PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICA
DEPARTAMENTO DE ESTADÍSTICA

PROPIEDADES Y APLICACIONES DE LA FAMILIA DE DISTRIBUCIONES SKEW-NORMAL MULTIVARIADA

POR GABRIELA VALDÉS

Tesis presentada a la Facultad de Matemática de la
Pontificia Universidad Católica de Chile para optar
al grado académico de Doctor en Estadística

Profesor Guía: Reinaldo Arellano.

Abril de 2011
Santiago. Chile.

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICA
DEPARTAMENTO DE ESTADÍSTICA

Título de Tesis: Propiedades y Aplicaciones de la Familia de Distribuciones Skew-Normal Multivariada.

Autor: Gabriela Valdés.

Fecha: Abril de 2011.

Santiago. Chile.

Profesor Guía: ______________________

Reinaldo Arellano

Comisión: ______________________

Manuel Galea

____

Francisco Torres

____

Mauricio Castro

Índice general

Resumen

El presente trabajo tiene por objetivo encontrar y mostrar propiedades no exploradas de la familia de distribuciones Skew-Normal.

El modelo utilizado en este trabajo fue planteado por Arellano-Valle y Genton (2005), para el que existe una representación jerárquica que fue utilizada para calcular ciertos momentos.

Además, se muestran ciertas condiciones de independencia de combinaciones lineales y cuadrátias de esta distribución. Para poder utilizar esta representación estocástica, fue necesario encontar ciertas propiedades de otras distribuciones, tales como, la distribución Normal-Truncada y la distribución Half-Normal.

Finalmente, se muestran resultados para la Matriz de Información de este modelo, señalándose también algunos casos particulares.

Capítulo 1

Introducción

La distribución Skew-Normal es una generalización de la distribución Normal, la cual se obtiene si se anula el parámetro de asimetría. El primer autor que escribió sobre esta familia de distribuciones, fue Azzalini (1985), quien planteó la siguiente definición:

Si la variable aletoria Z tiene función de distribución

$$\phi(z;\lambda) = 2\phi(z)\Phi(\lambda z) \qquad (-\infty < z < \infty)$$

donde ϕ y Φ son las funciones de densidad y distribución normal estándar, respectivamente, entonces podemos decir que Z es una variable skew-normal con parámetro λ.

A partir de esto, se han escrito una gran variedad de trabajos referentes a esta clase de distribuciones, en donde se han desarrollado algunas generalizaciones para el caso univariado y multivariado. Además se han desarrollado estudios con ciertas aplicaciones de esta distribución para el enfoque clásico y el enfoque Bayesiano. Arellano-Valle & Azzalini (2006) y Arellano-Valle & Genton (2005) realizan un resumen de estas variantes y extensiones.

Por otro lado, Sahu, Dey y Branco (2003) definen una nueva clase de la distribución Skew-Normal Multivariada, la que tiene la siguiente forma:

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{D}) = 2^m|\boldsymbol{\Sigma}+\boldsymbol{D}^2|^{-1/2}\phi_m\{(\boldsymbol{\Sigma}+\boldsymbol{D}^2)^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\}P(\boldsymbol{V}>\boldsymbol{0})$$

donde $\boldsymbol{\mu}$ es un vector m-dimensional, $\boldsymbol{\Sigma}$ es una matriz definida positiva de dimensiones $m \times m$ y $\boldsymbol{D}$ es una matriz diagonal con elementos $\delta_1, ..., \delta_m$.

Además, $\boldsymbol{V} \sim N_m\left(D(\boldsymbol{\Sigma}+\boldsymbol{D}^2)^{-1}(\boldsymbol{y}-\boldsymbol{\mu}), \boldsymbol{I}-\boldsymbol{D}(\boldsymbol{\Sigma}+\boldsymbol{D}^2)^{-1}\boldsymbol{D}\right)$, y ϕ_m es la densidad de la distribución normal m-variada estándar.

Ghosh et. al (2007) resaltan algunas diferencias entre los dos modelos y agregan que el modelo planteado por Sahu, Dey y Branco (2003) es muy conveniente para el enfoque bayesiano.

Un artículo importante para el desarrollo de este trabajo, es el presentado por Arellano-Valle y Genton (2005), señalado anteriormente, quienes propusieron la siguiente parametrización:

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Lambda}) \quad = \quad 2^m\phi_k(\boldsymbol{y}-\boldsymbol{\mu}|\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})\Phi_m(\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}(\boldsymbol{y}-\boldsymbol{\mu})|\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top)$$

donde $\boldsymbol{y}$ es un vector de dimensión $k \times 1$, $\boldsymbol{\Omega}$ es una matriz de dimensión $k \times k$ y $\boldsymbol{\Lambda}$ es una matriz de dimensión $m \times k$, y ϕ_k y Φ_m es la funión de densidad k-variada y la función de distribución m-variada de la distribución normal, respectivamente.

El objetivo de este trabajo es encontrar ciertas propiedades de esta parametrización de la distribución Skew-Normal, con el fin de que sean utilizadas en investigaciones, ya sea realizando estimaciones sobre los parámetros, desde un punto de vista clásico y Bayesiano, o realizar inferencias posteriores.

Para esto, en el segundo capítulo se presentan algunos resultados útiles para distribuciones relacionadas con la Skew-Normal, como son la distribución Normal, Half-Normal y Normal truncada. Luego, en el tercer capítulo se discuten algunos aspectos inferenciales para lo cual se calcula la matriz de información para el caso univariado, la que se extiende después al caso multivariado.

Capítulo 2

Propiedades

En este capítulo se presentan algunas propiedades no estándar de varias distribuciones relacionadas con la familia skew-normal, tales como la distribución normal, half-normal y normal truncada. Finalmente, se presentan algunas propiedades de la distribución skew-normal que serán necesarias en algunos cálculos de los capítulos posteriores.

2.1. Distribución Normal Truncada Multivariada

Un vector aleatorio n-dimensional $\boldsymbol{Z}$ tiene distribución normal multivariada, denotada por $\boldsymbol{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su función de densidad es

$$\phi_n(\boldsymbol{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})\right\}, \qquad \boldsymbol{z} \in R^n$$

donde $\boldsymbol{\mu} \in R^n$, $\boldsymbol{\Sigma} \in R^{n\times n}$

Denote por $TN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{b})$, $\boldsymbol{b} \in R^n$, a la distribución normal n-variada $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, truncada en una región del hiperplano izquierdo de R^n definido por $\{\boldsymbol{x} \in R^n : \boldsymbol{x} > \boldsymbol{b}\} = (\boldsymbol{b}, \infty)$.

Escribiremos $\boldsymbol{X} \sim TN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{b})$, para decir que la densidad de $\boldsymbol{X}$ tiene la forma

$$f(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{b}) = \frac{1}{\boldsymbol{\alpha}}\phi_n(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})\boldsymbol{I}_{(\boldsymbol{b},\infty)}(\boldsymbol{x})$$

donde $\boldsymbol{\alpha} = \Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b}, \boldsymbol{\Sigma}\right)$ e $\boldsymbol{I}_{(\boldsymbol{b},\infty)}(\boldsymbol{x})$ es el indicador del evento $(\boldsymbol{b}, \infty)$.

Teorema 2.1.1 Si $\boldsymbol{X} \sim TN_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{b})$, entonces

a) $E(\boldsymbol{X}) = \boldsymbol{\mu} + \boldsymbol{\alpha}^{-1}\boldsymbol{\Sigma}\boldsymbol{q}_n$

b) $E(\boldsymbol{X}\boldsymbol{X}^\top) = \boldsymbol{\mu}\boldsymbol{\eta}^\top + \boldsymbol{\eta}\boldsymbol{\mu}^\top - \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\Sigma} + \boldsymbol{\alpha}^{-1}\boldsymbol{\Sigma}\boldsymbol{Q}_n\boldsymbol{\Sigma}$

donde $\boldsymbol{\eta} = E(\boldsymbol{X})$ y

$$\boldsymbol{q}_n = \left.\frac{d\Phi_n(\boldsymbol{s}|b, \boldsymbol{\Sigma})}{d\boldsymbol{s}}\right|_{\boldsymbol{s}=\boldsymbol{\mu}} = \Phi_n'(\boldsymbol{\mu}|\boldsymbol{b}, \boldsymbol{\Sigma})$$

$$\boldsymbol{Q}_n = \left.\frac{d^2\Phi_n(\boldsymbol{s}|b, \boldsymbol{\Sigma})}{d\boldsymbol{s}d\boldsymbol{s}^\top}\right|_{\boldsymbol{s}=\boldsymbol{\mu}} = \Phi_n''(\boldsymbol{\mu}|\boldsymbol{b}, \boldsymbol{\Sigma})$$

Note que, según lo expuesto en Lin et al. (2009), $\boldsymbol{q}_n$ se puede expresar como un vector cuyo r-ésimo elemento es $f_r(b_r)G_{(r)}$ con

$$f_r(b_r) = \phi(b_r|\mu_r, \sigma_{rr})$$

y

$$G_{(r)} = \prod_{j\neq r} \int_{a_j}^{\infty} \phi_{n-1}\left(x_{(r)}|\mu_{2\cdot 1}^{(r)}, \Sigma_{22\cdot 1}^{(r)}\right) dx_{(r)}$$

y donde $\phi_{n-1}\left(x_{(r)}|\mu_{2\cdot 1}^{(r)}, \Sigma_{22\cdot 1}^{(r)}\right)$ denota la densidad condicional de las restantes $n-1$ variables dado $X_r = b_r$.

Además, Q_m se puede expresar como la suma de $\boldsymbol{H}$ y $\boldsymbol{A}$, las cuales son matrices de $p \times p$ dadas por

$$\begin{aligned} \boldsymbol{H} &= ((\delta_{rs} f_{r,s}(b_r, b_s) G_{(rs)})) \\ \boldsymbol{A} &= ((\sigma_{rr}^{-1}((b_r - \mu_r) f_r(b_r) - [\boldsymbol{\Sigma H}]_{rr}))) \end{aligned}$$

donde $[\boldsymbol{\Sigma H}]_{rr}$ denota el elemento (r, r) situado en la diagonal de la matriz $\boldsymbol{\Sigma H}$, $f_{r,s}(b_r, b_s)$ representa una densidad normal bivariada para la (r, s)-ésima variable de $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ evaluada en (b_r, b_s) y

$$G_{(rs)} = \prod_{j\neq r,s} \int_{a_j}^{\infty} \phi_{n-2}\left(x_{(r,s)}|\mu_{2\cdot 1}^{(r,s)}, \Sigma_{22\cdot 1}^{(r,s)}\right) dx_{(r,s)}.$$

Dem.

La función generadora de momentos de $\boldsymbol{X} \sim TN(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \boldsymbol{B})$ es

$$\begin{aligned}
M_{\boldsymbol{X}}(\boldsymbol{t}) &= \frac{1}{\alpha}\int_{\boldsymbol{x}>\boldsymbol{b}} \exp\left\{\boldsymbol{t}^\top x\right\} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} d\boldsymbol{x} \\
&= \frac{1}{\alpha}\int_{\boldsymbol{y}>\boldsymbol{b}-\boldsymbol{\mu}} \exp\left\{\boldsymbol{t}^\top(\boldsymbol{y}+\boldsymbol{\mu})\right\} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\boldsymbol{y}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{y}\right\} d\boldsymbol{y} \\
&= \frac{1}{\alpha}\exp\left\{\boldsymbol{t}^\top \boldsymbol{\mu}\right\}\int_{\boldsymbol{y}>\boldsymbol{b}-\boldsymbol{\mu}} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{y}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{y} - 2\boldsymbol{t}^\top y\right)\right\} d\boldsymbol{y} \\
&= \frac{1}{\alpha}\exp\left\{\boldsymbol{t}^\top \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}\right\}\int_{\boldsymbol{y}>\boldsymbol{b}-\boldsymbol{\mu}} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\left(\boldsymbol{y}-\boldsymbol{\Sigma}\boldsymbol{t}\right)^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{y}-\boldsymbol{\Sigma}\boldsymbol{t}\right)\right\} d\boldsymbol{y} \\
&= \frac{1}{\alpha}M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\int_{\boldsymbol{z}>\boldsymbol{b}-\boldsymbol{\mu}-\boldsymbol{\Sigma}\boldsymbol{t}} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{z}\right\} d\boldsymbol{y} \\
&= M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\frac{\Phi_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma}\boldsymbol{t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}
\end{aligned}$$

donde $M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t}) = \exp\left\{\boldsymbol{t}^\top\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Sigma}\boldsymbol{t}\right\}$ es la función generadora de momentos de una variable $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

La primera derivada de esta función es

$$\begin{aligned}
M'_{\boldsymbol{X}}(t) &= M'_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\frac{\Phi_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma}\boldsymbol{t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} + M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\frac{\Phi'_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma}\boldsymbol{t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} \\
&= M'_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\frac{\Phi_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma}\boldsymbol{t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} + M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{s}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}
\end{aligned}$$

Por lo que el primer momento tiene la siguiente forma

$$\begin{aligned}
E(\boldsymbol{X}) &= M_{\boldsymbol{X}}(0) \\
&= M'_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(0) + M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(0)\boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} \\
&= \boldsymbol{\mu} + \boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} \\
&= \boldsymbol{\mu} + \boldsymbol{\alpha}^{-1}\boldsymbol{\Sigma}\boldsymbol{q}_n
\end{aligned}$$

De la misma forma, para el segundo momento se tiene que

$$\begin{aligned} M''_{\boldsymbol{X}}(\boldsymbol{t}) &= M''_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\frac{\Phi_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)} + \boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\left(M'_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\right)^\top \\ &+ M'_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t})\left[\boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\right]^\top + \boldsymbol{\Sigma}\frac{\Phi''_n\left(\boldsymbol{\mu}+\boldsymbol{\Sigma t}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\Sigma}M_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{t}) \end{aligned}$$

y

$$\begin{aligned} E(\boldsymbol{X}\boldsymbol{X}^\top) &= M''_{\boldsymbol{X}}(0) \\ &= M''_{N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})}(0) + \boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\left[\boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\right]^\top + \boldsymbol{\Sigma}\frac{\Phi''_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\Sigma} \\ &= \boldsymbol{\Sigma} + \boldsymbol{\mu\mu}^\top + \boldsymbol{\Sigma}\frac{\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\frac{\left[\Phi'_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)\right]^\top}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\frac{\Phi''_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}{\Phi_n\left(\boldsymbol{\mu}|\boldsymbol{b},\boldsymbol{\Sigma}\right)}\boldsymbol{\Sigma} \\ &= \boldsymbol{\mu\eta}^\top + \boldsymbol{\eta\mu}^\top - \boldsymbol{\mu\mu}^\top + \boldsymbol{\Sigma} + \boldsymbol{\alpha}^{-1}\boldsymbol{\Sigma}\boldsymbol{Q}_n\boldsymbol{\Sigma} \end{aligned}$$

<u>Observación</u>:

Tomando $\boldsymbol{B} = \{\boldsymbol{x} = (x_1, ..., x_n)^\top | x_i > \mu_i; i = 1, ..., n\}$, se obtiene la distribución half-normal multivariada con densidad dada por

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\phi_n(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma})}{\Phi(\boldsymbol{\mu}|\boldsymbol{0},\boldsymbol{\Sigma})} \qquad \boldsymbol{x}-\boldsymbol{\mu} \in R^n_+$$

denotado por $\boldsymbol{X} \sim HN_m(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

En particular, cuando $\boldsymbol{\mu} = 0$ y $\boldsymbol{\Sigma} = \boldsymbol{I}$ se obtiene la distribución half-normal estándar multivariada.

2.2. Otras propiedades de la distribución Normal Multivariada

En esta sección se revisan algunos resultados no muy conocidos de la distribución normal multivariada que serán útiles para el desarrollo del presente trabajo.

Teorema 2.2.1 Sea $\boldsymbol{X} = (X_1, ..., X_n) \sim N_n(0, \boldsymbol{I}_n)$. Entonces

a) $E(X_p \boldsymbol{X}\boldsymbol{X}^\top) = 0$,

b) $E(X_p X_q \boldsymbol{X}\boldsymbol{X}^\top) = \boldsymbol{\Delta}_{pq} + \boldsymbol{\Delta}_{qp} + \delta_{pq}\boldsymbol{I}_n$,

c) $E(\boldsymbol{X} \otimes \boldsymbol{X}) = vec(\boldsymbol{I}_n)$,

d) $E(\boldsymbol{X} \otimes \boldsymbol{X}\boldsymbol{X}^\top) = 0$,

e) $E(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top) = \boldsymbol{I}_{n^2} + \boldsymbol{K}_{nn} + [vec(\boldsymbol{I}_n)][vec(\boldsymbol{I}_n)]^\top$,

f) $Cov(\boldsymbol{X} \otimes \boldsymbol{X}) = \boldsymbol{I}_{n^2} + \boldsymbol{K}_{nn}$,

donde p y q pueden asumir cualquier valor desde $1, ..., n$, δ_{pq} es una función indicadora que vale 1 si $p = q$ y 0 en caso contrario, y $\boldsymbol{\Delta}_{pq}$ y $\boldsymbol{K}_{nn}$ está definidas en la sección 5.1.4.

Dem. Ver Graybill (1983).

Teorema 2.2.2 Sea $\boldsymbol{X} \sim N_n(\mathbf{0}, \boldsymbol{V})$. Entonces

a) $E(\boldsymbol{X} \otimes \boldsymbol{X}) = vec(\boldsymbol{V})$

b) $Cov(\boldsymbol{X} \otimes \boldsymbol{X}) = (\boldsymbol{V} \otimes \boldsymbol{V})(\boldsymbol{I}_{n^2} + \boldsymbol{K}_{nn}) = (\boldsymbol{I}_{n^2} + \boldsymbol{K}_{nn})(\boldsymbol{V} \otimes \boldsymbol{V})$

Dem. Ver Graybill (1983).

Teorema 2.2.3 Sea $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{V})$. Entonces

a) $E(\boldsymbol{X} \otimes \boldsymbol{X}) = vec(\boldsymbol{V}) + (\boldsymbol{\mu} \otimes \boldsymbol{\mu})$,

b) $E(\boldsymbol{X} \otimes \boldsymbol{X}\boldsymbol{X}^\top) = vec(\boldsymbol{V})\boldsymbol{\mu}^\top + (\boldsymbol{V} \otimes \boldsymbol{\mu}) + (\boldsymbol{\mu} \otimes \boldsymbol{V})\boldsymbol{\mu} \otimes \boldsymbol{\mu}\boldsymbol{\mu}^\top$,

c) $Cov(\boldsymbol{X} \otimes \boldsymbol{X}) = (\boldsymbol{I}_{n^2} + \boldsymbol{K}_{nn})(\boldsymbol{V} \otimes \boldsymbol{V} + \boldsymbol{V} \otimes \boldsymbol{\mu}\boldsymbol{\mu}^\top + \boldsymbol{\mu}\boldsymbol{\mu}^\top \otimes \boldsymbol{V})$

Dem. Para a) y c) ver Graybill (1983). Para b) ver Arellano-Valle (2010)

Teorema 2.2.4 Sea $\Phi_m(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma})$ la función de distribución de una variable Normal m-variada con media $\boldsymbol{\beta}$ y matriz de varianzas-covarianzas $\boldsymbol{\Gamma}$, evaluada en el vector $\boldsymbol{\alpha}$. Entonces, el diferencial de $\Phi_m(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma})$ tiene la siguiente forma

$$d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) = \boldsymbol{q}_m^\top(d\boldsymbol{\alpha}) - \boldsymbol{q}_m^\top(d\boldsymbol{\beta}) + \frac{1}{2}(vec(d\boldsymbol{\Gamma}))^\top vec(Q_m)$$

donde, $\boldsymbol{q}_m = \Phi'_m(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma})$ y $\boldsymbol{Q}_m = \Phi''_m(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma})$

Sigue que,

$$\begin{aligned}
\frac{d}{d\alpha}\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= \boldsymbol{q}_m \\
\frac{d}{d\beta}\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= -\boldsymbol{q}_m \\
\frac{d}{d(vec(\boldsymbol{\Gamma}))}\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= \frac{1}{2}vec(Q_m)
\end{aligned}$$

Dem. Usando la definición de Φ_m se tiene que

$$\begin{aligned}
d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= d\left(\int_{-\infty}^{\boldsymbol{\alpha}} \phi\left(\boldsymbol{x}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)d\boldsymbol{x}\right) \\
&= d\left(\int_{-\infty}^{\boldsymbol{0}} \phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)d\boldsymbol{z}\right) \\
&= \int_{-\infty}^{\boldsymbol{0}} d\left(\phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right)d\boldsymbol{z} \\
&= \int_{-\infty}^{\boldsymbol{0}} \phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)d\left(log\left(\phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right)\right)d\boldsymbol{z} \\
&= \int_{-\infty}^{\boldsymbol{0}} \phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)d\left(-\frac{1}{2}log|\boldsymbol{\Gamma}| - \frac{1}{2}tr\left\{\boldsymbol{\Gamma}^{-1}(\boldsymbol{z}+\boldsymbol{\alpha}-\boldsymbol{\beta})(\boldsymbol{z}+\boldsymbol{\alpha}-\boldsymbol{\beta})^\top\right\}\right)d\boldsymbol{z} \\
&= \int_{-\infty}^{\boldsymbol{0}} \phi\left(\boldsymbol{z}+\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right) \\
&\qquad\left(-\frac{1}{2}tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\Gamma})\right\} - \frac{1}{2}tr\left\{-\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\Gamma})\boldsymbol{\Gamma}^{-1}(\boldsymbol{z}+\boldsymbol{\alpha}-\boldsymbol{\beta})(\boldsymbol{z}+\boldsymbol{\alpha}-\boldsymbol{\beta})^\top\right\}\right. \\
&\qquad\left. -tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\alpha}-d\boldsymbol{\beta})\left(\boldsymbol{z}+\boldsymbol{\alpha}-\boldsymbol{\beta}\right)^\top\right\}\right)d\boldsymbol{z} \\
&= -\frac{1}{2}\Phi\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\Gamma})\right\} - \frac{1}{2}tr\left\{-\boldsymbol{\Gamma}^{-2}(d\boldsymbol{\Gamma})\int_{-\infty}^{\boldsymbol{\alpha}}\phi\left(\boldsymbol{x}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)(\boldsymbol{x}-\boldsymbol{\beta})(\boldsymbol{x}-\boldsymbol{\beta})^\top d\boldsymbol{x}\right\} \\
&\quad -tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\alpha}-d\boldsymbol{\beta})\int_{-\infty}^{\boldsymbol{\alpha}}\phi\left(\boldsymbol{x}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)(\boldsymbol{x}-\boldsymbol{\beta})^\top d\boldsymbol{x}\right\}
\end{aligned}$$

Por otro lado, tenemos que si $\boldsymbol{X} \sim N_m(\boldsymbol{\beta}, \boldsymbol{\Gamma})$, entonces $\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{\beta} \sim N(\boldsymbol{0}, \boldsymbol{\Gamma})$

$$\begin{aligned}
\int_{-\infty}^{\boldsymbol{\alpha}} \phi\left(\boldsymbol{x}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)(\boldsymbol{x}-\boldsymbol{\beta})^\top d\boldsymbol{x} &= \int_{-\infty}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} \phi\left(\boldsymbol{y}|\boldsymbol{0}, \boldsymbol{\Gamma}\right)\boldsymbol{y}^\top d\boldsymbol{y} \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}^\top|\boldsymbol{Y} \leq \boldsymbol{\alpha} - \boldsymbol{\beta}\right) \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}^\top| -\boldsymbol{Y} \geq -\boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\
&= -\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(-\boldsymbol{Y}^\top| -\boldsymbol{Y} \geq -\boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\
&= -\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}^\top|\boldsymbol{Y} \geq -\boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\
&= -\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E(\boldsymbol{S}^\top)
\end{aligned}$$

donde $\boldsymbol{S} \sim TN_m\left(\boldsymbol{0}, \boldsymbol{\Gamma}; -\boldsymbol{\alpha} + \boldsymbol{\beta}\right)$ con

$$\begin{aligned}
E(\boldsymbol{S}) &= \Phi_m^{-1}\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}\boldsymbol{q}_m \\
E(\boldsymbol{S}\boldsymbol{S}^\top) &= \boldsymbol{\Gamma} + \Phi_m^{-1}\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)\boldsymbol{\Gamma}Q_m\boldsymbol{\Gamma}
\end{aligned}$$

y

$$\begin{aligned}
\boldsymbol{q}_m &= \left.\frac{d\Phi_m(\boldsymbol{s}|-\boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\Gamma})}{d\boldsymbol{s}}\right|_{\boldsymbol{s}=\boldsymbol{0}} = \Phi_m'(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}) \\
\boldsymbol{Q}_m &= \left.\frac{d^2\Phi_m(\boldsymbol{s}|-\boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\Gamma})}{d\boldsymbol{s}d\boldsymbol{s}^\top}\right|_{\boldsymbol{s}=\boldsymbol{0}} = \Phi_m''(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma})
\end{aligned}$$

por lo que

$$\int_{-\infty}^{\boldsymbol{\alpha}} \phi\left(\boldsymbol{x}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)\boldsymbol{x}^\top d\boldsymbol{x} = -\boldsymbol{q}_m^\top\boldsymbol{\Gamma}$$

De la misma forma

$$\begin{aligned}
\int_{-\infty}^{\boldsymbol{\alpha}} \phi\left(\boldsymbol{x}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)(\boldsymbol{x}-\boldsymbol{\beta})(\boldsymbol{x}-\boldsymbol{\beta})^\top d\boldsymbol{x} &= \int_{-\infty}^{\boldsymbol{\alpha}-\boldsymbol{\beta}} \phi\left(\boldsymbol{y}|\boldsymbol{0}, \boldsymbol{\Gamma}\right)\boldsymbol{y}\boldsymbol{y}^\top d\boldsymbol{y} \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}\boldsymbol{Y}^\top|\boldsymbol{Y} \leq \boldsymbol{\alpha} - \boldsymbol{\beta}\right) \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}\boldsymbol{Y}^\top| -\boldsymbol{Y} \geq -\boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E\left(\boldsymbol{Y}\boldsymbol{Y}^\top|\boldsymbol{Y} \geq -\boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right) E(\boldsymbol{S}\boldsymbol{S}^\top) \\
&= \Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\Gamma}\right)\boldsymbol{\Gamma} + \boldsymbol{\Gamma}Q_m\boldsymbol{\Gamma}
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= -\frac{1}{2}\Phi\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right) tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\Gamma})\right\} + \frac{1}{2} tr\left\{\boldsymbol{\Gamma}^{-2}(d\boldsymbol{\Gamma})\left[\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\boldsymbol{\Gamma} + \boldsymbol{\Gamma} Q_m \boldsymbol{\Gamma}\right]\right\} \\
&\quad + tr\left\{\boldsymbol{\Gamma}^{-1}(d\boldsymbol{\alpha} - d\boldsymbol{\beta})\boldsymbol{q}_m\boldsymbol{\Gamma}\right\} \\
&= \boldsymbol{q}_m^{\top}(d\boldsymbol{\alpha}) - \boldsymbol{q}_m^{\top}(d\boldsymbol{\beta}) + \frac{1}{2} tr\left\{(d\boldsymbol{\Gamma})Q_m\right\} \\
&= \boldsymbol{q}_m^{\top}(d\boldsymbol{\alpha}) - \boldsymbol{q}_m^{\top}(d\boldsymbol{\beta}) + \frac{1}{2}(vec(d\boldsymbol{\Gamma}))^{\top} vec(Q_m)
\end{aligned}$$

2.3. Distribución Half-Normal Multivariada

Un vector aleatorio n-dimensional $\boldsymbol{X}$ tiene distribución half-normal estándar multivariada, denotada por $\boldsymbol{X} \sim HN_n(0, \boldsymbol{I}_n)$ si su función de densidad es

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = 2^n \phi_n(\boldsymbol{x}), \qquad \boldsymbol{x} \in R_+^n.$$

Teorema 2.3.1 Si $\boldsymbol{W} \sim HN_n(\boldsymbol{0}, \boldsymbol{I}_n)$ representa la distribución half-normal n-variada, entonces sus dos primeros momentos multivariados son:

a) $E(\boldsymbol{W}) = \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{1}_n$,

b) $E(\boldsymbol{W}\boldsymbol{W}^\top) = \left(1 - \frac{2}{\pi}\right) \boldsymbol{I}_n + \frac{2}{\pi} \boldsymbol{1}_n \boldsymbol{1}_n^\top$.

Note también que $\boldsymbol{W} = |\boldsymbol{X}| \stackrel{d}{=} (\boldsymbol{X} | \boldsymbol{X} > \boldsymbol{0})$, donde $\boldsymbol{X} \sim N_n(\boldsymbol{0}, \boldsymbol{I}_n)$ y para
$\boldsymbol{x} = (x_1, ...x_n)^\top$, $|\boldsymbol{x}| = (|x_1|, ..., |x_n|)^\top$

Dem.

a)

$$\begin{aligned}
E(\boldsymbol{W}) &= E((W_1, ..., W_n)^\top) \\
&= (E(W_1), ..., E(W_n))^\top \\
&= \left(\left(\frac{2}{\pi}\right)^{1/2}, ..., \left(\frac{2}{\pi}\right)^{1/2} \right)^\top \\
&= \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{1}_n
\end{aligned}$$

b)

$$\begin{aligned}
E(\boldsymbol{W}\boldsymbol{W}^\top) &= E \begin{pmatrix} W_1^2 & W_1W_2 & \cdots & W_1W_n \\ \vdots & \vdots & & \vdots \\ W_1W_n & W_2W_n & \cdots & W_n^2 \end{pmatrix} \\
&= \begin{pmatrix} E(W_1^2) & E(W_1)E(W_2) & \cdots & E(W_1)E(W_n) \\ \vdots & \vdots & & \vdots \\ E(W_1)E(W_n) & E(W_2)E(W_n) & \cdots & E(W_n^2) \end{pmatrix} \\
&= \begin{pmatrix} 1 & \frac{2}{\pi} & \cdots & \frac{2}{\pi} \\ \vdots & \vdots & & \vdots \\ \frac{2}{\pi} & \frac{2}{\pi} & \cdots & 1 \end{pmatrix} \\
&= \left(1 - \frac{2}{\pi}\right) \boldsymbol{I}_n + \frac{2}{\pi} \boldsymbol{1}_n \boldsymbol{1}_n^\top
\end{aligned}$$

Teorema 2.3.2 Sea $\boldsymbol{W}$ un vector de $n \times 1$ que distribuye $HN(\boldsymbol{0}, \boldsymbol{I}_n)$. Entonces

a) $$E(\boldsymbol{W} \otimes \boldsymbol{W}) = \left(1 - \frac{2}{\pi}\right) vec(\boldsymbol{I}_n) + \frac{2}{\pi}\boldsymbol{1}_{n^2},$$

b) $$E(\boldsymbol{W} \otimes \boldsymbol{W}\boldsymbol{W}^\top) = \left(\frac{2}{\pi}\right)^{1/2} \left[2\sum_i (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i \neq j} [\boldsymbol{d}_i \otimes (\boldsymbol{\Delta}_{ij} + \boldsymbol{\Delta}_{ji} + \boldsymbol{\Delta}_{jj})] + \left(\frac{2}{\pi}\right) \sum_{i \neq j \neq k} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk})\right],$$

c) $$E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}) = \left(\frac{2}{\pi}\right)^{1/2} \Big[2\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i) + \sum_{i \neq j} [(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_j) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j)]$$
$$+ \frac{2}{\pi} \sum_{i \neq j \neq k} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k)\Big],$$

d) $$E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top) = 3\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) + \frac{4}{\pi} \sum_{i \neq j} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})]$$
$$+ \sum_{i \neq j} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji})$$
$$+ \frac{2}{\pi} \sum_{i \neq j \neq k} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kk}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ki})] + \frac{4}{\pi^2} \sum_{i \neq j \neq k \neq l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl}).$$

donde $\boldsymbol{\Delta}_{ij}$ se define en la sección 5.1.4.

Dem.

a) Sabemos que $\boldsymbol{W} \otimes \boldsymbol{W} = vec(\boldsymbol{W}\boldsymbol{W}^\top)$, por lo que

$$\begin{aligned}
E(\boldsymbol{W} \otimes \boldsymbol{W}) &= E(vec(\boldsymbol{W}\boldsymbol{W}^\top)) \\
&= vec(E(\boldsymbol{W}\boldsymbol{W}^\top)) \\
&= vec\left(\left(1 - \frac{2}{\pi}\right)\boldsymbol{I}_n + \frac{2}{\pi}\mathbf{1}_n\mathbf{1}_n^\top\right) \\
&= \left(1 - \frac{2}{\pi}\right) vec(\boldsymbol{I}_n) + \frac{2}{\pi} vec(\mathbf{1}_n\mathbf{1}_n^\top) \\
&= \left(1 - \frac{2}{\pi}\right) vec(\boldsymbol{I}_n) + \frac{2}{\pi}\mathbf{1}_{n^2}
\end{aligned}$$

b) Sabemos que $\boldsymbol{W} = \sum_i W_i \boldsymbol{d}_i$ y que $\boldsymbol{W}\boldsymbol{W}^\top = \sum_{jk} W_j W_k \boldsymbol{\Delta}_{jk}$.

Así,

$$\begin{aligned}
E(\boldsymbol{W} \otimes \boldsymbol{W}\boldsymbol{W}^\top) &= E\left(\sum_i W_i \boldsymbol{d}_i \otimes \sum_{jk} W_j W_k \boldsymbol{\Delta}_{jk}\right) \\
&= E\left(\sum_{ijk} W_i W_j W_k \boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk}\right) \\
&= \sum_{ijk} E(W_i W_j W_k)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk}) \\
&= \sum_{i=j=k} E(W_i^3)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i=j\neq k} E(W_i^2)E(W_k)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ik}) \\
&\quad + \sum_{i=k\neq j} E(W_i^2)E(W_j)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ji}) + \sum_{i\neq j=k} E(W_i)E(W_j^2)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jj}) \\
&\quad + \sum_{i\neq j\neq k} E(W_i)E(W_j)E(W_k)(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk}) \\
&= 2\left(\frac{2}{\pi}\right)^{1/2} \sum_i (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ii}) + \left(\frac{2}{\pi}\right)^{1/2} \left[\sum_{i\neq j} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ij}) + \sum_{i\neq j} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ji}) + \sum_{i\neq j} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jj})\right] \\
&\quad + \left(\frac{2}{\pi}\right)^{3/2} \sum_{i\neq j\neq k} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk}) \\
&= \left(\frac{2}{\pi}\right)^{1/2} \Big[2\sum_i (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i\neq j} [(\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jj})] \\
&\qquad + \left(\frac{2}{\pi}\right) \sum_{i\neq j\neq k} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk})\Big] \\
&= \left(\frac{2}{\pi}\right)^{1/2} \left[2\sum_i (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i\neq j} [\boldsymbol{d}_i \otimes (\boldsymbol{\Delta}_{ij} + \boldsymbol{\Delta}_{ji} + \boldsymbol{\Delta}_{jj})] + \left(\frac{2}{\pi}\right) \sum_{i\neq j\neq k} (\boldsymbol{d}_i \otimes \boldsymbol{\Delta}_{jk})\right]
\end{aligned}$$

c)

$$
\begin{aligned}
E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}) &= E\left(\sum_{ij} W_i W_j \boldsymbol{\Delta}_{ij} \otimes \sum_k W_k \boldsymbol{d}_k\right) \\
&= E\left(\sum_{ijk} W_i W_j W_k \boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k\right) \\
&= \sum_{ijk} E(W_i W_j W_k)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k) \\
&= \sum_{i=j=k} E(W_i^3)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i) + \sum_{i=j\neq k} E(W_i^2)E(W_k)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_k) \\
&\quad + \sum_{i=k\neq j} E(W_i^2)E(W_j)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i) + \sum_{i\neq j=k} E(W_i)E(W_j^2)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j) \\
&\quad + \sum_{i\neq j\neq k} E(W_i)E(W_j)E(W_k)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k) \\
&= 2\left(\frac{2}{\pi}\right)^{1/2} \sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i) + \left(\frac{2}{\pi}\right)^{1/2} \left[\sum_{i\neq j} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_j) + \sum_{i\neq j} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i) + \sum_{i\neq j} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j)\right] \\
&\quad + \left(\frac{2}{\pi}\right)^{3/2} \sum_{i\neq j\neq k} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k) \\
&= \left(\frac{2}{\pi}\right)^{1/2} \left[2\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i) + \sum_{i\neq j} [(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_j) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j)] + \frac{2}{\pi} \sum_{i\neq j\neq k} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_k)\right]
\end{aligned}
$$

d)

$$
\begin{aligned}
E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top) &= E(\sum_{ij} W_i W_j \boldsymbol{\Delta}_{ij} \otimes \sum_{kl} W_k W_l \boldsymbol{\Delta}_{kl}) \\
&= E(\sum_{ijkl} W_i W_j W_k W_l \boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl}) \\
&= \sum_{ijkl} E(W_i W_j W_k W_l)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl}) \\
&= \sum_{i=j=k=l} E(W_i^4)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i\neq j=k=l} E(W_i)E(W_j^3)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) \\
&\quad + \sum_{j\neq i=k=l} E(W_i^3)E(W_j)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + \sum_{k\neq i=j=l} E(W_i^3)E(W_k)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ki}) \\
&\quad + \sum_{l\neq i=j=k} E(W_i^3)E(W_l)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{il}) + \sum_{i=j\neq k=l} E(W_i^2)E(W_k^2)(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{kk}) \\
&\quad + \sum_{i=k\neq j=l} E(W_i^2)E(W_j^2)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + \sum_{i=l\neq j=k} E(W_i^2)E(W_j^2)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji}) \\
&\quad + \sum_{i\neq j\neq k=l} E(W_i)E(W_j)E(W_k^2)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kk}) + \sum_{i\neq k\neq l=j} E(W_i)E(W_j^2)E(W_k)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kj}) \\
&\quad + \sum_{j\neq k\neq l=i} E(W_i^2)E(W_j)E(W_k)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ki}) \\
&\quad + \sum_{i\neq j\neq k\neq l} E(W_i)E(W_j)E(W_k)E(W_l)(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl}) \\
&= 3\sum_{i=j=k=l} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) \\
&\quad + \frac{4}{\pi}\Bigg[\sum_{i\neq j=k=l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + \sum_{j\neq i=k=l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + \sum_{k\neq i=j=l} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ki}) \\
&\qquad + \sum_{l\neq i=j=k} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{il})\Bigg] \\
&\quad + \sum_{i=j\neq k=l} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{kk}) + \sum_{i=k\neq j=l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + \sum_{i=l\neq j=k} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji}) \\
&\quad + \frac{2}{\pi}\Bigg[\sum_{i\neq j\neq k=l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kk}) + \sum_{i\neq k\neq l=j} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kj}) + \sum_{j\neq k\neq l=i} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ki})\Bigg] \\
&\quad + \frac{4}{\pi^2}\sum_{i\neq j\neq k\neq l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl})
\end{aligned}
$$

$$
\begin{aligned}
E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top) &= 3\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) \\
&+ \frac{4}{\pi}\left[\sum_{i\neq j}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + \sum_{i\neq j}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + \sum_{i\neq j}(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + \sum_{i\neq j}(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})\right] \\
&+ \sum_{i\neq j}(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + \sum_{i\neq j}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + \sum_{i\neq j}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji}) \\
&+ \frac{2}{\pi}\left[\sum_{i\neq j\neq k}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kk}) + \sum_{i\neq j\neq k}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kj}) + \sum_{i\neq j\neq k}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ki})\right] \\
&+ \frac{4}{\pi^2}\sum_{i\neq j\neq k\neq l}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl}) \\
&= 3\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) + \frac{4}{\pi}\sum_{i\neq j}\left[(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})\right] \\
&+ \sum_{i\neq j}\left[(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji})\right] \\
&+ \frac{2}{\pi}\sum_{i\neq j\neq k}\left[(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kk}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ki})\right] + \frac{4}{\pi^2}\sum_{i\neq j\neq k\neq l}(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{kl})
\end{aligned}
$$

2.4. Normal con Half-Normal

En esta sección se calculan algunos momentos de productos de variables distribuidas normalmente con variables distribuidas half-normal. Estos resultados son muy útiles para el cálculo de las propiedades de la variable Skew-Nomal.

Teorema 2.4.1 Sea $\boldsymbol{W}$ un vector de $m \times 1$ que distribuye $HN_m(\mathbf{0}, \boldsymbol{I}_m)$ y $\boldsymbol{X}$ un vector de $n \times 1$ que distribuye $N_n(\mathbf{0}, \boldsymbol{\Sigma})$. Si $\boldsymbol{W}$ es independiente de $\boldsymbol{X}$, entonces

a) $E(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top) = \left(\frac{2}{\pi}\right)^{1/2} (\boldsymbol{\Sigma} \otimes \mathbf{1}_m^\top)$

b) $E(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}^\top) = \left(\frac{2}{\pi}\right)^{1/2} \left(\mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right)$

c) $E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top) = \left(\frac{2}{\pi}\right)^{1/2} \mathbf{1}_m vec(\boldsymbol{\Sigma})^\top$

d) $E(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top) = \left(1 - \frac{2}{\pi}\right) (\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m) + \frac{2}{\pi} \left(\boldsymbol{\Sigma} \otimes \mathbf{1}_m \mathbf{1}_m^\top\right)$

e) $E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top) = \left(1 - \frac{2}{\pi}\right) (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) + \frac{2}{\pi} \left(\mathbf{1}_m \mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right)$

f) $E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top) = \left(1 - \frac{2}{\pi}\right) vec\,(\boldsymbol{I}_m)\, vec(\boldsymbol{\Sigma})^\top + \frac{2}{\pi} vec\left(\mathbf{1}_m \mathbf{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top$

g) $E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top) = \left(1 - \frac{2}{\pi}\right) (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma})\, \boldsymbol{K}_{km} + \frac{2}{\pi} \left(\mathbf{1}_m \mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right) \boldsymbol{K}_{km}$

Dem.

a)

$$\begin{aligned}
E(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top) &= E(\boldsymbol{X}\boldsymbol{X}^\top) \otimes E(\boldsymbol{W}^\top) \\
&= \boldsymbol{\Sigma} \otimes \left(\frac{2}{\pi}\right)^{1/2} \mathbf{1}_m^\top \\
&= \left(\frac{2}{\pi}\right)^{1/2} (\boldsymbol{\Sigma} \otimes \mathbf{1}_m^\top)
\end{aligned}$$

b)

$$\begin{aligned}
E(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}^\top) &= E(\boldsymbol{X} \otimes \boldsymbol{X}^\top)E(\boldsymbol{W}^\top \otimes \boldsymbol{I}_k) \\
&= E(\boldsymbol{X}\boldsymbol{X}^\top)(E(\boldsymbol{W}^\top) \otimes \boldsymbol{I}_k) \\
&= \boldsymbol{\Sigma}\left(\left(\frac{2}{\pi}\right)^{1/2} \mathbf{1}_m^\top \otimes \boldsymbol{I}_k\right) \\
&= \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Sigma}\left(\mathbf{1}_m^\top \otimes \boldsymbol{I}_k\right) \\
&= \left(\frac{2}{\pi}\right)^{1/2} \left(\mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right)
\end{aligned}$$

c)

$$\begin{aligned}
E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top) &= E(\boldsymbol{W}(\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top)) \\
&= E(\boldsymbol{W})E(\boldsymbol{X} \otimes \boldsymbol{X})^\top \\
&= \left(\frac{2}{\pi}\right)^{1/2} \mathbf{1}_m vec(E(\boldsymbol{X}\boldsymbol{X}^\top))^\top \\
&= \left(\frac{2}{\pi}\right)^{1/2} \mathbf{1}_m vec(\boldsymbol{\Sigma})^\top
\end{aligned}$$

d)

$$\begin{aligned}
E(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top) &= E(\boldsymbol{X}\boldsymbol{X}^\top) \otimes E(\boldsymbol{W}\boldsymbol{W}^\top) \\
&= \left(\boldsymbol{\Sigma} \otimes \left(\left(1 - \frac{2}{\pi}\right) \boldsymbol{I}_m + \frac{2}{\pi}\mathbf{1}_m\mathbf{1}_m^\top\right)\right) \\
&= \left(1 - \frac{2}{\pi}\right)(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m) + \frac{2}{\pi}\left(\boldsymbol{\Sigma} \otimes \mathbf{1}_m\mathbf{1}_m^\top\right)
\end{aligned}$$

e)

$$\begin{aligned}
E(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top) &= E(\boldsymbol{W}\boldsymbol{W}^\top) \otimes E(\boldsymbol{X}\boldsymbol{X}^\top) \\
&= \left(\left(\left(1 - \frac{2}{\pi}\right)\boldsymbol{I}_m + \frac{2}{\pi}\mathbf{1}_m\mathbf{1}_m^\top\right) \otimes \boldsymbol{\Sigma}\right) \\
&= \left(1 - \frac{2}{\pi}\right)(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) + \frac{2}{\pi}\left(\mathbf{1}_m\mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right)
\end{aligned}$$

f)

$$\begin{aligned}
E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top) &= E(\boldsymbol{W} \otimes \boldsymbol{W})E(\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top) \\
&= E(vec(\boldsymbol{W}\boldsymbol{W}^\top))E(vec(\boldsymbol{X}\boldsymbol{X}^\top)^\top) \\
&= vec\left(\left(1 - \frac{2}{\pi}\right)\boldsymbol{I}_m + \frac{2}{\pi}\mathbf{1}_m\mathbf{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top \\
&= \left(1 - \frac{2}{\pi}\right) vec\left(\boldsymbol{I}_m\right) vec(\boldsymbol{\Sigma})^\top + \frac{2}{\pi} vec\left(\mathbf{1}_m\mathbf{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top
\end{aligned}$$

g)

$$\begin{aligned}
E(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top) &= E\left[(\boldsymbol{I}_m \otimes \boldsymbol{X})(\boldsymbol{W} \otimes \boldsymbol{W}^\top)(\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{I}_m \otimes \boldsymbol{X})\boldsymbol{W}\boldsymbol{W}^\top(\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X})(\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{I}_k)(\boldsymbol{I}_m \otimes \boldsymbol{X})(\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{I}_k)\boldsymbol{K}_{mk}(\boldsymbol{X} \otimes \boldsymbol{I}_m)(\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{I}_k)\boldsymbol{K}_{mk}(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= E\left[(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{I}_k)\right]\boldsymbol{K}_{mk}E\left[(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{I}_m)\right] \\
&= \left(\left(\left(1 - \frac{2}{\pi}\right)\boldsymbol{I}_m + \frac{2}{\pi}\mathbf{1}_m\mathbf{1}_m^\top\right) \otimes \boldsymbol{I}_k\right)\boldsymbol{K}_{mk}\left(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m\right) \\
&= \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{I}_m \otimes \boldsymbol{I}_k\right)\boldsymbol{K}_{mk}\left(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m\right) + \frac{2}{\pi}\left(\mathbf{1}_m\mathbf{1}_m^\top \otimes \boldsymbol{I}_k\right)\boldsymbol{K}_{mk}\left(\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m\right) \\
&= \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{I}_m \otimes \boldsymbol{I}_k\right)\left(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}\right)\boldsymbol{K}_{mk} + \frac{2}{\pi}\left(\mathbf{1}_m\mathbf{1}_m^\top \otimes \boldsymbol{I}_k\right)\left(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}\right)\boldsymbol{K}_{mk} \\
&= \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}\right)\boldsymbol{K}_{mk} + \frac{2}{\pi}\left(\mathbf{1}_m\mathbf{1}_m^\top \otimes \boldsymbol{\Sigma}\right)\boldsymbol{K}_{mk}
\end{aligned}$$

2.5. Distribución Skew-Normal Generalizada

En esta sección se presentan los resultados más relevantes de la familia de distribuciones Skew-Normal Generalizada. Para lograr obtener estos resultados se aplicarán muchos de los teoremas antes expuestos.

Sea Y un vector aleatorio distribuido Skew-Normal Generalizada con parámetros $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ y $\boldsymbol{\Lambda}$. Diremos que $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$, si su función de densidad es dada por

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = 2^m \phi_k \left(\boldsymbol{y} - \boldsymbol{\mu} | \boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \Phi_m \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) | \boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top\right).$$

donde $\boldsymbol{\mu} \in R^k$, $\boldsymbol{\Sigma}$ es una matriz definida positiva de dimensiones $k \times k$ y $\boldsymbol{\Lambda}$ es una matriz de dimensiones $m \times k$

Esta distribución corresponde a la familia CSN-2 planteada en Arellano-Valle y Azzalini (2006), en el caso en que $\gamma = 0$ y $\boldsymbol{\Gamma} = \boldsymbol{I}_m$. Vea también Arellano-Valle y Genton (2005).

Note que, a partir de esta función de densidad se pueden obtener los siguientes casos particulares:

1. Si $m = 1$ se obtiene el modelo propuesto por Azzalini & Dalla-Valle (1996) y estudiado con detalle en Azzalini & Capitanio (1999)

2. Si $k = m$ y $\boldsymbol{\Lambda}$ es una matriz diagonal, se obtiene el modelo plantedo por Sahu et al. (2003)

3. Si $k = m$ y $\boldsymbol{\Lambda}$ y $\boldsymbol{\Sigma}$ son matrices diagonales, se obtiene el modelo Skew-Normal de variables independientes.

2.5.1. Función Generadora de Momentos

Teorema 2.5.1 Sea $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$. Entonces, la función generadora de Momentos es

$$M_{\boldsymbol{Y}}(\boldsymbol{t}) = 2^m \exp\left\{\boldsymbol{\mu}^\top \boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^\top(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})\boldsymbol{t}\right\} \Phi_m\left(\boldsymbol{\Lambda t}|\boldsymbol{0}, \boldsymbol{I}_m\right)$$

Dem.

$$\begin{aligned}
M_{\boldsymbol{Y}}(\boldsymbol{t}) &= E(\exp\{\boldsymbol{t}^\top \boldsymbol{Y}\}) \\
&= \int_{R^k} \exp\{\boldsymbol{t}^\top \boldsymbol{y}\} 2^m \phi_k\left(\boldsymbol{y} - \boldsymbol{\mu}|\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right) \\
&\qquad \Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{y} - \boldsymbol{\mu})|\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) dy \\
&= 2^m \int |\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}|^{-1/2}(2\pi)^{k/2} \exp\left\{-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^\top(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) + \boldsymbol{t}^\top\boldsymbol{y}\right\} \\
&\qquad \Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{y} - \boldsymbol{\mu})|\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) dy \\
&= 2^m \int |\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}|^{-1/2}(2\pi)^{k/2} \\
&\qquad \exp\left\{-\frac{1}{2}\left(\boldsymbol{y}^\top(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{y} - 2\boldsymbol{\mu}^\top(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{y} + \boldsymbol{\mu}^\top(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\mu} - 2\boldsymbol{t}^\top\boldsymbol{y}\right)\right\} \\
&\qquad \Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{y} - \boldsymbol{\mu})|\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) dy
\end{aligned}$$

$$
\begin{aligned}
M_{\boldsymbol{Y}}(\boldsymbol{t}) &= 2^m \exp\left\{-\frac{1}{2}\boldsymbol{\mu}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\mu}\right\} \\
&\quad \exp\left\{\frac{1}{2}\left(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t}\right)^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\left(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t}\right)\right\} \\
&\quad \int |\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}|^{-1/2}(2\pi)^{k/2} \\
&\qquad \exp\Bigg\{-\frac{1}{2}\Bigg(\boldsymbol{y}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{y}-2\left(\boldsymbol{\mu}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}+\boldsymbol{t}^\top\right)\boldsymbol{y} \\
&\qquad\qquad +(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t})^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t})\Bigg)\Bigg\} \\
&\qquad \Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu})|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)dy \\
&= 2^m \exp\left\{\boldsymbol{\mu}^\top\boldsymbol{t}+\frac{1}{2}\boldsymbol{t}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})\boldsymbol{t}\right\} \\
&\quad \int |\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}|^{-1/2}(2\pi)^{k/2} \\
&\qquad \exp\left\{-\frac{1}{2}\left(\boldsymbol{y}^\top-\left(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t}\right)^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\left(\boldsymbol{y}^\top-\left(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t}\right)\right)\right)\right\} \\
&\qquad \Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{y}-\boldsymbol{\mu})|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)dy \\
&= 2^m \exp\left\{\boldsymbol{\mu}^\top\boldsymbol{t}+\frac{1}{2}\boldsymbol{t}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})\boldsymbol{t}\right\} \\
&\quad E_{\boldsymbol{X}}\left(\Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{x}-\boldsymbol{\mu})|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right)
\end{aligned}
$$

donde $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}+\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{t},\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})$.

Por otro lado, aplicando el lema 2.1 de Gupta et al. (2004) se tiene que

$$
E_{\boldsymbol{X}}\left(\Phi_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(\boldsymbol{x}-\boldsymbol{\mu})|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right)=\Phi_m\left(\boldsymbol{\Lambda}\boldsymbol{t}|\boldsymbol{0},\boldsymbol{I}_m\right)
$$

Por lo que

$$
M_{\boldsymbol{Y}}(\boldsymbol{t}) = 2^m \exp\left\{\boldsymbol{\mu}^\top\boldsymbol{t}+\frac{1}{2}\boldsymbol{t}^\top(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})\boldsymbol{t}\right\}\Phi_m\left(\boldsymbol{\Lambda}\boldsymbol{t}|\boldsymbol{0},\boldsymbol{I}_m\right)
$$

2.5.2. Representación estocástica

Para el vector aleatorio $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$ se puede usar la siguiente representación estocástica

$$\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}^\top \boldsymbol{W} + \boldsymbol{X}$$

donde $\boldsymbol{W} \sim HN_m(\boldsymbol{0}, \boldsymbol{I}_m)$ es independiente de $\boldsymbol{X} \sim N_k(\boldsymbol{0}, \boldsymbol{\Sigma})$

2.5.3. Representación jerárquica

Para el vector aleatorio $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$ se puede usar la siguiente representación jerárquica.

$$\begin{aligned} \boldsymbol{Y}|\boldsymbol{W} &\sim N_k(\boldsymbol{\mu} + \boldsymbol{\Lambda W}, \boldsymbol{\Sigma}) \\ \boldsymbol{W} &\sim HN_m(0, I_m) \end{aligned}$$

<u>Distribución a posteriori</u>:

Según la representación estocástica presentada anteriormente, se puede calcular la distribución a posteriori de W, la que tiene la siguiente forma

$$\begin{aligned} \pi(\boldsymbol{w}|\boldsymbol{Y}) &\propto L(\boldsymbol{w})\pi(\boldsymbol{w}) \\ &\propto \exp\left\{-\frac{1}{2}(\boldsymbol{y} - (\boldsymbol{\mu} + \boldsymbol{\Lambda w}))^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - (\boldsymbol{\mu} + \boldsymbol{\Lambda w}))\right\} \exp\left\{-\frac{1}{2}\boldsymbol{w}\boldsymbol{w}^\top\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[(-\boldsymbol{y}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda w} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda w} - \boldsymbol{w}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{y} + \boldsymbol{w}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \boldsymbol{w}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda w} + \boldsymbol{w}^\top \boldsymbol{w}\right]\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{w}^\top\left(\boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda} + \boldsymbol{I}_m\right)\boldsymbol{w} - (y - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda w} - \boldsymbol{w}^\top \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}(y - \boldsymbol{\mu})\right]\right\} \end{aligned}$$

Por lo tanto,

$$\boldsymbol{W}|\boldsymbol{Y} \sim N_m\left(\left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}), \left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1}\right)$$

y sus primeros momentos son

$$\begin{aligned} E(\boldsymbol{W}|\boldsymbol{Y}) &= \left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) \\ E(\boldsymbol{W}\boldsymbol{W}^\top|\boldsymbol{Y}) &= \left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \\ &\quad - \left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})(\boldsymbol{y} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \\ Var(\boldsymbol{W}|\boldsymbol{Y}) &= \left(\boldsymbol{I}_m + \boldsymbol{\Lambda}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\Lambda}\right)^{-1} \end{aligned}$$

2.5.4. Distribución Marginal

Sea $\boldsymbol{Y} \sim SNG_{km}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$. Por la representación estocástica se tiene que $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}^\top \boldsymbol{W} + \boldsymbol{X}$, donde $\boldsymbol{W} \sim HN_m(\boldsymbol{0}, \boldsymbol{I}_m)$ es independiente de $\boldsymbol{X} \sim N_k(\boldsymbol{0}, \boldsymbol{\Sigma})$, por lo tanto

$$\begin{aligned}
\begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix} &= \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} + \begin{pmatrix} \Lambda_{11} & \cdots & \Lambda_{m1} \\ \vdots & & \\ \Lambda_{1k} & \cdots & \Lambda_{mk} \end{pmatrix} \begin{pmatrix} W_1 \\ \vdots \\ W_m \end{pmatrix} + \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \\
&= \begin{pmatrix} \mu_1 + \Lambda_{11} W_1 + \cdots + \Lambda_{m1} W_m + X_1 \\ \vdots \\ \mu_k + \Lambda_{1k} W_1 + \cdots + \Lambda_{mk} W_m + X_k \end{pmatrix}
\end{aligned}$$

por lo que

$$Y_i = \mu_i + \boldsymbol{\Lambda}_i^\top W + X_i$$

donde μ_i es el i-ésimo elemento del vector $\boldsymbol{\mu}$, X_i es el i-ésimo elemento del vector $\boldsymbol{X}$ que distribuye $N(0, \sigma_i^2)$ y $\boldsymbol{\Lambda}_i$ es la i-ésima columna de la matriz $\boldsymbol{\Lambda}$.

Por otro lado, si estudiamos la función generadora de momentos, tenemos que

$$\begin{aligned}
M_{Y_i}(t) &= M_{Y_1,\dots,Y_i,\dots,Y_k}(0, ..., t, ..,0) \\
&= 2^m \exp\left\{ \boldsymbol{\mu}^\top (0, ..., t, ..,0)^\top + \frac{1}{2} (0, ..., t, ..,0) \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) (0, ..., t, ..,0)^\top \right\} \\
&\quad \Phi_m \left(\boldsymbol{\Lambda} (0, ..., t, ..,0)^\top | \boldsymbol{0}, \boldsymbol{I}_m \right) \\
&= 2^m \exp\left\{ \boldsymbol{\mu}_i t + \frac{1}{2} t^2 \left(\sigma_i^2 + \boldsymbol{\Lambda}_i^\top \boldsymbol{\Lambda}_i \right) \right\} \Phi_m \left(\boldsymbol{\Lambda}_i t | \boldsymbol{0}, \boldsymbol{I}_m \right)
\end{aligned}$$

donde σ_i es el i-ésimo componente de la diagonal de la matriz $\boldsymbol{\Sigma}$ y $\boldsymbol{\Lambda}_i$ es la i-ésima columna de la matriz $\boldsymbol{\Lambda}$

Por lo tanto, $Y_i \sim SNG_{1,m}(\mu_i, \sigma_i^2, \boldsymbol{\Lambda}_i)$

2.5.5. Formas Lineales

Teorema 2.5.2 : Sea $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$ y $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{b}$, con $\boldsymbol{X}$ vector de dimensiones $s \times 1$, $\boldsymbol{A}$ matriz de dimensiones $s \times k$, y $\boldsymbol{b}$ vector de dimensiones $s \times 1$. Entonces $\boldsymbol{X} \sim SNG_{s,m}(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top, \boldsymbol{\Lambda}\boldsymbol{A}^\top)$

Dem.

Usando la función generadora de momentos de $\boldsymbol{Y}$ se tiene que

$$\begin{aligned}
M_{\boldsymbol{X}}(t) &= M_{\boldsymbol{A}\boldsymbol{Y}+\boldsymbol{b}}(\boldsymbol{t}) \\
&= \exp\left\{\boldsymbol{b}^\top\boldsymbol{t}\right\} M_{\boldsymbol{Y}}(\boldsymbol{A}^\top\boldsymbol{t}) \\
&= 2^m \exp\left\{\boldsymbol{b}^\top\boldsymbol{t}\right\} \exp\left\{\boldsymbol{\mu}^\top\boldsymbol{A}^\top\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^\top\boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{A}^\top\boldsymbol{t}\right\} \Phi_m\left(\boldsymbol{\Lambda}\boldsymbol{A}^\top\boldsymbol{t}|\boldsymbol{0}, \boldsymbol{I}_m\right) \\
&= 2^m \exp\left\{\left(\boldsymbol{\mu}^\top\boldsymbol{A}^\top + \boldsymbol{b}^\top\right)\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^\top\boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)\boldsymbol{A}^\top\boldsymbol{t}\right\} \Phi_m\left(\boldsymbol{\Lambda}\boldsymbol{A}^\top\boldsymbol{t}|\boldsymbol{0}, \boldsymbol{I}_m\right)
\end{aligned}$$

Por lo tanto, $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Y} + \boldsymbol{b} \sim SNG_{s,m}(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top, \boldsymbol{\Lambda}\boldsymbol{A}^\top)$

Teorema 2.5.3 : Sea $\boldsymbol{Y} \sim SNG_{k,m}(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$, $\boldsymbol{A}$ una matriz constante de dimensiones $n_1 \times k$ y $\boldsymbol{B}$ una matriz constante de dimensiones $n_2 \times k$. Entonces, $\boldsymbol{A}\boldsymbol{Y}$ es independiente de $\boldsymbol{B}\boldsymbol{Y}$ si se cumple alguna de las siguientes condiciones

1. $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}^\top = \boldsymbol{0}$ y $\boldsymbol{\Lambda}\boldsymbol{A}^\top = \boldsymbol{0}$

2. $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}^\top = \boldsymbol{0}$ y $\boldsymbol{\Lambda}\boldsymbol{B}^\top = \boldsymbol{0}$

Dem.

Sabemos que

$$\boldsymbol{Y} \stackrel{d}{=} \boldsymbol{X}|\boldsymbol{X}_0 > \boldsymbol{0}$$

donde

$$\begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{X}_0 \end{pmatrix} \sim N_{k+m}\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} & \boldsymbol{\Lambda}^\top \\ \boldsymbol{\Lambda} & I_m \end{pmatrix}\right)$$

Note que la distribución conjunta de $\boldsymbol{A}\boldsymbol{X}, \boldsymbol{B}\boldsymbol{X}, \boldsymbol{X}_0$ es

$$\begin{pmatrix} \boldsymbol{A}\boldsymbol{X} \\ \boldsymbol{B}\boldsymbol{X} \\ X_0 \end{pmatrix} \sim N_{n_1+n_2+m}\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Omega}_{AB} & \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}\boldsymbol{\Lambda}^\top \\ \boldsymbol{\Lambda}\begin{pmatrix} \boldsymbol{A}^\top & \boldsymbol{B}^\top \end{pmatrix} & I_m \end{pmatrix}\right)$$

donde

$$\begin{aligned}
\boldsymbol{\Omega}_{AB} &= \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \begin{pmatrix} \boldsymbol{A}^\top & \boldsymbol{B}^\top \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \\ \boldsymbol{B}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^\top & \boldsymbol{B}^\top \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{A}^\top & \boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{B}^\top \\ \boldsymbol{B}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{A}^\top & \boldsymbol{B}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{B}^\top \end{pmatrix}
\end{aligned}$$

Por lo tanto,

$$\begin{pmatrix} \boldsymbol{AX} \\ \boldsymbol{BX} \\ X_0 \end{pmatrix} \sim N_{n_1+n_2+m}\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{A}^\top & \boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{B}^\top & \boldsymbol{A}\boldsymbol{\Lambda}^\top \\ \boldsymbol{B}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{A}^\top & \boldsymbol{B}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{B}^\top & \boldsymbol{B}\boldsymbol{\Lambda}^\top \\ \boldsymbol{\Lambda}\boldsymbol{A}^\top & \boldsymbol{\Lambda}\boldsymbol{B}^\top & I_m \end{pmatrix} \right)$$

Es decir,

$$\begin{pmatrix} \boldsymbol{AX} \\ \boldsymbol{BX} \end{pmatrix} \Bigg| X_0 > 0 \sim SNG_{n_1+n_2,m}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \boldsymbol{A\Sigma A}^\top & \boldsymbol{A\Sigma B}^\top \\ \boldsymbol{B\Sigma A}^\top & \boldsymbol{B\Sigma B}^\top \end{pmatrix}, \boldsymbol{\Lambda}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}^\top \right)$$

Sea $\boldsymbol{Y}_{AB} = \begin{pmatrix} \boldsymbol{Y}_A \\ \boldsymbol{Y}_B \end{pmatrix} = \begin{pmatrix} \boldsymbol{AY} \\ \boldsymbol{BY} \end{pmatrix}$

$$\begin{aligned}
f_{\boldsymbol{Y}_{AB}}\left(\boldsymbol{Y}_{AB}|0, \boldsymbol{\Sigma}_{AB}, \boldsymbol{\Lambda}_{AB}\right) &= 2^m \phi_{n_1+n_2}\left(\boldsymbol{Y}_{AB}|\boldsymbol{\Sigma}_{AB} + \boldsymbol{\Lambda}_{AB}^\top \boldsymbol{\Lambda}_{AB}\right) \\
&\quad \Phi_m\left(\boldsymbol{\Lambda}_{AB}\left(\boldsymbol{\Sigma}_{AB} + \boldsymbol{\Lambda}_{AB}^\top \boldsymbol{\Lambda}_{AB}\right)^{-1} \boldsymbol{Y}_{AB}|\boldsymbol{0}, \boldsymbol{\Omega}_{AB}\right)
\end{aligned}$$

donde $\boldsymbol{\Sigma}_{AB} = \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} \boldsymbol{A}^\top & \boldsymbol{B}^\top \end{pmatrix}$, $\boldsymbol{\Lambda}_{AB} = \boldsymbol{\Lambda}\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix}^\top$ y $\boldsymbol{\Omega}_{AB} = I_m - \boldsymbol{\Lambda}_{AB}\left(\boldsymbol{\Sigma}_{AB} + \boldsymbol{\Lambda}_{AB}^\top \boldsymbol{\Lambda}_{AB}\right)^{-1} \boldsymbol{\Lambda}_{AB}^\top$

Si estudiamos la parte de la densidad que depende de ϕ tenemos que,
si $\boldsymbol{A}\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)\boldsymbol{B}^\top = \boldsymbol{0}$, entonces

$$\phi_{n_1+n_2}\left(\boldsymbol{Y}_{AB}|\boldsymbol{\Sigma}_{AB} + \boldsymbol{\Lambda}_{AB}^\top \boldsymbol{\Lambda}_{AB}\right) = \phi_{n_1}\left(\boldsymbol{Y}_A|\boldsymbol{\Sigma}_A + \boldsymbol{\Lambda}_A^\top \boldsymbol{\Lambda}_A\right) \phi_{n_2}\left(\boldsymbol{Y}_B|\boldsymbol{\Sigma}_B + \boldsymbol{\Lambda}_B^\top \boldsymbol{\Lambda}_B\right)$$

donde $\boldsymbol{\Sigma}_A = \boldsymbol{A\Sigma A}^\top$, $\boldsymbol{\Lambda}_A = \boldsymbol{\Lambda A}^\top$, $\boldsymbol{\Sigma}_B = \boldsymbol{B\Sigma B}^\top$ y $\boldsymbol{\Lambda}_B = \boldsymbol{\Lambda B}^\top$

Ahora, si estudiamos la parte de la densidad que depende de Φ se tiene que

$$
\begin{aligned}
&\Phi_m\left(\boldsymbol{\Lambda}_{AB}\left(\boldsymbol{\Sigma}_{AB}+\boldsymbol{\Lambda}_{AB}^{\top}\boldsymbol{\Lambda}_{AB}\right)^{-1}\boldsymbol{Y}_{AB}\middle|\mathbf{0},\boldsymbol{\Omega}_{AB}\right)\\
&\qquad=\Phi_m\left(\boldsymbol{\Lambda}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}^{\top}\begin{pmatrix}\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{Y}_A\\ \boldsymbol{Y}_B\end{pmatrix}\middle|\mathbf{0},\boldsymbol{\Omega}_{AB}\right)\\
&\qquad=\Phi_m\left(\begin{pmatrix}\boldsymbol{\Lambda}\boldsymbol{A}^{\top} & \boldsymbol{\Lambda}\boldsymbol{B}^{\top}\end{pmatrix}\begin{pmatrix}\left(\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top}\right)^{-1} & \mathbf{0}\\ \mathbf{0} & \left(\boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\right)^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{Y}_A\\ \boldsymbol{Y}_B\end{pmatrix}\middle|\mathbf{0},\boldsymbol{\Omega}_{AB}\right)\\
&\qquad=\Phi_m\left(\begin{pmatrix}\boldsymbol{\Lambda}\boldsymbol{A}^{\top}\left(\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top}\right)^{-1} & \boldsymbol{\Lambda}\boldsymbol{B}^{\top}\left(\boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\right)^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{Y}_A\\ \boldsymbol{Y}_B\end{pmatrix}\middle|\mathbf{0},\boldsymbol{\Omega}_{AB}\right)\\
&\qquad=\Phi_m\left(\boldsymbol{\Lambda}\boldsymbol{A}^{\top}\left(\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top}\right)^{-1}\boldsymbol{Y}_A+\boldsymbol{\Lambda}\boldsymbol{B}^{\top}\left(\boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\right)^{-1}\boldsymbol{Y}_B\middle|\mathbf{0},\boldsymbol{\Omega}_{AB}\right)
\end{aligned}
$$

Por lo que si $\boldsymbol{\Lambda}\boldsymbol{A}^{\top}\left(\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top}\right)^{-1}=\mathbf{0}$ o $\boldsymbol{\Lambda}\boldsymbol{B}^{\top}\left(\boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\right)^{-1}=\mathbf{0}$, se logra la separación de la función de densidad.

Como $\left(\boldsymbol{A}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{A}^{\top}\right)^{-1}>\mathbf{0}$ y $\left(\boldsymbol{B}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)\boldsymbol{B}^{\top}\right)^{-1}>\mathbf{0}$ al ser matrices de varianzas-covarianzas, se tiene que la condición se reduce a $\boldsymbol{\Lambda}\boldsymbol{A}^{\top}=\mathbf{0}$ o $\boldsymbol{\Lambda}\boldsymbol{B}^{\top}=\mathbf{0}$, lo que produce que la primera condición se reduce a $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}^{\top}=\mathbf{0}$

2.5.6. Formas Cuadráticas

Teorema 2.5.4 : Sea $\boldsymbol{Y}_0 = \boldsymbol{Y} - \boldsymbol{\mu} \sim SNG_{k,m}(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$, y sean $\boldsymbol{A}$ y $\boldsymbol{B}$ dos matrices simétricas de dimensiones $k \times k$. Entonces, $\boldsymbol{Y}_0^\top \boldsymbol{A} \boldsymbol{Y}_0$ y $\boldsymbol{Y}_0^\top \boldsymbol{B} \boldsymbol{Y}_0$ son independientes si y solo si $\boldsymbol{\Lambda A} = \boldsymbol{\Lambda B} = \boldsymbol{A \Sigma B} = \mathbf{0}$

Dem.
Calculamos la función generadora de momentos conjunta de $Q_{\boldsymbol{A}} = \boldsymbol{Y}_0^\top \boldsymbol{A} \boldsymbol{Y}_0$ y $Q_{\boldsymbol{A}} = \boldsymbol{Y}_0^\top \boldsymbol{B} \boldsymbol{Y}_0$. Si llamamos $\boldsymbol{\Omega} = \boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}$. Por definición, tenemos que

$$
\begin{aligned}
M_{Q_{\boldsymbol{A}},Q_{\boldsymbol{B}}}(t,s) &= \int_{R^k} \exp\left\{t\boldsymbol{y}_0^\top \boldsymbol{A}\boldsymbol{y}_0 + s\boldsymbol{y}^\top \boldsymbol{B}\boldsymbol{y}_0\right\} 2^m \phi_k(\boldsymbol{y}_0|\boldsymbol{\Omega})\Phi_m\left(\boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{y}_0|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top\right) d\boldsymbol{y}_0 \\
&= 2^m |\boldsymbol{\Omega}|^{-1/2} \int_{R^k} \exp\left\{-\frac{1}{2}\boldsymbol{y}_0^\top \left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)\boldsymbol{y}_0\right\} (2\pi)^{-k/2} \\
&\quad \Phi_m\left(\boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{y}_0|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top\right) d\boldsymbol{y}_0 \\
&= 2^m |\boldsymbol{\Omega}|^{-1/2} \left|\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right|^{-1/2} \int_{R^k} \phi_k\left(\boldsymbol{y}_0|\left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)^{-1}\right) \\
&\quad \Phi_m\left(\boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{y}_0|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top\right) d\boldsymbol{y}_0
\end{aligned}
$$

Considerando el cambio de variable

$$
\boldsymbol{z} = \left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)^{1/2} \boldsymbol{y}_0 \qquad \Rightarrow \qquad d\boldsymbol{y}_0 = \left|\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right|^{-1/2} d\boldsymbol{z}
$$

tenemos que

$$
\begin{aligned}
M_{Q_{\boldsymbol{A}},Q_{\boldsymbol{B}}}(t,s) &= 2^m |\boldsymbol{I}_k - 2t\boldsymbol{A\Omega} - 2s\boldsymbol{B\Omega}|^{-1/2} \\
&\quad \int_{R^k} \phi_k(\boldsymbol{z})\,\Phi_m\left(\boldsymbol{\Lambda\Omega}^{-1}\left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)^{-1/2}\boldsymbol{z}|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top\right) d\boldsymbol{z} \\
&= 2^m |\boldsymbol{I}_k - 2t\boldsymbol{A\Omega} - 2s\boldsymbol{B\Omega}|^{-1/2} \\
&\quad E\left(\Phi_m\left(\boldsymbol{\Lambda\Omega}^{-1}\left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)^{-1/2}\boldsymbol{Z}|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top\right)\right) \\
&= 2^m |\boldsymbol{I}_k - 2t\boldsymbol{A\Omega} - 2s\boldsymbol{B\Omega}|^{-1/2} \\
&\quad \Phi_m\left(\mathbf{0}|\boldsymbol{I}_m - \boldsymbol{\Lambda\Omega}^{-1}\boldsymbol{\Lambda}^\top + \boldsymbol{\Lambda\Omega}^{-1}\left(\boldsymbol{\Omega}^{-1} - 2t\boldsymbol{A} - 2s\boldsymbol{B}\right)^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
M_{Q_{\boldsymbol{A}},Q_{\boldsymbol{B}}}(t,s) &= 2^m \left|\boldsymbol{I}_k - 2t\boldsymbol{A\Omega} - 2s\boldsymbol{B\Omega}\right|^{-1/2} \\
&\quad \Phi_m\left(\boldsymbol{0}|\boldsymbol{I}_m + 2\left(t\boldsymbol{\Lambda A} + s\boldsymbol{\Lambda B}\right)\boldsymbol{\Omega}\left(\boldsymbol{I}_k - 2t\boldsymbol{A\Omega} - 2s\boldsymbol{B\Omega}\right)^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}^\top\right)
\end{aligned} \tag{2.1}
$$

Como $M_{Q_{\boldsymbol{A}}} = M_{Q_{\boldsymbol{A}},Q_{\boldsymbol{B}}}(t,0)$ y $M_{Q_{\boldsymbol{B}}} = M_{Q_{\boldsymbol{A}},Q_{\boldsymbol{B}}}(0,s)$, (2.1) se factoriza como el producto de sus marginales si y solo si $\boldsymbol{\Lambda A} = \boldsymbol{\Lambda B} = \boldsymbol{A\Sigma B} = \boldsymbol{0}$

Corolario 2.5.1 : Sea $\boldsymbol{Y}_0 = \boldsymbol{Y} - \boldsymbol{\mu} \sim SNG_{k,m}(\boldsymbol{0},\boldsymbol{\Sigma},\boldsymbol{\Lambda})$, y sea $\boldsymbol{A}$ una matriz simétrica de rango r y tal que:

(i) $\boldsymbol{A\Lambda}^\top = \boldsymbol{\Lambda A} = \boldsymbol{0}$

(ii) $\boldsymbol{A\Sigma A} = \boldsymbol{A}$

Entonces $Q_{\boldsymbol{A}} = \boldsymbol{Y}_0^\top \boldsymbol{A}\boldsymbol{Y}_0 \sim \chi^2_r$

Dem.
De (2.1) se tiene que

$$
M_{Q_{\boldsymbol{A}}}(t) = 2^m \left|\boldsymbol{I}_k - 2t\boldsymbol{A\Omega}\right|^{-1/2} \Phi_m\left(\boldsymbol{0}|\boldsymbol{I}_m + 2t\boldsymbol{\Lambda A\Omega}\left(\boldsymbol{I}_k - 2t\boldsymbol{A\Omega}\right)^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}^\top\right)
$$

Por (i) se tiene que $M_{Q_{\boldsymbol{A}}}(t) = \left|\boldsymbol{I}_k - 2t\boldsymbol{A\Omega}\right|^{-1/2}$
Por (ii) se tiene que $M_{Q_{\boldsymbol{A}}}(t) = (1-2t)^{-1/2}$

2.5.7. Propiedades de la distribución Skew-Normal

El objetivo de esta sección es obtener y mostrar propiedades de esta parametrización de la distribución Skew-Normal, las que pueden ser útiles para realizar estimaciones e inferencia, ya sea clásica o Bayesiana.

Para obtener los primeros momentos de esta distribución, se utilizará la representación estocástica presentada en secciones anteriores, lo que genera los siguientes resultados:

$$
\begin{aligned}
E(\boldsymbol{Y}) &= \boldsymbol{\mu} + \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \mathbf{1}_m, \\
E(\boldsymbol{Y}\boldsymbol{Y}^\top) &= \boldsymbol{\mu}\boldsymbol{\mu}^\top + \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\mu}\mathbf{1}_m^\top\boldsymbol{\Lambda} + \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \mathbf{1}_m\boldsymbol{\mu}^\top + \left(1 - \frac{2}{\pi}\right)\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} + \frac{2}{\pi}\boldsymbol{\Lambda}^\top\mathbf{1}_m\mathbf{1}_m^\top\boldsymbol{\Lambda} + \boldsymbol{\Sigma}, \\
Var(\boldsymbol{Y}) &= \left(1 - \frac{2}{\pi}\right)\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} + \boldsymbol{\Sigma}
\end{aligned}
$$

Ahora bien, si $\boldsymbol{Y}_0 = \boldsymbol{Y} - \boldsymbol{\mu}$ y $\boldsymbol{S}_0 = \boldsymbol{Y}_0\boldsymbol{Y}_0^\top$. Entonces, se tiene que

$$
\begin{aligned}
E(\boldsymbol{Y}_0) &= \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \mathbf{1}_m, \\
Var(\boldsymbol{Y}_0) &= Var(\boldsymbol{Y}) = \left(1 - \frac{2}{\pi}\right)\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} + \boldsymbol{\Sigma}
\end{aligned}
$$

Por otro lado, se sabe que

$$
\begin{aligned}
\boldsymbol{S}_0 &= (\boldsymbol{\Lambda}^\top\boldsymbol{W} + \boldsymbol{X})(\boldsymbol{\Lambda}^\top\boldsymbol{W} + \boldsymbol{X})^\top \\
&= \boldsymbol{\Lambda}^\top\boldsymbol{W}\boldsymbol{W}^\top\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^\top\boldsymbol{W}\boldsymbol{X}^\top + \boldsymbol{X}\boldsymbol{W}^\top\boldsymbol{\Lambda} + \boldsymbol{X}\boldsymbol{X}^\top \\
E(\boldsymbol{S}_0) &= \left(1 - \frac{2}{\pi}\right)\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda} + \frac{2}{\pi}\boldsymbol{\Lambda}^\top\mathbf{1}_m\mathbf{1}_m^\top\boldsymbol{\Lambda} + \boldsymbol{\Sigma}
\end{aligned}
$$

Teorema 2.5.5 Sea $\boldsymbol{Y}$ un vector de $k \times 1$ que distribuye $SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$. Entonces,

$$
\begin{aligned}
a) \quad E\left(\boldsymbol{S}_0 \otimes \boldsymbol{Y}_0^\top\right) &= 2\left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \sum_{i \neq j} \left[(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_j)^\top + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i)^\top + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j)^\top\right] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \left(\frac{2}{\pi}\right)^{3/2} \boldsymbol{\Lambda}^\top \sum_{i \neq j \neq s} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_s)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\frac{2}{\pi}\right)^{1/2} \left(\boldsymbol{1}_m^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) + \left(\frac{2}{\pi}\right)^{1/2} \left(\boldsymbol{\Sigma} \otimes \boldsymbol{1}_m^\top \boldsymbol{\Lambda}\right)
\end{aligned}
$$

$$
\begin{aligned}
b) \quad E(\boldsymbol{S}_0 \otimes \boldsymbol{S}_0) &= 3\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \frac{4}{\pi}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j} \left[(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})\right] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j} \left[(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji})\right] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \frac{2}{\pi}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j \neq s} \left[(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ss}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{si})\right] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \frac{4}{\pi^2}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j \neq s \neq l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sl}) (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ \frac{2}{\pi}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \left(\boldsymbol{1}_m \boldsymbol{1}_m^\top \otimes \boldsymbol{\Sigma}\right) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) + \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{I}_m\right) vec(\boldsymbol{\Sigma})^\top \\
&+ \frac{2}{\pi}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \sum_i \sum_j \Sigma_{jj} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji}) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&+ \frac{2}{\pi}\left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \sum_{i \neq s} \sum_j \Sigma_{jj} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{js}) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&+ \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \sum_i \sum_j \Sigma_{jj} (\boldsymbol{\Delta}_{ji} \otimes \boldsymbol{\Delta}_{ij}) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ \frac{2}{\pi}\left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq s} \sum_j \Sigma_{jj} (\boldsymbol{\Delta}_{ji} \otimes \boldsymbol{\Delta}_{sj}) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) + \left(1 - \frac{2}{\pi}\right) vec(\boldsymbol{\Sigma}) vec\left(\boldsymbol{I}_m\right)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \frac{2}{\pi} vec(\boldsymbol{\Sigma}) vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(1 - \frac{2}{\pi}\right)\left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) (\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&+ \frac{2}{\pi}\left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \left(\boldsymbol{\Sigma} \otimes \boldsymbol{1}_m \boldsymbol{1}_m^\top\right) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) (\boldsymbol{I}_{k^2} + \boldsymbol{K}_{kk}) + vec(\boldsymbol{\Sigma}) vec(\boldsymbol{\Sigma})^\top
\end{aligned}
$$

Dem.

$$
\begin{aligned}
a) \quad E\left(\boldsymbol{S}_0 \otimes \boldsymbol{Y}_0^\top\right) &= E\left(\left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{X}^\top + \boldsymbol{X}\boldsymbol{W}^\top\boldsymbol{\Lambda} + \boldsymbol{X}\boldsymbol{X}^\top\right) \otimes \left(\boldsymbol{W}^\top\boldsymbol{\Lambda} + \boldsymbol{X}^\top\right)\right) \\
&= E\left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top\boldsymbol{\Lambda} \otimes \boldsymbol{W}^\top\boldsymbol{\Lambda}\right) + E\left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top\boldsymbol{\Lambda} \otimes \boldsymbol{X}^\top\right) + E\left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top\boldsymbol{\Lambda}\right) \\
&\quad + E\left(\boldsymbol{\Lambda}\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top\right) + E\left(\boldsymbol{X}\boldsymbol{W}^\top\boldsymbol{\Lambda} \otimes \boldsymbol{W}^\top\boldsymbol{\Lambda}\right) + E\left(\boldsymbol{X}\boldsymbol{W}^\top\boldsymbol{\Lambda} \otimes \boldsymbol{X}^\top\right) \\
&\quad + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top\boldsymbol{\Lambda}\right) + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top\right) \\
&= \boldsymbol{\Lambda}^\top E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \boldsymbol{\Lambda}^\top E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&\quad + \boldsymbol{\Lambda}^\top E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + \boldsymbol{\Lambda} E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top\right) + E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}^\top\right) \\
&= \boldsymbol{\Lambda}^\top E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&\quad + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&= 2\left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_i)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\frac{2}{\pi}\right)^{1/2} \boldsymbol{\Lambda}^\top \sum_{i \neq j} \left[(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{d}_j)^\top + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_i)^\top + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_j)^\top\right] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\frac{2}{\pi}\right)^{3/2} \boldsymbol{\Lambda}^\top \sum_{i \neq j \neq s} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{d}_s)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\frac{2}{\pi}\right)^{1/2} \left(\boldsymbol{1}_m^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) + \left(\frac{2}{\pi}\right)^{1/2} \left(\boldsymbol{\Sigma} \otimes \boldsymbol{1}_m^\top \boldsymbol{\Lambda}\right)
\end{aligned}
$$

b) Usando la representación estocástica $\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}^\top \boldsymbol{W} + \boldsymbol{X}$ se tiene que

$$
\begin{aligned}
\boldsymbol{S}_0 &= (\boldsymbol{\Lambda}^\top \boldsymbol{W} + \boldsymbol{X})(\boldsymbol{\Lambda}^\top \boldsymbol{W} + \boldsymbol{X})^\top \\
&= \boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{X}^\top + \boldsymbol{X}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{X}\boldsymbol{X}^\top \\
E(\boldsymbol{S}_0 \otimes \boldsymbol{S}_0) &= E\Bigg(\left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{X}^\top + \boldsymbol{X}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{X}\boldsymbol{X}^\top\right) \otimes \\
&\qquad \left(\boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{\Lambda}^\top \boldsymbol{W}\boldsymbol{X}^\top + \boldsymbol{X}\boldsymbol{W}^\top \boldsymbol{\Lambda} + \boldsymbol{X}\boldsymbol{X}^\top\right) \Bigg) \\
&= \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \cancel{E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right)} (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \cancel{E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)} (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \cancel{E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)} (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right) \\
&+ \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \cancel{E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right)} \\
&+ \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \cancel{E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)} (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \cancel{E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right)} (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \cancel{E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right)} \\
&+ \cancel{E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)} (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right) \\
&= \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) E\left(\boldsymbol{W}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&+ \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right) + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) E\left(\boldsymbol{W}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&+ \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{W}\boldsymbol{X}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) + E\left(\boldsymbol{X}\boldsymbol{W}^\top \otimes \boldsymbol{X}\boldsymbol{W}^\top\right)(\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&+ \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{W}\boldsymbol{W}^\top\right)(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) + E\left(\boldsymbol{X}\boldsymbol{X}^\top \otimes \boldsymbol{X}\boldsymbol{X}^\top\right)
\end{aligned}
$$

Luego, aplicando los teoremas 2.2.2, 2.4.2 y 2.5.1 se obtiene lo siguiente

$$
\begin{aligned}
E(\boldsymbol{S}_0 \otimes \boldsymbol{S}_0) &= \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \Bigg(3\sum_i (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ii}) + \frac{4}{\pi} \sum_{i \neq j} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})] \\
&\qquad + \sum_{i \neq j} (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji}) + \frac{4}{\pi^2} \sum_{i \neq j \neq s \neq l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sl}) \\
&\qquad + \frac{2}{\pi} \sum_{i \neq j \neq s} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ss}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{si})] \Bigg) (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \left(\left(1 - \frac{2}{\pi}\right) (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) + \frac{2}{\pi} \left(\boldsymbol{1}_m \boldsymbol{1}_m^\top \otimes \boldsymbol{\Sigma}\right) \right) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&\quad + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \left(\left(1 - \frac{2}{\pi}\right) vec\left(\boldsymbol{I}_m\right) vec(\boldsymbol{\Sigma})^\top + \frac{2}{\pi} vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top \right) \\
&\quad + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_k\right) \left(\left(1 - \frac{2}{\pi}\right) (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) \boldsymbol{K}_{mk} + \frac{2}{\pi} \left(\boldsymbol{1}_m \boldsymbol{1}_m^\top \otimes \boldsymbol{\Sigma}\right) \boldsymbol{K}_{mk} \right) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \left(\left(1 - \frac{2}{\pi}\right) \boldsymbol{K}_{mk} (\boldsymbol{I}_m \otimes \boldsymbol{\Sigma}) + \frac{2}{\pi} \boldsymbol{K}_{mk} \left(\boldsymbol{1}_m \boldsymbol{1}_m^\top \otimes \boldsymbol{\Sigma}\right) \right) (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_k) \\
&\quad + \left(\left(1 - \frac{2}{\pi}\right) vec\left(\boldsymbol{I}_m\right) vec(\boldsymbol{\Sigma})^\top + \frac{2}{\pi} vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top \right)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}^\top\right) \left(\left(1 - \frac{2}{\pi}\right) (\boldsymbol{\Sigma} \otimes \boldsymbol{I}_m) + \frac{2}{\pi} \left(\boldsymbol{\Sigma} \otimes \boldsymbol{1}_m \boldsymbol{1}_m^\top\right) \right) (\boldsymbol{I}_k \otimes \boldsymbol{\Lambda}) \\
&\quad + (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \left(\boldsymbol{I}_{k^2} + \boldsymbol{K}_{kk}\right) + vec(\boldsymbol{\Sigma}) vec(\boldsymbol{\Sigma})^\top \\
&= 3 \left(\boldsymbol{\Lambda}^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) K_{mm} \\
&\quad + \frac{4}{\pi} \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ii}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ji}) + (\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{ij})] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j} [(\boldsymbol{\Delta}_{ii} \otimes \boldsymbol{\Delta}_{jj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ji})] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \frac{2}{\pi} \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j \neq s} [(\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ss}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sj}) + (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{si})] (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \frac{4}{\pi^2} \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) \sum_{i \neq j \neq s \neq l} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{sl}) (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(1 - \frac{2}{\pi}\right) \left(\boldsymbol{\Lambda}^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) \\
&\quad + \frac{2}{\pi} \left(\boldsymbol{\Lambda}^\top \boldsymbol{1}_m \boldsymbol{1}_m^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) + \left(1 - \frac{2}{\pi}\right) \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{I}_m\right) vec(\boldsymbol{\Sigma})^\top \\
&\quad + \frac{2}{\pi} \left(\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right) vec(\boldsymbol{\Sigma})^\top + 2 \left(1 - \frac{2}{\pi}\right) \left(\boldsymbol{\Lambda}^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) K_{kk} \\
&\quad + \frac{4}{\pi} \left(\boldsymbol{\Lambda}^\top \boldsymbol{1}_m \boldsymbol{1}_m^\top \boldsymbol{\Lambda} \otimes \boldsymbol{\Sigma}\right) K_{kk} + \left(1 - \frac{2}{\pi}\right) vec(\boldsymbol{\Sigma}) vec\left(\boldsymbol{I}_m\right)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) \\
&\quad + \frac{2}{\pi} vec(\boldsymbol{\Sigma}) vec\left(\boldsymbol{1}_m \boldsymbol{1}_m^\top\right)^\top (\boldsymbol{\Lambda} \otimes \boldsymbol{\Lambda}) + \left(1 - \frac{2}{\pi}\right) \left(\boldsymbol{\Sigma} \otimes \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \\
&\quad + \frac{2}{\pi} \left(\boldsymbol{\Sigma} \otimes \boldsymbol{\Lambda}^\top \boldsymbol{1}_m \boldsymbol{1}_m^\top \boldsymbol{\Lambda}\right) + (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) \left(\boldsymbol{I}_{k^2} + \boldsymbol{K}_{kk}\right) + vec(\boldsymbol{\Sigma}) vec(\boldsymbol{\Sigma})^\top
\end{aligned}
$$

Teorema 2.5.6

Sea $\boldsymbol{Y}_0 \sim SNG_{d,1}(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ donde

$$f(\boldsymbol{y}_0) = 2\phi_d(\boldsymbol{y}_0; \boldsymbol{\Omega})\Phi(\boldsymbol{\eta}^\top \boldsymbol{y}_0)$$

y sea $\boldsymbol{\eta} = \boldsymbol{\omega}^{-1}\boldsymbol{\alpha}$, $\zeta_0(x) = \log(2\Phi(x))$, y ζ_1 y ζ_2 es la primera y la segunda derivada de ζ_0, repectivamente.

Entonces, siempre que las esperanzas existan,

$$E([\zeta_1(\boldsymbol{\eta}^\top \boldsymbol{Y}_0)]^s g(\boldsymbol{Y}_0)) = c_s E\left(\frac{g(\boldsymbol{W}_s)}{[\Phi(\boldsymbol{\eta}^\top \boldsymbol{W}_s)]^{s-1}}\right), \qquad (s = 1, 2, ...)$$

donde

$$c_s = \frac{2}{(2\pi)^{s/2}\sqrt{1 + s\boldsymbol{\eta}^\top \boldsymbol{\Omega}\boldsymbol{\eta}}}$$

y $\boldsymbol{W}_s \sim N_d(\mathbf{0}, \boldsymbol{\Omega}_s)$ donde

$$\boldsymbol{\Omega}_s = (\boldsymbol{\Omega}^{-1} + s\boldsymbol{\eta}\boldsymbol{\eta}^\top)^{-1} = \boldsymbol{\Omega} - s(1 + s\boldsymbol{\eta}^\top \boldsymbol{\Omega}\boldsymbol{\eta})^{-1}\boldsymbol{\Omega}\boldsymbol{\eta}\boldsymbol{\eta}^\top \boldsymbol{\Omega}$$

Por lo tanto, si $\boldsymbol{\eta} = \frac{\boldsymbol{\lambda}}{\boldsymbol{\sigma}}\frac{1}{\sqrt{\boldsymbol{\sigma}^2 + \boldsymbol{\lambda}^2}}$, $\boldsymbol{\Omega} = \boldsymbol{\sigma}^2 + \boldsymbol{\lambda}^2$ y $\boldsymbol{Y}_0 = \boldsymbol{y} - \mu$

$$\begin{aligned}
E(\zeta_2(\boldsymbol{\eta}\boldsymbol{Y}_0)) &= -c_2 a_0 \\
E(\zeta_1(\boldsymbol{\eta}\boldsymbol{Y}_0)) &= c_1 \\
E(\zeta_2(\boldsymbol{\eta}\boldsymbol{Y}_0)\boldsymbol{Y}_0) &= -c_1\frac{\boldsymbol{\sigma}^2\boldsymbol{\lambda}}{(\boldsymbol{\sigma}^2)^{1/2}(\boldsymbol{\sigma}^2 + \boldsymbol{\lambda}^2)^{1/2}} - c_2 a_1 \\
E(\zeta_1(\boldsymbol{\eta}\boldsymbol{Y}_0)\boldsymbol{Y}_0) &= 0 \\
E(\zeta_2(\boldsymbol{\eta}\boldsymbol{Y}_0)\boldsymbol{Y}_0^2) &= -c_2 A_2
\end{aligned}$$

donde

$$\begin{aligned}
a_0 &= E\left(\frac{1}{\Phi(\boldsymbol{\eta}^\top \boldsymbol{W}_2)}\right) \\
a_1 &= E\left(\frac{\boldsymbol{W}_2}{\Phi(\boldsymbol{\eta}^\top \boldsymbol{W}_2)}\right) \\
A_2 &= E\left(\frac{\boldsymbol{W}_2^2}{\Phi(\boldsymbol{\eta}^\top \boldsymbol{W}_2)}\right)
\end{aligned}$$

Note que es justamente el modelo Skew-Normal de Azzalini y Dalla-Valle (1996), el cual corresponde a un caso particualar del modelo Skew-Normal Generalizado.

Dem. Ver Arellano-Valle & Azzalini (2008).

Capítulo 3

Aspectos inferenciales

Ahora que ya se presentaron algunas propiedades de la familia de distribuciones utilizado en este trabajo, se muestran a continuación algunos aspectos inferenciales de este modelo.

Se calcula en este capítulo, la Matriz de Información de este modelo, la que no ha sido calculada anteriormente y puede resultar muy útil para algunas técnicas de estimación ya sea clásicas o Bayesiana.

Como se ha señalado anteriormente, este modelo es muy utilizado para aplicaciones y estimaciones con enfoque Bayesiano, por lo que resulta muy útil conocer las propiedades mostradas en el capítulo anterior, así como también es de interés discutir algunos aspectos inferenciales del mismo.

Es por esto, que en principio se presenta el desarrollo del cálculo de la Matriz de Información para el caso univariado, el que es equivalente al modelo planteado por Arellano-Valle y Azzalini (2008). Luego, se calcula la Matriz de Información para el caso multivariado. Finalmente, se obtiene el caso univariado en base a los resultados obtenidos para el caso multivariado.

3.1. Matriz de Información: Caso Univariado

El objetivo de esta sección es calcular la Matriz de Información para la distribución skew-normal, con la reparametrización planteada en la sección 2.5, para el caso univariado. La razón por la que se comenzará con este caso particular, es que se tiene el resultado de esta matriz con la parametrización de Arellano-Valle y Azzalini (2008), y es en el caso univariado en donde estas distruciones son equivalentes.

Al finalizar esta sección se hará la transformación de la matriz obtenida a la parametrización de Arellano-Valle y Azzalini (2008) y se confirmará su igualdad.

Teorema 3.1
Sea $Y \sim SNG_{1,1}(\mu, \sigma^2, \lambda)$, el caso univariado de la distribución Skew-Nomal presentado en la sección 2.5. Entonces, la Matriz de Información es:

$$I(\mu, \sigma^2, \lambda) \quad = \quad \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix}$$

donde,

$$\begin{aligned}
I_{11}(\nu) &= \frac{1}{\sigma^2+\lambda^2} + c_2 a_0 \frac{\lambda^2}{\sigma^2(\sigma^2+\lambda^2)} \\
I_{22}(\nu) &= \frac{1}{2}\frac{1}{(\sigma^2+\lambda^2)^2} + \frac{1}{4} c_2 A_2 \frac{\lambda^2(2\sigma^2+\lambda^2)^2}{(\sigma^2)^3(\sigma^2+\lambda^2)^3} \\
I_{33}(\nu) &= 2\frac{\lambda^2}{(\sigma^2+\lambda^2)^2} + c_2 A_2 \frac{\sigma^2}{(\sigma^2+\lambda^2)^3} \\
I_{12}(\nu) &= \left(\frac{2}{\pi}\right)^{1/2} \frac{\lambda}{(\sigma^2+\lambda^2)^2} - \frac{1}{2} c_1 \frac{\lambda(2\sigma^2+\lambda^2)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}} + \frac{1}{2} c_2 a_1 \frac{\lambda^2(2\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^2} \\
I_{13}(\nu) &= 2\left(\frac{2}{\pi}\right)^{1/2} \frac{\lambda^2}{(\sigma^2+\lambda^2)^2} + c_1 \frac{(\sigma^2)^{3/2}}{(\sigma^2+\lambda^2)^{5/2}} - c_2 a_1 \frac{\lambda}{(\sigma^2+\lambda^2)^2} \\
I_{23}(\nu) &= \frac{\lambda}{(\sigma^2+\lambda^2)^2} - \frac{1}{2} c_2 A_2 \frac{\lambda(2\sigma^2+\lambda^2)}{\sigma^2(\sigma^2+\lambda^2)^3}
\end{aligned}$$

y

$$
\begin{aligned}
c_1 &= \left(\frac{2}{\pi}\right)^{1/2}\left(\frac{\sigma^2}{\sigma^2+\lambda^2}\right)^{1/2} \\
c_2 &= \left(\frac{1}{\pi}\right)^{1/2}\left(\frac{\sigma^2}{\sigma^2+2\lambda^2}\right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
a_0 &= E\left(\Phi^{-1}\left(\frac{\lambda W}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\right) \\
a_1 &= E\left(W\Phi^{-1}\left(\frac{\lambda W}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\right) \\
A_2 &= E\left(W^2\Phi^{-1}\left(\frac{\lambda W}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\right)
\end{aligned}
$$

con $W \sim N\left(0, \frac{\sigma^2(\sigma^2+\lambda^2)}{\sigma^2+2\lambda^2}\right)$

Dem. Sea $Y \sim SNG_{1,1}(\mu, \sigma^2, \lambda)$, se tiene que

$$f(y|\mu,\sigma^2,\lambda) \quad = \quad 2\phi\left(y-\mu \middle| \sigma^2+\lambda^2\right)\Phi\left(\frac{\lambda}{\sigma^2+\lambda^2}(y-\mu)\middle| 1-\frac{\lambda^2}{\sigma^2+\lambda^2}\right)$$

Si llamamos $\zeta_0(x) = log(\Phi(x))$ entonces la log-verosimilitud es:

$$l(\mu,\sigma^2,\lambda) = cte - \frac{1}{2}log(\sigma^2+\lambda^2) - \frac{1}{2}\frac{(y-\mu)^2}{\sigma^2+\lambda^2} + \zeta_0\left(\frac{\lambda}{\sigma}\frac{y-\mu}{\sqrt{\sigma^2+\lambda^2}}\right)$$

Si llamamos $\nu = (\mu, \sigma^2, \lambda)$ y ζ_1 es la derivada de ζ_0, se tiene que

$$\begin{aligned}
dl(\nu) \quad = \quad & -\frac{1}{2(\sigma^2+\lambda^2)}(d\sigma+2\lambda d\lambda) - \frac{(y-\mu)}{\sigma^2+\lambda^2}(-d\mu) + \frac{(y-\mu)^2}{2(\sigma^2+\lambda^2)^2}(d\sigma+2\lambda d\lambda) \\
& +\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[\frac{(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\lambda) + \frac{\lambda}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(-d\mu)\right. \\
& \left. -\frac{1}{2}\frac{\lambda(y-\mu)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{1/2}}(d\sigma^2) - \frac{1}{2}\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2+2\lambda d\lambda)\right] \\
= \quad & -\frac{1}{2(\sigma^2+\lambda^2)}(d\sigma^2) - \frac{\lambda}{\sigma^2+\lambda^2}(d\lambda) + \frac{(y-\mu)}{\sigma^2+\lambda^2}(d\mu) + \frac{(y-\mu)^2}{2(\sigma^2+\lambda^2)^2}(d\sigma^2) + \frac{\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^2}(d\lambda) \\
& +\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[\frac{(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\lambda) - \frac{\lambda}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu)\right. \\
& -\frac{\lambda(y-\mu)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{1/2}}(d\sigma^2) - \frac{\lambda(y-\mu)}{2(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2) \\
& \left. -\frac{\lambda^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)\right] \\
= \quad & \frac{y-\mu}{\sigma^2+\lambda^2}(d\mu) + \left[\frac{(y-\mu)^2}{2(\sigma^2+\lambda^2)^2} - \frac{1}{2(\sigma^2+\lambda^2)}\right](d\sigma^2) + \left[\frac{\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^2} - \frac{\lambda}{\sigma^2+\lambda^2}\right](d\lambda) \\
& +\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[-\frac{\lambda}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu) - \frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)\right. \\
& \left. +\frac{\sigma^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)\right]
\end{aligned}$$

Ahora, si ζ_2 es la segunda derivada de ζ_0, al calcular la segunda derivada se tiene que

$$
\begin{aligned}
d^2 l(\nu) = & -\frac{1}{\sigma^2+\lambda^2}(d\mu)^2 - \frac{y-\mu}{(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2+2\lambda d\lambda) + \frac{y-\mu}{(\sigma^2+\lambda^2)^2}(-d\mu)(d\sigma^2) \\
& -\frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\sigma+2\lambda d\lambda) + \frac{1}{2(\sigma^2+\lambda^2)^2}(d\sigma^2)(d\sigma^2+2\lambda d\lambda) + \frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^2}(d\lambda)^2 \\
& +\frac{2\lambda(y-\mu)}{(\sigma^2+\lambda^2)^2}(d\lambda)(-d\mu) - \frac{2\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^2}(d\lambda)(d\sigma^2+2\lambda d\lambda) \\
& -\frac{1}{\sigma^2+\lambda^2}(d\lambda^2) + \frac{\lambda}{(\sigma^2+\lambda^2)^2}(d\lambda)(d\sigma^2+2\lambda d\lambda) \\
& +\zeta_2\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[-\frac{\lambda}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu) - \frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)\right. \\
& \qquad\qquad \left. +\frac{\sigma^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)\right]^2 \\
& +\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[-\frac{1}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu)(d\lambda) + \frac{\lambda}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu)(d\sigma^2)\right. \\
& \qquad +\frac{\lambda}{2(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\sigma^2+2\lambda d\lambda) - \frac{(y-\mu)(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)(d\sigma^2) \\
& \qquad -\frac{\lambda(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)(-d\mu) - \frac{\lambda(y-\mu)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(2d\sigma^2+2\lambda d\lambda)(d\sigma^2) \\
& \qquad +\frac{3}{4}\frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^{5/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)^2 + \frac{3}{4}\frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{5/2}}(d\sigma^2)(d\sigma^2+2\lambda d\lambda) \\
& \qquad +\frac{y-\mu}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)(d\sigma^2) + \frac{\sigma^2}{(\sigma^{1/2})(\sigma^2+\lambda^2)^{3/2}}(d\lambda)(-d\mu) \\
& \qquad \left. -\frac{\sigma^2}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)(d\sigma^2) - \frac{3}{2}\frac{\sigma^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}}(d\lambda)(d\sigma^2+2\lambda d\lambda)\right]
\end{aligned}
$$

$$
\begin{aligned}
d^2l(\nu) = &-\frac{1}{\sigma^2+\lambda^2}(d\mu)^2 - \frac{y-\mu}{(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2) - \frac{2\lambda(y-\mu)}{(\sigma^2+\lambda^2)}(d\mu)(d\lambda) - \frac{y-\mu}{(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma)^2 \\
&-\frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^3}(d\sigma^2)^2 - \frac{2\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\lambda) + \frac{1}{2(\sigma^2+\lambda^2)^2}(d\sigma^2)^2 + \frac{\lambda}{(\sigma^2+\lambda^2)^2}(d\sigma^2)(d\lambda) \\
&+\frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^2}(d\lambda)^2 - \frac{2\lambda(y-\mu)}{(\sigma^2+\lambda^2)^2}(d\mu)(d\lambda) - \frac{1}{\sigma^2+\lambda^2}(d\lambda)^2 + \frac{\lambda}{(\sigma^2+\lambda^2)^2}(d\sigma^2)(d\lambda) \\
&+\frac{2\lambda^2}{(\sigma^2+\lambda^2)^2}(d\lambda)^2 - \frac{2\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\lambda) - \frac{4\lambda^2(y-\mu)^2}{(\sigma^2+\lambda^2)^3}(d\lambda)^2 \\
&+\zeta_2\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\Bigg[\frac{\lambda^2}{\sigma^2(\sigma^2+\lambda^2)}(d\mu)^2 + \frac{\lambda^2(y-\mu)^2(2\sigma^2+\lambda^2)}{4(\sigma^2)^3(\sigma^2+\lambda^2)^3}(d\sigma^2)^2 \\
&\qquad +\frac{(\sigma^2)^2(y-\mu)^2}{\sigma^2(\sigma^2+\lambda^2)^3}(d\lambda)^2 + \frac{2\lambda^2(y-\mu)(\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2) \\
&\qquad -\frac{2\lambda\sigma^2(y-\mu)}{\sigma^2(\sigma^2+\lambda^2)^2}(d\mu)(d\lambda) - \frac{2\lambda\sigma^2(y-\mu)^2(\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\lambda)\Bigg] \\
&+\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\Bigg[-\frac{1}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu)(d\lambda) + \frac{\lambda}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{1/2}}(d\mu)(d\sigma^2) \\
&\qquad +\frac{\lambda}{2(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\sigma^2) + \frac{\lambda^2}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\lambda) \\
&\qquad -\frac{(y-\mu)(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\lambda)(d\sigma^2) + \frac{\lambda(2\sigma^2+\lambda^2)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\sigma^2) \\
&\qquad -\frac{\lambda(y-\mu)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)^2 - \frac{\lambda^2(y-\mu)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)(d\lambda) \\
&\qquad +\frac{3}{4}\frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^{5/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)^2 + \frac{3}{4}\frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{5/2}}(d\sigma^2)^2 \\
&\qquad +\frac{3}{2}\frac{\lambda^2(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{5/2}}(d\sigma^2)(d\lambda) + \frac{y-\mu}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)(d\lambda) \\
&\qquad -\frac{\sigma^2}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\lambda) - \frac{\sigma^2(y-\mu)}{2(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\sigma^2)(d\lambda) \\
&\qquad -\frac{3}{2}\frac{\sigma^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}}(d\sigma^2)(d\lambda) - \frac{3\lambda\sigma^2(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}}(d\lambda)^2\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
d^2 l(\nu) = & \left[-\frac{1}{\sigma^2+\lambda^2}\right](d\mu)^2 + \left[\frac{1}{2(\sigma^2+\lambda^2)^2} - \frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^3}\right](d\sigma^2)^2 \\
& + \left[\frac{(y-\mu)^2}{(\sigma^2+\lambda^2)^2} + \frac{2\lambda^2}{(\sigma^2+\lambda^2)^2} - \frac{1}{\sigma^2+\lambda^2} - \frac{4\lambda^2(y-\mu)^2}{(\sigma^2+\lambda^2)^3}\right](d\lambda)^2 + \left[-\frac{2(y-\mu)}{(\sigma^2+\lambda^2)^2}\right](d\mu)(d\sigma^2) \\
& + \left[-\frac{4\lambda(y-\mu)}{(\sigma^2+\lambda^2)^2}\right](d\mu)(d\lambda) + \left[\frac{2\lambda}{(\sigma^2+\lambda^2)^2} - \frac{2\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^3} - \frac{2\lambda(y-\mu)^2}{(\sigma^2+\lambda^2)^3}\right](d\sigma^2)(d\lambda) \\
& +\zeta_2\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[\frac{\lambda^2}{\sigma^2(\sigma^2+\lambda^2)}(d\mu)^2 + \frac{\lambda^2(y-\mu)^2(2\sigma^2+\lambda^2)^2}{4(\sigma^2)^3(\sigma^2+\lambda^2)^3}(d\sigma^2)^2\right. \\
& \qquad + \frac{(\sigma^2)^2(y-\mu)^2}{\sigma^2(\sigma^2+\lambda^2)^3}(d\lambda)^2 + \frac{\lambda^2(y-\mu)(2\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2) \\
& \qquad \left. - \frac{2\lambda(y-\mu)}{(\sigma^2+\lambda^2)^2}(d\mu)(d\lambda) - \frac{\lambda(y-\mu)^2(2\sigma^2+\lambda^2)}{\sigma^2(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\lambda)\right] \\
& +\zeta_1\left(\frac{\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\left[\left(\frac{3}{4}\frac{\lambda(y-\mu)(2\sigma^2+\lambda^2)^2}{(\sigma^2)^{5/2}(\sigma^2+\lambda^2)^{5/2}} - \frac{\lambda(y-\mu)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}\right)(d\sigma^2)^2\right. \\
& \qquad + \left(-\frac{3\sigma^2\lambda(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}}\right)(d\lambda)^2 + \left(-\frac{2\sigma^2}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{3/2}}\right)(d\mu)(d\lambda) \\
& \qquad \left. + \left(\frac{\lambda(2\sigma^2+\lambda^2)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}\right)(d\mu)(d\sigma^2) + \left(-\frac{(2\sigma^2+\lambda^2)(y-\mu)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}}\right)(d\sigma^2)(d\lambda)\right]
\end{aligned}
$$

Para el cálculo de la esparanza ocuparemos algunas de las propiedades enunciadas en el capítulo 2.

Así, tomando es cuenta que $\zeta_2(x) = -\zeta_1(x)(x + \zeta_1(x))$,

$$\begin{aligned}
E(-d^2(l(\nu))) &= \frac{1}{\sigma^2+\lambda^2}(d\mu)^2 + \frac{1}{2(\sigma^2+\lambda^2)^2}(d\sigma^2)^2 + \frac{2\lambda^2}{(\sigma^2+\lambda^2)^2}(d\lambda)^2 \\
&+2\left(\frac{2}{\pi}\right)^{1/2}\frac{\lambda}{(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2) + 4\left(\frac{2}{\pi}\right)^{1/2}\frac{\lambda^2}{(\sigma^2+\lambda^2)^2}(d\mu)(d\lambda) \\
&+c_2a_0\frac{\lambda^2}{\sigma^2(\sigma^2+\lambda^2)}(d\mu)^2 + \frac{2\lambda}{(\sigma^2+\lambda^2)^2}(d\sigma^2)(d\lambda) \\
&+c_2A_2\left(\frac{\lambda^2(2\sigma^2+\lambda^2)^2}{4(\sigma^2)^3(\sigma^2+\lambda^2)^3}(d\sigma^2)^2 + \frac{\sigma^2}{(\sigma^2+\lambda^2)^3}(d\lambda)^2 - \frac{\lambda(2\sigma^2+\lambda^2)}{\sigma^2(\sigma^2+\lambda^2)^3}(d\sigma^2)(d\lambda)\right) \\
&+\left(c_1\frac{\sigma^2\lambda}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}} + c_2a_1\right) \\
&\quad\left(\frac{\lambda^2(2\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^2}(d\mu)(d\sigma^2) - \frac{2\lambda}{(\sigma^2+\lambda^2)^2}(d\mu)(d\lambda)\right) \\
&-c_1\left(-\frac{2(\sigma^2)^{1/2}}{(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\lambda) + \frac{\lambda(2\sigma^2+\lambda^2)}{(\sigma^2)^{3/2}(\sigma^2+\lambda^2)^{3/2}}(d\mu)(d\sigma^2)\right)
\end{aligned}$$

Finalmente, las expresiones de las esperanzas que componen la Matriz de Información son:

$$\begin{aligned}
I_{11}(\nu) &= \frac{1}{\sigma^2+\lambda^2} + c_2a_0\frac{\lambda^2}{\sigma^2(\sigma^2+\lambda^2)} \\
I_{22}(\nu) &= \frac{1}{2}\frac{1}{(\sigma^2+\lambda^2)^2} + \frac{1}{4}c_2A_2\frac{\lambda^2(2\sigma^2+\lambda^2)^2}{(\sigma^2)^3(\sigma^2+\lambda^2)^3} \\
I_{33}(\nu) &= 2\frac{\lambda^2}{(\sigma^2+\lambda^2)^2} + c_2A_2\frac{\sigma^2}{(\sigma^2+\lambda^2)^3} \\
I_{12}(\nu) &= \left(\frac{2}{\pi}\right)^{1/2}\frac{\lambda}{(\sigma^2+\lambda^2)^2} - \frac{1}{2}c_1\frac{\lambda(2\sigma^2+\lambda^2)}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{5/2}} + \frac{1}{2}c_2a_1\frac{\lambda^2(2\sigma^2+\lambda^2)}{(\sigma^2)^2(\sigma^2+\lambda^2)^2} \\
I_{13}(\nu) &= 2\left(\frac{2}{\pi}\right)^{1/2}\frac{\lambda^2}{(\sigma^2+\lambda^2)^2} + c_1\frac{(\sigma^2)^{3/2}}{(\sigma^2+\lambda^2)^{5/2}} - c_2a_1\frac{\lambda}{(\sigma^2+\lambda^2)^2} \\
I_{23}(\nu) &= \frac{\lambda}{(\sigma^2+\lambda^2)^2} - \frac{1}{2}c_2A_2\frac{\lambda(2\sigma^2+\lambda^2)}{\sigma^2(\sigma^2+\lambda^2)^3}
\end{aligned}$$

En la sección 5.2 se comparan las matrices de información del modelo planteado por Arellano-Valle & Azzalini (2008) contra el utilizado en este trabajo (en el caso univariado), ya que para este caso, los modelos son equivalentes.

3.2. Matriz de Información: Caso Multivariado

En esta sección se calculará la Matriz de Información usando la parametrización señalada en la sección 2.5 en el caso multivariado. Para esto, se calculará el cuadrado de la primera derivada de la log-verosimilitud, para luego calcular la esperanza de esta expresión.

Teorema 3.2

Sea $Y \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$. Entonces, la Matriz de Información es:

$$\boldsymbol{I}(\boldsymbol{\mu}, v(\boldsymbol{\Sigma}), vec(\boldsymbol{\Lambda})) = \begin{pmatrix} \boldsymbol{I}_{11} & \boldsymbol{I}_{12} & \boldsymbol{I}_{13} \\ \boldsymbol{I}_{12} & \boldsymbol{I}_{22} & \boldsymbol{I}_{23} \\ \boldsymbol{I}_{13} & \boldsymbol{I}_{23} & \boldsymbol{I}_{33} \end{pmatrix}$$

donde,

$$\begin{aligned} \boldsymbol{I}_{11} &= \left(1 - \frac{2}{\pi}\right) (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda} (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} + \frac{2}{\pi} (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top \mathbf{1}_m \mathbf{1}_m^\top \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \\ &+ \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Sigma} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} - (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{R}_1 \boldsymbol{\Lambda} (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \\ &+ (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top \boldsymbol{R}_2 \boldsymbol{\Lambda} (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{I}_{22} \quad = \quad & \frac{1}{4} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right)^{\top} D \\
& +\frac{1}{4} D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) E_1\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& +\frac{1}{4} D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \boldsymbol{R}_3\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& +D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^{\top}\right) \boldsymbol{R}_4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\right) D \\
& -\frac{1}{2}\left(1-\frac{2}{\pi}\right) D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\frac{1}{\pi} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\boldsymbol{\Lambda}^{\top}\mathbf{1}_m\mathbf{1}_m^{\top}\boldsymbol{\Lambda}\right)^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\frac{1}{2} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{\Sigma})^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\frac{1}{2} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_5)^{\top}\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& +D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_1)^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\right) D \\
& +\frac{1}{2} D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_6\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_7\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\right) D \\
& -D^{\top}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \boldsymbol{R}_8\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\right) D
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}_{33} \quad = \quad & \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(I_k)vec(I_k)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)E_1\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}\right)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_3\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\boldsymbol{R}_3\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\boldsymbol{R}_4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(1-\frac{2}{\pi}\right)\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec\left(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&-\frac{4}{\pi}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec\left(\boldsymbol{\Lambda}^\top\mathbf{1}_m\mathbf{1}_m^\top\boldsymbol{\Lambda}\right)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec\left(\boldsymbol{\Sigma}\right)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec(\boldsymbol{R}_5)^\top\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec(\boldsymbol{R}_5)^\top\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec(\boldsymbol{R}_1)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k)vec(\boldsymbol{R}_1)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)\boldsymbol{R}_6\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)\boldsymbol{R}_6\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)\boldsymbol{R}_7\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&-4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)\boldsymbol{R}_7\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_3\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_8\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_8\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&+2\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\boldsymbol{R}_8\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&-4\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\boldsymbol{R}_8\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&-4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\boldsymbol{R}_4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}_{12} \;=\; & -\frac{1}{2}\left(\frac{2}{\pi}\right)^{1/2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\mathbf{1}_m vec\left(\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)^\top D\right. \\
& +\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}E_2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)D \\
& +\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_9\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)D \\
& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_{10}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\right)D \\
& +\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{11}vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)D \\
& -\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{12}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)D \\
& -\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{13}\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)D \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{14}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\right)D
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}_{13} \;=\; & -\left(\frac{2}{\pi}\right)^{1/2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\mathbf{1}_m vec(\boldsymbol{I}_k)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}^\top\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}E_2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_9\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_9\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_{10}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& -2\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{R}_{10}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{11}vec(\boldsymbol{I}_k)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}^\top\right) \\
& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{12}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{13}\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{13}\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{14}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& +2\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{R}_{14}\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}_{23} \quad = \quad & \frac{1}{2} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2}\left(1-\frac{2}{\pi}\right) D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{\pi} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\boldsymbol{\Lambda}^\top\boldsymbol{1}_m\boldsymbol{1}_m^\top\boldsymbol{\Lambda}\right)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{\Sigma})^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& +\frac{1}{2} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_5)^\top \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& -\frac{1}{2} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_5)^\top \left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_1)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& +D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{R}_1)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2}\left(1-\frac{2}{\pi}\right) D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec\left(\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{\pi} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{\Lambda}^\top\boldsymbol{1}_m\boldsymbol{1}_m^\top\boldsymbol{\Lambda}) vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{\Sigma}) vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) E_1 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_6 \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_6 \left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_7 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& -D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \boldsymbol{R}_7 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) vec(\boldsymbol{R}_5^\top) vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_6^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& -\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_3 \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_3 \left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \\
& +\frac{1}{2} D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_8 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right)
\end{aligned}
$$

$$
\begin{aligned}
&-D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) A_8 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&+D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) vec(\boldsymbol{R}_1)vec(\boldsymbol{I}_k)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^\top\right)\\
&-D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_7^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_8^\top \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right)\\
&-D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_4 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right)\\
&-D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_8^\top \left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+2D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) \boldsymbol{R}_4 \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)
\end{aligned}
$$

donde

$$\begin{aligned} E_1 &= E\left(\boldsymbol{S}_0 \otimes \boldsymbol{S}_0\right) \\ E_2 &= E\left(\boldsymbol{S}_0 \otimes \boldsymbol{y}_0^\top\right) \end{aligned}$$

fueron definidos en el Teorema 2.6.1, y

$$\begin{aligned}
\boldsymbol{R}_1 &= E\left(\Psi(\boldsymbol{\eta})\boldsymbol{y}_0\boldsymbol{q}_m^\top\right) \\
\boldsymbol{R}_2 &= E\left(\Psi^2(\boldsymbol{\eta})\boldsymbol{q}_m\boldsymbol{q}_m^\top\right) \\
\boldsymbol{R}_3 &= E\left(\Psi^2(\boldsymbol{\eta})vec(Q_m)vec(Q_m)^\top\right) \\
\boldsymbol{R}_4 &= E\left(\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{S}_0 \otimes \boldsymbol{q}_m\boldsymbol{q}_m^\top\right)\right) \\
\boldsymbol{R}_5 &= E\left(\Psi(\boldsymbol{\eta})Q_m\right) \\
\boldsymbol{R}_6 &= E\left(\Psi(\boldsymbol{\eta})vec(\boldsymbol{S}_0)vec(Q_m)^\top\right) \\
\boldsymbol{R}_7 &= E\left(\Psi(\boldsymbol{\eta})\left(\boldsymbol{S}_0 \otimes \boldsymbol{y}_0\boldsymbol{q}_m^\top\right)\right) \\
\boldsymbol{R}_8 &= E\left(\Psi^2(\boldsymbol{\eta})vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)^\top\right) \\
\boldsymbol{R}_9 &= E\left(\Psi(\boldsymbol{\eta})\boldsymbol{y}_0 vec(Q_m)^\top\right) \\
\boldsymbol{R}_{10} &= E\left(\Psi(\boldsymbol{\eta})\left(\boldsymbol{S}_0 \otimes \boldsymbol{q}_m^\top\right)\right) \\
\boldsymbol{R}_{11} &= E\left(\Psi(\boldsymbol{\eta})\boldsymbol{q}_m\right) \\
\boldsymbol{R}_{12} &= E\left(\Psi(\boldsymbol{\eta})\boldsymbol{q}_m vec(\boldsymbol{S}_0)^\top\right) \\
\boldsymbol{R}_{13} &= E\left(\Psi^2(\boldsymbol{\eta})\boldsymbol{q}_m vec(Q_m)^\top\right) \\
\boldsymbol{R}_{14} &= E\left(\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{y}_0^\top \otimes \boldsymbol{q}_m\boldsymbol{q}_m^\top\right)\right)
\end{aligned}$$

Dem. Sea $Y \sim SNG_{k,m}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda})$, se tiene que

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}) \quad = \quad 2^m \phi_k \left(\boldsymbol{y} - \boldsymbol{\mu}|\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right) \Phi_m \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}(\boldsymbol{y} - \boldsymbol{\mu})|\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)$$

Denotando $\boldsymbol{y}_0 = \boldsymbol{y} - \boldsymbol{\mu}$ y $\boldsymbol{S}_0 = \boldsymbol{y}_0 \boldsymbol{y}_0^\top$, la log-verosimilitud es:

$$\begin{aligned} l(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}) \quad = \quad & cte - \frac{1}{2} log|\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}| - \frac{1}{2} tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{S}_0\right\} \\ & + log \left(\Phi_m \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right) \end{aligned}$$

y la primera derivada de la log-verosimilitud es

$$\begin{aligned} dl(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Lambda}) \quad = \quad & -\frac{1}{2} tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}(d(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}))\right\} \\ & -\frac{1}{2} tr \left\{-(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}(d(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}))(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{S}_0 - (\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \left((d\boldsymbol{\mu})\boldsymbol{y}_0^\top + \boldsymbol{y}_0 (d\boldsymbol{\mu})^\top\right)\right\} \\ & + \Phi_m^{-1} \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) \\ & \ d \left(\Phi_m \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right) \\ = \quad & -\frac{1}{2} tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}(d\boldsymbol{\Sigma})\right\} - tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top (d\boldsymbol{\Lambda})\right\} \\ & + \frac{1}{2} tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}(d\boldsymbol{\Sigma})(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{S}_0\right\} + tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^\top (d\boldsymbol{\Lambda})(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{S}_0\right\} \\ & + tr \left\{(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 (d\boldsymbol{\mu})^\top\right\} \\ & + \Phi_m^{-1} \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) \\ & \ d \left(\Phi_m \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right) \\ = \quad & -\frac{1}{2} vec(d\boldsymbol{\Sigma})^\top vec \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1}\right) \\ & - vec(d\boldsymbol{\Lambda})^\top \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k) \\ & + \frac{1}{2} vec(d\boldsymbol{\Sigma})^\top \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0) \\ & + vec(d\boldsymbol{\Lambda})^\top \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0) \\ & + (d\boldsymbol{\mu})^\top \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 \\ & + \Phi_m^{-1} \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right) \\ & \ d \left(\Phi_m \left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 |\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^\top \boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right) \end{aligned}$$

Para calcular $d\left(\Phi_m\left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{y}_0|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)\right)$ usamos el teorema 2.2.4 con $\boldsymbol{\alpha}=\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0$, $\boldsymbol{\beta}=\boldsymbol{0}$ y $\boldsymbol{\Gamma}=\boldsymbol{I}_m-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top$ con lo que se obtiene que

$$\begin{aligned}
d\boldsymbol{\alpha} &= (d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\left(d\boldsymbol{\Sigma}+2\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\right)\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0\\
&\quad -\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})\\
&= (d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0\\
&\quad -2\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})\\
d\boldsymbol{\Gamma} &= -(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\left(d\boldsymbol{\Sigma}+2\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\right)\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\\
&\quad -\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Lambda})^\top\\
&= -2(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top+\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\\
&\quad +2\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top
\end{aligned}$$

y

$$\begin{aligned}
\boldsymbol{q}_m &= \Phi'_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0\Big|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
\boldsymbol{Q}_m &= \Phi''_m\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0\Big|\boldsymbol{0},\boldsymbol{I}_m-\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)
\end{aligned}$$

Por lo tanto,

$$
\begin{aligned}
d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= \frac{1}{2}vec\Bigg(-2(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top + \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \\
&\quad +2\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\Bigg)^\top vec(Q_m) \\
&\quad +\boldsymbol{q}_m^\top\Bigg((d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 - \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 \\
&\quad -2\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 - \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})\Bigg) \\
&= -vec\left((d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)^\top vec(Q_m) \\
&\quad +\frac{1}{2}vec\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)^\top vec(Q_m) \\
&\quad +vec\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)^\top vec(Q_m) \\
&\quad +\boldsymbol{q}_m^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 - \boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\Sigma})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 \\
&\quad -2\boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top(d\boldsymbol{\Lambda})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 - \boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu}) \\
&= -vec\left((d\boldsymbol{\Lambda})\right)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)vec(Q_m) \\
&\quad +\frac{1}{2}vec(d\boldsymbol{\Sigma})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
&\quad +vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
&\quad +vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
&\quad -vec(d\boldsymbol{\Sigma})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
&\quad -2vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) - \boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})
\end{aligned}
$$

$$
\begin{aligned}
d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right) &= -vec\left((d\boldsymbol{\Lambda})\right)^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_m\right) vec(Q_m)\\
&+\frac{1}{2}d(v(\boldsymbol{\Sigma}))^\top D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) vec(Q_m)\\
&+vec(d\boldsymbol{\Lambda})^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) vec(Q_m)\\
&+vec(d\boldsymbol{\Lambda})^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)\\
&-d(v(\boldsymbol{\Sigma}))^\top D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)\\
&-2vec(d\boldsymbol{\Lambda})^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) - \boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})
\end{aligned}
$$

donde D es la matriz de Duplicación definida en el capítulo 2.

Sigue que

$$
\begin{aligned}
\frac{d}{d\boldsymbol{\mu}}\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right) &= -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{q}_m\\
\frac{d}{dv(\boldsymbol{\Sigma})}\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right) &= \frac{1}{2}D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) vec(Q_m)\\
&-D^\top \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)\\
\frac{d}{d(vec(\boldsymbol{\Lambda}))}\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right) &= -\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \boldsymbol{I}_m\right) vec(Q_m)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right) vec(Q_m)\\
&+\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)\\
&-2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right) vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)
\end{aligned}
$$

con $\boldsymbol{\alpha} = \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0$, $\boldsymbol{\beta} = \boldsymbol{0}$ y $\boldsymbol{\Gamma} = \boldsymbol{I}_m - \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top$

Por lo que, si llamamos $\Psi(\boldsymbol{\eta}) = \Phi_m^{-1}\left(\boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{y}_0|\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{\Lambda}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}^\top\right)$, se tiene que

$$
\begin{aligned}
dl(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Lambda}) = & -\frac{1}{2}d(v(\boldsymbol{\Sigma}))^\top D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \\
& -vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k) \\
& +\frac{1}{2}d(v(\boldsymbol{\Sigma}))^\top D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)vec(\boldsymbol{S}_0) \\
& +vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)vec(\boldsymbol{S}_0) \\
& +(d\boldsymbol{\mu})^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 \\
& -\Psi(\boldsymbol{\eta})vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)vec(Q_m) \\
& +\frac{1}{2}\Psi(\boldsymbol{\eta})d(v(\boldsymbol{\Sigma}))^\top D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
& +\Psi(\boldsymbol{\eta})vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
& +\Psi(\boldsymbol{\eta})vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
& -\Psi(\boldsymbol{\eta})d(v(\boldsymbol{\Sigma}))^\top D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
& -2\Psi(\boldsymbol{\eta})vec(d\boldsymbol{\Lambda})^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)-\Psi(\boldsymbol{\eta})\boldsymbol{q}_m^\top\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}(d\boldsymbol{\mu})
\end{aligned}
$$

Las ecuaciones de verosimilitud son

$$
\begin{aligned}
\frac{d}{d\boldsymbol{\mu}} l(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Lambda}) &= \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{y}_0 - \Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\boldsymbol{q}_m \\
\frac{d}{d(v(\boldsymbol{\Sigma}))} l(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Lambda}) &= -\frac{1}{2}D^\top vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right) \\
&\quad +\frac{1}{2}D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)vec(\boldsymbol{S}_0) \\
&\quad +\frac{1}{2}\Psi(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
&\quad -\Psi(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
\frac{d}{dvec(\boldsymbol{\Lambda})} l(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Lambda}) &= -\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right)vec(\boldsymbol{I}_k) \\
&\quad +\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\right)vec(\boldsymbol{S}_0) \\
&\quad -\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{I}_m\right)vec(Q_m) \\
&\quad +\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m) \\
&\quad +\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right) \\
&\quad -2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)vec\left(\boldsymbol{q}_m\boldsymbol{y}_0^\top\right)
\end{aligned}
$$

Ahora, el cuadrado del Score para cada parámetro es

$$\begin{aligned}
S_{\boldsymbol{\mu}} S_{\boldsymbol{\mu}}^{\top} &= (d\boldsymbol{\mu})^{\top} \Big[(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} \boldsymbol{S}_0 (\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} - \Psi(\boldsymbol{\eta}) (\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} \boldsymbol{y}_0 \boldsymbol{q}_m^{\top} \boldsymbol{\Lambda} (\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} \\
&\quad + \Psi^2(\boldsymbol{\eta}) (\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} \boldsymbol{\Lambda}^{\top} \boldsymbol{q}_m \boldsymbol{q}_m^{\top} \boldsymbol{\Lambda} (\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda})^{-1} \Big] (d\boldsymbol{\mu})
\end{aligned}$$

$$\begin{aligned}
S_{\boldsymbol{\Sigma}} S_{\boldsymbol{\Sigma}}^{\top} &= d(v(\boldsymbol{\Sigma}))^{\top} \\
&\Bigg[\frac{1}{4} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right)^{\top} D \\
&+ \frac{1}{4} D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) D \\
&+ \frac{1}{4} \Psi^2(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \right) vec(Q_m) vec(Q_m)^{\top} \\
&\qquad \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) D \\
&+ \Psi^2(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}^{\top} \right) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top}) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda} \right) D \\
&- \frac{1}{2} D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) D \\
&- \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(Q_m)^{\top} \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) D \\
&+ \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda} \right) D \\
&+ \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(\boldsymbol{S}_0) vec(Q_m)^{\top} \\
&\qquad \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) D \\
&- \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda} \right) D \\
&- \Psi^2(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top} \right) vec(Q_m) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda} \right) D \Bigg] \\
&d(v(\boldsymbol{\Sigma}))
\end{aligned}$$

$$
\begin{aligned}
S_{\boldsymbol{\Lambda}}S_{\boldsymbol{\Lambda}}^{\top} \;=\;& vec(d\boldsymbol{\Lambda})^{\top} \\
& \Bigg[\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(I_k)vec(I_k)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) \\
& + \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0)vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}\right) \\
& + \Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\otimes\boldsymbol{I}_m\right) vec(Q_m)vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& + \Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) vec(Q_m)vec(Q_m)^{\top} \\
& \qquad \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) \\
& + \Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& + 4\Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& - 2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) \\
& + 2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& - 2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) \\
& - 2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& + 4\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\right) vec(\boldsymbol{I}_k)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& - 2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0)vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& + 2\Psi(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0)vec(Q_m)^{\top} \\
& \qquad \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) \\
& + 2\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& - 4\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) vec(\boldsymbol{S}_0)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& - 2\Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\otimes\boldsymbol{I}_m\right) vec(Q_m)vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) \\
& - 2\Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\otimes\boldsymbol{I}_m\right) vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& + 4\Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\otimes\boldsymbol{I}_m\right) vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& + 2\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) \\
& - 4\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^{\top}\right) vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& - 4\Psi^2(\boldsymbol{\eta})\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right) vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})vec(\boldsymbol{q}_m\boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right)\Bigg] vec(d\boldsymbol{\Lambda})
\end{aligned}
$$

$$
\begin{aligned}
S_{\boldsymbol{\mu}} S_{\boldsymbol{\Sigma}}^{\top} \;=\; (d\boldsymbol{\mu})^{\top} \Bigg[& -\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right)^{\top} D \\
& +\frac{1}{2}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& +\frac{1}{2}\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} y_0 vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\right) D \\
& +\frac{1}{2}\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\frac{1}{2}\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& -\frac{1}{2}\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1}\right) D \\
& +\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\right) D \Bigg] d(v(\boldsymbol{\Sigma}))
\end{aligned}
$$

$$
\begin{aligned}
S_{\boldsymbol{\mu}} S_{\boldsymbol{\Lambda}}^{\top} \;=\; (d\boldsymbol{\mu})^{\top} \Bigg[& -\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(\boldsymbol{I}_k)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^{\top}\right) \\
& +\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \\
& -\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& +\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \\
& +\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(\boldsymbol{q}_m \boldsymbol{y}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& -2\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{y}_0 vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \\
& +\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{I}_k)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}^{\top}\right) \\
& -\Psi(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \\
& +\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& -\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\right) \\
& -\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \otimes \boldsymbol{I}_m\right) \\
& +2\Psi^2(\boldsymbol{\eta})\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-1} \boldsymbol{\Lambda}^{\top}\boldsymbol{q}_m vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda}\right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top}\right) \Bigg] vec(d\boldsymbol{\Lambda})
\end{aligned}
$$

$$
\begin{aligned}
S_{\boldsymbol{\Sigma}} S_{\boldsymbol{\Lambda}}^{\top} \quad = \quad & d(v(\boldsymbol{\Sigma}))^{\top} \Bigg[\frac{1}{2} D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{I}_k)^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{\Lambda}^{\top} \right) \\
& - \frac{1}{2} D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) \\
& + \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(Q_m)^{\top} \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{I}_m \right) \\
& - \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(Q_m)^{\top} \left(\boldsymbol{\Lambda}^{\top} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) \\
& - \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{I}_m \right) \\
& + \Psi(\boldsymbol{\eta}) D^{\top} vec\left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top} \right) \\
& - \frac{1}{2} D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{I}_k)^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{\Lambda}^{\top} \right) \\
& + \frac{1}{2} D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{S}_0)^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) \\
& - \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(Q_m)^{\top} \left(\boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{I}_m \right) \\
& + \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(Q_m)^{\top} \\
& \qquad \left(\boldsymbol{\Lambda}^{\top} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{\Lambda} \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) \\
& + \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{I}_m \right) \\
& - \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \right) vec(\boldsymbol{S}_0) vec(\boldsymbol{q}_m \boldsymbol{y}_0^{\top})^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-2} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\top} \right) \\
& - \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) vec(Q_m) vec(\boldsymbol{I}_k)^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \boldsymbol{\Lambda}^{\top} \right) \\
& + \frac{1}{2} \Psi(\boldsymbol{\eta}) D^{\top} \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right) vec(Q_m) vec(\boldsymbol{S}_0)^{\top} \\
& \qquad \left(\left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \otimes \left(\boldsymbol{\Sigma} + \boldsymbol{\Lambda}^{\top}\boldsymbol{\Lambda} \right)^{-1} \boldsymbol{\Lambda}^{\top} \right)
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1}{2}\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m)vec(Q_m)^\top\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+\frac{1}{2}\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m)vec(Q_m)^\top\\
&\qquad\left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+\frac{1}{2}\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&-\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)vec(Q_m)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\\
&+\Psi(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(\boldsymbol{I}_k)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}^\top\right)\\
&-\Psi(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(\boldsymbol{S}_0)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&+\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(Q_m)^\top\left(\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&-\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(Q_m)^\top\left(\boldsymbol{\Lambda}^\top\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}^\top\right)\\
&-\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-1}\otimes\boldsymbol{I}_m\right)\\
&+2\Psi^2(\boldsymbol{\eta})D^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}^\top\right)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)vec(\boldsymbol{q}_m\boldsymbol{y}_0^\top)^\top\left(\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}^\top\boldsymbol{\Lambda}\right)^{-2}\otimes\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top\right)\Bigg]vec(d\boldsymbol{\Lambda})
\end{aligned}
$$

La Matriz de Información sigue de aplicar las esperanzas para cada cuadrado del Score.

En la sección 5.3 se obtiene con exito, el caso particular univariado de esta Matriz de Información.

Como se pudo apreciar a lo largo de esta sección, para calcular la Matriz de Información de este modelo en el caso multivariado, fue necesario conocer y aplicar todas las propiedades revisadas en el capítulo anterior, las que no solo comprometen a la distribución Skew-Normal, sino que también se refieren a otras distribuciones, tales como la Half-Normal y la Normla Truncada.

La obtención de la Matriz de Información para el caso multivariado en el modelo planteado en este trabajo, es una herramienta útil para el estudio de las propiedades de los estimadores de máxima verosimilitud desde el punto de vista clásico y también es un resultado útil para el desarrollo de modelos paramétricos con enfoque bayesiano.

Capítulo 4

Conclusiones

El desarrollo del presente trabajo consistió en la exploración de algunas propiedades no estudiadas de la familia de distribcuiones Skew-Normal basándose en el modelo presentado por Azzalini y Genton (2005).

Para lograr este objetivo se revisaron ciertos conceptos, tales como el producto directo entre dos matrices, la vectorización de una matriz y la matriz de duplicación y conmutación.

Luego, se estudiaron algunas propiedades para distribuciones relacionadas con la distribución Skew-Normal, tal como la distribución Normal truncada, para la cual se calculó la función generadora de momentos, con lo que se obtuvo los dos primeros momentos. Para la distribución Normal Multivariada se calcularon algunos momentos no conocidos y se obtuvo la forma de $d\left(\Phi_m\left(\boldsymbol{\alpha}|\boldsymbol{\beta},\boldsymbol{\Gamma}\right)\right)$. Para la distribución Half-Normal se calcularon algunos momentos necesarios para el cálculo de la matriz de información.

Al calcular la matriz de información para el caso univariado, se logró llegar a los mismos resultados que la parametrización usada en Arellano-Valle y Azzalini (2008) a través de la matriz jacobiana.

Al calcular la matriz de información para el caso multivariado se utilizó la mayoría de los resultados que se presentaron anteriormente, con lo que se llegó a la forma general de la matriz, la que tiene componentes que deben calcularse numericamente.

Este trabajo fue realizado con la distribución Skew-Normal, sin embargo en el futuro se podría estudiar estas propiedades para una familia mas general, como lo es la familia de distribuciones Skew-Eliptica. Por otro lado, la parametrización utilizada es un caso particular de la llamada CSN-2 (Close Skew Normal) planteada por Arellano-Valle y Azzalini (2006), por lo sería interesante estudiar estos resultados para esa familia.

Este trabajo puede servir de base para futuras aplicaciones que se deseen realizar con este modelo, ya sea desde el punto de vista clásico o Bayesiano.

Capítulo 5

Apéndice

5.1. Definiciones

5.1.1. Producto Directo

Sea $\boldsymbol{A}$ una matriz de dimensiones $m_2 \times n_2$ y sea $\boldsymbol{B}$ una matriz de dimensiones $m_1 \times n_1$; entonces el producto directo de $\boldsymbol{A}$ y $\boldsymbol{B}$, el cual escribimos como $\boldsymbol{A} \otimes \boldsymbol{B}$, es una matriz $\boldsymbol{C}$ de dimensiones $m_1 m_2 \times n_1 n_2$ definida por

$$\boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{C} = \begin{pmatrix} \boldsymbol{A}b_{11} & \boldsymbol{A}b_{12} & \cdots & \boldsymbol{A}b_{1n_1} \\ \boldsymbol{A}b_{21} & \boldsymbol{A}b_{22} & \cdots & \boldsymbol{A}b_{2n_1} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}b_{m_1 1} & \boldsymbol{A}b_{m_1 2} & \cdots & \boldsymbol{A}b_{m_1 n_1} \end{pmatrix} = \begin{pmatrix} b_{11}\boldsymbol{A} & b_{12}\boldsymbol{A} & \cdots & b_{1n_1}\boldsymbol{A} \\ b_{21}\boldsymbol{A} & b_{22}\boldsymbol{A} & \cdots & b_{2n_1}\boldsymbol{A} \\ \vdots & \vdots & & \vdots \\ b_{m_1 1}\boldsymbol{A} & b_{m_1 2}\boldsymbol{A} & \cdots & b_{m_1 n_1}\boldsymbol{A} \end{pmatrix}$$

5.1.2. Vectorización de una Matriz

Sea $\boldsymbol{A}$ una matriz de dimensiones $m \times n$, donde $\boldsymbol{A} = [a_1, a_2, ..., a_n]$; esto es, a_i es la i-ésima columna de $\boldsymbol{A}$. Entonces la vectorización de $\boldsymbol{A}$, denotada por $vec(\boldsymbol{A})$, está definida por

$$vec(\boldsymbol{A}) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

5.1.3. Matriz de Duplicación

Para una matriz simétrica $\boldsymbol{M}$ de dimensiones $d \times d$, entonces $v(\boldsymbol{M})$ es el vector formado al apilar el triángulo inferior de $\boldsymbol{M}$. Además, existe una matriz $\boldsymbol{D}$ de dimensiones $d^2 \times d(d+1)/2$, llamada matriz de duplicación tal que $vec(\boldsymbol{M}) = \boldsymbol{D}v(\boldsymbol{M})$ y $v(\boldsymbol{M}) = \boldsymbol{D}^{+}vec(\boldsymbol{M})$, donde $\boldsymbol{D}^{+} = (\boldsymbol{D}^{\top}\boldsymbol{D})^{-1}\boldsymbol{D}^{\top}$

5.1.4. Matriz de Conmutación

Sea $\boldsymbol{d}_i$ el vector unitario de diemnsión $m \times 1$, cuyo i-ésimo elemento es $+1$ y los elementos restantes son 0.

Sea $\boldsymbol{\Delta}_{ij} = \boldsymbol{d}_i\boldsymbol{d}_j^{\top}$ la matriz unitaria de dimensiones $m \times n$, cuyo ij-ésimo elemento es $+1$ y los elementos restantes son 0.

Entonces, la matriz $\boldsymbol{K}_{mn}$ de dimensiones $mn \times mn$ se define como la matriz de conmutación, donde

$$\boldsymbol{K}_{mn} = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\boldsymbol{\Delta}_{ij}^{\top} \otimes \boldsymbol{\Delta}_{ij}\right)$$

5.1.5. Algunas propiedades útiles

1. Sea α cualquier escalar y $\boldsymbol{A}$ y $\boldsymbol{B}$ matrices, entonces

$$(\alpha \boldsymbol{A}) \otimes \boldsymbol{B} = \boldsymbol{A} \otimes (\alpha \boldsymbol{B}) = \alpha(\boldsymbol{A} \otimes \boldsymbol{B})$$

2. Sea $\boldsymbol{A}$, $\boldsymbol{B}$ y $\boldsymbol{C}$ matrices cualquiera, entonces

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})$$

3. Sea $\boldsymbol{A}$ y $\boldsymbol{B}$ matrices cualquiera, entonces

$$(\boldsymbol{A} \otimes \boldsymbol{B})^\top = \boldsymbol{A}^\top \otimes \boldsymbol{B}^\top$$

4. Sea $\boldsymbol{A}$ una matriz de dimensión $m_1 \times n_1$, $\boldsymbol{B}$ una matriz de dimensión $m_2 \times n_2$, $\boldsymbol{F}$ una matriz de dimensión $n_1 \times k_1$ y $\boldsymbol{G}$ una matriz de dimensión $n_2 \times k_2$, entonces

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{F} \otimes \boldsymbol{G}) = (\boldsymbol{A}\boldsymbol{F}) \otimes (\boldsymbol{B}\boldsymbol{G})$$

5. Sea $\boldsymbol{A}$ y $\boldsymbol{B}$ matrices de dimensión $m \times m$ y sea $\boldsymbol{C}$ una matriz de dimensión $n \times n$, entonces

$$(\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} = (\boldsymbol{A} \otimes \boldsymbol{C}) + (\boldsymbol{B} \otimes \boldsymbol{C})$$

6. Si $\boldsymbol{x}$ es un vector de dimensión $m \times 1$ e y es un vector de diemnsión $n \times 1$, entonces

$$vec(\boldsymbol{x}\boldsymbol{y}^\top) = \boldsymbol{x} \otimes \boldsymbol{y}$$

7. Para cualquier vector $\boldsymbol{y}$ se tiene que $vec(\boldsymbol{y}) = vec(\boldsymbol{y}^\top) = \boldsymbol{y}$

8. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices de dimensión $n \times n$ y α y β son escalares, entonces

$$vec(\alpha \boldsymbol{A} + \beta \boldsymbol{B}) = \alpha vec(\boldsymbol{A}) + \beta vec(\boldsymbol{B})$$

9. Sea $\boldsymbol{A}$, $\boldsymbol{B}$ y $\boldsymbol{C}$ matrices de dimensión $m \times q$, $q \times s$ y $n \times s$ respectivamente. Entonces

$$vec(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}^\top) = (\boldsymbol{A} \otimes \boldsymbol{C}) vec(\boldsymbol{B})$$

10. Sea $\boldsymbol{A}$ una matriz de dimensión $m \times q$ y $\boldsymbol{B}$ una matriz de dimensión $q \times m$. Entonces

$$[vec(\boldsymbol{A}^\top)]^\top vec(\boldsymbol{B}) = tr(\boldsymbol{A}\boldsymbol{B})$$

11. Sea $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ y $\boldsymbol{D}$ matrices de dimensión $m \times n$, $n \times p$, $p \times q$ y $q \times m$, respectivamente. Entonces

$$tr(ABCD) = (vec(D^\top))^\top (C^\top \bigotimes A) vec(B) = (vec(D))^\top (A \bigotimes C^\top) vec(B^\top)$$

12. Sea $\boldsymbol{A} = [a_{ij}]$ una matriz de dimensión $m \times n$, y sea $\boldsymbol{B} = [b_{ij}]$ una matriz de dimensión $p \times q$. Entonces

$$\boldsymbol{A} \otimes \boldsymbol{B} = \sum_{i=1}^{m} \sum_{j=1}^{n} [\boldsymbol{A} \otimes (b_{ij} \boldsymbol{\Delta}_{ij})],$$

donde $\boldsymbol{\Delta}_{ij}$ es una matriz unitaria de dimensión $p \times q$

13. Sea $\boldsymbol{\Delta}_{ij}$ una natriz unitaria de dimensión $n \times n$. Entonces

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (\boldsymbol{\Delta}_{ij} \otimes \boldsymbol{\Delta}_{ij}) = [vec(\boldsymbol{I}_n)][vec(\boldsymbol{I}_n)]^\top$$

14. Sea $\boldsymbol{A}$ una matriz de dimensiones $m \times n$, $\boldsymbol{B}$ una matriz de dimensiones $p \times q$ y $\boldsymbol{b}$ un vector de diemensiones $p \times 1$. Entonces

$$\begin{aligned} \boldsymbol{K}_{pm} (\boldsymbol{A} \otimes \boldsymbol{B}) &= (\boldsymbol{B} \otimes \boldsymbol{A}) \boldsymbol{K}_{qn} \\ \boldsymbol{K}_{pm} (\boldsymbol{A} \otimes \boldsymbol{B}) \boldsymbol{K}_{qn} &= \boldsymbol{B} \otimes \boldsymbol{A} \\ \boldsymbol{K}_{pm} (\boldsymbol{A} \otimes \boldsymbol{b}) &= \boldsymbol{b} \otimes \boldsymbol{A} \\ \boldsymbol{K}_{mp} (\boldsymbol{b} \otimes \boldsymbol{A}) &= \boldsymbol{A} \otimes \boldsymbol{b} \end{aligned}$$

5.2. Comparación de las Matrices de Información

En esta sección se comparará la Matriz de Información obtenida mediante el modelo $Y \sim SNG_{1,1}(\mu, \sigma^2, \lambda)$, con la Matriz obtenida por Arellano-Valle & Azzalini (2008) ya que para el caso univariado los modelos son equivalentes. Para esto recordemos que la densidad enunciada por Arellano-Valle es:

$$
\begin{aligned}
f(y|\xi, \omega^2, \alpha) &= 2\omega^{-1}\phi\left(\frac{y-\xi}{\omega}\right)\Phi\left(\alpha\left(\frac{y-\xi}{\omega}\right)\right) \\
f(y|\xi, \Omega, \eta) &= 2\Omega^{-1/2}\phi\left(\frac{y-\xi}{\sqrt{\Omega}^{1/2}}\right)\Phi\left(\eta(y-\xi)\right)
\end{aligned}
$$

con $\Omega = \omega^2$ y $\eta = \frac{\alpha}{\omega}$

Es decir, la relación entre los parametros de los dos modelos es la siguiente:

$$
\begin{array}{lll}
\mu = \epsilon & \sigma^2 = \dfrac{\Omega}{\eta^2\Omega + 1} & \lambda = \dfrac{\eta\Omega}{(\eta^2\Omega+1)^{1/2}} \\
d_\epsilon \mu = 1 & d_\Omega \mu = 0 & d_\eta \mu = 0 \\
d_\epsilon \sigma^2 = 0 & d_\Omega \sigma^2 = \dfrac{1}{(\eta^2\Omega+1)^2} & d_\eta \sigma^2 = -\dfrac{2\Omega^2\eta}{(\eta^2\Omega+1)^2} \\
d_\epsilon \lambda = 0 & d_\Omega \lambda = \dfrac{\eta^3\Omega + 2\eta}{2(\eta^2\Omega+1)^{3/2}} & d_\eta \lambda = \dfrac{\Omega}{(\eta^2\Omega+1)^{3/2}}
\end{array}
$$

y la matriz jacobiana tiene la siguiente forma:

$$
D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{(\eta^2\Omega+1)^2} & -\frac{2\Omega^2\eta}{(\eta^2\Omega+1)^2} \\ 0 & \frac{\eta^3\Omega+2\eta}{2(\eta^2\Omega+1)^{3/2}} & \frac{\Omega}{(\eta^2\Omega+1)^{3/2}} \end{pmatrix}
$$

Finalmente, calculamos la matriz de información del modelo de Arellano-Valle & Azzalini (2008) a través de la siguiente tranformación:

$$
I(\theta) = D^\top I(\nu) D
$$

con lo que se obtiene lo siguiente:

$$
\begin{aligned}
I_{11}(\theta) &= I_{11} \\
&= \frac{1}{\Omega} + \eta^2 c_2 a_0 \\
I_{12}(\theta) &= D_{22} I_{12} + D_{32} I_{13} \\
&= \frac{1}{(\eta^2\Omega+1)^2}\Big(\frac{2^{1/2}\eta}{\pi^{1/2}\Omega(\eta^2\Omega+1)^{1/2}} + (\frac{1}{\Omega}+\frac{\eta^2}{2})\eta^2 c_2 a_1 - (\frac{1}{\Omega}+\frac{\eta^2}{2})\frac{\eta}{\eta^2\Omega+1}c_1\Big) \\
&\quad + \frac{\eta(\eta^2\Omega+2)}{2(\eta^2\Omega+1)^{3/2}}\Big(\frac{2^{3/2}\eta^2}{\pi^{1/2}(\eta^2\Omega+1)} + \frac{1}{\Omega(\eta^2\Omega+1)^{3/2}}c_1 - \frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}}c_2 a_1\Big) \\
&= \frac{\eta}{\Omega}c_1 \\
I_{13}(\theta) &= D_{23} I_{12} + D_{33} I_{13} \\
&= -\frac{2\Omega^2\eta}{(\eta^2\Omega+1)^2}\Big(\frac{2^{1/2}\eta}{\pi^{1/2}\Omega(\eta^2\Omega+1)^{1/2}} + (\frac{1}{\Omega}+\frac{\eta^2}{2})\frac{\eta}{\eta^2\Omega+1}c_1\Big) \\
&\quad + \frac{\Omega}{(\eta^2\Omega+1)^{3/2}}\Big(\frac{2^{3/2}\eta^2}{\pi^{1/2}(\eta^2\Omega+1)} + \frac{1}{(\eta^2\Omega+1)^{3/2}\Omega}c_1 - \frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}}c_1 a_1\Big) \\
&= c_1 - c_2 a_1 \eta \\
I_{22}(\theta) &= D_{22}^2 I_{22} + 2D_{22}D_{32}I_{23} + D_{32}^2 I_{33} \\
&= \frac{1}{(\eta^2\Omega+1)^4}\Big(\frac{1}{2\Omega^2} + (\frac{\eta^2}{\Omega^2}+\frac{\eta^4}{\Omega}+\frac{\eta^6}{4})c_2 A_2\Big) \\
&\quad + \frac{2}{(\eta^2\Omega+1)^2}\frac{\eta^3\Omega+2\eta}{2(\eta^2\Omega+1)^{3/2}}\Big(\frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}} - \frac{\eta(2+\eta^2\Omega)}{2\Omega^2(\eta^2\Omega+1)^{1/2}}c_2 A_2\Big) \\
&\quad + \Big(\frac{\eta^3\Omega+2\eta}{2(\eta^2\Omega+1)^{3/2}}\Big)^2\Big(\frac{2\eta^2}{\eta^2\Omega+1} + \frac{c_2 A_2}{\Omega^2(\eta^2\Omega+1)}\Big) \\
&= \frac{1}{2\Omega^2} \\
I_{23}(\theta) &= D_{23}(D_{22}I_{22} + D_{32}I_{32}) + D_{33}(D_{22}I_{23} + D_{32}I_{33}) \\
&= -\frac{2\Omega^2\eta}{(\eta^2\Omega+1)^2}\Big(\frac{1}{(\eta^2\Omega+1)^2}\Big(\frac{1}{2\Omega^2} + (\frac{\eta^2}{\Omega^2}+\frac{\eta^4}{\Omega}+\frac{\eta^6}{4})c_2 A_2\Big) \\
&\quad + \frac{\eta^3\Omega+2\eta}{2(\eta^2\Omega+1)^{3/2}}\Big(\frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}} - \frac{\eta(2+\eta^2\Omega)}{2\Omega^2(\eta^2\Omega+1)^{1/2}}c_2 A_2\Big)\Big) \\
&\quad + \frac{\Omega}{(\eta^2\Omega+1)^{3/2}}\Big(\frac{1}{(\eta^2\Omega+1)^2}\Big(\frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}} - \frac{\eta(2+\eta^2\Omega)}{2\Omega^2(\eta^2\Omega+1)^{1/2}}c_2 A_2\Big) \\
&\quad + \frac{\eta^3\Omega+2\eta}{2(\eta^2\Omega+1)^{3/2}}\Big(\frac{2\eta^2}{\eta^2\Omega+1} + \frac{c_2 A_2}{\Omega^2(\eta^2\Omega+1)}\Big)\Big) \\
&= 0
\end{aligned}
$$

$$\begin{aligned} I_{33}(\theta) &= D_{23}^2 I_{22} + 2D_{23}D_{33}I_{23} + D_{33}^2 I_{33} \\ &= (-\frac{2\Omega^2\eta}{(\eta^2\Omega+1)^2})^2(\frac{1}{2\Omega^2} + (\frac{\eta^2}{\Omega^2} + \frac{\eta^4}{\Omega} + \frac{\eta^6}{4})c_2A_2) \\ &\quad +2(-\frac{2\Omega^2\eta}{(\eta^2\Omega+1)^2})\frac{\Omega}{(\eta^2\Omega+1)^{3/2}}(\frac{\eta}{\Omega(\eta^2\Omega+1)^{1/2}} - \frac{\eta(2+\eta^2\Omega)}{2\Omega^2(\eta^2\Omega+1)^{1/2}}c_2A_2) \\ &\quad +(\frac{\Omega}{(\eta^2\Omega+1)^{3/2}})^2(\frac{2\eta^2}{\eta^2\Omega+1} + \frac{c_2A_2}{\Omega^2(\eta^2\Omega+1)}) \\ &= c_2A_2 \end{aligned}$$

lo que coincide con el resultado de Arellano-Valle y Azzalini (2008).

5.3. Obtención del caso univariado en base a la Matriz de Información multivariada

En este apartado se obtendrán los resultados obtenidos en la sección 3.1 con los obtenidos en la sección 3.2 para el caso particular en que $m = 1$, $k = 1$. Esto es, $\mathbf{\Lambda} = \lambda$, $\mathbf{\Sigma} = \sigma^2$, lo que genera los siguientes resultados:

$$
\begin{aligned}
q &= \frac{(\sigma^2+\lambda^2)^{1/2}}{(\sigma^2)^{1/2}}\phi\left(\frac{\lambda y_0}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)\\
Q &= -\frac{\lambda(\sigma^2+\lambda^2)^{1/2}}{(\sigma^2)^{3/2}}y_0\phi\left(\frac{\lambda y_0}{(\sigma^2)^{1/2}(\sigma^2+\lambda^2)^{1/2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
R_1 &= R_5 = R_6 = R_7 = 0\\
R_2 &= \frac{\sigma^2+\lambda^2}{\sigma^2}c_2a_0\\
R_3 &= \frac{\lambda^2(\sigma^2+\lambda^2)}{(\sigma^2)^3}c_2A_2\\
R_4 &= \frac{\sigma^2+\lambda^2}{\sigma^2}c_2A_2\\
R_8 &= -\frac{\lambda(\sigma^2+\lambda^2)}{(\sigma^2)^2}c_2A_2\\
R_9 &= -\left(\frac{2}{\pi}\right)^{1/2}\lambda\\
R_{10} &= \left(\frac{2}{\pi}\right)^{1/2}\sigma^2\\
R_{11} &= \left(\frac{2}{\pi}\right)^{1/2}\\
R_{12} &= \left(\frac{2}{\pi}\right)^{1/2}\sigma^2\\
R_{13} &= -\frac{\lambda(\sigma^2+\lambda^2)}{(\sigma^2)^2}c_2a_1\\
R_{14} &= \frac{\sigma^2+\lambda^2}{\sigma^2}c_2a_1
\end{aligned}
$$

donde c_2, a_0, a_1, A_2 fueron definidos en el lema 2.5.1 del capítulo 2.

Con estos resultados se replica la Matriz de Información del caso univariado, la que a su vez fue contrastada con la obtenida por Arellano-Valle & Azzalini (2008) en la sección 5.2.

Bibliografía

[1] R. Arellano-Valle and A. Azzalini (2006). On the Unification of Families of Skew-normal Distributions. *Scandinavian Journal of Statistics* **33**, 561-574.

[2] R. Arellano-Valle and A. Azzalini (2008). The centred parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis* **99**, 1362–1382.

[3] R. Arellano-Valle, M. Branco and M. Genton (2006). A unified view on skewed distributions arising from selections. *The Canadian Journal of Statistics* **34**, 581-601.

[4] R. Arellano-Valle and M. Genton (2005). On fundamental skew distributions. *Journal of Multivariate Analysis* **96**, 93–116.

[5] R. Arellano-Valle, H. Gómez and F. Quintana (2004). A New Class of Skew-Normal Distributions. *Communications in Statistics: Theory and Methods* **33**, 1465–1480.

[6] A. Azzalini (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics* **12**, 171-178.

[7] A. Azzalini and A. Capitanio (1999). Statistical Applications of the Multivariate Skew Normal Distribution. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **61**, (1999), 579-602.

[8] A. Azzalini and A. Dalla Valle (1996). The Multivariate Skew-Normal Distribution. *Biometrika* **83**, 715-726.

[9] P. Ghosh, M. Branco and H. Chakraborty (2007). Bivariate random effect model using skew-normal distribution with application to HIV-RNA. *Statistics in Medicine* **26**, 1255-1267.

[10] F. Graybill (1983). *Matrices with applications in Statistics.* Second Edition. Duxbury, Thomson Learning.

[11] A. Gupta, G. González-Farías and J. Domínguez-Molina (2004). A multivariate skew normal distribution. *Journal of Multivariate Analysis* **89**, 181–190.

[12] T. Lin, H. Ho and C. Chen (2009). Analysis of multivariate skew normal models with incomplete data. *Journal of Multivariate Analysis* **100**, 2337-2351.

[13] B. Liseo and N. Loperfido (2003). A Bayesian interpretation of the multivariate skew-normal distribution. *Statistics & Probability Letters* **61**, 395–401.

[14] B. Liseo and N. Loperfido (2006). A note on reference priors for the scalar skew-normal distribution. *Journal of Statistical Planning and Inference* **136**, 373–389.

[15] J.Magnus and H. Neudecker (1988). *Matix Differential Calculus.* J. Wiley & Sons, New York.

[16] S. Sahu, D. Dey and M. Branco (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics* **31**, 129-150.