Métodos robustos en modelo lineal y análisis multivariado

Adrover, Jorge Gabriel

1993

Tesis presentada para obtener el grado de Doctor en Ciencias Matemáticas de la Universidad de Buenos Aires

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Tema de Tesis

METODOS ROBUSTOS EN MODELO LINEAL Y ANÁLISIS MULTIVARIADO

Autor

Jorge Gabriel Adrover

Director

Dr. Víctor J. Yohai

Lugar de Trabajo

Dpto. de Matemática - Fac. de Ciencias Exactas y Naturales

Tesis presentada para optar al título de Doctor en Ciencias Matemáticas

1993

AGRADECIMIENTOS

Al Dr. Víctor J. Yohai, por su invalorable dirección y apoyo a lo largo de la realización de este trabajo. A mis compañeros del Laboratorio de Estadística, cuyo aprecio constante ha sido fundamental para sobrellevar las dificultades de la tarea. A mi familia, mis padres, hermana y sobrinos cuyo afecto y generosidad fueron un sostén insustituible durante este tiempo.

INDICE

ESTIMACIÓN ROBUSTA MINIMAX EN SESGO DE LA MATRIZ DE DISPERSION

1. INTRODUCCION

La aproximación clásica del análisis multivariado está basada en la distribución normal multivariada $\mathcal{N}_m(\mu, \mathbf{V})$ con vector de posición μ y matriz de dispersión $\mathbf{V}$. Los estimadores de máxima verosimilitud de estos parámetros son la media y matriz de covarianza muestrales respectivamente. Sin embargo es bien conocido que ligeros apartamientos de este modelo o la presencia de una pequeña proporción de datos atípicos, puede distorsionar completamente los estimadores. Los procedimientos robustos buscan proveer estimadores estables cuando los datos son contaminados por una proporción pequeña de tales observaciones. Entonces el sesgo asintótico máximo sobre una ϵ-vecindad de contaminación del modelo central da una medida de la robustez global del estimador. El nivel de contaminación más pequeño para el cual el sesgo asintótico máximo es infinito se llama *punto de ruptura del estimador*. Los estimadores clásicos tienen punto de ruptura 0.

Entre las diversas propuestas robustas que han surgido para posición y dispersión multivariada deben mencionarse las siguientes. Maronna (1976) define M-estimadores afínmente equivariantes de posición y dispersión multivariada, mostrando existencia, unicidad, consistencia y normalidad asintótica de dichos estimadores. Huber (1977, 1981) generaliza la propuesta anterior. Sin embargo estos estimadores de dispersión tienen punto de ruptura menor que $1/(m+1)$ donde m es la dimensión del vector multivariado, lo cual significa una deficiencia importante para altas dimensiones. Stahel(1981) y Donoho(1982) presentan un estimador para dispersión basado en una robustificación de la distancia de Mahalanobis que tiene punto de ruptura 1/2. Tyler (1987) estudia una forma límite de los M-estimadores de dispersión multivariada con distribución asintótica independiente de la forma funcional de la distribución elíptica subyacente y con ciertas propiedades minimax. Davies (1987) generaliza para el modelo multivariado el concepto de S-estimador definido en el contexto de modelo lineal por Rousseeuw and Yohai (1984). Los S-estimadores resultan asintóticamente normales con punto de ruptura independiente de la dimensión y tan alto como es posible para estimadores afínmente equivariantes de dispersión.

Sin embargo el punto de ruptura nada dice acerca del comportamiento del sesgo para niveles de contaminación menores que el mismo. Entonces interesaría tener un control global sobre el sesgo para niveles de contaminación menores que el punto de ruptura, obteniendo para cada nivel el estimador que minimiza el máximo sesgo asintótico dentro de una familia de estimadores (estimador minimax en sesgo). En ese sentido Huber(1964) establece para el modelo de posición que la mediana minimiza el máximo sesgo asintótico entre todos los estimadores equivariantes por traslación, cuando el máximo es tomado en vecindades de ϵ- contaminación de una distribución simétrica. Martin and Zamar (1987) obtienen estimadores de escala minimax en sesgo. Asimismo Martin and Zamar (1987) construyen estimadores de posición minimax en sesgo sujetos a una condición sobre la eficiencia en el modelo subyacente. Martin, Yohai and Zamar (1989) construyen estimadores de regresión minimax para dos clases diferentes de estimadores: (i) M-estimadores basados en funciones ρ acotadas y estimador de escala general para los residuos, y (ii) GM-estimadores, con curvas de influencia acotada (para el caso de dispersión conocida).

Trabajando en ese contexto de robustez global, Maronna and Yohai (1990) definen adecuadamente el concepto de sesgo asintótico para matriz de dispersión (se verá la definición inmediatamente), y estudian el sesgo asintótico máximo (restringido a un subconjunto de contaminaciones) en dos clases de estimadores robustos para la matriz de dispersión de un vector $\mathbf{z}$ bajo un modelo de contaminación de la forma $F = (1-\epsilon)F_0 + \epsilon\delta_{\mathbf{z}_0}$, donde F es la distribución de $\mathbf{z}$, F_0 es una distribución esféricamente simétrica y $\delta_{\mathbf{z}_0}$ es una masa puntual en $\mathbf{z}_0$. En particular ellos estudian el S-estimador resultante de tomar una función de salto (que contiene como caso especial al estimador de volumen mínimo de Rousseeuw) y el estimador de Tyler, el cual muestra un sesgo máximo más pequeño que el del S-estimador mencionado salvo para niveles de contaminación cercanos a $1/m$. Maronna,Stahel and Yohai (1992) generalizando ideas presentadas en Maronna and Yohai (1991) definen estimadores basados en proyecciones a los efectos de obtener un estimador de alto punto de ruptura cuyo sesgo máximo para contaminaciones puntuales es mucho menor que el del estimador de volumen mínimo de Rousseeuw y se compara favorablemente en varios casos con el estimador de Tyler.

La necesidad de estimar robustamente la matriz de dispersión multivariada aparece no sólo en los usos clásicos de la misma (análisis discriminante, componentes principales, correlaciones canónicas, modelo lineal multivariado) sino también asociada a la implementación de varios estimadores robustos para el modelo de regresión en los cuales es necesario estimar robustamente la matriz de covarianza de los efectos aleatorios. Tal situación se observa por ejemplo con los GM-estimadores de regresión, donde la falta de robustez en la estimación de dicha matriz puede hacer crecer los sesgos minimax de tales estimadores (ver Martin, Yohai and Zamar - 1989). Otro ejemplo lo constituyen los llamados estimadores de proyección para regresión (ver Maronna and Yohai - 1991), donde en una de las propuestas planteadas se utiliza un estimador robusto de la matriz de covarianza de los efectos, permitiendo obtener así un estimador con propiedades de optimalidad en sesgo importantes.

La restricción a contaminaciones puntuales para estudios de sesgos es intuitivamente plausible respecto de la definición de sesgo propuesta pues deberían proveer el mayor grado de asimetría del M-estimador comparado con el obtenido bajo la distribución central esféricamente simétrica y que resulta un múltiplo de la matriz identidad. La elección del estimador de Tyler en la clase de los M-estimadores como candidato a minimizar el sesgo máximo en el entorno de contaminación surge por analogía con los resultados obtenidos para GM-estimadores con escala conocida en Martin, Yohai and Zamar (1989).

En este trabajo se muestra que ambas suposiciones son verdaderas: las contaminaciones puntuales van a proveer sesgos tan cercanos al máximo como se desee, y se cumple la propiedad de optimalidad para el estimador de Tyler, esto es, en la clase de los M-estimadores de dispersión multivariada el estimador de Tyler minimiza el máximo sesgo asintótico en el entorno de ϵ-contaminación para ϵ fijo.

En la Sección 2. se presenta notación y se establece formalmente el problema. Al final de la misma se esboza el camino a seguir para llegar a los resultados fundamentales. En la Sección 3. se enuncian y demuestran los resultados.Primeramente se observa la forma de los M-estimadores para distribuciones puntuales. Luego se estudia punto de ruptura y comportamiento de los M-estimadores bajo contaminaciones puntuales para finalmente abordar el problema principal, esto es estudiar el máximo sesgo de los M-estimadores en la vecindad y luego encontrar el estimador minimax. En la Sección 4. se establecen resultados auxiliares de existencia y unicidad del estimador de Tyler en la ϵ-vecindad e identificabilidad del modelo central.

2. NOTACION Y DEFINICIONES

En lo que siga la función indicadora de un conjunto A se denotará por I_A. Por P se entenderá una medida de probabilidad en $\Re^m$ y sop(P) denotará al conjunto de puntos $\mathbf{z} \in \Re^m$ tal que existe un entorno de z, $V_0(\mathbf{z})$ satisfaciendo $P(V_0(\mathbf{z}) > 0$. $\delta_{(K^{1/2},0,\ldots,0)}$ con $K > 0$, denotará la función de distribución asociada con la medida $\delta_{(K^{1/2},0,\ldots,0)}(B) = I_B(K^{1/2},0,\ldots,0)$ donde B es un boreliano de $\Re^m$. Sea Δ el conjunto formado por tales distribuciones. La base canónica de $\Re^m$ se notará en la forma usual $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ e $\mathbf{I}$ será la matriz identidad $m \times m$.

Sea $\mathbf{z}$ un vector m-dimensional con distribución elipsoidal, i.e. $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{t}_0$ donde $\mathbf{A}$ es una matriz no singular $m \times m$, $\mathbf{t}_0$ es un vector m-dimensional y $\mathbf{x}$ tiene distribución esféricamente simétrica con densidad f_0. Si $\mathbf{V}_0 = \mathbf{A}\mathbf{A}'$ y f_0 es independiente de $\mathbf{t}_0$ y $\mathbf{V}_0$ la densidad de $\mathbf{z}$ resulta ser

$$f(\mathbf{z}, \mathbf{t}_0, \mathbf{V}_0, f_0) = |\mathbf{V}_0|^{-1/2} f_0\left((\mathbf{z} - \mathbf{t}_0)'\mathbf{V}_0^{-1}(\mathbf{z} - \mathbf{t}_0)\right)$$

La identificabilidad del modelo puede verse en la Sección 4.. Sea

$$\mathcal{E} = \left\{\mathbf{A} \in \Re^{m\times m} / \det \mathbf{A} \neq 0, \mathbf{A} = \mathbf{A}^{\mathbf{T}} \text{ y } \mathbf{A} > 0\right\}$$

Si $\mathbf{V} \in \mathcal{E}$, y $\left\{\lambda_{(j)}(\mathbf{V})\right\}_{j=1}^{m}$ denotan a los autovalores de $\mathbf{V}$ en orden descendente, entonces el *número de condición* de una matriz $\mathbf{V}$ se definirá como

$$\gamma(\mathbf{V}) = \frac{\lambda_{(1)}(\mathbf{V})}{\lambda_{(m)}(\mathbf{V})}$$

Si F_0 denota la distribución bajo el modelo central entonces se define el conjunto

$$\mathcal{F}_\epsilon = \{G = (1 - \epsilon)F_0 + \epsilon H \quad \text{donde H denota una distribucion arbitraria en } \Re^m\}$$

Sea $\mathbf{V}(G) : \mathcal{F}_\epsilon \to \mathcal{E}$ un funcional de dispersión multivariada afínmente equivariante, esto es si $\mathbf{z}$ tiene distribución G y $\mathbf{Cz}$ tiene distribución $G_{\mathbf{C}}$, entonces $\mathbf{V}(G_{\mathbf{C}}) = \mathbf{C}\mathbf{V}(G)\mathbf{C}'$. Se llamará *punto de ruptura asintótico* del funcional $\mathbf{V}$ al valor

$$\epsilon^*(\mathbf{V}, F_0) = \sup \left\{ \begin{array}{c} \epsilon > 0 : \exists \quad K(\epsilon) \text{ compacto en } \Re^+ \text{ tal que } \lambda_j(\mathbf{V}(G)) \in K(\epsilon) \\ \forall \quad 1 \leq j \leq m \text{ y } G \in \mathcal{F}_\epsilon \end{array} \right\}$$

Se define asimismo el *sesgo asintótico de la matriz de dispersión* $\mathbf{V}$ en $G \in \mathcal{F}_\epsilon$ como $\gamma(\mathbf{B}^{-1}\mathbf{V}(G)(\mathbf{B}')^{-1})$ y el *sesgo asintótico máximo de la matriz de dispersión en el entorno* $\mathcal{F}_\epsilon$ como

$$b(F_0, \epsilon) = \sup_{G \in \mathcal{F}_\epsilon} \gamma(\mathbf{B}^{-1}\mathbf{V}(G)(\mathbf{B}')^{-1})$$

con $\mathbf{BB}' = \mathbf{V}(F_0)$. Esta noción de sesgo resulta entonces invariante por transformaciones afines de los datos e independiente de la descomposición $\mathbf{V}(F_0) = \mathbf{BB}'$ usada. Esto último surge del siguiente resultado que puede encontrarse en Muirhead (1982) (pág. 589)

LEMA. *Si* $\mathbf{AA}' = \mathbf{BB}'$ *entonces* $\mathbf{A} = \mathbf{BQ}$ *donde* $\mathbf{Q}$ *es una matriz ortogonal de* $\Re^m$.

Este concepto de sesgo mide el grado de asimetría provocado por una contaminación sabiendo que bajo el modelo la distribución es esféricamente simétrica. Asimismo, y dado que la inversa de la matriz de dispersión estimada aparece en la definición de los M- y S-estimadores, el riesgo numérico que ello acarrea es medido por un número de condición (*sesgo*) muy grande.

Reescribiendo la densidad de (1) se obtiene que

$$f(\mathbf{z}, \mathbf{t}_0, f_o, \sigma, \mathbf{V}_1) = \sigma^{-m/2} f_0(\sigma^{-1}(\mathbf{z} - \mathbf{t}_0)'\mathbf{V}_1^{-1}(\mathbf{z} - \mathbf{t}_0))$$

donde $\sigma = |\mathbf{V}|^{1/m}$ y $|\mathbf{V}_1| = 1$.

$\mathbf{V}_1$ se llamará parámetro de FORMA y σ parámetro de TAMAÑO. Se pretende estudiar el efecto de la contaminación sobre el factor de forma en una cierta familia de estimadores. Los estimadores a tratar en este trabajo serán los M-estimadores de posición y dispersión multivariadas definidos en Maronna (1976) como las soluciones de las ecuaciones

$$E_P u_1\left(\left((\mathbf{z}-\mathbf{t})\mathbf{V}^{-1}(\mathbf{z}-\mathbf{t})\right)^{1/2}\right)(\mathbf{z}-\mathbf{t}) = 0$$
$$E_P u_2\left((\mathbf{z}-\mathbf{t})'\mathbf{V}^{-1}(\mathbf{z}-\mathbf{t})\right)(\mathbf{z}-\mathbf{t})(\mathbf{z}-\mathbf{t})' = \mathbf{V}$$

donde $(\mathbf{t}, \mathbf{V}) \in \Re^m \times \mathcal{E}$ y las funciones $u_1(s)$ y $u_2(s)$ están definidas para $s \geq 0$. En dicho trabajo se supone además que u_1, u_2 y P satisfacen las condiciones siguientes:

(A) $u_1(s)$ y $u_2(s)$ son funciones no negativas, decrecientes y continuas en $s \geq 0$.

(B) $\psi_i(s) = s u_i(s)$, i=1,2 son acotadas; sean $K_i = \sup_{s \geq 0} \psi_i(s)$

(C) ψ_2 es no decreciente, y es estrictamente creciente en el intervalo donde $\psi_2 < K_2$.

(D) Existe s_0 tal que $\psi_2(s_0^2) > m$ y $u_1(s) > 0$ para $s \leq s_0$ (y entonces $K_2 > m$).

(E) Existe $a > 0$ tal que para todo hiperplano H, $\mathrm{P}(H) \leq 1 - m/K_2 - a$.

Estas condiciones aseguran existencia de soluciones, unicidad en caso de $\mathbf{t}_0$ conocido, o bien en el caso de $\mathbf{t}_0$ desconocido y P la distribución bajo el modelo.

Huber (1981) (basado en Schonholzer(1979)) generaliza la definición de M-estimadores a las soluciones del sistema

$$E_P u_1 \left(\left((\mathbf{z} - \mathbf{t})\mathbf{V}^{-1}(\mathbf{z} - \mathbf{t}) \right)^{1/2} \right) (\mathbf{z} - \mathbf{t}) = 0$$

$$\frac{E_P u_2 \left((\mathbf{z} - \mathbf{t})'\mathbf{V}^{-1}(\mathbf{z} - \mathbf{t}) \right) (\mathbf{z} - \mathbf{t})(\mathbf{z} - \mathbf{t})' I_{\Re^m - \mathbf{t}}(\mathbf{z})}{E_P v \left((\mathbf{z} - \mathbf{t})'\mathbf{V}^{-1}(\mathbf{z} - \mathbf{t}) \right) I_{\Re^m - \mathbf{t}}(\mathbf{z})} = \mathbf{V}$$

Como el vector 0 no contiene información direccional, esto es, información sobre la matriz de dispersión, es que es omitido en la integral que define al M-estimador de dispersión, al integrar en $\mathbf{z} \neq \mathbf{t}$. Para el caso de posición fija (que es el que va a interesar en este trabajo) sólo se considera la segunda ecuación. En tal caso los supuestos para asegurar existencia de soluciones serán (ver Huber (1981) - pág. 216)

E1) $u_2(s)$ es monótona decreciente y $u_2(s) > 0$ para $s > 0$.

E2) La función $v(s)$ es monótona creciente y $v(s) > 0$ para $s > 0$.

E3) $\psi_2(s)$ y $v(s)$ son acotadas y continuas para $s \geq 0$. Sea $v(\infty) = \sup v(s)$.

E4) $\psi_2(0)/v(0) < m$.

E5) $\forall$ H hiperplano de $\Re^m$, $P(H) < 1 - mv(\infty)/K_2$

$\forall$ H hiperplano de $\Re^m$, $P(H) \leq 1/m$.

La condición (E5) permite asegurar la invertibilidad de la matriz de dispersión estimada. Asimismo las condiciones requeridas para asegurar unicidad son las siguientes (ver Huber (1981) - pág. 220)

U1) $u_2(s)$ es decreciente.

U2) $\psi_2(s)$ es continua, creciente y positiva en $s > 0$.

U3) $v(s)$ es continua, no negativa y decreciente. Además es positiva si $s \in [0, s_0]$.

U4) Para todo hiperplano $H \subset \Re^m$, $\mathrm{P}(H) < 1/2$.

Para estos estimadores "no monótonos" se prueba existencia y unicidad de soluciones en caso de tener $v =$ constante .

En lo que sigue $\mathbf{t}$ se supondrá conocido y por lo tanto sin pérdida de generalidad se asumirá $\mathbf{t} = 0$. Luego la familia de M-estimadores a considerar será la solución de la

$$E_P \psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}) \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z}) = \mathbf{V} \tag{1.1}$$

y las condiciones requeridas para P y $u_2(s)$ serán

A1) $u_2(s)$ es una función no negativa, decreciente y continua en $s > 0$.

A2) $\psi_2(s) = su_i(s)$ es no negativa, continua y acotada en $s \geq 0$; sea $K_2 = \sup_{s \geq 0} \psi_2(s)$.

A3) $\psi_2(s)$ es no decreciente y es estrictamente creciente en el intervalo donde $\psi_2 < K_2$

A4) Existe s_0 tal que $\psi_2(s_0^2) > m$.

A5) $\lim\limits_{s \to 0} \psi_2(s) = m_0 < m$.

A6) $\forall$ H hiperplano de $\Re^m$, $P(H) < 1 - m/K_2$

$\forall$ H hiperplano de $\Re^m$, $P(H) \leq 1/m$.

Como caso límite de estos estimadores aparece el estimador definido por Tyler (1987) para el parámetro de forma (más generalmente para funciones de $\mathbf{V}$ de la forma $H(\mathbf{V}) = H(c\mathbf{V})$), el cual resulta de tomar $\psi_2(s) = m$. Estos estimadores resultan de una forma límite de M-estimadores de dispersión de tipo Huber. Estos últimos estimadores están definidos elegiendo para algún r fijo, $u_2(s) = \alpha$ si $s \leq r^2$ y $u_2(s) = \alpha r^2/s$ para $s > r^2$, donde el factor de escala α está definido para conseguir consistencia en el sentido de Fisher bajo normalidad de los datos. En Tyler (1983) se muestra que cuando $r \to 0$, la varianza asintótica de los estimadores tipo Huber de dispersión no depende de la forma funcional de la distribución elíptica subyacente (en caso de haber obtenido la muestra aleatoria de tal modelo). Entonces el estimador resulta de resolver la ecuación

$$\frac{m}{P(\mathbf{z} \neq \mathbf{t})} E_P \frac{(\mathbf{z} - \mathbf{t})(\mathbf{z} - \mathbf{t})'}{(\mathbf{z} - \mathbf{t})'\mathbf{V}^{-1}(\mathbf{z} - \mathbf{t})} I_{\Re^m - \mathbf{t}}(\mathbf{z}) = \mathbf{V}$$

donde $\mathbf{t}$ es conocido o es un estimador de posición afínmente equivariante inicial robusto con ciertas condiciones dadas en Tyler (1987) para asegurar consistencia y normalidad asintótica. En dicho trabajo se prueba asimismo existencia de soluciones y unicidad salvo por un factor positivo en caso de tener la distribución empírica basada en la muestra aleatoria. Por dicha razón se debe establecer alguna condición adicional sobre la matriz $\mathbf{V}$ tal como $\mathrm{tr}(\mathbf{V}) = m$ o $\det(\mathbf{V}) = 1$ para lograr la unicidad. En la Sección 4. puede hallarse una generalización de esos hechos, mostrándose la existencia del estimador en el entorno de ϵ-contaminación, para $\epsilon < 1/(m+1)$ y la unicidad salvo por un factor positivo.

Sin pérdida de generalidad se supondrá en lo que siga que F_0 es esféricamente simétrica. Intuitivamente, contaminaciones del tipo $\delta_{(K^{1/2},0,\ldots,0)}$ deberían proveer el mayor grado de asimetría en el estimador que es lo detectado por el sesgo y por ello todo el estudio se centra en ellas. En caso de tomar la distribución $G = (1-\epsilon)F_0 + \epsilon\delta_{(K^{1/2},0,\ldots,0)}, 0 < K < \infty$, se verá que la solución del sistema a estudiar es una matriz diagonal definida positiva tal que los elementos diagonales v_{ii} son todos iguales $\forall\ i = 2,\ldots,m$. No se pasará estrictamente por el teorema general de existencia de soluciones (Maronna, 1976 - Huber, 1981) sino que se esbozará una prueba para la situación particular, para entonces restringir las condiciones pedidas en dicho teorema sobre la acotación de la probabilidad de hiperplanos y las condiciones sobre la función u_2. Ello es necesario pues para estudiar el sesgo se trabajará con una familia de funciones que no satisfará la primer condición de (A6).

Multiplicando y dividiendo a $\mathbf{V}$ por $s = |\mathbf{V}|^{1/m}$ el sistema (1.1) resultante es

$$E_{\mathrm{P}}\psi_2\left(\frac{1}{s}\mathbf{z}'\mathbf{V}_1^{-1}\mathbf{z}\right)\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}_1^{-1}\mathbf{z}}I_{\Re^m-\{0\}}(\mathbf{z}) = \mathbf{V}_1 \tag{1.2}$$

donde $|\mathbf{V}_1| = 1$.

Para la distribución de interés la ecuación (1.2) toma la forma

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\mathbf{z}'\mathbf{V}_1^{-1}\mathbf{z}\right)\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}_1^{-1}\mathbf{z}} + \epsilon\psi_2(K\mathbf{e}_1'\mathbf{V}_1^{-1}\mathbf{e}_1)\frac{\mathbf{e}_1\mathbf{e}_1{}'}{\mathbf{e}_1'\mathbf{V}_1^{-1}\mathbf{e}_1}I_{(0,\infty)}(K) = \mathbf{V}_1 \tag{1.3}$$

y como la búsqueda de soluciones del sistema (1.3) se restringe a matrices diagonales tal que sus elementos v_{ii} son iguales $\forall\ i = 2,\ldots,m$, al pre y posmultiplicar en (1.3) por $\mathbf{e}_\mathbf{j}'$ y $\mathbf{e}_\mathbf{j}$, $1 \le j \le m$ se obtiene el siguiente conjunto de m ecuaciones:

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+\frac{v}{a}\right)\right)\frac{u}{u/b+v/a} + \epsilon\psi_2\left(\frac{K}{sb}\right)bI_{(0,\infty)}(K) = b$$
$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+\frac{v}{a}\right)\right)\frac{v_j}{u/b+v/a} = a$$

donde $u = z_1^2$, $v_j = z_j^2, 2 \le j \le m$, $v = \sum_{i=2}^{m} z_i^2$ y $ba^{m-1} = 1$. Tras cancelar algunos factores el sistema se convierte en:

$$\begin{aligned}(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)\frac{u}{u+vb^{m/(m-1)}} + \epsilon\psi_2\left(\frac{K}{sb}\right)I_{(0,\infty)}(K) &= 1\\ (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v_j}{u+vb^{m/(m-1)}} &= 1\end{aligned} \tag{1.4}$$

El conjunto de soluciones de (1.4), sistema de m ecuaciones, es el mismo que el conjunto de soluciones del siguiente sistema de dos ecuaciones, dado que las (m-1) últimas ecuaciones son idénticas por ser la distribución F_0 esféricamente simétrica: la primer ecuación resultará de sumar las m ecuaciones y la segunda de sumar las m-1 últimas

$$\begin{aligned} g_1^1(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)+\epsilon\psi_2\left(\frac{K}{sb}\right)I_{(0,\infty)}(K)=m \\ g_2(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u+vb^{m/(m-1)}}=m-1 \end{aligned} \tag{1.5}$$

En caso de tomar $K=\infty$ o $K=0$ los sistemas resultantes son

$$\begin{aligned} g_1^2(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)+\epsilon K_2=m \\ g_2(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u+vb^{m/(m-1)}}=m-1 \end{aligned} \tag{1.6}$$

$$\begin{aligned} g_1^3(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)=m \\ g_2(s,b) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u+vb^{m/(m-1)}}=m-1 \end{aligned} \tag{1.7}$$

El interés en las ecuaciones (1.6) (y por ende en (1.5)) reside en que su solución va a proveer el máximo sesgo en la ϵ-vecindad de contaminación Las soluciones de (1.5) y (1.6) (cuya existencia y unicidad en cada caso surgirá tras el Teorema 1) se denotarán por $(s(K,\epsilon),b(K,\epsilon))$ y $(s(\epsilon),b(\epsilon))$ respectivamente. Cuando se tenga $\psi_2(x)\equiv m$ (definiendo el estimador de Tyler), la solución de (1.5) o (1.6) se denotará por $\mathbf{b_T}(\epsilon)$. Para llegar a los resultados fundamentales del Teorema 3 (donde se establece que $(b(\epsilon))^{m/(m-1)}$ resulta ser el máximo sesgo en la ϵ-vecindad de contaminación), y del Teorema 4 (donde se muestra que el máximo sesgo se minimiza cuando se toma el estimador de Tyler), el mecanismo de prueba será el siguiente. En el Lema 1 a) se encuentra la forma del M-estimador bajo el modelo lo cual es necesario para el cálculo de sesgo. Con ello resulta también que el M-estimador es consistente en el sentido de Fisher, esto es, la solución de (1.1) resulta ser un múltiplo de la matriz identidad cuando se tienen distribuciones esféricamente simétricas. Asimismo el Lema 1 b) permite decir que los sistemas (1.3) y (1.5) son realmente equivalentes y puede justificarse el haber deshechado las ecuaciones obtenidas de premultiplicar y posmultiplicar por $\mathbf{e}_j'$ y $\mathbf{e}_{j'}$ si $j\neq j'$ en (1.3). En el Teorema 1 se prueba la existencia de soluciones de tales sistemas sin la condición (A6) sobre hiperplanos, surgiendo la unicidad análogamente a los resultados de Maronna (1976). En el Lema 2 se establece el punto de ruptura para M-estimadores y el Corolario 1 es un resultado particular para analizar la ruptura de cada uno de las componentes en particular de la solución $(s(K,\epsilon), b(K,\epsilon))$ restringiendo las contaminaciones al conjunto Δ ya definido. El Lema 3 prueba un resultado de continuidad de la solución $(s(K,\epsilon), b(K,\epsilon))$ respecto de la contaminación K. En el Lema 4 se muestra que $b(K,\epsilon)$ es monótona creciente respecto de K, lo cual es importante para decir que se tienen contaminaciones que producen sesgos tan cercanos al máximo como se quiera. Además el resultado es fuertemente usado para mostrar que el máximo sesgo en el entorno de contaminación es creciente en el nivel de contaminación ϵ (Lema 9) y y con este último resultado mostrar el resultado general del Teorema 4. Un resultado de monotonía similar es probado vía el Teorema 2 y el Corolario 6 para el estimador de tamaño $s(K,\epsilon)$. Con el Corolario 4 se obtiene que $b(\epsilon)$ es mayor que 1 y por lo tanto puede usarse para comparar con el sesgo producido por otra contaminación, lo cual es básico para la demostración del Teorema 3. Otros lemas y corolarios que aparecen surgen por motivos técnicos de las demostraciones.

3. RESULTADOS Y DEMOSTRACIONES

3.1 Forma de los M-estimadores para contaminaciones puntuales.

LEMA 1.

a) *Si valen (A2)-(A5), la solución de la ecuación (1.1) es un múltiplo de la matriz identidad.*

b) *Si* $\mathbf{V}$ *es diagonal, entonces* $\mathbf{M}(\mathbf{V}) = E_{F_0}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}}$ *es diagonal.*

DEMOSTRACIÓN: a) Por la ecuación (1.1) basta mostrar que la matriz

$$\mathbf{M} = E_{F_0}\psi_2(\frac{1}{s}\mathbf{z}'\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{z}}$$

es un múltiplo de la matriz identidad. La matriz $\mathbf{M}$ es simétrica y definida positiva y por lo tanto su descomposición espectral resulta $\mathbf{M} = \mathbf{Q}\Delta\mathbf{Q}'$, donde $\mathbf{Q}$ es ortogonal y Δ es la matriz diagonal de autovalores. En consecuencia $\Delta = \mathbf{Q}'\mathbf{M}\mathbf{Q} = \mathbf{M}$, donde la última igualdad vale por simetría esférica. Por el mismo motivo de esfericidad, $\mathbf{M}$ tiene todos sus elementos diagonales idénticos, resultando así que los autovalores de $\mathbf{M}$ son iguales, con lo cual $\mathbf{M} = c\mathbf{I}$ con $c = E_{F_0}\psi_2(s^{-1}\mathbf{z}'\mathbf{z})/m$. Luego eligiendo s_0 tal que $c = 1$ resulta que $s_0\mathbf{I}$ es solución. Luego el resultado de unicidad de Maronna(1976) permite concluir que esta es la única solución.

b) Sea la matriz diagonal $\mathbf{B}^{(j)}$ tal que todos los elementos diagonales son 1 salvo el elemento j-ésimo que es -1. Entonces $\mathbf{B}^{(j)}$ es ortogonal y $\mathbf{B}^{(j)}\mathbf{z}$ tiene la misma distribución que $\mathbf{z}$. Por lo tanto

$$E_{F_0}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}} = E_{F_0}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\frac{\mathbf{B}^{(j)}\mathbf{z}\mathbf{z}'\mathbf{B}^{(j)}}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}}$$

Pero entonces los elementos de la fila j-ésima son nulos salvo el j-ésimo, resultando lo que se deseaba ver. ∎

NOTA 1: El Lema 1 b) permite asegurar que la ecuación matricial

$$\mathbf{M}(s) = (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\mathbf{z}'\mathbf{z}\right)\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{z}}$$

es un múltiplo de la matriz identidad. Simplemente tomando traza, la condición $\epsilon < 1 - m/K_2$ permite asegurar que existe s tal que $\mathbf{M}(s) = \mathbf{I}$ tal como se esperaba. Esto dice que el sistema (1.7) en realidad debe verse como la ecuación

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}(u+v)\right) = m$$

que es entonces equivalente a la ecuación vectorial $\mathbf{M}(s) = \mathbf{I}$. La condición $\epsilon < 1 - m/K_2$ es también necesaria para la existencia de solución en el caso $K = 0$ pues en caso de existir solución $\mathbf{V}$ del sistema (1.3), se cumple que $(1-\epsilon)E_{F_0}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}) = m$ y por lo tanto $\epsilon < 1 - m/K_2$.

NOTA 2: Las funciones $g_1^i(s,b)$ y $g_2(s,b)$ son estrictamente decrecientes en s para b fijo si la densidad de F_0 satisface que $0 \in \operatorname{sop}(F_0)$ o si ψ_2 es estrictamente creciente.

TEOREMA 1. *Supóngase que A2) - A5) valen, entonces*

a) *Si $\epsilon < 1/m$, el sistema (1.5) admite solución.*

b) *Si $\epsilon < \min\{1/K_2, (m-m_0)/(K_2-m_0)\}$, el sistema (1.6) admite solución.*

NOTA 3: En b) $(m-m_0)/(K_2-m_0) < 1/K_2$ si y sólo si $m_0 > K_2(m-1)/(K_2-1)$.

DEMOSTRACIÓN: Sean $s_1^i(b)$ y $s_2(b)$ las soluciones para $g_1^i(s,b) = m$ y $g_2(s,b) = m-1$ respectivamente. Debe notarse que $s_1^i(b)$ existe para todo b pues g_1^1 y g_1^2 son continuas como función de s en $(0,\infty)$, por (A1) y (A2).

$$\begin{aligned}
&\lim_{s\to\infty} g_1^1(s,b) = (1-\epsilon)m_0 + \epsilon m_0 = m_0 < m \text{ y}\\
&\lim_{s\to 0} g_1^1(s,b) = (1-\epsilon)K_2 + \epsilon K_2 = K_2 > m\\
&\lim_{s\to\infty} g_1^2(s,b) = (1-\epsilon)m_0 + \epsilon K_2 < m \left(\text{ sii } \epsilon < \frac{m-m_0}{K_2-m_0}\right) \text{ y}\\
&\lim_{s\to 0} g_1^2(s,b) = (1-\epsilon)K_2 + \epsilon K_2 = K_2 > m
\end{aligned}$$

Para estudiar cuando está definida $s_2(b)$, sea

$$h(b) = \int \frac{b^{m/(m-1)}v}{u + vb^{m/(m-1)}} \quad dF_0$$

la cual es una función creciente y continua de b tal que $\lim_{b\to\infty} h(b) = 1$ y $\lim_{b\to 0} h(b) = 0$. Sea b^* la solución de $(1-\epsilon)K_2h(b^*) = m-1$ y $b^{**} = \sup\{b \text{ tal que } (1-\epsilon)m_0h(b) \le m-1\}$. b^* existe pues $\epsilon < 1/m < 1-(m-1)/K_2$ con lo cual $(1-\epsilon)K_2 > m-1$. Luego por continuidad de h(b) existe b^* y $\lim_{s\to 0} g_2(s,b) = (1-\epsilon)K_2h(b) > m-1$ si $b > b^*$. Además $\lim_{s\to\infty} g_2(s,b) = (1-\epsilon)m_0h(b) < m-1$ si $b < b^{**}$. Como $b^* < b^{**}$, $s_2(b)$ existe si $b^* < b < b^{**}$ por continuidad de la función $g_2(s,b)$. Debe observarse que $b^{**} = \infty$ si $m_0 \le (m-1)/(1-\epsilon)$ (esta cota inferior es siempre menor que m por la suposición $\epsilon < 1/m$) y $b^{**} < \infty$ si $m > m_0 > (m-1)/(1-\epsilon)$. Nótese también que $b^* < 1$ si $\epsilon < 1 - m/K_2$ puesto que $(1-\epsilon)K_2h(1) = (1-\epsilon)K_2(m-1)/m > m-1$. Debe ahora probarse que para $i = 1$ o $i = 2$, $s_1^i(b)$ y $s_2(b)$ se intersectan en algún b y por lo tanto existirá solución para (1.5) y (1.6). Para ello se mostrará a continuación que

i) $\lim_{b \to b^{**}} s_2(b) = \infty$: supóngase por contradicción que $s_1 = \liminf_{b \to b^{**}} s_2(b) < \infty$. En caso que $m_0 \leq (m-1)/(1-\epsilon)$ entonces $\limsup_{b \to b^{**}} g_2(s_2(b), b) = (1-\epsilon)K_2 = m-1$ lo cual implica que $\epsilon = 1-(m-1)/K_2$ y ello es absurdo. Por otro lado, si $m_0 > (m-1)/(1-\epsilon)$ entonces $m-1 = (1-\epsilon)m_0h(b^{**}) < g_2(s_1, b^{**})$ y surge otro absurdo. En consecuencia $\lim_{b \to b^{**}} s_2(b) = \infty$ como se quería ver.

ii) Se desea probar que $s_2(b) > s_1^i(b)$ si $b^{**} > b > b_0$ para algún $b_0 > b^*$. Por contradicción se supondrán dos situaciones:

A) $s_1^i(b_n) \geq s_2(b_n)$ si $\lim_{n \to \infty} b_n = b^{**} = \infty$, obteniéndose en tal caso que

$$\begin{aligned}
m =& g_1^i(s_1^i(b_n), b_n) \leq g_1^i(s_2(b_n), b_n) \\
=& (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_2(b_n)}\left(\frac{u}{b_n} + vb_n^{1/(m-1)}\right)\right)\frac{b_n^{m/(m-1)}v}{u + b_n^{m/(m-1)}v} + \\
& (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_2(b_n)}\left(\frac{u}{b_n} + vb_n^{1/(m-1)}\right)\right)\frac{u}{u + b_n^{m/(m-1)}v} + \epsilon\psi_2\left(\frac{K}{s_2(b_n)b_n}\right) \\
=& (m-1) + (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_2(b_n)}\left(\frac{u}{b_n} + vb_n^{1/(m-1)}\right)\right)\frac{u}{u + b_n^{m/(m-1)}v} + \epsilon\psi_2\left(\frac{K}{s_2(b_n)b_n}\right)
\end{aligned}$$

El límite del miembro de la derecha cuando n tiende a infinito es $m-1+\epsilon m_0$ cuando se tiene (1.5) y $m-1+\epsilon K_2$ en caso de estar en (1.6) pues ψ_2 es acotada. La primer situación dice que $\epsilon m > \epsilon m_0 \geq 1$ y la última que $\epsilon K_2 \geq 1$ las cuales contradicen las suposiciones sobre ϵ respectivamente.

B) $s_1^i(b_n) \geq s_2(b_n)$ si $\limsup_{n \to \infty} b_n = b^{**} < \infty$. Como $\lim_{n \to \infty} s_2(b_n) = \infty$ entonces $\lim_{n \to \infty} s_1^i(b_n) = \infty$ pero ello implica que

$$m = \lim_{n \to \infty} g_1^1(s_1^1(b_n), b_n) = (1-\epsilon)m_0 + \epsilon m_0 < m \text{ en } (1.5)$$
$$m = \lim_{n \to \infty} g_1^2(s_1^2(b_n), b_n) = (1-\epsilon)m_0 + \epsilon K_2 < m \text{ en } (1.6)$$

Cualquier situación es absurda y por lo tanto vale $s_2(b) > s_1^i(b)$ si $b^{**} > b > b_0$ para algún $b_0 > b^*$ Debe observarse asimismo que $0 < \liminf_{b_n \to b} s_1^i(b_n) \leq \limsup_{b_n \to b} s_1^i(b_n) < \infty$ si $0 < b < \infty$ y que $0 < \liminf_{b_n \to b} s_2(b_n) \leq \limsup_{b_n \to b} s_2(b_n) < \infty$ si $b^* < b < b^{**}$. Además $g_1^i(s, b)$ y $g_2(s, b)$ son estrictamente decrecientes y continuas en s con lo cual $s_1^i(b)$ y $s_2(b)$

son continuas y dado que $s_2(b^*) = 0$ y $s_1(b^*) > 0$ debe deducirse que $s_1^i(b)$ y $s_2(b)$ deben cortarse en algún $b > b^*$. Luego queda probada la existencia de solución de los sistemas (1.5) y (1.6) ∎

El Teorema 1 y el Lema 1 b) dicen que la ecuación (1.3) con $\mathbf{V}$ diagonal admite solución. La unicidad de la solución de la ecuación (1.3) surge entonces de manera análoga a la prueba de Maronna (1976) y por lo tanto las soluciones de los sistemas (1.5) y (1.6) son únicas.

3.2 Punto de Ruptura de M-estimadores de dispersión.

Para poder estudiar el sesgo de un estimador es necesario determinar primeramente hasta que nivel de contaminación ϵ resiste el estimador sin escapar de todo conjunto compacto. O sea debe calcularse su punto de ruptura.

LEMA 2.

a) *Supóngase valen A1) - A6). Entonces el punto de ruptura de un M-estimador de dispersión (solución de (1.1)) es*

$$\epsilon^*(V, F_0) = \min\left\{\frac{1}{K_2}, \frac{m - m_0}{K_2 - m_0}, 1 - \frac{m}{K_2}\right\}$$

b) *Si el parámetro de posición es conocido, el punto de ruptura del estimador de Tyler es* $1/(m+1) \leq \epsilon_T^*(\mathbf{V}, F_0) \leq 1/m$.

NOTA 4: Al menos para $\epsilon < 1 \quad m/K_2$ puede asegurarse la existencia de solución para toda contaminación G (ver Huber-1981). Además, si $K_2 \geq m+1$ entonces $\epsilon^*(\mathbf{V}, F_0) \leq 1/(m+1)$, y si $K_2 < m+1$, $1 - m/K_2 < 1/(m+1)$ con lo cual $\epsilon^*(\mathbf{V}, F_0) \leq 1/(m+1)$ siempre.

DEMOSTRACIÓN: a) Para una sucesión de distribuciones $F_n = (1-\epsilon)F_0 + \epsilon G_n$, el sistema (1.1) resulta

$$(1-\epsilon)E_{F_0}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}} + \epsilon E_{G_n}\psi_2(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z}) = \mathbf{V}$$

$\mathbf{V}$ puede escribirse como $\mathbf{V} = \mathbf{Q}_n \Delta \mathbf{Q}_n'$ con $\mathbf{Q}_n$ una matriz ortogonal y Δ la matriz diagonal de autovalores de $\mathbf{V}$, por lo cual al pre y posmultiplicar por $\mathbf{Q}_n$ y $\mathbf{Q}_n'$ se obtiene que

$$(1-\epsilon)E_{F_0}\psi_2(\mathbf{z}'\Delta^{-1}\mathbf{z})\frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'\Delta^{-1}\mathbf{z}} + \epsilon E_{\tilde{G}_n}\psi_2(\tilde{\mathbf{z}}'\Delta^{-1}\tilde{\mathbf{z}})\frac{\tilde{\mathbf{z}}\tilde{\mathbf{z}}'}{\tilde{\mathbf{z}}'\Delta^{-1}\tilde{\mathbf{z}}} I_{\Re^m - \{0\}}(\mathbf{z}) = \Delta$$

donde $\tilde{\mathbf{z}} = \mathbf{Q}_n'\mathbf{z}$ tiene distribución $\tilde{G}_n$ si $\mathbf{z}$ tiene distribución G_n. Pre y posmultiplicando por el vector canónico $\mathbf{e}_s'$ y $\mathbf{e}_s$ la igualdad anterior se convierte en m ecuaciones de la forma siguiente

$$(1-\epsilon)E_{K_0}\psi_2\left(\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right)\frac{u_s}{\sum_{i=1}^{m}u_i\lambda_s(\lambda_i)^{-1}}+\epsilon E_{\tilde{G}_n}\psi_2(\tilde{\mathbf{z}}'\Delta^{-1}\tilde{\mathbf{z}})\frac{\tilde{z}_s^2}{\sum_{i=1}^{m}\tilde{z}_i^2\lambda_s(\lambda_i)^{-1}}I_{\Re^m-\{0\}}(\tilde{\mathbf{z}})=1$$

donde $u_s = z_s^2$ y K_0 es la distribución del vector $(u_1, u_2, \ldots, u_m)$ si $(z_1, z_2, \ldots, z_m)$ tiene distribución F_0. Se considerarán a continuación las siguientes situaciones donde el estimador escapa de cualquier conjunto compacto cuando se mueve la sucesión de contaminaciones (sin pérdida de generalidad se trabajará con la sucesión G_n aunque sea necesario restringir a subsucesiones)

(i) Si $\limsup_{n\to\infty}\dfrac{\lambda_{(1)}(F_n)}{\lambda_{(m)}(F_n)} = \infty$ y $a(G_n) = E_{\tilde{G}_n}\psi_2(\tilde{\mathbf{z}}'\Delta^{-1}\tilde{\mathbf{z}})\dfrac{\tilde{z}_{(1)}^2}{\sum_{i=1}^{m}\tilde{z}_i^2\lambda_{(1)}(\lambda_i)^{-1}}I_{\Re^m-\{0\}}(\mathbf{z})$ entonces resulta que

$$\epsilon K_2 \geq \epsilon \limsup_{n\to\infty} a(G_n) = 1 \qquad \text{y por lo tanto} \qquad \epsilon \geq 1/K_2$$

(ii) Si $\limsup_{n\to\infty}\lambda_{(1)}(F_n) = \infty$ pero $0 < \inf_n \dfrac{\lambda_{(1)}(F_n)}{\lambda_{(m)}(F_n)} < \sup_n \dfrac{\lambda_{(1)}(F_n)}{\lambda_{(m)}(F_n)} < \infty$ entonces

$$\begin{aligned}&(1-\epsilon)m_0+\epsilon K_2 \geq\\ &(1-\epsilon)\lim_{n\to\infty}E_{K_0}\psi_2\left(\frac{1}{\lambda_{(1)}(F_n)}\left(u_{(1)}+\sum_{(j)\neq(i)}\frac{\lambda_{(1)}}{\lambda_{(j)}}u_{(j)}\right)\right)+\\ &\epsilon\limsup_{n\to\infty}E_{\tilde{G}_n}\psi_2(\tilde{\mathbf{z}}'\Delta^{-1}\tilde{\mathbf{z}})I_{\Re^m-\{0\}} = m\end{aligned}$$

con lo cual surge que $\epsilon \geq (m-m_0)/(K_2-m_0)$. (iii) Si $\liminf_{n\to\infty}\lambda_{(1)}(G_n) = 0$ entonces

$$m \geq (1-\epsilon)\lim_{n\to\infty}E_{K_0}\psi_2\left(\frac{u_{(1)}}{\lambda_{(1)}(G_n)}+\sum_{(j)\neq(1)}\frac{u_{(j)}}{\lambda_{(j)}}\right) = (1-\epsilon)K_2$$

lo que implica que $\epsilon \geq 1 - m/K_2$.

Para obtener la otra desigualdad debe suponerse ahora que

$$0 < \inf_{G\in\mathcal{F}_\epsilon}\lambda_{(m)}(G) \leq \sup_{G\in\mathcal{F}_\epsilon}\lambda_{(1)}(G) < \infty$$

Si se toman contaminaciones del tipo $\delta_{(K^{1/2},0,\ldots,0)}, K \in \Re^+$, debe observarse que

$$(1-\epsilon)\lim_{K\to\infty} E_{F_0}\psi_2\left(\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right)\frac{u_1}{\sum_{i=1}^{m} u_i\lambda_1(\lambda_i)^{-1}} + \epsilon K_2 = 1$$

lo cual dice que $\epsilon \leq 1/K_2$. Por otro lado, si $K = 0$ se obtiene que

$$(1-\epsilon)K_2 \geq (1-\epsilon)E_{F_0}\psi_2\left(\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right) = m$$

lo cual implica que $\epsilon \leq 1 - m/K_2$. Si ahora se consideran contaminaciones G_K tal que su densidad es $g_K = K^{-m/2}f_0\left(1/K\sum_{i=1}^{m} u_i\right)$ resulta que

$$\begin{aligned}
m &= (1-\epsilon)E_{K_0}\psi_2\left(\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right) + \epsilon\int\psi_2\left(\sum_{i=1}^{m}\frac{z_i^2}{\lambda_i}\right)K^{-m/2}f_0\left(\frac{1}{K}\sum_{i=1}^{m} z_i^2\right)dz_1\ldots dz_m \\
&= (1-\epsilon)E_{K_0}\psi_2\left(\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right) + \epsilon\int\psi_2\left(K\sum_{i=1}^{m}\frac{z_i^2}{\lambda_i}\right)f_0\left(\sum_{i=1}^{m} z_i^2\right)dz_1\ldots dz_m \\
&\geq (1-\epsilon)m_0 + \epsilon E_{K_0}\psi_2\left(K\sum_{i=1}^{m}\frac{u_i}{\lambda_i}\right)
\end{aligned}$$

Al tomar límite para K tendiendo a ∞ resulta que $m \geq (1-\epsilon)m_0 + \epsilon K_2$ o equivalentemente $\epsilon \leq (m - m_0)/(K_2 - m_0)$. Por lo tanto surge que

$$\epsilon^*(V) = \min\left\{\frac{1}{K_2}, \frac{m-m_0}{K_2-m_0}, 1-\frac{m}{K_2}\right\}$$

b) En la Sección 4. se muestra que $\forall\ \epsilon < 1/(m+1)$ existe el estimador de Tyler. La condición adicional para unicidad, $\mathrm{tr}(\mathbf{V}) = m$ asegura que el supremo del máximo autovalor de la matriz $\mathbf{V}$ en el entorno de contaminación es finito. Asimismo el ínfimo del máximo autovalor de la matriz $\mathbf{V}$ en el entorno de contaminación es mayor que 0. Usando entonces la notación del ítem a) la única situación que podría darse es que $\limsup_{n\to\infty}\dfrac{\lambda_{(1)}(F_n)}{\lambda_{(m)}(F_n)} = \infty$ y con el mismo razonamiento de allí resulta que $\epsilon > 1/m$ en contra de lo que se había supuesto. Por lo tanto resulta que el ínfimo del mínimo autovalor es mayor que 0 en el entorno de contaminación, y luego $\epsilon_T^*(\mathbf{V}, F_0) \geq 1/(m+1)$ como se deseaba ver. En caso que la solución buscada sea la que satisface $\det(\mathbf{V}) = 1$, entonces basta tomar la solución que cumple $\mathrm{tr}(\mathbf{V}) = m$ y finalmente dividirla por $\det(\mathbf{V})^{1/m}$. Luego también se va a satisfacer que su punto de ruptura es mayor o igual que $1/(m+1)$ pues todos los autovalores permanecen en un compacto de $\Re^+$. ∎

NOTA 5: El punto de ruptura del estimador de Tyler es exactamente $1/m$ si se restringe al subconjunto de contaminaciones puntuales Δ.

COROLARIO 1. *Supóngamos A2)-A5) son satisfechas, y sean ϵ, ϵ_0 menores que $1/m$. Entonces resulta que*

(a) $\liminf_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) > 0$, $K_0 \in [0, \infty]$.

(b) $\limsup_{K \to K_0, \epsilon \to \epsilon_0} s(K, \epsilon) = \infty$ *sii* $K_0 = \infty$ *y* $1/m > \epsilon_0 \geq \min\{1/K_2, (m - m_0)/(K_2 - m_0\}$.

(c) $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) = \infty$ *sii* $K_0 = \infty$ *y* $1/m > \epsilon_0 \geq \min\{1/K_2, (m - m_0)/(K_2 - m_0)\}$.

(d) $\liminf_{K \to K_0, \epsilon \to \epsilon_0} s(K, \epsilon) = 0$ *sii* $K_0 = 0$ *y* $\epsilon_0 \geq 1 - (m - m_0)/(K_2 - m_0)$

DEMOSTRACIÓN:

(a) En caso que no fuera cierta la afirmación, surgiría un absurdo de la segunda ecuación del sistema (1.5) ya que en tal caso aparecería $0 = m - 1$ tras tomar límite.

(b) (SOLO SI) Si $\liminf_{K \to K_0, \epsilon \to \epsilon_1} s(K, \epsilon) = \infty$ con $K_0 \in [0, \infty)$ y $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) = \infty$ entonces la primer ecuación de (1.4) dice que $\epsilon_0 m_0 = 1$, y si $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) < \infty$ entonces de la primer ecuación de (1.5) surge que $m = (1 - \epsilon_0)m_0 + \epsilon_0 m_0$, siendo absurdas cualquiera de las situaciones. Por lo tanto $K_0 = \infty$. Nuevamente, si $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) < \infty$ se deduce de la primer ecuación de (1.5) que $(1 - \epsilon_0)m_0 + \epsilon_0 K_2 \geq m$, lo cual implica que $\epsilon_0 \geq (m - m_0)/(K_2 - m_0)$, y si $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K, \epsilon) = \infty$ la primer ecuación de (1.4) dice $\epsilon_0 \geq 1/K_2$.

(SI) En caso que $\limsup_{K \to \infty, \epsilon \to \epsilon_0} s(K, \epsilon) < \infty$ pero $\liminf_{K \to \infty, \epsilon \to \epsilon_0} b(K, \epsilon) = \infty$, entonces la segunda ecuación dice que $(1 - \epsilon_0)K_2 = m - 1$ y por lo tanto $\epsilon_0 = 1 - (m-1)/K_2 > 1/m$ lo cual es absurdo. Si $\limsup_{K \to \infty, \epsilon \to \epsilon_0} b(K, \epsilon) < \infty$ entonces la primer ecuación de (1.4) y de (1.5) dicen respectivamente que $\epsilon_0 K_2 \leq 1$ y $(1 - \epsilon_0)m_0 + \epsilon_0 K_2 < m$ pero ello contradice la desigualdad supuesta para ϵ_0.

(c) Sigue fácilmente del ítem anterior.

(d) Sigue en forma totalmente similar a lo hecho en b). ∎

NOTA 6: En medio de la demostración del ítem b) surge que nunca puede ocurrir simultáneamente que $\limsup_{K \to K_0, \epsilon \to \epsilon_0} s(K,\epsilon) < \infty$ y $\limsup_{K \to K_0, \epsilon \to \epsilon_0} b(K,\epsilon) = \infty$ o viceversa.

3.3 Comportamiento de los M-estimadores bajo contaminaciones puntuales.

LEMA 3. *Bajo A2)-A5) y si $\epsilon < 1/m$, la función $(s(K,\epsilon), b(K,\epsilon))$ es continua en $(K,\epsilon) \in (\Re^+ \times (0, 1/m)$.*

DEMOSTRACIÓN: Si $\lim_{n\to\infty} K_n = K_0 \in (0,\infty)$ y $\lim_{n\to\infty} \epsilon_n = \epsilon_0 \in (0, 1/m)$, por el Corolario 1 $(s(K_n,\epsilon), b(K_n,\epsilon))_{n\in\mathcal{N}}$ permanece acotada, y tomando una subsucesión si fuera necesario puede decirse que $\lim_{n\to\infty} s(K_n,\epsilon_n) = s_0$ y $\lim_{n\to\infty} b(K_n,\epsilon_n) = b_0$. Por continuidad de ψ_2, (s_0, b_0) es solución de (1.5) si $K = K_0$ y $\epsilon = \epsilon_0$. Luego por unicidad de soluciones $(s_0,b_0) = (s(K_0,\epsilon_0), b(K_0,\epsilon_0))$, con lo cual $(s(K,\epsilon), b(K,\epsilon))$ es continua, resultando la afirmación del lema. ∎

LEMA 4. *Bajo A2)-A5) y si $\epsilon < 1/m$, la función $b(K,\epsilon)$ es creciente en $K > 0$.*

DEMOSTRACIÓN: Sea $K_0 < K_1 < \infty$ entonces $g_1(s,b,K_0) \le g_1(s,b,K_1)$ y por lo tanto $m = g_1(s_1(b,K_0),b,K_0) \le g_1(s_1(b,K_0),b,K_1)$. Como $g_1(s,b,K)$ es decreciente en s, entonces

$$s_1(b,K_1) \ge s_1(b,K_0) \tag{1.8}$$

Sean (s_0,b_0) y (s_1,b_1) las soluciones para el sistema (1.5) cuando las contaminaciones son K_0 y K_1 respectivamente. Entonces $s_0 = s_1(b_0,K_0)$ y $s_1 = s_1(b_1,K_1)$. Por lo expuesto en la prueba de existencia de soluciones para (1.5) y por unicidad de las mismas resulta que

$$\begin{array}{ll} b > b_0 & s_1(b,K_0) < s_2(b,K_0) \\ b < b_0 & s_1(b,K_0) > s_2(b,K_0) \\ b > b_1 & s_1(b,K_1) < s_2(b,K_1) = s_2(b,K_0) \\ b < b_1 & s_1(b,K_1) > s_2(b,K_1) = s_2(b,K_0) \end{array} \tag{1.9}$$

En caso que $b_1 < b_0$ entonces por (1.8) y (1.9) resulta que

$$s_1 = s_1(b_1,K_1) \ge s_1(b_1,K_0) > s_2(b_1,K_0) = s_2(b_1,K_1) = s_1(b_1,K_1) = s_1$$

lo cual es absurdo. En consecuencia $b_0 \le b_1$ y luego $b(K_0) \le b(K_1)$ si $K_0 < K_1$. ∎

COROLARIO 2. *Si* $\epsilon < \min\{1/K_2, (m - m_0)/(K_2 - m_0)\}$ *entonces* $b(K,\epsilon) < b(\epsilon)$ *donde* $b(\epsilon)$ *soluciona (1.6) y* $\lim_{K\to\infty} b(K,\epsilon) = b(\epsilon)$.

En Maronna-Yohai(1990) aparece el lema siguiente, del cual será necesaria una generalización para el problema de interés.

LEMA. *Sea* $\mathbf{z} \in \Re^m$ *un vector aleatorio tal que su distribución admite densidad* $f(\mathbf{z})$ *decreciente en* $\|\mathbf{z}\|$. *Sea* $1 \le m' \le m$ *y sean* $u = u(z) = \sum_{i=1}^{m'} z_i^2$ *y* $v = v(z) = \sum_{i=m'+1}^{m} z_i^2$. *Si* F_b *denota la distribución de* $z(b) = u/b + v/a$ *donde* $b^{m'}a^{m-m'} = 1$, *entonces para cada* t, $F_b(t)$ *es una función decreciente de* b *si* $b \ge 1$ *y es creciente para* $b \le 1$

En todo lo que sigue se supondrá que

A7) La densidad del vector $\mathbf{z}$ bajo el modelo es decreciente en $\|\mathbf{z}\|$.

COROLARIO 3. *Si valen A2), A3) y A7) entonces la función*

$$p(b) = E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + b^{m'/(m-m')}v\right)\right)$$

es monótona creciente si $b \ge 1$ *y es monótona decreciente para* $b \le 1$

DEMOSTRACIÓN: Puede reescribirse el lema previo diciendo que

$$\int I_{(0,t)}(\mathbf{z}'\Lambda^{-1}\mathbf{z})f_0(\mathbf{z}'\mathbf{z})d\mathbf{z} < \int I_{(0,t)}(\mathbf{z}'\Delta^{-1}\mathbf{z})f_0(\mathbf{z}'\mathbf{z})d\mathbf{z}$$

donde Λ y Δ son matrices diagonales tal que $\lambda_{ii} = b$ si $1 \le i \le m'$ y $\lambda_{ii} = a$ si $m'+1 \le i \le m$ y $\delta_{ii} = b'$ si $1 \le i \le m'$ y $\delta_{ii} = a'$ si $m'+1 \le i \le m$ con $b > b' > 1$. Si ahora se toma una función escalera $g = \sum_{i=1}^{r} g_i I_{E_i}$ donde los conjuntos E_i son intervalos disjuntos tal que $\bigcup_{i=1}^{r} E_i = \Re^+, E_i = [b_i, c_i) \quad c_i = b_{i+1}, \quad b_1 = 0$, entonces resulta que

$$\int g(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} = \sum_{i=1}^{r} g_i \int I_{E_i}(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} =$$

$$\sum_{i=1}^{r} g_i \left(\int I_{[0,b_{i+1})}(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} - \int I_{[0,b_i)}(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \right) =$$

$$\sum_{i=1}^{r-1} \left(\int I_{[0,b_{i+1})}(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \right) (g_i - g_{i+1}) + g_r \int I_{[0,b_{r+1})}(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \leq$$

$$\sum_{i=1}^{r-1} \left(\int I_{[0,b_{i+1})}(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \right) (g_i - g_{i+1}) + g_r \int I_{[0,b_{r+1})}(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} =$$

$$\int g(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z}$$

Por lo tanto

$$\int g(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} < \int g(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z}$$

si g es una función escalera.

Si $g(x)$ es una función no creciente acotada en $\Re^+$, existe $\{s_n\}_{n \in \mathcal{N}}$ sucesión decreciente de funciones escalera tal que $\lim_{n \to \infty} s_n(x) = g(x)$ uniformemente. En consecuencia, por teorema de convergencia monótona surge que

$$\int g(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \leq \int g(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z}$$

Tomando ahora $\quad \psi_2(x) = g(0) - g(x)$ resulta que

$$\int \psi_2(\mathbf{z}'\Lambda^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z} \geq \int \psi_2(\mathbf{z}'\Delta^{-1}\mathbf{z}) f_0(\mathbf{z}'\mathbf{z}) d\mathbf{z}$$

surgiendo el resultado deseado. ■

Se ha obtenido el comportamiento de $b(K, \epsilon)$ bajo la acción de las contaminaciones puntuales. Si se quisiera ver que $s(K, \epsilon)$ posee las mismas propiedades de monotonía será necesario probar el Teorema 2 que sigue a continuación, el cual se prueba primeramente suponiendo que puede intercambiarse derivación respecto de s o b y el signo integral; tal situación se cumpliría, por ejemplo, en caso de ψ_2 con derivada continua y acotada. Entonces es necesario el siguiente el siguiente lema técnico, que se usará también en el Lema 9 y en el Teorema 4.

LEMA 5. *Si $\psi_2\colon \Re^+ \to \Re^+$ es una función monótona creciente, continua y acotada, entonces*

a) *Existe una sucesión de funciones $\psi_2^{(n)}\colon \Re^+ \to \Re^+$ crecientes, acotadas, diferenciables con derivada continua, positiva y acotada tal que $\lim_{n\to\infty} \psi_2^{(n)}(x) = \psi(x)$ uniformemente en x.*

b) *$\psi_2^{(n)}$ puede elegirse de forma de satisfacer la siguiente condición: existe una sucesión $\{x^{(n)}\} \in \Re^+$ tal que $\psi_2^{(n)}(x^{(n)}) = K_2 = \sup \psi_2(x)$.*

DEMOSTRACIÓN: a) Para ver la existencia de la sucesión de funciones aproximantes puede tomarse la convolución de ψ_2 con una sucesión de núcleos adecuados, por ejemplo $\psi_2^{(n)}(z) = (\psi_2 * f_n)(z) = \int \psi_2(z-x) f_n(x)dx$ donde $f_n(x) = 1/\sqrt{2\pi\sigma_n^2}\exp\left(-1/(2\sigma_n^2)x^2\right)$ y ψ_2 se ha extendido por paridad sobre los reales negativos. Como

$$(\psi_2 * f_n)(z) = \int \psi_2(z-x) f_n(x)dx = \int \psi_2(x) f_n(z-x)dx = \int \psi_2(z-\sigma_n x) f(x)dx$$

donde $f(x) = 1/\sqrt{2\pi}\exp\left(-1/2x^2\right)$, esta función resulta ser creciente con derivada continua y acotada (en realidad existirán derivadas de todos los órdenes). Si ahora se toma $\mathbf{C} = [-N, N]$ tal que $\int_{\mathbf{C}} f(x)dx > 1 - \delta/(2K_2)$ entonces

$$\int \psi_2(z-\sigma_n x) f(x)dx \le \int_{\mathbf{C}} \left(\psi_2(z-\sigma_n x) - \psi_2(z)\right) f(x)dx + \delta/2$$

Como una función continua, creciente y acotada en $\Re^+$ es uniformemente continua, entonces ψ_2 es uniformemente continua, y por lo tanto debe tomarse σ_n suficientemente pequeño para poder asegurar que $|\psi_2(z-\sigma_n x) - \psi_2(z)| < \delta/2$ si $x \in \mathbf{C}$ (ello independientemente de z). Por lo tanto se obtiene que

$$\lim_{n\to\infty} (\psi_2 * f_n)(z) - \psi_2(z) = \lim_{\sigma_n\to 0} \int \left(\psi_2(z-\sigma_n x) - \psi_2(z)\right) f(x)dx = 0$$

uniformemente en z. Como

$$(\psi_2 * f_n)'(\mathbf{z}) = \int \psi_2(x) f_n'(z-x)dx = -\int \psi_2(x) \frac{1}{\sigma_n^2\sqrt{2\pi\sigma_n^2}}(z-x)\exp\left(-\frac{1}{2\sigma_n^2}(z-x)^2\right)dx$$

entonces

$$\begin{aligned}(\psi_2 * f_n)'(\mathbf{z}) = &-\int \psi_2(x) \frac{1}{\sigma_n^2\sqrt{2\pi\sigma_n^2}}(z-x)\exp\left(-\frac{1}{2\sigma_n^2}(z-x)^2\right) I_{(-\infty,z]}(x)dx \\ &-\int \psi_2(x) \frac{1}{\sigma_n^2\sqrt{2\pi\sigma_n^2}}(z-x)\exp\left(-12\sigma_n^2(z-x)^2\right) I_{[z,\infty)}(x)dx > \\ &-\psi_2(z)\int \frac{1}{\sigma_n^2\sqrt{2\pi\sigma_n^2}}(z-x)\exp\left(-\frac{1}{2\sigma_n^2}(z-x)^2\right)dx = 0\end{aligned}$$

por imparidad de la función integrada.

b) Sea $f(z) = \exp(-1/z)I_{(0,\infty)}(z)$ y sea $g(z) = f(z)/(f(z) + f(1-z))$. Sea ahora $\phi_n(z) = 1 - g(z+n+1)g(n+1-z)$ la cual resulta ser una función con derivadas de todos los órdenes acotadas tal que $\phi_n(z) = 0$ si $|z| \leq n$, $\phi_n(z) = 1$ si $|z| \geq n+1$ y $0 \leq \phi_n(z) \leq 1$ $\forall\ z$. Defínase entonces $\psi_2^{(n)} = (1 - \phi_n)\psi_2 * f_n + \phi_n K_2$, la cual cumple $\psi_2^{(n)}(z) = K_2$ si $z \geq n+1$. Tal función es creciente a partir de un n suficientemente grande, pues el único sector que habría que chequear es el intervalo $[n, n+1]$ y la función allí toma la forma $\psi_2 * f_n + \phi_n(K_2 - \psi_2 * f_n)$. Si n es grande, $\psi_2 * f_n$ comanda el comportamiento pues el otro sumando es suficientemente pequeño, resultando así la función creciente. Para ver que $\lim_{n\to\infty} \psi_2^{(n)}(z) - \psi_2(z) = 0$ uniformemente en z, sea $\epsilon > 0$. Entonces $\exists\ n_0(\epsilon)$ tal que si $n \geq n_0$ se cumple que $(K_2 - \psi_2) \leq \epsilon/2$ si $|z| > n$ y $|(\psi_2 * f_n(z) - \psi_2(z))| \leq \epsilon/2$ (por parte a)). Luego

$$|\psi_2^{(n)}(z) - \psi_2(z)| = |(1-\phi_n)(\psi_2 * f_n - \psi_2) + \phi_n(K_2 - \psi_2)| \leq \begin{cases} |\psi_2 * f_n - \psi_2| & \text{si } |z| \leq n \\ \epsilon & \text{si } |z| > n \end{cases}$$

quedando así probada la afirmación. ∎

TEOREMA 2. *Si valen A2),A3) y A7), la función*

$$f(b) = E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u + vb^{m/(m-1)}}$$

es creciente en $b \geq 1$ para cada s fijo.

DEMOSTRACIÓN: Recuérdese que $u = u(z) = z_1^2$ y que $v = v(z) = \sum_{i=2}^m z_i^2$. Al derivar $f(b)$ respecto de b se obtiene que

$$\frac{1}{sb^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\left(-u + \frac{1}{m-1}vb^{m/(m-1)}\right)\frac{b^{1/(m-1)}v}{u/b + vb^{1/(m-1)}}f_0(u+v)dz_1\dots dz_m$$
$$+\int \psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\frac{b^{1/(m-1)}vum/(m-1)}{(u + vb^{m/(m-1)})^2}f_0(u+v)dz_1\dots dz_m \qquad (1.10)$$

El segundo sumando es positivo y se verá que el primero también lo es; para ello se hará cambio de variable a coordenadas elípticas:

$$
\begin{aligned}
z_1 &= \sqrt{b} r \cos\theta_1 \\
z_2 &= \sqrt{a} r \sin\theta_1 \cos\theta_2 \\
z_3 &= \sqrt{a} r \sin\theta_1 \sin\theta_2 \cos\theta_3 \\
\\
z_{m-2} &= \sqrt{a} r \sin\theta_1 \sin\theta_2 \dots \cos\theta_{m-2} \\
z_{m-1} &= \sqrt{a} r \sin\theta_1 \sin\theta_2 \dots \sin\theta_{m-2} \cos\theta_{m-1} \\
z_m &= \sqrt{a} r \sin\theta_1 \sin\theta_2 \dots \sin\theta_{m-2} \sin\theta_{m-1}
\end{aligned}
$$

donde $a = b^{-1/(m-1)}$, $0 \leq \theta_i \leq \pi$ si $1 \leq i \leq m-2$ y $0 \leq \theta_{m-1} \leq 2\pi$. El jacobiano de la transformación resulta ser $|J| = r^{m-1} \sin^{m-2}\theta_1 \sin^{m-3}\theta_2 \dots \sin\theta_{m-2}$ (ver Muirhead-1982-pág.36), si además

$$
h_0(r,\theta_1) = f_0(r^2(b\cos^2\theta_1 + b^{-1/(m-1)}\sin^2\theta_1)) = f_0(r^2(b + \sin^2\theta_1(b - b^{-1/(m-1)}))
$$

al hacer dicho cambio de variable en el primer sumando de (1.10) surge que

$$
\begin{aligned}
&\frac{1}{sb^2}\int \psi_2'\left(\frac{r^2}{s}\right)\left(\frac{-br^2\cos^2\theta_1 + (m-1)^{-1}b^{m/(m-1)}r^2 b^{-1/(m-1)}\sin^2\theta_1}{br^2\cos^2\theta_1 + br^2\sin^2\theta_1}\right) b^{m/(m-1)}b^{-1/(m-1)} \\
&\quad \times r^2\sin^2\theta_1 r^{m-1}\sin^{m-2}\theta_1\sin^{m-3}\theta_2\dots\sin\theta_{m-2}h_0(r,\theta_1)dr d\theta_1\dots d\theta_{m-1}
\end{aligned}
$$

o más precisamente que

$$
\begin{aligned}
&\frac{1}{sb(m-1)}\int \psi_2'\left(\frac{r^2}{s}\right)(-(m-1)\cos^2\theta_1 + \sin^2\theta_1) \\
&\qquad \times r^{m+1}\sin^m\theta_1\sin^{m-3}\theta_2\dots\sin\theta_{m-2}h_0(r,\theta_1)dr d\theta_1\dots d\theta_{m-1} \qquad (1.11)
\end{aligned}
$$

Las integrales de $\sin^j\theta_s, s = 2,\dots,m-2$ dan positivas pues se integra en $(0,\pi)$, resultando que (1.11) se convierte en la integral siguiente

$$
\frac{H(m)}{sb(m-1)}\int_0^\pi\int_0^\infty \psi_2'\left(\frac{r^2}{s}\right)(-(m-1) + m\sin^2\theta_1)r^{m+1}\sin^{m-2}\theta_1\sin^2\theta_1 h_0(r,\theta_1)dr d\theta_1 \tag{1.12}
$$

Si $\theta^* \in (0,\pi/2)$ satisface que $\sin\theta^* = (m-1)/m$ y se define $\mathbf{A} = \{\theta_1 : \theta^* < \theta_1 < \pi - \theta^*\}$, entonces el integrando de (1.12) es negativo sii $-(m-1) + m\sin^2\theta_1 < 0$ sii $\sin^2\theta_1 < (m-$

$1)/m = 1 - 1/m$ sii $\theta_1 \in \mathbf{A}^c = [0,\pi] - \mathbf{A} = \{\theta_1 : 0 \leq \theta_1 < \theta^* < \pi/2 < \pi - \theta^* < \theta_1 < \pi\}$, con lo cual la integral (1.12) puede reescribirse como

$$\frac{H(m)}{sb(m-1)} \int_{\mathbf{A}} \int_0^\infty \psi_2' \left(\frac{r^2}{s}\right) (-(m-1) + m\sin^2\theta_1) r^{m+1} \sin^{m-2}\theta_1 \sin^2\theta_1 h_0(r,\theta_1) dr d\theta_1$$
$$+ \int_{\mathbf{A}^c} \int_0^\infty \psi_2' \left(\frac{r^2}{s}\right) (-(m-1) + m\sin^2\theta_1) r^{m+1} \sin^{m-2}\theta_1 \sin^2\theta_1 h_0(r,\theta_1) dr d\theta_1 \quad (1.13)$$

Si $\theta_1 \in \mathbf{A}$ entonces $\sin^2\theta_1 \geq \sin^2\theta^* = (m-1)/m$ y si $\theta_1 \in \mathbf{A}^c$ resulta $\sin^2\theta_1 \leq \sin^2\theta^*$ y al multiplicar por algo negativo se da vuelta la desigualdad, con lo cual (1.13) es mayor o igual que

$$\frac{H(m)\sin^2\theta^*}{sb(m-1)} \int_{\mathbf{A}} \int_0^\infty \psi_2' \left(\frac{r^2}{s}\right) (-(m-1) + m\sin^2\theta_1) r^{m+1} \sin^{m-2}\theta_1 h_0(r,\theta_1) dr d\theta_1 \quad +$$
$$\frac{H(m)\sin^2\theta^*}{sb(m-1} \int_{\mathbf{A}^c} \int_0^\infty \psi_2' \left(\frac{r^2}{s}\right) (-(m-1) + m\sin^2\theta_1) r^{m+1} \sin^{m-2}\theta_1 h_0(r,\theta_1) dr d\theta_1$$
$$= \frac{H(m)\sin^2\theta_1}{sb(m-1)} \int_0^\pi \int_0^\infty \psi_2' \left(\frac{r^2}{s}\right) (-(m-1) + m\sin^2\theta_1) r^{m+1} \sin^{m-2}\theta_1 h_0(r,\theta_1) dr d\theta_1$$

y esta integral es la que corresponde a la derivada de la función $E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)$ la cual es creciente en $b \geq 1$ y por lo tanto tiene derivada no negativa. Entonces la integral estudiada es positiva y $f(b)$ es creciente en b como se quería ver. Por el lema 5 toda función continua, creciente y acotada puede ser aproximada uniformemente por funciones crecientes de derivada acotada siendo entonces el resultado válido para las funciones de interés. ∎

En lo que sigue se verá que el autovector asociado al máximo autovalor de la matriz **V** corresponderá siempre al eje donde ocurre la contaminación puntual. Ello equivale a probar que la solución $b(\epsilon)$ de (1.6) es mayor que 1, y para ver eso se tomará una familia de funciones indexadas en un parámetro c del tipo $\psi_2^c = (1-c)\psi_2 + cm$. Siendo las soluciones $(\bar{s}(c), \bar{b}(c))$ continuas en $c \in [0,1]$ resultará que $\bar{b}(c)$ tenderá a la solución de las ecuaciones de Tyler para contaminaciones puntuales, la cual es mayor que 1. Si $b(\epsilon) = \bar{b}(0) < 1$ entonces por continuidad existiría c^* tal que $\bar{b}(c^*) = 1$ y ello es imposible. La idea intuitiva de este absurdo radica en que un M-estimador monótono no debería ser un múltiplo de la matriz identidad (tal como resultaría en el caso de tener una distribución esféricamente simétrica) cuando la distribución es claramente asimétrica. En lo que sigue se desarrollarán estas ideas aquí expuestas.

Sean pues $m_0^{(c)} = \psi_2^{(c)}(0)$ y $K_2^{(c)} = \psi_2^{(c)}(\infty)$. Para la existencia de soluciones y cálculo de punto de ruptura en el sistema (1.6) se deben estudiar las cantidades

$$(K_2^{(c)})^{-1} = \frac{1}{(1-c)K_2 + cm} > \frac{1}{K_2}$$

$$\frac{m - m_0^{(c)}}{K_2^{(c)} - m_0^{(c)}} = \frac{m - [(1-c)m_0 + cm]}{[(1-c)K_2 + cm] - [(1-c)m_0 + cm]} = \frac{(1-c)(m-m_0)}{(1-c)(K_2 - m_0)} = \frac{m - m_0}{K_2 - m_0} \text{ si } c \neq 1$$

Por lo tanto $\epsilon^*(\psi_2^1) \geq 1/(m+1)$ y

$$\epsilon^*(\psi_2^{(c)}) = \min\left\{\frac{1}{(1-c)K_2 + cm}, \frac{m - m_0}{K_2 - m_0}, 1 - \frac{m}{(1-c)K_2 + cm}\right\} \text{ si } c \neq 1$$

Aún cuando $1 - m/((1-c)K_2 + cm)$ tiende a 0 si c tiende a 1, para la existencia de soluciones de (1.6) sólo interesa que

$$\epsilon < \min\{1/K_2, (m - m_0)/(K_2 - m_0)\} \leq \min\{1/((1-c)K_2 + cm), (m - m_0)/(K_2 - m_0)\}$$

El sistema (1.6) resulta para $\psi_2^{(c)}$ de la siguiente manera,

$$\begin{aligned}
&(1-c)\left[(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right) + \epsilon K_2\right] + cm = m \\
&(1-c)\left[(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u + vb^{m/(m-1)}}\right] + \\
&c(1-\epsilon)mE_{F_0}\frac{vb^{m/(m-1)}}{u + vb^{m/(m-1)}} = m - 1
\end{aligned}$$

lo cual equivale, si $c \neq 1$, al sistema:

$$\begin{aligned}
&(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right) + \epsilon K_2 = m \\
&(1-c)\left[(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u + vb^{m/(m-1)}}\right] + \\
&c(1-\epsilon)mE_{F_0}\frac{vb^{m/(m-1)}}{u + vb^{m/(m-1)}} = m - 1
\end{aligned} \tag{1.14}$$

Para ver la continuidad de las soluciones $(\bar{s}(c), \bar{b}(c))$ de (1.14) se usará el siguiente lema que muestra la continuidad de los M-estimadores respecto de las funciones ψ_2.

LEMA 6. *Supóngase se cumplen (A1) - (A6). Si* $\mathbf{V} = \sum_{i=1}^{m} \lambda_i(\mathbf{V}) t_i t_i'$ *es simétrica definida positiva en* $\Re^{m\times m}$, *con* $\lambda_i(\mathbf{V})$ *el i-ésimo autovalor de* $\mathbf{V}$, t_i *el autovector asociado y* $\lambda_1 \geq \ldots \geq \lambda_m$ *entonces*

(a) $\lambda_m(\mathbf{V}) \leq \mathbf{z}'\mathbf{z}/(\mathbf{z}'\mathbf{V}^{-1}\mathbf{z}) \leq \lambda_1(\mathbf{V}) \ \forall \ \ \mathbf{z} \in \Re^m - \{0\}$.

(b) Si $\lim_{n\to\infty} \psi_2^{(n)}(x) = \psi_2(x)$ *uniformemente en* x, $\mathbf{V}^{(n)}$ *y* $\mathbf{V}$ *denotan las soluciones de (1.1) para* $\psi_2^{(n)}$ *y* ψ_2 *respectivamente, y si* $\left\{\lambda_i(\mathbf{V}^{(n)})\right\}_{i=1}^{m}$ *permanecen en un compacto de* $(\Re^+)^m$ *entonces* $\lim_{n\to\infty} \lambda_i(\mathbf{V}^{(n)}) = \lambda_i(\mathbf{V})$.

c) Si $\psi_2^{(n)}$ *converge uniformemente a* ψ_2, *entonces vale la conclusión del punto b) para el sistema (1.6).*

DEMOSTRACIÓN:

(a) Dado que

$$\frac{\mathbf{z}'}{\|\mathbf{z}\|}\mathbf{V}^{-1}\frac{\mathbf{z}}{\|\mathbf{z}\|} = \sum_{i=1}^{n} \lambda_i^{-1}(\mathbf{V}) \left| t_i' \frac{\mathbf{z}}{\|\mathbf{z}\|}\right|^2$$

resulta que

$$\lambda_1^{-1}(\mathbf{V}) \leq \frac{\mathbf{z}'}{\|\mathbf{z}\|}\mathbf{V}^{-1}\frac{\mathbf{z}}{\|\mathbf{z}\|} \leq \lambda_m^{-1}(\mathbf{V})$$

de donde surge el resultado deseado.

(b) Si $\mathbf{V}^{(0)}$ denota un punto de acumulación de la sucesión de matrices $\mathbf{V}^{(c_n)}$, entonces puede escribirse que

$$\begin{aligned}
\mathbf{V}^{(c_n)} =& E_P \psi_2^{(c_n)}(\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}) \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z}) = \\
& E_P \left[\psi_2^{(c_n)}(\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}) - \psi_2^{(c_0)}(\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z})\right] \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z}) \\
& + E_P \psi_2(\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}) \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z})
\end{aligned}$$

Como $\psi_2^{(c_n)}$ converge uniformemente a ψ_2, entonces por parte (a) y teorema de convergencia dominada resulta que

$$\lim_{n\to\infty} E_P[\psi_2^{(c_n)} - \psi_2] \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'(\mathbf{V}^{(c_n)})^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z}) = 0$$

y en consecuencia

$$\mathbf{V}^{(0)} = E_P \psi_2(\mathbf{z}'(\mathbf{V}^{(0)})^{-1}\mathbf{z}) \frac{\mathbf{z}\mathbf{z}'}{\mathbf{z}'(\mathbf{V}^{(0)})^{-1}\mathbf{z}} I_{\Re^m - \{0\}}(\mathbf{z})$$

Por unicidad de la solución de (1.1) (ver Maronna - 1976) todos los puntos de acumulación coinciden, y $\lim_{n\to\infty} \mathbf{V}^{(n)} = \mathbf{V}$. Por lo tanto vale el resultado para la convergencia de los autovalores (ver Lemma 1 de Tyler - 1981).

c) Reescribiendo el sistema (1.6) en términos del máximo y mínimo autovalor $\lambda_1^{(n)}$ y $\lambda_m^{(n)}$ respectivamente resulta que

$$(1-\epsilon)E_{F_0}\psi_2^{(n)}\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)\frac{u/\lambda_1^{(n)}}{u/\lambda_1^{(n)}+v/\lambda_m^{(m)}}+\epsilon K_2=1$$
$$(1-\epsilon)E_{F_0}\psi_2^{(n)}\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)+\epsilon K_2=m$$

o equivalentemente

$$\begin{aligned}&(1-\epsilon)E_{F_0}(\psi_2^{(n)}-\psi_2)\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)\frac{u/\lambda_1^{(n)}}{u/\lambda_1^{(n)}+v/\lambda_m^{(n)}}+\\&(1-\epsilon)E_{F_0}\psi_2\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)\frac{u/\lambda_1^{(n)}}{u/\lambda_1^{(n)}+v/\lambda_m^{(n)}}+\epsilon K_2=1\\&(1-\epsilon)E_{F_0}(\psi_2^{(n)}-\psi_2)\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)+(1-\epsilon)E_{F_0}\psi_2\left(\frac{u}{\lambda_1^{(n)}}+\frac{v}{\lambda_m^{(n)}}\right)+\epsilon K_2=m\end{aligned}$$

Luego si $\lambda_1^{(n)}$ y $\lambda_m^{(n)}$ tienden a ∞ o $\lambda_m^{(n)}$ tiende a 0 entonces surge una contradicción de la segunda ecuación. Si $\lambda_m^{(n)}$ permanece en un compacto de $\Re^+$ y $\lambda_1^{(n)}$ tiende a ∞ entonces surge una contradicción de la primer ecuación. Luego la sucesión $(\lambda_1^{(n)},\lambda_m^{(n)})$ permanece acotada y en consecuencia la solución de (1.6) $(s^{(n)},b^{(n)})$ permanece acotada. Como consecuencia de b) aparece el resultado deseado. ∎

LEMA 7. *Supóngase que A2)-A5) valen. Si* $\epsilon < \min\{1/K_2,(m-m_0)/(K_2-m_0)\}$ *y* $c\in[0,1]$ *entonces las soluciones* $(\bar{s}(c),\bar{b}(c))$ *de (1.14) son continuas.*

DEMOSTRACIÓN: Por el Lema 6 se debe analizar que las soluciones de (1.14) permanecen acotadas si $c\in[0,1]$.

a) Si $\limsup_{c_n\to c_0}\bar{b}(c_n)=\infty$ entonces de la primer ecuación de (1.5) para $\psi_2^{(c_n)}$ y $K=\infty$ se obtiene que $\epsilon=1/[(1-c_0)K_2+c_0m]\geq 1/K_2$ y ello es absurdo.

b) Si $\liminf_{c_n\to c_0}\bar{b}(c_n)=0$ de la segunda ecuación de (1.14) resulta $0=m-1$.

Por lo tanto $0<\inf_{[0,1]}\bar{b}(c)\leq\sup_{[0,1]}\bar{b}(c)<\infty$

c) Si $\limsup_{c_n \to c_0} \bar{s}(c_n) = \infty$, la primer ecuación de (1.14) dice $\epsilon = (m - m_0)/(K_2 - m_0)$ y ello es absurdo.

d) Si $\liminf_{c_n \to c_0} \bar{s}(c_n) = 0$ entonces m sería igual a K_2 y ello es absurdo.
En consecuencia $(\bar{s}(c), \bar{b}(c))$ permanecen en un compacto de $\Re^+ \times \Re^+ \quad \forall c \in [0,1]$. Por el lema 6, $(\bar{s}(c), \bar{b}(c))$ son continuas en $[0,1]$. ∎

COROLARIO 4. *Bajo A2)-A5), la solución para el sistema (1.6) $b(\epsilon) = \bar{b}(0)$ es mayor que 1.*

DEMOSTRACIÓN: Sea h(b) definida como en el Teorema 1. $h(1) = (1-\epsilon)m(m-1)/m < m-1$ y $h(\bar{b}(1)) = m - 1$. Por lo tanto $\bar{b}(1) > 1$ y si $\bar{b}(0) < 1$ existiría $c^* \in (0,1)$ tal que $\bar{b}(c^*) = 1$. Al mirar el sistema (1.4) para $\psi_2^{(c^*)}$ y $K = \infty$ resultaría que $K_2^{(c^*)} = 0$ y ello es absurdo. Por lo tanto $\bar{b}(0) > 1$ como se quería ver. ∎

COROLARIO 5. *Bajo A2)-A5), $b(K,\epsilon) > 1 \quad \forall \quad K > 0$*

DEMOSTRACIÓN: Sea $\tilde{K} = \inf\{K : b(K,\epsilon) \geq 1\}$; si $\tilde{K} \in (0,\infty)$ entonces $b(\tilde{K}) = 1$ por continuidad de $b(K,\epsilon)$, y el sistema (1.4) resulta igual a:

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{u+v}{s}\right)\frac{u}{u+v} + \epsilon\psi_2\left(\frac{\tilde{K}}{sb}\right) = 1$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{u+v}{s}\right)\frac{v_j}{u+v} = 1$$

En consecuencia se deduce que $\psi_2\left(\frac{\tilde{K}}{s}\right) = 0$ y ello es absurdo pues $\tilde{K} \in (0,\infty)$ con lo cual $\tilde{K} = 0$ y $b(K,\epsilon) > 1 \quad \forall \quad K$ o bien $\tilde{K} = \infty$ y $b(K,\epsilon) < 1 \quad \forall \quad K$. Por el corolario 4, $b(\epsilon) > 1$ y $\lim_{K\to\infty} b(K,\epsilon) = b(\epsilon)$, resultando que sólo la primer conclusión puede valer. ∎

COROLARIO 6. *Bajo A2)-A5) y A7), la función $s(K,\epsilon)$ es creciente en $K > 0$.*

DEMOSTRACIÓN: La segunda ecuación de (1.5) es decreciente en s y por el Teorema 2 es creciente en $b \geq 1$. Por el Corolario 5, $b(K,\epsilon) > 1$ y entonces para satisfacer tal ecuación de (1.5) $s(K,\epsilon)$ sólo puede ser creciente en K. ∎

NOTA 7: Si $m_0 > 0$ entonces las funciones $s(K,\epsilon)$ y $b(K,\epsilon)$ son discontinuas en el 0 ya que si $s^0 = \lim_{K\to 0} s(K,\epsilon)$ y $b^0 = \lim_{K\to 0} b(K,\epsilon)$ entonces se cumple tras tomar límite en la primer ecuación de (1.5) que

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s^0}(u+v)\right) < (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s^0}\left(\frac{u}{b^0} + (b^0)^{1/(m-1)}v\right)\right) + \epsilon m_0 = m$$

lo cual dice que $s^0 > s(\delta_{(0,\ldots,0)})$. Luego, si b^0 fuera 1, la segunda ecuación de (1.5) diría un absurdo.

3.4 Estudio de Sesgo para M-estimadores de Dispersión.

A continuación se calculará el sesgo máximo de un M-estimador de dispersión. Si $c_0\mathbf{I}$ denota la solución bajo el modelo central, entonces sin pérdida de generalidad se supondrá en lo que sigue $c_0 = 1$.

TEOREMA 3. *Si ψ_2 satisface condiciones (A1)-(A6) y F_0 la condición (A7), entonces el sesgo máximo para un M-estimador aparece en las soluciones del sistema (1.6).*

DEMOSTRACIÓN: Se utilizará la misma notación que en el Lema 2 para la descomposición espectral de una matriz definida. Sea G una distribución en $\Re^m$ y $\mathbf{V} = \mathbf{Q_G}\Delta_\mathbf{G}\mathbf{Q}'_\mathbf{G}$ el M-estimador que soluciona el sistema (1.6), o en la notación del Lema 1, $\mathbf{M(V)} = \mathbf{V}$ lo cual dice que en particular se tienen m ecuaciones $\mathbf{e_j}'\mathbf{M(V)}\mathbf{e_j} = \mathbf{e_j}'\mathbf{V}\mathbf{e_j}$, $1 \le j \le m$, las cuales pueden escribirse en forma equivalente como

$$(1-\epsilon)E_{F_0}\psi_2\left(\sum_{i=1}^m \frac{u_{(i)}}{\lambda_i(G)}\right)\frac{u_h\lambda_{(1)}(G)/\lambda_{(h)}(G)}{\sum_{i=1}^m u_{(i)}\lambda_{(1)}/\lambda_{(i)}}+$$
$$\epsilon E_{\tilde{G}}\psi_2\left(\sum_{i=1}^m \frac{\tilde{u}_{(i)}}{\lambda_i(G)}\right)\frac{\tilde{u}_h\lambda_{(1)}(G)/\lambda_{(h)}(G)}{\sum_{i=1}^m \tilde{u}_{(i)}\lambda_{(1)}(G)/\lambda_{(i)}(G)} I_{\Re^m-\{0\}}(\mathbf{z}) = 1 \qquad (1.15)$$

donde $1 \le h \le m$, $\lambda_{(i)}(G)$ denota el i-ésimo autovalor ordenado de la matriz $\mathbf{V}$ y $u_{(i)} = z_{(i)}^2$ es la correspondiente variable asociada con el autovalor ordenado (i). Denominando entonces $s = \lambda_{(1)}$ y $\mathbf{t} = (t_1, \ldots, t_m)$ tal que $t_1 = 1, t_j = \lambda_{(1)}/\lambda_{(j)})$ $(t_j \le t_{j+1})$, puede construirse el siguiente sistema de m ecuaciones

$$g^1(\mathbf{t}, s, G) =$$
$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\sum_{i=1}^m u_{(i)}t_i\right) + \epsilon E_{\tilde{G}}\psi_2\left(\frac{1}{s}\sum_{i=1}^m \tilde{u}_{(i)}t_i\right) I_{\Re^m-\{0\}}(\mathbf{z}) \qquad (1.16)$$
$$g^h(\mathbf{t}, s, G) =$$
$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\sum_{i=1}^m u_{(i)}t_i\right)\frac{u_{(h)}t_h}{\sum_{i=1}^m u_{(i)}t_i} + \epsilon E_{\tilde{G}}\psi_2\left(\frac{1}{s}\sum_{i=1}^m \tilde{u}_{(i)}t_i\right)\frac{\tilde{u}_{(h)}t_h}{\sum_{i=1}^m \tilde{u}_{(i)}t_i} I_{\Re^m-\{0\}}(\mathbf{z})$$

donde $2 \leq h \leq m$, y además

$$g^1(\mathbf{t}(G), s(G), G) = m \quad \text{y} \quad g^h(\mathbf{t}(G), s(G), G) = 1 \quad \forall \quad 2 \leq h \leq m$$

En caso de trabajar con contaminación $\delta_{(\infty,0,\ldots,0)}$ (que por razones de simplicidad se denotará δ_∞), el sistema (1.16) se define como

$$\begin{aligned}
g^1(\mathbf{t}, s, \delta_\infty) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right)t_m\right)\right) + \epsilon K_2 \\
g^h(\mathbf{t}, s, \delta_\infty) &= (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right)t_m\right)\right)\frac{u_{(h)}t_h}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right)t_m}
\end{aligned} \tag{1.17}$$

El sistema (1.17) es una reparametrización del sistema (1.6) y por lo tanto valen las mismas propiedades de existencia y unicidad de soluciones. Consíderese ahora un vector

$$\mathbf{t}^h_{s,G} = (1, t^h_{2,s,G}, \ldots, t^h_{m,s,G})$$

tal que $t^h_{j,s,G} \leq t^h_{j+1,s,G}$ $\forall$ h y que satisface

$$\begin{cases} g^1(\mathbf{t}^1_{s,G}, s, G) = m & \text{si } h = 1 \\ g^h(\mathbf{t}^h_{s,G}, s, G) = 1 & \text{si } 2 \leq h \leq m \end{cases}$$

Tal vector existe al menos para algún s, por ejemplo si $s = \lambda_{(1)}(G)$ entonces la elección particular de dicho vector recaerá en $(1, \lambda_{(1)}(G)/\lambda_{(2)}(G), \ldots, \lambda_{(1)}(G)/\lambda_{(m)}(G)$. Para el sistema (1.17) la definición de este vector toma la forma particular

$$\begin{aligned}
\mathbf{t}^1_{s,\delta_\infty} &= (1, t^1_{s,\delta_\infty}, \ldots, t^1_{s,\delta_\infty}) \\
\mathbf{t}^h_{s,\delta_\infty} &= (1, t^m_{s,\delta_\infty}, \ldots, t^m_{s,\delta_\infty}) \quad \forall \quad h > 1
\end{aligned}$$

Se mostrará a continuación que $\mathbf{t}^1_{s,\delta_\infty}$ y $\mathbf{t}^m_{s,\delta_\infty}$ existen para todo s a partir de un cierto momento. Sean $1 = (1, \ldots, 1)$. Como

$$\lim_{s\to\infty} g^m(1, s, \delta_\infty) = (1-\epsilon)m_0/m < 1 \quad \text{y} \quad \lim_{s\to 0} g^m(1, s, \delta_\infty) = (1-\epsilon)K_2/m > 1$$

por continuidad existe $s_1 < 1$ tal que $g^m(1, s_1, \delta_\infty) = 1$. Si ahora se toma el vector $\mathbf{f} = (1, t, \ldots, t)$, $t \geq 1$, entonces se cumple que $\lim_{t\to\infty} g^m(\mathbf{t}, s, \delta_\infty) = (1-\epsilon)K_2/(m-1) > 1$

y $g^m(1,s,\delta_\infty) < 1$ si $s > s_1$, con lo cual por continuidad existe $\mathbf{t}^m_{s,\delta_\infty}$ para todo $s > s_1$. Asimismo, para todo s se tiene que $\lim\limits_{t\to\infty} g^1(\mathbf{t},s,\delta_\infty) = (1-\epsilon)K_2 + \epsilon K_2 > m$. Por otro lado, se supuso que bajo el modelo con distribución esféricamente simétrica la solución es la matriz $\mathbf{I}$ con lo cual $E_{F_0}\psi_2(\mathbf{z}'\mathbf{z}) = m$ y $g^1(\mathbf{t},1,\delta_\infty) \geq g^1(1,1,\delta_\infty) = (1-\epsilon)m + \epsilon K_2 > m$. Además $\lim\limits_{s\to\infty} g^1(1,s,\delta_\infty) = (1-\epsilon)m_0 + \epsilon K_2 < m$, con lo cual puede tomarse $s_0 > 1$ tal que $g^1(1,s_0,\delta_\infty) = m$ y en consecuencia existe $\mathbf{t}^1_{s,\delta_\infty}$ $\forall\ s > s_0$. La condición (A4) para ψ_2 y la condición (A7) para F_0 permiten afirmar que $g^1(\mathbf{f},s,\delta_\infty)$ y $g^m(\mathbf{f},s,\delta_\infty)$ son estrictamente crecientes en t para s fijo y por lo tanto existe un único valor t^1_{s,δ_∞} y t^m_{s,δ_∞} para cada s satisfaciendo que $g^1(\mathbf{t}^1_{s,\delta_\infty},s,\delta_\infty) = m$ y $g^m(\mathbf{t}^m_{s,\delta_\infty},s,\delta_\infty) = 1$. Luego las funciones t^1_{s,δ_∞} y t^m_{s,δ_∞} son estrictamente crecientes y continuas.

Dado el vector $\mathbf{t}^1_{s,G}$ constrúyase el vector $\mathbf{f}^1_{s,G} = (1, t^1_{m,s,G}, \ldots, t^1_{m,s,G})$. Luego, para todo $s > s_0$ puede concluirse que

$$m = g^1(\mathbf{t}^1_{s,G},s,G) \leq (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\sum_{i=1}^m u_{(i)}t^1_{i,s,G}\right) + \epsilon K_2 \leq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\left(\sum_{i=2}^m u_{(i)}\right)t^1_{m,s,G} + u_{(1)}\right)\right) + \epsilon K_2 = g^1(\mathbf{f}^1_{s,G},s,\delta_\infty)$$

Usando la monotonía de ψ_2 puede inferirse que

$$t^1_{m,s,G} \geq t^1_{s,\delta_\infty} \quad \forall \quad s > s_0 \tag{1.18}$$

Por otro lado defínase $\mathbf{f}^m_{s,G} = (1, t^m_{m,s,G}, \ldots, t^m_{m,s,G})$. Por lo tanto

$$1 = g^m(\mathbf{t}^m_{s,G},s,G) \geq (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\sum_{i=1}^m u_{(i)}t^m_{i,s,G}\right)\frac{u_{(m)}t^m_{m,s,G}}{\sum\limits_{i=1}^m u_{(i)}t^m_{i,s,G}} \geq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\left(\sum_{i=1}^m u_{(i)}\right)t^m_{m,s,G} + u_{(1)}\right)\right)\frac{u_{(m)}t^m_{m,s,G}}{u_{(1)} + \left(\sum\limits_{i=2}^m u_{(i)}\right)t^m_{m,s,G}} =$$

$$g^m(\mathbf{f}^m_{s,G},s,\delta_\infty)$$

donde la última desigualdad vale por la monotonía de $u_2(s) = \psi_2(s)/s$ (propiedad (A1)). Como $g^m(\mathbf{f}^m_{s,G},s,\delta_\infty) \geq g^m(1,s,\delta_\infty) > 1$ si $s < s_1$ entonces de las acotaciones obtenidas $t^m_{m,s,G}$ sólo puede existir para $s > s_1$, resultando entonces que

$$t^m_{m,s,G} \leq t^m_{s,\delta_\infty} \quad \forall \quad s > s_1 \tag{1.19}$$

Se verá a continuación que

$$\lim_{s\to\infty} \|\mathbf{t}^1_{s,\delta_\infty}\| = \infty \tag{1.20}$$

Supóngase por el contrario que existe una sucesión $\{s_n\}$ satisfaciendo que $\lim_{n\to\infty} s_n = \infty$ y que $\sup_n \|\mathbf{t}^1_{s_n,\delta_\infty}\| < \infty$, entonces $\lim_{n\to\infty} g^1(\mathbf{t}^1_{s_n,\delta_\infty}, s_n, \delta_\infty) = (1-\epsilon)m_0 + \epsilon K_2$ pero ello diría que $\epsilon > (m - m_0)/(K_2 - m_0)$ lo cual es absurdo.

Si $(s(\epsilon), b(\epsilon)) = (s^*, b^*)$ denota la solución del sistema (1.6) entonces se chequeará que

$$\begin{cases} t^1_{s,\delta_\infty} \geq t^m_{s,\delta_\infty} & \text{si } s > s^*b^* \\ t^1_{s,\delta_\infty} \leq t^m_{s,\delta_\infty} & \text{si } s < s^*b^* \end{cases} \tag{1.21}$$

Supóngase que la primera afirmación de (1.21) no fuera cierto. Entonces existe una sucesión $\{s_n\}$ tal que $\lim_{n\to\infty} s_n = \infty$ y $t^1_{s_n,\delta_\infty} \leq t^m_{s_n,\delta_\infty}$, y se obtendría que

$$m = (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty} + u_{(1)}\right)\right) + \epsilon K_2 =$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty} + u_{(1)}\right) \frac{u_{(1)}}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty}} +$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty} + u_{(1)}\right)\right) \frac{\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty}}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty}} + \epsilon K_2 \leq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty} + u_{(1)}\right)\right) \frac{u_{(1)}}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta\infty}} +$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\left(\sum_{i=2}^m u_{(i)}\right) t^m_{s_n,\delta_\infty} + u_{(1)}\right)\right) \frac{\left(\sum_{i=2}^m u_{(i)}\right) t^m_{s_n,\delta\infty}}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right) t^m_{s_n,\delta_\infty}} + \epsilon K_2 =$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_n}\left(\left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta_\infty} + u_{(1)}\right)\right) \frac{u_{(1)}}{u_{(1)} + \left(\sum_{i=2}^m u_{(i)}\right) t^1_{s_n,\delta\infty}} + m - 1 + \epsilon K_2$$

Por (1.20) el último miembro tiende a $m - 1 + \epsilon K_2$ con lo cual se obtendría que $\epsilon > 1/K_2$. Por unicidad de soluciones del sistema (1.17) las curvas t^1_{s,δ_∞} y t^m_{s,δ_∞} sólo pueden intersectarse una vez, con lo cual debe ocurrir la primer ecuación de (1.21). Para llegar a la conclusión sobre la segunda ecuación de (1.21) debe notarse que si $s^*b^* < \lambda_{(1)}(G)$ entonces (1.19) y (1.18) permiten afirmar que

$$t^m_{\lambda_{(1)}(G),\delta_\infty} \geq t^m_{m,\lambda_{(1)}(G),G} = t_m(G) = t^1_{m,\lambda_{(1)}(G),G} \geq t^1_{\lambda_{(1)}(G),\delta_\infty} > t^1_{s^*b^*,\delta_\infty}$$

contradiciendo así la primer ecuación de (1.21), por lo cual se tiene que

$$\lambda_{(1)}(G) \leq s^*b^* \tag{1.22}$$

Entonces de no ser válida la segunda ecuación de (1.21) se cumpliría que $t^1_{s,\delta_\infty} > t^m_{s,\delta_\infty}$ para todo $s < s * b*$, en especial para $\lambda_{(1)}(G)$,pero entonces por (1.18) y (1.19) resultaría que

$$t_m(G) = t^1_{m,\lambda_{(1)}(G),G} \geq t^1_{\lambda_{(1)}(G),\delta_\infty} > t^m_{\lambda_{(1)}(G),\delta_\infty} \geq t^m_{m,\lambda_{(1)}(G),G} = t_m(G)$$

se intersectan $t^1_{m,s,G}$ y $t^m_{m,s,G}$ lo cual es imposible. Luego por (1.19), (1.21) y la monotonía de t^m_{s,δ_∞} se obtiene que

$$(b^*)^{m/(m-1)} = t^m_{s^*b^*,\delta_\infty} > t^m_{s,\delta_\infty} \geq t^m_{m,s,G} \quad \forall \quad s < s^*b^* \tag{1.23}$$

Si $(b^*)^{m/(m-1)} < t_m(G) = t^m_{m,\lambda_{(1)}(G),G}$ entonces la monotonía de t^m_{s,δ_∞} permite afirmar que $t^m_{\lambda_{(1)}(G),\delta_\infty} \leq t^m_{s^*b^*,\delta_\infty} = (b^*)^{m/(m-1)} < t^m_m(G) = t^m_{m,\lambda_{(1)}(G),G}$ lo cual contradice (1.19). Luego, por (1.23) $t^m_{m,\lambda_{(1)}(G),G} = t_m(G) \leq b^*$ que es lo que se deseaba ver, y en consecuencia

$$t_m(G) = \frac{\lambda_{(1)}}{\lambda_{(m)}}(G) \leq (b^*)^{m/(m-1)}$$

O sea que el máximo sesgo se alcanza cuando la contaminación puntual sobre un eje tiende a infinito. ∎

NOTA 8: Obviamente el teorema anterior se aplica a estimadores resultantes de funciones de la forma $\tilde{\psi}_2(x) = m\psi_2(cx)(E_{F_0}\psi_2(cz'z))^{-1}$ donde ψ_2 satisface las condiciones ya establecidas para existencia y unicidad y c es alguna constante mayor que 0. Sean $(s_c(\psi_2), b_c(\psi_2))$ las correspondientes soluciones para el sistema (1.6).Si se consideran ahora funciones de la forma $\tilde{\psi}_2(x) = m\psi_2(x)(E_{F_0}\psi_2(cz'z))^{-1}$. la solución de (1.6) para $\tilde{\psi}_2$ es de la forma $(c^{-1}s_c(\psi_2), b_c(\psi_2))$, por lo tanto ambas poseen el mismo sesgo máximo en el infinito.

NOTA 9: En caso de trabajar con $m = 2$ la demostración del teorema previo podría haber sido realizada siguiendo exactamente el mismo lineamiento pero parametrizando en términos de $s = |\mathbf{V}|^{1/2}$ en vez de como se hizo en términos del máximo autovalor: en tal caso la propiedad de que $u_2(s)$ es no creciente sólo se usa para asegurar la existencia de soluciones para cualquier contaminación G y no explícitamente como se utiliza para conseguir (1.19). Asimismo, para obtener una desigualdad del tipo (1.21) basta con usar el Teorema 2, y las acotaciones empleadas no son necesarias. Entonces allí puede afirmarse que

$$s(0) \leq s(G) \leq s(\infty) \tag{1.24}$$

donde las cotas superior e inferior son obtenidas contaminando con $\delta_{(\infty,0)}$ y $\delta_{0,0}$ respectivamente. Ambas desigualdades surgen con un razonamiento absolutamente análogo al realizado en la demostración anterior. La desigualdad $s(G) \leq s(\infty)$ no pudo generalizarse para $m > 2$ pero el siguiente lema muestra que la primer desigualdad de (1.24) vale para cualquier dimensión.

LEMA 8. *Si valen A1)-A6), entonces $s(0) \leq s(G)$ para todo $m \geq 2$ ($s = |\mathbf{V}|^{1/m}$).*

DEMOSTRACIÓN: Sea $\delta_{(0,\ldots,0)}$ la masa puntual en $0 \in \Re^m$. Con la misma notación del Teorema 3 se verá que $\lambda_{(i)}(G) \geq \lambda_{(i)}(\delta_{(0,\ldots,0)})$ para todo $1 \leq i \leq m$ y cualquier contaminación G. Tal como se hizo en el Teorema 3 para construir el sistema (1.16) tómese la parametrización $s = \lambda_{(1)}$. El equivalente al sistema (1.17) para la contaminación $\delta_{(0,\ldots,0)}$ es la ecuación

$$g_1^0(s, \delta_{(0,\ldots,0)}) = (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\left(\sum_{j=2}^{m} u_{(j)}\right) + u_{(1)}\right)\right)\frac{u_{(m)}}{u_{(1)} + \left(\sum_{j=2}^{m} u_{(j)}\right)}$$

Sean $s_G = \lambda_{(1)}(G)$ y s_0 la solución de $g_1^0(s_0, \delta_{(0,\ldots,0)}) = 1$ entonces en caso de $s_G < s_0$

puede escribirse que

$$1 = g^m(\mathbf{t}^m_{s_G,G}, s_G, G) \geq (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G}\sum_{j=1}^m u_{(j)}t^m_{j,s_G,G}\right)\frac{u_{(m)}t^m_{m,s_G,G}}{\sum_{j=1}^m u_{(j)}t^m_{m,s_G,G}} \geq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G}\sum_{j=1}^m u_{(j)}\right)\frac{u_{(m)}}{\sum_{j=1}^m u_{(j)}} >$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_0}\sum_{j=1}^m u_{(j)}\right)\frac{u_{(m)}}{\sum_{j=1}^m u_{(j)}} = g_1^0(s_0, \delta_{(0,\ldots,0)}) = 1$$

lo cual es claramente absurdo. Tómese ahora la parametrización $s = \lambda_{(i)}$ para un $i > 1$ fijo y $\mathbf{t} = (t_1, \ldots, t_m)$ tal que $t_i = 1$ y $t_j = \lambda_{(i)}/\lambda_{(j)}$. $t_j \leq 1$ si $j \leq i$ y $t_j \geq 1$ si $j \geq i$. Formalmente se obtiene un sistema idéntico al (1.16). Para la contaminación $\delta_{(0,\ldots,0)}$ puede asumirse ahora que la única ecuación a resolver toma la forma

$$g_2^0(t^0, s, \delta_{(0,\ldots,0)}) = (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\left(\sum_{j=2}^m u_{(j)}\right) + u_{(1)}t^0\right)\right)\frac{u_{(m)}}{u_{(1)}t^0 + \left(\sum_{j=2}^m u_{(j)}\right)}$$

Sea como antes un vector $\mathbf{t}^m_{s,G} = (t^m_{1,s,G}, \ldots, t^m_{i-1,s,G}, 1, t^m_{i+1,s,G}, \ldots, t^m_{m,s,G})$ tal que $g^m(\mathbf{t}^m_{s,G}, s, G) = 1$. Además $t^m_{j,s,G} \leq 1$ si $j \leq i$ y $t^m_{j,s,G} \geq 1$ si $j \geq i$. Si $s_G = \lambda_{(i)}(G)$ entonces $t^m_{j,s_G^{(i)},G} = \lambda_{(i)}(G)/\lambda_{(j)}(G)$. En caso que $\lambda_{(i)}(G) < \lambda_{(i)}(\delta_{(0,\ldots,0)})$, y si $s_0 = \lambda_{(i)}(\delta_{(0,\ldots,0)})$ (este valor es idéntico para todo i) puede concluirse que

$$1 = g^m(\mathbf{t}^m_{s_G^{(i)},G}, s_G^{(i)}, G) \geq (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G^{(i)}}\sum_{j=1}^m u_{(j)}t^m_{j,s_G^{(i)},G}\right)\frac{u_{(m)}t^m_{m,s_G^{(i)},G}}{\sum_{j=1}^m u_{(j)}t^m_{m,s_G^{(i)},G}} \geq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G^{(i)}}\left(\left(\sum_{j=2}^m u_{(j)}\right)t^m_{m,s_G^{(i)},G} + u_{(1)}t^m_{1,s_G^{(i)},G}\right)\right)\times$$

$$\frac{u_{(m)}t^m_{m,s_G^{(i)},G}}{u_{(1)}t^m_{1,s_G^{(i)},G} + \left(\sum_{j=2}^m u_{(j)}\right)t^m_{m,s_G^{(i)},G}} \geq$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G^{(i)}}\left(\left(\sum_{j=2}^m u_{(j)}\right) + u_{(1)}t^m_{1,s_G^{(i)},G}\right)\right)\frac{u_{(m)}}{u_{(1)}t_1^m + \left(\sum_{j=2}^m u_{(j)}\right)} >$$

$$(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_0}\left(\left(\sum_{j=2}^m u_{(j)}\right) + u_{(1)}t^m_{1,s_G^{(i)},G}\right)\right)\frac{u_{(m)}}{u_{(1)}t_1^m + \left(\sum_{j=2}^m u_{(j)}\right)} =$$

$$g^0(t^m_{1,s_G^{(i)},G}, s_0, \delta_{(0,\ldots,0)})$$

Luego se obtiene que $1 \geq t^{m}_{1,s^{(i)}_{G},G} > t^{m}_{1,s_0,\delta_{0,\dots,0}} = 1$ y ello es obviamente absurdo. En consecuencia $\lambda_{(i)}(0) \leq \lambda_{(i)}(G)$ para todo $1 \leq i \leq m$ y luego se cumple que $s(\delta_{(0,\dots,0)}) \leq s(G)$ como se deseaba ver. ∎

NOTA 10: Para ver que el determinante de la matriz de covarianza se minimiza en la contaminación $\delta_{(0,\dots,0)}$, se ha visto que cada autovalor se minimiza en dicha contaminación. De (1.22) se sabe que $\lambda_{(1)}(G) \leq \lambda_{(1)}(\delta_\infty)$, donde $\lambda_{(1)}(\delta_\infty)$ denota el máximo autovalor del M-estimador cuando la contaminación es $\delta_{(\infty,0,\dots,0)}$. Ello no se generaliza para cualquier autovalor, como se puede observar estudiando $\lambda_{(m)}(G)$, el cual se maximiza cuando se toma la contaminación $\delta_{(\infty,\infty,\dots,\infty)}$. Para mostrar eso se procederá en la forma habitual. Denotando por $s_\infty = \lambda_{(m)}(\overline{\infty})$ al mínimo autovalor asociado a esa contaminación. Parametrizando respecto de $s = \lambda_{(m)}$ el equivalente al sistema (1.17) para la contaminación en cuestión es

$$\overline{g}(t,s) = (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\left(\sum_{j=2}^{m} u_{(j)}\right) + u_{(1)}t\right)\right) + \epsilon K_2$$

Sea como antes un vector $\mathbf{t}^1_{s,G} = (t^1_{1,s,G},\dots,t^1_{m-1,s,G},1)$ tal que $g^1(\mathbf{t}^1_{s,G},s,G) = m$ y $t^1_{j,s,G} \leq 1$ si $j \leq m-1$. Si $s_G = \lambda_{(m)}(G)$ entonces $t^1_{j,s_G,G} = \lambda_{(m)}(G)/\lambda_{(j)}(G)$. Si $s_G > s_\infty$ entonces puede escribirse que

$$\begin{aligned}
m = g^1(\mathbf{t}^1_{s,G},s,G) &\leq (1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G}\sum_{j=1}^{m} u_{(j)}t^1_{j,s_G,G}\right) + \epsilon K_2 \leq \\
&(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_G}\left(\left(\sum_{j=2}^{m} u_{(j)}\right) + u_{(1)}t^1_{1,s_G,G}\right)\right) + \epsilon K_2 \leq \\
&(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_\infty}\left(\left(\sum_{j=2}^{m} u_{(j)}\right) + u_{(1)}t^1_{1,s_G,G}\right)\right) + \epsilon K_2 = \overline{g}(t^1_{1,s_G,G},s_\infty)
\end{aligned}$$

Por lo tanto se obtiene que

$$1 \geq t^1_{1,s_G,G} > t^1_{1,s_\infty,\delta_{(\infty,\dots,\infty)}} = 1$$

y en consecuencia $\lambda_{(m)}(G) \leq \lambda_m(\overline{\infty})$ que era lo que se deseaba ver.

3.5 Optimalidad del estimador de Tyler.

Por el Teorema 3 se tiene que $b^{m/(m-1)}(\epsilon) = \sup_{G\in\mathcal{F}_\epsilon} \lambda_{(1)}(G)/\lambda_{(m)}(G)$ que resulta el máximo sesgo en el entorno de contaminación. Por lo tanto si $\epsilon_1 < \epsilon_2 < \epsilon^*(\mathbf{V}, F_0)$ entonces $\mathcal{F}_{\epsilon_1} \subseteq \mathcal{F}_{\epsilon_2}$ con lo cual

$$b^{m/(m-1)}(\epsilon_1) = \sup_{G\in\mathcal{F}_{\epsilon_1}} \lambda_{(1)}(G)/\lambda_{(m)}(G) \leq \sup_{G\in\mathcal{F}_{\epsilon_2}} \lambda_{(1)}(G)/\lambda_{(m)}(G) = b^{m/(m-1)}(\epsilon_2)$$

Sin embargo se verá que la monotonía de $b(\epsilon)$ es válida sobre el conjunto más grande dado en el enunciado del lema siguiente.

LEMA 9. *La función* $b(\epsilon) : [0, \min\{1/K_2, (m-m_0)/(K_2-m_0)\}) \to [1,\infty)$ *es monótona creciente.*

DEMOSTRACIÓN: Supóngase primeramente que ψ_2 es diferenciablemente continua con derivada acotada estrictamente positiva en $\Re$. El Lema 5 a) dice como construir una función ψ_2 satisfaciendo tales características. Entonces en el sistema (1.5) pueden obtenerse las derivadas de las funciones involucradas derivando bajo el signo integral, donde las derivadas parciales respecto de s, b y K existen y son continuas:

$$g_1^s(s,b,K,\epsilon) =$$
$$-\frac{1-\epsilon}{s^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(\frac{u}{b}+b^{1/(m-1)}v\right)dF_0 - \epsilon\psi_2'(\frac{K}{sb})\frac{K}{s^2b} < 0$$

$$g_2^s(s,b,\epsilon) =$$
$$-\frac{1-\epsilon}{s^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)b^{1/(m-1)}v\,dF_0 < 0$$

$$g_1^b(s,b,K,\epsilon) =$$
$$\frac{1-\epsilon}{(m-1)sb^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(b^{m/(m-1)}v-(m-1)u\right)dF_0 - \epsilon\psi_2'(\frac{K}{sb})\frac{K}{sb^2}$$

$$g_2^b(s,b,\epsilon) =$$
$$\frac{1-\epsilon}{(m-1)sb^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(b^{m/(m-1)}v-(m-1)u\frac{b^{m/(m-1)}v}{u+b^{m/(m-1)}v}\right)dF_0$$
$$+(1-\epsilon)\frac{m}{m-1}\int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(\frac{uvb^{1/(m-1)}}{(u+b^{m/(m-1)}v)^2}\right)dF_0 \geq 0$$

$$g_1^\epsilon(s,b,K,\epsilon) = \psi_2(\frac{K}{sb}) - \int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)dF_0$$

$$g_2^\epsilon(s,b,\epsilon) = -\int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\frac{b^{m/(m-1)}v}{u+b^{m/(m-1)}v}dF_0 < 0$$

$$g_1^K(s,b,K,\epsilon) = \frac{\epsilon}{sb}\psi_2(\frac{K}{sb}) > 0$$

Por el Lema 3, Lema 4 y Corolario 6 las soluciones de (1.5) $(s(K,\epsilon), b(K,\epsilon))$ son continuas y crecientes en K para ϵ fijo. Luego ambas son diferenciables salvo un conjunto de medida cero. Sea $\mathbf{C}_0(\epsilon)$ el conjunto donde ambas son diferenciables, y por lo tanto vale el siguiente conjunto de ecuaciones

$$0 = g_1^s \frac{d}{dK} s(K,\epsilon) + g_1^b \frac{d}{dK} b(K,\epsilon) + g_1^K$$
$$0 = g_2^s \frac{d}{dK} s(K,\epsilon) + g_2^b \frac{d}{dK} b(K,\epsilon)$$

En caso que el determinante de dicho sistema se anule al valuar en las soluciones, i.e. $(g_1^s g_2^b - g_1^b g_2^s)(s(K,\epsilon), b(K,\epsilon), K, \epsilon) = 0$ resultaría que

$$g_1^b(s(K,\epsilon), b(K,\epsilon), K, \epsilon) = \lambda(K) g_1^s(s(K,\epsilon), b(K,\epsilon), K, \epsilon)$$
$$g_2^b(s(K,\epsilon), b(K,\epsilon), K, \epsilon) = \lambda(K) g_2^s(s(K,\epsilon), b(K,\epsilon), K, \epsilon)$$

y reemplazando en el sistema anterior surge que

$$0 = g_1^s \frac{d}{dK} s(K,\epsilon) + \lambda(K) g_1^s \frac{d}{dK} b(K,\epsilon) + g_1^K$$
$$0 = g_2^s \frac{d}{dK} s(K,\epsilon) + \lambda(K) g_2^s \frac{d}{dK} b(K,\epsilon)$$

Como $\psi_2' > 0$ por suposición, entonces $(g_1^s g_2^b - g_1^b g_2^s)(s(K,\epsilon), b(K,\epsilon), K, \epsilon)$ no se anula si $K \in \mathbf{C}_o(\epsilon)$. Luego para todo $K_0 \in \mathbf{C}_0(\epsilon)$ existe un entorno $V^\epsilon(K_0)$ tal que $(g_1^s g_2^b - g_1^b g_2^s)(s(K,\epsilon), b(K,\epsilon), K, \epsilon) \neq 0$ si $K \in V^\epsilon$ por continuidad de las derivadas. Consecuentemente si K pertenece al abierto denso de $(0,\infty)$ $\mathbf{C}(\epsilon) = \bigcup_{K_0 \in \mathbf{C}_0(\epsilon)} V^\epsilon(K_0)$ se tiene que $(g_1^s g_2^b - g_1^b g_2^s)(s(K,\epsilon), b(K,\epsilon), K, \epsilon)$ no se anula en dicho conjunto. Como un abierto de $\Re$ es unión numerable de intervalos abiertos y el conjunto que se tiene es denso, entonces es unión de intervalos abiertos adyacentes, i.e. $\mathbf{C}(\epsilon) = (0,\infty) - \{a_i\}_{i=1}^\infty$ con $\{a_i\}_{i=1}^\infty$ una sucesión creciente. Sea ahora $D = \{\epsilon_j\}_{j=1}^\infty$ un conjunto denso numerable de $[0, \min\{1/K_2, (m - m_0)/(K_2 - m_0)\}]$. Luego $U = \bigcap_{j=1}^\infty \mathbf{C}(\epsilon_j)$ es denso en $(0,\infty)$ y puede afirmarse que $(g_1^s g_2^b - g_1^b g_2^s)(s(K,\epsilon), b(K,\epsilon), K, \epsilon) \neq 0$ si $K \in U$ y $\epsilon \in D$. En consecuencia el teorema de la función implícita permite afirmar que si $K \in U$ y $\epsilon \in D$

$$\frac{d}{dK} b(K,\epsilon) = -\frac{-g_1^K g_2^s}{g_1^s g_2^b - g_2^s g_1^b} \qquad \frac{d}{d\epsilon} b(K,\epsilon) = \frac{g_1^\epsilon g_2^s - g_2^\epsilon g_1^s}{g_1^s g_2^b - g_2^s g_1^b}$$

Para ϵ fijo, $b(K,\epsilon)$ es creciente en K y por lo tanto, $\frac{d}{dK}b(K,\epsilon) \geq 0$. Sea $K_0 > 0$ e $I_\delta = (0, \min\{1/K_2, (m-m_0)/(K_2-m_0)\} - \delta]$ tal que si $K \in J = [K_0, \infty)$ y $\epsilon \in I_\delta$ entonces por el Corolario 1 y todos sus incisos, $\text{Im}(s(K,\epsilon), b(K,\epsilon)) \subseteq \mathbf{F}$ donde $\mathbf{F}$ es un conjunto compacto de $\Re^+ \times \Re^+$. Entonces existe $\delta_1 > 0$ tal que $E_{F_0}\psi_2 < K_2 - \delta_1$ si $(K,\epsilon) \in (U \cap J) \times (D \cap I_\delta)$. Puede elegirse $K_0 = K_0(\delta, \delta_1)$ tal que

$$g_1^\epsilon(s(K,\epsilon), b(K,\epsilon), K, \epsilon) = \psi_2(K/(s(K,\epsilon)b(K,\epsilon))) - E_{F_0}\psi_2 > K_2 - \delta_1/2 - K_2 + \delta_1 = \delta_1/2$$

si $(K,\epsilon) \in (U \cap [K_0,\infty)) \times (D \cap I_\delta)$. Luego $g_1^\epsilon g_2^s < 0$, $-g_1^s g_2^\epsilon < 0$ y $g_1^s g_2^b - g_2^s g_1^b < 0$ y por lo tanto $\frac{d}{d\epsilon}b(K,\epsilon) > 0$ si $(K,\epsilon) \in (U \cap [K_0,\infty)) \times (D \cap I_\delta)$. Por el teorema de la función implícita y las condiciones pedidas sobre ψ_2 esta derivada es continua y en consecuencia $\frac{d}{dK}b(K,\epsilon)$ es positiva sobre un abierto denso de $D \cap I_\delta$ el cual como antes debe ser unión de intervalos adyacentes,luego es creciente sobre cada intervalo. Por el Lema 3, $b(K,\epsilon)$ es función continua de (K,ϵ) con lo cual resulta que $b(K,\epsilon)$ es creciente en $K \in (0,\infty)$ si ϵ está fijo en D y $b(K,\epsilon)$ es creciente en ϵ si $\epsilon \in I_\delta$ y K está fijo en U. Luego $b(\epsilon) = \lim_{K\to\infty} b(K,\epsilon)$ si $K \in U \cap [K_0,\infty)$ (los K son comunes a todo ϵ). Como el límite puntual de funciones crecientes es creciente entonces $b(\epsilon)$ es creciente en I_δ. Como el análisis se repite para todo $\delta > 0$, resulta que $b(\epsilon)$ es creciente en $[0, \min\{1/K_2, (m-m_0)/(K_2-m_0)\}]$. Para una ψ_2 general el resultado surge aplicando el Lema 6 c). ∎

Se ha encontrado entonces el máximo sesgo y el mismo ocurre con una contaminación en el infinito. Ello sugeriría elegir para hacer mínimo a tal sesgo un estimador resultante de una ψ_2 que penalice lo más posible a dicha contaminación. Luego el candidato sería $\psi_2(x) = m$ pues toda otra ψ_2 satisface que su supremo es estrictamente mayor que m. Luego el candidato será el estimador de Tyler. Asimismo, Tyler (1987) señala que el estimador propuesto posee la propiedad "minimax siguiente: minimiza la varianza asintótica entre todos los estimadores consistentes y asintóticamente normales en la clase de las distribuciones elípticas.

TEOREMA 4. *Si valen A2)-A7), entonces el estimador de Tyler* $\mathbf{b_T}(\epsilon)$ *(solución de (1.5) para* $\psi_2 \equiv m$*) minimiza el máximo sesgo para M-estimadores de forma.*

DEMOSTRACIÓN: Supóngase primeramente que ψ_2 satisface condiciones (A3) y (A4), $\psi_2(x) = K_2$ para $x \geq K_0$ y $\psi_2(x) < K_2$ si $x < K_2$. Además ψ_2 se supondrá diferenciablemente continua con derivada acotada estrictamente positiva en $(0, K_0)$. El Lema 5 permite construir adecuadamente una función con tales características. Entonces en el sistema (1.6) pueden obtenerse las derivadas de las funciones involucradas derivando bajo el signo integral, donde las derivadas parciales respecto de s, b y ϵ existen y son continuas:

$$g_1^s(s,b,\epsilon) = -\frac{1-\epsilon}{s^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(\frac{u}{b}+b^{1/(m-1)}v\right)dF_0 < 0$$

$$g_2^s(s,b,\epsilon) = -\frac{1-\epsilon}{s^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)b^{1/(m-1)}v\,dF_0 < 0$$

$$g_1^b(s,b,\epsilon) = \frac{1-\epsilon}{(m-1)sb^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(b^{m/(m-1)}v-(m-1)u\right)dF_0 \geq 0$$

$$g_2^b(s,b,\epsilon) =$$

$$\frac{1-\epsilon}{(m-1)sb^2}\int \psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\left(b^{m/(m-1)}v-(m-1)u\frac{b^{m/(m-1)}v}{u+b^{m/(m-1)}v}\right)dF_0 +$$

$$(1-\epsilon)\frac{m}{m-1}\int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\frac{uvb^{1/(m-1)}}{(u+b^{m/(m-1)}v)^2}dF_0 \geq 0$$

$$g_1^\epsilon(s,b,\epsilon) = K_2 - \int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)dF_0 > 0$$

$$g_2^\epsilon(s,b,\epsilon) = -\int \psi_2\left(\frac{1}{s}\left(\frac{u}{b}+b^{1/(m-1)}v\right)\right)\frac{b^{m/(m-1)}v}{u+b^{m/(m-1)}v} < 0$$

Como ocurre en el Lema 9 las desigualdades estrictas colocadas para g_1^s y g_2^s son válidas. Sea $s_2(b,\epsilon)$ tal que $g_2(s_2(b,\epsilon),b,\epsilon) = m - 1$. Por el teorema de la función implícita resulta que

$$\frac{\partial}{\partial b}s_2(b,\epsilon) = -g_2^b/g_2^s \text{ y } \frac{\partial}{\partial \epsilon}s_2(b,\epsilon) = -g_2^\epsilon/g_2^s$$

Si $(s(\epsilon), b(\epsilon))$ denotan las soluciones de (1.6), entonces el Lema 9 dice que $b(\epsilon)$ es monótona creciente en $\mathbf{E} = (0, \min\{1/K_2, (m-m_0)/(K_2-m_0)\})$ y por lo tanto es diferenciable salvo un conjunto de medida de Lebesgue cero en dicho conjunto; sea $\mathbf{E}_0$ dicho conjunto. Como $s(\epsilon) = s_2(b(\epsilon),\epsilon)$ resulta que $s(\epsilon)$ es diferenciable ppϵ y en consecuencia puede escribirse para $\epsilon \in \mathbf{E}_0$ el siguiente sistema de ecuaciones

$$0 = g_1^s s'(\epsilon) + g_1^b b'(\epsilon) + g_1^\epsilon$$

$$0 = g_2^s s'(\epsilon) + g_2^b b'(\epsilon) + g_2^\epsilon$$

En caso que el determinante de dicho sistema se anule al valuar en las soluciones, i.e. $(g_1^s g_2^b - g_1^b g_2^s)(s(\epsilon),b(\epsilon),\epsilon) = 0$ resultaría que

$$g_1^b(s(\epsilon),b(\epsilon),\epsilon) = \lambda(\epsilon)g_1^s(s(\epsilon),b(\epsilon),\epsilon) \quad \text{y} \quad g_2^b(s(\epsilon),b(\epsilon),\epsilon) = \lambda(\epsilon)g_2^s(s(\epsilon),b(\epsilon),\epsilon)$$

y reemplazando en el sistema anterior surge que

$$0 = g_1^s s'(\epsilon) + \lambda(\epsilon) g_1^s b'(\epsilon) + g_1^\epsilon$$
$$0 = g_2^s s'(\epsilon) + \lambda(\epsilon) g_2^s b'(\epsilon) + g_2^\epsilon$$

$$0 > -(K_2 - E_{F_0}\psi_2) = g_1^s(s'(\epsilon) + \lambda(\epsilon)b'(\epsilon)) \text{ lo cual implica } s' + \lambda b' > 0$$
$$0 < -g_2^\epsilon = g_2^s(s'(\epsilon) + \lambda(\epsilon)b'(\epsilon)) \text{ lo cual implica } s' + \lambda b' < 0$$

Luego el determinante no se anula si $\epsilon \in \mathbf{E}_0$ y por continuidad del determinante y de las funciones involucradas resulta que no se anula si $\epsilon \in \mathbf{O} = \bigcup_{\tilde{\epsilon} \in \mathbf{E}_0} I(\tilde{\epsilon})$ y tal conjunto es abierto (por ser unión de abiertos) y denso (pues $\mathbf{E}_0$ es denso). En consecuencia $\mathbf{O} = \bigcup_{i=1}^{\infty}(a_i, a_{i+1})$ donde $\{a_i\}_{i=1}^{\infty}$ es una sucesión estrictamente creciente en $\mathbf{E}$. Aplicando nuevamente el teorema de la función implícita surge que

$$s'(\epsilon) = -\frac{g_2^b g_1^\epsilon - g_1^b g_2^\epsilon}{g_1^s g_2^b - g_2^s g_1^b} \qquad b'(\epsilon) = \frac{g_2^s g_1^\epsilon - g_1^s g_2^\epsilon}{g_1^s g_2^b - g_2^s g_1^b}$$

Como $b(\epsilon)$ es creciente, $b'(\epsilon) \geq 0$ y $g_2^s g_1^\epsilon - g_1^s g_2^\epsilon < 0$ con lo cual resulta que $(g_1^s g_2^b - g_2^s g_1^b)(s(\epsilon), b(\epsilon), \epsilon) < 0$. Esto dice que $s(\epsilon)$ es también creciente.

Tomando ahora la familia de funciones $\psi_2^c = m\psi_2(E_{F_0}\psi_2(c\mathbf{z}'\mathbf{z}))^{-1}$ $c \geq 1$, el sistema (1.6) resulta ser

$$\begin{aligned}
&(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right) + \epsilon K_2 = E_{F_0}\psi_2(c\mathbf{z}'\mathbf{z}) \\
&(1-\epsilon)mE_{F_0}\psi_2\left(\frac{1}{s}\left(\frac{u}{b} + vb^{1/(m-1)}\right)\right)\frac{b^{m/(m-1)}v}{u + vb^{m/(m-1)}} = (m-1)E_{F_0}\psi_2(c\mathbf{z}'\mathbf{z})
\end{aligned} \tag{1.25}$$

Las soluciones de (1.25) se denotarán por $(s(\epsilon, c), b(\epsilon, c))$ (para $c = 1, b(\epsilon, 1) = b(\epsilon)$); si ϵ está fijo, dichas funciones son continuas como funciones de c (se chequea en la forma usual). $b(\epsilon, c)$ es creciente en ϵ y razonando como se hizo para una ψ_2 general se llamará $\mathbf{O}(c)$ al abierto denso donde $b(\epsilon, c)$ es diferenciable (como función de ϵ). Sea ahora $\{c_n\}_{n=1}^{\infty}$ un denso numerable en $[1,\infty)$, entonces $\bigcap_{n=1}^{\infty}\mathbf{O}(c_n)$ es denso y por lo tanto $b(\epsilon, c_n)$ es diferenciable en $\bigcap_{t=1}^{\infty}\mathbf{O}(c_t)$ $\forall$ n y en consecuencia puede asegurarse que

$$(g_{1,c_n}^s g_{2,c_n}^b - g_{2,c_n}^s g_{1,c_n}^b)(s(\epsilon, c_n), b(\epsilon, c_n), \epsilon) < 0 \quad \forall \quad c_n \text{ y } \epsilon \in \bigcap_{n=1}^{\infty}\mathbf{O}(c_n)$$

Fijando ϵ, surge por el teorema de la función implícita que existen abiertos $\mathbf{V}_1(c_n)$ y $\mathbf{V}_2(c_n)$ en $\Re$ tal que $b(\epsilon,c):\mathbf{V}_1(c_n)\to\mathbf{V}_2(c_n)$ es diferenciablemente continua y su derivada resulta

$$\frac{d}{dc}b(\epsilon,c)=\frac{g^c_{1,c}g^s_{2,c}-g^c_{2,c}g^s_{1,c}}{g^s_{1,c}g^b_{2,c}-g^s_{2,c}g^b_{1,c}} \tag{1.26}$$

Como $g^c_{2,c}=-(m-1)\frac{d}{dc}E_{F_0}\psi_2(c\mathbf{z}'\mathbf{z})=(m-1)g^c_{1,c}$ el numerador de (1.26) resulta ser igual a $g^c_{1,c}(g^s_{2,c}-(m-1)g^s_{1,c})$. Como

$$\begin{aligned}
g^c_{1,c}&=-\frac{d}{dc}E_{F_0}\psi_2(c\mathbf{z}'\mathbf{z})<0 \text{ y además}\\
g^s_{2,c}-(m-1)g^s_{1,c}&=\frac{1-\epsilon}{s^2b}\int\psi_2'\left(\frac{1}{s}\left(\frac{u}{b}+vb^{1/(m-1)}\right)\right)(-b^{m/(m-1)}v+(m-1)u)dF_0\\
&=-\frac{(m-1)s}{b}g^b_1(s,b,\epsilon)\le 0
\end{aligned}$$

se obtiene que $\frac{d}{dc}b(\epsilon,c)\le 0$ si $c\in\mathbf{V}_1(c_n)$. Como $\bigcup_{n=1}^{\infty}\mathbf{V}_1(c_n)$ es un abierto denso en $[1,\infty)$, entonces resulta ser de la forma $\bigcup_{i=1}^{\infty}(a_i,a_{i+1})\cup[1,a_1)]$ con lo cual $b(\epsilon,c)$ es diferenciable salvo probablemente en $\{a_i\}$ y como es continua debe ser decreciente en c. Llámese $b_*=b_*(\epsilon)=\lim_{c\to\infty}b(\epsilon,c)=\inf_{c\ge 1}b(\epsilon,c)$. Obsérvese ahora que

a) $\limsup_{c\to\infty}s(\epsilon,c)=\infty$ o $\liminf_{c\to\infty}s(\epsilon,c)=0$ pues caso contrario, si $0<\liminf_{c\to\infty}s(\epsilon,c)\le\limsup_{c\to\infty}s(\epsilon,c)<\infty$ tras tomar límite surge que

$$K_2=(1-\epsilon)E_{F_0}\psi_2\left(\frac{1}{s_*}\left(\frac{u}{b_*}+vb_*\right)\right)+\epsilon K_2<(1-\epsilon)K_2+\epsilon K_2$$

y ello es absurdo. En caso que

b) $\liminf_{c\to\infty}s(\epsilon,c)=0$ resulta que

$$(1-\epsilon)mK_2E_{F_0}\frac{b_*^{m/(m-1)}v}{u+vb_*^{m/(m-1)}}=(m-1)K_2$$

con lo cual $b_*=\mathbf{b_T}(\epsilon)$ y como $b(\epsilon,c)$ es decreciente en c implica que

$$b(\epsilon)\le\mathbf{b_T}(\epsilon)$$

c) Si $\limsup_{c\to\infty}s(\epsilon,c)=\infty$ entonces

$$(1-\epsilon)mm_0E_{F_0}\frac{b_*^{m/(m-1)}v}{u+vb_*^{m/(m-1)}}=(m-1)K_2$$

y en caso de $b_* < \mathbf{b_T}(\epsilon)$ resultaría que

$$1 < (1-\epsilon)\frac{m}{K_2} m_0 E_{F_0} \frac{\mathbf{b_T}(\epsilon)^{m/(m-1)} v}{u + v\mathbf{b_T}(\epsilon)^{m/(m-1)}} \leq (m-1)\frac{m_0}{K_2}$$

lo que diría que $m_0 \geq K_2$ y en consecuencia

$$b(\epsilon) \leq b_* \leq \mathbf{b_T}(\epsilon)$$

Entonces el sesgo del estimador de Tyler es menor que el obtenido para un M-estimador si $\epsilon \in \bigcap_{t=1}^{\infty} \mathbf{O}(c_t)$. Si $\epsilon_0 \in (\bigcap_{t=1}^{\infty} \mathbf{O}(c_t))^c \cap \mathbf{E}$, y $b(\epsilon_0) < \mathbf{b_T}(\epsilon_0)$ resulta, por continuidad de $b(\epsilon)$ y $\mathbf{b_T}(\epsilon)$ que $b(\epsilon) < \mathbf{b_T}(\epsilon) \quad \forall \quad \epsilon \in (\epsilon_0 - \delta, \epsilon + \delta)$ para algún δ suficientemente pequeño, pero como $\bigcap_{t=1}^{\infty} \mathbf{O}(c_t)$ es denso, $\bigcap_{t=1}^{\infty} \mathbf{O}(c_t) \cap (\epsilon_0 - \delta, \epsilon + \delta) \neq \emptyset$ y por lo anteriormente visto no puede mejorarse el máximo sesgo. Por lo tanto $b(\epsilon) \geq \mathbf{b_T}(\epsilon) \quad \forall \quad \epsilon \in \mathbf{E}$.

El Lema 6 c) permite concluir que la solución $(s(\epsilon), b(\epsilon))$ para ψ_2 satisface que $b(\epsilon) = \lim_{n\to\infty} b^{(n)} \geq \mathbf{b_T}$. ∎

4. ALGUNOS RESULTADOS COMPLEMENTARIOS.

4.1 Identificabilidad del modelo. .

Sea $\mathbf{z}$ un vector m-dimensional con distribución elipsoidal, i.e. $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{t}_0$ donde $\mathbf{A}$ es una matriz no singular $m \times m$, $\mathbf{t}_0$ es un vector m-dimensional y $\mathbf{x}$ tiene distribución esféricamente simétrica con densidad. Si $\mathbf{V}_0 = \mathbf{A}\mathbf{A}'$ y f_0 es una función positiva, conocida, decreciente en $\|\mathbf{z}\|$ entonces $\mathbf{V}_0 = \mathbf{A}\mathbf{A}'$ y $\mathbf{t}_0$ son identificables, i.e. si $\mathbf{A}_0\mathbf{z} + \mathbf{t}_0$ tiene la misma distribución de $\mathbf{A}_1\mathbf{z} + \mathbf{t}_1$ entonces $\mathbf{A}_0\mathbf{A}_0' = \mathbf{A}_1\mathbf{A}_1'$ y $\mathbf{t}_0 = \mathbf{t}_1$. Igualando las funciones características de ambas distribuciones se tendría que (Muirhead - 1982)

$$\exp(-i\mathbf{t}\mathbf{t}_o)\psi(\mathbf{t}'\mathbf{A}_0\mathbf{A}_0'\mathbf{t}) = \exp(-i\mathbf{t}\mathbf{t}_1)\psi(\mathbf{t}'\mathbf{A}_1\mathbf{A}_1'\mathbf{t})$$

Como ψ es una función real entonces $\exp(-i\mathbf{t}(\mathbf{t}_o - \mathbf{t}_1)) \in \Re$ y por lo tanto $\mathbf{t}_0 = \mathbf{t}_1$. Entonces para todo conjunto $\mathbf{B} \in \Re^m$ medible y $\mathbf{T} = \mathbf{A}_1^{-1}\mathbf{A}_0$ se cumple que

$$\int f_0(\|\mathbf{z}\|^2)I_{\mathbf{B}}d\mathbf{z} = \int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}(\mathbf{B})}d\mathbf{z}$$

o igualmente

$$\int f_0(\|\mathbf{z}\|^2)I_{\mathbf{B}}d\mathbf{z} = \int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}^n(\mathbf{B})}d\mathbf{z}$$

Como $\mathrm{Vol}(\mathbf{T}^n(\mathbf{B})) = |\det(\mathbf{T}^n)|\mathrm{Vol}(\mathbf{B})$ al tomar límite para n tendiendo a infinito resulta que $|\det(\mathbf{T})| = 1$ y por lo tanto $\mathrm{Vol}(\mathbf{T}(\mathbf{B})) = \mathrm{Vol}(\mathbf{B})$ para todo $\mathbf{B}$. Además

$$\int f_0(\|\mathbf{z}\|^2)I_{\mathbf{B}}d\mathbf{z} = \int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}(\mathbf{B})\cap\mathbf{B}}d\mathbf{z} + \int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}(\mathbf{B})\cap\mathbf{B}^c}d\mathbf{z} =$$
$$\int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}(\mathbf{B})\cap\mathbf{B}}d\mathbf{z} + \int f_0(\|\mathbf{z}\|^2)I_{(\mathbf{T}(\mathbf{B}))^c\cap\mathbf{B}}d\mathbf{z}$$

Sea $\mathbf{B}_{0,\mathbf{r}} = B(0,r) = \{\mathbf{z}/\|\mathbf{z}\| < r\}$. Entonces $\mathrm{Vol}(\mathbf{T}(\mathbf{B}_{0,\mathbf{r}}) \cap \mathbf{B}_{0,\mathbf{r}}^c) = \mathrm{Vol}((\mathbf{T}(\mathbf{B}_{0,\mathbf{r}}))^c \cap \mathbf{B}_{0,\mathbf{r}})$. En caso que $\mathbf{T}(\mathbf{B}_{0,\mathbf{r}}) \cap \mathbf{B}_{0,\mathbf{r}}^c \neq \emptyset$, $\int f_0(\|\mathbf{z}\|^2)I_{\mathbf{T}(\mathbf{B}_{0,\mathbf{r}})\cap\mathbf{B}_{0,\mathbf{r}}^c}d\mathbf{z} < \int f_0(\|\mathbf{z}\|^2)I_{(\mathbf{T}(\mathbf{B}_{0,\mathbf{r}}))^c\cap\mathbf{B}_{0,\mathbf{r}}}d\mathbf{z}$ pero eso contradice la igualdad conseguida. Luego resulta que $\mathbf{T}(\mathbf{B}_{0,\mathbf{r}}) = \mathbf{B}_{0,\mathbf{r}}$ para cualquier $\mathbf{B}_{0,\mathbf{r}}$. En consecuencia $\mathbf{T}(\mathbf{B}_{0,\mathbf{r}} - \mathbf{B}_{0,\mathbf{r}'}) = \mathbf{B}_{0,\mathbf{r}} - \mathbf{B}_{0,\mathbf{r}'}$ para todo $r > r'$ y por lo tanto $\mathbf{T}(S^{m-1}) = S^{m-1}$ si S^{m-1} es la esfera $(m-1)$-dimensional en $\Re^m$. Luego $\mathbf{T}$ es ortogonal, y en consecuencia el parámetro es identificable.

4.2 Unicidad del estimador de Tyler en el entorno $\mathcal{F}_\epsilon$.

LEMA. *Sea $F \in \mathcal{F}_\epsilon$. Si $\mathbf{M}(\Gamma_1) = \mathbf{M}(\Gamma_2)$ entonces $\Gamma_1 = c\Gamma_2$ con $c > 0$.*

DEMOSTRACIÓN: Basta suponer que $\Gamma_2 = \mathbf{I}$. Sea λ_1 el máximo autovalor de Γ_1 y $\mathbf{P}_1$ su proyección asociada. Entonces $\mathbf{P}_1\mathbf{M}(\Gamma_1) = \mathbf{P}_1\mathbf{M}(\mathbf{I})$. En especial

$$\mathbf{P}_1\mathbf{M}(\Gamma_1) = mE_F\frac{\gamma_1^{1/2}\mathbf{P}_1\mathbf{z}\mathbf{z}'\Gamma_1}{\mathbf{z}'\Gamma\mathbf{z}}$$

$$\mathbf{P}_1\mathbf{M}(\mathbf{I}) = mE_F\frac{\mathbf{P}_1\mathbf{z}\mathbf{z}'}{\mathbf{z}'\Gamma\mathbf{z}}$$

$$\mathrm{tr}(\mathbf{P}_1\mathbf{M}(\Gamma_1)) = m\gamma_1E_F\frac{\mathrm{tr}(\mathbf{P}_1\mathbf{z}\mathbf{z}')}{\mathbf{z}'\Gamma\mathbf{z}} = m\gamma_1E_F\frac{\mathbf{z}'\mathbf{P}_1\mathbf{z}}{\mathbf{z}'\Gamma\mathbf{z}}$$

$$\mathrm{tr}(\mathbf{P}_1\mathbf{M}(\mathbf{I})) = mE_F\frac{\mathrm{tr}(\mathbf{P}_1\mathbf{z}\mathbf{z}')}{\mathbf{z}'z} = mE_F\frac{\mathbf{z}'\mathbf{P}_1\mathbf{z}}{\mathbf{z}'\mathbf{z}}$$

Por lo tanto

$$E_F\frac{\mathbf{z}'\mathbf{P}_1\mathbf{z}}{\mathbf{z}'\mathbf{z}} = m\gamma_1E_F\frac{\mathbf{z}'\mathbf{P}_1\mathbf{z}}{\mathbf{z}'\Gamma\mathbf{z}} \tag{4.1}$$

Sean $\gamma_1 = \ldots = \gamma_{j_1}$. Si $\oplus$ denota la suma directa de subespacios, sea $\mathrm{C} = (\mathrm{Im}(\mathbf{P}_1)\oplus\ldots\oplus \mathrm{Im}(\mathbf{P}_{j_1})^c$. Supóngase que $\mathrm{C} \neq \emptyset$. $\mathrm{C}\cap \mathrm{sop}(F) \neq \emptyset$ pues F_0 tiene distribución elíptica, y por el mismo motivo

$$\mathrm{C}\cap \mathrm{sop}(F)\cap(\mathrm{Nu}(\mathbf{P}_1))^c \neq \emptyset \tag{4.2}$$

Si $\mathbf{z} \in \mathrm{C}$ entonces $(\gamma_1)^{-1}\mathbf{z}'\Gamma\mathbf{z} < \mathbf{z}'\mathbf{z}$. Luego para preservar la igualdad (4.1), si $\mathbf{z} \in \mathrm{C}\cap\mathrm{sop}(F)$, entonces $\mathbf{z}'\mathbf{P}_1\mathbf{z} = 0$ y por lo tanto $\mathbf{z} \in \mathrm{Nu}(\mathbf{P}_1)$. Luego $\mathrm{C}\cap\mathrm{sop}(F) \subseteq \mathrm{Nu}(\mathbf{P}_1)$. Pero ello contradice (4.2). Luego $\mathrm{C} = \emptyset$, o sea $j_1 = m$ y por lo tanto $\Gamma = c\mathbf{I}$. ∎

Este lema permite entonces deducir la unicidad de la solución de las ecuaciones de Tyler en caso de fijar la condición $\mathrm{tr}(\Gamma) = m$ o $\det(\Gamma) = 1$.

4.3 Existencia del estimador de Tyler en el entorno $\mathcal{F}_\epsilon$.

Para ver la existencia de soluciones para el estimador de Tyler en el entorno de ϵ-contaminación, la prueba sigue muy cercamente a la demostración de existencia de soluciones para distribuciones empíricas que se muestra en Tyler (1987).

Sea el proceso iterativo $\mathbf{M}_k = \mathbf{M}(\Gamma_k) = m'E\mathbf{C}_k\mathbf{z}\mathbf{z}'\mathbf{C}_k'/\mathbf{z}'\Gamma_k\mathbf{z}I_{\Re^m-\{0\}}(\mathbf{z})$ donde $m' = m/P(z \neq 0)$, $\Gamma_k = \mathbf{C}_k\mathbf{C}_k'$ y $\Gamma_{k+1} = \mathbf{C}_k'\mathbf{M}_k^{-1}\mathbf{C}_k/\mathrm{tr}(\Gamma_k\mathbf{M}_k^{-1})$. Sea $\mathbf{H}_k$ una matriz ortogonal tal que $\mathbf{M}_k^{-1/2}\mathbf{C}_k = \mathbf{H}_k\mathbf{C}_{k+1}/(\mathrm{tr}(\Gamma_k\mathbf{M}_k^{-1}))^{1/2}$. Sea $\lambda_{1,k} = \lambda_{(1)}(\mathbf{M}_k)$ el máximo autovalor de $\mathbf{M}_k$ y $\lambda_{m,k} = \lambda_{(m)}(\mathbf{M}_k)$ el mínimo autovalor de $\mathbf{M}_k$. Para definir bien el proceso es necesario tener una matriz inicial tal que $\mathbf{M}_1$ sea inversible. Entonces tomando $\Gamma_1 = \mathbf{I}$ se tiene que si la invertibilidad no se cumple entonces existe $\mathbf{z}_0$ tal que $E|\mathbf{z}'\mathbf{z}_0|^2/(\mathbf{z}'\mathbf{z})I_{\Re^m-\{0\}}(\mathbf{z}) = E\mathbf{z}_0'\mathbf{z}\mathbf{z}'\mathbf{z}_0/(\mathbf{z}'\mathbf{z})I_{\Re^m-\{0\}}(\mathbf{z}) = 0$ con lo cual $E_{F_0}|\mathbf{z}'\mathbf{z}_0|^2/(\mathbf{z}'\mathbf{z})I_{\Re^m-\{0\}}(\mathbf{z}) = 0$ y $F_0(\mathbf{z}'\mathbf{z}_0 = 0) = 1$ pero F_0 tiene densidad y ello es absurdo. Por lo tanto va a existir $\mathbf{M}_k^{-1}$ para todo k. Si G_1 denota la contaminación usada, entonces $\mathbf{M}_1 = (1-\epsilon)(m'/m)\mathbf{I} + \epsilon m E_{G_1}\mathbf{z}\mathbf{z}'/(\mathbf{z}'\mathbf{z}) \geq (1-\epsilon)\mathbf{I}$. Por lo tanto $\lambda_{m,1} \geq m/(m+1)$.

LEMA. *La sucesión de autovalores máximos $\lambda_{1,k}$ es decreciente y la sucesión de autovalores mínimos $\lambda_{m,k}$ es creciente. Además $\lambda_{m,k} \leq 1 \leq \lambda_{1,k}$ para todo k.*

DEMOSTRACIÓN:

$$\begin{aligned}
\mathbf{I} &= m'E\mathbf{M}_k^{-1/2}\mathbf{C}_k\mathbf{z}\mathbf{z}'\mathbf{C}_k'\mathbf{M}_k^{-1/2}/\mathbf{z}'\Gamma_k\mathbf{z}I_{\Re^m-\{0\}}(\mathbf{z}) \\
&= m'E\left(\mathbf{H}_k\mathbf{C}_{k+1}\mathbf{z}\mathbf{z}'\mathbf{C}_{k+1}'\mathbf{H}_k'/\mathbf{z}'\Gamma_{k+1}\mathbf{z}\right)\left(\mathbf{z}'\Gamma_k^{1/2}\mathbf{M}_k^{-1}\Gamma_k^{1/2}\mathbf{z}/\mathbf{z}'\Gamma_k\mathbf{z}\right)I_{\Re^m-\{0\}}(\mathbf{z}) \\
&= m'E\left(\mathbf{H}_k\mathbf{C}_{k+1}\mathbf{z}\mathbf{z}'\mathbf{C}_{k+1}'\mathbf{H}_k'/\mathbf{z}\Gamma_{k+1}\mathbf{z}\right)\left(\mathbf{z}'\Gamma_k^{1/2}\mathbf{M}_k^{-1}\Gamma_k^{1/2}\mathbf{z}/\mathbf{z}'\Gamma_k\mathbf{z}\right)I_{\Re^m-\{0\}}(\mathbf{z}) \\
&= m'E\left(\mathbf{H}_k\mathbf{C}_{k+1}\mathbf{z}/\|\mathbf{C}_{k+1}\mathbf{z}\|\right)\left(\mathbf{H}_k\mathbf{C}_{k+1}\mathbf{z}/\|\mathbf{C}_{k+1}\mathbf{z}\|\right)'\left(\mathbf{z}'\Gamma_k^{1/2}\mathbf{M}_k^{-1}\Gamma_k^{1/2}\mathbf{z}/\mathbf{z}'\Gamma_k\mathbf{z}\right)I_{\Re^m-\{0\}}(\mathbf{z}) \\
&= E\left(\mathbf{M}_{k+1}^{o}\right)'\left(\mathbf{z}'\Gamma_k^{1/2}\mathbf{M}_k^{-1}\Gamma_k^{1/2}\mathbf{z}/\mathbf{z}'\Gamma_k\mathbf{z}\right)
\end{aligned}$$

donde $E\mathbf{M}_{k+1}^{o}$ tiene los mismos autovalores (aunque distintos autovectores) que $\mathbf{M}_{k+1}$. Como

$$\sup_{\mathbf{z}\neq 0}\frac{\mathbf{z}'\Gamma^{1/2}\mathbf{M}_k^{-1}\Gamma^{1/2}\mathbf{z}}{\mathbf{z}'\Gamma^{1/2}\Gamma^{1/2}\mathbf{z}} = \lambda_{(1)}(\mathbf{M}_k^{-1}) = \lambda_{(m)}^{-1}(\mathbf{M}_k) \text{ y } \inf_{\mathbf{z}\neq 0}\frac{\mathbf{z}'\Gamma^{1/2}\mathbf{M}_k^{-1}\Gamma^{1/2}\mathbf{z}}{\mathbf{z}'\Gamma^{1/2}\Gamma^{1/2}\mathbf{z}} = \lambda_{(1)}^{-1}(\mathbf{M}_k)$$

resulta que

$$\lambda_{1,k}^{-1}E\mathbf{M}_{k+1}^{o} = \lambda_{(1)}^{-1}(\mathbf{M}_k)E\mathbf{M}_{k+1}^{o} \leq \mathbf{I} \leq \lambda_{(m)}^{-1}(\mathbf{M}_k)E\mathbf{M}_{k+1}^{o} = \lambda_{m,k}^{-1}E\mathbf{M}_{k+1}^{o}$$

La primer desigualdad dice que $\lambda_{1,k} \geq \lambda_{1,k+1}$ y la segunda desigualdad que $\lambda_{m,k} \leq \lambda_{m,k+1}$, o sea que la sucesión de autovalores máximos es decreciente y la sucesión de autovalores mínimos es creciente. Además $\mathrm{tr}E(\mathbf{M}_k^{o}) = m$ con lo cual $\lambda_{m,k} \leq 1 \leq \lambda_{1,k}$ para todo k y por lo tanto vale el resultado deseado. ∎

TEOREMA. *Si* $\epsilon < 1/(m+1)$ *entonces el estimador de Tyler existe para toda distribución en el entorno de contaminación* $\mathcal{F}_\epsilon$.

DEMOSTRACIÓN: $\mathrm{tr}(\Gamma_j) = 1$, entonces existe una subsucesión $\Gamma_j = \Gamma_{k_j} \to \Gamma = \mathbf{C}'\mathbf{C}$, una matriz simétrica definida positiva o semidefinida positiva. Γ_j puede elegirse para que $\mathbf{C}_j\mathbf{z}/(\mathbf{z}'\Gamma_j\mathbf{z})^{1/2} \to \theta(\mathbf{z})$ si $\mathbf{z} \in \mathrm{Nu}(\Gamma)$. Para ello se elige la subsucesión para converger en la base de $\mathrm{Nu}(\Gamma)$ y por lo tanto converge en todo el subespacio. Luego si $\{\mathbf{z}_i\}_{i=1}^{s}$ denota una base de $\mathrm{Nu}(\Gamma)$ y $\mathbf{z} = \sum_{i=1}^{s} a_i(\mathbf{z})\mathbf{z}_i$ entonces $\theta(\mathbf{z}) = \sum_{i=1}^{s} b_i(\mathbf{z})\theta(\mathbf{z}_i)$ donde $b_i(\mathbf{z}) = a_i(\mathbf{z})/\|\sum_{i=1}^{s} a_i(\mathbf{z})\theta(\mathbf{z}_i)\|$. Luego la función $\theta(\mathbf{z})$ es continua en $\mathrm{Nu}(\Gamma) - \{0\}$. Entonces la subsucesión $\mathbf{M}_j$ converge a

$$\mathbf{M} = m'E\frac{\mathbf{Czz'C'}}{\mathbf{z'}\Gamma\mathbf{z}}I_{(\mathrm{Nu}(\Gamma))^c} + m'E\theta(\mathbf{z})\theta(\mathbf{z})'I_{\mathrm{Nu}(\Gamma)-\{0\}}$$

que es una matriz no singular por el resultado del lema anterior. Si $\Gamma_{j+1} = \Gamma_{k_j+1}$ entonces esta sucesión converge a $\Gamma_0 = \mathbf{C}'\mathbf{M}^{-1}\mathbf{C}/\mathrm{tr}(\mathbf{M}^{-1}\Gamma) = \mathbf{C}_0'\mathbf{C}_0$ y por lo tanto $\mathrm{Nu}(\Gamma) = \mathrm{Nu}(\Gamma_0)$. Como $\mathbf{C}_{j+1}^{-1}\mathbf{M}_j^{-1/2}\mathbf{C}_j = \mathbf{H}_j$ es una sucesión de matrices ortogonales convergente a una matriz ortogonal $\mathbf{H}$, entonces la sucesión $\mathbf{M}_{j+1} = \mathbf{M}_{k_j+1}$ converge a

$$\mathbf{M}_0 = m'E\mathbf{C}_0'\mathbf{zz}'\mathbf{C}_0 I_{(\mathrm{Nu}(\Gamma))^c} + m'E\phi(\mathbf{z})\phi'(\mathbf{z})I_{\mathrm{Nu}(\Gamma)-\{0\}}$$

donde $\phi(\mathbf{z}) = \mathbf{HM}^{-1/2}\theta(\mathbf{z})/(\theta(\mathbf{z})'\mathbf{M}^{-1}\theta(\mathbf{z}))^{1/2}$. El lema previo asegura que $\mathbf{M}$ y $\mathbf{M}_0$ poseen los mismos autovalores pues la sucesión de autovalores son convergentes. Sean λ_1 y λ_m el mayor y menor autovalor de ambas matrices. Sean $\mathbf{P}$ y $\mathbf{P}_0$ las proyecciones asociadas con el autovalor máximo de $\mathbf{M}$ y $\mathbf{M}_0$ respectivamente, y sean $r = \mathrm{rango}(\mathbf{P})$ y $r_0 = \mathrm{rango}(\mathbf{P}_0)$. Siempre puede asumirse que $r_0 \geq r$, pues si la condición no vale entonces en vez de considerarse el par (Γ_j, Γ_{j+1}) puede tomarse $(\Gamma_{j+1}, \Gamma_{j+2})$ o en general $(\Gamma_{j+i}, \Gamma_{j+i+1})$. Si r_i denota el rango de la proyección correspondiente al máximo autovalor para la sucesión Γ_{j+i} entonces no puede ocurrir permanentemente que $r_i < r_{i+1}$ para todo i ($r_i \leq m$).

$$\mathbf{I} = m'E\frac{\mathbf{M}^{-1/2}\mathbf{Czz'C'M}^{-1/2}}{\mathbf{z'}\Gamma\mathbf{z}}I_{(\mathrm{Nu}(\Gamma))^c} + m'E\mathbf{M}^{-1/2}\theta(\mathbf{z})\theta(\mathbf{z})'\mathbf{M}^{-1/2}I_{\mathrm{Nu}(\Gamma)-\{0\}} \qquad (4.3)$$

o equivalentemente

$$\begin{aligned}\mathbf{I} =&\lambda_1^{-1}\mathbf{M}_0 + m'E\frac{\mathbf{C}_0\mathbf{zz}'\mathbf{C}_0'}{\mathbf{z}'\Gamma_0\mathbf{z}}\left(\frac{\mathbf{z'CM}^{-1}\mathbf{C'z}}{\mathbf{z'}\Gamma\mathbf{z}} - \lambda_1^{-1}\right)I_{(\mathrm{Nu}(\Gamma))^c} + \\ &m'E\phi(\mathbf{z})\phi(\mathbf{z})'(\theta(\mathbf{z})'\mathbf{M}^{-1}\theta(\mathbf{z}) - \lambda_1^{-1})I_{\mathrm{Nu}(\Gamma)-\{0\}} \qquad (4.4)\end{aligned}$$

Al pre y posmultiplicar en la última igualdad por $\mathbf{P}_0$ y dado que que $\lambda^{-1}\mathbf{P}_0\mathbf{M}_0\mathbf{P}_0 = \mathbf{P}_0$, resulta que

$$0 = m'E\frac{\mathbf{P}_0\mathbf{C}_0\mathbf{z}\mathbf{z}'\mathbf{C}_0'\mathbf{P}_0}{\mathbf{z}'\Gamma_0\mathbf{z}}\left(\frac{\mathbf{z}'\mathbf{C}\mathbf{M}^{-1}\mathbf{C}'\mathbf{z}}{\mathbf{z}'\Gamma\mathbf{z}} - \lambda_1^{-1}\right)I_{(\mathrm{Nu}(\Gamma))^c} +$$
$$m'E\left(\mathbf{P}_0\phi(\mathbf{z})\phi(\mathbf{z})'\mathbf{P}_0(\theta(\mathbf{z})'\mathbf{M}^{-1}\theta(\mathbf{z}) - \lambda_1^{-1}\right)I_{\mathrm{Nu}(\Gamma)-\{0\}}$$

Como $\theta(\mathbf{z})'\theta(\mathbf{z}) = 1$ entonces $\theta(\mathbf{z})'\mathbf{M}^{-1}\theta(\mathbf{z}) - \lambda_1^{-1} \geq 0$ y el mismo razonamiento vale para decir que $((\mathbf{z}'\mathbf{C}\mathbf{M}^{-1}\mathbf{C}'\mathbf{z}/\mathbf{z}'\Gamma\mathbf{z}) - \lambda_1^{-1}) \geq 0$ por lo que ambas sumandos definen matrices no negativas definidas y debe entonces ocurrir que $\mathbf{P}_0\mathbf{H}'\mathbf{M}^{-1/2}\mathbf{C}\mathbf{z} = 0$ o $\mathbf{P}\mathbf{C}\mathbf{z} = 0$ si $\mathbf{z} \in (\mathrm{Nu}(\Gamma))^c$ y que $\mathbf{P}_0\mathbf{H}'\mathbf{M}^{-1/2}\theta(\mathbf{z}) = 0$ o $\mathbf{P}\theta(\mathbf{z}) = \theta(\mathbf{z})$ si $\mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$. Esta conclusión sugiere premultiplicar ambos miembros de (4.3) por $\mathbf{P}_0\mathbf{H}'$ y posmultiplicar por $\mathbf{I}-\mathbf{P}$ con lo cual el miembro derecho se anula y resulta $\mathbf{P}_0\mathbf{H}'(\mathbf{I}-\mathbf{P}) = 0$ o equivalentemente $\mathbf{H}\mathbf{P}_0\mathbf{H}' = \mathbf{H}\mathbf{P}_0\mathbf{H}'\mathbf{P}$. Que $\mathbf{P} = \mathbf{H}\mathbf{P}_0\mathbf{H}'$ resulta del hecho más general siguiente: si se tienen dos proyecciones ortogonales $\mathbf{P}_1$ y $\mathbf{P}_2$ tal que $\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_1$ y $\mathrm{rg}(\mathbf{P}_2) \geq \mathrm{rg}(\mathbf{P}_1)$ entonces $\mathbf{P}_2 = \mathbf{P}_1$. Para ver esto nótese que si $\mathbf{z} \in \mathrm{Nu}(\mathbf{P}_2)$ entonces $\mathbf{z} \in \mathrm{Nu}(\mathbf{P}_1)$ y por lo tanto $\mathrm{Nu}(\mathbf{P}_2) \subseteq \mathrm{Nu}(\mathbf{P}_1)$ y $\mathrm{Im}(\mathbf{P}_1) \subseteq \mathrm{Im}(\mathbf{P}_2)$. Luego $\mathrm{rg}(\mathbf{P}_2) = \mathrm{rg}(\mathbf{P}_1)$, $\mathrm{Im}(\mathbf{P}_1) = \mathrm{Im}(\mathbf{P}_2)$ y $\mathbf{P} = \mathbf{H}\mathbf{P}_0\mathbf{H}'$. A partir de esto se puede inferir que $\forall\, \mathbf{z} \in (\mathrm{Nu}(\Gamma))^c$ se cumple que $\mathbf{P}\mathbf{M}^{-1/2}\mathbf{C}\mathbf{z} = 0$ o $\mathbf{P}\mathbf{C}\mathbf{z} = \mathbf{C}\mathbf{z}$ o equivalentemente $\mathbf{P}\mathbf{C}\mathbf{z} = 0$ o $\mathbf{P}\mathbf{C}\mathbf{z} = \mathbf{C}\mathbf{z}$, y $\forall\, \mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$ se tiene igualmente que $\mathbf{P}\theta(\mathbf{z}) = 0$ o $\mathbf{P}\theta(\mathbf{z}) = \theta(\mathbf{z})$. Nótese asimismo que, dado $\mathbf{z}_0 \in S^{m-1}$, W un subespacio de $\Re^m$ tal que $\{0\} \subsetneq W \subsetneq \Re^m$, y $h(\mathbf{z}) : W - \{0\} \to S^{m-1}$ una función continua entonces

$$m'E\mathbf{z}_0'h(\mathbf{z})h(\mathbf{z})'\mathbf{z}_0 I_{W-\{0\}} < 1 \tag{4.5}$$

puesto que $m\mathrm{P}(W - \{0\})/\mathrm{P}(\mathbf{z} \neq 0) \leq m\epsilon/(1-\epsilon)$. La última cota es creciente en ϵ y por lo tanto puede decirse que es menor que valuar en $1/(m+1)$ obteniéndose así el valor 1 y resultando la desigualdad deseada.

A) Si $\dim(\mathrm{Nu}(\Gamma)) = m - 1$ entonces $\Re^m - \mathrm{Nu}(\Gamma)$ tiene dos componentes conexas (para dimensión $m = 2$ es el único caso a tratar). Entonces $\Gamma = \mathbf{t}\mathbf{t}'$ donde $\mathbf{t} \in \Re^m$ y $\|\mathbf{t}\| = 1$. Luego

$$\mathbf{M} = m'\mathbf{t}\mathbf{t}'\mathrm{P}(\mathrm{Nu}(\Gamma))^c) + \epsilon m' E\theta(\mathbf{z})\theta(\mathbf{z})' I_{\mathrm{Nu}(\Gamma)-\{0\}} \tag{4.6}$$

Como $\theta(\mathbf{z}) : \mathrm{Nu}(\Gamma) - \{0\} \to S^{m-1}$ es continua y $\mathrm{Nu}(\Gamma) - \{0\}$ es conexo entonces d1)$\theta(\mathbf{z}) \in \mathrm{Nu}(\mathbf{P})\forall\ \mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$ o d2)$\theta(\mathbf{z}) \in \mathrm{Im}(\mathbf{P})\forall\ \mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$. Luego, si ocurre la primer situación al componer en (4.6) con $\mathbf{P}$ surge que

$$\lambda_1 \mathbf{P} = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c)\mathbf{P}\mathbf{t}\mathbf{t}' \tag{4.7}$$

por lo cual $\mathbf{Pt} \neq 0$ o sea que $\mathbf{t}$ no pertenece a $\mathrm{Nu}(\mathbf{P})$. Si $\mathbf{z}_0$ denota cualquier vector en $\mathrm{Nu}(\mathbf{P})$ entonces $0 = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c)\mathbf{P}\mathbf{t}\mathbf{t}'\mathbf{z}_0$ lo que dice que $\mathbf{t} \in \mathrm{Im}(\mathbf{P})$. Luego $\mathbf{t}$ es un autovector asociado al máximo autovalor λ_1 y por lo tanto $\lambda_1 = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c) \geq m/\mathrm{P}(\mathbf{z} \neq 0)(1-\epsilon) > m(m/(m+1)) > m-1$. Luego $(m-1)\lambda_m < m - m(m/(m+1))$ o $\lambda_m < (m/(m+1))1/(m-1)$ lo cual contradice la suposición sobre el λ_m inicial. Por lo tanto, entonces podría estar ocurriendo que $\mathrm{Nu}(\mathbf{P}) = \{0\}$ con lo cual $\mathbf{P} = \mathbf{M} = \mathbf{I}$, pero esto junto a (4.7) es absurdo.

Luego resta la posibilidad que $\theta(\mathbf{z}) \in \mathrm{Im}(\mathbf{P}) \forall\ \mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$. Como recién sea $\mathbf{z}_0$ cualquier vector en $\mathrm{Nu}(\mathbf{P})$. Luego $\mathbf{M}\mathbf{z}_0 = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c)\mathbf{t}'\mathbf{z}_0\mathbf{t}$ con lo cual $\mathbf{t}'\mathbf{z}_0 \neq 0\ \forall\ \mathbf{z}_0 \in \mathrm{Nu}(\mathbf{P})$. Por lo tanto $0 = \mathbf{PM}\mathbf{z}_0 = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c)\mathbf{t}'\mathbf{z}_0\mathbf{Pt}$ y $\mathbf{t} \in \mathrm{Nu}(\mathbf{P})$. Consecuentemente $\mathbf{Mt} = m'\mathrm{P}((\mathrm{Nu}(\Gamma))^c)\mathbf{t}$ y es entonces un autovector asociado con el máximo autovalor. Pero eso contradice que $\mathbf{t} \in \mathrm{Nu}(\mathbf{P})$, y sólo cabría que $\mathbf{P} = \mathbf{M} = \mathbf{I}$, pero en ese caso tomando un $\mathbf{z}_0 \in S^{m-1}$ tal que $\mathbf{t}'\mathbf{z}_0 = 0$ se contradice (4.5). Luego Γ debe ser inversible, y entonces $\mathbf{PCz} = 0$ o $\mathbf{PCz} = \mathbf{z}\ \forall\ \mathbf{z} \in \Re^m - \{0\}$. Por un argumento de conexión y continuidad $\mathbf{Cz} \in \mathrm{Nu}(\mathbf{P})$ o $\mathbf{Cz} \in \mathrm{Im}(\mathbf{P})$, pero ello es absurdo si ambos son subespacios propios de $\Re^m$ pues $\mathrm{rg}(\mathbf{C}) = m$. Por lo tanto sólo puede ser $\mathrm{Nu}(\mathbf{P}) = \{0\}$ y por lo tanto $\mathbf{M} = \mathbf{I}$. Luego cualquier otra subsucesión convergente converge a una matriz inversible y por unicidad difiere de Γ por una constante multiplicativa, pero como la traza es 1 entonces $\Gamma_k \to \Gamma$ satisfaciendo $\mathbf{M}(\Gamma) = \mathbf{I}$.

B) Si $1 < \dim(\mathrm{Nu}(\Gamma)) < m-1$ entonces $\Re^m - \mathrm{Nu}(\Gamma)$ es conexo y como $\mathbf{PC}$ es continua entonces ocurre sólo una de las dos situaciones siguientes en $(\mathrm{Nu}(\Gamma))^c$: c1)$\mathbf{PC} = 0$ o c2)$\mathbf{PC} = \mathbf{C}$. Por el mismo argumento sólo puede ocurrir d1) o d2) si $\mathbf{z} \in \mathrm{Nu}(\Gamma) - \{0\}$. Si ocurren c1) y d1) resulta $\mathbf{P} = 0$ y es absurdo. c2) y d1) dicen al componer en (4.4) con $\mathbf{P}$ y tomar traza que $\lambda_1 r = m\mathrm{P}((\mathrm{Nu}(\Gamma))^c)/\mathrm{P}(\mathbf{z} \neq 0)$. Si $r = m$ entonces $\lambda_1 < 1$, y si $r \leq m-1$ surge que

$$\lambda_1 \geq m(1-\epsilon)/r \geq m^2/((m-1)(m+1)) \text{ y entonces } \lambda_m \leq m^2/((m-1)^2(m+1)) \tag{4.8}$$

pero eso contradice la suposición inicial del primer paso de la iteración pues $m/(m+1) \geq m^2/((m-1)^2(m+1))$ si $m \geq 3$. Si ocurre c1) y d2) entonces por (4.5) resultaría que

$\lambda_1 < 1$. c2) y d2) juntos dicen que si $\mathbf{z}_0 \in \mathrm{Nu}(\mathbf{P})$ entonces $\mathbf{M}\mathbf{z}_0 = 0$ y luego $\mathbf{M} = \mathbf{P} = \mathbf{I}$. Nuevamente se toma un $\mathbf{z}_0 \in \mathrm{Nu}(\Gamma) \cap S^{m-1}$ y entonces se contradice (4.5). La conclusión sigue como en A).

C) Si $\dim(\mathrm{Nu}(\Gamma)) = 1$ entonces $E\theta(\mathbf{z})\theta(\mathbf{z})' I_{\mathrm{Nu}(\Gamma)-\{0\}} = \theta(\mathbf{z}_1)\theta(\mathbf{z}_1)' \mathrm{P}(\mathrm{Nu}(\Gamma) - \{0\})$ donde $\mathbf{z}_1$ es el generador de $\mathrm{Nu}(\Gamma)$. Luego si ocurre c1) nuevamente se llega a una contradicción por (4.5). Luego ocurre c2) y si existe $\mathbf{z}_0 \in \mathrm{Nu}(\mathbf{P})$ entonces $\mathbf{M}\mathbf{z}_0 = \theta(\mathbf{z}_1)\theta(\mathbf{z}_1)'\mathbf{z}_0 \neq 0$. Al aplicar $\mathbf{P}$ resulta que $\mathbf{P}(\theta(\mathbf{z}_1)) = 0$ y entonces al componer $\mathbf{P}$ con $\mathbf{M}$ surge una conclusión como en (4.8). Luego sólo queda que $\mathbf{M} = \mathbf{I}$ y todo sigue como en A) y B). ■

NOTA 11: Para contaminaciones en Δ se puede mostrar la existencia de soluciones si y sólo si $\epsilon < 1/m$.

Referencias

Davies,P.L.(1987) *Asymptotic Behavior of S-estimates of Multivariate Location Parameters and Dispersion Matrices* **Ann. Statist.** 15, **1269-1292**

Hampel,F.R.,Ronchetti,E.M.,Rousseeuw,P.J. and Stahel,W.A. (1986) *Robust Statistics: The Approach Based on Influence Functions,* **New York:John Wiley.**

Huber,P.J. (**1964**)*Robust Estimation of a location parameter* **Ann. Math. Stat.**35, **73-101.**

Huber,P.J. (1981) *Robust Statistics,* **John Wiley and Sons, New York.**

Maronna,R.A. (**1976**) *Robust M-estimates of Multivariate Location and Scatter* **Ann. Statist.** 4, **51-67.**

Maronna,R.A., Yohai. V.J. (1990) *The Maximum Bias of Robust Covariances* **Comm. Statist.Theory Meth.** 19, **3925-3933**

Maronna,R.A., Yohai, V.J. (1990) *A new class of Bias-Robust Estimates of Regression* **Manuscript.**

Maronna,R.A.,Yohai,V.J. and Stahel W.A. *Bias-Robust Estimator of Multivariate Scatter based on Projections* **To appear in Journal of Multivariate Analysis.**

Martin,R.D. and Zamar, R.H. (1989) *Asymptotically Min-Max Bias Robust M-Estimates of Scale for Positive Random Variables* **J.A.S.A.** 84,, **494-501**

Martin,R.D., Yohai, V.J. and Zamar, R.H. (1989) *Min-max Bias Robust Regression* **Ann. Statist.** 17, **1608-1630**

Muirhead, Robb(1982) *Aspects of Multivariate Statistical Theory* **John Wiley and Sons, New York.**

Rousseeuw,P.J. and Yohai,V.J. (1984) *Robust Regression by means of S-estimators* **ROBUST AND NONLINEAR TIME SERIES ANALYSIS Lecture Notes in Statistics** 26, **256-272.**

Tyler,D.E. (1981) *Asymptotic Inference for Eigenvectors* **Ann. Statist.** 9, **725-736.**

Tyler,D.E. (1983) *Robustness and Efficiency Properties of Scatter Matrices* **Biometrika,** 70, **411-420.**

Tyler,D.E. (1987) *A Distribution Free M-estimate of Multivariate Scatter* **Ann. Statist.** 15, **234-251**