HOLOS
ISSN 1807 - 1600

SENSIBILIDAD DE LOS ESTIMADORES (A,U,Θ)

ARÍSTIDES ALEJANDRO LEGRÁ LOBAINA
Instituto Superior Minero Metalúrgico de Moa, Holguín, Cuba
alegra@ismm.edu.cu

Submetido em 09/05/2018 e aceito em 26/07/2019

DOI: 10.15628/holos.2020.7282

RESUMO

La eficacia de una estimación puntual (A,U,Θ) depende de la cantidad y calidad de un conjunto de datos W y de la capacidad que tenga el modelo para: sacarle el máximo a las características positivas de W; y evitar sus características negativas.

En este trabajo es analizada la sensibilidad de estos estimadores y en particular la causada por el mal condicionamiento de A, el cual ocasionalmente provoca que si hay pequeños cambios en el valor de los datos, se producen grandes cambios en los resultados. Se describe un procedimiento para detectar y evaluar estos problemas usando: Números de Condición, los errores de estimación y las pruebas de Validación Cruzada.

Se proponen criterios para resolver el mal condicionamiento del estimador mediante el uso de suficientes cifras decimales significativas en los cálculos y algoritmos eficientes; también se recomienda la selección de modelos adecuados pero simples y el uso de soportes compactos.

Finalmente se presenta un diagrama que facilita la comprensión del algoritmo para resolver los problemas de mal condicionamiento de los estimadores (A,U,Θ).

PALAVRAS-CHAVE: Estimador (A,U,Θ), Sensibilidad, Mal condicionamiento, Validación cruzada, Soporte de estimación.

SENSIBILITY OF THE ESTIMATORS (A,U,Θ)

ABSTRACT

The effectiveness of a point estimate (A, U, Θ) depends on the quantity and quality of a W data set and the ability of the model to: maximize the positive characteristics of W; and avoid its negative characteristics.

In this work the sensitivity of these estimators is analyzed and in particular the one caused by the ill conditioning of A, which occasionally causes that if there are small changes in the value of the data, great changes are produced in the results. A procedure is described to detect and evaluate these problems using: Condition Numbers, estimation errors and Cross Validation tests.

Criteria are proposed to solve the ill conditioning of the estimator by using enough significant decimal places in the calculations and efficient algorithms; the selection of suitable but simple models and the use of compact supports are also recommended.

Finally, a diagram that facilitates the understanding of the algorithm to solve the problems of ill conditioning of the estimators (A, U, Θ) is presented.

KEYWORDS: Estimator (A,U,Θ), Sensitivity, Ill conditioned, Cross validation, Estimation support

HOLOS ISSN 1807 - 1600

1 INTRODUCCIÓN

El conocido matemático francés Jacques S. Hadamard definió en el contexto de los Problemas de Cauchy (Ecuación Diferencial Ordinaria de Primer Orden con una Condición Inicial) el concepto de Problema Bien Planteado como aquel que, para los datos dados, cumpliera tres condiciones:

a. Tenga solución.

b. La solución sea única.

c. La solución depende de manera continua de las condiciones iniciales.

En el texto de Isaacson (1994) se explica que bajo la concepción de Hadamard un Problema de Cálculo X=Ɲ(b) está Bien Planteado si se cumplen las condiciones siguientes:

a. Para b dado existe un X tal que X= Ɲ(b)

b. Para cada valor de b el valor de X= Ɲ(b) es único

c. A pequeños cambios de b y de otros parámetros que intervienen en el cálculo, se producen pequeños cambios de X.

Estas condiciones son de interés en el problema de cálculo que permite obtener la solución de un sistema de ecuaciones lineales (SEL) que se escribe de forma matricial: **A X = b**, donde A es una matriz de m filas y m columnas; b y X son vectores de dimensión m. Si existe la inversa de A: A^{-1} entonces se puede calcular el vector incógnita $\mathbf{X=A^{-1}\ b}$

Un SEL cumple con las condiciones a. y b. en correspondencia con el cumplimiento del Teorema de Kronecker-Capelli (Polyanin, 2017).

Cuando un SEL no cumple la tercera condición se dice que es Mal Condicionado y detectar estos casos es un asunto de primera importancia. Si un problema matemático conduce a un SEL que para los datos dados no tiene solución, o esta existe pero no es única, o existe la solución única pero el SEL es mal condicionado, se dice que se trata de un Problema Mal Planteado.

Según (Legrá Lobaina, 2017) para obtener estimaciones puntuales (A,U,Θ) es necesario obtener los valores de cierto ponderadores resolviendo un SEL cuyos datos (la matriz A y el vector b) pueden tener elementos medidos o calculados con cierto nivel de error (falta de exactitud). Estos ponderadores son, además, elementos claves para aproximar $\mathcal{L}_e$, el error de las estimaciones (A,U,Θ), según se ha argumentado en el trabajo de Legrá Lobaina (2018).

En cualquier caso, quienes estiman deberán comprobar que el SEL tiene solución única y que no sea Mal Condicionado porque en ese último caso, en dependencia de la exactitud de los elementos de A y de b, los ponderadores buscados podrían tener diferencias importantes respecto a sus valores correctos. En lo que sigue se le denominará Análisis de Sensibilidad al estudio del condicionamiento de un SEL cuando se estima (A,U,Θ).

El objetivo del presente trabajo es sistematizar las bases teóricas y prácticas para realizar el Análisis de la Sensibilidad durante las estimaciones (A,U,Θ).

2 ALGORITMO GENERAL PARA ESTIMAR (A,U,Θ)

Las estimaciones (A,U,Θ) pueden realizarse de dos maneras que son equivalentes respecto al resultado. Sin perder generalidad, en la Figura 1, siguiendo la notación de Legrá Lobaina (2017,2018), se muestra el algoritmo correspondiente a la Clase ΘU.

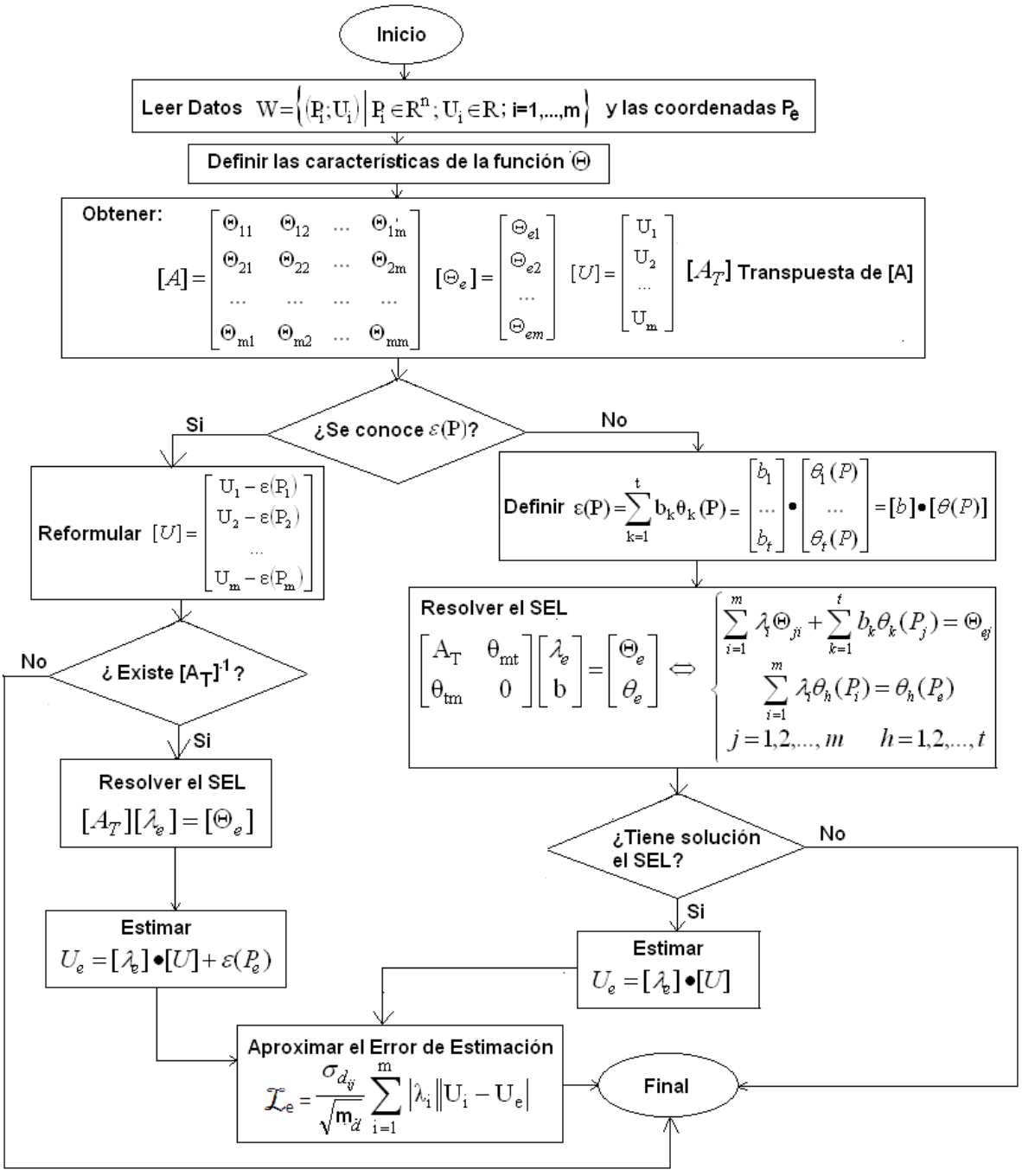

Figura 1: Algoritmo para estimar U_e y aproximar $\mathcal{L}_e$ para la Clase ΘU de los estimadores (A,U,Θ).

Para la Clase UΘ el algoritmo es semejante pero cambian algunas fórmulas tal como se explica en el trabajo de Legrá Lobaina (2017). Particularmente en los SEL la matriz $[A_T]$ es sustituida por la matriz [A], el vector independiente del SEL se tomar de otra manera y U_e se calcula mediante otra fórmula.

HOLOS ISSN 1807 - 1600

3 APUNTES SOBRE LOS SISTEMAS DE ECUACIONES LINEALES MAL CONDICIONADOS

Diversos autores presentan en sus textos la cuestión del mal condicionamiento (en idioma inglés: *ill-conditioned*) de un SEL del tipo A X = b. Siguiendo el texto de Isaacson (1994), serán recordados algunos conocimientos básicos sobre estos temas.

Norma de un Vector de R^n

Sea el vector $X \in R^n$. Se define la norma de ||X|| como un único valor real que cumple las propiedades:

a. $||X|| \geq 0$. Además, $||X||=0$ si y solo si X es el vector nulo.

b. $|| k X || = |k| \, ||X||$ para $k \in R$.

c. $||X+Y|| \leq ||X|| + ||Y||$ para $Y \in Rn$.

Un ejemplo de norma de vector es:

$$||X||_{oo} = \underset{i=1..n}{Max} |x_i| \quad (1)$$

Norma de una Matriz Real Cuadrada de orden n

Sea M_n el conjunto de todas las matrices reales y cuadradas de orden n y sean $A, B \in M_n$. Se define la norma ||A|| como un único valor real que cumple las mismas condiciones que la norma de un vector y además se debe cumplir la propiedad:

d. $||A \, B|| \leq ||A|| \, ||B||$.

Un ejemplo de norma de matriz es:

$$||A||_{oo} = \underset{j=1..n}{Max} \left(\sum_{i=1}^{n} |a_{ij}| \right) \quad (2)$$

Comportamiento de la Solución de un SEL Perturbado

Sea el sistema de ecuaciones lineales:

$$A \, X = b \quad (3)$$

Considérese que el vector b es perturbado mediante la adición del vector δ_b cuya norma es pequeña y la matriz A se perturba adicionándole la perturbación δ_A cuya norma también es pequeña. Surge la pregunta: ¿en qué medida se perturba la solución X?

Si el SEL Perturbado se describe por la expresión:

$$(A+ \delta_A) (X+ \delta_X) = b+ \delta_b \quad (4)$$

Entonces, la pregunta queda redactada de otra forma ¿Será pequeña la perturbación δ_X?

En el Teorema 3 de la página 37, Isaacson (1994) demuestra que si se define a:

$$\mu(A) = \left\| A^{-1} \right\| \, \|A\| \quad (5)$$

Y además:

$$\|\delta_A\| < \frac{1}{\|A^{-1}\|} \tag{6}$$

Entonces se cumple que:

$$\frac{\|\delta_X\|}{\|X\|} \leq \frac{\mu(A)}{1-\mu(A)\frac{\|\delta_A\|}{\|A\|}}\left(\frac{\|\delta_b\|}{\|b\|}+\frac{\|\delta_A\|}{\|A\|}\right) \tag{7}$$

A $\mu(A)$ se le denomina Número de Condición de la matriz A.

De la expresión (7) se demuestra que la magnitud de los cambios relativos de X dada por:

$$\frac{\|\delta_X\|}{\|X\|} \tag{8}$$

está acotada por el resultado de multiplicar a (9), la suma de los cambios de b y A:

$$\frac{\|\delta_b\|}{\|b\|}+\frac{\|\delta_A\|}{\|A\|} \tag{9}$$

por la expresión (10):

$$\frac{\mu(A)}{1-\mu(A)\frac{\|\delta_A\|}{\|A\|}} \tag{10}$$

Es obvio que en la medida en que la expresión (10) sea menor entonces (8) también será menor, esto es si (9) permanece constante.

El lector puede notar que se garantiza que (10) es no negativa ya que la expresión (6) es equivalente a la expresión:

$$\mu(A)\frac{\|\delta_A\|}{\|A\|} < 1 \tag{11}$$

Se puede demostrar entonces que $\mu(A)$ tiene valores admisibles en el intervalo $\left[1; \frac{\|A\|}{\|\delta_A\|}\right)$ y los valores de (10) están en el intervalo $\left[\frac{\|A\|}{\|A\|-\|\delta_A\|}; \infty\right)$.

En muchos problemas prácticos se tiene que la magnitud $\|\delta_A\|$ es nula y la expresión (10) se reduce a $\mu(A)$ por lo que (7) se escribe:

$$\frac{\|\delta_X\|}{\|X\|} \leq \mu(A)\frac{\|\delta_b\|}{\|b\|} \tag{12}$$

Para propósitos prácticos es necesario responder a la siguiente pregunta: ¿cuál es el valor límite de $\mu(A)$ para decidir si un SEL es bien o mal condicionado?

Isaacson (1994) asegura que si el Número de Condición, que se calcule con una norma dada, no es demasiado grande (*is not too large*), entonces el SEL es bien condicionado.

En opinión de Burden (2005), para establecer el mal condicionamiento de un SEL se calcula $\mu(A)$ y se comprueba que esté cerca de 1; para este autor si se comprueba que $\mu(A)$ es significativamente mayor que 1, entonces el SEL es mal condicionado.

Álvarez (2007) establece que el límite entre un SEL bien condicionado y un SEL mal condicionado es el valor $\mu(A)$ =10 utilizando la norma (2).

En el texto de Faddeev (1963) señala que una matriz A es mal condicionada cuando le corresponde A^{-1} inestable lo cual significa que cuando se producen pequeños cambios en algunos elementos de A entonces se producen grandes cambios en algunos elementos de A^{-1}. En este texto se describen (páginas 126 y 127) los dos números de condición de Turing y los dos números de condición de Todd obtenidos para diferentes normas matriciales.

Lapidus (1962) explica que una matriz bien condicionada es aquella que tiene su número de condición N cerca de 1 y define: $N = \frac{N(A)N(A^{-1})}{n}$ donde $N(A) = \sqrt{traza\left(A^T A\right)}$ y la traza de una matriz A se calcula como la suma de los elementos de la diagonal principal de A.

Es conveniente que el lector conozca que algunos autores (Lapidus, 1962 y Suárez Alonzo, 1986) señalan que valores pequeños del determinante de la matriz A (denotado |A|) indican que posiblemente AX=b es mal condicionada.

Suárez Alonso (1986) explica, además, que valores pequeños del determinante normalizado, denotado $|A|_N = \frac{|A|}{a_1 a_2 .. a_n}$ y donde a_i es el módulo de la fila i de la matriz A, indican que AX=b es mal condicionado. Señala que si $|A|_N$ es un infinitésimo del orden 10^{-k} entonces un cambio de la k-ésima o anterior cifra significativa de cualquiera de los coeficientes de A puede producir cambios del orden 10^k en la solución (página 75).

Sarmiento Almeida (2010) da una demostración de la expresión (12) y además asegura que para el caso en que las perturbaciones se producen en A y no en b se cumple la desigualdad:

$$\frac{\|\delta_X\|}{\|X + \delta_X\|} \leq \mu(A) \frac{\|\delta_A\|}{\|A\|} \qquad (13)$$

Y refiere que cuanto más cercano a 1 sea $\mu(A)$, mejor condicionado estará el SEL.

Como puede verse, este es un asunto donde la teoría, la práctica, la intuición y el sentido común, tienen que intervenir. El autor del presente trabajo sugiere que un sistema satisfactorio de criterios se puede establecer usando la norma matricial (2):

Criterio de Clasificación del Condicionamiento de un SEL (CCC)

- Excelentemente condicionado (EXC) si: $1 \leq \mu(A) \leq 10$
- Muy Bien condicionado (MBC) si: $10 < \mu(A) \leq 100$
- Bien condicionado (BIC) si: $100 < \mu(A) \leq 500$
- Aceptablemente condicionado (ACC) si: $500 < \mu(A) \leq 1000$

HOLOS ISSN 1807 - 1600

- Regularmente condicionado (REC) si: 1000< $\mu(A)$ ≤ 10000
- Casi Mal condicionado (CMC) si: 10000< $\mu(A)$ ≤ 100000
- Mal condicionado (MAC) si: 100000< $\mu(A)$ ≤ 1000000
- Pésimamente condicionado (PEC) si: 1000000< $\mu(A)$

Ejemplo A:

Sean los siguientes datos:

Tabla 1: Un ejemplo ilustrativo de 7 puntos para n=2

X_1	X_2	U
0	0	1
1	0	1
0	1	1
1	1	2
0,5	0,5	1,5
0.1	0,8	1,3
0,7	0,2	1,6

Asumiendo que δ=0 y ε(P)=0, se analizan para varios modelos de Θ(d) el valor del número de condición $\mu(A)$ correspondiente y se clasifica mediante el CCC descrito anteriormente.

Tabla 2: Clasificación de los SEL de diferentes estimadores para datos de la Tabla 1. Se usa la norma (2)

Estimador	$\mu(A)$ (aproximado)	Clasificación
• Inverso de potencia 1,45 de la distancia	1	EXC
• FBR Spline de Placa Delgada, R= 0,0114285	22,314	MBC
• UPD con potencia 1,45	110,025	BIC
• Kriging Puntual (Variograma Potencial con pendiente 1 y potencia 1,95 y efecto pepita nulo)	1886,956	REC
• Estimador m-Funcional donde $\upsilon(P) = 0,1 + 0,01\ X_1 + 0,07\ X_2$ y $T_W = \{\psi_1(P)=1;\ \psi_2(P)=\upsilon(P);\ \psi_3(P)= \upsilon(P)^2;\ \psi_4(P)= \upsilon(P)^3;\ \psi_5(P)= \upsilon(P)^4;\ \psi_6(P)= \upsilon(P)^{0,5}\ \psi_6(P)= \upsilon(P)^{1,5}\}$.	7674508316	PEC

A continuación, para el caso del estimador m-funcional de la Tabla 2 se estimarán mallas de 40×40 nodos, tomando (mediante redondeo) 12, 10, 9 y 8 el número de cifras decimales de los elementos del vector [b]. En particular en la Tabla 3 se muestran las variaciones de los valores del vector $[\lambda_e]$ según disminuyen las cifras decimales de [b] cuando se estima U_e para las coordenadas $(X_1;X_2)=(0,23;0,15)$ lo cual confirma que el SEL es Mal Condicionado.

Las variaciones observadas del vector $[\lambda_e]$ determinan la calidad de los modelos de malla tal como se ilustra en la Figura 1 donde se observa que en la medida en que disminuyen las cifras decimales disponibles, se pierde la continuidad de los modelos lo cual significa precisamente que en estos modelos defectuosos, a pequeños cambios de las coordenadas de los puntos de la malla se producen cambios importantes en los valores estimados.

HOLOS ISSN 1807 - 1600

Tabla 3: Variación del vector [λ_e] cuando al estimar Ue en las coordenadas (0,23;0,15) las cifras decimales (CD) disminuyen en los elementos del vector [b]

[λ_e]	CD=12	CD=10	CD=9	CD=8
λ_{e1}	-0,032335962448	-0,030910616786	-0,053179172746	-0,10501347527
λ_{e2}	-0,063405582388	-0,07307796058	0,074286074308	0,433607444099
λ_{e3}	0,114853493833	0,128984114521	-0,087267785517	-0,611716041592
λ_{e4}	-0,15811026193	-0,170809082768	0,024693907582	0,494658655747
λ_{e5}	0,409730726135	0,420373392663	0,255398744977	-0,13611035239
λ_{e6}	0,713291349957	0,706687453789	0,809473493933	1,051134733111
λ_{e7}	0,015697251547	0,018473712234	-0,023684239685	-0,126839950632

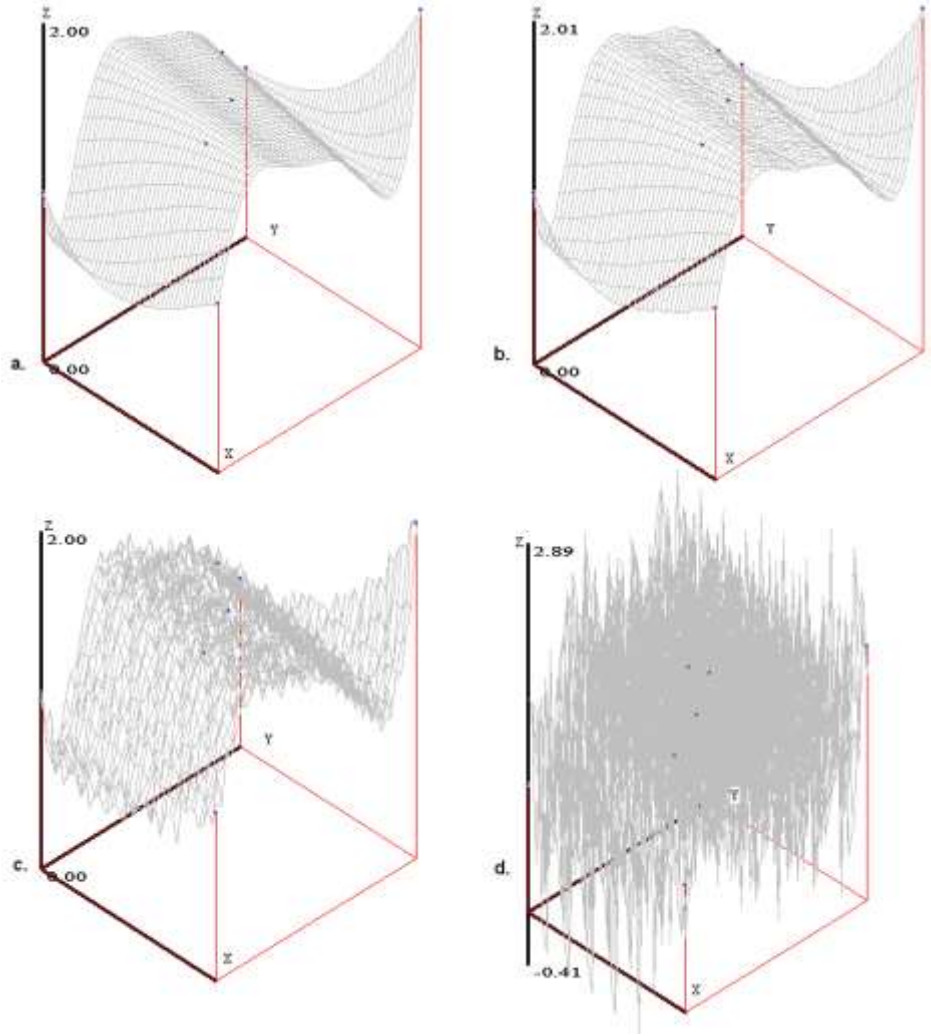

Figura 1: Modelos de malla para el estimador m-Funcional de la Tabla 1.

a. **Con 12 cifras decimales en el vector [b]. μ_e = 0,0205604; σ_e = 0,0194669; C_V=94,68%**

b. **Con 10 cifras decimales en el vector [b]. μ_e = 0,0205719; σ_e = 0,0194936; C_V=94,76%**

c. **Con 9 cifras decimales en el vector [b]. μ_e = 0,0242491; σ_e = 0,0230198; C_V=94,93%**

d. **Con 8 cifras decimales en el vector [b]. μ_e = 0,201245; σ_e = 0,239676; C_V=119,1%**

Como puede observarse, en el pié de la Figura 1 también se informan los valores de la media aritmética $\boldsymbol{\mu_e}$ y de la desviación estándar $\boldsymbol{\sigma_e}$ de los errores $\mathcal{L}_e$ calculados al estimar cada punto de la malla para cada caso. Nótese que como tendencia, a medida que el modelo se deteriora estas medidas μ_e, σ_e y el coeficiente porcentual de variación $C_V = 100\frac{\sigma_e}{\mu_e}$ aumentan.

La prueba que se ha realizado para el quinto modelo de la Tabla 2 usando los datos de la Tabla 1, puede ser realizada para cualquier modelo que el investigador desee analizar. Por ejemplo, puede comprobarse que para los cuatro primeros modelos de la Tabla 2, las cifras decimales pueden ser bajadas hasta 3 sin que el modelo sufra cambios visualmente perceptibles.

Es conveniente precisar explícitamente que: cuando el número de condición $\mu(A)$ es cercano a 1 se **asegura** un buen condicionamiento del SEL; en el caso de que $\mu(A)$ es muy grande respecto a 1, solo se puede esperar que para pequeñas variaciones no nulas $\frac{\|\delta_b\|}{\|b\|}$ y $\frac{\|\delta_A\|}{\|A\|}$ existe la **posibilidad** de que se produzcan variaciones importantes de $\frac{\|\delta_X\|}{\|X\|}$.

Es por esa razón que puede proponerse que también sea realizada una Prueba Exploratoria de Sensibilidad (PES) que consiste en determinar la Media Aritmética y la Desviación Estándar de los errores absolutos relativos porcentuales de los valores de los nodos de las mallas obtenidos para, por ejemplo, CD=10, CD=9 y CD=8 con respecto a CD=12. Para el caso que se estudia se obtienen los resultados de la Tabla 4:

Tabla 4: Prueba Exploratoria de Sensibilidad usando Mallas de 40×40. Se calcula $e_i=100|U_{12i}-U_{ki}|/|U_{12i}|$

K	Media aritmética de e_i en %	Desviación Estándar de e_i en %
10	0,343251472	0,260778996
9	4,21156101	3,23820814
8	38,2546527	30,6651898

Los resultados de la Tabla 4 muestran que la exactitud de los resultados se deteriora rápidamente si los elementos del vector [b] pierden por redondeo algunas pocas de sus últimas cifras decimales.

¿Qué hacer para estimar correctamente (A,U,Θ)?

En cualquier caso las **medidas básicas** para garantizar que la estimación es precisa son:

a. Seleccionar y utilizar variables y constantes con un número alto de cifras significativas. Este aspecto ha sido considerado notablemente por quienes diseñan los software desarrolladores y matemáticos. Por ejemplo, desde la década de los 80 del siglo XX, BORLAND INTERNATIONAL INC (1988, página 34) puso a disposición de los desarrolladores los tipos de números *Real* con 11 cifras significativas. En el siglo actual los software ponen a disposición de sus usuarios una importante cantidad de cifras significativas tal como muestra la aplicación Mathematica que reporta el uso en sus cálculos de tantas cifras decimales como necesite el usuario (Wolfram, 2003, página 33).

Sin embargo el lector debe saber, además, que en una aplicación informática de cálculo: a medida que se usan más cifras significativas, entonces se requiere de mayor cantidad de recursos computacionales, incluyendo el tiempo de cálculo; esta situación obliga a que se analicen los pro y los contra de la decisión que se tome.

b. Seleccionar algoritmos matemáticos eficientes con respecto al número de operaciones aritméticas y que eviten en lo posible la propagación de errores. Para los SEL es muy popular y eficiente el Método de Gauss; en la página 352 del texto de Burden (2005) se presenta una versión completa (proceso directo y proceso inverso) de este método y se precisa que para resolver un SEL de matriz A cuadrada de orden n se realizan:

- $\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}$ sumas y restas
- $\frac{n^3}{3}+n^2-\frac{n}{3}$ productos y cocientes

El lector avisado puede darse cuenta de que resolver un SEL de alto orden n puede ser una tarea larga y arriesgada en el sentido de la propagación de errores. Por ejemplo si n=1000 se realizarán 333832500 sumas y restas, y 334333000 productos y cocientes que suman más de 668 millones de operaciones aritméticas.

Cabe precisar que cuando se aplica el proceso directo del Método de Gauss a un SEL con valores nulos en la diagonal principal se hacen necesarias transformaciones controladas de intercambios de filas y columnas. También se realizan estas transformaciones cuando se aplica una Estrategia de Pivote que según Álvarez (2007) contribuye a minimizar la propagación de errores que se pueden producir cuando se calculan cocientes donde el divisor tiene un valor absoluto muy pequeño. En el texto citado se describen tres tipos de Estrategia de Pivote, a saber: Elemental, Parcial y Total (páginas 150-151) que también aumentarán el tiempo de cálculo.

c. Seleccionar modelos matemáticos que, respondiendo a los requerimientos del problema planteado, tengan las menores, complejidad y dimensiones posibles. Para el problema de los estimadores (A,U,Θ) hay dos cuestiones importantes:
   - La selección de la función Θ, cuestión que se explicará detalladamente en un próximo trabajo.
   - En vez de usar todos los datos de W en cada estimación (en este caso se dice que se utiliza un Soporte Global para la estimación), se seleccione para cada estimación un subconjunto específico de puntos adecuados (en este caso se dice que se utiliza un Soporte Local o Compacto para la estimación). En el punto 5 del presente trabajo se explicarán algunas técnicas para seleccionar el Soporte Local.

Después de tomar las medidas anteriores, para cada estimación (A,U,Θ) se deberá calcular su número de condición $\mu(A)$ y si su valor es apreciado como significativamente mayor que 1 (por ejemplo, mayor que 1000) entonces se tienen tres opciones:

1. Volver a analizar las medidas básica a, b, c.
2. Redefinir el Soporte de la Estimación
3. Hacer Pruebas Exploratorias de Sensibilidad

4. Hacer Pruebas de Validación Cruzada que permiten estudiar la eficacia y eficiencia de un estimador.

4 APUNTES SOBRE LAS PRUEBAS DE VALIDACIÓN CRUZADA

Ya se ha comentado (Legrá, 2018) que la Validación Cruzada es una vía para aproximar el error general de las estimaciones mediante la calificación del estimador. Una Prueba Simple de Validación Cruzada de un estimador (A,U,Θ) sobre los datos W comienza con establecer los conjuntos no vacíos de puntos W_D (datos para estimar) y W_C (datos donde se estima y se comprueba la eficacia del estimador) tal que: $W_D \cap W_C = \varnothing$ y $W_D \cup W_C = W$. Estos conjuntos W_D y W_C deben tomarse adecuadamente de manera que las coordenadas de los m_D puntos de W_D y las coordenadas de los $t=m_C$ puntos de W_C representen local y globalmente al conjunto W.

La técnica simple denominada ***Holdout*** (Lendasse, 2003; Arlot, 2010) consiste en obtener, usando un único modelo del estimador (A,U,Θ), todos los valores U_{ei} (i=1,…,m_C) de la variable dependiente U en las coordenadas P_i de los puntos de W_C utilizando como datos a los puntos de W_D. Los resultados de las estimaciones se comparan con los valores de U conocidos en los puntos de W_C y en particular con las $t=m_c$ diferencias e_i entre U_i y U_{ei} (i=1,2,…,t) obtenidas para calificar la calidad zonal y general del estimador.

Existen otras técnicas más complejas para realizar Pruebas de Validación Cruzada entre las cuales se destacan (Chilés, 1999; Díaz Viera, 2002; Lendasse, 2003; Arlot, 2010; Zhang, 2015):

Técnica *leave one out*: Se trata de un algoritmo recursivo donde el conjunto W_C es unitario y cambia en cada una de las t≤m pruebas o iteraciones donde para cada una de ellas se particulariza un modelo basado en los datos del $W_D = W/W_C$ correspondiente a cada iteración.

En otras palabras, se realiza la estimación U_{ei} para i=1,2,…,t tomando en cada caso el conjunto unitario $W_C=\{(P_i;U_i)\}$ y, al finalizar el cálculo de las t estimaciones, para cada coordenada P_i de los puntos seleccionados de W, se tienen dos valores: el valor medido o conocido U_i y el valor estimado U_{ei} cuya cercanía caracteriza localmente la calidad de la estimación (A,U,Θ).

El conjunto de las t diferencias e_i entre U_i y U_{ei} (i=1,2,…,t) permite estudiar la calidad zonal y general del estimador.

La implementación más usual de esta técnica es el caso donde t=m o sea la prueba se realiza para todos los elementos de W.

Técnica *leave one out* aleatoria: Es similar al caso anterior donde $W_C=\{(P_i;U_i)\}$ es un conjunto unitario pero solo se estiman los valores de U_i en t<m puntos P_i de W los cuales son tomados aleatoriamente (generalmente en un muestreo sin reposición). El número de modelos que se analizan es t.

Técnica K-Fold o K-Pliegues: Del conjunto W se toman K subconjuntos disjuntos (pliegues) W_1, W_2,… , W_K donde al menos uno de ellos está formado por más de un elemento aunque por razones de representatividad es práctica común tomar aproximadamente el mismo cardinal para todos los K subconjuntos; estos se toman al azar o se definen mediante criterios de clasificación vinculados con el objeto que se investiga. Los subconjuntos W_1, W_2,… , W_K pueden cumplir que $\bigcup_{j=1}^{K} W_j = W$ en cuyo caso t=m. Por el contrario puede cumplirse que $W / \bigcup_{j=1}^{K} W_j \neq \varnothing$.

Para esta técnica solo se realizan K iteraciones donde en aquella de orden j se estiman todos los valores U_e para las coordenadas de los puntos de W_j a partir de los datos de W que no pertenezcan a W_j, de manera que después de realizar todas las iteraciones se tienen t≤m diferencias e_i (i=1,…,t) que permiten estudiar la calidad del estimador.

Para cada pliegue j pueden ser calculadas las medias de sus valores e y en este caso la calidad del estimador se estudia con estos t=K valores medios.

Para cualquiera de las técnicas descritas se calculan los residuos $e_i = U_{ei} - U_i$ y con el fin de calificar la calidad del modelo se analizan el histograma de la variable $\{e_i\}$, el gráfico de dispersión entre $\{U_{ei}\}$ y $\{U_i\}$ y el cumplimiento de los siguientes criterios (Díaz Viera, 2002):

a. La media aritmética $\frac{\sum_{i=1}^{t} e_i}{t}$ es cercana a 0.

b. El error medio cuadrático $\frac{\sum_{i=1}^{t} e_i^2}{t}$ es pequeño.

c. El coeficiente de correlación lineal entre las variables $\{U_i; U_{ei}\}$ donde i=1,…,t debe ser cercano a 1.

Para ilustrar las técnicas de validación cruzada se aplicará la denominada ***leave one out*** para los datos de la Tabla 1 y los estimadores descritos en la primera columna de la Tabla 2. El lector ahora puede usar los resultados que se obtengan como parámetros de comparación de la eficacia de pronóstico de cada uno de los modelos.

Tabla 5: Resultados de la Validación Cruzada para cinco estimadores

Estimador	$\frac{\sum_{i=1}^{t} e_i}{t}$	$\frac{\sum_{i=1}^{t} e_i^2}{t}$	Coeficiente de Correlación de $\{U_i; U_{ei}\}$
Inverso de potencia 1,45 de la distancia	0.00203584	0.166107	-0.532235
FBR Spline de Placa Delgada, R= 0,0114285	0.177964	6.35871	-0.540256
UPD con potencia 1,45	-0.13247	0.298966	-0.0939778
Kriging Puntual (Variograma Potencial con pendiente 1 y potencia 1,95 y efecto pepita nulo)	0.115359	0.380534	-0.0894656
Estimador m-Funcional donde $\upsilon(P) = 0,1 + 0,01\ X_1 + 0,07\ X_2$ y $T_W = \{\psi_1(P)=1; \psi_2(P)=\upsilon(P); \psi_3(P)= \upsilon(P)^2; \psi_4(P)= \upsilon(P)^3; \psi_5(P)= \upsilon(P)^4; \psi_6(P)= \upsilon(P)^{0,5}\ \psi_6(P)= \upsilon(P)^{1,5}\}$.	0.904078	4.85504	0.049611

5 TÉCNICAS PARA DEFINIR SOPORTES LOCALES

Existen tres razones importantes para usar soporte local en lugar de soporte global cuando se realizan estimaciones (A,U,Θ):

1. Para los problemas que se modelan mediante mallas con estimadores (A,U,Θ), el comportamiento de la variable U en cierta coordenada P_e podría depender significativamente solo de los puntos (P_i;U_i) que cumplan ciertas propiedades.
2. Para obtener estimaciones (A,U,Θ) se hace necesario resolver SEL de orden n×n donde n es igual o mayor que el número de datos que conforman el soporte. Los soportes locales conducen a SEL de menor orden que desde el punto de vista numérico se manejan con menos operaciones y, probablemente, menos errores.
3. Con una mayor cantidad de puntos en el soporte de estimación, aumenta la probabilidad de aparición de valores negativos en los coeficientes λ_i debido al fenómeno que en la literatura geoestadística se le conoce como Apantallamiento o Efecto Pantalla (Legrá Lobaina, 2010).

Por ejemplo, sean los m=79 datos (X_i;Y_i;N_i) de un conjunto W que se muestra en la Figura 2. A continuación se obtendrá el valor estimado de N en el punto P_0=(X_0;Y_0)=(165;165) mediante el método UPD tomando p=1,45 y factor de suavización nulo.

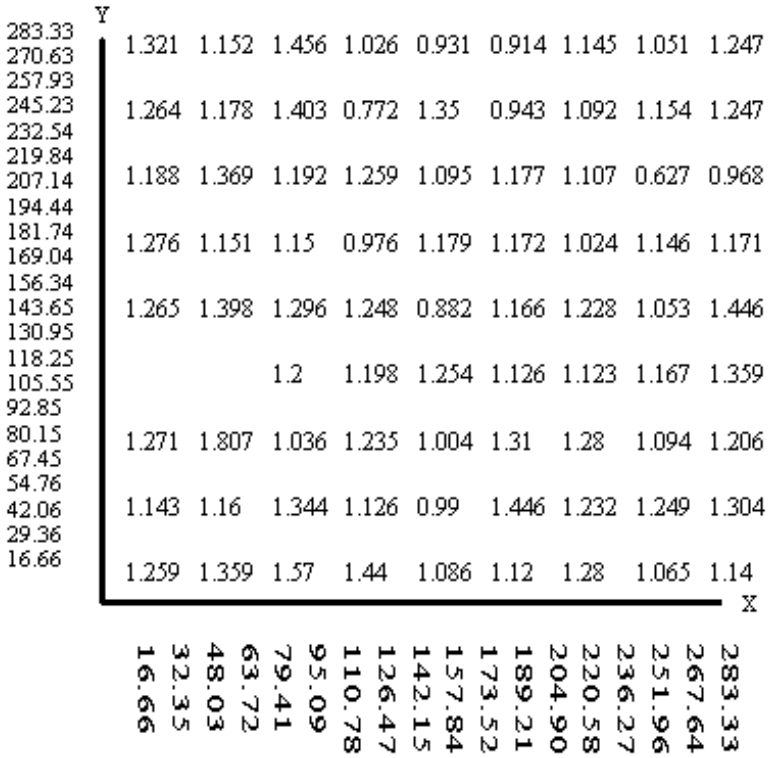

Figura 2: Valores de N en las coordenadas (X;Y) que definen al conjunto de datos W

Para ilustrar el comportamiento del Efecto Pantalla se tomarán varios soportes de estimación formado por los K puntos de W más cercanos a (X_0;Y_0) donde K toma los valores: 2, 3, 4, 5, 6, 16 y 20. En cada caso los valores calculados de λ_i se muestran en la Figura 3.

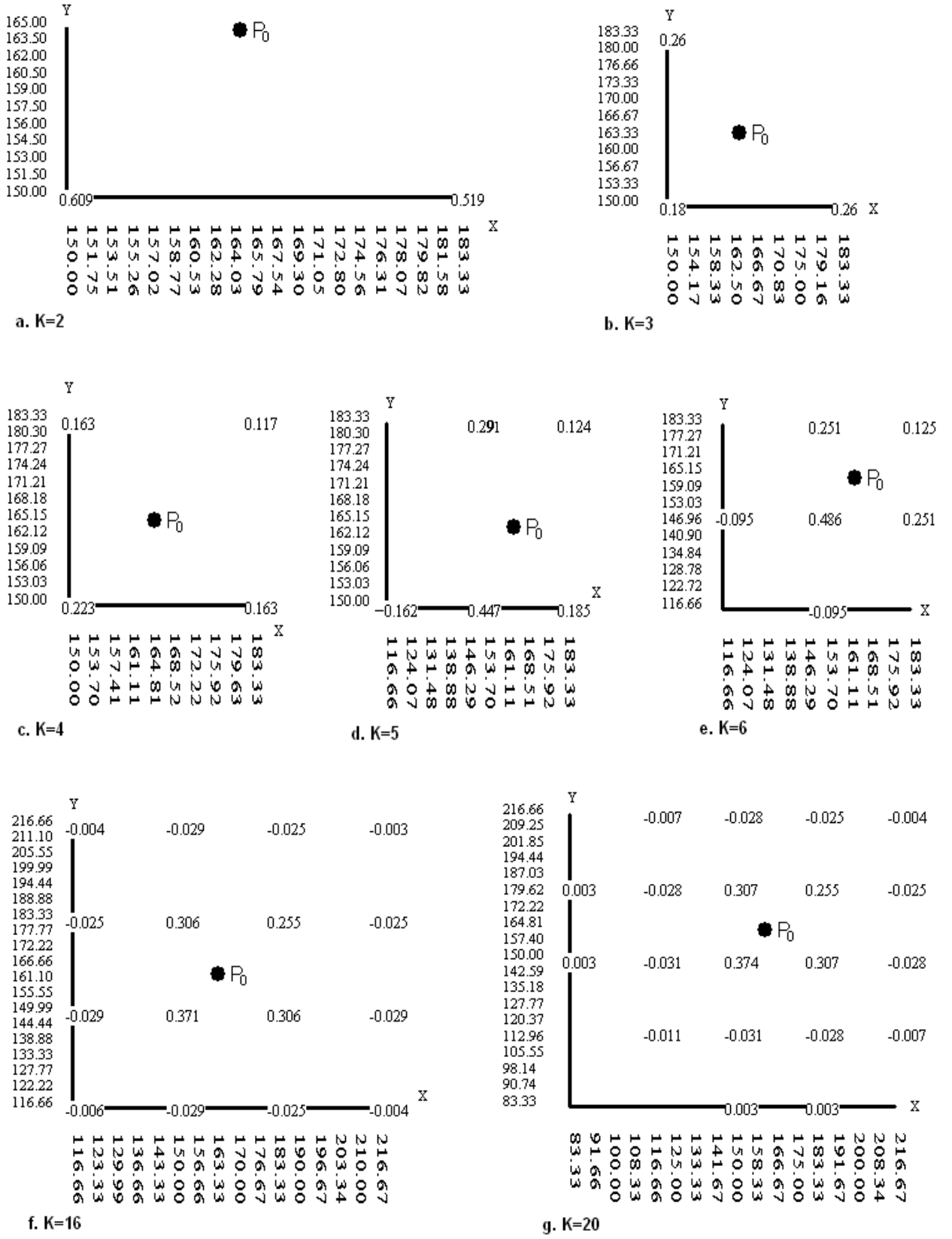

Figura 3: Valores de λ_i en cada punto de soporte cuando se realiza la estimación de N en P_0

Puede notarse que cuando se consideran dentro del soporte algunos puntos alejados de P_0 (apantallados por los puntos más cercanos a P_0), entonces una buena parte de los valores de λ_i para los puntos apantallados son negativos.

En sentido general, la selección de datos del soporte local (*search*) se basa en reglas que permitan que el soporte de cada estimación puntual se obtenga como el conjunto de datos de W que pertenecen a una "ventana móvil" o "vecindad de búsqueda" o "zona de influencia". Las reglas para crear estos soportes locales generalmente se basan en propiedades de vecindad geométrica, en propiedades de otro tipo que cumple el objeto que se investiga o en combinaciones de estos dos casos.

Criterios Geométricos:

1. Se toman los n_{sop} datos más cercanos al punto de coordenadas P_e.
   Una variante es dividir toda la vecindad de P_e en sectores geométricos y se toman cantidades de datos del soporte más o menos iguales en cada sector de manera que la suma de los datos en cada sector sea n_{sop}.
2. Se toman los puntos que están dentro de una vecindad de centro en P_e y de radio R_{sop} denominada Bola de un Espacio Métrico (Hinrichsen y Fernández., 1977).
   Nuevas variantes se obtienen al tomar vecindades diferentes a las Bolas de un Espacio Métrico, por ejemplo:
   a. Para el caso de n=2, en lugar de un círculo tomar una elipse, un triángulo, un cuadrado, un rectángulo, etc.
   b. Para n=3, en lugar de la esfera tomar una elipsoide, un paralelepípedo, etc.

Criterios No Geométricos

Se toman en el soporte local solo aquellos puntos $(P_i;U_i)$ que cumplen una o varias propiedades de interés. Por ejemplo, si U es la concentración del Ni en mineral laterítico entonces puede ser de interés estimar U teniendo en cuenta solo aquellas datos que pertenezcan a muestras minerales semejantes. En ocasiones también es conveniente dejar fuera del soporte aquellos datos donde U_i sea considerado atípico.

Criterios Mixtos

Es el caso de incluir en los soportes puntos de interés por su cercanía geométrica pero que además cumplan con las propiedades no geométricas exigidas.

El lector seguramente se habrá percatado de que cuando se usa un soporte global solo es necesario resolver un único SEL de orden igual o superior a la cantidad m de datos de W y que cuando se usan soportes locales entonces se resuelven varios SEL distintos de pequeño orden. Entonces, queda establecido que en el proceso de obtención del modelo es necesario valorar expresamente el costo-beneficio de seleccionar uno u otro tipo de soporte para lograr estimaciones de calidad de manera eficiente.

6 PROPUESTA DE UN ALGORITMO PARA OBTENER MALLAS MEDIANTE ESTIMACIONES (A,U,Θ)

En la Figura 2 se muestra un algoritmo básico para estimar todos los nodos de una malla el cual tiene en cuenta los aspectos que se han tratado en este trabajo.

El lector puede notar que la instrucción denominada ***Diseñar y ejecutar pruebas de validación cruzada*** no está descrita detalladamente pero debe asumirse que se trata de aplicar el algoritmo que sigue después de la instrucción ***¿Resultados aceptables?*** para ciertas mallas

especiales que dependen de la técnica seleccionada para validar. En ese caso, para los resultados obtenidos se realizan los análisis escritos en el epígrafe 4.

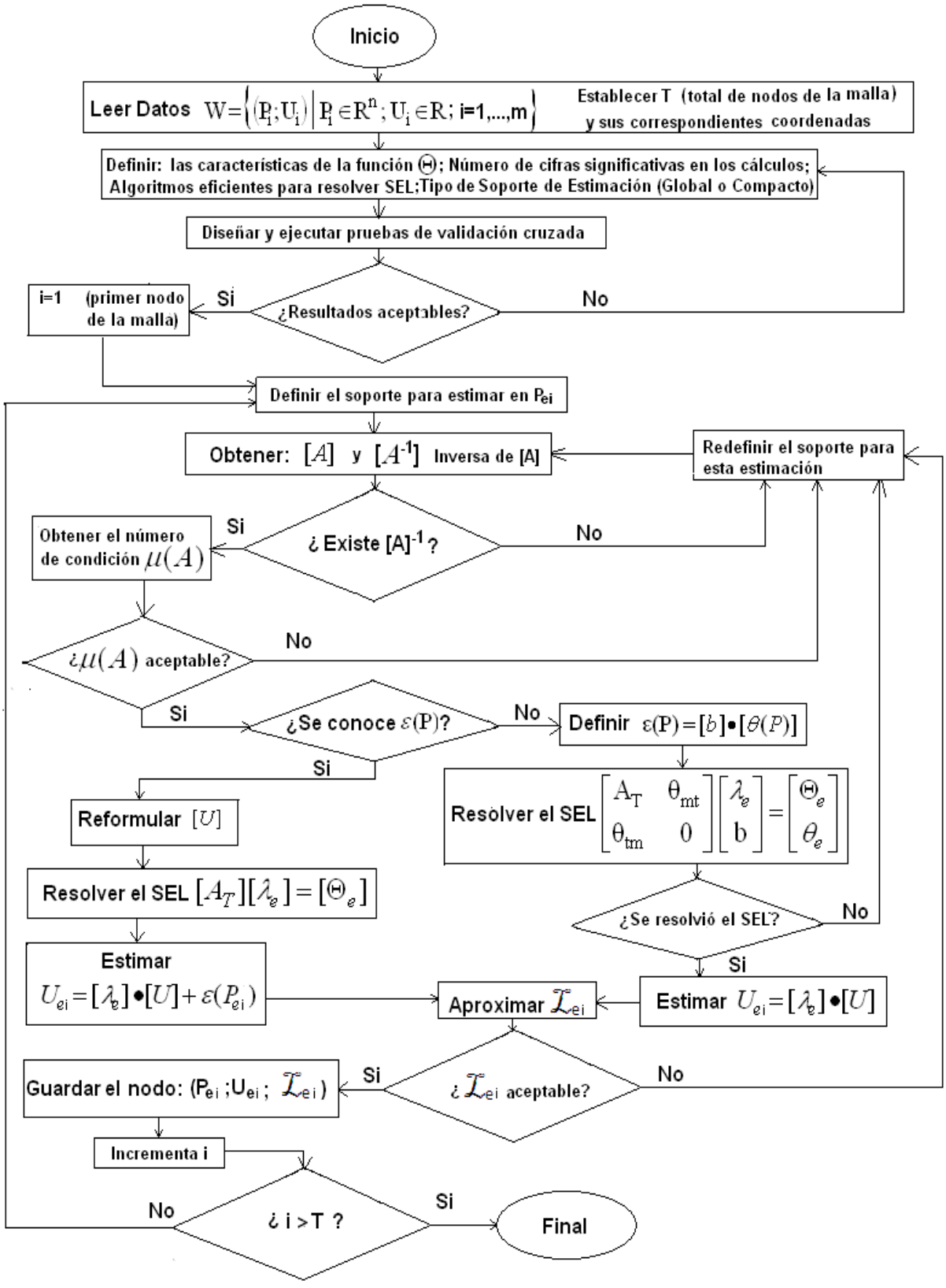

Figura 2: Algoritmo para estimar U en nodos de una malla mediante estimadores (A,U,Θ) de la Clase ΘU

7 CONCLUSIONES

Se ha comprobado que el Mal Condicionamiento de la matriz A aumenta la sensibilidad de los estimadores (A,U,Θ) y esto ocasionalmente provoca que pequeños cambios en el valor de los datos, produzcan grandes cambios en los resultados.

Ha sido determinado que mediante los Números de Condición, los errores de estimación y las pruebas de Validación Cruzada pueden ser detectados y evaluados los problemas de mal condicionamiento.

Se han propuesto criterios para resolver los problemas de mal condicionamiento de los estimadores (A,U,Θ) mediante la selección de modelos adecuados y el uso de suficientes cifras decimales significativas en los cálculos, algoritmos eficientes y soportes compactos.

Como resumen es presentado un diagrama que facilita la comprensión del algoritmo para resolver adecuadamente los problemas de mal condicionamiento de los estimadores (A,U,Θ).

8 REFERENCIAS BIBLIOGRÁFICAS

Álvarez Blanco, M., Guerra Hernández, A. y Lau Fernández, R. (2007). *Matemática Numérica*. La Habana: Editorial Félix Varela.

Arlot, S. y Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40-79. DOI: 10.1214/09-SS054.

Borland International Inc (1988). *Turbo Pascal Reference Manual. Version 3.0*. California: ALPHA SYSTEMS CORPORATION. [Recuperado de: http://bitsavers.informatik.uni-stuttgart.de/pdf/borland/turbo_pascal/TURBO_Pascal_Reference_Manual_CPM_Version_3_Dec88.pdf.

Burden, R. L. y Faires, J. D. (2005). *Análisis Numérico*. Séptima Edición. Ciudad Méjico: Editorial Thomson Learning.

Chilés, Jean-Paul y Delfiner, Pierre (1999). *Geostatistics. Modeling Spatial Uncertainty*. Canadá: John Wiley & Sons, Inc.

Díaz Viera, M. A. (2002). *Geoestadística Aplicada*. México: Instituto de Geofísica, UNAM e Instituto de Geofísica y Astronomía, CITMA de Cuba. Recuperado de http://mmc2.geofisica.unam.mx/cursos/geoest/GeoEstadistica.pdf.

Faddeev, D. K. y Faddeeva, V. N. (1963). *Computational Methods of Linear Algebra*. USA: W.H.Freeman & Co Ltd.

Hinrichsen, D. Y Fernández, J.L. (1977). *Topología General*. La Habana: Editorial Pueblo y Educación.

Isaacson, E. y Keller, H.B. (1994). *Analysis of Numerical Method*. New York: Dover Publications, Inc.

Lapidus, L (1962). *Digital Computation for Chemical Engineers*. New York: McGraw Hill Series in Chemical Engineering.

Legrá Lobaina, A. A. (2017). Modelos de malla basados en estimadores (A,U,Θ). *Revista HOLOS*, 33(4), 88-110. DOI:10.15628/holos.2017.5351.

Legrá Lobaina, A. A. (2018). Evaluación del error en estimaciones (A,U,Θ). *Revista HOLOS*, 33(4), 88-110. DOI:10.15628/holos.2017.5351.

Legrá Lobaina, A. A. y Atanes Beatón, D. M. (2010). Variogramas adaptativos: un método práctico para aumentar la utilidad del error de estimación por kriging. *Revista Minería y Geología*, 26(4), 53-78. Recuperado de http://revista.ismm.edu.cu/index.php/revistamg/article/download/63/69

Lendasse, L.; Wertz, V. y Verleysen, M. (2003). *Model Selection with Cross-Validations and Bootstraps-Application to Time Series Prediction with RBFN Models*. Berlín: Springer-Verlag Berlin. Recuperado de: http://www.di.ens.fr/willow/pdfs/2010_Arlot_Celisse_SS.pdf.

Miller, I., Freund, J. y Johnson, R. (2005). *Probabilidades y Estadísticas para ingenieros*. Cuarta Edición. México: Prentice-Hall Hispanoamericana S.A.

Polyanin, A. D. y Manzhirov, A. V. (2007). *Handbook of mathematics for engineers and scientists*. Florida: Chapman & Hall/CRC, Taylor & Francis Group.

Wolfram, S. (2003). *Mathematica book, 5th Edition*. USA: Wolfram Media Inc, USA. Recuperado de: http://deptche.ccu.edu.tw/Chemistry/Chem_Math/Mathematica_V5_Book.pdf.

Zhang, Y. y Yang, Y. (2015). Cross-Validation for Selecting a Model Selection Procedure. *Journal of Econometrics,* 187, 95-112. Recuperado de: http://users.stat.umn.edu/~yangx374/papers/ACV_v30.pdf.