j} (\mathbf{u}_j^t(\mathbf{B}) \otimes \mathbf{u}_j^t(\mathbf{B})) \right) \right) \right] \mathbf{D}_q \\
 &= \text{vec}^t(\boldsymbol{\Sigma}^{-1}) s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \mathbf{D}_q + q \left[\sum_{i=1}^n \psi_1(d_i^*) \frac{1}{2d_i} (-\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) s \right. \\
 &\quad \left. \times \left\{ 2 \sum_{k=1}^n \rho_2(d_k^*) - \sum_{i=1}^n \psi_2(d_i^*) \frac{d_i}{s} \right\} - \sum_{i=1}^n \psi_2(d_i^*) \left(\frac{1}{2d_i} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \right) \sum_{j=1}^n \psi_1(d_j^*) d_j \right] \mathbf{D}_q.
 \end{aligned}$$

Por definición de C_n y D_n , y usando que $\text{vec}^t(\boldsymbol{\Sigma}) = \text{vec}^t(\boldsymbol{\Sigma}^{-1})(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma})$, la igualdad anterior se simplifica a,

$$\begin{aligned}
 \mathbf{0} &= \text{vec}^t(\boldsymbol{\Sigma}) s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) \mathbf{D}_q + q \left[\sum_{i=1}^n \psi_1(d_i^*) \frac{1}{2d_i} (-\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) sn C_n \right. \\
 &\quad \left. - \sum_{i=1}^n \psi_2(d_i^*) \left(\frac{1}{2d_i} (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \right) sn D_n \right] \mathbf{D}_q \\
 &= \text{vec}^t(\boldsymbol{\Sigma}) s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) \mathbf{D}_q - \frac{qn}{2} \sum_{i=1}^n \left[\frac{\psi_1(d_i^*)}{d_i^*} C_n + \frac{\psi_2(d_i^*)}{d_i^*} D_n \right] (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q \\
 &= \text{vec}^t(\boldsymbol{\Sigma}) s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) \mathbf{D}_q - \frac{qn}{2} \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q, \quad \text{por (1.17)}.
 \end{aligned}$$

Multiplicando por 2 la ecuación anterior, sumando y restando el término $s^2 \sum_{i=1}^n \psi_2(d_i^*) d_i^*$

$\times \sum_{j=1}^n \psi_1(d_j^*) d_j^* \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q$, se tiene que

$$\begin{aligned}
 \mathbf{0} &= \text{vec}^t(\boldsymbol{\Sigma}) 2s \sum_{j=1}^n \psi_1(d_j^*) d_j \sum_{k=1}^n \rho_2(d_k^*) \mathbf{D}_q - qn \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q \\
 &\quad + s^2 \sum_{i=1}^n \sum_{j=1}^n \psi_2(d_i^*) d_i^* \sum_{j=1}^n \psi_1(d_j^*) d_j^* \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q - s^2 \sum_{i=1}^n \psi_2(d_i^*) d_i^* \sum_{j=1}^n \psi_1(d_j^*) d_j^* \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q \\
 &= s^2 \sum_{j=1}^n \psi_1(d_j^*) d_j^* \left[2 \sum_{k=1}^n \rho_2(d_k^*) - \sum_{i=1}^n \psi_2(d_i^*) d_i^* \right] \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q - qn \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q \\
 &\quad + s^2 \sum_{i=1}^n \psi_2(d_i^*) d_i^* \sum_{j=1}^n \psi_1(d_j^*) d_j^* \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q \\
 &= s^2 \sum_{j=1}^n \psi_1(d_j^*) d_j^* [nC_n] \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q - qn \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q \\
 &\quad + s^2 \sum_{i=1}^n \psi_2(d_i^*) d_i^* nD_n \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q, \quad \text{por (1.19) y (1.20)} \\
 &= s^2 n \sum_{i=1}^n [\psi_1(d_i^*) d_i^* C_n + \psi_2(d_i^*) d_i^* D_n] \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q - qn \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q \\
 &= s^2 n \sum_{i=1}^n [\psi_n^*(d_i^*) d_i^*] \text{vec}^t(\boldsymbol{\Sigma}) \mathbf{D}_q - qn \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})) \mathbf{D}_q, \quad \text{por (1.18)}
 \end{aligned}$$

así, resulta

$$s^2 \sum_{i=1}^n [\psi_n^*(d_i^*) d_i^*] \text{vec}^t(\boldsymbol{\Sigma}) = q \sum_{i=1}^n w_n^*(d_i^*) (\mathbf{u}_i^t(\mathbf{B}) \otimes \mathbf{u}_i^t(\mathbf{B})),$$

que es equivalente a,

$$s^2 \sum_{i=1}^n [\psi_n^*(d_i^*) d_i^*] \boldsymbol{\Sigma} = q \sum_{i=1}^n w_n^*(d_i^*) \mathbf{u}_i \mathbf{u}_i^t(\mathbf{B}),$$

obteniendo la segunda ecuación del sistema (1.22). Por último, cabe recordar que se desea que las soluciones sigan satisfaciendo la condición impuesta en el problema (1.15), en consecuencia, la tercera ecuación vendrá dada por la ecuación de la condición. \square

B Apéndice del Capítulo 2

B.1 Demostración del Lema 2.1.1

Demostración. En esta demostración cuando no exista ambigüedad, se denotará $s_n := s_n(\mathbf{B}, \Sigma)$, $C_n := C_n(\mathbf{B}, \Sigma)$, $D_n := D_n(\mathbf{B}, \Sigma)$, $u_i = \mathbf{u}_i(\mathbf{B})$ y $d_i^* := d_i^*(\mathbf{B}, \Sigma)$, que fueron definidas en (1.11), (1.19), (1.20) y (1.21), respectivamente. Por definición de $\mathbf{A}_i(\cdot)$ y $\mathbf{M}_{(C_n, D_n)}(\cdot)$ se tiene que,

$$\begin{aligned} \mathbf{M}_{(C_n, D_n)}(\mathbf{B}, \Sigma) \mathbf{A}_i(\mathbf{B}, \Sigma, s_n) &= \begin{pmatrix} C_n \mathbf{I}_{pq} & \mathbf{0} & D_n \mathbf{I}_{pq} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_n \mathbf{I}_{q(q+1)/2} & \mathbf{0} & D_n \mathbf{I}_{q(q+1)/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) \\ \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, s_n)) \\ \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) \\ \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, s_n)) \\ \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n) \end{pmatrix} \\ &= \begin{pmatrix} C_n \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) + D_n \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) \\ C_n \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, s_n)) + D_n \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, s_n)) \\ \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n) \end{pmatrix}, \end{aligned}$$

utilizando las definiciones de $\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)$, $\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n)$, y $\Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n)$ dadas en (2.2) para $k = 1, 2$ con $t = s_n$, y las definiciones de $w_n^*(\cdot)$, $\psi_n^*(\cdot)$, y $d_i^*(\cdot)$ dadas en (1.17), (1.18) y (1.21), respectivamente, se tiene que

$$\begin{aligned} C_n \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) + D_n \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \Sigma, s_n)) &= C_n \text{vec} \left(\frac{\psi_1(d_i^*)}{d_i^*} \mathbf{u}_i \mathbf{x}_i^t \right) + D_n \text{vec} \left(\frac{\psi_2(d_i^*)}{d_i^*} \mathbf{u}_i \mathbf{x}_i^t \right) \\ &= \left[C_n \frac{\psi_1(d_i^*)}{d_i^*} + D_n \frac{\psi_2(d_i^*)}{d_i^*} \right] \text{vec}(\mathbf{u}_i \mathbf{x}_i^t) \\ &= w_n^*(d_i^*) \text{vec}(\mathbf{u}_i \mathbf{x}_i^t), \\ C_n \text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}, \Sigma, s_n)) + D_n \text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}, \Sigma, s_n)) &= q C_n \text{vech} \left(\frac{\psi_1(d_i^*)}{d_i^*} \mathbf{u}_i \mathbf{u}_i^t \right) - s_n^2 C_n \text{vech}(\psi_1(d_i^*) d_i^* \Sigma) \\ &\quad + q D_n \text{vech} \left(\frac{\psi_2(d_i^*)}{d_i^*} \mathbf{u}_i \mathbf{u}_i^t \right) - s_n^2 D_n \text{vech}(\psi_2(d_i^*) d_i^* \Sigma) \\ &= q \left[C_n \frac{\psi_1(d_i^*)}{d_i^*} + D_n \frac{\psi_2(d_i^*)}{d_i^*} \right] \text{vech}(\mathbf{u}_i \mathbf{u}_i^t) \\ &\quad - s_n^2 [C_n \psi_1(d_i^*) + D_n \psi_2(d_i^*)] d_i^* \text{vech}(\Sigma) \\ &= q w_n^*(d_i^*) \text{vech}(\mathbf{u}_i \mathbf{u}_i^t) - s_n^2 \psi_n^*(d_i^*) d_i^* \text{vech}(\Sigma), \end{aligned}$$

y,

$$\Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n) = s_n^2 \rho_2(d_i^*) - \kappa_2.$$

Por lo tanto, $\sum_{i=1}^n \mathbf{M}_{(C_n, D_n)}(\mathbf{B}, \boldsymbol{\Sigma}) A_i(\mathbf{B}, \boldsymbol{\Sigma}, s_n) = \mathbf{0}$ es

$$\sum_{i=1}^n \begin{pmatrix} w_n^*(d_i^*) \text{vec}(\mathbf{u}_i \mathbf{x}_i^t) \\ qw_n^*(d_i^*) \text{vech}(\mathbf{u}_i \mathbf{u}_i^t) - s_n^2 \psi_n^*(d_i^*) d_i^* \text{vech}(\boldsymbol{\Sigma}) \\ s_n^2 \rho_2(d_i^*) - \kappa_2 \end{pmatrix} = \mathbf{0},$$

que es equivalente al sistema de ecuaciones (1.22). \square

B.2 Demostración del Teorema 5

Antes de la demostración del Teorema se presentan resultados auxiliares junto con sus respectivas demostraciones.

Lema B.2.1. *Sea (\mathbf{x}, \mathbf{y}) un vector que cumple con el MLM (1). Sea ρ una ρ -función que satisface A1)-A6). Si P satisface (\mathbf{S}_ϵ) para algún $0 < \epsilon < 1 - \frac{\kappa}{a}$, entonces el M -funcional de escala para el MLM, $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}})$ que está basado en ρ , es continuo en $H_{P, \mathbf{B}_1, \boldsymbol{\Sigma}_1}$ para todo $(\mathbf{B}_1, \boldsymbol{\Sigma}_1) \in \Theta$.*

Demostración del Lema B.2.1. Se usa el mismo enfoque del Teorema 3.5 de (Fasano, 2009). Dado $k_0^* = s(H_{P, \mathbf{B}_1, \boldsymbol{\Sigma}_1})$ y $0 < \delta < k_0^*$. Se quiere probar que si $(\mathbf{B}, \boldsymbol{\Sigma})$ está en una vecindad de $(\mathbf{B}_1, \boldsymbol{\Sigma}_1)$ entonces

$$k_0^* - \delta < s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}}) < k_0^* + \delta$$

para ello se definen

$$g_1(\mathbf{x}, \mathbf{y}, \gamma) = \sup \left\{ \rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* + \delta} \right) : (\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta \text{ y } \|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_1, \boldsymbol{\Sigma}_1)\| \leq \gamma \right\}, \quad (\text{B.1})$$

$$g_2(\mathbf{x}, \mathbf{y}, \gamma) = \inf \left\{ \rho \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* - \delta} \right) : (\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta \text{ y } \|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_1, \boldsymbol{\Sigma}_1)\| \leq \gamma \right\}. \quad (\text{B.2})$$

Como ρ es continua y acotada los límites de $g_1(\mathbf{x}, \mathbf{y}, \gamma)$ y $g_2(\mathbf{x}, \mathbf{y}, \gamma)$ son

$$\begin{aligned} \lim_{\gamma \rightarrow 0} g_1(\mathbf{x}, \mathbf{y}, \gamma) &= \rho \left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* + \delta} \right), \\ \lim_{\gamma \rightarrow 0} g_2(\mathbf{x}, \mathbf{y}, \gamma) &= \rho \left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta} \right). \end{aligned}$$

Como ρ es acotada, por el Teorema de Convergencia Dominada resulta,

$$\lim_{\gamma \rightarrow 0} E_P[g_1(\mathbf{x}, \mathbf{y}, \gamma)] = E_P \left[\rho \left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* + \delta} \right) \right], \quad (\text{B.3})$$

$$\lim_{\gamma \rightarrow 0} E_P[g_2(\mathbf{x}, \mathbf{y}, \gamma)] = E_P \left[\rho \left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta} \right) \right]. \quad (\text{B.4})$$

Por otro lado,

$$\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta} > \frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^*} > \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{k_0^* + \delta},$$

luego

$$\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta}\right) \geq \rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^*}\right) \geq \rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* + \delta}\right),$$

y, por tanto,

$$E_P\left(\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta}\right)\right) \geq \kappa \geq E_P\left(\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* + \delta}\right)\right), \quad (\text{B.5})$$

donde $\kappa = E_P\left(\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^*}\right)\right)$. Por el Lema 1.2.2 que garantiza la existencia y unicidad del M-funcional de escala para el MLM para todo $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$, en particular para $(\mathbf{B}_1, \boldsymbol{\Sigma}_1)$ resulta que las desigualdades involucradas en (B.5) son estrictas, es decir,

$$E_P\left(\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* - \delta}\right)\right) > \kappa > E_P\left(\rho\left(\frac{d(\mathbf{B}_1, \boldsymbol{\Sigma}_1)}{k_0^* + \delta}\right)\right). \quad (\text{B.6})$$

Usando (B.3), (B.4) y (B.6) se observa que existen $\gamma_0 > 0$ y $\eta > 0$ tales que

$$E_P[g_1(\mathbf{x}, \mathbf{y}, \gamma_0)] \leq \kappa - \eta, \quad (\text{B.7})$$

$$E_P[g_2(\mathbf{x}, \mathbf{y}, \gamma_0)] \geq \kappa + \eta. \quad (\text{B.8})$$

Por otro lado, por definición de las funciones g_1 y g_2 se tiene que para todo $(\mathbf{B}, \boldsymbol{\Sigma})$ que satisfaga $\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| \leq \gamma_0$ resulta

$$\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* + \delta}\right) \leq g_1(\mathbf{x}, \mathbf{y}, \gamma_0),$$

$$\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* - \delta}\right) \geq g_2(\mathbf{x}, \mathbf{y}, \gamma_0),$$

para todo \mathbf{x}, \mathbf{y} . Luego usando (B.7) y (B.8) resulta

$$E_P\left(\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* + \delta}\right)\right) \leq E_P[g_1(\mathbf{x}, \mathbf{y}, \gamma_0)] \leq \kappa - \eta, \quad (\text{B.9})$$

$$E_P\left(\rho\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{k_0^* - \delta}\right)\right) \geq E_P[g_2(\mathbf{x}, \mathbf{y}, \gamma_0)] \geq \kappa + \eta. \quad (\text{B.10})$$

Por la existencia y unicidad de $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}})$ para $(\mathbf{B}, \boldsymbol{\Sigma}) \in \Theta$ garantizada por el Lema 1.2.2 se tiene que para todo $(\mathbf{B}, \boldsymbol{\Sigma})$ tal que $\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_1, \boldsymbol{\Sigma}_1)\| < \gamma_0$ se tiene que $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}}) < k_0^* + \delta$ por (B.9) y $s(H_{P, \mathbf{B}, \boldsymbol{\Sigma}}) > k_0^* - \delta$ por (B.10), lo cual prueba el resultado. \square

Lema B.2.2 (Lema 1 de (Yohai, 1974)). *Sea U_1, \dots, U_n, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas. Sea $(f_k), k \in C$, donde C es un compacto, una familia de funciones reales medible Borel tal que*

1. $|f_k| \leq f$ donde $E(f(U_1)) < \infty$.
2. $\lim_{t \rightarrow k} f_t(U_1) = f_k(U_1)$ c.s. para todo k en C .
3. $|E(f_k(U_1))| \leq A$ para todo k en C .

Entonces,

$$\limsup_{n \rightarrow \infty} \sup_{k \in C} \left| \sum_{j=1}^n \frac{f_k(U_j)}{n} \right| \leq A \text{ c.s.}$$

El siguiente lema se sigue del Lema B.2.2, sustituyendo para nuestro caso particular $k = (\mathbf{B}, \Sigma)$, $U_i = \mathbf{u}_i$, $f_k(\mathbf{u}) = f(\mathbf{u}, (\mathbf{B}, \Sigma))$, $f = f^*$, y $A = 0$.

Lema B.2.3. Sean $g : \mathbb{R}^{p+q} \times C \times I \rightarrow \mathbb{R}$ una función boreliana con C un conjunto compacto en $\mathbb{R}^{p+q} \times S_+^q$ e $I = [h_1, h_2] \subset \mathbb{R}_{>0}$, y $g^* : \mathbb{R}^{p+q} \rightarrow \mathbb{R}$ otra función boreliana. Sean $(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_n, \mathbf{y}_n), \dots$ vectores aleatorios independientes e idénticamente distribuidos con valores en \mathbb{R}^{p+q} y distribución P . Se asume que

1. $\sup_{(\mathbf{B}, \Sigma) \in C, s \in I} |g((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{B}, \Sigma, s))| \leq g^*(\mathbf{x}_1, \mathbf{y}_1)$ con $E_P(g^*(\mathbf{x}_1, \mathbf{y}_1)) < \infty$.
2. $\lim_{(\mathbf{B}, \Sigma, s) \rightarrow (\tilde{\mathbf{B}}, \tilde{\Sigma}, \tilde{s})} g((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{B}, \Sigma, \hat{s})) = g((\mathbf{x}_1, \mathbf{y}_1), (\tilde{\mathbf{B}}, \tilde{\Sigma}, \tilde{s}))$ c.s. para todo $(\tilde{\mathbf{B}}, \tilde{\Sigma}, \tilde{s}) \in C \times I$.
3. $E_P(g((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{B}, \Sigma, s))) = 0$ para todo $(\mathbf{B}, \Sigma, s) \in C \times I$.

Entonces

$$\lim_{n \rightarrow \infty} \sup_{(\mathbf{B}, \Sigma) \in C, s \in I} \frac{1}{n} \left| \sum_{i=1}^n g((\mathbf{x}_i, \mathbf{y}_i), (\mathbf{B}, \Sigma, s)) \right| = 0 \text{ c.s.}$$

Este lema se sigue del Lema B.2.2 sustituyendo para este caso en particular $k = (\mathbf{B}, \Sigma, s)$, $U_i = (\mathbf{x}_i, \mathbf{y}_i)$, $f_k(U_1) = g((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{B}, \Sigma, s))$, $f = g^*$ y $A = 0$.

Lema B.2.4. Sean $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^p \times \mathbb{R}^q$, $i = 1, \dots, n$, una muestra aleatoria que satisface el MLM (1). Sea ρ una función que satisface A1)-A6), C un conjunto compacto en $\mathbb{R}^{p \times q} \times S_+^q$ y $[h_1, h_2]$ un intervalo cerrado con $h_1 > 0$. Entonces

$$\sup_{(\mathbf{B}, \Sigma) \in C, s \in [h_1, h_2]} \left| \frac{1}{n} \sum_{i=1}^n \left(\rho \left(\frac{d_i(\mathbf{B}, \Sigma)}{s} \right) \right) - E_P \left(\rho \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s} \right) \right) \right| \xrightarrow{\text{c.s.}} 0.$$

Demostración del Lema B.2.4. Se desea utilizar el Lema B.2.3 con g definida como

$$g((\mathbf{x}, \mathbf{x}), (\mathbf{B}, \Sigma), s) = \rho \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s} \right) - E_P \left(\rho \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s} \right) \right) \quad (\text{B.11})$$

Para ello se deben probar las condiciones 1.-3. de dicho lema: por A5) la condición 1. se cumple mientras que la 2. se verifica usando A4) y $s \in [h_1, h_2]$. Finalmente la condición 3. es cierta por la definición de f dada en (B.11). En consecuencia, por el Lema B.2.3, se obtiene la convergencia deseada. \square

Lema B.2.5 (Lema 4.5 de (Yohai and Zamar, 1986) y Lema 6.3 de (Fasano, 2009)). *Sea $(\mathbf{x}_i, \mathbf{y}_i)_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Sea ρ_1 que satisface A1)-A6), \mathbf{u}_i definidos en (1.2) que satisfacen A8). Sea C un conjunto compacto en $\mathbb{R}^{p \times q} \times S_+^q$. Entonces*

$$\sup_{(\mathbf{B}, \mathbf{\Sigma}) \in C} |s_n(\mathbf{B}, \mathbf{\Sigma}) - s(\mathbf{B}, \mathbf{\Sigma})| \xrightarrow{c.s.} 0, \quad (\text{B.12})$$

donde $s(\mathbf{B}, \mathbf{\Sigma})$ es el M -funcional de escala para la distribución $H_{P, \mathbf{B}, \mathbf{\Sigma}}$ con $\rho = \rho_1$ y $\kappa = \kappa_1 > a$.

Demostración Lema B.2.5. Se quiere probar que para todo $\epsilon > 0$, existe $n_0 > 0$ tal que para todo $n > n_0$ y para todo $(\mathbf{B}, \mathbf{\Sigma}) \in C$, se cumple

$$s(\mathbf{B}, \mathbf{\Sigma}) - \epsilon < s_n(\mathbf{B}, \mathbf{\Sigma}) < s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon.$$

Para demostrar dicha desigualdad, se encontrará un $\delta > 0$ y un $n_0 = n_0(\delta) > 0$ tales que, para todo $n \geq n_0$ se verifique que

$$\inf_{(\mathbf{B}, \mathbf{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) - \epsilon} \right) \geq \kappa_1 + \frac{\delta}{2}, \quad (\text{B.13})$$

$$\sup_{(\mathbf{B}, \mathbf{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon} \right) \leq \kappa_1 - \frac{\delta}{2}. \quad (\text{B.14})$$

Pues con esto, se obtiene que para todo $(\mathbf{B}, \mathbf{\Sigma}) \in C$, se cumple

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) - \epsilon} \right) &\geq \inf_{(\mathbf{B}, \mathbf{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) - \epsilon} \right) > \kappa_1, \\ \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon} \right) &\leq \sup_{(\mathbf{B}, \mathbf{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon} \right) < \kappa_1. \end{aligned}$$

Ya que $\kappa_1 = \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s_n(\mathbf{B}, \mathbf{\Sigma})} \right)$, para todo $(\mathbf{B}, \mathbf{\Sigma}) \in C$ vale que

$$\frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) - \epsilon} \right) > \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s_n(\mathbf{B}, \mathbf{\Sigma})} \right) > \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon} \right). \quad (\text{B.15})$$

De aquí se deduce la desigualdad deseada puesto que si, por ejemplo, valiera $s_n(\mathbf{B}, \mathbf{\Sigma}) \geq s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon$, entonces, por A3) se tendría que $\frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \mathbf{\Sigma})}{s(\mathbf{B}, \mathbf{\Sigma}) + \epsilon} \right) \geq \kappa_1$ lo que contradice (Ecuación (B.15)).

Ahora para encontrar el $\delta > 0$ y el $n_0 > 0$ que verifiquen (B.13) y (B.14) se definen

$$h_1 := \inf_{(\mathbf{B}, \mathbf{\Sigma}) \in C} s(\mathbf{B}, \mathbf{\Sigma}) \quad \text{y} \quad h_2 := \sup_{(\mathbf{B}, \mathbf{\Sigma}) \in C} s(\mathbf{B}, \mathbf{\Sigma}).$$

Como los $\mathbf{u}_i(\mathbf{B}_0)$ satisfacen A8) entonces P cumple la propiedad (S_ϵ) para todo $\epsilon > 0$ y por lo tanto se verifican las hipótesis de los Lemas B.2.1, B.2.4 y 1.2.2. Por Lema B.2.1 se tiene que $s(\mathbf{B}, \Sigma)$ es continua y positiva y ya que C es compacto, se tiene que $h_1 > 0$ y que $h_2 < \infty$. Además, de acuerdo con el Lema B.2.4 se tiene que,

$$\lim_{n \rightarrow \infty} \sup_{(\mathbf{B}, \Sigma) \in C, s \in [h_1/2, 2h_2]} \left| \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s} \right) \right) - E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s} \right) \right) \right| = 0 \text{ c.s.} \quad (\text{B.16})$$

Para $0 < \epsilon \leq h_1/2$, sean

$$g_1(\mathbf{B}, \Sigma) := E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s(\mathbf{B}, \Sigma) + \epsilon} \right) \right),$$

y,

$$g_2(\mathbf{B}, \Sigma) := E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s(\mathbf{B}, \Sigma) - \epsilon} \right) \right).$$

Dado que $s(\mathbf{B}, \Sigma)$ es el M -funcional de escala para la distribución $H_{P, \mathbf{B}, \Sigma}$ con $\rho = \rho_1$, y y de la unicidad del M -funcional de escala para el MLM garantizada en el Lema 1.2.2 $\kappa = \kappa_1$, se tiene que

$$\kappa_1 = E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s(\mathbf{B}, \Sigma)} \right) \right).$$

Como $\epsilon > 0$, a partir de A3) se tiene que, $g_1(\mathbf{B}, \Sigma) < \kappa_1$ y $g_2(\mathbf{B}, \Sigma) > \kappa_1$ para todo $(\mathbf{B}, \Sigma) \in C$. Además, g_1 y g_2 son continuas por A4), entonces

$$\gamma_1 := \sup_{(\mathbf{B}, \Sigma) \in C} g_1(\mathbf{B}, \Sigma) < \kappa_1, \quad \gamma_2 := \inf_{(\mathbf{B}, \Sigma) \in C} g_2(\mathbf{B}, \Sigma) > \kappa_1.$$

Sea $\delta := \min(\kappa_1 - \gamma_1, \gamma_2 - \kappa_1) > 0$. Por (B.16), existe $n_0 = n_0(\delta)$ para $n \geq n_0$ resulta,

$$\sup_{(\mathbf{B}, \Sigma) \in C, s \in [h_1/2, 2h_2]} \left| \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s} \right) \right) - E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s} \right) \right) \right| \leq \frac{\delta}{2}. \quad (\text{B.17})$$

Para demostrar (B.13), se observa que para todo $(\mathbf{B}, \Sigma) \in C$, como $h_1 = \inf_{(\mathbf{B}, \Sigma) \in C} s(\mathbf{B}, \Sigma) \leq s(\mathbf{B}, \Sigma) \leq h_2 = \sup_{(\mathbf{B}, \Sigma) \in C} s(\mathbf{B}, \Sigma) \leq (\mathbf{B}, \Sigma)$ y $0 < \epsilon \leq h_1/2$ se tiene,

$$\frac{h_1}{2} \leq h_1 - \epsilon \leq s(\mathbf{B}, \Sigma) - \epsilon \leq h_2 - \epsilon \leq h_2,$$

es decir, $s(\mathbf{B}, \Sigma) - \epsilon \in [h_1/2, 2h_2]$. De esta forma, para todo $n \geq n_0$, por propiedades de supremo e ínfimo¹ y (B.17) se tiene que,

$$\begin{aligned} & \left| \inf_{(\mathbf{B}, \Sigma) \in C} \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma) - \epsilon} \right) \right) - \inf_{(\mathbf{B}, \Sigma) \in C} E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s(\mathbf{B}, \Sigma) - \epsilon} \right) \right) \right| \\ & \leq \sup_{(\mathbf{B}, \Sigma) \in C} \left| \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma) - \epsilon} \right) \right) - E_P \left(\rho_1 \left(\frac{((\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \Sigma^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x}))^{1/2}}{s(\mathbf{B}, \Sigma) - \epsilon} \right) \right) \right| \leq \frac{\delta}{2}. \end{aligned}$$

¹ Si $f, g : A \rightarrow \mathbb{R}$ son funciones acotadas, entonces $|\sup_A f - \sup_A g| \leq \sup_A |f - g|$, y $|\inf_A f - \inf_A g| \leq \sup_A |f - g|$.

De esto se sigue que,

$$\inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) - \epsilon} \right) - \inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) - \epsilon} \right) \right) \geq -\frac{\delta}{2}$$

o equivalentemente,

$$\begin{aligned} \inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) - \epsilon} \right) &\geq \inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) - \epsilon} \right) \right) - \frac{\delta}{2} \\ &= \inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} g_2(\mathbf{B}, \boldsymbol{\Sigma}) - \frac{\delta}{2} \\ &= \gamma_2 - \frac{\delta}{2}. \end{aligned} \quad \text{por definición de } \gamma_2$$

Dado que $\delta \leq \gamma_2 - \kappa_1$, entonces $\kappa_1 + \delta \leq \gamma_2$, así para todo $n \geq n_0$ se tiene que

$$\inf_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) - \epsilon} \right) \geq \kappa_1 + \delta - \frac{\delta}{2} \geq \kappa_1 + \frac{\delta}{2}. \quad (\text{B.18})$$

Similarmente, para demostrar (B.14), se considera que para todo $(\mathbf{B}, \boldsymbol{\Sigma}) \in C$ como $h_1 = \inf_{(\mathbf{B}, \boldsymbol{\Sigma})} s(\mathbf{B}, \boldsymbol{\Sigma}) \leq s(\mathbf{B}, \boldsymbol{\Sigma}) \leq \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} s(\mathbf{B}, \boldsymbol{\Sigma})$ y $0 < \epsilon \leq h_1/2$ se tiene,

$$h_1 \leq h_1 + \epsilon \leq s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon \leq h_2 + \epsilon \leq 2h_2,$$

así, $s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon \in [h_1/2, 2h_2]$, usando propiedades de supremo e ínfimo y (B.17) se tiene, para todo $n \geq n_0$, que

$$\begin{aligned} &\left| \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) - \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) \right| \\ &\leq \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \left| \frac{1}{n} \sum_{i=1}^n \left(\rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) - E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) \right| \leq \frac{\delta}{2}, \end{aligned}$$

en consecuencia,

$$\sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) - \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) \leq \frac{\delta}{2}$$

o equivalentemente,

$$\begin{aligned} \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) &\leq \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} E_P \left(\rho_1 \left(\frac{(\mathbf{y} - \mathbf{B}^t \mathbf{x})^t \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{B}^t \mathbf{x})^{1/2}}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \right) + \frac{\delta}{2} \\ &= \sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} g_1(\mathbf{B}, \boldsymbol{\Sigma}) + \frac{\delta}{2} \\ &= \gamma_1 + \frac{\delta}{2}. \end{aligned} \quad \text{por definición de } \gamma_1$$

Como $\delta \leq \kappa_1 - \gamma_1$ resulta que $\gamma_1 \leq \kappa_1 - \delta$, así para todo $n \geq n_0$ se tiene que

$$\sup_{(\mathbf{B}, \boldsymbol{\Sigma}) \in C} \frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma}) + \epsilon} \right) \leq \kappa_1 - \delta + \frac{\delta}{2} \leq \kappa_1 - \frac{\delta}{2}, \quad (\text{B.19})$$

con lo cual B.14 queda probada. \square

Lema B.2.6. *Bajo los supuestos del Lema B.2.5, sea (\mathbf{B}_n, Σ_n) sucesión de estimadores que converge a $(\mathbf{B}_0, \Sigma_0^*)$. Entonces,*

$$s_n(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} s(\mathbf{B}_0, \Sigma_0^*).$$

Demostración del Lema B.2.6. Se quiere probar que para todo $\epsilon > 0$ existe un natural n_0 tal que para $n \geq n_0$, con probabilidad 1 vale que,

$$|s_n(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_0, \Sigma_0^*)| < \epsilon. \quad (\text{B.20})$$

Por la desigualdad triangular se tiene que,

$$|s_n(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_0, \Sigma_0^*)| \leq |s_n(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_n, \Sigma_n)| + |s(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_0, \Sigma_0^*)|. \quad (\text{B.21})$$

Como por hipótesis $(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} (\mathbf{B}_0, \Sigma_0^*)$ y como el M-funcional de escala es continuo (Lema B.2.1), entonces se tiene que $s(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} s(\mathbf{B}_0, \Sigma_0^*)$, y por lo tanto, existe un natural n_1 tal que para todo $n \geq n_1$, con probabilidad 1 vale que,

$$|s(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_0, \Sigma_0^*)| < \frac{\epsilon}{2}. \quad (\text{B.22})$$

Sea A un abierto alrededor de $(\mathbf{B}_0, \Sigma_0^*)$. Como $(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} (\mathbf{B}_0, \Sigma_0^*)$ existe $n_2 > 0$ tal que $(\mathbf{B}_n, \Sigma_n) \in A$ para todo $n \geq n_2$ con probabilidad 1. Sea ahora C un compacto que contiene a A . Luego, para todo $n \geq n_2$ se tiene que

$$|s_n(\mathbf{B}_n, \Sigma_n) - s(\mathbf{B}_n, \Sigma_n)| \leq \sup_{(\mathbf{B}, \Sigma) \in C} |s_n(\mathbf{B}, \Sigma) - s(\mathbf{B}, \Sigma)|. \quad (\text{B.23})$$

Por el Lema B.2.5, existe $n_3 > 0$ tal que para todo $n \geq n_3$ vale que

$$\sup_{(\mathbf{B}, \Sigma) \in C} |s_n(\mathbf{B}, \Sigma) - s(\mathbf{B}, \Sigma)| < \frac{\epsilon}{2}. \quad (\text{B.24})$$

Considerando $n_0 = \max\{n_1, n_2, n_3\}$ con (B.22), (B.23) y (B.24) resulta (B.20). \square

Lema B.2.7 (Lema 4.2 (Yohai, 1985)). *Sea $g : \mathbb{R}^k \times \mathbb{R}^h \rightarrow \mathbb{R}$ continua y sea Q una distribución de probabilidad sobre \mathbb{R}^k tal que para algún $\delta > 0$,*

$$E \left(\sup_{\|\lambda - \lambda_0\| < \delta} |g(z, \lambda)| \right) < \infty. \quad (\text{B.25})$$

Sea λ_m una sucesión de estimadores en \mathbb{R}^h tal que $\lim_{m \rightarrow \infty} \lambda_m = \lambda_0$ c.s. Entonces si z_i , $i = 1, \dots, n$ son variables aleatorias independientes e idénticamente distribuidas en \mathbb{R}^k con distribución Q , se tiene

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n g(z_i, \lambda_n) = E_Q(g(z, \lambda_0)) \quad c.s. \quad (\text{B.26})$$

Lema B.2.8. *Bajo los supuestos del Lema B.2.5, sea $f : \mathbb{R}^{p+q} \times \mathbb{R}^{p \times q} \times S_+^q \times \mathbb{R} \rightarrow \mathbb{R}$ una función continua tal que, para algún $\delta > 0$,*

$$E_P \left(\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} |f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))| \right) < \infty. \quad (\text{B.27})$$

Entonces si $\mathbf{z}_i := (\mathbf{x}_i, \mathbf{y}_i)$, con $i \geq 1$ son vectores aleatorios independientes e idénticamente distribuidos en \mathbb{R}^{p+q} que satisfacen el MLM (1) y sea $\{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\}_{n \geq 1}$ con $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$, resulta que,

$$\frac{1}{n} \sum_{i=1}^n f(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} E_P(f(\mathbf{z}, \mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))). \quad (\text{B.28})$$

Demostración del Lema B.2.8. Sea $\lambda_n = (\mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n))$, por hipótesis $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$ y por Lema B.2.6 se tiene $(\mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$, es decir, $\lambda_n \xrightarrow{c.s.} \lambda_0$ con $\lambda_0 = (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$. Entonces por el Lema B.2.7 resulta,

$$\frac{1}{n} \sum_{i=1}^n f(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} E_P(f(\mathbf{z}, \mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))),$$

que era lo requerido. \square

Lema B.2.9. *Bajo las hipótesis del Lema B.2.5, sea $g : \mathbb{R}^{p+q} \times \mathbb{R}^{p \times q} \times S_+^q \times \mathbb{R} \rightarrow \mathbb{R}^{l \times s}$ una función continua tal que, para algún $\delta > 0$,*

$$E_P \left(\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \|g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))\| \right) < \infty. \quad (\text{B.29})$$

Entonces si $\mathbf{z}_i := (\mathbf{x}_i, \mathbf{y}_i)$, con $i \geq 1$ son vectores aleatorios independientes e idénticamente distribuidos en \mathbb{R}^{p+q} que satisfacen el MLM (1) y sea $\{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\}_{n \geq 1}$ una sucesión en Θ tal que $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$, resulta que,

$$\frac{1}{n} \sum_{i=1}^n g(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} E_P(g(\mathbf{z}, \mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))).$$

Demostración del Lema B.2.9. Sea $g_{jk}(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$ el elemento (j, k) de la función $g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$. De (B.29) se deduce que (B.27) es válido para g_{jk} pues $|g_{jk}(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))| \leq \|g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))\|$ para todo $(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))$. El Lema B.2.8 garantiza la convergencia casi segura de cada componente de la función $g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))$, es decir,

$$\frac{1}{n} \sum_{i=1}^n g_{jk}(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} E_P(g_{jk}(\mathbf{z}, \mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))).$$

Y esto ocurre para $j = 1, \dots, l$ y $k = 1, \dots, s$. Tanto la convergencia de sucesiones de matrices, como la esperanza de matrices están basadas en el comportamiento de cada coordenada, en consecuencia queda probado. \square

Demostración del Teorema 5. En la siguiente prueba se usará una idea similar a la presentada en la demostración del Teorema 5.1 de (Yohai and Zamar, 1986).

Dado que por hipótesis $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$ y por Lema B.2.6 $s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$, la siguiente convergencia

$$C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} C(\mathbf{B}_0, \boldsymbol{\Sigma}_0)$$

es consecuencia del Lema B.2.8, tomando como $f(\cdot)$

$$f((\mathbf{x}, \mathbf{y}), \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma})) = 2\rho_2 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) - \psi_2 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) \frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})},$$

pues, en este caso el Lema B.2.8 implica que,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \left[2\rho_2 \left(\frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) - \psi_2 \left(\frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] \\ & \xrightarrow{c.s.} E_P \left(2\rho_2 \left(\frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right) - \psi_2 \left(\frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right) \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right), \end{aligned}$$

luego usando que $\boldsymbol{\Sigma}_0^* = r\boldsymbol{\Sigma}_0$ y

$$\frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} = \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0)} \quad \text{por (1.23) y el Lema 1.2.3}$$

se obtiene lo requerido. Similarmente, la consistencia

$$D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} D(\mathbf{B}_0, \boldsymbol{\Sigma}_0),$$

es consecuencia del Lema B.2.8 tomando como $f(\cdot)$

$$f((x, y), \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma})) = \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) \frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})},$$

pues, en este caso el Lema B.2.8 implica que,

$$\frac{1}{n} \sum_{i=1}^n \left[\psi_1 \left(\frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] \xrightarrow{c.s.} E_P \left(\psi_1 \left(\frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right) \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right).$$

□

B.3 Demostración del Lema 2.1.2

La convergencia $\frac{1}{\sqrt{n}} \sum_{i=1}^n [\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - \mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, k_0)] \xrightarrow{P} 0$ es consecuencia inmediata del siguiente lema.

Lema B.3.1. Sea $(\mathbf{x}_i, \mathbf{y}_i)_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Asuma que ρ_1 y ρ_2 satisface A10), M_0 satisface A9), entonces

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n [\text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*))) - \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))] \xrightarrow{p} \mathbf{0}, \\ & \frac{1}{\sqrt{n}} \sum_{i=1}^n [\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*))) - \text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \xrightarrow{p} \mathbf{0}, \\ & \frac{1}{\sqrt{n}} \sum_{i=1}^n [\text{vec}(\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s_n(\mathbf{B}_0, \Sigma_0^*))) - \text{vec}(\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)))] \xrightarrow{p} \mathbf{0}, \end{aligned}$$

donde $\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, \cdot)$, $\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0^*, \cdot)$ y $\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, \cdot)$ fueron definidas en (2.2), para $k = 1, 2$.

Demostración del Lema B.3.1. Se usa a continuación el mismo enfoque de Yohai (1985) para el conjunto de funciones continuas definidas en el intervalo $[0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]$, es decir, $C[0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]$.

Para $t \in [0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]$ sea

$$G_{n,j}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, t)_j \quad (\text{B.30})$$

donde $\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, t)_j$ es el j -ésimo elemento del vector $\text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, t))$ para cada $i = 1, \dots, n$. Por el Lema B.2.5 $s_n(\mathbf{B}_0, \Sigma_0^*) \xrightarrow{p} s(\mathbf{B}_0, \Sigma_0^*)$ y por lo tanto es suficiente mostrar que las sucesiones $G_{n,j}(t)_{n \geq 1}, j = 1, \dots, pq$ son tight en $C[0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]$. Esta prueba es análoga a las pruebas de los Lema 5.2 de Fasano (2009) y Lema 5.1 de Yohai (1985) por lo que será omitida.

Se quiere ver que $G_{n,j}(s_n(\mathbf{B}_0, \Sigma_0^*)) - G_{n,j}(s(\mathbf{B}_0, \Sigma_0^*)) \xrightarrow{p} 0$, es decir, para todo $\epsilon, \eta > 0$, existe N tal que si $n \geq N$ entonces

$$P(|G_{n,j}(s_n(\mathbf{B}_0, \Sigma_0^*)) - G_{n,j}(s(\mathbf{B}_0, \Sigma_0^*))| \geq \epsilon) \leq \eta.$$

Sean ϵ y $\eta > 0$ dados, como $s_n(\mathbf{B}_0, \Sigma_0^*) \xrightarrow{p} s(\mathbf{B}_0, \Sigma_0^*)$, existe n_0 tal que

$$n \geq n_0 \Rightarrow P(|s_n(\mathbf{B}_0, \Sigma_0^*) - s(\mathbf{B}_0, \Sigma_0^*)| > 0.5s(\mathbf{B}_0, \Sigma_0^*)) < \frac{\eta}{2}$$

en consecuencia, para $n \geq n_0$, se tiene $P(s_n(\mathbf{B}_0, \Sigma_0^*) \in [0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]) \geq 1 - \frac{\eta}{2}$.

Ya que se supone probado que $G_{n,j}(\cdot)$ es tight, entonces por el Teorema 8.2 de (Billingsley, 1968), existe δ_1 y n_1 tal que

$$n \geq n_1 \Rightarrow P \left(\sup_{\substack{|r-t| < \delta_1; \\ r,t \in [0.5s(\mathbf{B}_0, \Sigma_0^*), 1.5s(\mathbf{B}_0, \Sigma_0^*)]}} |G_{n,j}(r) - G_{n,j}(t)| \geq \epsilon \right) \leq \frac{\eta}{2}. \quad (\text{B.31})$$

Sea $\delta = \min\{\delta_1, 0.5s(\mathbf{B}_0, \Sigma_0^*)\}$, por la convergencia de $s_n(\mathbf{B}_0, \Sigma_0^*)$, existe $n_2 \geq n_0$ tal que

$$n \geq n_2 \Rightarrow P(s_n(\mathbf{B}_0, \Sigma_0^*) - s(\mathbf{B}_0, \Sigma_0^*) > \delta) \leq \frac{\eta}{2}, \quad (\text{B.32})$$

además para todo $n \geq n_2$ se tiene que $s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) \in [0.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*), 1.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)]$ con alta probabilidad, entonces

$$|G_{n,j}(s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - G_{n,j}(s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))| \leq \sup_{\substack{|r-t| < |s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) - s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)|; \\ r, t \in [0.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*); 1.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)]}} |G_{n,j}(r) - G_{n,j}(t)|. \quad (\text{B.33})$$

Por último, sea $N = \max\{n_1, n_2\}$, usando (B.31), (B.32) y (B.33) para $n \geq N$,

$$\begin{aligned} & P(|G_{n,j}(s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - G_{n,j}(s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))| \geq \epsilon) \\ &= P(|G_{n,j}(s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - G_{n,j}(s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))| \geq \epsilon \wedge |s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) - s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)| < \delta) \\ &\quad + P(|G_{n,j}(s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - G_{n,j}(s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))| \geq \epsilon \wedge |s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) - s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)| \geq \delta) \\ &\leq P\left(\sup_{\substack{|r-t| < \delta; \\ r, t \in [0.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*); 1.5s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)]}} |G_{n,j}(r) - G_{n,j}(t)| \geq \epsilon\right) + P(|s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) - s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)| \geq \delta) \\ &\leq \frac{\eta}{2} + \frac{\eta}{2} = \eta, \end{aligned}$$

así $G_{n,j}(s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) - G_{n,j}(s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) \xrightarrow{p} 0$. Luego, se puede notar que

$$\begin{aligned} \Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) &= \frac{\psi_k\left(\frac{d_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}\right)}{\left(\frac{d_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}\right)} \mathbf{u}_i(\mathbf{B}_0) \mathbf{x}_i^t \quad \text{por (2.2)} \\ &= \frac{\psi_k\left(\frac{1/\sqrt{r}d_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{1/\sqrt{r}k_0}\right)}{\left(\frac{1/\sqrt{r}d_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{1/\sqrt{r}k_0}\right)} \mathbf{u}_i(\mathbf{B}_0) \mathbf{x}_i^t \quad \text{por (1.23) y el Lema 1.2.3} \\ &= \Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0, k_0), \end{aligned}$$

en consecuencia, $\frac{1}{\sqrt{n}} \sum_{i=1}^n [\text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s_n(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))) - \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0, k_0))] \xrightarrow{p} \mathbf{0}$. De manera similar para

$$G_{n,j}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, t)_j,$$

donde $\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, t)_j$ es el j -ésimo elemento del vector $\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, t)$ para cada $i = 1, \dots, n$. En el caso $\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, t)$, la demostración es directa por ser continua en $t \neq 0$ para cada $i = 1, \dots, n$.

□

B.4 Demostración del Teorema 6

Lema B.4.1. *Sea $(\mathbf{x}_i, \mathbf{y}_i)_{i \geq 1}$ una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1). Para cualquier función arbitraria $g : \mathbb{R}^q \rightarrow \mathbb{R}$ para la cual la esperanza esté*

definida, se cumple que

$$E \left[\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)g(\mathbf{u}_i(\mathbf{B}_0)) \right] = \mathbf{0}, \quad (\text{B.34})$$

donde vec denota la operación de vectorización.

Demostración del Lema B.4.1. Usando la Ley de las Esperanzas Iteradas se tiene,

$$\begin{aligned} E \left[\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)g(\mathbf{u}_i(\mathbf{B}_0)) \right] &= E \left[E \left[\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)g(\mathbf{u}_i(\mathbf{B}_0)) \mid \mathbf{u}_i(\mathbf{B}_0) \right] \right] \\ &= E \left[g(\mathbf{u}_i(\mathbf{B}_0)) E \left[\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \mid \mathbf{u}_i(\mathbf{B}_0) \right] \right] \\ &= E \left[g(\mathbf{u}_i(\mathbf{B}_0)) \text{vec} \left(E[\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t \mid \mathbf{u}_i(\mathbf{B}_0)] \right) \right] \quad E[\cdot] \text{ y } \text{vec}(\cdot) \text{ conmutan} \\ &= E \left[g(\mathbf{u}_i(\mathbf{B}_0)) \text{vec} \left(\mathbf{u}_i(\mathbf{B}_0) E[\mathbf{x}_i^t \mid \mathbf{u}_i(\mathbf{B}_0)] \right) \right] \\ &= E \left[g(\mathbf{u}_i(\mathbf{B}_0)) \text{vec} \left(\mathbf{u}_i(\mathbf{B}_0) E[\mathbf{x}_i^t \mid \mathbf{u}_i(\mathbf{B}_0)] \right) \right] \quad \text{porque } \mathbf{x}_i \perp\!\!\!\perp \mathbf{u}_i(\mathbf{B}_0) \\ &= \mathbf{0}. \end{aligned}$$

□

Demostración del Teorema. Como $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ es una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^p \times \mathbb{R}^q$ que siguen el MLM (1), resulta que la sucesión $\{(A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))\}_{i \geq 1}$ es de vectores aleatorios independiente idénticamente distribuidos con media $E[A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))]$ y varianza $\text{Var}[A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))]$. Por las condiciones sobre ρ_1 , ρ_2 , y sobre las distribuciones de $\mathbf{u}_i(\mathbf{B}_0)$ y de \mathbf{x}_i , la sucesión $\{(A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))\}_{i \geq 1}$ verifica las hipótesis del Teorema Central del Límite, por lo que la distribución límite de la sucesión involucrada en el enunciado de este teorema es normal. Para calcular la media y la varianza de $A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$ por simplicidad se denota a $d_{(i)0}^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0) = d_{(i)0}^*$ definido en (2.11). Por Definición 2.1.1 la esperanza de $\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, k_0)$ para $k = 1, 2$ viene dada por las siguientes componentes,

$$\begin{aligned} E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))] &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \boldsymbol{\Sigma}_0, k_0))] \\ &= E \left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \right] \\ &= E \left[E \left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \mid \mathbf{u}_i(\mathbf{B}_0) \right] \right] \\ &= E \left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} E \left[\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \mid \mathbf{u}_i(\mathbf{B}_0) \right] \right] \\ &= E \left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \text{vec} \left(\mathbf{u}_i(\mathbf{B}_0) E[\mathbf{x}_i^t \mid \mathbf{u}_i(\mathbf{B}_0)] \right) \right] \\ &= E \left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \text{vec} \left(\mathbf{u}_i(\mathbf{B}_0) E[\mathbf{x}_i^t] \right) \right] \\ &= \mathbf{0} \quad \text{ya que } E[\mathbf{x}_i] = \mathbf{0} \text{ y } \mathbf{u}_i(\mathbf{B}_0) \perp\!\!\!\perp \mathbf{x}_i. \end{aligned}$$

De forma análoga se tiene,

$$\begin{aligned} E[\text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)))] &= E \left[\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \text{vec}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{x}_i^t) \right] \\ &= \mathbf{0}. \end{aligned}$$

Asimismo, se calculará el siguiente término,

$$\begin{aligned} E[\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)))] &= E[\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\ &= E \left[q \frac{\psi_k(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) - k_0^2 \psi_k(d_{(i)0}^*) d_{(i)0}^* \right. \\ &\quad \left. \times \text{vech}(\Sigma_0) \right] \end{aligned} \tag{B.35}$$

$$\begin{aligned} &= q E \left[\frac{\psi_k(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0) \mathbf{u}_i^t(\mathbf{B}_0)) \right] - k_0^2 E \left[\psi_k(d_{(i)0}^*) d_{(i)0}^* \right] \\ &\quad \times \text{vech}(\Sigma_0). \end{aligned} \tag{B.36}$$

Dado que la distribución de $\mathbf{u}_i(\mathbf{B}_0)$ se asume elíptica, tomando $\mathbf{z}_i = \Sigma_0^{-1/2} \mathbf{u}_i(\mathbf{B}_0)$ resulta que \mathbf{z}_i también tiene distribución elíptica con matriz de covarianza \mathbf{I}_q obteniendo que,

$$\begin{aligned} d_{(i)0}^* &= \frac{(\mathbf{u}_i^t(\mathbf{B}_0) \Sigma_0^{-1} \mathbf{u}_i(\mathbf{B}_0))^{1/2}}{k_0} \\ &= \frac{(\mathbf{z}_i^t \mathbf{z}_i)^{1/2}}{k_0} \\ &= \frac{\|\mathbf{z}_i\|}{k_0}. \end{aligned}$$

Sustituyendo en la ecuación (B.36) resulta,

$$\begin{aligned} E[\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] &= q E \left[\frac{\psi_k \left(\frac{\|\mathbf{z}_i\|}{k_0} \right)}{\frac{\|\mathbf{z}_i\|}{k_0}} \text{vech}(\Sigma_0^{1/2} \mathbf{z}_i \mathbf{z}_i^t \Sigma_0^{1/2}) \right] \\ &\quad - k_0^2 E \left[\psi_k \left(\frac{\|\mathbf{z}_i\|}{k_0} \right) \frac{\|\mathbf{z}_i\|}{k_0} \right] \text{vech}(\Sigma_0) \\ &= q \text{vech} \left(\Sigma_0^{1/2} E \left[\frac{\psi_k \left(\frac{\|\mathbf{z}_i\|}{k_0} \right)}{\frac{\|\mathbf{z}_i\|}{k_0}} \mathbf{z}_i \mathbf{z}_i^t \right] \Sigma_0^{1/2} \right) \\ &\quad - k_0^2 E \left[\psi_k \left(\frac{\|\mathbf{z}_i\|}{k_0} \right) \frac{\|\mathbf{z}_i\|}{k_0} \right] \text{vech}(\Sigma_0). \end{aligned}$$

Si $\mathbf{z} = (z_1, \dots, z_q)^t$ tiene distribución elíptica, entonces $\|\mathbf{z}\|$ y $\frac{\mathbf{z}}{\|\mathbf{z}\|}$ son independientes. Esto implica que para cualquier función h medible tal que $E[h(\|\mathbf{z}\|)z_j z_k] < \infty$ para $1 \leq j, k \leq q$

vale que

$$E[h(\|\mathbf{z}\|)z_j z_k] = 0 \text{ si } 1 \leq j \neq k \leq q,$$

y

$$E[h(\|\mathbf{z}\|)z_j^2] = \frac{E[h(\|\mathbf{z}\|)\|\mathbf{z}\|^2]}{q} \text{ para todo } 1 \leq j \leq q.$$

Utilizando estas propiedades y tomando $h(\|\mathbf{z}\|) = \frac{\psi\left(\frac{\|\mathbf{z}\|}{k_0}\right)}{\|\mathbf{z}\|}$ se tiene,

$$\begin{aligned} E[\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] &= \text{vech}\left(\Sigma_0^{1/2} E\left[\psi_k\left(\frac{\|\mathbf{z}_i\|}{k_0}\right) k_0 \|\mathbf{z}_i\|\right] \Sigma_0^{1/2}\right) \\ &\quad - k_0^2 E\left[\psi_k\left(\frac{\|\mathbf{z}_i\|}{k_0}\right) \frac{1}{k_0} \|\mathbf{z}_i\|\right] \text{vech}(\Sigma_0) \\ &= E\left[\psi_k\left(\frac{\|\mathbf{z}_i\|}{k_0}\right) k_0 \|\mathbf{z}_i\|\right] \text{vech}(\Sigma_0) \\ &\quad - E\left[\psi_k\left(\frac{\|\mathbf{z}_i\|}{k_0}\right) k_0 \|\mathbf{z}_i\|\right] \text{vech}(\Sigma_0) \\ &= \mathbf{0}. \end{aligned}$$

Por último,

$$\begin{aligned} E_{2,\kappa_2} &:= E[\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] = E\left[s^2(\mathbf{B}_0, \Sigma_0^*) \rho_2(d_{(i)0}^*) - \kappa_2\right] \\ &= s^2(\mathbf{B}_0, \Sigma_0^*) E\left[\rho_2(d_{(i)0}^*)\right] - \kappa_2 \\ &= 0. \end{aligned} \quad \text{por (1.27)}$$

Dado que $E[\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] = \mathbf{0}$, para calcular la matriz de covarianza de $\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))$ es suficiente calcular

$$\text{Var}(\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))) = E\left[\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*)) \mathbf{A}_i^t(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))\right].$$

Para realizar este cálculo, se emplearán las siguientes propiedades de matrices:

B1) $\text{vec}(ABC) = (C^t \otimes A)\text{vec}(C)$,

B2) Si A , B , C y D son matrices de tal tamaño que se pueden formar los productos matriciales AC y BD , entonces $(A \otimes B)(C \otimes D) = AC \otimes BD$,

B3) $(A \otimes B)^t = A^t \otimes B^t$.

A continuación, se procederá a calcular este valor en detalle

$$E[\mathbf{A}_i(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))\mathbf{A}_i^t(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] = \begin{pmatrix} L_{11}\{pq \times pq\} & L_{12}\{pq \times q(q+1)/2\} & L_{13}\{pq \times pq\} \\ L_{12}^t & L_{22}\{q(q+1)/2 \times q(q+1)/2\} & L_{23}\{q(q+1)/2 \times pq\} \\ L_{13}^t & L_{23}^t & L_{33}\{pq \times pq\} \\ L_{14}^t & L_{24}^t & L_{34}^t \\ L_{15}^t & L_{25}^t & L_{35}^t \\ \\ L_{14}\{pq \times q(q+1)/2\} & L_{15}\{pq \times 1\} \\ L_{24}\{q(q+1)/2 \times q(q+1)/2\} & L_{25}\{q(q+1)/2 \times 1\} \\ L_{34}\{pq \times q(q+1)/2\} & L_{35}\{pq \times 1\} \\ L_{44}\{q(q+1)/2 \times q(q+1)/2\} & L_{45}\{q(q+1)/2 \times 1\} \\ L_{45}^t & L_{55}\{1 \times 1\} \end{pmatrix},$$

donde

$$\begin{aligned} L_{11} &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vec}^t(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))] \\ &= E \left[\left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vec}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \right] \\ &= E[\mathbf{x}_i\mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0) \right], \end{aligned} \quad \mathbf{u}_i(\mathbf{B}_0) \perp \mathbf{x}_i,$$

$$\begin{aligned} L_{12} &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\ &= E \left[q \left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right. \\ &\quad \left. - k_0^2 \psi_1^2(d_{(i)0}^*)\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\Sigma_0) \right], \\ &= \mathbf{0}, \end{aligned} \quad \text{por el Lema B.4.1,}$$

$$\begin{aligned} L_{13} &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vec}^t(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))] \\ &= E \left[\frac{\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}} \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vec}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \right] \\ &= E[\mathbf{x}_i\mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}} \right) \mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0) \right] \end{aligned} \quad \mathbf{u}_i(\mathbf{B}_0) \perp \mathbf{x}_i,$$

$$\begin{aligned} L_{14} &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\ &= E \left[q \frac{\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}} \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right. \\ &\quad \left. - k_0^2 \psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\Sigma_0) \right] \\ &= \mathbf{0}, \end{aligned} \quad \text{por el Lema B.4.1,}$$

$$\begin{aligned}
 L_{15} &= E[\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \\
 &= E\left[\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*}\{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\}\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\right] \\
 &= \mathbf{0},
 \end{aligned}$$

por el Lema B.4.1,

$$\begin{aligned}
 L_{22} &= E[\text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= q^2 E\left[\left(\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*}\right)^2 \text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad - qk_0^2 E\left[\psi_1^2(d_{(i)0}^*)\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \text{vech}^t(\Sigma_0) \\
 &\quad - qk_0^2 \text{vech}(\Sigma_0) E\left[\psi_1^2(d_{(i)0}^*)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad + k_0^4 E\left[\psi_1^2(d_{(i)0}^*)d_{(i)0}^{*2}\right] \text{vech}(\Sigma_0)\text{vech}^t(\Sigma_0),
 \end{aligned}$$

$$\begin{aligned}
 L_{23} &= E[\text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\text{vec}^t(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= E\left[q\frac{\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}}\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\text{vec}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\right. \\
 &\quad \left.- k_0^2\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)\text{vech}(\Sigma_0)\text{vec}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\right] \\
 &= \mathbf{0},
 \end{aligned}$$

por el Lema B.4.1,

$$\begin{aligned}
 L_{24} &= E[\text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= q^2 E\left[\frac{\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)}{d_{(i)0}^{*2}}\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad - qk_0^2 E\left[\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \text{vech}^t(\Sigma_0) \\
 &\quad - qk_0^2 \text{vech}(\Sigma_0) E\left[\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad + k_0^4 E\left[\psi_1(d_{(i)0}^*)\psi_2(d_{(i)0}^*)d_{(i)0}^{*2}\right] \text{vech}(\Sigma_0)\text{vech}^t(\Sigma_0),
 \end{aligned}$$

$$\begin{aligned}
 L_{25} &= E[\text{vech}(\Psi_{i,1,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \\
 &= E\left[\left\{q\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*}\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) - k_0^2\psi_1(d_{(i)0}^*)d_{(i)0}^{*2}\text{vech}(\Sigma_0)\right\}\right. \\
 &\quad \left.\times \{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\}\right]
 \end{aligned}$$

de (2.2)

$$\begin{aligned}
 &= qE\left[\left\{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\right\}\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*}\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad - E\left[\left\{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\right\}k_0^2\psi_1(d_{(i)0}^*)d_{(i)0}^{*2}\right] \text{vech}(\Sigma_0), \\
 &= qE\left[\left\{\frac{k_0^2}{r}\rho_2(d_{(i)0}^*) - \kappa_2\right\}\frac{\psi_1(d_{(i)0}^*)}{d_{(i)0}^*}\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\right] \\
 &\quad - k_0^2 E\left[\left\{\frac{k_0}{r}\rho_2(d_{(i)0}^*) - \kappa_2\right\}\psi_1(d_{(i)0}^*)d_{(i)0}^{*2}\right] \text{vech}(\Sigma_0),
 \end{aligned}$$

por (1.23) y Lema 1.2.3

$$\begin{aligned}
 L_{33} &= E[\text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vec}^t(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= E \left[\left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vec}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \right] \\
 &= E[\mathbf{x}_i\mathbf{x}_i^t] \otimes E \left[\left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0) \right], \quad \mathbf{u}_i(\mathbf{B}_0) \perp \mathbf{x}_i
 \end{aligned}$$

$$\begin{aligned}
 L_{34} &= E[\text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= E \left[q \left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right. \\
 &\quad \left. - k_0^2 \psi_2^2(d_{(i)0}^*)\text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t)\text{vech}^t(\Sigma_0) \right] \\
 &= \mathbf{0}
 \end{aligned}$$

por el Lema B.4.1,

$$\begin{aligned}
 L_{35} &= E[\text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}_0, \Sigma_0, k_0))\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \\
 &= E \left[\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\} \text{vec}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{x}_i^t) \right] \\
 &= \mathbf{0}
 \end{aligned}$$

por el Lema B.4.1,

$$\begin{aligned}
 L_{44} &= E[\text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\text{vech}^t(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))] \\
 &= q^2 E \left[\left(\frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \right)^2 \text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0))\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right] \\
 &\quad - qk_0^2 E \left[\psi_2^2(d_{(i)0}^*)\text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right] \text{vech}^t(\Sigma_0) \\
 &\quad - qk_0^2 \text{vech}(\Sigma_0) E \left[\psi_2^2(d_{(i)0}^*)\text{vech}^t(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right] \\
 &\quad + k_0^4 E \left[\psi_2^2(d_{(i)0}^*)d_{(i)0}^{*2} \right] \text{vech}(\Sigma_0)\text{vech}^t(\Sigma_0),
 \end{aligned}$$

$$\begin{aligned}
 L_{45} &= E[\text{vech}(\Psi_{i,2,\Sigma}(\mathbf{B}_0, \Sigma_0, k_0))\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \Sigma_0^*, s(\mathbf{B}_0, \Sigma_0^*))] \\
 &= E \left[\left\{ q \frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) - k_0^2 \psi_2(d_{(i)0}^*)d_{(i)0}^* \text{vech}(\Sigma_0) \right\} \right. \\
 &\quad \left. \times \{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\} \right]
 \end{aligned}$$

por (2.2)

$$\begin{aligned}
 &= qE \left[\{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\} \frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right] \\
 &\quad - E \left[\{s^2(\mathbf{B}_0, \Sigma_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\} k_0^2 \psi_2(d_{(i)0}^*)d_{(i)0}^* \right] \text{vech}(\Sigma_0), \\
 &= qE \left[\left\{ \frac{k_0}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \frac{\psi_2(d_{(i)0}^*)}{d_{(i)0}^*} \text{vech}(\mathbf{u}_i(\mathbf{B}_0)\mathbf{u}_i^t(\mathbf{B}_0)) \right] \\
 &\quad - k_0^2 E \left[\left\{ \frac{k_0}{r} \rho_2(d_{(i)0}^*) - \kappa_2 \right\} \psi_2(d_{(i)0}^*)d_{(i)0}^* \right] \text{vech}(\Sigma_0),
 \end{aligned}$$

por (1.23) y Lema 1.2.3

$$\begin{aligned}
 L_{55} &= E \left[(\Psi_{i,2,\kappa_2}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)))^2 \right] && \text{por (2.2)} \\
 &= E \left[\{s^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\rho_2(d_{(i)0}^*) - \kappa_2\}^2 \right] \\
 &= E \left[s^4(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\rho_2^2(d_{(i)0}^*) - 2\kappa_2 s^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\rho_2(d_{(i)0}^*) + \kappa_2^2 \right] \\
 &= E \left[s^4(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\rho_2^2(d_{(i)0}^*) \right] - 2\kappa_2 E \left[s^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\rho_2(d_{(i)0}^*) \right] + \kappa_2^2 \\
 &= s^4(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) E \left[\rho_2^2(d_{(i)0}^*) \right] - 2\kappa_2 s^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) E \left[\rho_2(d_{(i)0}^*) \right] + \kappa_2^2 \\
 &= \frac{\kappa_0^4}{r^2} E \left[\rho_2^2(d_{(i)0}^*) \right] - 2\kappa_2 \frac{\kappa_0^2}{r} E \left[\rho_2(d_{(i)0}^*) \right] + \kappa_2^2. && \text{por (1.23) y Lema 1.2.3}
 \end{aligned}$$

En consecuencia, la matriz de covarianza de $A_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))$ se reduce a,

$$\text{Var}(\mathbf{A}_i(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*, s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))) = \begin{pmatrix} L_{11} & \mathbf{0} & L_{13} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_{22} & \mathbf{0} & L_{24} & L_{25} \\ L_{13}^t & \mathbf{0} & L_{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_{24}^t & \mathbf{0} & L_{44} & L_{45} \\ \mathbf{0} & L_{25}^t & \mathbf{0} & L_{45}^t & L_{55} \end{pmatrix}$$

Sustituyendo los valores de L_{ij} para $i, j = 1, \dots, 5$ se obtiene el resultado deseado. \square

B.5 Demostración del Lema 2.1.3

Demostración del Lema 2.1.3. Se denota $\boldsymbol{\theta}_n^* = (\mathbf{B}_n^*, \boldsymbol{\Sigma}_n^*)$ y $s_n := s_n(\mathbf{B}, \boldsymbol{\Sigma})$, por definición,

$$\begin{aligned}
 \mathbf{V}_n &= \mathbf{M}_{(C_n, D_n)}(\mathbf{B}_n, \boldsymbol{\Sigma}_n) (F_{1,n}(\boldsymbol{\theta}_n^*), F_{2,n}(\boldsymbol{\theta}_n^*)) \\
 &= \sum_{i=1}^n \begin{pmatrix} C_n(\mathbf{B}, \boldsymbol{\Sigma}) \mathbf{I}_{pq} & \mathbf{0} & D_n(\mathbf{B}, \boldsymbol{\Sigma}) \mathbf{I}_{pq} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_n(\mathbf{B}, \boldsymbol{\Sigma}) \mathbf{I}_{q(q+1)/2} & \mathbf{0} & D_n(\mathbf{B}, \boldsymbol{\Sigma}) \mathbf{I}_{q(q+1)/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix} \\
 &\quad \times \begin{pmatrix} \frac{\partial \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial [\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))]}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\ \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\ \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\ \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \\ \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vec}^t(\mathbf{B}^t)} & \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \end{pmatrix} \Bigg|_{\boldsymbol{\theta}_n^*} \\
 &= \sum_{i=1}^n \begin{pmatrix} \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} \right] \Bigg|_{\boldsymbol{\theta}_n^*} \\ \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} \right] \Bigg|_{\boldsymbol{\theta}_n^*} \\ \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vec}^t(\mathbf{B}^t)} \Bigg|_{\boldsymbol{\theta}_n^*} \end{pmatrix} \\
 &\quad \begin{pmatrix} \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial [\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))]}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \right] \Bigg|_{\boldsymbol{\theta}_n^*} \\ \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \right] \Bigg|_{\boldsymbol{\theta}_n^*} \\ \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Bigg|_{\boldsymbol{\theta}_n^*} \end{pmatrix} \\
 &= \begin{pmatrix} f_{11(n)} & f_{12(n)} \\ f_{21(n)} & f_{22(n)} \\ f_{31(n)} & f_{32(n)} \end{pmatrix}, \tag{B.37}
 \end{aligned}$$

donde

$$\begin{aligned}
 f_{11(n)} &:= \sum_{i=1}^n \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} \right] \Big|_{\boldsymbol{\theta}_n^*} \\
 f_{12(n)} &:= \sum_{i=1}^n \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial [\text{vec}(\Psi_{i,1,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))]}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vec}(\Psi_{i,2,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \right] \Big|_{\boldsymbol{\theta}_n^*} \\
 f_{21(n)} &:= \sum_{i=1}^n \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vec}^t(\mathbf{B}^t)} \right] \Big|_{\boldsymbol{\theta}_n^*} \\
 f_{22(n)} &:= \sum_{i=1}^n \left[C_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,1,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} + D_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial \text{vech}(\Psi_{i,2,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \right] \Big|_{\boldsymbol{\theta}_n^*} \\
 f_{31(n)} &:= \sum_{i=1}^n \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{\boldsymbol{\theta}_n^*} \\
 f_{32(n)} &:= \sum_{i=1}^n \frac{\partial \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n)}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{\boldsymbol{\theta}_n^*}.
 \end{aligned}$$

La matriz de Moore Penrose (denotada por \mathbf{V}_n^\dagger) para columnas linealmente independientes es de la forma $(\mathbf{V}_n^t \mathbf{V}_n)^{-1} \mathbf{V}_n^t$ (Maruskin, 2012). A continuación se calcula,

$$\begin{aligned}
 \mathbf{V}_n^t \mathbf{V}_n &= \begin{pmatrix} f_{11(n)}^t & f_{21(n)}^t & f_{31(n)}^t \\ f_{12(n)}^t & f_{22(n)}^t & f_{32(n)}^t \end{pmatrix} \begin{pmatrix} f_{11(n)} & f_{12(n)} \\ f_{21(n)} & f_{22(n)} \\ f_{31(n)} & f_{32(n)} \end{pmatrix} \\
 &= \begin{pmatrix} f_{11(n)}^t f_{11(n)} + f_{21(n)}^t f_{21(n)} + f_{31(n)}^t f_{31(n)} & f_{11(n)}^t f_{12(n)} + f_{21(n)}^t f_{22(n)} + f_{31(n)}^t f_{32(n)} \\ f_{12(n)}^t f_{11(n)} + f_{22(n)}^t f_{21(n)} + f_{32(n)}^t f_{31(n)} & f_{12(n)}^t f_{12(n)} + f_{22(n)}^t f_{22(n)} + f_{32(n)}^t f_{32(n)} \end{pmatrix}
 \end{aligned}$$

luego, se busca la inversa $(\mathbf{V}_n^t \mathbf{V}_n)^{-1}$ particionando la matriz anterior de la siguiente manera

$$\left(\begin{array}{c|c} f_{11(n)}^t f_{11(n)} + f_{21(n)}^t f_{21(n)} + f_{31(n)}^t f_{31(n)} & f_{11(n)}^t f_{12(n)} + f_{21(n)}^t f_{22(n)} + f_{31(n)}^t f_{32(n)} \\ \hline f_{12(n)}^t f_{11(n)} + f_{22(n)}^t f_{21(n)} + f_{32(n)}^t f_{31(n)} & f_{12(n)}^t f_{12(n)} + f_{22(n)}^t f_{22(n)} + f_{32(n)}^t f_{32(n)} \end{array} \right) := \begin{pmatrix} A & B \\ B^t & G \end{pmatrix}$$

considerando A y $E = G - B^t A^{-1} B$ no singulares por Petersen and Pedersen (2012) se tiene,

$$\begin{pmatrix} A & B \\ B^t & G \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1} B E^{-1} B^t A^{-1} & -A^{-1} B E^{-1} \\ -E^{-1} B^t A^{-1} & E^{-1} \end{pmatrix}. \quad (\text{B.38})$$

Por último, falta multiplicar por derecha por \mathbf{V}_n^t en (B.38),

$$(\mathbf{V}_n^t \mathbf{V}_n)^{-1} \mathbf{V}_n^t = \begin{pmatrix} A^{-1} + A^{-1} B E^{-1} B^t A^{-1} & -A^{-1} B E^{-1} \\ -E^{-1} B^t A^{-1} & E^{-1} \end{pmatrix} \begin{pmatrix} f_{11(n)}^t & f_{21(n)}^t & f_{31(n)}^t \\ f_{12(n)}^t & f_{22(n)}^t & f_{32(n)}^t \end{pmatrix}$$

obteniendo,

$$\begin{aligned}
 \mathbf{V}_n^\dagger &= \begin{pmatrix} A^{-1} [(\mathbf{I} + B E^{-1} B^t A^{-1}) f_{11(n)}^t - B E^{-1} f_{12(n)}^t] & A^{-1} [(\mathbf{I} + B E^{-1} B^t A^{-1}) f_{21(n)}^t - B E^{-1} f_{22(n)}^t] \\ E^{-1} [-B^t A^{-1} f_{11(n)}^t + f_{12(n)}^t] & E^{-1} [-B^t A^{-1} f_{21(n)}^t + f_{22(n)}^t] \\ A^{-1} [(\mathbf{I} + B E^{-1} B^t A^{-1}) f_{31(n)}^t - B E^{-1} f_{32(n)}^t] & \\ E^{-1} [-B^t A^{-1} f_{31(n)}^t + f_{32(n)}^t] & \end{pmatrix}. \quad (\text{B.39})
 \end{aligned}$$

□

B.6 Demostración del Teorema 8

En esta sección, se emplea la notación $\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})$ para $k = 1, 2$, definida como:

$$\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma}) := \psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) - \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})). \quad (\text{B.40})$$

Aquí, d_i y d_i^* se utilizan como notación abreviada para $d_i(\mathbf{B}, \boldsymbol{\Sigma})$ y $d_i^*(\mathbf{B}, \boldsymbol{\Sigma})$ con $d_i^*(\mathbf{B}, \boldsymbol{\Sigma})$ definida en (1.21), respectivamente, siempre y cuando no haya confusión en su uso.

Para establecer la convergencia de \mathbf{V}_n^\dagger , se procederá a analizar la convergencia de cada uno de sus elementos, representados por las derivadas asociadas a $\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$, $\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$ y $\Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))$, $k = 1, 2$, como indica la expresión (2.15). Por consiguiente, el primer paso consistirá en calcular estas derivadas, para posteriormente examinar su convergencia. Una vez establecida la convergencia de las derivadas, se procederá a derivar la demostración correspondiente de la convergencia de \mathbf{V}_n^\dagger en su totalidad.

Lema B.6.1. *Las derivadas de $\text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))$, con respecto a $\text{vec}^t(\mathbf{B}^t)$ y $\text{vech}^t(\boldsymbol{\Sigma})$ son,*

$$\begin{aligned} \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) &= \frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\ &\quad - \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{x}_i\mathbf{x}_i^t \otimes \mathbf{I}_q), \\ \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) &= \frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})}, \end{aligned}$$

para $k = 1, 2$.

Demostración del Lema B.6.1. Utilizando la definición de la función $\text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))$ y la igualdad (B.40), junto con las siguientes derivadas

$$\frac{\partial \mathbf{u}_i(\mathbf{B})}{\partial \text{vec}^t(\mathbf{B}^t)} = -(\mathbf{x}_i^t \otimes \mathbf{I}_q) \quad \text{y} \quad \frac{\partial \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t)}{\partial \text{vec}^t(\mathbf{B}^t)} = -(\mathbf{x}_i\mathbf{x}_i^t \otimes \mathbf{I}_q),$$

las cuales son válidas según Lütkepohl (1996, p. 183, 10.4.1(5)), se tiene que

$$\begin{aligned} \frac{\partial \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))}{\partial \text{vec}^t(\mathbf{B}^t)} &= \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \\ &= \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \left[\frac{\psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) - \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)}}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \\ &\quad + \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} \frac{\partial \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t)}{\partial \text{vec}^t(\mathbf{B}^t)} \\ &= \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \left[\frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)}}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] - \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{x}_i\mathbf{x}_i^t \otimes \mathbf{I}_q) \\ &= \left[\frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} - \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{x}_i\mathbf{x}_i^t \otimes \mathbf{I}_q). \end{aligned}$$

De manera similar para la otra derivada,

$$\begin{aligned}
 \frac{\partial \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))}{\partial \text{vech}^t(\boldsymbol{\Sigma})} &= \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \\
 &= \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} \\
 &= \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \left[\frac{[\psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) - \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))] \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})}}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \\
 &= \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \left[\frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})}}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \quad \text{por (B.40)} \\
 &= \left[\frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})},
 \end{aligned}$$

obteniendo el valor deseado. \square

Lema B.6.2. Las derivadas de $\text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))$, con respecto a $\text{vec}^t(\mathbf{B}^t)$ y $\text{vec}^t(\boldsymbol{\Sigma})$ son,

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) &= -q \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{u}_i(\mathbf{B})\mathbf{x}_i^t \otimes \mathbf{I}_q) \\
 &\quad - 2s_n(\mathbf{B}, \boldsymbol{\Sigma})\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma})\text{vec}(\boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} + \left(q \frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \right. \\
 &\quad \left. - \left[s_n^2(\mathbf{B}, \boldsymbol{\Sigma})\psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) + s_n^2(\mathbf{B}, \boldsymbol{\Sigma})\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right] \text{vec}(\boldsymbol{\Sigma}) \right) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)},
 \end{aligned}$$

y,

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) &= \left\{ q \frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \right. \\
 &\quad \left. - s_n^2(\mathbf{B}, \boldsymbol{\Sigma})d_i^*(\mathbf{B}, \boldsymbol{\Sigma})\psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))\text{vec}(\boldsymbol{\Sigma}) - s_n^2(\mathbf{B}, \boldsymbol{\Sigma})\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))\text{vec}(\boldsymbol{\Sigma}) \right\} \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\boldsymbol{\Sigma})} \\
 &\quad - 2s_n(\mathbf{B}, \boldsymbol{\Sigma})\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma})\text{vec}(\boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\boldsymbol{\Sigma})} - s_n^2(\mathbf{B}, \boldsymbol{\Sigma})\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))d_i^*(\mathbf{B}, \boldsymbol{\Sigma})\mathbf{I}_{q^2},
 \end{aligned}$$

para $k = 1, 2$, donde $\mathbf{K}_{q,q}$ se define como una matriz de conmutación de tamaño $q^2 \times q^2$, que transforma $\text{vec}(\mathbf{A})$ en $\text{vec}(\mathbf{A}^t)$ para cualquier matriz $\mathbf{A}_{q \times q}$.

Demostración del Lema B.6.2. Usando la definición de la función $\text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))$

dada en (2.2),

$$\begin{aligned}
 & \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} [\text{vec}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \\
 &= \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \left[q \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \text{vec}(\Sigma) \right] \\
 &= q \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \left[\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} \right] + q \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} \frac{\partial \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B}))}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &\quad - \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \text{vec}(\Sigma). \tag{B.41}
 \end{aligned}$$

Por propiedades de (Magnus and Neudecker, 1999) se obtiene

$$\begin{aligned}
 \frac{\partial \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B}))}{\partial \text{vec}^t(\mathbf{B}^t)} &= \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vec}((\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i)(\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i)^t) \\
 &= -(\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{y}_i - \mathbf{B}^t \mathbf{x}_i) \mathbf{x}_i^t \otimes \mathbf{I}_q \\
 &= -(\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \otimes \mathbf{I}_q), \tag{B.42}
 \end{aligned}$$

entonces sustituyendo en (B.41), calculando la derivada de $\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)}$, y aplicando propiedades de derivadas, se tiene

$$\begin{aligned}
 & \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} [\text{vec}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \\
 &= q \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \left[\frac{[\psi_k'(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) - \psi_k(d_i^*(\mathbf{B}, \Sigma))] \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}(\mathbf{B}^t)}}{d_i^{*2}(\mathbf{B}, \Sigma)} \right] \\
 &\quad + q \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} [-(\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \otimes \mathbf{I}_q)] \\
 &\quad - \text{vec}(\Sigma) \left[\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \right] \\
 &= q \text{vec}(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \left[\frac{\phi_{i,k}(\mathbf{B}, \Sigma)}{d_i^{*2}(\mathbf{B}, \Sigma)} \right] \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}(\mathbf{B}^t)} - q \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \otimes \mathbf{I}_q) \\
 &\quad - \text{vec}(\Sigma) \left[\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \right], \tag{B.43}
 \end{aligned}$$

la última igualdad se obtuvo por (B.40). Luego, se observa que

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) &= \frac{\partial s_n^2(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \\
 &+ s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{\partial \text{vec}^t(\mathbf{B}^t)} d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \\
 &+ s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &= 2s_n(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &+ s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &+ s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &= 2s_n(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \\
 &+ \left[s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \right. \\
 &\left. + s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right] \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)},
 \end{aligned}$$

aplicando propiedades algebraicas y sustituyendo la igualdad anterior en (B.43), resulta

$$\begin{aligned}
 &\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} [\text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))] \\
 &= q \text{vec}(\mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B})) \left[\frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \right] \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} - q \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \otimes \mathbf{I}_q) \\
 &\quad - \text{vec}(\boldsymbol{\Sigma}) \left[2s_n(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} + \left[s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \right. \right. \\
 &\quad \left. \left. + s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right] \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \right] \quad \text{por (B.40)} \\
 &= -q \frac{\psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma}))}{d_i^*(\mathbf{B}, \boldsymbol{\Sigma})} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}) \mathbf{x}_i^t \otimes \mathbf{I}_q) - 2s_n(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \\
 &\quad \times \text{vec}(\boldsymbol{\Sigma}) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} + \left(q \frac{\phi_{i,k}(\mathbf{B}, \boldsymbol{\Sigma})}{d_i^{*2}(\mathbf{B}, \boldsymbol{\Sigma})} \text{vec}(\mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B})) - \left[s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi'_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) d_i^*(\mathbf{B}, \boldsymbol{\Sigma}) \right. \right. \\
 &\quad \left. \left. + s_n^2(\mathbf{B}, \boldsymbol{\Sigma}) \psi_k(d_i^*(\mathbf{B}, \boldsymbol{\Sigma})) \right] \text{vec}(\boldsymbol{\Sigma}) \right) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)}.
 \end{aligned}$$

Ahora, para conseguir la otra derivada se realiza un procedimiento similar al anterior, por la

definición de $vec(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))$ dada en (2.2),

$$\begin{aligned}
 & \frac{\partial}{\partial vec^t(\Sigma)} [vec(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \\
 &= \frac{\partial}{\partial vec^t(\Sigma)} \left[q \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} vec(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \right] \\
 &= q vec(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \frac{\partial}{\partial vec^t(\Sigma)} \frac{\psi_k(d_i^*(\mathbf{B}, \Sigma))}{d_i^*(\mathbf{B}, \Sigma)} - \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \frac{\partial s_n^2(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \frac{\partial \psi_k(d_i^*(\mathbf{B}, \Sigma))}{\partial vec^t(\Sigma)} - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) vec(\Sigma) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \frac{\partial vec(\Sigma)}{\partial vec^t(\Sigma)} \\
 &= q vec(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \left[\frac{[\psi_k'(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) - \psi_k(d_i^*(\mathbf{B}, \Sigma))] \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)}}{d_i^{*2}(\mathbf{B}, \Sigma)} \right] \\
 &\quad - 2s_n(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \left(\psi_k'(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \right) - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) vec(\Sigma) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \frac{\partial vec(\Sigma)}{\partial vec^t(\Sigma)} \\
 &= q vec(\mathbf{u}_i(\mathbf{B})\mathbf{u}_i^t(\mathbf{B})) \left[\frac{\phi_{i,k}(\mathbf{B}, \Sigma)}{d_i^{*2}(\mathbf{B}, \Sigma)} \right] \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} - 2s_n(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) d_i^*(\mathbf{B}, \Sigma) vec(\Sigma) \left(\psi_k'(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \right) - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) vec(\Sigma) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \\
 &\quad - s_n^2(\mathbf{B}, \Sigma) \psi_k(d_i^*(\mathbf{B}, \Sigma)) d_i^*(\mathbf{B}, \Sigma) \mathbf{I}_{q^2}, \quad \text{por (B.40)}
 \end{aligned}$$

ya que

$$\frac{\partial vec(\Sigma)}{\partial vec^t(\Sigma)} = \mathbf{I}_{q^2}, \quad (\text{B.44})$$

las derivadas obtenidas quedan compuestas por derivadas descritas anteriormente, específicamente en los Lemas A.5.2 y A.5.3. \square

Lema B.6.3. *Las derivadas de $vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))$ con respecto a $vec^t(\mathbf{B}^t)$ y $vech^t(\Sigma)$ son,*

$$\frac{\partial}{\partial vec^t(\mathbf{B}^t)} vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) = \mathbf{L}_q \frac{\partial}{\partial vec^t(\mathbf{B}^t)} [vec(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] ,$$

y

$$\frac{\partial}{\partial vech^t(\Sigma)} vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) = \mathbf{L}_q \frac{\partial}{\partial vec^t(\Sigma)} [vec(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \mathbf{D}_q ,$$

para $k = 1, 2$, donde \mathbf{L}_q es la matriz de eliminación que proporciona la siguiente igualdad, $\mathbf{L}_q vec(A) = vech(A)$ para cualquier matriz simétrica $A_{q \times q}$. Y \mathbf{D}_q es la matriz de duplicación que proporciona la igualdad $\mathbf{D}_q vech(A) = vec(A)$ para cualquier matriz simétrica $A_{q \times q}$.

Demostración del Lema B.6.3.

$$\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} [\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] = \mathbf{L}_q \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} [\text{vec}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] ,$$

y usando

$$\frac{\partial \text{vec}(\Sigma)}{\partial \text{vech}^t(\Sigma)} = \mathbf{D}_q,$$

resulta

$$\frac{\partial}{\partial \text{vech}^t(\Sigma)} [\text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] = \mathbf{L}_q \frac{\partial}{\partial \text{vec}^t(\Sigma)} [\text{vec}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \mathbf{D}_q.$$

□

Lema B.6.4. *Las derivadas de $\Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))$ con respecto a $\text{vec}^t(\mathbf{B}^t)$, $\text{vec}^t(\Sigma)$ y $\text{vech}^t(\Sigma)$ son,*

$$\begin{aligned} \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)) &= 2s_n(\mathbf{B}, \Sigma) \rho_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \\ &\quad + s_n^2(\mathbf{B}, \Sigma) \psi_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)}, \end{aligned} \quad (\text{B.45})$$

$$\begin{aligned} \frac{\partial}{\partial \text{vec}^t(\Sigma)} \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)) &= 2s_n(\mathbf{B}, \Sigma) \rho_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\Sigma)} \\ &\quad + s_n^2(\mathbf{B}, \Sigma) \psi_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\Sigma)}, \end{aligned} \quad (\text{B.46})$$

$$\begin{aligned} \frac{\partial}{\partial \text{vech}^t(\Sigma)} \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)) &= 2s_n(\mathbf{B}, \Sigma) \rho_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)} \\ &\quad + s_n^2(\mathbf{B}, \Sigma) \psi_2(d_i^*(\mathbf{B}, \Sigma)) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\Sigma)}, \end{aligned} \quad (\text{B.47})$$

Demostración del Lema B.6.4. Usando la regla de la cadena y propiedades de derivadas se alcanza el resultado deseado. Estas derivadas se obtienen explícitamente usando el Lema A.5.2 y Lema A.5.3. □

Para conocer la convergencia de cada entrada de $(F_{1,n}(\mathbf{B}, \Sigma), F_{2,n}(\mathbf{B}, \Sigma))$ definida en (2.14) y que compone a \mathbf{V}_n^\dagger , se presentan los siguientes lemas

Lema B.6.5. *Sea $\{\mathbf{z}_i = (\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios que siguen el MLM (1) y sea $\{(\mathbf{B}_n, \Sigma_n)\}_{n \geq 1}$ una sucesión contenida en Θ que satisface $(\mathbf{B}_n, \Sigma_n) \xrightarrow{\text{c.s.}} (\mathbf{B}_0, \Sigma_0^*)$. Entonces*

$$\left. \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n(\mathbf{B}, \Sigma) \right|_{(\mathbf{B}_n, \Sigma_n)} \xrightarrow{\text{c.s.}} \mathbf{0}, \quad (\text{B.48})$$

$$\frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} s_n(\mathbf{B}, \boldsymbol{\Sigma}) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} - \frac{1}{2r^{3/2}} \frac{E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right]}{E[\psi_1(d_0^*) d_0]} \mathbf{D}_q. \quad (\text{B.49})$$

Con d_0 , d_0^* , y r definidos en (2.9), (2.10) y (1.23), respectivamente.

Demostración del Lema B.6.5. Por el Lema A.5.2,

$$\frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} = - \frac{s_n(\mathbf{B}, \boldsymbol{\Sigma}) \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s_n(\mathbf{B}, \boldsymbol{\Sigma})} \right) \left(\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})}{d_i(\mathbf{B}, \boldsymbol{\Sigma})} \right)}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s_n(\mathbf{B}, \boldsymbol{\Sigma})} \right) d_i(\mathbf{B}, \boldsymbol{\Sigma})}. \quad (\text{B.50})$$

Sean $f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma})) = s(\mathbf{B}, \boldsymbol{\Sigma}) \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) \left(\frac{\mathbf{x}^t \otimes (\mathbf{u}^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})}{d(\mathbf{B}, \boldsymbol{\Sigma})} \right)$ y $g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma})) = \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) d(\mathbf{B}, \boldsymbol{\Sigma})$, para $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ vector aleatorio en \mathbb{R}^{p+q} . Ambas funciones continuas, por ser composición de funciones continuas, además

$$E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \|f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))\| \right] < \infty.$$

Por los Lemas B.2.6 y B.2.9, resulta que

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_1(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \left(\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1})}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) \\ & \xrightarrow{c.s.} E \left[s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) \psi_1(d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) \left(\frac{\mathbf{x}^t \otimes (\mathbf{u}^t(\mathbf{B}_0) \boldsymbol{\Sigma}_0^{*-1})}{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \right) \right], \end{aligned}$$

donde $d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$ se define como $d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) := \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}$ y como se conoce, $d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) = d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) / s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)$.

Luego usando que $\boldsymbol{\Sigma}_0^* = r \boldsymbol{\Sigma}_0$ y

$$\begin{aligned} d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) &= \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \\ &= \frac{d(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{s(\mathbf{B}_0, \boldsymbol{\Sigma}_0)} && \text{por (1.23) y el Lema 1.2.3} \\ &= d_0^*, && \text{por (2.10)} \end{aligned} \quad (\text{B.51})$$

se tiene

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_1(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \left(\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1})}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) \xrightarrow{c.s.} E \left[\frac{k_0}{r} \psi_1(d_0^*) \left(\frac{\mathbf{x}^t \otimes (\mathbf{u}^t(\mathbf{B}_0) \boldsymbol{\Sigma}_0^{-1})}{d_0} \right) \right] \\ & = \mathbf{0} \text{ porque } \mathbf{u} \perp \mathbf{x} \text{ y } E[\mathbf{x}] = \mathbf{0}. \end{aligned} \quad (\text{B.52})$$

Por otro lado, para la función $g(\cdot)$, se tiene que

$$E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} |g(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))| \right] = E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \left| \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) d(\mathbf{B}, \boldsymbol{\Sigma}) \right| \right] < \infty,$$

y por los Lemas B.2.6 y B.2.8 y la ecuación (B.51), resulta que

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \psi_1(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) &\xrightarrow{c.s.} E[\psi_1(d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)) d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)] \\ &= E \left[\frac{1}{r^{1/2}} \psi_1(d_0^*) d_0 \right]. \end{aligned} \quad (\text{B.53})$$

Usando (B.52) y (B.53) en (B.50), se tiene

$$\left. \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} s_n(\mathbf{B}, \boldsymbol{\Sigma}) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} \mathbf{0}.$$

Para la otra derivada, recordar que por el Lema A.5.2, y por $\text{vec}^t(\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1}) = (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1} \otimes \mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})$ resulta,

$$\frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\boldsymbol{\Sigma})} = - \frac{s_n(\mathbf{B}, \boldsymbol{\Sigma}) \sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s_n(\mathbf{B}, \boldsymbol{\Sigma})} \right) \left(\frac{\text{vec}^t(\boldsymbol{\Sigma}^{-1} \mathbf{u}_i(\mathbf{B}) \mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})}{2d_i(\mathbf{B}, \boldsymbol{\Sigma})} \right) \mathbf{D}_q}{\sum_{i=1}^n \psi_1 \left(\frac{d_i(\mathbf{B}, \boldsymbol{\Sigma})}{s_n(\mathbf{B}, \boldsymbol{\Sigma})} \right) d_i(\mathbf{B}, \boldsymbol{\Sigma})}. \quad (\text{B.54})$$

Realizando un procedimiento similar al anterior, sea $f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma})) = s(\mathbf{B}, \boldsymbol{\Sigma}) \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) \times \left(\frac{\text{vec}^t(\mathbf{u}(\mathbf{B}) \mathbf{u}^t(\mathbf{B}))}{d(\mathbf{B}, \boldsymbol{\Sigma})} \right)$ una función continua, donde

$$E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} |f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))| \right] = E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \left| s(\mathbf{B}, \boldsymbol{\Sigma}) \psi_1 \left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})} \right) \left(\frac{\text{vec}^t(\mathbf{u}(\mathbf{B}) \mathbf{u}^t(\mathbf{B}))}{d(\mathbf{B}, \boldsymbol{\Sigma})} \right) \right| \right] < \infty,$$

y por los Lemas B.2.6 y B.2.8, resulta que

$$\frac{1}{n} \sum_{i=1}^n s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_1(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \left(\frac{\text{vec}^t(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n))}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right) \xrightarrow{c.s.} E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\mathbf{u}(\mathbf{B}_0) \mathbf{u}^t(\mathbf{B}_0)) \right]. \quad (\text{B.55})$$

Usando (B.53), (B.55), (B.51) y $\boldsymbol{\Sigma}_n^{-1} \xrightarrow{c.s.} \boldsymbol{\Sigma}_0^{*-1}$ en (B.54), se tiene

$$\begin{aligned} \left. \frac{\partial}{\partial \text{vec}^t(\boldsymbol{\Sigma})} s_n(\mathbf{B}, \boldsymbol{\Sigma}) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &\xrightarrow{c.s.} - \frac{1}{2} \frac{E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{*-1} \mathbf{u}(\mathbf{B}_0) \mathbf{u}^t(\mathbf{B}_0) \boldsymbol{\Sigma}_0^{*-1}) \right]}{E[\psi_1(d_0^*) d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)]} \mathbf{D}_q \\ &= - \frac{1}{2r^{3/2}} \frac{E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}(\mathbf{B}_0) \mathbf{u}^t(\mathbf{B}_0) \boldsymbol{\Sigma}_0^{-1}) \right]}{E[\psi_1(d_0^*) d_0]} \mathbf{D}_q \end{aligned}$$

□

Lema B.6.6. Sea $\{\mathbf{z}_i = (\mathbf{x}_i, \mathbf{y}_i)\}_{i \geq 1}$ una sucesión de vectores aleatorios que siguen el MLM (1) y sea $\{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\}_{n \geq 1}$ una sucesión contenida en Θ que satisface $(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)$, entonces para $k = 1, 2$ vale que

$$\left. \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} M_{\mathbf{B},\mathbf{B},k}$$

donde

$$M_{\mathbf{B},\mathbf{B},k} := -E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{u}\mathbf{x}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{x}^t) \right] - E \left[\frac{\psi_k(d_0^*)}{d_0^*} (\mathbf{x}\mathbf{x}^t \otimes \mathbf{I}_q) \right],$$

y

$$\left. \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} M_{\boldsymbol{\Sigma},\boldsymbol{\Sigma},k}$$

donde

$$\begin{aligned} M_{\boldsymbol{\Sigma},\boldsymbol{\Sigma},k} := & -\frac{q}{2r} \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{u}\mathbf{u}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\ & + \frac{1}{2r} \mathbf{L}_q E \left[\frac{[\psi'_k(d_0^*) d_0^* + \psi_k(d_0^*)]}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\ & + \frac{q}{2r E[\psi_1(d_0^*) d_0]} \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0} \text{vec}(\mathbf{u}\mathbf{u}^t) \right] E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\ & - \frac{E[d_0 \phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)]}{2r E[\psi_1(d_0^*) d_0]} \mathbf{L}_q E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\ & - \frac{k_0^2}{r} E[\psi_k(d_0^*) d_0^*] \mathbf{I}_{q(q+1)/2}. \end{aligned}$$

Además,

$$\left. \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} M_{\kappa_2,\boldsymbol{\Sigma},2}$$

donde,

$$\begin{aligned}
 M_{\kappa_2, \Sigma, 2} &:= -\frac{k_0 E[\rho_2(d_0^*)]}{r^2 E[\psi_1(d_0^*) d_0]} E\left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\Sigma_0^{-1} \mathbf{u} \mathbf{u}^t \Sigma_0^{-1})\right] \mathbf{D}_q \\
 &\quad -\frac{1}{2r^2} E\left[\frac{\psi_2(d_0^*)}{d_0^*} \text{vec}^t(\Sigma_0^{-1} \mathbf{u} \mathbf{u}^t \Sigma_0^{-1})\right] \mathbf{D}_q \\
 &\quad +\frac{E[\psi_2(d_0^*) d_0]}{2r^2 E[\psi_1(d_0^*) d_0]} E\left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\Sigma_0^{-1} \mathbf{u} \mathbf{u}^t \Sigma_0^{-1})\right] \mathbf{D}_q.
 \end{aligned}$$

Por último,

$$\begin{aligned}
 &\frac{\partial}{\partial \text{vech}^t(\Sigma)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} \xrightarrow{c.s.} \mathbf{0}, \\
 &\frac{\partial}{\partial \text{vech}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} \xrightarrow{c.s.} \mathbf{0}, \\
 &\frac{\partial}{\partial \text{vech}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)) \Big|_{(\mathbf{B}_n, \Sigma_n)} \xrightarrow{c.s.} \mathbf{0}.
 \end{aligned}$$

Donde ϕ_k está definido por: $\phi_k(\mathbf{B}_0, \Sigma_0^*) = \psi_k' \left[\frac{d(\mathbf{B}_0, \Sigma_0^*)}{s(\mathbf{B}_0, \Sigma_0^*)} \right] \frac{d(\mathbf{B}_0, \Sigma_0^*)}{s(\mathbf{B}_0, \Sigma_0^*)} - \psi_k \left[\frac{d(\mathbf{B}_0, \Sigma_0^*)}{s(\mathbf{B}_0, \Sigma_0^*)} \right]$, d_0 y d_0^* están definidos en (2.9) y (2.10), respectivamente.

Demostración del Lema B.6.6. Por el Lema B.6.1 y por la definición de $\phi_{i,k}$, ecuación (B.40) se tiene que,

$$\frac{\partial}{\partial \text{vech}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^{*2}(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \Sigma)}{\partial \text{vech}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} - \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q) \right\}, \quad (\text{B.56})$$

usando la ecuación (A.14) del Lema A.5.3 en la ecuación anterior, resulta

$$\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \left[-\frac{\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1})}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right. \right. \\ \left. \left. - \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] - \frac{\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q) \right\},$$

aplicando propiedades algebraicas se obtiene

$$\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} = \frac{1}{n} \sum_{i=1}^n \left\{ - \underbrace{\frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) (\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1}))}_{(1)} \right. \\ \left. - \underbrace{\frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t)}_{(2)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} - \underbrace{\frac{\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q)}_{(3)} \right\}. \quad (\text{B.57})$$

Para obtener la convergencia de (B.57), se analiza cada uno de los elementos que lo componen, es decir, se analizan los elementos (1), (2) y (3).

Para (1): Sea la función $f(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) = \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1} (\mathbf{x}_i^t \otimes \mathbf{I}_q)$, esta función es continua además para algún $\delta > 0$,

$$E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \|f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))\| \right] < \infty,$$

por los Lemas B.2.6 y B.2.8,

$$\frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1} (\mathbf{x}_i^t \otimes \mathbf{I}_q) \xrightarrow{c.s.} E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}{d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} \text{vec}(\mathbf{u} \mathbf{x}^t) \mathbf{u}^t \boldsymbol{\Sigma}_0^{*-1} (\mathbf{x}^t \otimes \mathbf{I}_q) \right], \quad (\text{B.58})$$

entonces se puede simplificar la expresión del límite de (B.58) utilizando que $\Sigma_0^* = r\Sigma_0$, $\mathbf{u}^t \Sigma_0^{*-1} (\mathbf{x}^t \otimes \mathbf{I}_q) = \text{vec}^t(\Sigma_0^{*-1} \mathbf{u} \mathbf{x}^t)$ y (B.51) obteniendo

$$E \left[\frac{\phi_k(\mathbf{B}_0, \Sigma_0^*)}{d^2(\mathbf{B}_0, \Sigma_0^*) d^*(\mathbf{B}_0, \Sigma_0^*)} \text{vec}(\mathbf{u} \mathbf{x}^t) \mathbf{u}^t \Sigma_0^{*-1} (\mathbf{x}^t \otimes \mathbf{I}_q) \right] = E \left[\frac{\phi_k(\mathbf{B}_0, \Sigma_0^*)}{d_0^2 d_0^*} \text{vec}(\mathbf{u} \mathbf{x}^t) \text{vec}^t(\Sigma_0^{-1} \mathbf{u} \mathbf{x}^t) \right],$$

así

$$\frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^2(\mathbf{B}_n, \Sigma_n) d_i^*(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1} (\mathbf{x}_i^t \otimes \mathbf{I}_q) \xrightarrow{c.s.} E \left[\frac{\phi_k(\mathbf{B}_0, \Sigma_0^*)}{d_0^2 d_0^*} \text{vec}(\mathbf{u} \mathbf{x}^t) \text{vec}^t(\Sigma_0^{-1} \mathbf{u} \mathbf{x}^t) \right]. \quad (\text{B.59})$$

Para (2): se realiza un procedimiento similar al anterior, sea la función $f(\mathbf{z}_i, \mathbf{B}_n, \Sigma_n, s_n(\mathbf{B}_n, \Sigma_n)) = \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t)$, esta función es continua, y

$$E \left[\sup_{\|(\mathbf{B}, \Sigma) - (\mathbf{B}_0, \Sigma_0^*)\| < \delta} \|f(\mathbf{z}, \mathbf{B}, \Sigma, s(\mathbf{B}, \Sigma))\| \right] = E \left[\sup_{\|(\mathbf{B}, \Sigma) - (\mathbf{B}_0, \Sigma_0^*)\| < \delta} \left| \frac{\phi_k(\mathbf{B}, \Sigma)}{d(\mathbf{B}, \Sigma)} \text{vec}(\mathbf{u}(\mathbf{B}) \mathbf{x}^t) \right| \right] < \infty$$

por los Lemas B.2.6 y B.2.8, resulta que

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) &\xrightarrow{c.s.} E \left[\frac{\phi_k(\mathbf{B}_0, \Sigma_0)}{d(\mathbf{B}_0, \Sigma_0^*)} \text{vec}(\mathbf{u} \mathbf{x}^t) \right] \\ &= \mathbf{0} \text{ porque } \mathbf{u} \perp \mathbf{x} \text{ y } E[\mathbf{x}] = \mathbf{0}. \end{aligned} \quad (\text{B.60})$$

Por último, para (3): Sea la función $f(\mathbf{z}_i, \mathbf{B}_n, \Sigma_n, s_n(\mathbf{B}_n, \Sigma_n)) = -\frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q)$, se observa que es una función continua al ser composición de funciones continuas, además para algún $\delta > 0$,

$$E \left[\sup_{\|(\mathbf{B}, \Sigma) - (\mathbf{B}_0, \Sigma_0^*)\| < \delta} \|f(\mathbf{z}, \mathbf{B}, \Sigma, s(\mathbf{B}, \Sigma))\| \right] = E \left[\sup_{\|(\mathbf{B}, \Sigma) - (\mathbf{B}_0, \Sigma_0^*)\| < \delta} \left| \frac{\psi_k\left(\frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)}\right)}{\frac{d(\mathbf{B}, \Sigma)}{s(\mathbf{B}, \Sigma)}} \right| \|(\mathbf{x} \mathbf{x}^t \otimes \mathbf{I}_q)\| \right],$$

se puede asegurar que esta esperanza es finita, puesto que $\frac{\psi_k\left(\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})}\right)}{\frac{d(\mathbf{B}, \boldsymbol{\Sigma})}{s(\mathbf{B}, \boldsymbol{\Sigma})}}$, es continua al ser composición de funciones continuas, y porque $\boldsymbol{\Sigma}_{\mathbf{x}}$ es finita al usar la condición A9), usando los Lemas B.2.6 y B.2.8 se obtiene la siguiente convergencia

$$-\frac{1}{n} \sum_{i=1}^n \frac{\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q) \xrightarrow{c.s.} -E \left[\frac{\psi_k(d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))}{d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)} (\mathbf{xx}^t \otimes \mathbf{I}_q) \right],$$

en consecuencia, por (B.51)

$$-\frac{1}{n} \sum_{i=1}^n \frac{\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i \mathbf{x}_i^t \otimes \mathbf{I}_q) \xrightarrow{c.s.} -E \left[\frac{\psi_k(d_0^*)}{d_0^*} (\mathbf{xx}^t \otimes \mathbf{I}_q) \right]. \quad (\text{B.61})$$

Entonces usando (B.61), (B.59), (B.60) y (B.48), resulta que (B.57) converge a

$$\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} -E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{ux}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{ux}^t) \right] - E \left[\frac{\psi_k(d_0^*)}{d_0^*} (\mathbf{xx}^t \otimes \mathbf{I}_q) \right].$$

Realizando el mismo procedimiento anterior para la derivada $\frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}$, usando el Lema B.6.1 y la ecuación (A.15) del Lema A.5.3,

$$\begin{aligned} \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\ &= \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \left\{ -\frac{1}{2d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}^t(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) (\boldsymbol{\Sigma}_n^{-1} \otimes \boldsymbol{\Sigma}_n^{-1}) \mathbf{D}_q \right. \\ &\quad \left. - \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right\}, \end{aligned}$$

donde $\text{vec}^t(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) = \mathbf{u}_i^t(\mathbf{B}_n) \otimes \mathbf{u}_i^t(\mathbf{B}_n)$, usando propiedades algebraicas, se obtiene

$$\left. \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} = - \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t) \left\{ \frac{1}{2d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}^t(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n))(\boldsymbol{\Sigma}_n^{-1} \otimes \boldsymbol{\Sigma}_n^{-1})\mathbf{D}_q + \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \right\} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}, \quad (\text{B.62})$$

que se puede segmentar de la siguiente manera,

$$\left. \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} = \underbrace{- \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{2d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t) \text{vec}^t(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n))(\boldsymbol{\Sigma}_n^{-1} \otimes \boldsymbol{\Sigma}_n^{-1})\mathbf{D}_q}_{(1)} - \underbrace{\frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t)}_{(2)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}.$$

A continuación, se analizan las convergencias del promedio de los elementos (1) y (2):

Sea $f(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) = - \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{2d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t) \text{vec}^t(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n))$, esta función es continua, y

$$E \left[\sup_{\|(\mathbf{B}, \boldsymbol{\Sigma}) - (\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\| < \delta} \|f(\mathbf{z}, \mathbf{B}, \boldsymbol{\Sigma}, s(\mathbf{B}, \boldsymbol{\Sigma}))\| \right] < \infty$$

por los Lemas B.2.6, B.2.8 y por (B.51) se tiene la siguiente convergencia

$$\frac{1}{n} \sum_{i=1}^n f(\mathbf{z}_i, \mathbf{B}_n, \boldsymbol{\Sigma}_n, s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} -E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{2d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)d_0^*} \text{vec}(\mathbf{u}\mathbf{x}^t) \text{vec}^t(\mathbf{u}\mathbf{u}^t) \right],$$

además como $\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{2d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)d_0^*}$ depende solamente de \mathbf{u} , resulta,

$$E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{2d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)d_0^*} \text{vec}(\mathbf{u}\mathbf{x}^t) \text{vec}^t(\mathbf{u}\mathbf{u}^t) \right] = \mathbf{0} \quad \text{porque } \mathbf{u} \perp \mathbf{x} \text{ y } E[\mathbf{x}] = \mathbf{0}.$$

En consecuencia,

$$-\frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{2d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t) \text{vec}^t(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \xrightarrow{c.s.} \mathbf{0}. \quad (\text{B.63})$$

La convergencia a cero del promedio de (1) es consecuencia de (B.63) y de la hipótesis que implica que $(\boldsymbol{\Sigma}_n^{-1} \otimes \boldsymbol{\Sigma}_n^{-1})\mathbf{D}_q$ converge c.s. a $((\boldsymbol{\Sigma}_0^*)^{-1} \otimes (\boldsymbol{\Sigma}_0^*)^{-1})\mathbf{D}_q$.

Cabe notar que el elemento (2) ya fue analizado en la derivada anterior, en la ecuación (B.60), entonces utilizando (B.49), (B.60), y (B.63) en (B.62) se obtiene,

$$\frac{\partial}{\partial \text{vec}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\mathbf{B}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} \mathbf{0}.$$

Se ha obtenido hasta el momento la convergencia de dos derivadas, para las restantes el procedimiento es similar, se reescribirán las derivadas de manera que dependan de la derivada de $s_n(\mathbf{B}, \boldsymbol{\Sigma})$ que corresponda. Para $\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}$ se usa los resultados de los Lemas B.6.2 y B.6.3,

$$\begin{aligned} \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vec}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= \mathbf{L}_q \left\{ -q \frac{\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q})(\mathbf{u}_i(\mathbf{B}_n)\mathbf{x}_i^t \otimes \mathbf{I}_q) \right. \\ &\quad - 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \text{vec}(\boldsymbol{\Sigma}_n) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\ &\quad + \left(q \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \right. \\ &\quad - \left[s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k'(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \right. \\ &\quad \left. \left. + s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \right] \text{vec}(\boldsymbol{\Sigma}_n) \right) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \left. \right\}, \end{aligned}$$

luego usando la derivada (A.14), y propiedad distributiva resulta,

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \mathbf{L}_q \left\{ -q \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q) \right. \\
 &\quad - 2d_i(\mathbf{B}_n, \Sigma_n) \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \text{vec}(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^*(\mathbf{B}_n, \Sigma_n) d_i^2(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) [\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1})] \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &\quad + \psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \text{vec}(\Sigma_n) [\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1})] \\
 &\quad + \psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i^*(\mathbf{B}_n, \Sigma_n) d_i(\mathbf{B}_n, \Sigma_n) \text{vec}(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &\quad + \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} \text{vec}(\Sigma_n) [\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1})] \\
 &\quad \left. + d_i(\mathbf{B}_n, \Sigma_n) \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \text{vec}(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\},
 \end{aligned}$$

extrayendo factor común

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \mathbf{L}_q \left\{ -q \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q) \right. \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^*(\mathbf{B}_n, \Sigma_n) d_i^2(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \left[\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1}) \right] \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &\quad + \left[\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) + \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} \right] \text{vec}(\Sigma_n) \left[\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1}) \right] \\
 &\quad + [\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i^*(\mathbf{B}_n, \Sigma_n) d_i(\mathbf{B}_n, \Sigma_n) + d_i(\mathbf{B}_n, \Sigma_n) \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \\
 &\quad \left. - 2d_i(\mathbf{B}_n, \Sigma_n) \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))] \text{vec}(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\},
 \end{aligned}$$

usando la definición de $\phi_{i,k}$ dada en la ecuación (B.40), la igualdad anterior se reduce a

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \mathbf{L}_q \left\{ -q \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q) \right. \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^*(\mathbf{B}_n, \Sigma_n) d_i^2(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \left[\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1}) \right] \\
 &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &\quad + \frac{\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i^*(\mathbf{B}_n, \Sigma_n) + \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} \text{vec}(\Sigma_n) \left[\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1}) \right] \\
 &\quad \left. + d_i(\mathbf{B}_n, \Sigma_n) \phi_{i,k}(\mathbf{B}_n, \Sigma_n) \text{vec}(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\}.
 \end{aligned}$$

Como $vec^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) = \mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}_n) \Sigma_n^{-1})$, entonces

$$\begin{aligned} \frac{\partial}{\partial vec^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \frac{1}{n} \sum_{i=1}^n \mathbf{L}_q \left\{ -q \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q) \right. \\ &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^*(\mathbf{B}_n, \Sigma_n) d_i^2(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) vec^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \\ &\quad - q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\ &\quad + \frac{\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i^*(\mathbf{B}_n, \Sigma_n) + \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} vec(\Sigma_n) vec^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \\ &\quad \left. + d_i(\mathbf{B}_n, \Sigma_n) \phi_{i,k}(\mathbf{B}_n, \Sigma_n) vec(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\}. \end{aligned}$$

A continuación, se desea analizar cada uno de los elementos que componen la derivada, para facilitar su tratamiento, se enumeran los

términos involucrados,

$$\begin{aligned}
 \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vech}(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \frac{1}{n} \sum_{i=1}^n \mathbf{L}_q \left\{ \underbrace{-q \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q)}_{(1)} \right. \\
 &\quad - \underbrace{q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^*(\mathbf{B}_n, \Sigma_n) d_i^2(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \text{vec}^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t)}_{(2)} \\
 &\quad - \underbrace{q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)}}_{(3)} \\
 &\quad + \underbrace{\frac{[\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) + \psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i^*(\mathbf{B}_n, \Sigma_n)]}{d_i^*(\mathbf{B}_n, \Sigma_n)} \text{vec}(\Sigma_n) \text{vec}^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t)}_{(4)} \\
 &\quad \left. + \underbrace{d_i(\mathbf{B}_n, \Sigma_n) \phi_{i,k}(\mathbf{B}_n, \Sigma_n) \text{vec}(\Sigma_n)}_{(5)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\}. \tag{B.64}
 \end{aligned}$$

Se analiza la convergencia del promedio de cada cada término por separado, que serán consecuencia de los Lemas B.2.6 y B.2.8. Para (1) nótese que

$$-\frac{q}{n} \mathbf{L}_q \sum_{i=1}^n \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t \otimes \mathbf{I}_q) \xrightarrow{c.s.} -q \mathbf{L}_q E \left[\frac{\psi_k(d_0^*)}{d_0^*} (\mathbf{I}_{q^2} + \mathbf{K}_{q,q}) (\mathbf{u} \mathbf{x}^t \otimes \mathbf{I}_q) \right] = \mathbf{0},$$

ya que $\mathbf{u} \perp \mathbf{x}$ y $E[\mathbf{x}] = \mathbf{0}$. Después, la convergencia para el promedio de (2) se expresa como

$$\begin{aligned}
 -\frac{1}{n} \mathbf{L}_q \sum_{i=1}^n q \frac{\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i^2(\mathbf{B}_n, \Sigma_n) d_i^*(\mathbf{B}_n, \Sigma_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \text{vec}^t(\Sigma_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) &\xrightarrow{c.s.} -q \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \Sigma_0)}{d^2(\mathbf{B}_0, \Sigma_0^*) d^*(\mathbf{B}_0, \Sigma_0^*)} \text{vec}(\mathbf{u} \mathbf{u}^t) \text{vec}^t(\Sigma_0^{*-1} \mathbf{u} \mathbf{x}^t) \right] \\
 &= \mathbf{0},
 \end{aligned}$$

la igualdad de la convergencia anterior se debe a que el factor $\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)}$ depende de \mathbf{u} y no de \mathbf{x} , $\mathbf{u} \perp \perp \mathbf{x}$ y $E[\mathbf{x}] = \mathbf{0}$.

Para (3) se tiene,

$$-\frac{1}{n} \mathbf{L}_q \sum_{i=1}^n q \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \xrightarrow{c.s.} -qr^{1/2} \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0} \text{vec}(\mathbf{u} \mathbf{u}^t) \right],$$

luego, para (4)

$$\begin{aligned} & \mathbf{L}_q \text{vec}(\boldsymbol{\Sigma}_n) \frac{1}{n} \sum_{i=1}^n \frac{[\psi'_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) + \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))]}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}^t(\boldsymbol{\Sigma}_n^{-1} \mathbf{u}_i(\mathbf{B}_n) \mathbf{x}_i^t) \\ & \xrightarrow{c.s.} \mathbf{L}_q \text{vec}(\boldsymbol{\Sigma}_0^*) E \left[\frac{[\psi'_k(d_0^*) d_0^* + \psi_k(d_0^*)]}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{*-1} \mathbf{u}(\mathbf{B}_0) \mathbf{x}^t) \right] \\ & = \mathbf{0} \quad \text{porque } \mathbf{u} \perp \perp \mathbf{x} \text{ y } \mathbf{x} \text{ es centrado,} \end{aligned}$$

y para el elemento (5),

$$\begin{aligned} & \mathbf{L}_q \text{vec}(\boldsymbol{\Sigma}_n) \frac{1}{n} \sum_{i=1}^n d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \xrightarrow{c.s.} \mathbf{L}_q E [\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0) d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)] \text{vec}(\boldsymbol{\Sigma}_0^*) \\ & = r^{1/2} \mathbf{L}_q E [\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0) d_0] \text{vec}(\boldsymbol{\Sigma}_0). \end{aligned}$$

Usando las convergencias obtenidas para (1)-(5), y (B.48) en (B.64) resulta,

$$\left. \frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vech} [\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \right|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} \mathbf{0}.$$

Para la otra derivada utilizando la regla de la cadena y el Lema B.6.2, se obtiene

$$\begin{aligned} \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} [\text{vech}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= \mathbf{L}_q \left[\left\{ q \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \right. \right. \\ &\quad - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi'_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \text{vec}(\boldsymbol{\Sigma}_n) \\ &\quad \left. \left. - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \text{vec}(\boldsymbol{\Sigma}_n) \right\} \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right. \\ &\quad \left. - 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \text{vec}(\boldsymbol{\Sigma}_n) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] \\ &\quad - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \mathbf{I}_{q(q+1)/2}, \end{aligned}$$

al sustituir $\frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}$ por su expresión dada en (A.15), se obtiene

$$\begin{aligned} \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} [\text{vech}(\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma})))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= \mathbf{L}_q \left[\left\{ q \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i^{*2}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n) \mathbf{u}_i^t(\mathbf{B}_n)) \right. \right. \\ &\quad - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi'_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \text{vec}(\boldsymbol{\Sigma}_n) \\ &\quad \left. \left. - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \text{vec}(\boldsymbol{\Sigma}_n) \right\} \left[- \frac{(\mathbf{u}_i^t(\mathbf{B}_n) \otimes \mathbf{u}_i^t(\mathbf{B}_n)) (\boldsymbol{\Sigma}_n^{-1} \otimes \boldsymbol{\Sigma}_n^{-1})}{2d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \mathbf{D}_q \right. \right. \\ &\quad \left. \left. - \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] \right. \\ &\quad \left. - 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \text{vec}(\boldsymbol{\Sigma}_n) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right] \\ &\quad - s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \mathbf{I}_{q(q+1)/2}. \end{aligned}$$

Usando que $vec^t(\Sigma^{-1}\mathbf{u}_i\mathbf{u}_i^t\Sigma^{-1}) = (\mathbf{u}_i^t \otimes \mathbf{u}_i^t)(\Sigma^{-1} \otimes \Sigma^{-1})$, aplicando propiedad distributiva y extrayendo factor común resulta,

$$\begin{aligned}
 \frac{\partial}{\partial vec^t(\Sigma)} [vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma)))] \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \mathbf{L}_q \left[-\frac{q\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{2d_i^*(\mathbf{B}_n, \Sigma_n)d_i^2(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n))vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})\mathbf{D}_q \right. \\
 &+ \frac{1}{2} \left\{ \psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) + \frac{\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))}{d_i^*(\mathbf{B}_n, \Sigma_n)} \right\} vec(\Sigma_n)vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})\mathbf{D}_q \\
 &- \frac{q\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &+ \left\{ d_i(\mathbf{B}_n, \Sigma_n)d_i^*(\mathbf{B}_n, \Sigma_n)\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n)) + d_i(\mathbf{B}_n, \Sigma_n)\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \right. \\
 &\left. - 2s_n(\mathbf{B}_n, \Sigma_n)d_i^*(\mathbf{B}_n, \Sigma_n)\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n)) \right\} vec(\Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\
 &- s_n^2(\mathbf{B}_n, \Sigma_n)\psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))d_i^*(\mathbf{B}_n, \Sigma_n)\mathbf{I}_{q(q+1)/2},
 \end{aligned}$$

usando la definición de $\phi_{i,k}$, ver la ecuación (B.40), y segmentando el promedio de la derivada resulta,

$$\begin{aligned}
 \frac{\partial}{\partial vec^t(\Sigma)} \frac{1}{n} \sum_{i=1}^n vech(\Psi_{i,k,\Sigma}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))) \Big|_{(\mathbf{B}_n, \Sigma_n)} &= \frac{1}{n} \sum_{i=1}^n \mathbf{L}_q \left\{ \left[\underbrace{-\frac{q\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{2d_i^2(\mathbf{B}_n, \Sigma_n)d_i^*(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n))vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})}_{(1)} \right. \right. \\
 &+ \left. \underbrace{\frac{[\psi'_k(d_i^*(\mathbf{B}_n, \Sigma_n))d_i^*(\mathbf{B}_n, \Sigma_n) + \psi_k(d_i^*(\mathbf{B}_n, \Sigma_n))]}{2d_i^*(\mathbf{B}_n, \Sigma_n)} vec(\Sigma_n)vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})}_{(2)} \right] \mathbf{D}_q \\
 &- \underbrace{\frac{q\phi_{i,k}(\mathbf{B}_n, \Sigma_n)}{d_i(\mathbf{B}_n, \Sigma_n)} vec(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)}}_{(3)} \Big|_{(\mathbf{B}_n, \Sigma_n)}
 \end{aligned}$$

$$\left. \begin{aligned} & \underbrace{+d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\text{vec}(\boldsymbol{\Sigma}_n)}_{(4)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\ & \underbrace{-\frac{1}{n} \sum_{i=1}^n s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}_{(5)} \mathbf{I}_{q(q+1)/2}. \end{aligned} \right\}$$

Con los Lemas B.2.6 y B.2.8 se estudia la convergencia c.s. de la derivada anterior. Para ello se analiza la convergencia del promedio de (1),

$$\begin{aligned} -\frac{1}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{2d_i^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \text{vec}^t(\boldsymbol{\Sigma}_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\boldsymbol{\Sigma}_n^{-1}) &\xrightarrow{c.s.} -\frac{1}{2}E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)d_0^*} \text{vec}(\mathbf{u}\mathbf{u}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{*-1}\mathbf{u}\mathbf{u}^t\boldsymbol{\Sigma}_0^{*-1}) \right] \\ &= -\frac{1}{2r}E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{u}\mathbf{u}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1}\mathbf{u}\mathbf{u}^t\boldsymbol{\Sigma}_0^{-1}) \right], \end{aligned}$$

la última igualdad se obtuvo al utilizar que $\boldsymbol{\Sigma}_0^* = r\boldsymbol{\Sigma}_0$. Luego, para el término (2) se obtiene la siguiente convergencia,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \frac{[\psi_k'(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) + \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))] \text{vec}(\boldsymbol{\Sigma}_n) \text{vec}^t(\boldsymbol{\Sigma}_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\boldsymbol{\Sigma}_n^{-1})}{2d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\ & \xrightarrow{c.s.} \frac{1}{2}E \left[\frac{(\psi_k'(d_0^*)d_0^* + \psi_k(d_0^*)) \text{vec}(\boldsymbol{\Sigma}_0^*) \text{vec}^t(\boldsymbol{\Sigma}_0^{*-1}\mathbf{u}\mathbf{u}^t\boldsymbol{\Sigma}_0^{*-1})}{d_0^*} \right] \\ & = \frac{1}{2r}E \left[\frac{(\psi_k'(d_0^*)d_0^* + \psi_k(d_0^*)) \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1}\mathbf{u}\mathbf{u}^t\boldsymbol{\Sigma}_0^{-1})}{d_0^*} \right]. \end{aligned}$$

Para (3), usando nuevamente que $\boldsymbol{\Sigma}_0^* = r\boldsymbol{\Sigma}_0$ se obtiene

$$-\frac{q}{n} \sum_{i=1}^n \frac{\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \text{vec}(\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)) \xrightarrow{c.s.} -qr^{1/2}E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0} \text{vec}(\mathbf{u}\mathbf{u}^t) \right],$$

de igual manera para (4) se tiene,

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\phi_{i,k}(\mathbf{B}_n, \boldsymbol{\Sigma}_n)\text{vec}(\boldsymbol{\Sigma}_n) \xrightarrow{c.s.} E[d(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*)\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)]\text{vec}(\boldsymbol{\Sigma}_0^*) \\ & = r^{1/2}E[d_0\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)]\text{vec}(\boldsymbol{\Sigma}_0), \end{aligned}$$

y para (5) resulta

$$\begin{aligned} -\frac{1}{n} \sum_{i=1}^n s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_k(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n) &\xrightarrow{c.s.} -s^2(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) E[\psi_k(d_0^*) d_0^*] \\ &= -\frac{k_0^2}{r} E[\psi_k(d_0^*) d_0^*], \quad \text{por (1.25), } \boldsymbol{\Sigma}_0^* = r\boldsymbol{\Sigma}_0 \text{ y por el Lema 1.2.3.} \end{aligned}$$

Entonces, usando todos estos límites c.s., incluida la convergencia dada en (B.49), es decir

$$\frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} s_n(\mathbf{B}, \boldsymbol{\Sigma}) \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} -\frac{1}{2r^{3/2}} \frac{E\left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1})\right]}{E[\psi_1(d_0^*) d_0]} \mathbf{D}_q$$

usando $\boldsymbol{\Sigma}_n^{-1} \xrightarrow{c.s.} \boldsymbol{\Sigma}_0^{*-1}$ y extrayendo factor común se obtiene

$$\begin{aligned} \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \text{vech}[\Psi_{i,k,\boldsymbol{\Sigma}}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &\xrightarrow{c.s.} \frac{1}{2r} \mathbf{L}_q \left\{ \begin{aligned} &-qE\left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{u}\mathbf{u}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1})\right] \\ &+ E\left[\frac{[\psi'_k(d_0^*) d_0^* + \psi_k(d_0^*)]}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1})\right] \\ &+ \frac{q}{E[\psi_1(d_0^*) d_0]} E\left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0} \text{vec}(\mathbf{u}\mathbf{u}^t)\right] E\left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1})\right] \\ &- \frac{E[d_0 \phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)]}{E[\psi_1(d_0^*) d_0]} E\left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1})\right] \end{aligned} \right\} \mathbf{D}_q \\ &- \frac{k_0^2}{r} E[\psi_k(d_0^*) d_0^*] \mathbf{I}_{q(q+1)/2} \end{aligned}$$

$$\begin{aligned}
&= -\frac{q}{2r} \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0^2 d_0^*} \text{vec}(\mathbf{u}\mathbf{u}^t) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
&+ \frac{1}{2r} \mathbf{L}_q E \left[\frac{[\psi'_k(d_0^*) d_0^* + \psi_k(d_0^*)]}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
&+ \frac{q}{2r E[\psi_1(d_0^*) d_0]} \mathbf{L}_q E \left[\frac{\phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)}{d_0} \text{vec}(\mathbf{u}\mathbf{u}^t) \right] E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
&- \frac{E[d_0 \phi_k(\mathbf{B}_0, \boldsymbol{\Sigma}_0)]}{2r E[\psi_1(d_0^*) d_0]} \mathbf{L}_q E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}(\boldsymbol{\Sigma}_0) \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u}\mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
&- \frac{k_0^2}{r} E[\psi_k(d_0^*) d_0^*] \mathbf{I}_{q(q+1)/2}.
\end{aligned}$$

Luego, para $\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vech} [\Psi_{i,k,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}$ se usan el Lema B.6.4 y la ecuación (A.14) del Lema A.5.3,

$$\begin{aligned}
\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \text{vech} [\Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \rho_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\
&+ s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\
&= 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \rho_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\
&+ s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \left[-\frac{1}{s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i^t \otimes (\mathbf{u}_i^t(\mathbf{B}) \boldsymbol{\Sigma}^{-1})) \right. \\
&\left. - \frac{d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}{s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \right]
\end{aligned}$$

$$\begin{aligned}
 &= 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \rho_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\
 &\quad - \frac{\psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i^t \otimes \mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1}) \\
 &\quad - \psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vec}^t(\mathbf{B}^t)} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}.
 \end{aligned}$$

Usando convergencias de los Lemas B.2.6, B.2.8 se tiene que

$$\begin{aligned}
 &2 \frac{1}{n} \sum_{i=1}^n s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \rho_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} 2E[s(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) \rho_2(d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))] \\
 &= \frac{2k_0}{r^{1/2}} E[\rho_2(d_0^*)] \quad \text{porque } \boldsymbol{\Sigma}_0^* = r\boldsymbol{\Sigma}_0 \text{ y por el Lema 1.2.3}
 \end{aligned}$$

y

$$\begin{aligned}
 &-\frac{1}{n} \sum_{i=1}^n \frac{\psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n))}{d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} (\mathbf{x}_i^t \otimes \mathbf{u}_i^t(\mathbf{B}_n) \boldsymbol{\Sigma}_n^{-1}) \xrightarrow{c.s.} -\frac{1}{r} E \left[\frac{\psi_2(d_0^*)}{d_0^*} (\mathbf{x}^t \otimes \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] = \mathbf{0} \quad \text{porque } \mathbf{u} \perp \mathbf{x} \\
 &-\frac{1}{n} \sum_{i=1}^n d_i(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \xrightarrow{c.s.} -E[d_0^* \psi_2(d^*(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*))].
 \end{aligned}$$

Utilizando las convergencias anteriores y la convergencia presente en el Lema B.6.5 resulta

$$\frac{\partial}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{1}{n} \sum_{i=1}^n \text{vech} [\Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \xrightarrow{c.s.} \mathbf{0}.$$

Por último, para obtener la convergencia de $\frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \text{vech} [\Psi_{i,2,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}$ se utiliza el Lema B.6.4 dada por

$$\begin{aligned}
 \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \text{vech} [\Psi_{i,k,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &= 2s_n(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \rho_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial s_n(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} \\
 &\quad + s_n^2(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \psi_2(d_i^*(\mathbf{B}_n, \boldsymbol{\Sigma}_n)) \frac{\partial d_i^*(\mathbf{B}, \boldsymbol{\Sigma})}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)}
 \end{aligned}$$

aplicando (A.15) y $vec^t(\Sigma^{-1}\mathbf{u}_i\mathbf{u}_i^t\Sigma^{-1}) = (\mathbf{u}_i^t \otimes \mathbf{u}_i^t)(\Sigma^{-1} \otimes \Sigma^{-1})$ en la igualdad anterior, se obtiene

$$\begin{aligned} \frac{\partial}{\partial vec^t(\Sigma)} vech [\Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))] \Big|_{(\mathbf{B}_n, \Sigma_n)} &= 2s_n(\mathbf{B}_n, \Sigma_n) \rho_2(d_i^*(\mathbf{B}_n, \Sigma_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\ &+ s_n^2(\mathbf{B}_n, \Sigma_n) \psi_2(d_i^*(\mathbf{B}_n, \Sigma_n)) \left\{ -\frac{vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})\mathbf{D}_q}{2d_i(\mathbf{B}_n, \Sigma_n)s_n(\mathbf{B}_n, \Sigma_n)} \right. \\ &\left. - \frac{d_i(\mathbf{B}_n, \Sigma_n)}{s_n^2(\mathbf{B}_n, \Sigma_n)} \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \right\} \end{aligned}$$

al aplicar la propiedad distributiva se llega a

$$\begin{aligned} \frac{\partial}{\partial vec^t(\Sigma)} vech [\Psi_{i,2,\kappa_2}(\mathbf{B}, \Sigma, s_n(\mathbf{B}, \Sigma))] \Big|_{(\mathbf{B}_n, \Sigma_n)} &= 2s_n(\mathbf{B}_n, \Sigma_n) \rho_2(d_i^*(\mathbf{B}_n, \Sigma_n)) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)} \\ &- \frac{\psi_2(d_i^*(\mathbf{B}_n, \Sigma_n))}{2d_i^*(\mathbf{B}_n, \Sigma_n)} vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})\mathbf{D}_q \\ &- \psi_2(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i(\mathbf{B}_n, \Sigma_n) \frac{\partial s_n(\mathbf{B}, \Sigma)}{\partial vec^t(\Sigma)} \Big|_{(\mathbf{B}_n, \Sigma_n)}, \end{aligned}$$

usando los Lemas B.2.6, B.2.8 resulta

$$\begin{aligned} &\frac{1}{n} \sum_{i=1}^n 2s_n(\mathbf{B}_n, \Sigma_n) \rho_2(d_i^*(\mathbf{B}_n, \Sigma_n)) \xrightarrow{c.s.} \frac{2k_0}{r^{1/2}} E[\rho_2(d_0^*)] \\ &- \frac{1}{n} \sum_{i=1}^n \frac{\psi_2(d_i^*(\mathbf{B}_n, \Sigma_n))}{2d_i^*(\mathbf{B}_n, \Sigma_n)} vec^t(\Sigma_n^{-1}\mathbf{u}_i(\mathbf{B}_n)\mathbf{u}_i^t(\mathbf{B}_n)\Sigma_n^{-1})\mathbf{D}_q \xrightarrow{c.s.} -\frac{1}{2r^2} E\left[\frac{\psi_2(d_0^*)}{d_0^*} vec^t(\Sigma_0^{-1}\mathbf{u}\mathbf{u}^t\Sigma_0^{-1})\mathbf{D}_q\right] \\ &- \frac{1}{n} \sum_{i=1}^n \psi_2(d_i^*(\mathbf{B}_n, \Sigma_n)) d_i(\mathbf{B}_n, \Sigma_n) \xrightarrow{c.s.} -\frac{1}{r^{1/2}} E[\psi_2(d_0^*)d_0]. \end{aligned}$$

Y luego al usar B.6.5 (ecuación (B.49)) resulta la siguiente convergencia

$$\begin{aligned}
 \frac{\partial}{\partial \text{vech}^t(\boldsymbol{\Sigma})} \frac{1}{n} \sum_{i=1}^n \text{vech} [\Psi_{i,k,\kappa_2}(\mathbf{B}, \boldsymbol{\Sigma}, s_n(\mathbf{B}, \boldsymbol{\Sigma}))] \Big|_{(\mathbf{B}_n, \boldsymbol{\Sigma}_n)} &\xrightarrow{c.s.} -\frac{k_0}{r^2} E[\rho_2(d_0^*)] \frac{E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right]}{E[\psi_1(d_0^*) d_0]} \mathbf{D}_q \\
 &- \frac{1}{2r^2} E \left[\frac{\psi_2(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
 &+ \frac{1}{2r^2} E[\psi_2(d_0^*) d_0] \frac{E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right]}{E[\psi_1(d_0^*) d_0]} \mathbf{D}_q \\
 &= -\frac{k_0 E[\rho_2(d_0^*)]}{r^2 E[\psi_1(d_0^*) d_0]} E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
 &- \frac{1}{2r^2} E \left[\frac{\psi_2(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q \\
 &+ \frac{E[\psi_2(d_0^*) d_0]}{2r^2 E[\psi_1(d_0^*) d_0]} E \left[\frac{\psi_1(d_0^*)}{d_0^*} \text{vec}^t(\boldsymbol{\Sigma}_0^{-1} \mathbf{u} \mathbf{u}^t \boldsymbol{\Sigma}_0^{-1}) \right] \mathbf{D}_q.
 \end{aligned}$$

□

Demostración del Teorema 8. Usando el Lema 2.1.3 y Lema B.6.6 el resultado es directo. En efecto, empleando la notación presente en (B.37), y usando los resultados de Lema B.6.6 y del Teorema 5, resulta

$$\begin{aligned}
 \frac{1}{n} f_{11(n)} &\xrightarrow{c.s.} f_{11}^* && \text{con } f_{11}^* := C(\mathbf{B}_0, \boldsymbol{\Sigma}_0) M_{\mathbf{B}, \mathbf{B}, 1} + D(\mathbf{B}_0, \boldsymbol{\Sigma}_0) M_{\mathbf{B}, \mathbf{B}, 2}, \\
 \frac{1}{n} f_{12(n)} &\xrightarrow{c.s.} f_{12}^* && \text{con } f_{12}^* := \mathbf{0}, \\
 \frac{1}{n} f_{21(n)} &\xrightarrow{c.s.} f_{21}^* && \text{con } f_{21}^* := \mathbf{0}, \\
 \frac{1}{n} f_{22(n)} &\xrightarrow{c.s.} f_{22}^* && \text{con } f_{22}^* := C(\mathbf{B}_0, \boldsymbol{\Sigma}_0) M_{\boldsymbol{\Sigma}, \boldsymbol{\Sigma}, 1} + D(\mathbf{B}_0, \boldsymbol{\Sigma}_0) M_{\boldsymbol{\Sigma}, \boldsymbol{\Sigma}, 2}, \\
 \frac{1}{n} f_{31(n)} &\xrightarrow{c.s.} f_{31}^* && \text{con } f_{31}^* := \mathbf{0}, \\
 \frac{1}{n} f_{32(n)} &\xrightarrow{c.s.} f_{32}^* && \text{con } f_{32}^* := M_{\kappa_2, \boldsymbol{\Sigma}, 2}.
 \end{aligned}$$

Obteniendo que

$$\frac{1}{n} \mathbf{V}_n \xrightarrow{c.s.} \mathbf{V} \tag{B.65}$$

con

$$\mathbf{V} = \begin{pmatrix} f_{11}^* & \mathbf{0} \\ \mathbf{0} & f_{22}^* \\ \mathbf{0} & f_{32}^* \end{pmatrix}. \tag{B.66}$$

Se busca encontrar el límite casi seguro de la matriz de Moore Penrose \mathbf{V}_n^\dagger , definida por Maruskin (2012) para matrices \mathbf{V} con columnas linealmente independientes, como $\mathbf{V}^\dagger := (\mathbf{V}^t \mathbf{V})^{-1} \mathbf{V}^t$, la cual se aplica tanto a la matriz \mathbf{V} como a \mathbf{V}_n , ambas con columnas linealmente independientes. Esta definición es fundamental para analizar la convergencia de $n \mathbf{V}_n^\dagger$, es decir, la convergencia de $n \mathbf{V}_n^\dagger = n (\mathbf{V}_n^t \mathbf{V}_n)^{-1} \mathbf{V}_n^t$. Para abordar este análisis, se utiliza (B.65) obteniendo,

$$\left(\frac{1}{n} \mathbf{V}_n^t \right) \left(\frac{1}{n} \mathbf{V}_n \right) \xrightarrow{c.s.} \mathbf{V}^t \mathbf{V},$$

donde

$$\mathbf{V}^t \mathbf{V} = \begin{pmatrix} f_{11}^{*t} f_{11}^* & \mathbf{0} \\ \mathbf{0} & f_{22}^{*t} f_{22}^* + f_{32}^{*t} f_{32}^* \end{pmatrix},$$

ya que $\mathbf{V}^t \mathbf{V}$ es invertible,

$$\left(\frac{1}{n} \mathbf{V}_n^t \frac{1}{n} \mathbf{V}_n \right)^{-1} \xrightarrow{c.s.} (\mathbf{V}^t \mathbf{V})^{-1} = \begin{pmatrix} (f_{11}^{*t} f_{11}^*)^{-1} & \mathbf{0} \\ \mathbf{0} & (f_{22}^{*t} f_{22}^* + f_{32}^{*t} f_{32}^*)^{-1} \end{pmatrix},$$

utilizando nuevamente (B.65), se obtiene

$$\left(\frac{1}{n} \mathbf{V}_n^t \frac{1}{n} \mathbf{V}_n \right)^{-1} \frac{1}{n} \mathbf{V}_n^t \xrightarrow{c.s.} (\mathbf{V}^t \mathbf{V})^{-1} \mathbf{V}^t$$

con

$$(\mathbf{V}^t \mathbf{V})^{-1} \mathbf{V}^t = \begin{pmatrix} ((f_{11}^*)^t f_{11}^*)^{-1} (f_{11}^*)^t & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & ((f_{22}^*)^t f_{22}^* + (f_{32}^*)^t f_{32}^*)^{-1} (f_{22}^*)^t & ((f_{22}^*)^t f_{22}^* + (f_{32}^*)^t f_{32}^*)^{-1} (f_{32}^*)^t \end{pmatrix} \quad (\text{B.67})$$

de tamaño $[pq + q(q + 1)/2] \times [pq + q(q + 1)/2 + 1]$. □

C Apéndice del Capítulo 3

C.1 Demostración de la Observación 3.2.4

Demostración. Obsérvese que para $\mathbf{F} = \mathbf{GR}$ se tiene

$$\begin{aligned} h_M \left(h_M^{-1} \left(\begin{pmatrix} \mathbf{R} \\ \mathbf{F} \end{pmatrix}, \boldsymbol{\Sigma} \right) \right) &= h_M \left(h_M^{-1} \left(\begin{pmatrix} \mathbf{R} \\ \mathbf{GR} \end{pmatrix}, \boldsymbol{\Sigma} \right) \right) \\ &= h_M \left(\mathbf{GRR}^t (\mathbf{RR}^t)^{-1}, \mathbf{R}, \boldsymbol{\Sigma} \right) \\ &= \begin{pmatrix} \mathbf{R} \\ \mathbf{GR} \end{pmatrix}, \boldsymbol{\Sigma} \\ &= \begin{pmatrix} \mathbf{R} \\ \mathbf{F} \end{pmatrix}, \boldsymbol{\Sigma} \end{aligned}$$

□

C.2 Demostración del Teorema 14

Demostración del Teorema. Por el Teorema 13, la sucesión de estimadores $(\mathbf{B}_n^t, \boldsymbol{\Sigma}_n)$ resulta fuertemente consistente a $(\mathbf{B}_0^t, \boldsymbol{\Sigma}_0^*)$. Como $\mathbf{B}_0^t \in \mathbb{R}_{first,d}^{q \times p}$, entonces, existe un n_0 a partir del cual $\mathbf{B}_n^t \in \mathbb{R}_{first,d}^{q \times p}$. Dado que la función h_M es biyectiva, con $(\mathbf{B}_n^t, \boldsymbol{\Sigma}_n)$ en su imagen, se puede afirmar que existen sucesiones $\mathbf{Q}_n \in \mathbb{R}^{(q-d) \times d}$ y $\mathbf{R}_n \in \mathbb{R}_d^{d \times p}$ tales que \mathbf{B}_n puede ser expresada como $\mathbf{B}_n^t = \begin{pmatrix} \mathbf{R}_n \\ \mathbf{Q}_n \mathbf{R}_n \end{pmatrix}$.

Considerando las sucesiones \mathbf{Q}_n y \mathbf{R}_n por la continuidad y biyectividad de la función h_M (ver Lema 3.1.1) se garantiza que la convergencia se mantiene con los nuevos parámetros y se obtiene

$$\begin{aligned} \lim_{n \rightarrow \infty} (\mathbf{Q}_n, \mathbf{R}_n, \boldsymbol{\Sigma}_n) &= \lim_{n \rightarrow \infty} h_M^{-1}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \boldsymbol{\Sigma}_n)) \\ &= \lim_{n \rightarrow \infty} h_M^{-1}(\mathbf{B}_n, \boldsymbol{\Sigma}_n) \\ &= h_M^{-1}(\mathbf{B}_0, \boldsymbol{\Sigma}_0^*) \\ &= h_M^{-1}(h_M(\mathbf{Q}_0, \mathbf{R}_0, \boldsymbol{\Sigma}_0^*)) \\ &= (\mathbf{Q}_0, \mathbf{R}_0, \boldsymbol{\Sigma}_0^*). \end{aligned}$$

□

C.3 Demostración del Teorema 15

Se quiere obtener el sistema de ecuaciones correspondiente al caso de rango reducido, para ello es necesario comprobar las derivadas a utilizar. Teniendo en cuenta las propiedades presentadas a continuación, se pueden hallar las derivadas parciales requeridas.

Lema C.3.1. *Las derivadas de $\begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}$ con respecto a \mathbf{Q} y \mathbf{R} , vienen dadas por las siguientes matrices,*

$$(a) \frac{\partial \text{vec} \begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}}{\partial \text{vec}^t(\mathbf{Q})} = \mathbf{R}^t \otimes \begin{pmatrix} \mathbf{0}_{d \times (q-d)d} \\ \mathbf{I}_{(q-d)} \end{pmatrix} \text{ de dimensiones } pq \times (q-d)d,$$

$$(b) \frac{\partial \text{vec} \begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}}{\partial \text{vec}^t(\mathbf{R})} = \mathbf{I}_p \otimes \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q} \end{pmatrix} \text{ de dimensiones } pq \times pd.$$

La demostración de este Lema se encuentra en (Bura et al., 2018).

A partir de la Definición 3.1.2 y la ecuación (3.7) el parámetro \mathbf{B} viene dado por la imagen del elemento (\mathbf{Q}, \mathbf{R}) a través de h_1 . En consecuencia, el MLM (1) con $\text{rango}(\mathbf{B}_0) < \min\{p, q\}$ depende ahora de tres parámetros $(\mathbf{Q}_0, \mathbf{R}_0, \mathbf{\Sigma}_0)$, y por ende el sistema de ecuaciones para las τ -estimaciones, será diferente a (1.22).

Demostración del Teorema 15. Se usará un procedimiento similar al presentado en la demostración del Teorema 2, ya que minimizar la función $\log(g(\mathbf{B}, \mathbf{\Sigma})) = \log(|\mathbf{\Sigma}| \tau^{2q} (H_{P_n, \mathbf{B}, \mathbf{\Sigma}}))$ es equivalente a minimizar la función objetivo del τ -estimador, sólo que ahora se busca optimizarla sujeto a la restricción del rango en \mathbf{B} . Para tomar en cuenta esta restricción, se minimiza la composición $\log(g(h_M(\mathbf{Q}, \mathbf{R}, \mathbf{\Sigma})))$. Primero se hallan las derivadas parciales de esta función compuesta. Con h_1 y h_2 definidas en (3.7), las derivadas de la composición con respecto a \mathbf{Q} , \mathbf{R} y $\mathbf{\Sigma}$ quedan expresadas de la siguiente manera,

$$\begin{aligned}
 \frac{\partial \log(g(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)))}{\partial \text{vec}^t(\mathbf{Q})} &= \frac{\partial(\log g)}{\partial \text{vec}^t(h_1)} \Big|_{(h_1=h_1(\mathbf{Q}, \mathbf{R}), h_2=\Sigma)} \frac{\partial \text{vec}(h_1)}{\partial \text{vec}^t(\mathbf{Q})} \\
 &= \frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{\partial \text{vec} \begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}}{\partial \text{vec}^t(\mathbf{Q})}, \\
 \frac{\partial \log(g(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)))}{\partial \text{vec}^t(\mathbf{R})} &= \frac{\partial(\log g)}{\partial \text{vec}^t(h_1)} \Big|_{(h_1=h_1(\mathbf{Q}, \mathbf{R}), h_2=\Sigma)} \frac{\partial \text{vec}(h_1)}{\partial \text{vec}^t(\mathbf{R})} \\
 &= \frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vec}^t(\mathbf{B}^t)} \frac{\partial \text{vec} \begin{pmatrix} \mathbf{R} \\ \mathbf{QR} \end{pmatrix}}{\partial \text{vec}^t(\mathbf{R})}, \\
 \frac{\partial \log(g(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)))}{\partial \text{vech}^t(\Sigma)} &= \frac{\partial \log(g(\mathbf{B}, \Sigma))}{\partial \text{vech}^t(\Sigma)}.
 \end{aligned}$$

Igualando a cero las tres derivadas parciales, y utilizando las expresiones de las derivadas de la función objetivo $\log(g(\mathbf{B}, \Sigma))$ que se obtuvieron para el caso de rango completo (A.22) en la demostración del Teorema 2, y las derivadas (a) y (b) del Lema C.3.1, se pueden obtener las ecuaciones que satisfacen los puntos críticos. Las primeras dos resultan

$$\begin{aligned}
 \mathbf{0} &= \sum_i w_n^* (d_i^* h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \text{vec}^t(\mathbf{u}_i(h_1^t(\mathbf{Q}, \mathbf{R})) \mathbf{x}_i^t) (\mathbf{I}_p \otimes \Sigma^{-1}) \left(\mathbf{R}^t \otimes \begin{pmatrix} \mathbf{0}_{d \times (q-d)} \\ \mathbf{I}_{(q-d)} \end{pmatrix} \right), \\
 \mathbf{0} &= \sum_i w_n^* (d_i^* h_M(\mathbf{Q}, \mathbf{R}, \Sigma)) \text{vec}^t(\mathbf{u}_i(h_1^t(\mathbf{Q}, \mathbf{R})) \mathbf{x}_i^t) (\mathbf{I}_p \otimes \Sigma^{-1}) \left(\mathbf{I}_p \otimes \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q} \end{pmatrix} \right).
 \end{aligned}$$

La ecuación que corresponde a la derivada con respecto a Σ se mantiene igual que en el caso de rango completo presentado en el Capítulo 2. Lo mismo ocurre con la ecuación correspondiente a la restricción de la escala τ .

□

C.4 Demostración del Teorema 16

A continuación se presenta un resultado necesario, para la demostración del Teorema 16.

Lema C.4.1. *Considerando las sucesiones las sucesiones $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ en las hipótesis del Teorema 15 y $(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*) \in \Theta_d^*$ que converge a.s. a $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$, la sucesión $n(\mathbf{V}_{R,n})^\dagger$ de la inversa Moore-Penrose de la matriz $\mathbf{V}_{R,n}$ converge casi seguramente a $(\mathbf{V}_R)^\dagger$, donde $\mathbf{V}_{R,n}$ se define como $\mathbf{G}(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) \mathbf{V}_n J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*)$ y \mathbf{V}_R según la definición (3.13), donde las matrices \mathbf{V}_n , \mathbf{G} y J_R fueron definidas en (2.16), (3.9), y (3.14), respectivamente.*

Demostración del Lema C.4.1. Para establecer la convergencia de la sucesión de matrices de Moore-Penrose $(\mathbf{V}_{R,n})^\dagger$ para columnas linealmente independientes (Maruskin, 2012), se utiliza la convergencia de \mathbf{V}_n según se muestra en (B.65), la consistencia casi segura de las

sucesiones $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ y $(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*)$ a $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$, y la continuidad de \mathbf{G} y J_R . Así se tiene,

$$\begin{aligned} \frac{1}{n} \mathbf{V}_n &\xrightarrow{c.s.} \mathbf{V}, \\ \mathbf{G}(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n) &\xrightarrow{c.s.} \mathbf{G}(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), \\ J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*) &\xrightarrow{c.s.} J_R(\mathbf{Q}_0, \mathbf{R}_0). \end{aligned}$$

De donde se deduce,

$$\frac{1}{n} \mathbf{V}_{R,n} \xrightarrow{c.s.} \mathbf{V}_R. \quad (\text{C.1})$$

Ya que $\mathbf{V}_{R,n}$ tiene columnas independientes se tiene que $(\mathbf{V}_{R,n})^\dagger = ((\mathbf{V}_{R,n})^t \mathbf{V}_{R,n})^{-1} (\mathbf{V}_{R,n})^t$. Luego de (C.1),

$$\left(\frac{1}{n} (\mathbf{V}_{R,n})^t \right) \left(\frac{1}{n} \mathbf{V}_{R,n} \right) \xrightarrow{c.s.} (\mathbf{V}_R)^t \mathbf{V}_R,$$

luego, ya que $(\mathbf{V}_R)^t \mathbf{V}_R$ y $(\mathbf{V}_{R,n})^t \mathbf{V}_{R,n}$ son invertibles,

$$\left(\frac{1}{n} (\mathbf{V}_{R,n})^t \frac{1}{n} \mathbf{V}_{R,n} \right)^{-1} \xrightarrow{c.s.} ((\mathbf{V}_R)^t \mathbf{V}_R)^{-1}.$$

Por último, multiplicando por derecha por $\frac{1}{n} (\mathbf{V}_{R,n})^t$, se tiene

$$n (\mathbf{V}_{R,n})^\dagger = \left(\frac{1}{n} (\mathbf{V}_{R,n})^t \frac{1}{n} \mathbf{V}_{R,n} \right)^{-1} \frac{1}{n} (\mathbf{V}_{R,n})^t \xrightarrow{c.s.} ((\mathbf{V}_R)^t \mathbf{V}_R)^{-1} (\mathbf{V}_R)^t = (\mathbf{V}_R)^\dagger.$$

□

Demostración del Teorema 16. Para esta demostración se sigue el procedimiento realizado en la prueba del Teorema 9, hasta la ecuación (2.18), es decir, es decir, como $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ es una sucesión que satisface el sistema de ecuaciones (3.10) y converge a.s. a $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$, se tiene que

$$\begin{aligned} &\sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n), s_n(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n))) - \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))) \\ &= (F_{1,n}(h_M(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*)), F_{2,n}(h_M(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*))) J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*) \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix}, \end{aligned} \quad (\text{C.2})$$

donde, el producto $(F_{1,n}(\cdot), F_{2,n}(\cdot)) J_R(\cdot)$ es una matriz cuyas columnas son las derivadas de $\sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}, \mathbf{R}, \Sigma), s_n(h_M(\mathbf{Q}, \mathbf{R}, \Sigma)))$ con respecto a $\text{vec}(\mathbf{Q}^t)$, $\text{vec}(\mathbf{R}^t)$ y $\text{vech}(\Sigma)$ de tamaño $[2pq + q(q+1) + 1] \times [pd + (q-d)d + q(q+1)/2]$, además

$$J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*) = \begin{pmatrix} (\mathbf{R}_n^*)^t \otimes \begin{pmatrix} \mathbf{0}_{d \times (q-d)} \\ \mathbf{I}_{(q-d)} \end{pmatrix} & \mathbf{I}_p \otimes \begin{pmatrix} \mathbf{I}_d \\ \mathbf{Q}_n^* \end{pmatrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{q(q+1)/2} \end{pmatrix},$$

y $(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*)$ son valores que se encuentran entre $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ y $(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)$. Se multiplica por el producto matricial $\mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n))$ por izquierda en ambos miembros de (C.2), el resultado obtenido fue

$$\begin{aligned}
 & \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n), s_n(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n))) \\
 & - \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))) \\
 & = \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) (F_{1,n}(h_M(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*)), F_{2,n}(h_M(\mathbf{Q}_n^*, \mathbf{R}_n^*, \Sigma_n^*))) \\
 & \times J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*) \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix}. \tag{C.3}
 \end{aligned}$$

Ya que $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$ cumple con el sistema de ecuaciones (3.8), por el Lema 3.4.1 se tiene

$$\mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n), s_n(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n))) = \mathbf{0}.$$

Sustituyendo en la ecuación (C.3), resulta

$$\begin{aligned}
 & - \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))) \\
 & = \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{V}_n J_R(\mathbf{Q}_n^*, \mathbf{R}_n^*) \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix},
 \end{aligned}$$

donde \mathbf{V}_n se define en (2.16). Según la notación del Lema C.4.1, se obtiene

$$\begin{aligned}
 & - \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))) \\
 & = \mathbf{V}_{R,n} \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix}. \tag{C.4}
 \end{aligned}$$

Luego premultiplicando a ambos lados de la ecuación (C.4) por $\sqrt{n} (\mathbf{V}_{R,n})^\dagger$, definida en el Lema C.4.1 resulta que

$$\begin{aligned}
 & -n (\mathbf{V}_{R,n})^\dagger \mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)} \mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)) \frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))) \\
 & = \sqrt{n} \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix}.
 \end{aligned}$$

Ahora, para los términos del lado izquierdo de esta igualdad se tienen los siguientes resultados:

- (a) Por el Corolario 2.1.1: la sucesión de matrices $\mathbf{M}_{(C_n, D_n)}(h_M(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n))$ converge casi seguramente a $\mathbf{M}_{(C, D)}(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*))$ definida en (2.3).

- (b) Por el Teorema 7: $-\frac{1}{\sqrt{n}} \sum_{i=1}^n \mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s_n(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)))$ converge en distribución a una normal con media cero y matriz de covarianza (2.12).
- (c) Por la consistencia de la sucesión de $(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)$, se tiene que $\mathbf{G}_{(\mathbf{Q}_n, \mathbf{R}_n, \Sigma_n)}$ converge casi seguramente a $\mathbf{G}_{(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)}$.
- (d) Por el Lema C.4.1: la sucesión $n(\mathbf{V}_{R,n})^\dagger$ converge en probabilidad a $(\mathbf{V}_R)^\dagger$.

Por lo tanto, la distribución asintótica de

$$\sqrt{n} \begin{pmatrix} \text{vec}(\mathbf{Q}_n) - \text{vec}(\mathbf{Q}_0) \\ \text{vec}(\mathbf{R}_n) - \text{vec}(\mathbf{R}_0) \\ \text{vech}(\Sigma_n) - \text{vech}(\Sigma_0^*) \end{pmatrix},$$

se logra empleando (a), (b), (c), (d) y el Teorema de Slutsky. Así la convergencia en distribución es a una normal con media igual a

$$\mathbf{W}_R E[\mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)))] = \mathbf{0},$$

con

$$\mathbf{W}_R = (\mathbf{V}_R)^\dagger \mathbf{G}_{(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)} \mathbf{M}_{(C,D)}(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)),$$

y matriz de covarianza

$$\mathbf{W}_R \text{Var}[\mathbf{A}_i(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*), s(h_M(\mathbf{Q}_0, \mathbf{R}_0, \Sigma_0^*)))] (\mathbf{W}_R)^t.$$

□

Notación

Este capítulo proporciona las notaciones y sus definiciones correspondientes que serán utilizadas a lo largo de la tesis. Si es necesario, se dan otras definiciones en los capítulos correspondientes.

- c.s. denota en las convergencias el caso de casi segura.
- $\mathbb{R}_{\geq 0}$ es el conjunto de números reales no negativos.
- $\mathbb{R}_d^{p \times q}$ es el conjunto de matrices de tamaño $p \times q$ con rango d .
- $\mathbb{R}_{first,d}^{p \times q}$ representa el conjunto de matrices $\mathbb{R}_d^{p \times q}$ cuyas primeras d filas son linealmente independientes.
- $\perp\!\!\!\perp$ denota independencia entre dos vectores aleatorios.
- \mathbf{D}_q matriz de duplicación de tamaño $q^2 \times q(q+1)/2$, que conectan el operador vec y el operador $vech$ como sigue $\mathbf{D}_q vech(\mathbf{A}) = vec(\mathbf{A})$.
- \mathbf{L}_q matriz de eliminación, de tamaño $q(q+1)/2 \times q^2$, que conectan el operador vec y el operador $vech$ como sigue $\mathbf{L}_q vec(\mathbf{A}) = vech(\mathbf{A})$, para cualquier matriz $\mathbf{A}_{q \times q}$ simétrica.
- $\mathbf{K}_{p,q}$ matriz de conmutación de tamaño $pq \times pq$, que transforma $vec(\mathbf{A})$ en $vec(\mathbf{A}^t)$, es decir, $\mathbf{K}_{p,q} vec(\mathbf{A}) = vec(\mathbf{A}^t)$, para cualquier matriz $\mathbf{A}_{p \times q}$.
- \mathbf{N} matriz de tamaño $q^2 \times q^2$ se define como $\mathbf{N} = \frac{1}{2}(\mathbf{I}_q^2 + \mathbf{K}_{q,q})$, con la propiedad de que $\mathbf{N}vec(\mathbf{A}) = \frac{1}{2}vec(\mathbf{A} + \mathbf{A}^t)$ para cualquier matriz $\mathbf{A}_{q \times q}$.
- \cong indica que un componente del espacio matricial en un espacio producto se identifica con su imagen a través del operador vec .
- \xrightarrow{w} denota la convergencia débil.
- \xrightarrow{d} denota la convergencia en distribución.
- $\xrightarrow{c.s.}$ denota la convergencia casi segura.