

# Estimación Robusta de Parámetros en Distribuciones con Datos SAR

Luis H. Pizarro Quiroz

Tesis desarrollada como requisito parcial  
para la obtención del título y grado de

Ingeniero Civil en Informática  
Magister en Ciencias de la Ingeniería Informática

Departamento de Informática  
Universidad Técnica Federico Santa María  
Valparaíso, Chile.

© Luis H. Pizarro Quiroz, Diciembre 2003.

Universidad Técnica Federico Santa María  
Departamento de Informática  
Valparaíso - Chile

TITULO DE LA TESIS:  
**ESTIMACIÓN ROBUSTA DE PARÁMETROS  
EN DISTRIBUCIONES CON DATOS SAR**

AUTOR  
**LUIS H. PIZARRO QUIROZ**

TRABAJO DE GRADO, presentado en cumplimiento parcial de los requisitos para el título de **Ingeniero Civil en Informática** y el grado de **Magister en Ciencias de la Ingeniería Informática** de la Universidad Técnica Federico Santa María.

Dr. Héctor Allende \_\_\_\_\_  
Profesor Guía

Dr. Hubert Hoffmann \_\_\_\_\_  
Correferente

Dr. Jorge Galbiati \_\_\_\_\_  
Correferente Externo

Valparaíso - Chile.  
Diciembre 2003.

## ABSTRACT

Many applications consecrate the use of asymmetric distributions to fit real data, and practical situations often require robust parametric inference. This thesis presents the derivation of M-estimators with asymmetric influence functions, motivated by the  $\mathcal{G}_A^0$  distribution. This law, based on the multiplicative model for speckled imagery, can be highly skewed and maximum likelihood estimation can be severely hampered by small percentages of outliers. These outliers appear mainly because the hypothesis of independence and equal distribution of observations are seldom satisfied in practice; for instance, in the process of filtering, some pixels within a window frequently come from regions with different underlying distributions. Traditional robust estimation methods, based on symmetric robustifying functions, assume that the distribution is symmetric, but when the data distribution is asymmetric these methods yield biased estimators. Empirical influence functions for maximum likelihood estimators are computed, and based on this information this thesis propose the Asymmetric M-estimator (AM-estimator), an M-estimator with asymmetric redescending functions.

## RESUMEN

Muchas aplicaciones utilizan distribuciones asimétricas para ajustar datos reales, y frecuentemente las situaciones prácticas requieren técnicas de inferencia paramétrica robusta. Esta tesis presenta la derivación de los M-estimadores con funciones de influencia asimétricas, motivados en la distribución  $\mathcal{G}_A^0$ . Esta ley, basada en el modelo multiplicativo para imágenes contaminadas con ruido speckle, puede ser altamente sesgada, y la estimación por máxima verosimilitud puede ser severamente obstaculizada por pequeños porcentajes de outliers. Los outliers aparecen principalmente debido a que las hipótesis de independencia e igual distribución de las observaciones rara vez son satisfechas en la práctica; por ejemplo, en el proceso de filtraje, algunos píxeles dentro de una ventana a menudo provienen de regiones con diferentes distribuciones subyacentes. Los métodos tradicionales de estimación robusta, basados en funciones de robustez simétricas, asumen que la distribución es simétrica, pero cuando la distribución de los datos es asimétrica tales métodos producen estimadores sesgados. Las funciones de influencia empírica son calculadas, y basada en esta información, esta tesis propone el Asymmetric M-estimator (AM-estimator), un M-estimator con funciones redescendentes asimétricas.

# Índice general

Resumen	3
Índice de figuras	6
Índice de tablas	7
<b>1. Introducción</b>	<b>1</b>
<b>2. El Modelo Multiplicativo</b>	<b>7</b>
2.1. Introducción . . . . .	7
2.2. Modelo . . . . .	9
2.3. La Distribución $\mathcal{G}_A^0$ . . . . .	12
2.4. Métodos de Estimación Clásicos . . . . .	15
2.4.1. Estimación por Momentos . . . . .	16
2.4.2. Estimación por Máxima Verosimilitud . . . . .	17
<b>3. Métodos de Estimación Robustos</b>	<b>21</b>
3.1. M-Estimadores . . . . .	22
3.2. AM-Estimadores . . . . .	25
<b>4. Estudio de Simulación</b>	<b>28</b>

<i>ÍNDICE GENERAL</i>	5
<b>5. Conclusiones y Trabajo Futuro</b>	<b>42</b>
<b>A. Implementación de los Estimadores Robustos</b>	<b>44</b>
A.1. Funciones Correctoras del M-estimador . . . . .	46
A.2. Funciones Correctoras del AM-estimador . . . . .	47
<b>B. Cómputo de los Parámetros de Ajuste</b>	<b>49</b>
B.1. Varianza Asintótica del ML-estimador . . . . .	50
B.2. Varianza Asintótica del M-estimador . . . . .	50
B.3. Varianza Asintótica del AM-estimador . . . . .	52
<b>Bibliografía</b>	<b>56</b>

# Índice de figuras

1.1.	Imagen SAR de una mina de cobre en el norte de Chile. . . . .	2
1.2.	Imagen SAR afectada por ruido speckle . . . . .	3
1.3.	Interpretación del parámetro de rugosidad $\alpha$ . . . . .	5
2.1.	Diferentes representaciones de una imagen SAR: (a) componente real, (b) componente imaginaria, (c) fase, (d) amplitud, (e) intensidad y (f) log-intensidad. . . . .	8
3.1.	SEIF para el estimador de máxima verosimilitud bajo la distribución $\mathcal{G}_A^0(\alpha, 1)$ , con $-\alpha \in \{1, 3, 6, 15\}$ para $N = 9$ (línea sólida), $N = 25$ (línea punteada), $N = 49$ (línea segmentada) and $N = 81$ (línea punteada-segmentada). . . . .	26
3.2.	Función de robustificación redescendente asimétrica lineal definida por partes. . . . .	27
4.1.	M-estimadores versus AM-estimadores para $\alpha = -1$ . . . . .	31
4.2.	M-estimadores versus AM-estimadores para $\alpha = -10$ . . . . .	32
4.3.	AM-estimadores para $\alpha = -1$ versus tamaño muestral, con nivel de contaminación: $\epsilon = 0,00$ (línea sólida), $\epsilon = 0,01$ (línea punteada), $\epsilon = 0,05$ (línea segmentada) and $\epsilon = 0,10$ (línea punteada-segmentada). . . . .	40
4.4.	AM-estimadores para $\alpha = -6$ versus tamaño muestral, con nivel de contaminación: $\epsilon = 0,00$ (línea sólida), $\epsilon = 0,01$ (línea punteada), $\epsilon = 0,05$ (línea segmentada) and $\epsilon = 0,10$ (línea punteada-segmentada). . . . .	40
4.5.	AM-estimadores para $\alpha = -10$ versus tamaño muestral, con nivel de contaminación: $\epsilon = 0,00$ (línea sólida), $\epsilon = 0,01$ (línea punteada), $\epsilon = 0,05$ (línea segmentada) and $\epsilon = 0,10$ (línea punteada-segmentada). . . . .	41

# Índice de tablas

4.1. Comparación numérica de la media entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 5$ . . . . .	34
4.2. Comparación numérica de la media entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 15$ . . . . .	35
4.3. Comparación numérica del error cuadrático medio entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 5$ . . . . .	36
4.4. Comparación numérica del error cuadrático medio entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 15$ . . . . .	37
4.5. Comparación numérica del sesgo relativo absoluto entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 5$ . . . . .	38
4.6. Comparación numérica del sesgo relativo absoluto entre los estimadores ML, M y AM, variando $\alpha$ , el tamaño muestral $N$ y el nivel de contaminación $\epsilon$ , con $v = 15$ . . . . .	39

# Capítulo 1

## Introducción

La última década estuvo marcada por el fortalecimiento de las imágenes de radar de apertura sintética (SAR: *Synthetic Aperture Radar*) como una herramienta para el monitoreo de nuestro planeta. Numerosos estudios fueron realizados confirmando su relevancia, desarrollándose una serie de técnicas de procesamiento de imágenes especialmente dedicadas a esta tecnología.

La mayoría de las técnicas de procesamiento de imágenes SAR se basan en las propiedades estadísticas de los datos. El conocimiento preciso de estas propiedades juega un rol central en los procedimientos de extracción de información desde éstos. En teledetección (*remote sensing*), por ejemplo, ciertas características pueden ser utilizadas para discriminar distintos tipos y usos de suelo, así como también para desarrollar filtros especializados para la reducción del ruido contenido en las imágenes, entre otras aplicaciones.

Se debe notar que un sensor SAR emite y recibe ondas electromagnéticas de naturaleza compleja, por ende, la señal receptionada puede ser almacenada en diferentes formatos: complejo, intensidad, amplitud y fase (Oliver and Quegan, 1998). No obstante, el 90 % de las aplicaciones hacen uso del formato amplitud, puesto que otorga



Figura 1.1: Imagen SAR de una mina de cobre en el norte de Chile.

una mejor discriminación de las características presentes en la imagen, ver figura 1.1. Es por esta razón que a lo largo de esta tesis se trabaja con el formato amplitud.

La figura 1.1 es representativa de una típica imagen SAR, en ella es posible distinguir varios tipos de rugosidad o textura, característica según la cual se puede clasificar los diferentes tipos de cobertura en:

- **Áreas homogéneas:** Superficies de muy poca o inexistente textura, por ejemplo zonas de cultivo, zonas con deforestación, nieve o agua, entre otros.
- **Áreas heterogéneas:** Superficies que presentan cierto grado de textura, por ejemplo zonas boscosas y relieves no muy pronunciados, entre otros.
- **Áreas extremadamente heterogéneas:** Superficies donde la textura es intensa, por ejemplo zonas urbanas e industriales y de relieves bien pronunciados, entre otros.

Diversos procedimientos estadísticos de filtraje, segmentación y clasificación de imágenes dependen fuertemente de metodologías de inferencia confiables (Allende



Figura 1.2: Imagen SAR afectada por ruido speckle

*et al.*, 2001), siendo el modelamiento estadístico de particular relevancia cuando se trabaja con datos ruidosos. En toda imagen SAR se puede apreciar un cierto tipo de ruido denominado *speckle*, el que aparece cuando se emplea iluminación coherente, como es el caso de las tecnologías de microondas, sonar, láser y ultrasonido, entre otros; y se debe al fenómeno de interferencia constructiva y destructiva entre las señales incidente y reflejada. Este tipo de degradación deteriora severamente la habilidad humana, y la de cualquier máquina, de discriminar objetivos en la imagen. Por otra parte, es sabido que este tipo de ruido viola los supuestos clásicos de aditividad y gaussianidad. En la figura 1.2 se aprecia el efecto que tiene el ruido speckle sobre una imagen SAR.

Existen muchos modelos estadísticos para representar datos con ruido speckle (Oliver and Quegan, 1998). En este trabajo se estudia uno de los más utilizados, el denominado modelo multiplicativo, el cual declara bajo ciertas condiciones (Tur *et al.*, 1982) que el campo aleatorio observado  $Z$  es resultado del producto entre dos campos aleatorios independientes y no observados directamente:  $X$  e  $Y$ . El campo aleatorio  $X$  modela el terreno sensado (*backscatter*), entonces depende sólo del tipo de área a la cual pertenece cada píxel; mientras que el campo aleatorio  $Y$  describe el

ruido speckle, tomando en cuenta que  $n$  (idealmente) imágenes independientes (*looks*) son promediadas con el objeto de reducir el ruido (procesamiento *multilook*). Cabe recordar que de aquí en adelante la señal retrodispersada  $Z$  se asumirá en formato amplitud.

El fenómeno físico de generación del ruido speckle permite asumir una ley  $\Gamma^{1/2}$  para el campo  $Y$  (Goodman, 1985), mientras que hay varias formas de modelar el campo aleatorio  $X$ , entre ellas la más clásica resulta al considerar también el backscatter como una distribución  $\Gamma^{1/2}$ , parametrización que produce un modelo para el retorno  $Z$  correspondiente a la distribución  $K_A$  (Jakeman and Pusey, 1976; Jao, 1984; Oliver and Quegan, 1998; Blacknell, 1994). No obstante, este modelo falla en situaciones donde el retorno representa áreas extremadamente heterogéneas, siendo además una distribución difícil de trabajar analítica y computacionalmente. Por otra parte, en (Frery *et al.*, 1997a; Mejail *et al.*, 2001) se propuso la distribución  $\Gamma^{-1/2}$  para describir el backscatter de amplitud  $X$ , produciendo la distribución  $\mathcal{G}_A^0$  para el retorno. Una de las ventajas de la distribución  $\mathcal{G}_A^0$  sobre la clásica distribución  $K_A$  es que modela (ajusta) muy bien todos los tipos de terreno, independiente de su grado de homogeneidad; al igual que permite un mejor tratamiento analítico y computacional. Esta ley también ha sido utilizada para describir diferentes tipos de tejido en imágenes B-scan de ultrasonido (Navarrete *et al.*, 2002).

La distribución  $\mathcal{G}_A^0$  está caracterizada por tantos parámetros como la distribución  $K_A$ : el número de looks ( $n$ ), el parámetro de escala ( $\gamma$ ) y el parámetro de rugosidad ( $\alpha$ ). Ambas distribuciones permiten una adecuada interpretación de sus parámetros. El parámetro de rugosidad es de particular interés en algunas aplicaciones, puede ser utilizado como un indicador que permite discriminar diferentes tipos de cobertura terrestre. El parámetro de escala  $\gamma$  concierne a la potencia relativa entre las señales reflejada e incidente. En (Mejail, 1999) se verifica la factibilidad práctica y teórica de reemplazar la distribución  $K_A$  por la distribución  $\mathcal{G}_A^0$ , en el modelamiento de datos SAR.

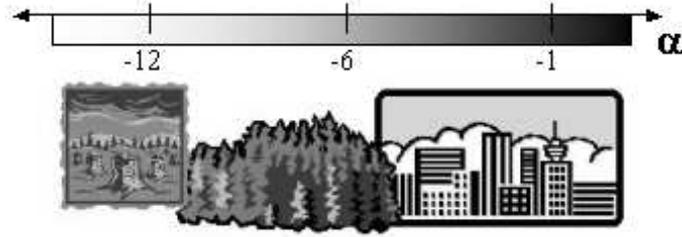


Figura 1.3: Interpretación del parámetro de rugosidad  $\alpha$

En la figura 1.3 se aprecia la interpretabilidad del parámetro de rugosidad  $\alpha$ . Aunque es difícil establecer límites bien definidos entre cada tipo de cobertura, es posible aseverar con base en estudios previos (Frery *et al.*, 1997a; Mejail, 1999; Bustos *et al.*, 2002) que, aproximadamente, en áreas homogéneas  $\alpha \in (-\infty, -10)$ , en áreas heterogéneas  $\alpha \in (-10, -4)$  y, en áreas extremadamente heterogéneas  $\alpha \in (-4, 0)$ .

En esta tesis se discute el problema de estimar los parámetros de la distribución  $\mathcal{G}_A^0$  para el caso de un look. Este es el caso en que la imagen se encuentra con el más alto contenido de ruido speckle y, por lo tanto, las imágenes con  $n = 1$  dependen mucho más de procedimientos de inferencia confiables, por ejemplo en filtraje (Frery *et al.*, 1997b; Jin *et al.*, 2003) y clasificación (Lucini *et al.*, 2003; Mejail *et al.*, 2003).

La robustez es una propiedad deseable de los estimadores, ya que permite su utilización incluso en situaciones donde la calidad de los datos es poco confiable. Una situación donde esto ocurre es cuando aparecen puntos de control terrestre (GCP: *Ground Control Points*) en la imagen SAR, los que son esenciales para la calibración de los sensores. Esos puntos producen un retorno muy alto respecto del resto de la imagen, reflejando gran parte de la energía incidente, razón por la cual se les denomina reflectores de esquina (*corner reflectors*). Si datos provenientes de corner reflectors se incluyen en un análisis con procedimientos de estimación no robustos, los resultados pueden ser completamente desastrosos. Otra típica situación se origina en la aplicación de filtros (Polzehl and Spokoiny, 2003), los datos se van procesando en ventanas, pero no existe forma de asegurar que ellos conforman una muestra libre de

contaminación.

Aunque no existe una definición formal, en la literatura se denomina *outliers* a las observaciones de una muestra que se alejan del grueso de los datos, como es el caso de los corner reflectors, constituyendo por lo tanto observaciones que no obedecen completamente los supuestos establecidos sobre la muestra. En el caso aquí presentado, los corner reflectors violan el supuesto de caracterizarse según la distribución  $\mathcal{G}_A^0$ .

Cuando la distribución es simétrica, los problemas debidos a la aparición de outliers pueden ser reducidos usando técnicas clásicas de estimación robusta (Huber, 1981; Hampel *et al.*, 1986; Rousseeuw and Verboven, 2002), las cuales tienden a ignorar o a ponderar levemente la influencia de aquellas observaciones a ambos lados de la media de la distribución.

No obstante lo anterior, es frecuente encontrar en muchas aplicaciones distribuciones de datos asimétricas con colas pesadas, ver por ejemplo (Nikias and Shao, 1995), como la distribución  $\mathcal{G}_A^0$ . Trabajar con estos datos es esencialmente difícil, puesto que las muestras provenientes de las colas de la distribución tienen una fuerte influencia en los procedimientos de estimación de parámetros, y la subponderación de dichas observaciones puede introducir sesgo en la estimación (Reed III and Stark, 1996).

El capítulo 2 de esta tesis examina las propiedades fundamentales del modelo aquí considerado, así como también las técnicas tradicionales de estimación de parámetros para la distribución  $\mathcal{G}_A^0$ . En el capítulo 3 se introduce la estimación robusta, presentando los clásicos M-estimadores, y los nuevos AM-estimadores propuestos en esta tesis. El capítulo 4 muestra un estudio de simulación comparativo entre los estimadores clásicos y la técnica propuesta. Finalmente, en el capítulo 5 se exponen las conclusiones de este trabajo y las extensiones futuras de éste.

# Capítulo 2

## El Modelo Multiplicativo

### 2.1. Introducción

El modelo multiplicativo es una estructura frecuentemente utilizada para explicar el comportamiento estocástico de datos obtenidos mediante iluminación coherente. Bajo este paradigma se asume que las observaciones (píxeles) de este tipo de imágenes son el resultado del producto de dos variables aleatorias independientes: una ( $X$ ) modela el terreno (*backscatter*), y la otra ( $Y$ ) modela el ruido *speckle* (Tur *et al.*, 1982; Goodman, 1985). Por lo tanto, el valor observado se modela como

$$Z = X \cdot Y. \tag{2.1}$$

En la ecuación (2.1), el backscatter se asume habitualmente real y positivo, mientras que el speckle se considera complejo. De esta forma, el retorno corresponde a una variable aleatoria compleja  $Z(r, s) = a(r, s) + ib(r, s)$ , donde el conjunto de imágenes  $a(r, s) = A(r, s) \cos \phi(r, s)$  y  $b(r, s) = A(r, s) \sin \phi(r, s)$  representan la proyección de la amplitud en dos planos perpendiculares. La fase de la señal percibida en la posición  $(r,s)$  se denota como  $\phi(r, s)$  y  $A(r, s)$  es su amplitud.

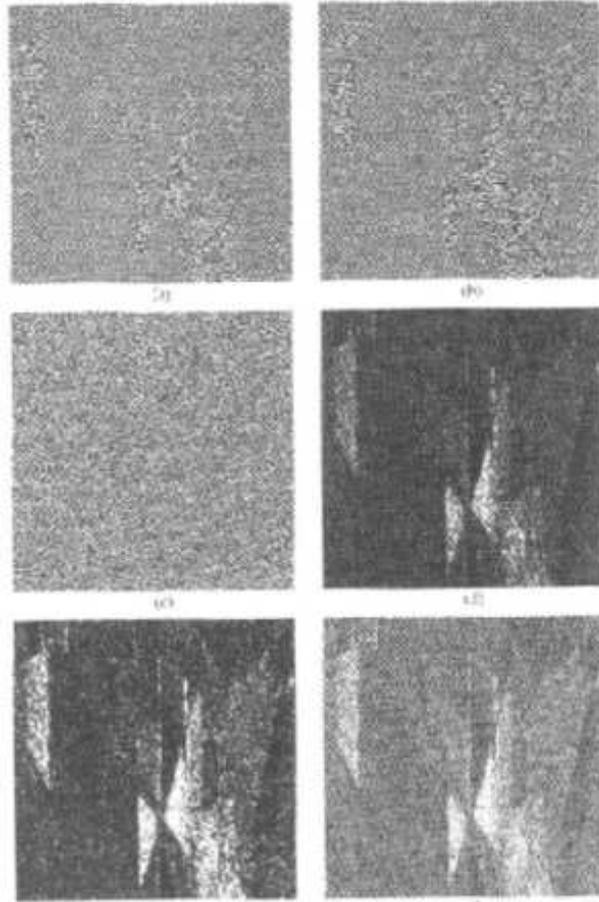


Figura 2.1: Diferentes representaciones de una imagen SAR: (a) componente real, (b) componente imaginaria, (c) fase, (d) amplitud, (e) intensidad y (f) log-intensidad.

A partir del formato complejo para representar la retrodispersión, es posible derivar otros formatos de imagen:  $Z_I = \|Z\|^2$  (intensidad),  $Z_A = \|Z\| = \sqrt{Z_I}$  (amplitud) y  $Z_\phi = \arctan(Im(Z)/Re(Z))$  (fase). En la figura 2.1 es posible notar a simple vista que el formato amplitud provee una mejor discriminación de la textura o rugosidad de la superficie, hecho por el cual el 90% de las aplicaciones utilizan datos bajo este formato. Por esta razón en esta tesis se considerará el modelo de amplitud para imágenes SAR.

## 2.2. Modelo

Bajo el modelo de amplitud multilook (Frery *et al.*, 1997a), el ruido speckle sigue una distribución raíz cuadrada de Gamma:  $Y_A \sim \Gamma^{1/2}(n, n)$ , caracterizada por la densidad

$$f_{Y_A}(y) = \frac{2n^n}{\Gamma(n)} y^{2n-1} \exp(-ny^2), \quad n \geq 1, y > 0, \quad (2.2)$$

donde el número de looks ( $n$ ) denota la cantidad de imágenes independientes que fueron registradas (en tiempo infinitesimal) sobre el terreno, tomando el promedio con el objeto de reducir el ruido; esta práctica se conoce con el nombre de *procesamiento multilook*.

El número de looks es un parámetro que permanece constante sobre toda la imagen y es controlado en la etapa de generación de una imagen SAR, por lo que en teoría debe corresponder a un número entero conocido de antemano; no obstante, esto rara vez ocurre debido a que la media fue tomada sobre observaciones correlacionadas. Es por ello que en la práctica debe ser estimado para obtener lo que se denomina el *número equivalente de looks* (Yanasse *et al.*, 1995).

Por su parte, en (Frery *et al.*, 1997a) se propuso la raíz cuadrada de la distribución Gaussiana inversa generalizada, como el modelo estadístico general para el backscatter de amplitud:  $X_A \sim \mathcal{N}^{-1/2}(\alpha, \gamma, \lambda)$ , distribución caracterizada por la densidad

$$f_{X_A}(x) = \frac{(\lambda/\gamma)^{\alpha/2}}{K_\alpha(2\sqrt{\gamma\lambda})} x^{2\alpha-1} \exp\left(-\frac{\gamma}{x^2} - \lambda x^2\right), \quad x > 0, \quad (2.3)$$

donde  $K_\alpha$  denota la función de Bessel modificada de tercer tipo y orden  $\alpha$  (Gradshteyn and Ryzhik, 1998) dada por

$$K_\alpha(2\sqrt{\gamma\lambda}) = \frac{1}{2} \left(\frac{\lambda}{\gamma}\right)^{\frac{\alpha}{2}} \int_0^\infty t^{\alpha-1} \exp\left(-\frac{\gamma}{t} - \lambda t\right) dt, \quad (2.4)$$

donde su espacio paramétrico está dado por

$$\left\{ \begin{array}{ll} \gamma > 0, \lambda \geq 0 & \text{si } \alpha < 0 \\ \gamma > 0, \lambda > 0 & \text{si } \alpha = 0 \\ \gamma \geq 0, \lambda > 0 & \text{si } \alpha > 0. \end{array} \right. \quad (2.5)$$

Esta distribución tiene dos casos particulares de gran interés en análisis de datos SAR: la raíz cuadrada de Gamma ( $\Gamma^{1/2}$ ) y la recíproca de una raíz cuadrada de Gamma ( $\Gamma^{-1/2}$ ). La primera provee un modelo bastante estudiado para superficies homogéneas y heterogéneas. La segunda, además de mantener las habilidades de la primera ley, permite modelar superficies de texturas extremadamente heterogéneas. Ambos casos, combinados con el modelo del ruido speckle para amplitud (2.2), producen la distribuciones  $K_A$  y  $\mathcal{G}_A^0$ , respectivamente, para el retorno de amplitud  $Z_A$ .

La distribución general para la retrodispersión de amplitud se origina entonces, mediante el producto  $Z_A = X_A \cdot Y_A$ . Si  $X_A \sim \mathcal{N}^{-1/2}(\alpha, \gamma, \lambda)$  e  $Y_A \sim \Gamma^{1/2}(n, n)$ , asumiendo ambas variables aleatorias independientes, se dice que el retorno sigue una distribución denominada *amplitud*  $\mathcal{G}$ :  $Z_A \sim \mathcal{G}_A(\alpha, \gamma, \lambda, n)$ , caracterizada por la densidad

$$f_{Z_A}(z) = \frac{2n^n(\lambda/\gamma)^{\alpha/2}}{\Gamma(n)K_\alpha(2\sqrt{\gamma\lambda})} z^{2n-1} \left( \frac{\gamma + nz^2}{\lambda} \right)^{\frac{\alpha-n}{2}} K_{\alpha-n} \left( 2\sqrt{\lambda(\gamma + nz^2)} \right), \quad (2.6)$$

donde  $n \geq 1$ ,  $z > 0$  y el espacio paramétrico de  $(\alpha, \gamma, \lambda)$  es el mismo definido en (2.5).

Una forma de describir la distribución es a través de sus momentos. Los momentos de orden  $r$  de la distribución  $\mathcal{G}_A$  están dados por

$$E[Z_A^r] = \left( \frac{\gamma}{n^2\lambda} \right)^{r/4} \frac{K_{\alpha+r/2}(2\sqrt{\gamma\lambda})\Gamma(n+r/2)}{K_\alpha(2\sqrt{\gamma\lambda})\Gamma(n)}, \quad r > 0. \quad (2.7)$$

Esta distribución para la respuesta en amplitud es bastante general. Sin embargo, es difícil obtener estimadores de sus parámetros debido a su complejidad analítica. Es por ello que se intenta trabajar con casos particulares de esta distribución:

- **Distribución  $K_A$ :** Si  $\alpha, \lambda > 0$  y  $\gamma \downarrow 0$ , la distribución  $\mathcal{G}_A(\alpha, \gamma, \lambda, n)$  converge en distribución a la distribución  $K_A(\alpha, \lambda, n)$ , la cual ha sido estudiada ampliamente en la literatura desde mediados de los años 70.
- **Distribución  $\mathcal{G}_A^0$ :** Si  $-\alpha, \gamma > 0$  y  $\lambda \downarrow 0$ , la distribución  $\mathcal{G}_A(\alpha, \gamma, \lambda, n)$  converge en distribución a la distribución  $\mathcal{G}_A^0(\alpha, \gamma, n)$ , cuyo estudio se ha desarrollado hasta hoy día desde sus orígenes a mediados de los años 90.

En (Frery *et al.*, 1997a; Mejail, 1999) se puede ver en detalle la convergencia de la distribución  $\mathcal{G}_A(\alpha, \gamma, \lambda, n)$  a sus casos particulares.

Como se mencionó anteriormente, la distribución  $\mathcal{G}_A^0$  tiene las mismas propiedades interpretativas de sus parámetros que la distribución  $K_A$ , pero además posee una mayor flexibilidad en el modelamiento de superficies con diferente grado de homogeneidad. Por otro lado, la ley  $\mathcal{G}_A^0$  es más tratable numérica y computacionalmente que la ley  $K_A$ , lo cual que la hace muy atractiva como alternativa general para el análisis de datos SAR. De hecho, esta distribución fue desarrollada originalmente para describir zonas extremadamente heterogéneas en (Frery *et al.*, 1997a), y más tarde fue propuesta y evaluada como un modelo universal para datos con ruido speckle en (Mejail *et al.*, 2001). La distribución  $\mathcal{G}_A^0$  constituye el tema central de esta tesis, la que se estudia con mayor detalle en la sección 2.3.

Por otra parte, es posible obtener la distribución del retorno en formato intensidad mediante la transformación  $Z_I = Z_A^2$ , o bien la distribución para el retorno en formato log-intensidad mediante la transformación  $Z_{\log I} = \log(Z_I)$ .

### 2.3. La Distribución $\mathcal{G}_A^0$

Sean  $X_A$  e  $Y_A$  dos variables aleatorias independientes tales que  $X_A \sim \Gamma^{-1/2}(\alpha, \gamma)$  e  $Y_A \sim \Gamma^{1/2}(n, n)$  con  $-\alpha, \gamma, n > 0$ , entonces la variable aleatoria producto  $Z_A = X_A \cdot Y_A$  tiene una distribución  $\mathcal{G}_A^0 : Z_A \sim \mathcal{G}_A^0(\alpha, \gamma, n)$  (Mejail, 1999), caracterizada por la densidad

$$f_{Z_A}(z) = \frac{2n^n \Gamma(n - \alpha)}{\gamma^\alpha \Gamma(n) \Gamma(-\alpha)} \frac{z^{2n-1}}{(\gamma + nz^2)^{n-\alpha}}, \quad z > 0, \quad (2.8)$$

donde  $-\alpha > 0$  se refiere al parámetro de rugosidad,  $\gamma > 0$  es el parámetro de escala y  $n \geq 1$  es el número de looks. Por su parte, la función de distribución acumulada de la variable aleatoria  $Z_A$  viene dada por

$$F_{Z_A}(z) = \frac{n^{n-1} \Gamma(n - \alpha)}{\gamma^n \Gamma(n) \Gamma(-\alpha)} z^{2n} H(n, n - \alpha; n + 1; -nz^2/\gamma), \quad (2.9)$$

donde  $-\alpha, \gamma > 0, n \geq 1, z > 0$  y  $H$  es la función hipergeométrica (Gradshteyn and Ryzhik, 1998) dada por

$$H(a, b; c; t) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)t^k}{\Gamma(c+k)k!}. \quad (2.10)$$

Es posible demostrar (Mejail, 1999) que la función de distribución acumulada indicada en la ecuación (2.9) también puede ser descrita según

$$F_{Z_A}(z) = \Upsilon_{2n, -2\alpha}(-\alpha z^2/\gamma), \quad -\alpha, \gamma > 0, n \geq 1, z > 0, \quad (2.11)$$

donde  $\Upsilon_{a,b}(S)$  es la función de distribución acumulada de una variable aleatoria  $S$  con distribución  $\mathcal{F}_{a,b}$  de Snedecor con  $a$  y  $b$  grados de libertad. Esta relación tiene especial importancia por tres razones:

1. De ser necesario evaluar la función de distribución acumulada de variables aleatorias  $\mathcal{G}_A^0$ , bastará con evaluar la función  $\Upsilon$  de la relación (2.11), la cual se encuentra disponible en prácticamente todas las plataformas computacionales.

2. Es posible obtener muestras de la distribución  $\mathcal{G}_A^0$  a partir de muestras con distribución  $\mathcal{F}_{2n, -2\alpha}$ , utilizando el método de la transformada inversa (Banks, 1998):

$$Z_A = \sqrt{-\gamma \Upsilon_{2n, -2\alpha}^{-1}(U)/\alpha}, \quad (2.12)$$

donde  $U$  sigue una ley uniforme en el intervalo  $(0, 1)$ , y  $\Upsilon_{2n, -2\alpha}^{-1}$  es la función de distribución acumulada de la ley  $\mathcal{F}_{2n, -2\alpha}$  de Snedecor inversa ( $\mathcal{F}_{2n, -2\alpha}^{-1}$ ), la que también se encuentra disponible computacionalmente.

3. Es posible calcular la función hipergeométrica de la ecuación (2.9) en función de la distribución acumulada de la ley  $\mathcal{F}_{2n, -2\alpha}$  de Snedecor:

$$H(n, n - \alpha; n + 1; -nz^2/\gamma) = \frac{\gamma^n \Gamma(n) \Gamma(-\alpha)}{n^{n-1} \Gamma(n - \alpha)} \Upsilon_{2n, -2\alpha}(-\alpha z^2/\gamma). \quad (2.13)$$

Los momentos de orden  $r$  de la distribución  $\mathcal{G}_A^0$  están dados por

$$E[Z_A^r] = \begin{cases} \left(\frac{\gamma}{n}\right)^{r/2} \frac{\Gamma(n+r/2)\Gamma(-\alpha-r/2)}{\Gamma(n)\Gamma(-\alpha)} & \text{si } -\alpha > r/2 \\ \infty & \text{si } 0 < -\alpha \leq r/2 \end{cases}, r > 0. \quad (2.14)$$

La media  $\mu_A$  de la distribución está dada por

$$\mu_A = E[Z_A] = \begin{cases} \sqrt{\frac{\gamma}{n}} \frac{\Gamma(n+1/2)\Gamma(-\alpha-1/2)}{\Gamma(n)\Gamma(-\alpha)} & \text{si } -\alpha > 1/2 \\ \infty & \text{si } 0 < -\alpha \leq 1/2. \end{cases} \quad (2.15)$$

La varianza  $\sigma_A^2 = E[Z_A^2] - E^2[Z_A]$  de la distribución está dada por

$$\sigma_A^2 = \begin{cases} \left(\frac{\gamma}{n}\right) \frac{[n\Gamma^2(n)(-\alpha-1)\Gamma^2(-\alpha-1) - \Gamma^2(n+1/2)\Gamma^2(-\alpha-1/2)]}{\Gamma^2(n)\Gamma^2(-\alpha)} & \text{si } -\alpha > 1 \\ \infty & \text{si } 0 < -\alpha \leq 1, \end{cases} \quad (2.16)$$

La función de densidad de la distribución  $\mathcal{G}_A^0$  es unimodal, y su moda está dada por

$$z_m = \sqrt{\frac{\gamma(2n-1)}{n(1-2\alpha)}}, \quad -\alpha, \gamma > 0, n \geq 1. \quad (2.17)$$

Otras propiedades de esta distribución pueden ser revisadas en (Mejail, 1999).

El caso de un look ( $n = 1$ ) es de particular interés, puesto que corresponde a imágenes a las cuales no se ha aplicado procesamiento multilook y, por lo tanto, describe aquellas imágenes con mayor contenido de ruido speckle, lo que hace más difícil el tratamiento digital y el análisis de este tipo de imágenes. Por tal motivo, el estudio de las propiedades y los estimadores de los parámetros de la distribución  $\mathcal{G}_A^0$ , considerando  $n = 1$ , constituye el núcleo de esta investigación.

La distribución de interés, que se denotará como  $\mathcal{G}_A^0(\alpha, \gamma) \equiv \mathcal{G}_A^0(\alpha, \gamma, 1)$ , está caracterizada por la densidad

$$f(z) = -\frac{2\alpha z}{\gamma^\alpha(\gamma + z^2)^{1-\alpha}} = -\frac{2\alpha z}{\gamma(1 + z^2/\gamma)^{1-\alpha}}, \quad -\alpha, \gamma, z > 0. \quad (2.18)$$

Su función de distribución acumulada está dada por

$$F(z) = 1 - (1 + z^2/\gamma)^\alpha, \quad -\alpha, \gamma, z > 0, \quad (2.19)$$

y su inversa, útil para la generación de muestras aleatorias y el cálculo de cuantiles, está dada por

$$Z = F^{-1}(U) = \sqrt{\gamma((1 - U)^{1/\alpha} - 1)}. \quad (2.20)$$

Similarmente, los momentos de orden  $r$  de la distribución  $\mathcal{G}_A^0(\alpha, \gamma)$ , están dados por

$$E[Z^r] = \begin{cases} \gamma^{r/2} \frac{r\Gamma(r/2)\Gamma(-\alpha-r/2)}{2\Gamma(-\alpha)} & \text{if } -\alpha > r/2 \\ \infty & \text{si } 0 < -\alpha \leq r/2 \end{cases}, r > 0. \quad (2.21)$$

Por su parte, los momentos de orden  $r$  basados en una realización  $\mathbf{z} = (z_1, z_2, \dots, z_N)$  de la variable aleatoria  $\mathbf{Z}$ , están dados por

$$m_r = \frac{1}{N} \sum_{i=1}^N z_i^r, \quad r > 0. \quad (2.22)$$

A continuación, en la sección 2.4 se presentan las técnicas clásicas de estimación de parámetros para esta distribución. En el capítulo 3 se exponen los estimadores robustos clásicos conocidos como M-estimadores y se proponen los nuevos AM-estimadores.

## 2.4. Métodos de Estimación Clásicos

Como se mencionó anteriormente, el modelo propuesto para el retorno de las imágenes SAR depende de tres parámetros: rugosidad, escala y número de looks. La variación de éstos dentro de un espacio de parámetros adecuadamente definido, indicará las diferentes clases de datos presentes en las imágenes.

Es evidente la importancia de estimar al parámetro  $\alpha$ , ya que permite estimar la rugosidad, que es una cantidad que discrimina diferentes áreas independientemente de la señal de incidencia. El parámetro de escala  $\gamma$  está asociado con la amplitud del backscatter y, sin pérdida de generalidad, se puede escoger  $\gamma$  para cada  $\alpha$  de interés, de tal forma que  $E[Z] = 1$ . De esta manera, es posible realizar el estudio de inferencia del parámetro  $\alpha$  suponiendo  $\gamma$  conocido. Por lo tanto, mediante la relación (2.14), la estimación del parámetro de escala podría obtenerse a partir de un estimador  $\hat{\alpha}$  del parámetro de rugosidad y la cantidad conocida de looks  $n$  según

$$\hat{\gamma}^* = n \left( \frac{\Gamma(-\hat{\alpha})\Gamma(n)}{\Gamma(-\hat{\alpha} - 1/2)\Gamma(n + 1/2)} \right)^2. \quad (2.23)$$

Varias técnicas de inferencia de parámetros están disponibles para estimar  $\theta = (\alpha, \gamma)$ , siendo las más importantes aquellas basadas en momentos muestrales y máxima verosimilitud. En las secciones siguientes se describen ambas metodologías, considerando el parámetro  $\gamma$  conocido o no.

Mejores estimaciones para los parámetros  $\alpha$  and  $\gamma$  de esta distribución fueron encontrados usando bootstrap (Cribari-Neto *et al.*, 2002), mientras que la robustez

para el caso  $n = 1$  es estudiada en (Bustos *et al.*, 2002; Allende and Pizarro, 2003) usando M-estimadores.

### 2.4.1. Estimación por Momentos

El método de los momentos es, quizás, el método más antiguo de estimación de parámetros y tiene la virtud de ser bastante simple. Éste consiste en igualar los momentos apropiados de la distribución de la población con sus correspondientes momentos muestrales. No obstante, sus estimaciones no garantizan consistencia y otras propiedades deseables en los estimadores (Kendall and Stuart, 1979; Casella and Berger, 1990).

Sea  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$  un vector de variables aleatorias independientes e idénticamente distribuidas con función de densidad de probabilidad  $f(\mathbf{z}|\theta)$ , parametrizada por  $\theta = (\theta_1, \dots, \theta_k)$ . Dada una realización  $\mathbf{z} = (z_1, z_2, \dots, z_N)$  del vector  $\mathbf{Z}$ , usando los momentos definidos por (2.21) y (2.22), el estimador de momentos de  $\theta$  es un valor  $\hat{\theta}_{MO} = (\hat{\theta}_1, \dots, \hat{\theta}_k)$  que se encuentra resolviendo el siguiente sistema de  $k$  ecuaciones simultáneas

$$E[Z^{r_i}] = m_{r_i}, \quad r_i > 0, i = 1, \dots, k. \quad (2.24)$$

### MO-estimadores de la Distribución $\mathcal{G}_A^0$

Para el caso multilook, asumiendo  $n$  conocido, la estimación conjunta de los parámetros  $\theta = (\alpha, \gamma)$  de la distribución  $\mathcal{G}_A^0$  puede obtenerse a partir de (2.24) utilizando los momentos de orden  $r = 1/2, 1$ . Resolviendo el siguiente sistema de ecuaciones

se encuentran los estimadores de momentos  $\hat{\theta}_{MO} = (\hat{\alpha}_{MO}, \hat{\gamma}_{MO})$

$$\left(\frac{\gamma}{n}\right)^{1/4} \frac{\Gamma(n+1/4)\Gamma(-\alpha-1/4)}{\Gamma(n)\Gamma(-\alpha)} = m_{1/2} \quad (2.25)$$

$$\left(\frac{\gamma}{n}\right)^{1/2} \frac{\Gamma(n+1/2)\Gamma(-\alpha-1/2)}{\Gamma(n)\Gamma(-\alpha)} = m_1, \quad (2.26)$$

asumiendo que  $-\alpha > 1/2$ . Es posible llevar este sistema a una expresión en una sola incógnita,  $\alpha$ , haciendo el cociente entre la ecuación (2.26) y el cuadrado de la ecuación (2.25), de manera que el estimador de la rugosidad  $\hat{\alpha}_{MO}$  se obtiene resolviendo la ecuación (2.27) por métodos numéricos (Burden and Faires, 1985; Nieves and Domínguez, 1995)

$$\frac{\Gamma^2(-\alpha-1/4)\Gamma^2(n+1/4)}{\Gamma(-\alpha-1/2)\Gamma(n+1/2)\Gamma(-\alpha)\Gamma(n)} - \frac{m_{1/2}^2}{m_1} = 0, \quad (2.27)$$

para luego estimar la escala por medio de (2.25) ó bien (2.26). Bajo esta última relación, el estimador de momentos de la escala  $\hat{\gamma}_{MO}$  está dado por

$$\hat{\gamma}_{MO} = n \left( \frac{m_1 \Gamma(-\hat{\alpha}_{MO}) \Gamma(n)}{\Gamma(-\hat{\alpha}_{MO} - 1/2) \Gamma(n + 1/2)} \right)^2. \quad (2.28)$$

Por otra parte, considerando el parámetro de escala conocido según (2.23), la estimación del parámetro de rugosidad puede efectuarse mediante la expresión (2.25).

### 2.4.2. Estimación por Máxima Verosimilitud

Este es uno de los métodos de estimación más utilizados en la literatura (Kendall and Stuart, 1979; Casella and Berger, 1990), debido a que los estimadores obtenidos mediante esta metodología poseen importantes propiedades estadísticas: consistencia, normalidad asintótica, eficiencia asintótica e invariancia; sin embargo, los estimadores máximo verosímiles suelen ser sesgados. El método consiste en encontrar los parámetros desconocidos a partir de una muestra dada tal que se maximice la probabilidad de que ésta se distribuya según la ley supuesta.

Sea  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$  un vector de variables aleatorias independientes e idénticamente distribuidas con función de densidad de probabilidad  $f(\mathbf{z}|\theta)$ , parametrizada por  $\theta = (\theta_1, \dots, \theta_k) \in \Theta$ , donde  $\Theta$  es el espacio paramétrico. Dada una realización  $\mathbf{z} = (z_1, z_2, \dots, z_N)$  del vector  $\mathbf{Z}$ , se define la función *verosimilitud* como

$$L(\theta|\mathbf{z}) = \prod_{i=1}^N f(z_i|\theta). \quad (2.29)$$

El estimador de máxima verosimilitud (ML: *maximum likelihood*) de  $\theta$  es un valor  $\hat{\theta}_{ML}$  que maximiza la ecuación (2.29) sobre  $\mathbf{z}$ , formalmente

$$\hat{\theta}_{ML} \equiv \hat{\theta}_{ML}(\mathbf{z}) = \arg \max_{\theta \in \Theta} L(\theta|\mathbf{z}). \quad (2.30)$$

De forma equivalente, y con el objeto de simplificar la aritmética, se efectúa la optimización sobre la función *log-verosimilitud* definida por

$$\begin{aligned} \ell(\theta|\mathbf{z}) &= \log L(\theta|\mathbf{z}) \\ &= \log \prod_{i=1}^N f(z_i|\theta) \\ &= \sum_{i=1}^N \log f(z_i|\theta). \end{aligned} \quad (2.31)$$

Luego, la ecuación (2.30) tiene su analogía como

$$\hat{\theta}_{ML} \equiv \hat{\theta}_{ML}(\mathbf{z}) = \arg \max_{\theta \in \Theta} \ell(\theta|\mathbf{z}). \quad (2.32)$$

Finalmente, el estimador de máxima verosimilitud  $\hat{\theta}_{ML}$  es solución del siguiente sistema de  $k$  ecuaciones simultáneas

$$\frac{\partial}{\partial \theta_i} \ell(\theta|\mathbf{z}) = 0, \quad i = 1, \dots, k. \quad (2.33)$$

**ML-estimadores de la Distribución  $\mathcal{G}_A^0$** 

De la función de densidad de probabilidad de la distribución  $\mathcal{G}_A^0$  multilook definida en (2.8) se tiene que

$$\begin{aligned} \ln f(z) &= \ln \frac{2n^n}{\Gamma(n)} + \ln \Gamma(n - \alpha) - \ln \Gamma(-\alpha) - \alpha \ln \gamma + \\ &+ (2n - 1) \ln z - (n - \alpha) \ln(\gamma + nz^2). \end{aligned} \quad (2.34)$$

Derivando esta expresión con respecto al parámetro  $\alpha$  se obtiene

$$\begin{aligned} \frac{\partial}{\partial \alpha} \ln f(z) &= -\Psi(n - \alpha) + \Psi(-\alpha) - \ln \gamma + \ln(\gamma + nz^2) \\ &= \Psi(-\alpha) - \Psi(n - \alpha) + \ln \left( 1 + \frac{n}{\gamma} z^2 \right), \end{aligned} \quad (2.35)$$

donde  $\Psi(t) = \frac{\partial}{\partial t} \ln \Gamma(t)$  es la función *Psi* o *Digamma*. Por otro lado, derivando la relación (2.34) con respecto al parámetro  $\gamma$ , se obtiene

$$\frac{\partial}{\partial \gamma} \ln f(z) = \frac{-\alpha}{\gamma} - \frac{n - \alpha}{\gamma + nz^2}. \quad (2.36)$$

Luego, para el caso multilook ( $n$  conocido), la estimación conjunta de máxima verosimilitud  $\hat{\theta}_{ML} = (\hat{\alpha}_{ML}, \hat{\gamma}_{ML})$  de los parámetros  $\theta = (\alpha, \gamma)$  de la distribución  $\mathcal{G}_A^0$  son solución del siguiente sistema de ecuaciones

$$\sum_{i=1}^N s(z_i; \theta) = \mathbf{0}, \quad (2.37)$$

donde  $s(z; \theta)$  es la función *score* determinada por

$$s(z; \theta) = \begin{pmatrix} s_1(z; \theta) \\ s_2(z; \theta) \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \alpha} \ln f(z) \\ \frac{\partial}{\partial \gamma} \ln f(z) \end{pmatrix}. \quad (2.38)$$

El sistema de ecuaciones (2.37) también debe ser resuelto por métodos numéricos (Burden and Faires, 1985; Nieves and Domínguez, 1995). En forma explícita, para encontrar el estimador de máxima verosimilitud  $\hat{\theta}_{ML}$  a partir de una realización  $\mathbf{z} =$

$(z_1, z_2, \dots, z_N)$  de la variable aleatoria  $\mathbf{Z}$ , se debe resolver

$$\begin{aligned} N[\Psi(-\alpha) - \Psi(n - \alpha)] + \sum_{i=1}^N \ln \left( 1 + \frac{n}{\gamma} z_i^2 \right) &= 0 \\ \frac{-N\alpha}{\gamma} - (n - \alpha) \sum_{i=1}^N \frac{1}{\gamma + n z_i^2} &= 0. \end{aligned} \quad (2.39)$$

Para el caso de interés,  $n = 1$ , la función score queda determinada por

$$s(z; \theta) = \begin{pmatrix} \frac{1}{\alpha} + \ln \left( 1 + \frac{z^2}{\gamma} \right) \\ \frac{-\alpha}{\gamma} - \frac{1-\alpha}{\gamma+z^2} \end{pmatrix}, \quad (2.40)$$

la cual produce el estimador de máxima verosimilitud  $\hat{\theta}_{ML} = (\hat{\alpha}_{ML}, \hat{\gamma}_{ML})$

$$\begin{aligned} \hat{\alpha}_{ML} &= - \left( \frac{1}{N} \sum_{i=1}^N \ln \left( 1 + \frac{z_i^2}{\hat{\gamma}_{ML}} \right) \right) \\ \hat{\gamma}_{ML} &= \left[ \left( 1 + \frac{1}{N} \sum_{i=1}^N \ln \left( 1 + \frac{z_i^2}{\hat{\gamma}_{ML}} \right) \right) \frac{1}{N} \sum_{i=1}^N (\hat{\gamma}_{ML} + z_i^2)^{-1} \right]^{-1}. \end{aligned} \quad (2.41)$$

En (Oliveira, 2002) se estudia los aspectos numéricos de la inferencia por máxima verosimilitud para la distribución  $\mathcal{G}_A^0$ .

Por otra parte, los estimadores de máxima verosimilitud de la distribución  $K_A$  pueden encontrarse en (Joughin *et al.*, 1993).

## Capítulo 3

# Métodos de Estimación Robustos

La estimación robusta ha llegado a ser más predominante en teledetección con la emergencia de una nueva generación de sensores. Mientras la nueva tecnología de sensores provee una más alta resolución espectral y espacial, permitiendo identificar un número más grande de clases espectralmente separables; etiquetar las muestras para el diseño de clasificadores sigue siendo una tarea difícil y costosa.

La presencia de outliers en los datos no es poco común en aplicaciones prácticas (Andrews *et al.*, 1972; Huber, 1981; Hampel *et al.*, 1986; Rousseeuw and Leroy, 1987). En teledetección, una escena usualmente contiene pixeles de origen desconocido, por ejemplo, en áreas agrícolas y boscosas. La distribución estadística de esos pixeles puede ser significativamente diferente de las clases de entrenamiento, y pueden constituir outliers estadísticos. Desafortunadamente, esos pixeles anómalos están frecuentemente esparcidos a través de la escena y son pequeños en número y, por lo tanto, identificarlos podría ser una tarea tediosa. Un enfoque común para eliminar el efecto de esos pixeles es usar técnicas robustas (Frery *et al.*, 1997b; Lucini *et al.*, 2003).

### 3.1. M-Estimadores

Si se considera el problema de estimación de la media incondicional desde una muestra finita de una distribución de colas pesadas, el promedio empírico puede ser un pobre estimador en este caso, ya que sólo bastan unas pocas observaciones provenientes de las colas para introducir gran variabilidad en el promedio. En el caso de las distribuciones simétricas, es posible subponderar o ignorar el efecto de esos outliers con el objeto de reducir dicha variabilidad.

Uno de los mejores estimadores robustos conocidos son los denominados M-estimadores, los cuales fueron propuestos por (Huber, 1964). Pertenecen a una clase de estimadores de influencia acotada, siendo una generalización de los estimadores de máxima verosimilitud.

Considere el vector  $\mathbf{z} = (z_1, z_2, \dots, z_N)$  de  $N$  muestras independientes de la misma distribución simétrica con función de densidad de probabilidad  $f(\mathbf{z}|\theta)$ , parametrizada por  $\theta = (\theta_1, \dots, \theta_k) \in \Theta$ , donde  $\Theta$  es el espacio paramétrico. Se define el M-estimador  $\hat{\theta}_M$  como el mínimo global de la función de energía  $E(\theta)$

$$\hat{\theta}_M = \arg \min_{\theta \in \Theta} E(\theta), \quad (3.1)$$

donde la función de energía es definida en términos de una función de pérdida  $\rho$  como

$$E(\theta) = \sum_{i=1}^N \rho(z_i; \theta). \quad (3.2)$$

Equivalentemente, es posible encontrar el M-estimador resolviendo la ecuación

$$\sum_{i=1}^N \psi(z_i; \theta) = 0, \quad (3.3)$$

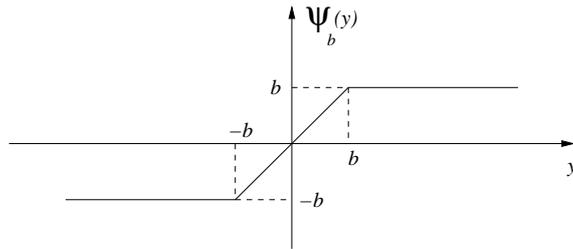
donde  $\psi(z_i; \theta) = \partial \rho(z_i; \theta) / \partial \theta$  se denomina *función de robustificación*.

Típicamente la función  $\rho$  es escogida simétrica; algunos casos particulares son:  $\rho(y) = y^2/2$ , que da origen al estimador de Mínimos Cuadrados (*LS: Least Square*),

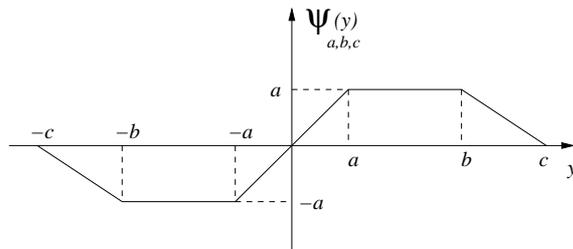
mientras que  $\rho(y) = |y|$  origina el estimador de mediana. La ecuación (3.3) es una generalización de la ecuación (2.37), ya que los ML-estimadores se obtienen tras considerar  $\rho(z; \theta) = -\ln f(z|\theta)$  y  $\psi(z; \theta) = s(z; \theta)$ .

Las funciones de robustificación  $\psi$  se construyen como composición entre las funciones score y ciertas funciones simétricas acotadas, donde estas últimas son usualmente definidas por partes, por ejemplo según sus autores

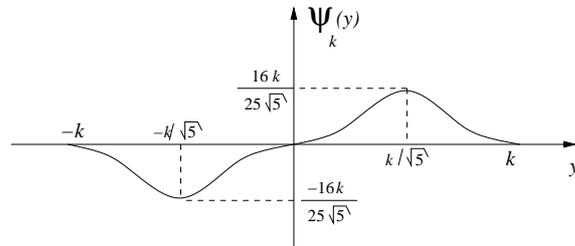
$$\text{Huber: } \psi_b(y) = \min\{b, \max\{y, -b\}\} = \begin{cases} -b & , y \leq -b \\ y & , -b < y < b \\ b & , y \geq b \end{cases} \quad (3.4)$$



$$\text{Hampel: } \psi_{a,b,c}(y) = \begin{cases} y & , 0 \leq |y| < a \\ a \cdot \text{sign}(y) & , a \leq |y| < b \\ a \frac{c-|y|}{c-b} \text{sign}(y) & , b \leq |y| < c \\ 0 & , c \leq |y| \end{cases} \quad (3.5)$$



$$\text{Tuckey: } \psi_k(y) = \begin{cases} y \left(1 - \left(\frac{y}{k}\right)^2\right)^2, & |y| \leq k \\ 0, & |y| > k \end{cases} \quad (3.6)$$



Los valores  $a, b, c, k$  se denominan parámetros de ajuste, y se obtienen de manera tal que la *eficiencia relativa asintótica* (ARE: *asymptotic relative efficiency*) del M-estimador con respecto al ML-estimador se encuentre, por ejemplo, en el rango 90 % a 95 %.

La importancia de las funciones  $\psi$  es que ellas truncan la función *score* para aquellas observaciones que se encuentran por debajo o por encima de ciertos valores umbrales. De esta forma, las funciones de robustificación permiten reducir la influencia de los datos atípicos (outliers) en el proceso de estimación.

Con el objeto de obtener estimadores robustos asintóticamente insesgados, en la ecuación (3.3) se introducen funciones correctoras  $c : \Theta \times (0, +\infty) \rightarrow \mathbb{R}$ . De esta forma, se redefine el M-estimador  $\hat{\theta}_M$  como solución de la ecuación

$$\sum_{i=1}^N \psi[s(z_i; \theta) - c] = 0. \quad (3.7)$$

Las funciones correctoras  $c$  deben cumplir con la consistencia de Fisher, por lo que se definen implícitamente como

$$\int_{-\infty}^{\infty} \psi[s(z_i; \theta) - c] dF_{\theta}(z) = 0. \quad (3.8)$$

Muchos resultados teóricos concernientes a las propiedades asintóticas y de robustez de los M-estimadores pueden encontrarse en (Andrews *et al.*, 1972; Huber, 1981; Hampel *et al.*, 1986; Rousseeuw and Leroy, 1987).

## 3.2. AM-Estimadores

Una manera cualitativa de describir la robustez de los estimadores es por medio de la *función de influencia empírica* (EIF: *empirical influence function*) (Hampel *et al.*, 1986). La EIF muestra lo que sucede con un estimador  $T_N$  cuando una observación  $z$  se mueve sobre el soporte de la distribución. Se define como  $\text{EIF}(z) = T_N(z_1, z_2, \dots, z_{N-1}, z)$ . Con el objeto de hacer el valor de  $\text{EIF}(z)$  independiente de una muestra en particular, se recurre a la denominada *función de influencia empírica estilizada* (SEIF: *stylized empirical influence function*) propuesta en (Andrews *et al.*, 1972), que consiste en usar el  $i$ -ésimo cuantil de la distribución subyacente

$$z_i = F^{-1}\left(\frac{i - 1/3}{N + 1/3}\right). \quad (3.9)$$

Cuando se trabaja con distribuciones simétricas sus SEIF también son simétricas, y es precisamente debido a esta razón que típicamente se escogen funciones de robustificación  $\psi$  simétricas (Rousseeuw and Verboven, 2002). En el presente caso, la distribución  $\mathcal{G}_A^0$  es no simétrica, tal como se aprecia en la SEIF de su ML-estimador en la figura 3.1 para muestras de tamaño variable.

En la figura 3.1 se observa que la pérdida de robustez, esto es, el grado en que el estimador  $\hat{\alpha}$  se va alejando del verdadero valor  $\alpha$ , depende del tamaño del outlier  $z$  y del tipo de área (recordar la figura 1.3). En áreas homogéneas ( $\alpha = -15$ , abajo-derecha) la pérdida de robustez se hace más crítica para grandes valores de outliers, mientras que en áreas extremadamente heterogéneas ( $\alpha = -1$ , arriba-izquierda) esto ocurre para valores pequeños de outliers. En cuanto a las áreas heterogéneas, la pérdi-

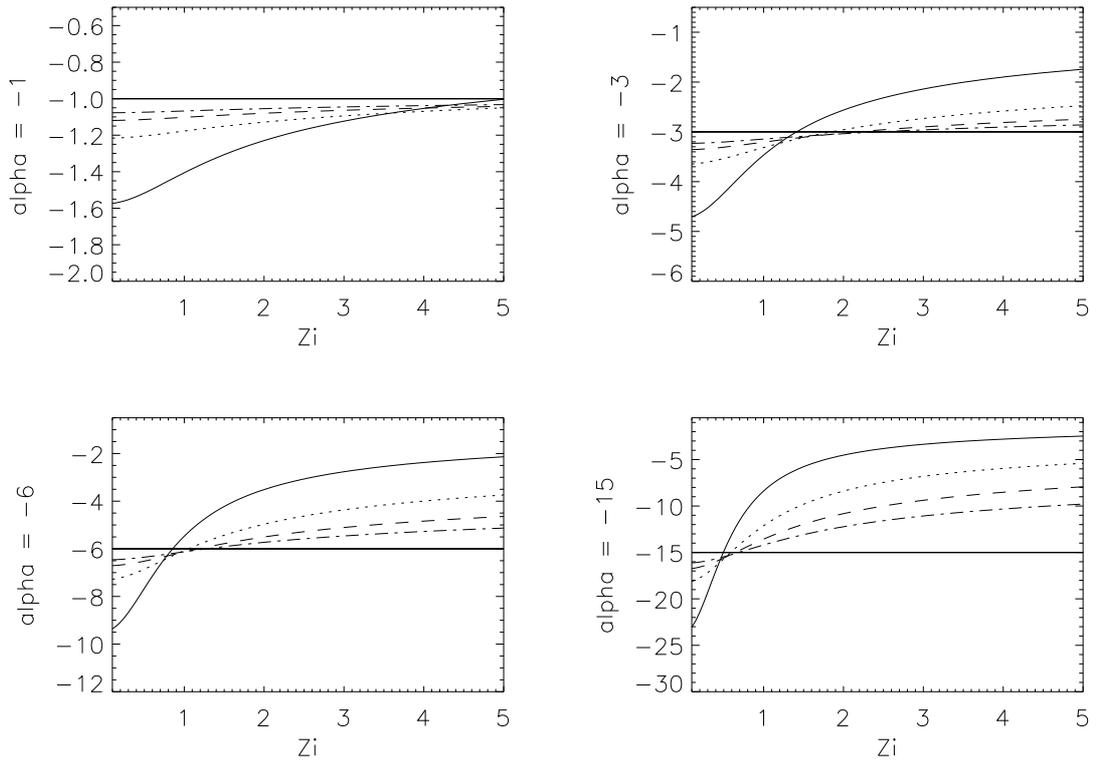


Figura 3.1: SEIF para el estimador de máxima verosimilitud bajo la distribución  $\mathcal{G}_A^0(\alpha, 1)$ , con  $-\alpha \in \{1, 3, 6, 15\}$  para  $N = 9$  (línea sólida),  $N = 25$  (línea punteada),  $N = 49$  (línea segmentada) and  $N = 81$  (línea punteada-segmentada).

da de robustez oscila entre ambos extremos. Por otro lado es importante notar que, independientemente de la rugosidad, si el tamaño de la muestra decrece la asimetría de la SEIF se hace más pronunciada.

Con base en los argumentos precedentes, se propone enfrentar el proceso de inferencia robusta en distribuciones no simétricas a través del uso de una familia  $\Psi_{r_1, r_2}$  de funciones de robustificación redescendentes asimétricas, donde  $0 < r_1, r_2 < \infty$  son los parámetros de ajuste. Toda función asimétrica  $\psi_{r_1, r_2} \in \Psi_{r_1, r_2}$  que pertenezca a esta familia será nula fuera del intervalo  $[r_1, r_2]$ .

Una función que pertenece a esta familia, ya sea para  $r_1 \leq r_2$  ó  $r_1 \geq r_2$ , está dada por

$$\psi_{r_1, r_2}(y) = \begin{cases} 0 & , y < -r_1 \\ -y - r_1 & , -r_1 \leq y < -r_1/2 \\ y & , -r_1/2 \leq y < r_2/2 \\ -y + r_2 & , r_2/2 \leq y < r_2 \\ 0 & , y \geq r_2 \end{cases} \quad (3.10)$$

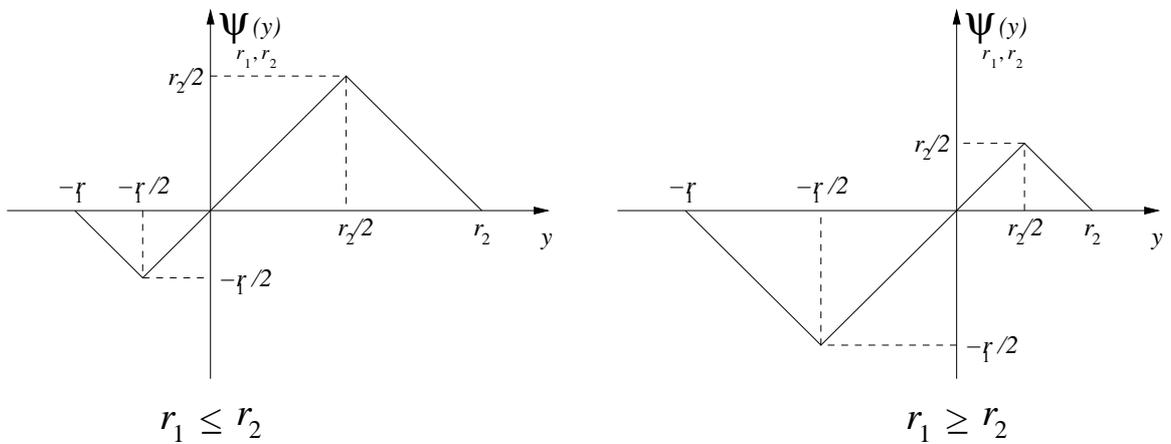


Figura 3.2: Función de robustificación redescendente asimétrica lineal definida por partes.

Por lo tanto, un AM-estimador es definido como en la ecuación (3.7), pero utilizando una función asimétrica del tipo  $\psi_{r_1, r_2} \in \Psi_{r_1, r_2}$ , como la presentada en la figura 3.2. Como ya se ha mencionado, los parámetros de ajuste  $r_1$  y  $r_2$  son escogidos según el criterio de eficiencia descrito en el apéndice B.

# Capítulo 4

## Estudio de Simulación

Con el objeto de evaluar el comportamiento de los estimadores robustificados con la familia de funciones  $\Psi_{r_1, r_2}$  (AM-estimadores), se considerará un modelo sin contaminación y varios modelos con outliers. El patrón de contaminación es definido como una secuencia de variables aleatorias independientes e idénticamente distribuidas  $z_1, z_2, \dots, z_N$ , con función de distribución común dada por

$$F(z; (\alpha, \gamma); \epsilon; \tau_v) = (1 - \epsilon) F(z; (\alpha, \gamma)) + \epsilon \delta_{\tau_v}(z), \quad (4.1)$$

donde  $\delta_{\tau_v}(z) = 1_{[\tau_v; +\infty)}(z)$  con  $\tau_v$  un valor muy grande comparado con la mayoría de los datos de la muestra, considerado como un factor de la media muestral:  $\tau_v = vE[Z]$ .  $\epsilon \in [0, 1]$  es la probabilidad de que una observación sea un outlier. Por lo tanto, en una muestra de  $N$  datos habrá en promedio  $(1 - \epsilon)N$  observaciones con distribución  $\mathcal{G}_A^0$  y  $\epsilon N$  outliers.

Las muestras fueron generadas mediante la relación (2.20), y sin pérdida de generalidad, se escogió el parámetro de escala  $\gamma$  en términos de  $\alpha$ , de manera tal que  $E[Z] = 1$ . Usando la ecuación (2.21) se tiene

$$\gamma_\alpha = \frac{4}{\pi} \left[ \frac{\Gamma(-\alpha)}{\Gamma(-\alpha - 1/2)} \right]^2. \quad (4.2)$$

A continuación se estima el parámetro de rugosidad  $\alpha$  de la distribución  $\mathcal{G}_A^0(\alpha, \gamma_\alpha, 1)$ , recordando que este es el caso que contiene mayor ruido speckle. Se comparan el AM-estimador, usando la función de robustez descrita en (3.10), con respecto al ML-estimador y al M-estimador presentado en (Bustos *et al.*, 2002), este último basado en la función de Huber (3.4). El ML-estimador es calculado según (2.41), mientras que el M-estimador y los AM-estimador se infieren a partir de (3.7), cuya solución se obtiene con la estrategia empleada por (Bustos and Ling, 1996), detallada en el apéndice A.

Por simplicidad, los parámetros de ajuste  $r_1$  y  $r_2$  del AM-estimador se vinculan entre sí, y se relacionan a su vez con el parámetro de ajuste  $b$  de la función de Huber. Esta correspondencia puede efectuarse bajo dos escenarios posibles:

$$\text{Caso (i) : } r_1 = 2b, r_2 = \eta r_1,$$

$$\text{Caso (ii) : } r_2 = 2b, r_1 = \eta r_2,$$

considerando  $b = \{1/2, 1, 2, 3, 4\}$  y  $\eta = \{1/2, 1, 3/2, 2\}$ , donde  $\eta$  es usado para controlar la amplitud de la asimetría de la función de robustificación (3.10).

La regla para escoger los parámetros de ajuste  $b$  y consecuentemente  $r_1$  y  $r_2$ , consiste en requerir que la eficiencia relativa asintótica (ARE) del AM-estimador con respecto al ML-estimador, considerando el modelo sin outliers, varíe en el rango  $90\% \leq \text{ARE} \leq 95\%$ . No obstante, pueden encontrarse varias tuplas  $r = (r_1, r_2)$  con la misma eficiencia relativa asintótica, por lo tanto, esta regla no determina unicidad. Para remediar este problema es posible escoger la tupla que minimice la aproximación de la máxima varianza asintótica como función de  $r$ . En el apéndice B se pueden observar estos cálculos.

En las figuras 4.1 y 4.2 se compara gráficamente el M-estimador y el AM-estimador bajo todas las posibles parametrizaciones indicadas anteriormente. En las estimaciones se consideró el modelo de contaminación (4.1), variando la probabilidad y la

magnitud de la contaminación (outliers),  $\epsilon = \{0,01, 0,05, 0,10\}$  y  $v = \{5, 10, 20, 40\}$ , respectivamente.

El primer hecho que se aprecia en las gráficas es que los AM-estimadores de tipo (ii), es decir, aquellos con parametrización  $r_2 = 2b$  y  $r_1 = \eta r_2$ , exhiben similar comportamiento en todos los casos, indicando que para  $\psi_{r_1, r_2}$  en  $\mathbb{R}^-$  sería posible trabajar con una única parametrización para todo  $\alpha$ , por ejemplo  $\eta = 1$ . Esto puede apreciarse también en el comportamiento de la SEIF del ML-estimador (figura 3.1): su bias negativo es pequeño y la pérdida de robustez se vuelve crítica para valores grandes de outliers. Por lo tanto, es suficiente estudiar el comportamiento del AM-estimador de tipo (i), aquellos con parametrización  $r_1 = 2b$  y  $r_2 = \eta r_1$ .

En las figuras 4.1 y 4.2 se observa también que a medida que los datos se hacen más homogéneos el M-estimador de Huber pierde precisión, esto es, en áreas homogéneas ( $\alpha = -10$ ) es menos preciso que en áreas extremadamente heterogéneas ( $\alpha = -1$ ), llegando a ser mucho peor a medida que la proporción y magnitud de los outliers crece, lo que ocurre debido a la influencia constante que la función de Huber (3.4) asigna a los valores extremos. Por otro lado, los AM-estimadores alcanzan buena precisión independientemente de la homogeneidad de los datos, especialmente cuando la proporción y magnitud de la contaminación aumenta, debido a la influencia decreciente que la función de robustez  $\psi_{r_1, r_2}$  asigna a los valores extremos.

En cuanto a la asimetría de la función  $\psi_{r_1, r_2}$ , se observa que ésta produce mejores resultados para áreas homogéneas cuando  $\eta = \{1/2, 1\}$ , mientras que para áreas extremadamente heterogéneas esto ocurre cuando  $\eta = \{3/2, 2\}$ , y en áreas heterogéneas todas las parametrizaciones presentan un comportamiento similar. En este punto sería interesante relacionar las propiedades de la distribución  $\mathcal{G}_A^0$  con la asimetría de la función de robustez, con el objeto de encontrar la parametrización más adecuada para cada tipo de área; en este sentido se podría recurrir a los momentos de orden  $r$ , u otras medidas robustas de tendencia central y dispersión, lo cual puede ser objeto de

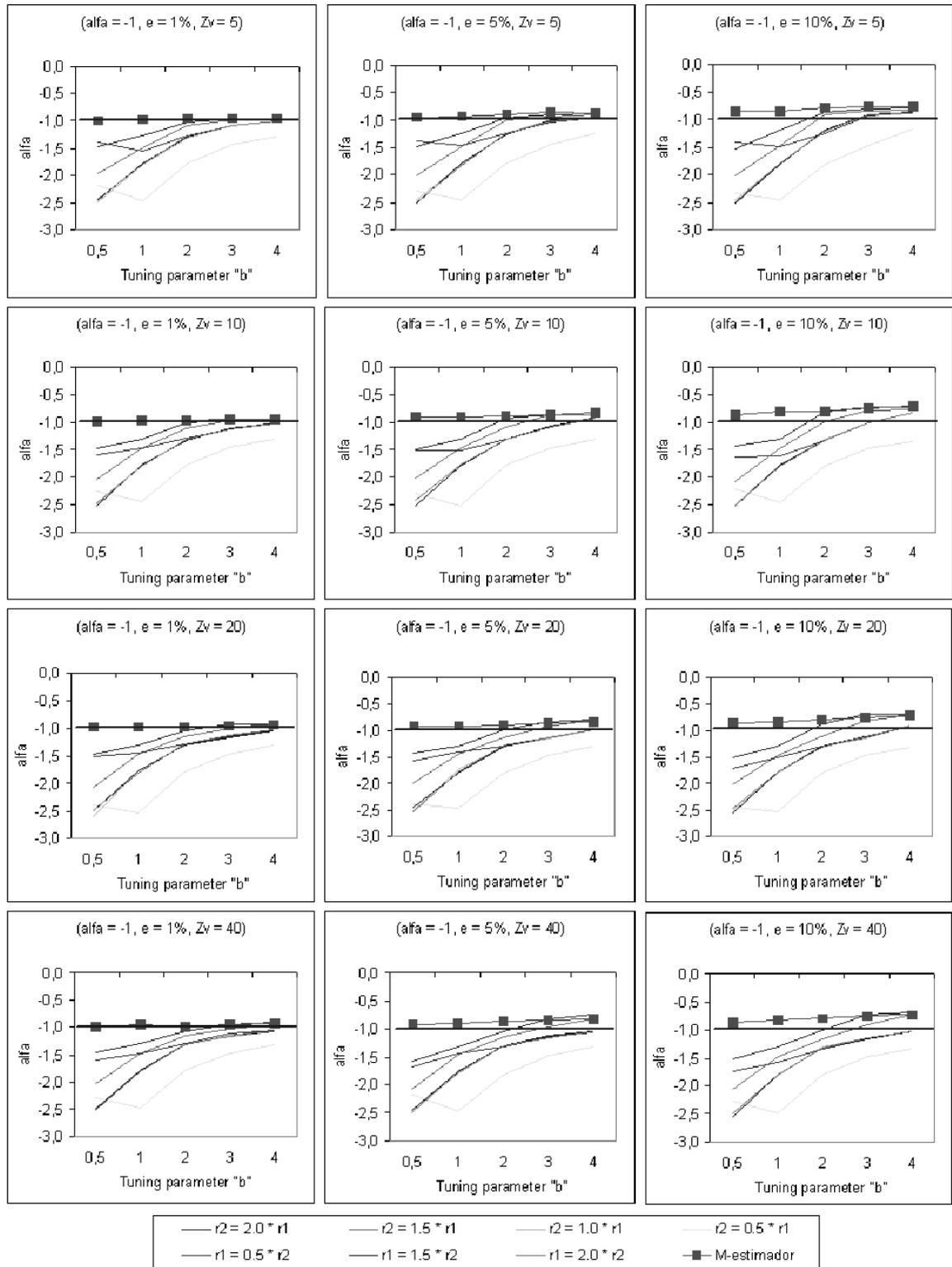


Figura 4.1: M-estimadores versus AM-estimadores para  $\alpha = -1$ .

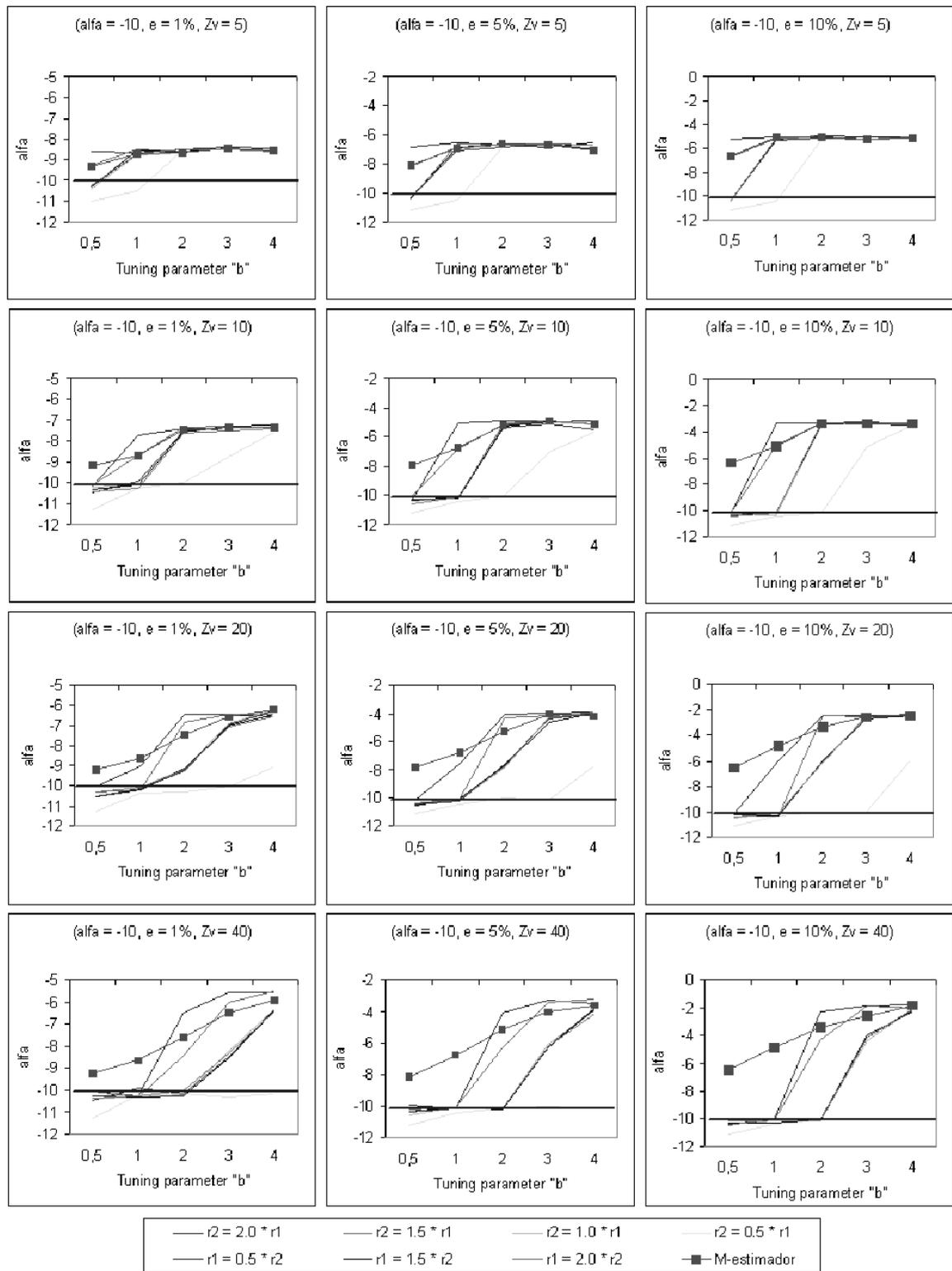


Figura 4.2: M-estimadores versus AM-estimadores para  $\alpha = -10$ .

estudios ulteriores.

Los estimadores son comparados por medio de una experiencia de Monte Carlo usando  $R = 1000$  replicaciones, tomando las siguientes medidas: media del estimador  $\widehat{E}[\hat{\alpha}] = R^{-1} \sum_{i=1}^R \hat{\alpha}_i$ , error cuadrático medio del estimador  $\widehat{MSE} = \widehat{E}[\hat{\alpha} - \alpha]^2 = \widehat{V}(\hat{\alpha}) + (\widehat{E}[\hat{\alpha}] - \alpha)^2$  y el sesgo relativo absoluto del estimador  $\widehat{B}[\hat{\alpha}] = \alpha^{-1} |\widehat{E}[\hat{\alpha}] - \alpha|$ , donde  $\alpha$  es el verdadero valor del parámetro y  $\hat{\alpha}$  es su estimador.

Las tablas 4.1, 4.2, 4.3, 4.4, 4.5 y 4.6 muestran una comparación cuantitativa entre los estimadores ML, M y AM, basada en dichas medidas. El estudio se efectúa para cada tipo de área: homogénea ( $\alpha = -10$ ), heterogénea ( $\alpha = -6$ ) y extremadamente heterogénea ( $\alpha = -1$ ), y bajo diferentes escenarios, variando el tamaño de la muestra  $N = \{9, 25, 49, 81\}$ , la probabilidad y magnitud de los outliers  $\epsilon = \{0,01, 0,05, 0,10\}$  y  $v = \{5, 15\}$ .

Los resultados en las tablas muestran que todos los estimadores exhiben casi el mismo comportamiento cuando la muestra está libre de contaminación. Además, cuando el tamaño de la muestra crece se obtienen estimaciones más precisas. No obstante, cuando el porcentaje de outliers aumenta, los estimadores ML y M pierden precisión más rápido que el estimador AM. En resumen, los AM-estimadores muestran ya sea el mismo o mejor desempeño que los estimadores ML y M en todos los casos de estudio.

Por otro lado, es importante investigar el comportamiento de los AM-estimadores para diferentes tamaños de muestra ( $N$ ) y frecuencia de contaminación ( $\epsilon$ ). Las figuras 4.3, 4.4, y 4.5 muestran los AM-estimadores para  $\alpha = \{-1, -6, -10\}$ , considerando  $N = \{9, 25, 49, 81\}$  y  $\epsilon = \{0,00, 0,01, 0,05, 0,10\}$ , para outliers de magnitud 40. En las figuras no se aprecia una diferencia significativa entre las estimaciones basadas en diferentes tamaños de muestra, lo que permite concluir que los AM-estimadores son eficientes en un amplio rango de situaciones.





Tabla 4.3: Comparación numérica del error cuadrático medio entre los estimadores ML, M y AM, variando  $\alpha$ , el tamaño muestral  $N$  y el nivel de contaminación  $\epsilon$ , con  $v = 5$ .

$\epsilon$	$N$	$\alpha = -1$			$\alpha = -6$			$\alpha = -10$			
		$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	
0,00	9	0,218	0,218	0,234	5,265	5,267	5,274	5,824	5,825	5,825	5,886
	25	0,046	0,052	0,058	1,801	1,802	1,804	3,516	3,516	3,516	3,529
	49	0,021	0,024	0,028	0,799	0,799	0,799	2,207	2,207	2,207	2,207
	81	0,014	0,016	0,018	0,477	0,476	0,476	1,377	1,377	1,377	1,377
0,01	9	0,056	0,078	0,129	8,471	8,555	8,645	29,109	29,140	29,140	17,442
	25	0,033	0,042	0,052	2,775	2,786	2,798	10,350	10,352	10,352	5,106
	49	0,020	0,024	0,029	1,461	1,465	1,469	4,934	4,934	4,934	2,596
	81	0,012	0,014	0,016	0,915	0,917	0,919	2,988	2,989	2,989	1,573
0,05	9	0,058	0,072	0,115	9,673	9,794	9,928	32,389	32,450	32,450	22,094
	25	0,043	0,045	0,053	4,778	4,812	4,846	16,932	16,941	16,941	9,114
	49	0,032	0,031	0,036	3,550	3,568	3,586	12,859	12,862	12,862	6,129
	81	0,026	0,022	0,024	3,224	3,237	3,251	12,259	12,262	12,262	5,378
0,10	9	0,069	0,078	0,125	11,223	11,403	11,609	37,020	37,130	37,130	28,813
	25	0,060	0,053	0,058	7,754	7,836	7,925	26,177	26,213	26,213	17,022
	49	0,058	0,048	0,051	7,091	7,152	7,217	24,965	24,988	24,988	14,636
	81	0,058	0,045	0,048	7,308	7,367	7,429	24,873	24,891	24,891	13,960

Tabla 4.4: Comparación numérica del error cuadrático medio entre los estimadores ML, M y AM, variando  $\alpha$ , el tamaño muestral  $N$  y el nivel de contaminación  $\epsilon$ , con  $v = 15$ .

$\epsilon$	$N$	$\alpha = -1$			$\alpha = -6$			$\alpha = -10$			
		$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	$\widehat{mse}[\hat{\alpha}_{ML}]$	$\widehat{mse}[\hat{\alpha}_M]$	$\widehat{mse}[\hat{\alpha}_{AM}]$	
0,00	9	0,218	0,218	0,234	5,316	5,320	5,325	6,036	6,036	6,036	6,101
	25	0,046	0,052	0,058	1,647	1,647	1,648	3,636	3,636	3,636	3,650
	49	0,021	0,024	0,028	0,782	0,782	0,782	2,189	2,189	2,189	2,189
	81	0,014	0,016	0,018	0,444	0,444	0,444	1,415	1,415	1,415	1,417
0,01	9	0,114	0,078	0,310	17,546	10,494	6,079	59,377	44,509	44,509	6,389
	25	0,045	0,042	0,061	7,815	3,225	1,946	31,439	17,195	17,195	3,458
	49	0,026	0,024	0,030	4,219	1,637	0,861	17,194	8,014	8,014	2,208
	81	0,016	0,014	0,016	2,777	1,051	0,488	11,325	4,962	4,962	1,461
0,05	9	0,124	0,072	0,314	18,690	12,054	5,538	62,132	48,426	48,426	5,854
	25	0,074	0,045	0,061	11,444	5,793	2,143	40,882	26,000	26,000	3,386
	49	0,058	0,031	0,032	9,027	4,165	0,862	34,517	20,700	20,700	2,298
	81	0,050	0,022	0,018	8,690	3,835	0,490	33,603	19,612	19,612	1,402
0,10	9	0,148	0,078	0,480	19,958	13,810	5,859	64,993	52,696	52,696	6,463
	25	0,110	0,052	0,068	15,480	9,404	2,177	52,624	38,525	38,525	3,826
	49	0,114	0,047	0,032	15,378	8,976	0,943	51,837	36,880	36,880	2,221
	81	0,120	0,046	0,019	15,725	9,086	0,544	53,149	37,672	37,672	1,486

Tabla 4.5: Comparación numérica del sesgo relativo absoluto entre los estimadores ML, M y AM, variando  $\alpha$ , el tamaño muestral  $N$  y el nivel de contaminación  $\epsilon$ , con  $v = 5$ .

$\epsilon$	$N$	$\alpha = -1$						$\alpha = -6$						$\alpha = -10$					
		$\hat{B}[\hat{\alpha}_{ML}]$	$\hat{B}[\hat{\alpha}_M]$	$\hat{B}[\hat{\alpha}_{AM}]$															
0,00	9	0,162	0,140	0,154	0,087	0,086	0,086	0,086	0,086	0,086	0,086	0,086	0,008	0,008	0,008	0,010			
	25	0,048	0,041	0,045	0,047	0,047	0,047	0,047	0,047	0,047	0,047	0,047	0,031	0,031	0,031	0,030			
	49	0,013	0,004	0,002	0,020	0,020	0,020	0,020	0,020	0,020	0,020	0,020	0,017	0,017	0,017	0,017			
	81	0,014	0,012	0,013	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011			
0,01	9	0,174	0,080	0,018	0,476	0,478	0,478	0,481	0,481	0,481	0,481	0,535	0,538	0,538	0,383				
	25	0,089	0,057	0,042	0,250	0,250	0,250	0,251	0,251	0,251	0,302	0,302	0,302	0,163					
	49	0,060	0,043	0,036	0,168	0,168	0,168	0,168	0,168	0,168	0,192	0,192	0,192	0,094					
	81	0,044	0,033	0,029	0,122	0,123	0,123	0,123	0,123	0,123	0,140	0,140	0,140	0,064					
0,05	9	0,185	0,091	0,035	0,507	0,510	0,510	0,513	0,513	0,513	0,562	0,562	0,562	0,430					
	25	0,143	0,100	0,082	0,335	0,336	0,336	0,337	0,337	0,387	0,387	0,387	0,245						
	49	0,128	0,095	0,086	0,284	0,284	0,284	0,285	0,285	0,332	0,332	0,332	0,201						
	81	0,125	0,092	0,085	0,276	0,277	0,277	0,278	0,278	0,329	0,329	0,329	0,199						
0,10	9	0,208	0,111	0,050	0,545	0,549	0,549	0,553	0,553	0,553	0,599	0,599	0,600	0,495					
	25	0,200	0,152	0,137	0,439	0,441	0,441	0,443	0,443	0,490	0,490	0,491	0,363						
	49	0,216	0,173	0,168	0,426	0,427	0,427	0,429	0,429	0,484	0,484	0,485	0,353						
	81	0,226	0,185	0,183	0,440	0,441	0,441	0,443	0,443	0,490	0,490	0,490	0,355						

Tabla 4.6: Comparación numérica del sesgo relativo absoluto entre los estimadores ML, M y AM, variando  $\alpha$ , el tamaño muestral  $N$  y el nivel de contaminación  $\epsilon$ , con  $v = 15$ .

$\epsilon$	$N$	$\alpha = -1$						$\alpha = -6$						$\alpha = -10$					
		$\hat{B}[\hat{\alpha}_{ML}]$	$\hat{B}[\hat{\alpha}_M]$	$\hat{B}[\hat{\alpha}_{AM}]$															
0,00	9	0,162	0,140	0,154	0,085	0,085	0,085	0,084	0,000	0,000	0,000	0,001	0,001	0,000	0,000	0,001			
	25	0,048	0,041	0,045	0,044	0,044	0,044	0,044	0,030	0,030	0,030	0,029	0,029	0,030	0,030	0,029			
	49	0,013	0,004	0,002	0,019	0,019	0,019	0,019	0,018	0,018	0,018	0,017	0,017	0,018	0,018	0,017			
	81	0,014	0,012	0,013	0,010	0,010	0,010	0,010	0,012	0,012	0,012	0,012	0,012	0,012	0,012	0,012			
0,01	9	0,318	0,080	0,258	0,697	0,533	0,110	0,770	0,666	0,666	0,666	0,016	0,016	0,666	0,666	0,016			
	25	0,167	0,057	0,044	0,459	0,274	0,049	0,557	0,404	0,404	0,404	0,021	0,021	0,404	0,404	0,021			
	49	0,106	0,043	0,013	0,326	0,177	0,021	0,404	0,262	0,262	0,262	0,024	0,024	0,262	0,262	0,024			
	81	0,078	0,033	0,006	0,256	0,135	0,008	0,319	0,197	0,197	0,197	0,019	0,019	0,197	0,197	0,019			
0,05	9	0,333	0,091	0,285	0,718	0,568	0,101	0,787	0,692	0,692	0,692	0,006	0,006	0,692	0,692	0,006			
	25	0,233	0,100	0,052	0,550	0,369	0,062	0,631	0,492	0,492	0,492	0,021	0,021	0,492	0,492	0,021			
	49	0,204	0,095	0,022	0,481	0,309	0,027	0,571	0,430	0,430	0,430	0,023	0,023	0,430	0,430	0,023			
	81	0,198	0,092	0,016	0,474	0,303	0,018	0,565	0,423	0,423	0,423	0,015	0,015	0,423	0,423	0,015			
0,10	9	0,362	0,114	0,354	0,741	0,606	0,109	0,804	0,720	0,720	0,720	0,005	0,005	0,720	0,720	0,005			
	25	0,299	0,139	0,085	0,642	0,482	0,055	0,717	0,603	0,603	0,603	0,007	0,007	0,603	0,603	0,007			
	49	0,319	0,170	0,034	0,644	0,482	0,021	0,712	0,593	0,593	0,593	0,020	0,020	0,593	0,593	0,020			
	81	0,334	0,186	0,019	0,655	0,491	0,012	0,725	0,606	0,606	0,606	0,015	0,015	0,606	0,606	0,015			

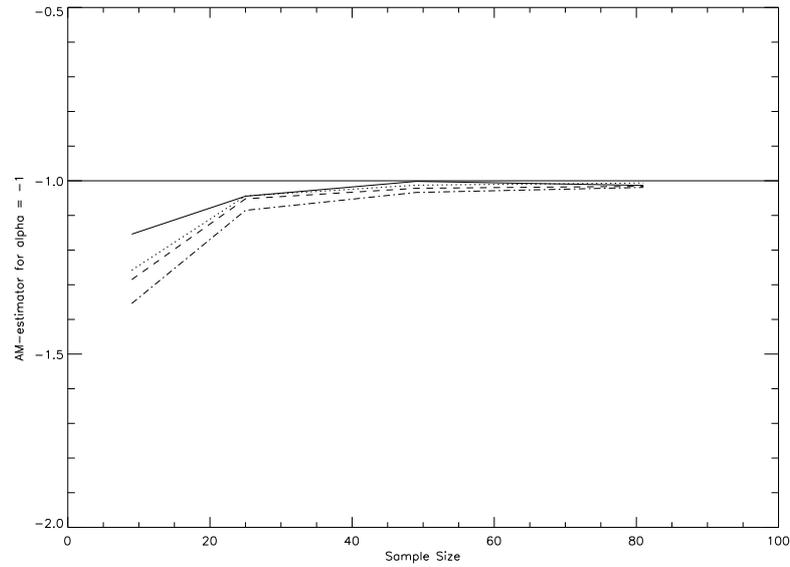


Figura 4.3: AM-estimadores para  $\alpha = -1$  versus tamaño muestral, con nivel de contaminación:  $\epsilon = 0,00$  (línea sólida),  $\epsilon = 0,01$  (línea punteada),  $\epsilon = 0,05$  (línea segmentada) and  $\epsilon = 0,10$  (línea punteada-segmentada).

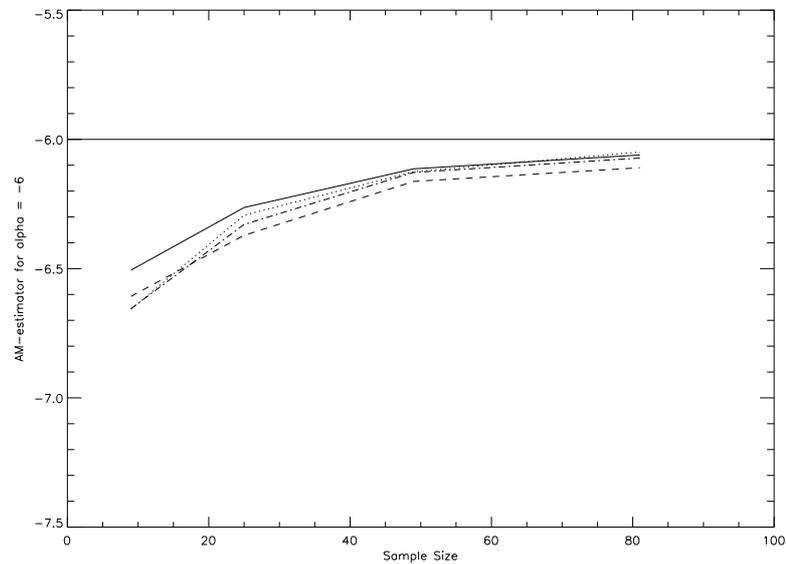


Figura 4.4: AM-estimadores para  $\alpha = -6$  versus tamaño muestral, con nivel de contaminación:  $\epsilon = 0,00$  (línea sólida),  $\epsilon = 0,01$  (línea punteada),  $\epsilon = 0,05$  (línea segmentada) and  $\epsilon = 0,10$  (línea punteada-segmentada).

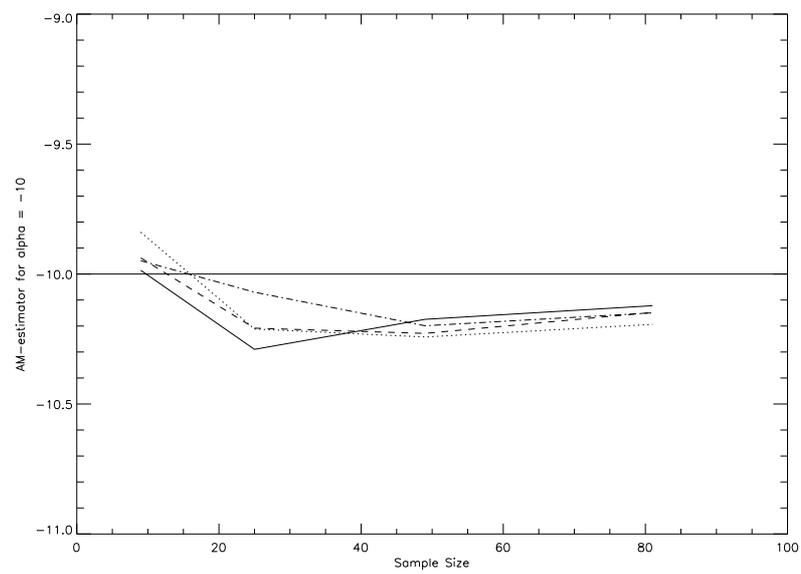


Figura 4.5: AM-estimadores para  $\alpha = -10$  versus tamaño muestral, con nivel de contaminación:  $\epsilon = 0,00$  (línea sólida),  $\epsilon = 0,01$  (línea punteada),  $\epsilon = 0,05$  (línea segmentada) and  $\epsilon = 0,10$  (línea punteada-segmentada).

# Capítulo 5

## Conclusiones y Trabajo Futuro

Esta tesis se origina con la intención de proveer un mecanismo alternativo a los métodos clásicos de estimación robusta, los cuales fueron establecidos bajo supuestos de distribuciones simétricas. No obstante, en innumerables situaciones de la vida real los datos no obedecen estas leyes, sino por el contrario, es frecuente descubrir procesos que obedecen leyes distribuciones no simétricas. En estos casos, los procedimientos de estimación robusta no son adecuados, por lo que se hace necesario contar con metodologías de estimación diseñadas especialmente para trabajar con este tipo de datos, de manera tal que las estimaciones alcancen la más alta precisión como sea posible.

Si bien es cierto este estudio nace en el contexto del modelamiento de imágenes SAR, el procedimiento de inferencia robusta propuesto puede ser aplicado a cualquier fenómeno cuyas observaciones obedezcan una distribución no simétrica.

Se han propuesto los AM-estimadores (Asymmetric M-estimators), un nuevo método robusto basado en una familia de funciones de robustez asimétricas. Se desarrolló un estudio de Monte Carlo para investigar las propiedades de robustez de estos estimadores, comparándolos con los clásicos estimadores de máxima verosimilitud y

los M-estimadores robustos. Se estimó el parámetro de rugosidad  $\alpha$  de la distribución  $\mathcal{G}_A^0$ , y los resultados mostraron que el estimador  $\hat{\alpha}_{AM}$  es muy favorable con respecto a los estimadores  $\hat{\alpha}_{ML}$  y  $\hat{\alpha}_M$  bajo el criterio del error cuadrático medio, en todas las situaciones bajo estudio. Además, se puede ver que  $\hat{\alpha}_{AM}$  tiene un sesgo más pequeño en presencia de contaminación. Al igual que se constató que el AM-estimador tiene un comportamiento similar tanto en grandes muestras como en pequeñas muestras.

Es importante mencionar que el esfuerzo computacional requerido para computar  $\hat{\alpha}_{AM}$  es comparable con el esfuerzo de computar  $\hat{\alpha}_M$  bajo la metodología descrita en (Bustos and Ling, 1996), gracias a la cual tampoco es vital contar con una solución inicial antes del proceso de estimación robusto. No obstante, el cómputo de los AM-estimadores presenta algunos problemas numéricos cuando se trabaja con pequeñas muestras.

Las extensiones de este trabajo son muchas. Se debe continuar con la estimación simultánea de los parámetros  $\alpha$  y  $\gamma$ . También es importante extender estos resultados para el caso multilook ( $n > 1$ ), al igual que para imágenes multivariadas polarimétricas. Asimismo, se pretende desarrollar filtros para imágenes SAR basados en los AM-estimadores.

Por otra parte, pueden estudiarse estimadores robustos para intervalos de confianza basados en funciones de robustez asimétricas, los cuales pueden ser combinados con técnicas de bootstrap u otros métodos de remuestreo.

En el ámbito de la aplicaciones, utilizando los AM-estimadores pueden desarrollarse algoritmos más precisos de clasificación y segmentación para imágenes SAR.

# Apéndice A

## Implementación de los Estimadores Robustos

Como se estableció en el capítulo 3, los estimadores robustos son solución de la ecuación

$$\sum_{i=1}^N \psi[s(z_i; \theta) - c] = 0, \quad (\text{A.1})$$

donde  $c$  son funciones correctoras que permiten obtener estimadores asintóticamente insesgados (consistencia de Fisher), y están definidas implícitamente por

$$\int_{-\infty}^{\infty} \psi[s(z_i; \theta) - c] dF_{\theta}(z) = 0. \quad (\text{A.2})$$

En el caso del M-estimador de Huber  $\psi = \psi_b$  de la ecuación (3.4), y en el caso del AM-estimador  $\psi = \psi_{r_1, r_2}$  de la ecuación (3.10), donde  $b, r_1$  y  $r_2$  son los parámetros de ajuste.

En (Bustos and Ling, 1996) se propone el siguiente procedimiento de dos pasos para resolver el problema de estimación:

1. Fijar los parámetros de ajuste  $p$  de la función de robustez  $\psi_p$  y resolver la ecuación (A.2) con respecto a  $c$  para un conjunto discreto de valores  $\alpha$  y almacenar las soluciones. Más precisamente, escoger un intervalo  $[\alpha_1, \alpha_2]$  y un entero  $k$  (por ejemplo,  $k = 100$ ), y tabular las soluciones  $c(\alpha)$  para  $\alpha = \alpha_1 + (i - 1)(\alpha_2 - \alpha_1)/(k - 1)$ ,  $i = 1, \dots, k$ .
2. Resolver la ecuación (A.1) con respecto a  $\alpha$ , usando una interpolación lineal de la tabla obtenida en el paso 1 con el objeto de computar  $c(\alpha)$  para los valores de  $\alpha$  requeridos.

En la práctica, es necesario fijar  $\alpha_1$  y  $\alpha_2$  tal que  $[\alpha_1, \alpha_2]$  contenga la solución  $\hat{\alpha}$  de la ecuación (A.1). Sin embargo, la computación preliminar del paso 1 acelera sustancialmente el paso 2.

A continuación se desarrollarán los aspectos numéricos relacionados con la solución de la ecuación (A.2) tanto para el M-estimador de Huber como para el AM-estimador propuesto bajo la estimación del parámetro de rugosidad  $\alpha$ . En ambos casos se quiere encontrar el valor de la función  $c$  tal que

$$I_\alpha = \int_{-\infty}^{\infty} \psi[s_1(z; \theta) - c]f(z)dz = 0, \quad (\text{A.3})$$

donde  $s_1$  es el primer componente de la función score definida en (2.38) y  $f$  es la función de densidad de probabilidad de la distribución  $\mathcal{G}_A^0$  definida en (2.18). Para resolver  $I_\alpha$  es necesario recurrir a la definición por partes de las funciones  $\psi$ ; siendo conveniente introducir en este punto los cambios de variable  $u = \frac{z^2}{\gamma}$  e  $y = \alpha^{-1} + \ln(1 + u) - c$ . De esta forma, la función de robustificación para cada estimador puede expresarse como

$$\psi_b(y) = \begin{cases} -b & , u \leq A \\ y & , A < u < B \\ b & , u \geq B \end{cases} \quad \text{donde} \quad \begin{cases} A = \exp(c - b - \alpha^{-1}) - 1 \\ B = \exp(c + b - \alpha^{-1}) - 1 \end{cases} \quad (\text{A.4})$$

$$\psi_{r_1, r_2}(y) = \begin{cases} 0 & , u < A \\ -y - r_1 & , A \leq u < B \\ y & , B \leq u < C \\ -y + r_2 & , C \leq u < D \\ 0 & , u \geq D \end{cases} \quad \text{donde} \quad \begin{cases} A = \exp(c - r_1 - \alpha^{-1}) - 1 \\ B = \exp(c - r_1/2 - \alpha^{-1}) - 1 \\ C = \exp(c + r_2/2 - \alpha^{-1}) - 1 \\ D = \exp(c + r_2 - \alpha^{-1}) - 1 \end{cases} \quad (\text{A.5})$$

## A.1. Funciones Correctoras del M-estimador

Desarrollando por partes la integral  $I_\alpha$  en el caso del M-estimador de Huber se tiene

$$\begin{aligned} I_\alpha &= \int_{-\infty}^{\infty} \psi_b(y) f(u) du \\ &= \underbrace{\int_{-\infty}^A -bf(u) du}_{I_1} + \underbrace{\int_A^B yf(u) du}_{I_2} + \underbrace{\int_B^{\infty} bf(u) du}_{I_3} , \end{aligned} \quad (\text{A.6})$$

con

$$I_1 = -bF(A), \quad (\text{A.7})$$

$$I_2 = (1+u)^\alpha (c - \ln(1+u)) \Big|_A^B, \quad (\text{A.8})$$

$$I_3 = b[1 - F(B)], \quad (\text{A.9})$$

donde  $F$  es la función de distribución acumulada de la distribución  $\mathcal{G}_A^0$  definida en (2.19).

Dependiendo del dominio de integración, la integral  $I_\alpha = I_1 + I_2 + I_3$  puede definirse bajo tres casos:

**Caso 1:**  $0 < A < B \implies c \geq b + \alpha^{-1}$  .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (A.7), (A.8) y (A.9), respectivamente.

**Caso 2:**  $A < 0 < B \implies -b + \alpha^{-1} < c < b + \alpha^{-1}$  .

Entonces

$$\begin{aligned} I_1 &= 0, \\ I_2 &= (1+u)^\alpha (c - \ln(1+u)) \Big|_0^B, \\ I_3 &\text{ como definido anteriormente en (A.9).} \end{aligned} \tag{A.10}$$

**Caso 3:**  $A < B < 0 \implies c < -b + \alpha^{-1}$  .

Entonces

$$\begin{aligned} I_1 &= I_2 = 0, \\ I_3 &= b. \end{aligned} \tag{A.11}$$

Lo cual no tiene sentido puesto que debe cumplirse  $I_\alpha = 0$ , y  $b > 0$ . Por lo tanto, este caso no es factible.

La solución de  $I_\alpha = 0$  con respecto a  $c$  puede encontrarse por métodos numéricos, por ejemplo mediante el algoritmo de Broyden (Burden and Faires, 1985), al cual se debe proporcionar una solución inicial  $c_0$ , fijada en cero en este estudio.

## A.2. Funciones Correctoras del AM-estimador

Al desarrollar por partes la integral  $I_\alpha$  de la ecuación (A.3) en el caso del AM-estimador se tiene

$$\begin{aligned} I_\alpha &= \int_{-\infty}^{\infty} \psi_{r_1, r_2}(y) f(u) du \\ &= \underbrace{\int_A^B -(y+r_1)f(u) du}_{I_1} + \underbrace{\int_B^C yf(u) du}_{I_2} + \underbrace{\int_C^D -(y-r_2)f(u) du}_{I_3}, \end{aligned} \tag{A.12}$$

con

$$I_1 = -(1+u)^\alpha(c-r_1-\ln(1+u))\Big|_A^B, \quad (\text{A.13})$$

$$I_2 = (1+u)^\alpha(c-\ln(1+u))\Big|_B^C, \quad (\text{A.14})$$

$$I_3 = -(1+u)^\alpha(c+r_2-\ln(1+u))\Big|_C^D. \quad (\text{A.15})$$

Dependiendo del dominio de integración, la integral  $I_\alpha = I_1 + I_2 + I_3$  puede definirse bajo cinco casos:

**Caso 1:**  $0 < A < B < C < D \implies c > r_1 + \alpha^{-1}$  .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (A.13), (A.14) y (A.15), respectivamente.

**Caso 2:**  $A < 0 < B < C < D \implies r_1/2 + \alpha^{-1} < c \leq r_1 + \alpha^{-1}$  .

Entonces

$$I_1 = -(1+u)^\alpha(c-r_1-\ln(1+u))\Big|_0^B,$$

$$I_2, I_3 \text{ como definidos anteriormente en (A.14) y (A.15), respectivamente.} \quad (\text{A.16})$$

**Caso 3:**  $A < B < 0 < C < D \implies -r_2/2 + \alpha^{-1} < c \leq r_1/2 + \alpha^{-1}$  .

Entonces

$$I_1 = 0,$$

$$I_2 = (1+u)^\alpha(c-\ln(1+u))\Big|_0^C, \quad (\text{A.17})$$

$$I_3 \text{ como definido anteriormente en (A.15).}$$

**Caso 4:**  $A < B < C < 0 < D \implies -r_2 + \alpha^{-1} < c \leq -r_2/2 + \alpha^{-1}$  .

Entonces

$$I_1 = I_2 = 0,$$

$$I_3 = -(1+u)^\alpha(c+r_2-\ln(1+u))\Big|_0^D. \quad (\text{A.18})$$

**Caso 5:**  $A < B < C < D < 0 \implies c \leq -r_2 + \alpha^{-1}$  .

Entonces

$$I_1 = I_2 = I_3 = 0. \quad (\text{A.19})$$

# Apéndice B

## Cómputo de los Parámetros de Ajuste

Es bien sabido que bajo condiciones de regularidad muy poco restrictivas los estimadores de máxima verosimilitud son asintóticamente normales y óptimos con respecto a varianza. Es por ello que la definición de estimadores robustos eficientes requiere el ajuste de los parámetros  $p$  de las funciones de robustificación  $\psi_p$ , los cuales son escogidos de manera que la eficiencia de los estimadores robustos con respecto a los de máxima verosimilitud no sea demasiado baja en el modelo sin outliers. Esta restricción impone que la varianza asintótica de los estimadores robustos no sobrepase la de los estimadores de máxima verosimilitud en más de un cierto factor controlable. De esta forma, se define la eficiencia relativa asintótica como

$$ARE = \frac{V(\hat{\theta}_{ML})}{V(\hat{\theta}_{robusto})}, \quad (\text{B.1})$$

y se requiere controlar este factor en el rango  $90\% < ARE < 95\%$ .

A continuación se desarrollarán los aspectos numéricos relacionados con la búsqueda de los parámetros de ajuste, tanto para el M-estimador de Huber como para el AM-estimador propuesto bajo la estimación del parámetro de rugosidad  $\alpha$  (caso uni-

dimensional). Para el caso de estimación conjunta revisar la extensión al caso multi-dimensional en (Hampel *et al.*, 1986).

## B.1. Varianza Asintótica del ML-estimador

Se define la varianza asintótica del estimador de máxima verosimilitud del parámetro  $\alpha$  como  $V(\hat{\alpha}_{ML}) = J^{-1}$ , donde  $J$  es la matriz de información de Fisher definida por

$$J = \int_{-\infty}^{\infty} s_1(z; \theta)^2 dF, \quad (\text{B.2})$$

donde  $s_1$  es la primera componente de la función score definida en (2.38) y  $F$  es la función de distribución acumulada de la distribución  $\mathcal{G}_A^0$  definida en (2.19).

Al integrar la ecuación (B.2), la varianza asintótica del estimador de máxima verosimilitud viene dada por

$$V(\hat{\alpha}_{ML}) = \alpha^2. \quad (\text{B.3})$$

## B.2. Varianza Asintótica del M-estimador

Se define la varianza asintótica del M-estimador de Huber del parámetro  $\alpha$  como

$$V(\hat{\alpha}_M) = \frac{\int_{-\infty}^{\infty} \psi_b^2(y) dF}{\left( \int_{-\infty}^{\infty} \psi_b(y) s_1(z; \theta) dF \right)^2}. \quad (\text{B.4})$$

Desarrollando por partes el numerador, considerando los cambios de variable  $u = \frac{z^2}{\gamma}$  e  $y = \alpha^{-1} + \ln(1 + u) - c$ , se tiene

$$\begin{aligned} T_1 &= \int_{-\infty}^{\infty} \psi_b^2(y) dF \\ &= \underbrace{\int_{-\infty}^A (-b)^2 f(u) du}_{I_1} + \underbrace{\int_A^B y^2 f(u) du}_{I_2} + \underbrace{\int_B^{\infty} b^2 f(u) du}_{I_3}, \end{aligned} \quad (\text{B.5})$$

con

$$I_1 = b^2 F(A), \quad (\text{B.6})$$

$$I_2 = -(1+u)^\alpha \{ \alpha^{-2} + c^2 + \ln(1+u)[-2c + \ln(1+u)] \} \Big|_A^B, \quad (\text{B.7})$$

$$I_3 = b^2 [1 - F(B)], \quad (\text{B.8})$$

donde  $f$  es la función de densidad de probabilidad de la distribución  $\mathcal{G}_A^0$  definida en (2.18) y  $A, B$  y  $C$  son los definidos en (A.4). La ecuación (B.5) posee dos casos de solución:

**Caso 1:**  $0 < A < B \implies c \geq b + \alpha^{-1}$  .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (B.6), (B.7) y (B.8), respectivamente.

**Caso 2:**  $A < 0 < B \implies -b + \alpha^{-1} < c < b + \alpha^{-1}$  .

Entonces

$$\begin{aligned} I_1 &= 0, \\ I_2 &= -(1+u)^\alpha \{ \alpha^{-2} + c^2 + \ln(1+u)[-2c + \ln(1+u)] \} \Big|_0^B, \\ I_3 &\text{ como definido anteriormente en (B.8).} \end{aligned} \quad (\text{B.9})$$

Desarrollando por partes el denominador de la ecuación (B.4), se tiene

$$\begin{aligned} T_2 &= \int_{-\infty}^{\infty} \psi_b(y) s_1(z; \theta) dF \\ &= \underbrace{\int_{-\infty}^A -b s_1(u; \theta) f(u) du}_{I_1} + \underbrace{\int_A^B y s_1(u; \theta) f(u) du}_{I_2} + \underbrace{\int_B^{\infty} b s_1(u; \theta) f(u) du}_{I_3}, \end{aligned} \quad (\text{B.10})$$

con

$$I_1 = b(1+A)^\alpha \ln(1+A), \quad (\text{B.11})$$

$$I_2 = - \{ (1+u)^\alpha [ \alpha^{-2} - c \ln(1+u) + \ln^2(1+u) ] \} \Big|_A^B, \quad (\text{B.12})$$

$$I_3 = b(1+B)^\alpha \ln(1+B). \quad (\text{B.13})$$

La ecuación (B.10) posee dos casos de solución:

**Caso 1:**  $0 < A < B \implies c \geq b + \alpha^{-1}$  .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (B.11), (B.12) y (B.13), respectivamente.

**Caso 2:**  $A < 0 < B \implies -b + \alpha^{-1} < c < b + \alpha^{-1}$  .

Entonces

$$\begin{aligned} I_1 &= 0, \\ I_2 &= - \left\{ (1+u)^\alpha [\alpha^{-2} - c \ln(1+u) + \ln^2(1+u)] \right\} \Big|_0^B, \\ I_3 &\text{ como definido anteriormente en (B.13).} \end{aligned} \quad (\text{B.14})$$

Finalmente, la varianza asintótica del M-estimador de Huber está dada por

$$V(\hat{\alpha}_M) = \frac{T_1}{(T_2)^2}. \quad (\text{B.15})$$

### B.3. Varianza Asintótica del AM-estimador

Se define la varianza asintótica del AM-estimador del parámetro  $\alpha$  propuesto en (3.10) como

$$V(\hat{\alpha}_{AM}) = \frac{\int_{-\infty}^{\infty} \psi_{r_1, r_2}^2(y) dF}{\left( \int_{-\infty}^{\infty} \psi_{r_1, r_2}(y) s_1(z; \theta) dF \right)^2}. \quad (\text{B.16})$$

Desarrollando por partes el numerador, considerando los cambios de variable  $u = \frac{z^2}{\gamma}$  e  $y = \alpha^{-1} + \ln(1+u) - c$ , se tiene

$$\begin{aligned} T_1 &= \int_{-\infty}^{\infty} \psi_{r_1, r_2}^2(y) dF \\ &= \underbrace{\int_A^B [-(y+r_1)]^2 f(u) du}_{I_1} + \underbrace{\int_B^C y^2 f(u) du}_{I_2} + \underbrace{\int_C^D [-(y-r_2)]^2 f(u) du}_{I_3}, \end{aligned} \quad (\text{B.17})$$

con

$$I_1 = -(1+u)^\alpha \left\{ \alpha^{-2} + (c-r_1)^2 + \ln(1+u)[-2(c-r_1) + \ln(1+u)] \right\} \Big|_A^B, \quad (\text{B.18})$$

$$I_2 = -(1+u)^\alpha \left\{ \alpha^{-2} + c^2 + \ln(1+u)[-2c + \ln(1+u)] \right\} \Big|_B^C, \quad (\text{B.19})$$

$$I_3 = -(1+u)^\alpha \left\{ \alpha^{-2} + (c+r_2)^2 + \ln(1+u)[-2(c+r_2) + \ln(1+u)] \right\} \Big|_C^D. \quad (\text{B.20})$$

donde  $A, B, C$  y  $D$  son los definidos en (A.5). La ecuación (B.17) posee cinco casos de solución:

**Caso 1:**  $0 < A < B < C < D \implies c > r_1 + \alpha^{-1}$  .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (B.18), (B.19) y (B.20), respectivamente.

**Caso 2:**  $A < 0 < B < C < D \implies r_1/2 + \alpha^{-1} < c \leq r_1 + \alpha^{-1}$  .

Entonces

$$\begin{aligned}
 I_1 &= -(1+u)^\alpha \{ \alpha^{-2} + (c-r_1)^2 + \ln(1+u)[-2(c-r_1) + \ln(1+u)] \} \Big|_0^B, \\
 I_2, I_3 &\text{ como definidos anteriormente en (B.19) y (B.20), respectivamente.}
 \end{aligned} \tag{B.21}$$

**Caso 3:**  $A < B < 0 < C < D \implies -r_2/2 + \alpha^{-1} < c \leq r_1/2 + \alpha^{-1}$  .

Entonces

$$\begin{aligned}
 I_1 &= 0, \\
 I_2 &= -(1+u)^\alpha \{ \alpha^{-2} + c^2 + \ln(1+u)[-2c + \ln(1+u)] \} \Big|_0^C, \\
 I_3 &\text{ como definido anteriormente en (B.20).}
 \end{aligned} \tag{B.22}$$

**Caso 4:**  $A < B < C < 0 < D \implies -r_2 + \alpha^{-1} < c \leq -r_2/2 + \alpha^{-1}$  .

Entonces

$$\begin{aligned}
 I_1 &= I_2 = 0, \\
 I_3 &= -(1+u)^\alpha \{ \alpha^{-2} + (c+r_2)^2 + \ln(1+u)[-2(c+r_2) + \ln(1+u)] \} \Big|_0^D.
 \end{aligned} \tag{B.23}$$

**Caso 5:**  $A < B < C < D < 0 \implies c \leq -r_2 + \alpha^{-1}$  .

Entonces

$$I_1 = I_2 = I_3 = 0. \tag{B.24}$$

Desarrollando por partes el denominador de la ecuación (B.16), se tiene

$$\begin{aligned}
 T_2 &= \int_{-\infty}^{\infty} \psi_{r_1, r_2}(y) s_1(z; \theta) dF \\
 &= \underbrace{\int_A^B -(y + r_1) s_1(u; \theta) f(u) du}_{I_1} + \underbrace{\int_B^C y s_1(u; \theta) f(u) du}_{I_2} + \underbrace{\int_C^D -(y - r_2) s_1(u; \theta) f(u) du}_{I_3},
 \end{aligned} \tag{B.25}$$

con

$$I_1 = -(1+u)^\alpha \left\{ -\alpha^{-2} + \ln(1+u)[c - r_1 - \ln(1+u)] \right\} \Big|_A^B, \tag{B.26}$$

$$I_2 = -(1+u)^\alpha \left\{ \alpha^{-2} - \ln(1+u)[c - \ln(1+u)] \right\} \Big|_B^C, \tag{B.27}$$

$$I_3 = -(1+u)^\alpha \left\{ -\alpha^{-2} + \ln(1+u)[c + r_2 - \ln(1+u)] \right\} \Big|_C^D. \tag{B.28}$$

La ecuación (B.25) posee cinco casos de solución:

**Caso 1:**  $0 < A < B < C < D \implies c > r_1 + \alpha^{-1}$ .

Entonces  $I_1, I_2$  e  $I_3$  se expresan como en (B.26), (B.27) y (B.28), respectivamente.

**Caso 2:**  $A < 0 < B < C < D \implies r_1/2 + \alpha^{-1} < c \leq r_1 + \alpha^{-1}$ .

Entonces

$$\begin{aligned}
 I_1 &= -(1+u)^\alpha \left\{ -\alpha^{-2} + \ln(1+u)[c - r_1 - \ln(1+u)] \right\} \Big|_0^B, \\
 I_2, I_3 &\text{ como definidos anteriormente en (B.27) y (B.28), respectivamente.}
 \end{aligned} \tag{B.29}$$

**Caso 3:**  $A < B < 0 < C < D \implies -r_2/2 + \alpha^{-1} < c \leq r_1/2 + \alpha^{-1}$ .

Entonces

$$\begin{aligned}
 I_1 &= 0, \\
 I_2 &= -(1+u)^\alpha \left\{ \alpha^{-2} - \ln(1+u)[c - \ln(1+u)] \right\} \Big|_0^C, \\
 I_3 &\text{ como definido anteriormente en (B.28).}
 \end{aligned} \tag{B.30}$$

**Caso 4:**  $A < B < C < 0 < D \implies -r_2 + \alpha^{-1} < c \leq -r_2/2 + \alpha^{-1}$  .

Entonces

$$\begin{aligned} I_1 &= I_2 = 0, \\ I_3 &= -(1+u)^\alpha \left\{ -\alpha^{-2} + \ln(1+u)[c + r_2 - \ln(1+u)] \right\} \Big|_0^D. \end{aligned} \quad (\text{B.31})$$

**Caso 5:**  $A < B < C < D < 0 \implies c \leq -r_2 + \alpha^{-1}$  .

Entonces

$$I_1 = I_2 = I_3 = 0. \quad (\text{B.32})$$

Finalmente, la varianza asintótica del AM-estimador está dada por

$$V(\hat{\alpha}_{AM}) = \frac{T_1}{(T_2)^2}. \quad (\text{B.33})$$

# Bibliografía

- ALLENDE, H., GALBIATI, J., AND VALLEJOS, R. (2001). Robust image modelling on image processing. *Pattern Recognition Letters*, 22(11):1219–1231.
- ALLENDE, H. AND PIZARRO, L. (2003). Robust estimation of roughness parameter in SAR amplitude images. In SANFELIU, A. AND RUIZ-SHULCLOPER, J., editors, *Progress in Pattern Recognition, Speech and Image Analysis*, volume 2905, pages 129–136, Havana, Cuba. Springer-Verlag. Proceedings of LNCS.
- ANDREWS, D., BICKEL, P., HAMPEL, F., HUBER, P., ROGERS, W., AND TUCKEY, J. (1972). *Robust Estimates of Location: Survey and Advances*. Princeton University Press, Princeton, NJ.
- BANKS, J. (1998). *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*. Wiley, New York.
- BLACKNELL, D. (1994). Comparison of the parameter estimator for  $K$  distribution. In *IEE Proceedings - Radar, Sonar Navigation*, volume 141, pages 45–52.
- BURDEN, R. AND FAIRES, J. (1985). *Análisis Numérico*. Grupo Editorial Iberoamericana, México.
- BUSTOS, O. AND LING, K. (1996). *Procedimientos Robustos*. Springer-Verlag, Berlin. Lecture Notes in Statistics 109.
- BUSTOS, O., LUCINI, M., AND FRERY, A. (2002). M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, 2002(1):105–114.
- CASELLA, G. AND BERGER, R. (1990). *Statistical Inference*. Brooks/Cole, California.
- CRIBARI-NETO, F., FRERY, A., AND SILVA, M. (2002). Improved estimation of clutter properties in speckled imagery. *Computational Statistics and Data Analysis*, 40(4):801–824.
- FRERY, A., MÜLLER, H.-J., YANASSE, C., AND SANT’ANNA, S. (1997a). A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):648–659.

- FRERY, A., SANT'ANNA, S., MASCARENHAS, N., AND BUSTOS, O. (1997b). Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images. *Applied Signal Processing*, 4:61–76.
- GOODMAN, J. (1985). *Statistical Optics*. Wiley, New York.
- GRADSHTEYN, I. AND RYZHIK, I. (1998). *Table of Integrals, Series and Products*. Academic Press, New York.
- HAMPEL, F., RONCHETTI, E., ROUSSEEUW, P., AND STAHEL, W. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.
- HUBER, P. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35:73–101.
- HUBER, P. (1981). *Robust Statistics*. Wiley, New York.
- JAKEMAN, E. AND PUSEY, P. (1976). A model for non-Rayleigh sea echo. *IEEE Transactions on Antennas and Propagation*, AP-24:806–814.
- JAO, K. (1984). Amplitud distribution of composite terrain radar clutter and the  $K$  distribution. *IEEE Transactions on Antennas and Propagation*, AP-32:1049–1061.
- JIN, J., SILVA, G., AND FRERY, A. (2003). SAR despeckling filters in ultrasound imaging. *Latin American Applied Research*. In Press.
- JOUGHIN, I., PERCIVAL, D., AND WINEBRENNER, D. (1993). Maximum likelihood estimation of  $K$  distribution parameters for SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):989–999.
- KENDALL, M. AND STUART, A. (1979). *The Advanced Theory of Statistics: Inference and Relationship*, volume II. Macmillan, New York, 4th edition.
- LUCINI, M., RUIZ, V., FRERY, A., AND BUSTOS, O. (2003). Robust classification of SAR imagery. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 557–560, IEEE, Hong Kong.
- MEJAIL, M. (1999). *La distribución  $\mathcal{G}_A^0$  en el Modelado y Análisis de Imágenes de Radar de Apertura Sintética*. Ph.D. Thesis, Universidad Nacional de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Computación, Buenos Aires, Argentina.
- MEJAIL, M., FRERY, A., JACOBO-BERLLES, J., AND BUSTOS, O. (2001). Approximation of distributions for SAR images: proposal, evaluation and practical consequences. *Latin American Applied Research*, 31:83–92.

- MEJAIL, M., JACOBO-BERLLES, J., FRERY, A., AND BUSTOS, O. (2003). Classification of SAR images using a general and tractable multiplicative model. *International Journal of Remote Sensing*. In Press.
- NAVARRETE, H., FRERY, A., SÁNCHEZ, F., AND ANTÓ, J. (2002). Ultrasound images filtering using the multiplicative model. In INSANA, M. AND WALKER, W., editors, *Progress in Biomedical Optics and Imaging*, volume 4687, pages 345–353, San Diego, USA. Proceedings of SPIE.
- NIEVES, A. AND DOMÍNGUEZ, F. (1995). *Métodos Numéricos Aplicados a la Ingeniería*. CECSA, México.
- NIKIAS, C. AND SHAO, M. (1995). *Signal Processing with Alpha-Stable Distributions and Applications*. Wiley.
- OLIVEIRA, M. (2002). *Aspectos Numéricos na Inferência por Máxima Verossimilhança da Distribuição  $\mathcal{G}_A^0$* . Master Thesis, Universidade Federal de Pernambuco, Departamento de Estatística, Recife, PE, Brasil.
- OLIVER, C. AND QUEGAN, S. (1998). *Understanding Synthetic Aperture Radar Images*. Artech House, Boston.
- POLZEHL, J. AND SPOKOINY, V. (2003). Image denoising: pointwise adaptive approach. *Annals of Statistics*, 31:30–57.
- REED III, J. AND STARK, D. (1996). Hinge estimators of location: robust to asymmetry. *Computer Methods and Programs in Biomedicine*, 49:11–17.
- ROUSSEEUW, P. AND LEROY, A. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.
- ROUSSEEUW, P. AND VERBOVEN, S. (2002). Robust estimation in very small samples. *Computational Statistics and Data Analysis*, 40(4):741–758.
- TUR, M., CHIN, K., AND GOODMAN, J. (1982). When is the speckle noise multiplicative? *Applied Optics*, 21:1157–1159.
- YANASSE, C., FRERY, A., AND SANT’ANNA, S. (1995). Stochastic distributions and the multiplicative model: relations, properties, estimators and applications to SAR image analysis. Report 5630-NTC/318, INPE, São José dos Campos, SP, Brasil.