

00584
1
2ci



Universidad Nacional Autónoma de México

FACULTAD DE CIENCIAS

**PRINCIPIO DE SEUDO-OPTIMALIDAD
PARA EL METODO DE REGULARIZACION DE TIONOV
EN LA RESOLUCION DE PROBLEMAS MAL-PLANTEADOS**

por

JESUS LOPEZ ESTRADA

T E S I S

Para obtener el Grado de:

DOCTOR EN CIENCIAS

(MATEMATICAS)

**TESIS CON
FALLA DE ORIGEN**

México, D. F.

1988



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1
24

PRINCIPIO DE SEUDO-OPTIMALIDAD PARA EL METODO DE REGULARIZACION DE TIJONOV EN LA RESOLUCION DE PROBLEMAS MAL-PLANTEADOS

por

Jesús López E.

Resumen. En la resolución de problemas Mal-Planteados a la Hadamard, la metodología de Regularización de Tifonov ha venido jugando, desde su aparición en 1963, un papel central. Tanto es así, que hoy día constituye toda una rama de la Matemática.

Una de las dificultades centrales de esta metodología, no del todo resuelta es la elección del parámetro de regularización. El propósito central del presente trabajo es la propuesta de un criterio nuevo, Principio de Seudo-optimalidad, para la selección de dicho parámetro. Se demuestra que con este principio se obtiene un algoritmo de regularización a la Tifonov con propiedades de convergencia asintóticas "casi" óptimas. Se muestra su superioridad con respecto al Principio de Quasi-optimalidad de Tifonov para el problema del cálculo estable de la matriz inversa generalizada de Moore-Penrose, pues en tal caso el referido criterio de Tifonov no es aplicable.

Por último, se presenta una formulación estadística de nuestro principio, cuya versión en dimensión finita, nos permite finalizar con la propuesta de un algoritmo automático para la resolución numérica de ecuaciones integrales de Fredholm de primera especie, el cual es también un algoritmo automático para la estimación "ridge" en la regresión lineal con presencia de colinealidad.

PROLOGO

"La computadora digital es uno de los descubrimientos más grandes de la humanidad comparable con la máquina de vapor, la electricidad y la energía nuclear. Pero lo único notable de la computadora es que ella ensancha el poder intelectual del hombre, extendiendo su potencial para aprehender las leyes de la naturaleza".

A.N. Tikhonov y A.V. Goncharsky (*)

En la Matemática Aplicada de las últimas tres décadas el estudio y la resolución de los problemas *Mal-planteados* a la Hadamard ha venido mereciendo una sorprendente cada vez más creciente atención. Se dice que un problema es *Mal-planteado* cuando su solución no depende continuamente de sus datos, siendo por ello precisamente que su resolución tropiece con serias dificultades, sobre todo desde el punto de vista de la Matemática Numérica.

Ahora bien, a la clase de los problemas *Mal-planteados* pertenecen una gran parte de los problemas *Inversos* que con frecuencia creciente tienen lugar en la ciencia y la tecnología de nues

(*) A.N. TIKHONOV y A.V. GONCHARSKY, *Ill-Posed Problems in the Nature Sciences*, MIR (1987), prólogo.

tros días (*). En términos generales, resolver un problema *Inverso* significa determinar (descubrir) las causas que dieron lugar a ciertos efectos observados. He aquí el proceso contrario al de la lógica de causa-efecto.

En la actualidad, el método más sobresaliente para la resolución de los problemas *Mal-planteados* es el método de *Regularización* de Tijonov, el cual nos permite obtener soluciones aproximadas a tales problemas que dependen continuamente de sus datos. Cabe decir que en ello juegan un papel relevante el Análisis Funcional, la Estadística, la Matemática Numérica y sobre todo la computadora digital.

Se tiene, sin embargo, que una de las principales dificultades que se presenta en la realización práctica del método de *Regularización* es la buena elección del llamado *parámetro de regularización*. A este respecto está dedicado el presente trabajo, haciendo una nueva propuesta para efectos de dicha elección, la cual hemos llamado *Principio de Seudo-optimalidad*, demostrando que ella define un algoritmo de regularización a la Tijonov (Capítulo 3), y exhibiendo su efectividad práctica en la resolución numérica de la ecuación integral de Fredholm de primer tipo en conexión con la estimación Gauss-Markov regularizada en la regresión lineal (Capítulo 4).

(*) . Tales problemas aparecen en Análisis Numérico (diferenciación numérica y suavamiento de datos entre otros), en Ingeniería (reconstrucción de señales), en Geofísica (prospección magnetométrica y gravimétrica), en Medicina (tomografía), etc.

Con el objetivo de hacer el presente trabajo autocontenido, fue preciso: i) desarrollar una amplia discusión sobre el concepto de un problema *Mal-planteado* a la Hadamard ilustrándolo - con varios ejemplos (Capítulo I); y ii) presentar los elementos básicos de la teoría sobre el método de *Regularización* de Tijonov (Capítulo II).

No se puede concluir este prólogo sin antes agradecer el apoyo cabal y franco que me brindó mi tutor y amigo el Dr. José L. Farah I., sin el cual no hubiera sido posible llevar a cabo esta tesis. También quiero agradecer el estímulo desinteresado que hasta ahora me ha ofrecido mi profesor y amigo el Dr. Pablo Barrera S.. Además, es para mi grato el poder agradecer el apoyo económico que me brindó el Dr. Richard Griego para hacer posible mi visita al Departamento de Matemática y Estadística de la Universidad de Nuevo México, (Albuquerque), en donde se llevó a cabo gran parte del estudio experimental que en este trabajo se reporta. Finalmente, agradezco el profesional apoyo de mecanografía que la Sra. Lucina Parra A. tuvo la bondad de otorgarme.

Julio de 1988.

INDICE

RESUMEN	1
PRÓLOGO	11
CAPÍTULO 1.	
INTRODUCCIÓN.	
§1.1 Introducción Histórica	1
§1.2 Problemas Mal-Planteados a la Hadamard	6
1. Problemas de Hadamard	6
2. Problema Retrógrado del Calor	8
3. Continuación Armónica	10
4. Diferenciación Numérica	11
5. Ecuación Integral de Fredholm de la. especie	12
6. La Inversa Generalizada Moore-Penrose de una Matriz	13
7. Sistemas Lineales Algebraicos	16
8. Estimación Lineal Gauss-Markov	19
§1.3 Problemas Mal-Planteados y la Inversa Generalizada de Moore-Penrose	21
§1.4 Método de Regularización de Tíjonov: Introducción al Principio de Seudo-optimalidad	25
CAPÍTULO 2.	
MÉTODO DE REGULARIZACIÓN DE TÍJONOV.	
§2.1 Antecedentes Históricos.....	31
§2.2 Método de Regularización de Tíjonov	35

§2.3 Teorema de Picard y Convergencia del Método de Regularización.	44
--	-----------

CAPÍTULO 3.

ELECCIÓN DEL PARÁMETRO DE REGULARIZACIÓN:

PRINCIPIO DE SEUDO-OPTIMALIDAD.

§3.1 Selección del Parámetro de Regularización	50
§3.2 Principio de Seudo-Optimalidad: I. Versión Métrica	56
§3.3 Regularización de la Inversa Generalizada (Moore-Penrose) de una Matriz	62
§3.4 Principio de Seudo-Optimalidad: II. Versión Estadística	72
Apéndice. Vectores Aleatorios en Espacios de Hilbert.	77

CAPÍTULO 4.

ESTIMACIÓN GAUSS-MARKOV REGULARIZADA CON SELECCIÓN SEUDO-ÓPTIMA DEL PARÁMETRO DE REGULARIZACIÓN.

§4.1 Resolución Numérica de la Ecuación Integral de Fredholm de Primer Tipo y Colinealidad en la Regresión Lineal.	84
4.1.1 Colinealidad en la Regresión Lineal.	84
4.1.2 Convergencia Discreta del Método de Regularización.	88
§4.2 Estimación Gauss-Markov Regularizada	95
§4.3 Principio de Seudo-optimalidad: Realización Práctica.	107
§4.4 El Algoritmo: Descripción.	118
§4.5 Ejemplos y Resultados Numéricos.	126

CONCLUSIONES	149
REFERENCIAS	152
REFERENCIAS ADICIONALES	160

CAPITULO 1

INTRODUCCION

El objetivo de este capítulo es presentar una introducción a: (1) el concepto de *Problema Mal-Plantado* a la Hadamard, (2) el *Método de Regularización* de Tijonov para la resolución de problemas mal-plantados, y (3) las contribuciones del presente trabajo.

§1.1 Introducción Histórica.

Propuesto un problema, para la Matemática es fundamental la cuestión sobre la existencia y la unicidad de la solución. En el campo de las Ecuaciones Diferenciales Parciales (EDP) o de la Física Matemática, esta cuestión se remonta a los trabajos de Cauchy y Kovalevskaya. Pero realmente es J. Hadamard [40], en su clásico reporte de 1902, quien introduce a este respecto los términos, tanto de un *Problema Bien-Plantado* (B-Pdo) como el de uno *Mal-Plantado* (M-Pdo); entendiéndose por lo primero, un problema que en general es *posible* (i.e. que su solución

existe) y *determinado* (i.e. que su solución es única); y por lo segundo, un problema que en general no es B-Pdo. Hace notar que las tres ecuaciones clásicas de la Física Matemática (de onda, de calor y de potencial) con condiciones iniciales o de Cauchy, o bien con condiciones de frontera del tipo de Dirichlet dictadas por la Física según el caso, son ejemplos típicos de problemas B-Pdos. Por otro lado, muestra que la ecuación de - Laplace en el espacio con condiciones de Cauchy sobre el plano yz , hoy llamado *problema de Hadamard*, es un problema M-Pdo, - pues en general es *imposible*. (*) También muestra que lo mismo ocurre con la ecuación de onda en el espacio con condiciones de Cauchy sobre el plano $x=0$. Es importante anotar que para Hadamard estos dos problemas carecen de significado físico, en el sentido de no corresponder a alguna situación física.

En 1917, en una reunión de la sociedad matemática Suiza (Zurich), J. Hadamard ([41], p. 33) da a conocer su clásico - ejemplo de problema de Cauchy para la ecuación de Laplace:

$$\begin{aligned} \partial_x^2 u + \partial_y^2 u &= 0, \text{ en el semi-plano } x > 0, \\ \text{con condiciones iniciales} & \\ u(0,y) &= 0 \text{ y } \partial_x u(0,y) = A_n \operatorname{sen}(ny), \end{aligned}$$

cuya solución es $u(x,y) = (A_n/n) \operatorname{sh}(nx) \operatorname{sen}(ny)$. De lo cual se sigue que la solución de este problema no depende continuamente de sus datos $u(0,y)$ y $\partial_x u(0,y)$. En efecto, para datos de

(*) Siendo determinado cuando es posible.

clase C^{p-1} con topología C^{p-1} se observa que $A_n \operatorname{sen}(ny)$ con $A_n = n^{-p}$ tiende a la función idénticamente cero, mientras que para cualquier $x > 0$ dado, $(A_n/n) \operatorname{sh}(nx) \operatorname{sen}(ny)$ tiende a infinito cuando $n \rightarrow \infty$. Luego, la solución $u(x, y) \equiv 0$ correspondiente a los datos $u(0, y) = \sum_x u(0, y) \equiv 0$, no depende continuamente de sus datos. Más aún, si se toma $A_n = e^{-\sqrt{n}}$, se concluye lo mismo para datos de clase C^∞ .

Hoy día, por un problema B-Pdo a la Hadamard se entiende un problema cuya solución existe, es única y depende continuamente de sus datos. ⁽¹⁾ A este respecto, en su libro ([42], p. 146) de edición póstuma (Pekín, 1954), J. Hadamard menciona que esta noción, la cual él comparte, se debe a R. Courant y D. Hilbert, quienes la proponen por primera vez en el clásico libro de "Métodos Matemáticos de la Física" ([19], vol. 2).

Desde el punto de vista práctico y/o de la Física, la dependencia continua de la solución de un problema con respecto de sus datos es una condición natural y fundamental; pues en la vida práctica es usual que los datos del problema se obtengan por medición. Si un cierto fenómeno en estudio no cumple con esta condición, entonces muy pequeños cambios en sus datos bastarían para cambiar radicalmente su marcha. Así, no parecería estar gobernado por leyes precisas, sino más bien por

(1) En EDP, por datos de un problema puede entenderse valores iniciales, valores de frontera, coeficientes del operador, geometría del dominio de definición, etc.

el azar. (2) En este sentido, el problema de Hadamard carece de significado físico, aparte de no corresponder a alguna situación física concreta.

Hasta antes de los 50's, se cuentan pocos trabajos sobre el estudio de los problemas M-Pdos, y con excepción de uno de Tijonov ([87], 1943), ellos tratan sobre los aspectos de existencia y unicidad, como ocurre con Bourgin [14] y Bourgin-Duffin [13] con relación al problema de Dirichlet para la ecuación de onda.

Es durante la década de los 50's cuando aparecen trabajos en los que: (1). Se proponen esquemas numéricos para el cálculo de soluciones aproximadas del problema de Hadamard, que bajo razonables supuestos a priori sobre la solución, convergen a su solución (Pucci [76] (3) y Lavrentiev [54]); (2). Se discuten problemas de incuestionable significado físico que matemáticamente dan lugar al problema de Hadamard (Pucci [77] y Courant-Hilbert [19], entre otros); y, (3). Se presentan resultados, fundamentales para efectos de La Matemática Numérica, sobre la dependencia continua de la solución para varios problemas M-Pdos, bajo razonables supuestos a priori sobre la solución. (John [50] y Pucci [77]). La importancia de estos trabajos es capital, pues se muestra que una gran variedad de pro--

(2) Éste es precisamente el concepto de azar de Poincaré ([41], p. 38).

(3) Este trabajo es mencionado por Hadamard ([42], p. 146).

blemas M-Pdos, con ayuda de la computadora resultan ser tan tr**á** bajables como los B-Pdos, echando abajo severos prejuicios al respecto (véase a John [50]). Cabe mencionar, por otro lado, - que de tiempo atrás para ese entonces la ciencia y la tecnología venían planteando problemas relevantes que resultaban ser - M-Pdos; en especial, ciertos problemas inversos en Geofísica.

Ya para la década de los 60's se observa el inicio de un crecimiento acelerado de trabajos sobre el estudio de problemas M-Pdos de variada naturaleza en diversas ramas de la ciencia y la tecnología (Payne [72], Lavrentiev [55], Arsenin-Tikhonov [2], Twomey [93] y Morozov [65]), sobre todo a raíz de los tr**á** bajos de Phillips [74] y Tikhonov ([88] y [89]) con los cuales se abre propiamente una área nueva de la Matemática: la *Teoría de Regularización de Tijonov para La Resolución de Problemas Mal-Planteados a La Hadamard*.

Para los 70's, la Metodología de Regularización es mundialmente aceptada y cultivada, convirtiéndose en una de las ramas de gran importancia para el Análisis y la Matemática Numérica.

Terminamos este relato histórico diciendo que el estudio y resolución de los problemas M-Pdos, es un claro ejemplo en donde la exigencia del desarrollo de los medios de producción ayudan al hombre de ciencia a barrer prejuicios y obstáculos en el ataque de los retos que el indisoluble binomio teoría-práctica le plantea.

§1.2. Problemas Mal-Planteados a la Hadamard.

Para nuestros propósitos es de interés el presentar varios ejemplos, empezando por los clásicos, de problemas M-Fdos, los cuales se pueden llevar a la formulación operacional

$$(1.2.1) \quad Au = f, \quad f \in F \text{ dada,}$$

siendo $A : U \rightarrow F$ un operador lineal compacto sobre ciertos espacios de funciones (vectoriales normados) U y F .

1. **Problema de Hadamard.** La importancia de este ejemplo radica en que a pesar de haber sido ignorado y visto sólo como un ejemplo patológico, éste tiene lugar en diversas situaciones de interés para las ciencias como ocurre en la determinación del biopotencial dentro de un cuerpo a partir de valores dados sobre la superficie del cuerpo; en la continuación analítica del potencial de corrientes eléctricas que fluyen al interior de la tierra a partir de los valores de éste tomados sobre la superficie (Nedelkov [70]), y en la determinación del potencial gravitacional con propósitos de exploración Geofísica, a partir de información tomada sobre la superficie terrestre (Dmitriev et. al. [25]). Así pues, a la luz de nuestros días, la opinión de Hadamard con relación a la carencia de significado físico de su clásico ejemplo carece de justificación (Pucci [77]).

Ahora bien, considérese el problema de Hadamard siguiente

$$(1.2.2) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad u = u(x, y),$$

en el rectángulo semi-infinito $D = \{(x, y) | 0 < x < a, y > 0\}$ con condiciones de Cauchy

(1.2.2') $u(x, 0) = f(x)$ y $\partial_y u(x, 0) = g(x)$, $0 < x < a$
y con condiciones de frontera

(1.2.2'') $u(0, y) = u(a, y) = 0$, $y > 0$.

Por aplicación directa del método de separación de variables se obtiene que la solución, para el rectángulo $0 < x < a$, $0 < y < y_0$, $y_0 > 0$ dado, viene dada por

$$(1.2.3) \quad u(x, y) = \int_0^a G_1(x, y; \xi) f(\xi) d\xi + \int_0^a G_2(x, y; \xi) g(\xi) d\xi$$

donde

$$G_1(x, y; \xi) = \frac{2}{a} \sum_{n=1}^{\infty} \operatorname{ch} \frac{n\pi}{a} y \operatorname{sen} \frac{n\pi}{a} x \operatorname{sen} \frac{n\pi}{a} \xi,$$

$$G_2(x, y; \xi) = 2 \sum_{n=1}^{\infty} \frac{1}{n\pi} \operatorname{sh} \frac{n\pi}{a} y \operatorname{sen} \frac{n\pi}{a} x \operatorname{sen} \frac{n\pi}{a} \xi.$$

siempre que f y g admitan desarrollo de Fourier en senos, y que las series

$$(1.2.3') \quad \sum_{n=1}^{\infty} A_n \operatorname{ch} \frac{n\pi}{a} y_0 \operatorname{sen} \frac{n\pi}{a} x \quad (A_n = \frac{2}{a} \int_0^a f(\xi) \operatorname{sen} \frac{n\pi}{a} \xi d\xi),$$

$$\sum_{n=1}^{\infty} B_n \operatorname{sh} \frac{n\pi}{a} y_0 \operatorname{sen} \frac{n\pi}{a} x \quad (B_n = \frac{2}{n\pi} \int_0^a g(\xi) \operatorname{sen} \frac{n\pi}{a} \xi d\xi)$$

sean en algún sentido ($L^2[0, a]$ por ejemplo) convergentes. Es

importante anotar que estas condiciones de convergencia son extremadamente exigentes.

Para el caso particular en que $g(x) \equiv 0$, de ocurrencia práctica (Nedelkov [70]), de (1.2.3) se sigue, para $y = y_0$, que

$$\int_0^a u(n, y_0) \operatorname{sen} \frac{n\pi}{a} n \, dn = \operatorname{ch} \frac{n\pi}{a} y_0 \int_0^a f(n) \operatorname{sen} \frac{n\pi}{a} n \, dn.$$

Así, escribiendo $u_{y_0}(x) \equiv u(x, y_0)$, se tiene que la solución del problema (1.2.2 - 2' - 2'') con $g(x) \equiv 0$ y para $y = y_0$, viene a ser la solución de la ecuación integral

$$(1.2.4) \quad \int_0^a k_{y_0}(x, \xi) u_{y_0}(\xi) \, d\xi = f(x), \quad 0 < x < a,$$

donde

$$k_{y_0}(x, \xi) = \sum_{n=1}^{\infty} \lambda_n \operatorname{sen} \frac{n\pi}{a} x \operatorname{sen} \frac{n\pi}{a} \xi$$

$$(\lambda_n = \frac{2}{a} (\operatorname{ch} \frac{n\pi}{a} y_0)^{-1}).$$

Ecuación que podemos escribir, en términos operacionales, como

$$(1.2.4') \quad Au_{y_0} = f,$$

donde $A: L^2[0, a] \rightarrow L^2[0, a]$ (p. ej.) viene dado por

$$\lambda[v](x) = \int_0^a k_{y_0}(x, \xi) v(\xi) \, d\xi.$$

2. Problema Retrógrado del Calor. Este problema es otro de los clásicos problemas M-Pdos de especial interés en Geofísica

ca, el cual consiste en, conociendo la distribución de temperaturas $f(x) = u(x, T)$ ($T > 0$) de una cierta barra de longitud a , determinar su distribución de temperaturas $u_0(x) = u(x, 0)$ inicial, manteniendo los extremos de la barra a temperatura, por ejemplo, constante cero. Bajo supuestos de homogeneidad y de escalamiento adecuado del tiempo, la formulación matemática del problema consiste en hallar $u_0(x) = u(x, 0)$, $0 \leq x \leq a$, dado que

$$(1.2.5) \quad \partial_t^2 u = \partial_x^2 u$$

$$(1.2.5') \quad u(x, T) = f(x), \quad T > 0 \text{ dado, y}$$

$$(1.2.5'') \quad u(0, t) = u(a, t) \equiv 0, \quad 0 \leq t \leq T.$$

Mediante la aplicación directa del método de separación de variables, se obtiene que la solución del problema de Cauchy siguiente

$$\partial_t^2 u = \partial_x^2 u$$

$$u(x, 0) = u_0(x), \quad 0 \leq x \leq a$$

$$u(0, t) = u(a, t) \equiv 0, \quad t > 0$$

viene dada por [90]

$$(1.2.6) \quad u(x, t) = \int_0^a G(x, t; \xi) u_0(\xi) d\xi$$

donde

$$G(x, t; \xi) = \frac{2}{a} \sum_{n=1}^{\infty} e^{-\left(\frac{n\pi}{a}\right)^2 t} \sin \frac{n\pi}{a} x \sin \frac{n\pi}{a} \xi$$

siempre que $u_0(x)$ admita en algún sentido el desarrollo de -
Fourier $u_0(x) = \sum_{n=1}^{\infty} C_n \operatorname{sen} \frac{n\pi}{a} x$. (*)

Ahora, como $u(x, T) = f(x)$, se tiene que la solución del -
problema bajo discusión se obtiene resolviendo la ecuación inte-
gral

$$(1.2.7) \quad \int_0^a G_T(x, \xi) u_0(\xi) d\xi = f(x), \quad 0 < x < a,$$

donde $G_T(x, \xi) \equiv G(x, T; \xi)$. La cual bien podemos describir en for-
ma operacional (1.2.1) como ya se indicó en el ejemplo anterior.†

3. Extensión Armónica. Este es un problema clásico en -
Análisis. Pensando en geometría circular, el problema consiste
en extender la función armónica $u(\rho, \theta)$ del disco $D_r = \{(\rho, \theta) | 0 < \rho < r, 0 < \theta < 2\pi\}$ al disco $D_R = \{(\rho, \theta) | 0 < \rho < R, 0 < \theta < 2\pi\}$ ($R > r$).
De la representación integral de Poisson para $u(\rho, \theta)$ en D_r , da-
da $u_r(\theta) = u(r, \theta)$

$$u(\rho, \theta) = \int_0^{2\pi} k_r(\rho, \theta; \varphi) u_r(\varphi) d\varphi,$$

donde

$$k_r(\rho, \theta; \varphi) = \frac{1}{2\pi} \frac{R^2 - \rho^2}{R^2 + \rho^2 - 2R\rho \cos(\varphi - \theta)},$$

se sigue que la solución del problema se reduce a resolver, da-
da $f(\theta) = u(r, \theta)$, la ecuación integral

$$(1.2.8) \quad \int_0^{2\pi} k(\theta, \varphi) u_r(\varphi) d\varphi = f(\theta), \quad 0 < \theta < 2\pi,$$

(*) Siendo C_n el correspondiente coeficiente de Fourier de $u_0(x)$ con respecto a $\operatorname{sen} \frac{n\pi}{a} x$.

donde $k(\theta, \varphi) = k_{R,r}(\theta, \varphi) = k_R(\rho, \theta, \varphi)$. Que bien podemos reescribir en la formulación (1.2.1). #

4. **Diferenciación Numérica.** Este problema del Análisis Numérico ocurre con frecuencia en las aplicaciones, y consiste en hallar la k -ésima derivada $u(x) = f^{(k)}(x)$ de una función $f(x)$ dada (usualmente a través de una tabla T_x de valores de f). Suponiendo, sin perder generalidad, que $f^{(i)}(0) = 0$ para $i = 0, 1, \dots, k-1$, del desarrollo de Taylor con su residuo en forma integral se tiene que

$$\begin{aligned} f(x) &= f(0) + f'(0)x + \dots + \frac{1}{(k-1)!} f^{(k-1)}(0) x^{k-1} \\ &+ \int_0^x \frac{(x-\xi)^{k-1}}{(k-1)!} u(\xi) d\xi \\ &= \int_0^a g(x, \xi) u(\xi) d\xi, \quad 0 < x < a, \end{aligned}$$

dónde

$$g(x, \xi) = \begin{cases} 0, & \xi > x \\ (x-\xi)^{k-1}/(k-1)!, & \xi \leq x. \end{cases}$$

Esto es, el problema se reduce a resolver la ecuación integral

$$(1.2.9) \quad \int_0^a k(x, \xi) u(\xi) d\xi = f(x), \quad 0 < x < a,$$

que bien se escribe en la formulación (1.2.1). #

5. Ecuación Integral de Fredholm de Primera Especie. En diversas aplicaciones en Ingeniería (Audley et. al. [3], Cabayan et. al. [17]), en Físico-Química (Provencher [75]), en Inmunología (Hanson [43]), en Ciencias Atmosféricas (Twomey [93]), entre muchas otras, tiene lugar la ecuación integral de Fredholm de primera especie

$$(1.2.10) \quad \int_a^b k(x,t)u(t) dt = f(x), \quad a \leq x \leq b, \quad (*)$$

donde su función núcleo $k(x,t)$ puede estar en L_1 o en L_2 , con frecuencia continuo o bien de convolución (i.e. $k(x,t) = k(x-t)$).

Este problema, en términos de operadores, juega un papel central en Análisis Funcional pues, bajo ciertas condiciones, resulta ser un ejemplo prototipo de operador (lineal) compacto. (4)

Un resultado clásico del Análisis Funcional [3] dice que para todo operador lineal compacto con inverso sobre espacios de Banach (de dimensión infinita), el operador inverso no puede

- (*) . Aquí, $a < x < b$, si $a = -\infty$ y $b = +\infty$; $a \leq x < b$, si $b = \infty$; etc.
- (4) . Siendo U y F espacios de Banach, $A: U \rightarrow F$ se dice que es compacto, si manda conjuntos acotados de U en conjuntos pre-compactos de F .
- (5) . El problema de diferenciación numérica es B-Pdo si se piensa de $C^n[a,b] \rightarrow C[a,b]$ con topología C^n para $C^n[a,b]$, y es N-Pdo si $C^n[a,b]$ se piensa con topología C ; en ambos casos $C[a,b]$ se piensa con topología C .

ser continuo. En consecuencia, si se tiene que

$$\int_a^b k(x,t) u(t) dt \equiv 0 \longrightarrow u(t) \equiv 0,$$

entonces la ecuación integral (1.2.10) es un problema M-Pdo. (*)

6. La inversa Generalizada de Moore-Penrose de una Matriz. La inversa generalizada Moore-Penrose A^+ de una matriz (real) A de $m \times n$, se define como la solución del sistema matricial - (Penrose [73] y Barrera et. al. [7] para la su presentación de esta tesis)

$$(1.2.11) \quad AX = P, \quad y$$

$$XA = Q$$

donde $P: \mathbb{R}^m \rightarrow \text{Im}(A)$ y $Q: \mathbb{R}^n \rightarrow N(A)^\perp$ son las proyecciones ortogonales sobre $\text{Im}(A)$ y $N(A)^\perp$ respectivamente.

Es inmediato ver que si A , en el caso $m=n$, tiene inversa entonces $A^+ = A^{-1}$.

Si por $M_{m \times n}$ denotamos al espacio vectorial de todas las matrices reales $m \times n$ y por $M_{n \times m}$ a las de $n \times m$, entonces podemos definir el mapeo *inversión generalizado* de Moore-Penrose (M-P)

$$(1.2.12)$$

$$\psi^+: M_{m \times n} \rightarrow M_{n \times m},$$

dado por

$$(1.2.12')$$

$$A \longmapsto A^+$$

(*) En la Sec. 3 de este capítulo se demostrará, en un sentido generalizado de solución por precisar, que la ecuación integral (1.2.10) es un problema a M-Pdo.

Si denotamos por I al conjunto de todas las matrices invertibles en $M_{n \times n}$, entonces se tiene, como es bien sabido, que el mapeo *Inversión* $\psi : I \subset M_{n \times n} \rightarrow M_{n \times n}$ dado por $A \mapsto A^{-1}$, es continuo sobre I ([79]).

Siendo claro que $\psi^+ : M_{n \times n} \rightarrow M_{n \times n}$ es una extensión de ψ , es, es, que $\psi^+|_I = \psi$, resulta sorprendente que (en general) ψ^+ no sea continuo.

En efecto, si el rango de A ($\text{rgo}(A)$) $< \min(m, n)$, para A en $M_{m \times n}$ dada, entonces se tiene que

$$(1.2.13) \quad \| (A + \delta A)^+ - A^+ \| \geq \frac{1}{\|\delta A\|}, \quad (6)$$

siempre que $\text{rgo}(A + \delta A) > \text{rgo}(A)$ (Stewart [85]). En consecuencia, $\| (A + \delta A)^+ - A^+ \| \rightarrow +\infty$ cuando $\|\delta A\| \rightarrow 0$. Con ello, queda demostrado que el problema de hallar (calcular) A^+ , dada A , es un problema M-Pdo a la Hadamard. Así, se tiene que el cálculo numérico de A^+ puede resultar catastróficamente sensible a pequeñas perturbaciones en A , en particular al redondeo.

Veamos más de cerca este punto. Supongamos que se dispone, teóricamente hablando, de un método *numéricamente estable* $\psi^+ : M_{m \times n} \rightarrow M_{n \times n}$ para calcular a A^+ numéricamente, en el sentido que la resolución numérica de (1.2.11) que produce, es la solución exacta para una cierta matriz "vecina" $A + \delta A$. En términos precisos, que

(6) Aquí, $\| \cdot \| : M_{n \times n} \rightarrow R^+$ es una norma matricial *compatible*, véase a Stewart [84].

$$(1.2.14) \quad \tilde{\varphi}^+(A) = \varphi^+(A + \delta A) \text{ con } \frac{\|\delta A\|}{\|A\|} < \mu,$$

siendo μ la precisión de la computadora digital anfitrión.

De (1.2.13) y (1.2.14) se tiene que

$$\frac{\|\tilde{\varphi}^+(A) - \varphi^+(A)\|}{\|\varphi^+(A)\|} > \frac{1}{\|\varphi^+(A)\|} \cdot \frac{1}{\|\delta A\|} \\ = \frac{1}{\|\varphi^+(A)\| \|A\|} \frac{\|A\|}{\|\delta A\|} > \frac{1}{K(A)\mu} > \frac{1}{\mu},$$

donde $K(A) = \|A^+\| \|A\|$ (siempre > 1) es el número de condición numérica de A . Este resultado es crucial, pues nos dice que con nuestro método hipotético numéricamente estable, pensando $1/\mu$ como medida de la precisión de la computadora anfitrión, aún en una máquina de precisión infinita (como caso límite), no se puede calcular numéricamente a A^+ , en el mejor de los casos, con siquiera un dígito (en norma) significativo. Nótese que la situación empeora al aumentar la precisión.

El asunto se torna diferente si sólo se permiten perturbaciones δA de manera que $\text{rgo}(A + \delta A) = \text{rgo}(A) = r \leq \min(m, n)$, ya que si $M_{m \times n}^r$ denota a todas las matrices en $M_{m \times n}$ de rango r entonces $\varphi^+|_{M_{m \times n}^r}$ es continuo sobre $M_{m \times n}^r$, pues se tiene que

$$\|\varphi^+(A + \delta A) - \varphi^+(A)\| < \frac{3\|A^+\|^2}{1 - \|A^+\| \|\delta A\|} \|\delta A\|,$$

siempre que $\|A^+ \| \| \Delta A \| < 1$ (Wedin [10]). En consecuencia, para nuestro método hipotético $\tilde{\varphi}$ se tiene que

$$\frac{\| \tilde{\varphi}^+(A) - \varphi^+(A) \|}{\| \varphi^+(A) \|} < \frac{3K(A)\mu}{1 - K(A)\mu},$$

siempre que $\text{rgo}(A + \delta A) = \text{rgo}(A)$ y $K(A)\mu < 1$. Luego, si $K(A)\mu = 10^{-p}$ (p un entero mayor que 1) entonces en el peor de los casos, se tendrán por lo menos p dígitos significativos (en norma).

El gran pezo de este último resultado es la condición de só lo aceptar perturbaciones (por redondeo) δA de A con $\text{rgo}(A + \delta A) = \text{rgo}(A)$, la cual es muy restrictiva; pues desde un punto de vista real, y hablando en lenguaje "probabilístico", con probabilidad (*) uno se tiene que $\text{rgo}(A + \delta A) > \text{rgo}(A)$.

En el capítulo 3, de esta tesis se verá cómo calcular numéricamente a A^+ de manera estable a perturbaciones en A , sin la hipótesis $\text{rgo}(A + \delta A) > \text{rgo}(A)$, por el método de regularización de Tíjonov. #

7. Sistemas Lineales Algebraicos. Sean A en M_{\max}^m y b en R^m dados. En las aplicaciones del Álgebra Matricial Numérica, el problema de hallar x en R^n tal que

$$(1.2.14) \quad Ax = b,$$

es de capital importancia.

(*) En el sentido de que dada cualquier vecindad V_R de radio $R(>0)$ de A , $V_R \cap M_{\max}^m$ tiene medida cero.

Ahora bien, en el sentido clásico de solución, el problema (1.2.14) no tiene solución, en general; y de tenerla, ésta no tiene porque ser única.

Para la teoría como para las aplicaciones, es conveniente considerar el concepto siguiente de solución generalizada para $Ax = b$.

Definición 1.2. Por la *solución generalizada de Moore-Penrose* x^+ de $Ax = b$, se entiende la solución de menor norma euclídeana del problema de mínimo de cuadrados

$$\text{Mín } \|b - Ax\|_2^{(*)}. \quad \#$$

Se demuestra que x^+ siempre existe y que es única (Golub-Van Loan [37']). Más aún,

$$(1.2.15) \quad x^+ = A^+ b,$$

Así pues, el problema $Ax = b$, en sentido de solución generalizada M-P, es siempre posible y determinado, no siendo así en sentido clásico. Sin embargo es importante observar, para $m = n$ en particular, que si $Ax = b$ tiene solución clásica entonces $x^+ = A^{-1} b$. Esto es, las soluciones clásica y generalizada coinciden.

Si por datos del problema $Ax = b$ se entiende la pareja (A, b) , entonces el problema $Ax = b$ es en general M-Pdo, pues

(*) En \mathbb{R}^n (norma euclídeana): $\|y\|_2 = (\sum_{i=1}^n |y_i|^2)^{1/2}$

$x^+ = \varphi^+(A) \cdot b$, y como ya hemos visto, φ^+ no es continuo en general.

Si por datos del problema $Ax = b$ se entiende sólo al vector b (en \mathbb{R}^m) entonces $Ax = b$ es un problema B-Pdo, ya que

$$\frac{\|A^+(b + \delta b) - A^+b\|_2}{\|A^+b\|_2} < K_2(A) \frac{\|\delta b\|_2}{\|b\|_2}$$

$$(K_2(A) = \|A^+\|_2 \|A\|_2).^{(*)}$$

Cabe dejar bien claro que el conjunto de las matrices de rango deficiente (o singulares, según el caso) sobre las que en general φ^+ no es continuo, es un conjunto de medida cero, en el sentido que su intersección con cualquier conjunto acotado de $M_{m \times n}$ tiene medida^(**) cero. Esto es, para una gran mayoría de los casos, $Ax = b$ se puede resolver numéricamente de manera aceptable. Sin embargo, como iremos viendo en los capítulos siguientes, la resolución numérica de problemas (lineales) M-Pdos constituye una de las principales fuentes que dan lugar al problema $Ax = b$ con A "muy próxima" a una de rango deficiente (matrices muy mal-comportadas, en Inglés "ill-conditioned"). "muy próximas" significa indistinguibles de una de rango deficiente de acuerdo a la precisión de la computadora anfitriona. \square

(*) $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$.

(**) En sentido de Lebesgue.

8. Estimación Lineal Gauss-Markov. En Estadística juega un papel central (regresión lineal) el problema siguiente. Dado el modelo lineal

$$(1.2.16) \quad \mathbf{X}\underline{\beta} + \underline{\epsilon} = \mathbf{y},$$

en donde \mathbf{X} es una matriz real $n \times p$, y \mathbf{y} es un vector en \mathbb{R}^n , ambos dados; y $\underline{\epsilon}$ es un vector aleatorio en \mathbb{R}^n (no observable) con $\mathbb{E}[\underline{\epsilon}] = 0$ y $\mathbb{E}[\underline{\epsilon} \underline{\epsilon}^t] = \sigma^2 \mathbf{I}^{(n)}$. El problema consiste en hallar una "buena estimación" $\hat{\underline{\beta}}$ de $\underline{\beta}$.

Como bien se sabe, cuando la matriz de diseño \mathbf{X} es de rango máximo ($\text{rango}(\mathbf{X}) = p$), el estimador Gauss-Markov

$$(1.2.17) \quad \hat{\underline{\beta}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y},$$

es de mínima varianza dentro de todos los estimadores lineales insesgados de $\underline{\beta}$ (Seber [80]).

Sin embargo, si \mathbf{X} es muy cercana a ser de rango deficiente, o sea mal-comportada en términos numéricos (i.e. $K_2(\mathbf{X}) \mu \approx 1$), o con presencia del fenómeno de colinealidad en términos estadísticos, entonces el cálculo numérico de $\hat{\underline{\beta}}$ resulta en general ser insatisfactorio, si no es que hasta inservible. La explicación a esta situación la hemos dado en el ejemplo anterior; para verlo, basta con considerar que $\hat{\underline{\beta}} = \mathbf{X}^+ \mathbf{y}$.

(*) \mathbb{E} denota al operador *esperanza* (Seber [80]).

Ahora, veamos como la resolución numérica de la ecuación integral de Fredholm de 1a. clase da lugar al modelo lineal (1.2.18), con presencia de colinealidad como consecuencia hereditaria del ser dicha ecuación un problema M-Pdo.

En las aplicaciones es usual que, para la ecuación integral (1.2.10), a $f(x)$ se le conozca por medio de una tabla de sus valores:

$$T_f : (x_i, y_i), y_i = f(x_i) + \tilde{c}_i, i = 1, 2, \dots, n;$$

obtenidos por observación. En estos términos, la ecuación integral (1.2.10) toma la formulación siguiente

$$(1.2.18) \quad \int_a^b k(x_i, t)u(t) dt + \tilde{c}_i = y_i, i = 1, 1, \dots, n.$$

Luego, discretizando el término integral, mediante una regla de cuadratura del tipo: $\int_a^b g(t) dt \approx \sum_{j=1}^p \omega_j g(t_j)$; se obtiene el modelo lineal

$$(1.2.19) \quad \mathbb{K} \underline{\beta} + c = y$$

en donde

$$\mathbb{K}_{n \times p} = \mathbb{K}D; \quad \mathbb{K} = (k_{ij}), k_{ij} = k(x_i, t_j); \quad D = \text{diag}(\omega_1, \dots, \omega_p)$$

$$\underline{\beta} = (u_1, u_2, \dots, u_p)^T; \quad u_j = u(t_j)$$

$$c = (c_1, c_2, \dots, c_n)^T; \quad c_i = \tilde{c}_i + n_i \quad (n_i \ll \tilde{c}_i). \quad \theta$$

11.3. Problemas Mal-Planteados y la Inversa Generalizada de Moore-Penrose.

Para la siguiente definición, aún en términos vagos, U y F son espacios de funciones (normados o métricos), y $A: U \rightarrow F$ un cierto operador no necesariamente lineal y/o continuo.

Definición 1.3.1 (Hadamard) Dado $f \in F$, el problema de hallar u en U tal que

$$Au = f,$$

se dice *Bien-Planteadado*, si éste tiene una única solución u^0 , la cual depende continuamente de sus datos. Y se dice que es *Mal-Planteadado*, si no es Bien-Planteadado. Por *datos*, en sentido *amplio* se entiende A y f , y en sentido *restringido* se entiende sólo a f . #

El estudio de los problemas $Au = f$, en sentido *amplio* M-Pdos, resulta ser más difícil y complicado que en su sentido *restringido*. Y suele ocurrir que un cierto problema sea B-Pdo en sentido *restringido* y M-Pdo en sentido *amplio* (Sec. 1.2., ejemplo 7).

En este trabajo, salvo el caso del ejemplo 6 de la sección anterior, se hablará fundamentalmente de la resolución (por el método de regularización) de problemas M-Pdos en sentido *restringido*.

Con el propósito, en un sentido generalizado de solución por precisar a continuación, de demostrar que la ecuación integral de Fredholm de 1a. especie con núcleo no-degenerado (7) $k(x,t)$ es un problema M-Pdo, se hace necesaria la introducción de la siguiente definición, pensada en espacios de Hilbert.

Definición 1.3.2. Se dirá que u^+ en U es una *solución generalizada Moore-Penrose (M-P)* de $Au = f$, si

- (a) $\|f - Au^+\|_F \leq \|f - Au\|_F$, para toda $u \in U$; y
 (b) Si $\|f - Au^0\|_F \leq \|f - Au\|_F$, para toda $u \in U$, entonces $\|u^0\|_U \geq \|u^+\|_U$. (*) #

Es importante observar que si $Au = f$ tiene solución única u^0 (en sentido clásico) entonces $u^+ = u^0$.

Proposición 1.3.1. Sea $A : U \rightarrow F$ lineal continuo y P_f la proyección ortogonal de f sobre $\overline{\text{Im}(A)}$ (8). Se tiene que $Au = f$ tiene solución generalizada (M-P) u^+ si y sólo si $P_f \in \text{Im}(A)$. En cuyo caso, u^+ es única.

Demostración. Se obtiene directamente al aplicar el teorema de la proyección ortogonal a $\overline{\text{Im}(A)}$ y a $N(A)$. #

(7) Un núcleo $k(x,t)$ se dice que es *degenerado*, si $k(x,t) = \sum_{i=1}^n a_i(x)b_i(t)$. Y *no-degenerado*, si no es degenerado.

(*) Aquí, y en lo sucesivo, la norma sobre el espacio vectorial normado V , se denotará por $\|\cdot\|_V$.

(8). Por $\overline{\text{Im}(A)}$ se entiende la cerradura de $\text{Im}(A)$.

Así pues, para $A : U \rightarrow F$ lineal continuo y en sentido generalizado (M-P), si el problema $Au = f$ tiene solución entonces esta es única (2).

Para $A : U \rightarrow F$ lineal continuo, sea

$$D(A^+) = \{f \in F \mid Au = f \text{ tiene solución generalizada (M-P)}\}$$

se tiene que el mapeo $A^+ : D(A^+) \subset F \rightarrow U$ dado por $f \mapsto u^+$, llamado *inverso generalizado* de Moore-Penrose, es lineal; que $D(A^+) = \text{Im}(A) \oplus \text{Im}(A)^\perp$ (luego $D(A^+)$ denso en F). Más aún, se tiene que

$$(1.3.1) \quad a) \quad AA^+ = P|_{\text{Im}(A)}$$

$$b) \quad A^+A = Q$$

siendo $P : F \rightarrow \overline{\text{Im}(A)}$ y $Q : U \rightarrow N(A)^\perp$ las proyecciones ortogonales de F y U sobre $\overline{\text{Im}(A)}$ y $N(A)^\perp$, respectivamente. Así, si A tiene inverso entonces $A^+ = A^{-1}$ sobre $D(A^{-1})$.

En estos términos, es conveniente hacer la siguiente reformulación de la definición 1.3.1.

Definición. 1.3.3. (Hadamard)^(*) Para $A : U \rightarrow F$ lineal, dada $f \in F$, se dirá que el problema $Au = f$ es *Bien-Plantado* si y sólo si

$$A^+ : D(A^+) (\subset F) \rightarrow U$$

es continuo. Y se dirá que es *Mal-Plantado*, si no es Bien-Plantado.

(2) En el lenguaje de Hadamard, si el problema $Au = f$ es posible entonces es determinado.

(*) En sentido restringido, esto es, con perturbaciones sólo en f .

Proposición 1.3.2. Si $A: U \rightarrow F$ es lineal y continuo entonces $A^+: D(A^+) \subset F \rightarrow U$ es lineal y cerrado. (8)

Demostración: Sea (f, u) un punto de acumulación de $Gr(A^+) = \{(f, A^+f) \mid f \in D(A^+)\}$, la gráfica de A^+ . Por ver que $(f, u) \in Gr(A^+)$. Sea $\{(f_n, A^+f_n)\}$ tal que $(f_n, A^+f_n) \rightarrow (f, u)$. De la continuidad de $A, \{(f_n - A(A^+f_n)) \in C \text{Im}(A)^\perp\}$ converge a $f - Au \in \text{Im}(A)^\perp$. Luego, $A^+(f - Au) = 0$, o bien que $A^+f = u$. Así, $(f, u) \in Gr(A^+)$. #

Por tanto, para $A: U \rightarrow F$ lineal continuo, A^+ no es continuo en general

Proposición 1.3.3. Sea $A: U \rightarrow F$ lineal y continuo. Si $\overline{\text{Im}(A)} = \text{Im}(A)$ entonces $A^+: F \rightarrow U$ es lineal y continuo (i.e. $Au = f$ es B-Pdo).

Demostración: En este caso $A^+ = \tilde{A}^{-1} P$, siendo $P: F \rightarrow \text{Im}(A)$ la proyección ortogonal de F sobre $\overline{\text{Im}(A)} = \text{Im}(A)$, y $\tilde{A} = A|_{N(A)^\perp}$. Y como \tilde{A}^{-1} y P son continuos, A^+ es continuo. #

Entre los ejemplos típicos de esta situación están: (1). $A: U \rightarrow F$ con $\text{Im}(A)$ de dimensión finita; (2). $A: U \rightarrow F$ inferiormente acotado (i.e. $\|Au\|_F \geq k\|u\|_U$ ($k > 0$), para toda $u \in U$); y (3) $A = (\lambda I - K): U \rightarrow F$ ($U = F$) con K lineal compacto.

(8). Siendo X y Y espacios vectoriales normados, $T: D(T) \subset X \rightarrow Y$ se dice cerrado si su gráfica $Gr(T) = \{(x, Tx) \mid x \in D(T)\}$ es un conjunto cerrado de $X \times Y$.

Teorema 1.3.1. Si $A:U \rightarrow F$ es lineal y compacto entonces para que $D(A^+) = F$ y $A^+:F \rightarrow U$ sea continuo es necesario y suficiente que $\text{Im}(A) \subset F$ sea de dimensión finita.

Demostración: Si A^+ es continuo entonces $AA^+ = P|_{\text{Im}(A)}$ es compacto, y $P|_{\text{Im}(A)}$ es compacto sii $\text{Im}(A) \subset F$ es de dimensión finita. #

En consecuencia, si $A:U \rightarrow F$ es lineal compacto y $\text{Im}(A)$ es de dimensión infinita entonces $A^+:D(A^+) \subset F \rightarrow U$ no puede ser continuo. Esto es, $Au = f$ es M-Pdo. El caso típico es la ecuación integral de Fredholm de 1a. clase con función núcleo $k(x,t)$ no-degenerado. Luego, hemos logrado el objetivo de esta sección. En particular, se tiene que los cinco primeros ejemplos de la sección anterior son, en el presente sentido de solución generalizada, problemas M-Pdos a la Hadamard.

§1.4. Método de Regularización de Tíjonov: Introducción al Principio de Seudo-optimalidad.

La metodología de regularización de Tíjonov nos proporciona un procedimiento bien fundamentado y sistemático para hallar soluciones aproximadas, estables a perturbaciones en f , de la solución en sentido generalizado Moore-Penrose del problema $Au = f$ ($f \in D(A^+)$), donde $A:U \rightarrow F$ es lineal y compacto sobre espacios de Hilbert U y F .

Considérese el problema de hallar una *solución aproximada* para $Au = f$, a partir de la relación operacional $Au = f_\delta$, siendo $f_\delta \in F$ tal que $\|f_\delta - f\|_F < \delta$, $\delta > 0$ dada, estable a perturbaciones en f .

Una de las ideas centrales de Tijonov con respecto a este problema está contenida en la siguiente

Definición 1.4.1. (Tijonov) Bajo el supuesto que $f \in D(A^+)$, se dirá que la familia de operadores $\{R_Y: F \rightarrow U\}_{Y > 0}$ es una *familia regularizante* para $Au = f$, si se puede dar un *criterio* $\gamma = \gamma(\delta)$ de manera que $\gamma \rightarrow 0^+$ cuando $\delta \rightarrow 0^+$, y

$$\|u_Y^\delta - u^+\|_U \rightarrow 0 \text{ cuando } \delta \rightarrow 0^+,$$

en donde $u_Y^\delta = R_Y f_Y$. Y, se dirá que un problema M-Pdo $Au = f$ es *regularizable* (a la Tijonov), si para él existe una familia regularizante $\{R_Y\}_{Y > 0}$. #

Uno de los objetivos centrales del capítulo siguiente será demostrar que el problema $Au = f$, para $A: U \rightarrow F$ *lineal compacto* y $f \in D(A^+)$ es *regularizable*, por medio de la construcción explícita de una *familia regularizante*.

Es muy importante observar que en la definición anterior, no se dice nada sobre cómo definir (o determinar) el *criterio* $\gamma = \gamma(\delta)$. A este problema se le conoce por el *problema de selección del parámetro de regularización*. De hecho, este problema resulta ser una de las serias dificultades, sobre todo prácticas,

del método de regularización. Y si bien es cierto que hoy se cuenta con varios principios para la selección de dicho parámetro (Principio de Discrepancia ([6], [65]), Principio de Quasi-optimalidad ([34], [35] y [32], Criterio de Validación Cruzada Generalizada [100], y otros), también es cierto que no se puede decir que el problema se tiene totalmente resuelto.

El propósito central de la presente tesis es la propuesta de un *criterio nuevo* para la elección de dicho parámetro, el cual nosotros llamaremos Principio de Seudo-Optimalidad, presentando un sistemático primer estudio al respecto.

La génesis de la idea emergió en la búsqueda de la elección óptima del parámetro en cuestión. Específicamente, para el problema de la regresión lineal (Sec. 1.2, ejemplo 8):

$$\mathbf{X}\underline{\beta} + \underline{\epsilon} = \mathbf{Y} \quad \text{con} \quad \underline{\epsilon} \sim (\mathbf{0}, \delta^2 \mathbf{I})^{(*)}$$

bajo presencia de colinealidad se tiene que la estimación (segunda) propuesta por Hoerl y Kennard ([46], [47])

$$\hat{\underline{\beta}}_{\gamma} (= \mathbf{R}_{\gamma} \mathbf{Y}) = (\mathbf{X}'\mathbf{X} + \gamma \mathbf{I})^{-1} \mathbf{X}' \mathbf{Y}, \quad \gamma > 0,$$

la cual nosotros llamaremos *estimación de Gauss-Markov regularizada* por claras razones que veremos después en el Capítulo 4, es una alternativa a la estimación clásica $\hat{\underline{\beta}}$ de Gauss-Markov para $\underline{\beta}$. Así, con base a la función de *error cuadrático medio*

(*). $\underline{\epsilon} \sim (\mathbf{0}, \delta^2 \mathbf{I})$ significa que $\mathbf{E}(\underline{\epsilon}) = \mathbf{0}$ y que $\text{Cov}(\underline{\epsilon}) = \delta^2 \mathbf{I}$.

$$ecm(\gamma) := ecm(\hat{\beta}_\gamma; \underline{\beta}, \delta^2) = E\{(\hat{\beta}_\gamma - \underline{\beta})^2\}$$

es natural el elegir al parámetro de regularización (o "ridge") $\gamma = \gamma_{op}(\delta)$ como aquel en donde la función $ecm(\gamma)$ se minimiza. Ahora bien, cabe aclarar que en la práctica la determinación de γ_{op} se efectúa mediante el uso de una "buena" estimación $C(\gamma)$ para $ecm(\gamma)$ sin el requerimiento del conocimiento explícito de $\underline{\beta}$ y δ^2 .

Por otro lado, considerando que

$$ecm(\gamma) = vt(\gamma) + b^2(\gamma),$$

en donde $vt := vt(\hat{\beta}; \delta^2)$ es la *varianza total* y $b^2(\gamma) := b^2(\hat{\beta}_\gamma; \underline{\beta})$ es el *sesgo cuadrado* para $\hat{\beta}_\gamma$, y que las gráficas de las funciones $ecm(\gamma)$, $vt(\gamma)$ y $b^2(\gamma)$, cuando X es de rango máximo, son como las que se ilustran a continuación en la figura 1, se sugiere (también de manera natural) el elegir a $\gamma = \gamma_{so}(\delta)$ como la solución de la ecuación de *seudo-optimalidad*

$$vt(\gamma) = b^2(\gamma), \quad \gamma > 0,$$

como una aproximación para $\gamma = \gamma_{op}(\delta)$.

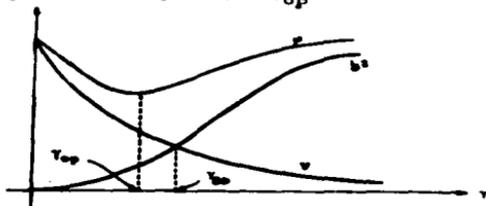


Fig. 1. Ilustración Geométrica del Principio de Sseudo-optimalidad.

Así pues, de existir $\gamma = \gamma_{op}(\delta)$, es de esperar que $\hat{\gamma}_{so} = \gamma_{op}$; y por ello, que las propiedades de convergencia de $\hat{\gamma}_Y$ a $\underline{\gamma}$ con $\gamma = \gamma_{so}(\delta)$ cuando $\delta \rightarrow 0^+$, sean semejantes a los de $\hat{\gamma}_Y$ con $\gamma = \gamma_{op}(\delta)$.

En el capítulo 3, para el problema en cuestión $Au = f_\delta$, $\|f_\delta - f\|_F \leq \delta$ con $A:U \rightarrow F$ lineal compacto se demuestra, si $PF = A(A'A)^v h^{(*)}$, $0 < v \leq 1$, para alguna $h \in U$, que

$$\|R_Y f_\delta - u^+\|_U = O(\delta^{2v/(2v+1)}),$$

para $\gamma = \gamma_{so}(\delta)$. En particular, para $v=1$, la razón de convergencia es inmejorable, pues lo mismo ocurre para $\gamma = \gamma_{op}(\delta)$.

En el mismo capítulo 3 se prueba, si $u \in \text{Im}(A'A)^v$, $0 < v \leq 1$, que

$$\sup_{\xi, \text{vt}(\xi) < \delta} \|\hat{u}_Y^+ - u\|_E = O(\delta^{2v/(2v+1)})$$

$$(\|\hat{u}_Y^+ - u\|_E = O(\|\hat{u}_Y^+ - u\|_U^2)^{v/2})$$

para $\gamma = \gamma_{so}(\delta)$, siendo $\hat{u}_Y^+ = R_Y f$ la estimación Gauss-Markov regularizada para u en el modelo lineal $Au = \xi = f$ con $A:U \rightarrow F$ como antes lineal compacto y en donde ξ es un vector aleatorio sobre F con $E(\xi) = 0$, en el sentido que de que para toda $w \in F$ dada, $E(\xi_w) = 0$ donde $\xi_w = \langle \xi, w \rangle_F$, y con varianza total

$$\text{vt}(\xi) = \sum_{i=1}^{\infty} E(\xi_{w_i})^2 < \delta^2$$

(*) Aquí, $A':F \rightarrow U$ denota al adjunto de A .

con respecto a cualquier conjunto ortogonal a lo más numerable $\{w_i\}_{i=1}^{\infty}$ de F dado. Y viendo que se tiene el mismo resultado para $\gamma = \gamma_{op}(\delta)$, es que se justifica el nombre de Seudo-optimalidad para nuestro principio para la construcción del criterio $\gamma = \gamma_{so}(\delta)$.

En el capítulo 4, se analiza la discretización de la regularización de la ecuación integral de Fredholm de primer tipo, viendo su liga con la estimación de Gauss-Markov regularizada, se discute el Principio de Seudo-optimalidad con base a las estimaciones insesgadas y sesgadas para la varianza total y el sesgo cuadrado, y finalmente se presenta un algoritmo "automático" con ejemplos numéricos para el cálculo de $\hat{\beta}_{\gamma}$ con $\gamma = \gamma_{so}(\delta)$, el cual se aplica a la resolución numérica vía regularización de la ecuación de Fredholm antes referida.

El capítulo siguiente tiene por objeto el estudio de las ideas centrales del Método de Regularización de Tijonov.

CAPITULO 2

METODO DE REGULARIZACION DE TIJONOV.

El objetivo de este capítulo es el presentar las ideas y los resultados básicos del método de regularización de Tijonov para la resolución del problema $Au = f$, $A:U \rightarrow F$ lineal compacto sobre los espacios de Hilbert U y F ; al interesado en los desarrollos teóricos y prácticos de este método en situaciones más generales y en diferentes direcciones, se le recomiendan los libros de Arsenin-Tikhonov [2] y Morozov [65].

§2.1. Antecedentes Históricos.

Aun cuando desde finales de los años 30's, en opinión de Moler (Miller [60] p. 57), se han venido usando ciertos métodos de "regularización" en la resolución de sistemas lineales algebraicos *mal-comportados* (ill-conditioned), no es sino hasta los 50's cuando aparece el trabajo pionero de Fox-Goodwin [29], en el cual se dan algunas explicaciones y ciertos remedios

a las dificultades que se presentan en la resolución de la ecuación integral (lineal) de Fredholm de 1a. clase

$$(2.1.1) \quad \int_a^b k(x,t) u(t) dt = f(x), \quad a \leq x \leq b,$$

vía la discretización por cuadratura numérica de la integración. Otro trabajo pionero de la década de los 60's, es el de Baker et al. [4], en el cual se discute el método por truncamiento del desarrollo espectral de la matriz resultante de la discretización de la ecuación integral (2.1.1). Para ver la extensión de esta idea en términos de la descomposición singular de una matriz, se recomiendan los trabajos de Hanson [43], y de Varah ([96], [97]).

Los trabajos claves para nuestros propósitos tienen lugar en 1962 y 1963, los cuales constituyen el origen de los que hoy se conocen como métodos de Phillips-Twomey ([74] y [32]), y de regularización de Tijonov ([88] y [89]).

Como hemos visto en la sec. 1.2, la discretización por cuadratura de la ecuación (ec.) integral (2.1.1) da origen, despreciando errores de discretización y de observación, a un sistema lineal algebraico $Bu = b$; en el cual la matriz B, como herencia de ser la ec. (2.1.1) un problema M-Pdo, resulta ser Malcomportada en general. En consecuencia, la solución $u^+ = B^+b$ viene a ser muy sensible a pequeñas perturbaciones, en particular al redondeo, en B y b. Lo que explica el por qué de la presencia

de las bruscas oscilaciones de la solución numérica, careciendo ésta, usualmente, de alguna utilidad práctica.

La idea de Phillips para superar tal situación es la siguiente.

Si se considera que en la práctica es muy común que a $f(x)$ se le conozca sólo aproximadamente, esto es, que se conozca $f_\delta(x) = f(x) + c(x)$ (con $\int_a^b |c(x)|^2 dx$ menor o igual a δ), entonces resulta natural buscar al miembro de la familia

$$F_\delta = \{u(x) \text{ en } W_2^1[a,b] \mid \int_a^b k(x,t)u(t) dt = f_\delta(x) \\ \text{y } \int_a^b |c(x)|^2 dx \leq \delta^2\}, \delta > 0 \text{ dada,}$$

lo menos oscilante posible. Una forma de lograr tal propósito consiste en determinar a aquel miembro de F_δ que minimice a la funcional

$$(2.1.2) \quad S[u] = \int_a^b |u''(s)|^2 ds.$$

Lo cual da lugar al problema variacional siguiente:

$$(2.1.3) \quad \min S[u]$$

sujeto a

$$(2.1.3') \quad \int_a^b |r(s)|^2 ds \leq \delta^2,$$

donde $r(x) = \int_a^b k(x,t) u(t) dt - f_\delta(x)$.

La discretización (con partición uniforme) de este problema toma la forma

$$(2.1.4) \quad \min_{j=1}^p (u_{j+1} - 2u_j + u_{j-1})^2$$

sujeto a

$$(2.1.4') \quad \sum_{i=1}^n \left(\sum_{j=1}^p k_{ij} w_j u_j - f_i^{\delta} \right)^2 \leq \delta^2.$$

Cuyo lagrangiano viene a ser

$$(2.1.5) \quad L(u; \gamma^{-1}) = \sum_{j=1}^p (u_{j+1} - 2u_j + u_{j-1})^2 + \gamma^{-1} (H u - b)^2 - \delta^2,$$

donde

$$u = (u_1, \dots, u_p)^t \quad \text{y} \quad b = (f_1^{\delta}, \dots, f_n^{\delta})^t;$$

$$B = KD; \quad K = (k_{ij}), \quad k_{ij} = k(x_i, t_j);$$

$$D = \text{diag}(w_1, w_2, \dots, w_p).$$

Derivando y tomando las condiciones de frontera naturales $\Delta^2 u_j = \Delta^2 u_j = 0$, $j=1$ y $j=p$; se tiene que la solución de (2.1.4-4'), se obtiene resolviendo el sistema lineal

$$(2.1.6) \quad (B^t B + \gamma H) u = B^t b.$$

donde H es la matriz (pentadiagonal) de cuartas diferencias con condiciones de frontera $\Delta^2 u_j = \Delta^2 u_j = 0$, para $j=1$ y $j=p$. El parámetro (de regularización) se determina mediante el método universal de

prueba y error con base en ciertas heurísticas de carácter práctico.

Es conveniente dejar asentado, que si bien ya se usaban tiempo atrás procedimientos similares (en apariencia) al método de regularización de Tijonov, corresponde reconocer a Tijonov el haber puesto la primera piedra sobre la fundamentación matemática de la que hoy es toda una rama del Análisis, que en sus aspectos numéricos constituye uno de los principales pivotes del desarrollo de la Matemática Numérica actual (Courant-Hilbert [19], p. 231; Morozov [65], p. 6; y Nashed [67], p. 290).

En la sección que sigue se bosquejan las ideas iniciales sobre el método de regularización de Tijonov ([88], [89] y [2]).

§2.2. Método de Regularización de Tijonov.

Considérese a la ecuación integral (2.1.1) en su forma operacional

$$(2.2.1) \quad Au = f,$$

donde $A: H \rightarrow F(H = F = L_2[a, b])$ está dado por

$$A[u](x) = \int_a^b k(x, t) u(t) dt, \quad a < x < b,$$

con $k(x, t)$ un elemento en $L_2([a, b] \times [a, b])$. Como es bien sabido A es un operador lineal compacto, el cual supondremos invertivo, por el momento.

La primera de las ideas centrales de Tijonov ([88] y [89]) está inmersa en el siguiente concepto de familia de operadores regularizantes, la cual está pensada en términos de un operador $A: H \rightarrow F$ no necesariamente lineal y/o continuo (compárese con la definición 1.4.1).

Definición 2.2.1 (Tijonov) Sea $u^0 (\in U)$ la solución de la ecuación $Au = f$. La familia de operadores $\{R_\gamma\}_{\gamma > 0}$ donde cada miembro R_γ está definido sobre una vecindad de $f (\in F)$ en H , se llamará *regularizante* para $Au = f$, si se tiene que:

- (a). Existe $\delta_1 > 0$ de manera que todo R_γ está definido sobre la vecindad V_{δ_1} de f . Y,
 (b). Se cuenta con un criterio $\gamma = \gamma(\delta)$ de forma que, para toda $\varepsilon > 0$ dada, existe $\delta = \delta(\varepsilon) < \delta_1$ tal que

$$\|R_{\gamma(\delta)} f_\delta - u^0\|_H < \varepsilon, \text{ siempre que } \|f_\delta - f\|_F < \delta. \quad \#$$

Nótese que: (1). la condición (a) se satisface trivialmente si $R_\gamma: F \rightarrow H$, para toda $\gamma > 0$; y (2) la condición (b) implica, en particular, que $\{R_\gamma\}_{\gamma > 0}$ es una familia de operadores continuos con (*) $R_{\gamma(\delta)} Au^0 \xrightarrow{H} u^0$ cuando $\delta \rightarrow 0^+$, e inversamente.

La segunda de las ideas claves va en la dirección de la construcción de una familia regularizante. Para tal efecto, considérese una funcional que llamaremos *estabilizante*

$$S: U \subset H \rightarrow \mathbb{R}, \quad (U \subset H \text{ denso}),$$

no-negativa continua y tal que, para toda $c > 0$ dada,

$$(*) \quad R_{\gamma(\delta)} Au^0 \xrightarrow{H} u^0 \text{ significa que } \|R_{\gamma(\delta)} Au^0 - u^0\|_H \rightarrow 0.$$

$$S_c = \{u \in U \mid S(u) < c\}$$

es un conjunto compacto en H . Como un ejemplo típico, se tiene a la funcional dada por

$$(2.2.3) \quad S(u) = \int_a^b \sum_{j=0}^m p_j(t) |u^{(j)}(t)|^2 dt$$

en donde $p_j(t)$ es continua no-negativa para $j=0,1,\dots,m-1$; $p_m(t)$ es continua positiva, y definida sobre el espacio de Sobolev $(@)$ $U = W_2^m[a,b]$. Por los teoremas de inmersión de Sobolev (Smirnov [82] vol. 5, p. 353), se tiene que $S_c (c > 0)$ es un conjunto compacto en $C^{m-1}[a,b]$ $(*)$; y con ello en H , pues la inyección $C^{m-1}[a,b] \subset H$ es continua.

Ahora, considérese a la funcional

$$(2.2.4) \quad \Phi_\gamma[u; f] = \|f - Au\|_p^2 + \gamma S(u), \quad \gamma > 0 \text{ dada,}$$

la cual llamaremos *regularizante* de Tijonov.

Para nuestro caso, A lineal sobre espacios de Hilbert y S como en (2.2.3), es directo verificar la identidad que sigue, tipo ley del paralelogramo:

(@). Aquí, $U = W_2^m[a,b]$ es el espacio de completación de $C^m[a,b]$ con respecto a la norma $\|u\|_U^2 = S(u)$.

(*). $C^{m-1}[a,b]$ es espacio de Banach con norma:

$$\|g\|_{\infty, m-1} = \max_{0 \leq i < m-1} \left\{ \max_{[a,b]} |g^{(i)}(t)| \right\}.$$

$$(2.2.4) \quad \Phi_Y\left[\frac{u-v}{2}, 0\right] = \frac{1}{2} \Phi_Y[u; f] + \frac{1}{2} \Phi_Y[v; f] \\ - \Phi_Y\left[\frac{u+v}{2}, f\right]$$

Teorema 2.2.1. Para toda $f \in F$ dada, existe una única u_Y en U tal que $\Phi_Y[u_Y; f] \leq \Phi_Y[u; f]$, para toda $u \in U$.

Demostración: (*) Por ser $\Phi_Y[u] \equiv \Phi_Y[u; f]$ no-negativa, existe $m_Y > 0$ tal que $m_Y = \inf \Phi_Y[u]$ sobre U . Sea $\{u_n\} \subset U$ una sucesión minimizante dada.

Por (2.2.4'), se tiene que

$$\Phi_Y\left[\frac{u_n - u_m}{2}, 0\right] \leq \frac{1}{2} \Phi_Y[u_n; f] + \frac{1}{2} \Phi_Y[u_m; f] - m_Y,$$

que junto con la desigualdad

$$S[u_n - u_m] \leq \frac{4}{Y} \Phi_Y\left[\frac{u_n - u_m}{2}, 0\right],$$

(*) La siguiente demostración de existencia (Tijonov) sólo supone a $S: U \subset \mathbb{H} \rightarrow \mathbb{R}$ no negativa con la propiedad de que, para toda $c > 0$, S_c es un compacto en U , y no hace uso explícito de la linealidad de A : Sean $0 < m_Y = \inf \Phi_Y[u; f]$ sobre U , y $\{u_n\} \subset U$ una sucesión minimizante con $\Phi_Y[u_1; f] > \Phi_Y[u_2; f] > \dots$; como $\{u_n\} \subset S_c$, $c = \frac{1}{Y} \Phi_Y[u_1; f]$, existen $u_Y \in S_c(\subset U)$ y $\{u_{n'}\}$ subsucesión de $\{u_n\}$ tal que $u_{n'} \rightarrow u_Y$ en U (con ello, en \mathbb{H} , si la inyección $U \subset \mathbb{H}$ es continua). De la continuidad de A , se tiene que $m_Y = \Phi_Y[u_Y; f]$.

Nótese que S , dada por (2.2.3), hace a S_c compacto en U , pensando en $C^{m-1}[a, b]$.

implican que $\{u_n\}$ es de Cauchy en U . Luego, existe $u_Y \in U$ tal que $u_n \rightarrow u_Y$ en U .

De la continuidad de A se obtiene que $m_Y = \Phi_Y[u_Y; f]$.

La unicidad se sigue de la desigualdad

$$\Phi_Y\left[\frac{u-v}{2}; 0\right] \leq \frac{1}{2}\Phi_Y[u; f] + \frac{1}{2}\Phi_Y[v; f] - m_Y,$$

válida para toda u y v en U . \square

A u_Y se le puede hallar resolviendo la correspondiente ecuación de Euler del problema $\text{Min } \Phi_Y[u; f]$ sobre U , la cual resulta ser la ecuación integro-diferencial:

$$(2.2.5) \quad (A^*A + \gamma L) u = A^*f,$$

donde $A^* : F \rightarrow H$ es el adjunto de A , el cual está dado por

$$A^*[u](x) = \int_a^b k(t, x) u(t) dt,$$

y $L : D(L) \subset U \rightarrow F$ es el operador diferencial dado por

$$(2.2.6) \quad L[u](x) = \sum_{j=0}^m (-1)^j \frac{d^j}{dx^j} (p_j(x) u^{(j)}(x))$$

con las condiciones de frontera, por ejemplo, naturales:

$$(2.2.6') \quad u^{(j)}(a) = u^{(j)}(b) = 0, \quad j = m, m+1, \dots, 2m-1,$$

bajo las cuales L resulta ser un operador auto-adjunto con subespacio núcleo trivial. Luego, L tiene inverso, el cual viene dado por

$$L^{-1}[g](x) = \int_a^b G(x, t) g(t) dt,$$

siendo $G(x,t)$ la correspondiente función de Green de L .

Ahora, aplicando L^{-1} por la izquierda en (2.2.5), se obtiene la ecuación integral de Fredholm de 2a. especie.

$$(2.2.7) \quad (K + \gamma I)u = h, \quad K = L^{-1}A^*A, \quad h = L^{-1}A^*f,$$

que por la alternativa de Fredholm [78], siempre tiene una única solución u_γ en U , la cual depende continuamente de $f \in F$, pues $K + \gamma I$ tiene inverso continuo debido a que su correspondiente ecuación homogénea sólo admite la solución trivial. (*)

En conclusión, para $\gamma > 0$ dada, el operador

$$(2.2.8) \quad R_\gamma : F \rightarrow U, \text{ dado por } f \mapsto u_\gamma,$$

es lineal y continuo; y con ello cumplen la primera condición de la definición 2.2.1. Para ver que también se cumple la segunda condición para la familia $\{R_\gamma\}_{\gamma > 0}$ determinada por (2.2.8), veamos que

$$(2.2.9) \quad \|R_\gamma f_\delta - u^0\|_U \leq \|R_\gamma(f_\delta - f)\|_U + \|R_\gamma f - u^0\|_U \\ < \|R_\gamma\| \cdot \delta + \|R_\gamma f - u^0\|_U,$$

para $f_\delta \in F$ tal que $\|f_\delta - f\|_F \leq \delta$, siendo u^0 la solución de $Au = f$. De la cual se sigue que $\{R_\gamma\}_{\gamma > 0}$ es una familia regularizante pa

(*). Si $(K + \gamma I)u = 0 \implies \gamma \langle Lu, u \rangle_H = - \langle A^*Au, u \rangle_H \implies \gamma \|u\|_U^2 = - \|Au\|_F^2 \implies u = 0$, al ser A inyectivo.

ra $Au = f$, si se toman γ y $\delta \rightarrow 0^+$ de manera que $\|R_\gamma\| \cdot \delta \rightarrow 0^+$; y se prueba el siguiente:

Teorema 2.2.2. Para toda $u \in U$, $\|R_\gamma Au - u\|_U \rightarrow 0$ cuando $\gamma \rightarrow 0^+$ (De hecho, se tiene que $\|R_\gamma Au - u\|_{U, m-1} \rightarrow 0$ cuando $\gamma \rightarrow 0^+$).

Demostración: Como

$$(*) \quad \Phi_\gamma[u_\gamma; Au] \leq \Phi_\gamma[u; Au] = \gamma S[u],$$

se tiene que $u_\gamma \in S_C (C = S[u])$, ya que (por $(*)$) $S[u_\gamma] \leq \frac{1}{\gamma} \Phi_\gamma[u_\gamma; Au] \leq S[u]$. Y como $A|_{S_C} : S_C \subset U \rightarrow AS_C$ es continuo y biyectivo sobre S_C un conjunto compacto de $C^{m-1}[a, b]$, se sabe que $(A|_{S_C})^{-1} : AS_C \rightarrow S_C (CU) \subset C^{m-1}[a, b]$ es continuo. Así, la demostración estará completa al ver que $\|Au_\gamma - Au\|_F \rightarrow 0$ cuando $\gamma \rightarrow 0^+$. Lo cual se sigue de que $\|Au_\gamma - Au\|_F^2 \leq \Phi_\gamma[u_\gamma; Au] \leq \gamma C$, que se obtiene directamente de $(*)$. #

Ahora, para ver que es posible hacer γ y $\delta \rightarrow 0^+$ de manera que $\|R_\gamma\| \cdot \delta \rightarrow 0$, se hace necesario estimar $\|R_\gamma\|$. Para ello, observando que $S[u] = \|u\|_U^2$, de la funcional regularizante (2.2.4) escrita como

$$(2.2.10) \quad \Phi_\gamma[u; f] = \|f - Au\|_F^2 + \gamma \|u\|_U^2,$$

se obtiene la reformulación siguiente de la ecuación de Euler

$$(2.2.11) \quad (A^*A + \gamma I)u = A^*f,$$

pensando a $A = A|_U : U(CH) \rightarrow F^{(*)}$ y siendo $A' : F \rightarrow U$ el adjunto de A con respecto a U y F . De donde se sigue que

$$(2.2.12) \quad R_Y = (A'A + \gamma I)^{-1} A',$$

y se obtiene, como puede verse en Riesz-Nagy [78] (apartado 118), que

$$(2.2.13) \quad \|R_Y\| \leq \frac{1}{\sqrt{\gamma}}.$$

En resumen, si se toman γ y $\delta \rightarrow 0^+$ de manera que $\frac{\delta}{\sqrt{\gamma}} \rightarrow 0$ entonces se tiene que $R_Y f_\delta \rightarrow u^*$ en H (de hecho en U). Con lo cual queda demostrado que la familia $\{R_Y\}_{\gamma > 0}$ construida a partir del problema variacional $\text{Min } \Phi_Y[u; f]$ sobre U , es una familia regularizante de Tikonov para el problema $Au = f$, para $A : H \rightarrow F$ lineal y compacto.

Ahora, de (2.2.7) y (2.2.11) se sugiere que $A' = L^{-1} A^*$. Para demostrar tal hecho, ^(**) considérese el producto interior (...) sobre $D(L) \subset H$ definido por $(u, v) = \langle Lu, v \rangle_H$ con L dado por (2.2.6-6'). Se tiene que: (a) $\Omega[u] = \langle Lu, u \rangle_H$; y (b) la completación de $D(L) \subset H$ con respecto al producto interior (...) es precisamente $U (= W_2^m[a, b])$. En estos términos, para $u \in D(L)$ y $v \in F$ dados, se tiene que

(*) Como la inyección UCH es continua, de hecho compacta, $A = A|_U : U \rightarrow F$ es compacto.

(**) Véase a Oganessian et. al. [71]

$$\begin{aligned} \langle u, A^* v \rangle_H &= \langle Au, v \rangle_F = \langle u, A' v \rangle_U \\ &= \langle Lu, A' v \rangle_H = \langle u, LA' v \rangle_H, \end{aligned}$$

y consecuentemente, que

$$(2.2.14) \quad A^* = LA'.$$

Para el cálculo numérico de u_γ , puede resultar conveniente considerar la ecuación integral (2.2.11) (véase Cullum [22]), en vez de la ecuación integro-diferencial (2.2.5) como inicialmente propone Tijonov. También conviene decir que la determinación de $A' : F \rightarrow U$ puede resultar bastante más laboriosa que la de $A^* : F \rightarrow H$, pues la primera involucra el cálculo de la función de Green de L .

Ahora bien, como hemos visto, el método variacional propuesto por Tijonov para la construcción de la familia regularizante $\{R_\gamma\}_{\gamma>0}$ para $Au = f$ ($A : H \rightarrow F$ lineal y compacto) consiste en: (a) considerar un operador lineal (usualmente diferencial e inyectivo) $L : D(L) \subset U \rightarrow H$, $\overline{D(L)} = U$ respecto a U , siendo U un espacio de Hilbert denso en H con la inyección natural $U \subset H$ compacta, U es elegido por supuestos a priori sobre las propiedades de suavidad deseables con respecto a la solución de $Au = f$; (b) pensar $A (= A|_U) : U \rightarrow F$; y (c) tomar a $\Phi_\gamma[u; f]$ como en (2.2.10), como funcional regularizante de Tijonov, o lo que es lo mismo, a tomar $\{R_\gamma\}_{\gamma>0}$ con $R_\gamma = (A'A + \gamma I)^{-1} A'$ como

familia regularizante para $Au = f$. En pocas palabras, en pensar a $A : U \rightarrow F$ y considerar la funcional regularizante (2.2.10). Es en estos últimos términos como procederemos en lo que resta del presente trabajo.

Cabe mencionar que existen otros procedimientos para la construcción de las familias regularizantes $\{R_\gamma\}_{\gamma > 0}$, entre ellos se tiene al método operacional (universal) de Bakushinskii ([5] y [6]), donde $R_\gamma = \psi_\gamma(A^*A)A^*$, siendo $\psi_\gamma(\lambda)$ una función real con dominio en $[0, \infty)$ conteniendo al espectro $\sigma(A^*A)$ de A y cumpliendo con ciertas propiedades que hacen a $\psi_\gamma(\lambda) = (\lambda + \gamma)^{-1}$, uno de sus casos representativos. Nótese que la función $\psi_\gamma(\lambda) = (\lambda + \gamma)^{-1}$ corresponde al procedimiento variacional de Tijonov (véase a (2.2.12)). El procedimiento de Bakushinskii permite considerar el caso en el cual $A : D(A) \subset U \rightarrow F$ es lineal no-acotado con A^*A (actuando en U) auto-adjunto positivo.

§2.3. Teorema de Picard y Convergencia del Método de Regularización.

Como ya hemos mencionado, (sec. 1.3), la hipótesis de inyectividad de $A : U \rightarrow F$ no es indispensable, si la solución de $Au = f$ se piensa en el sentido generalizado de Moore-Penrose (M-P). Con el propósito de presentar el método de regularización en dicho contexto, se vuelve conveniente el hacer un breve recuento sobre la solubilidad de $Au = f$, en donde $A : U \rightarrow F$ es lineal y

compacto, siendo U y F espacios de Hilbert (*) con productos interiores $\langle \dots \rangle_U$ y $\langle \dots \rangle_F$ respectivamente.

Empecemos por recordar que todo operador lineal compacto sobre espacios de Hilbert admite, como puede verse en Groersch [39], el desarrollo siguiente, llamado de valores singulares

$$(2.3.1) \quad A = \sum_{i=1}^{\infty} s_i \langle \cdot, v_i \rangle_U w_i,$$

donde $s_i > 0$, $s_i \rightarrow 0$ cuando $i \rightarrow \infty$, son los llamados *valores singulares* de A ; $\{v_i\}$ un sistema ortonormal en U , llamado de *vectores singulares derechos*, que es una base de Schauder para $N(A)^\perp = \overline{\text{Im}(A')}$; y $\{w_i\}$ un sistema ortonormal en F , llamado de *vectores singulares izquierdos*, que es una base de Schauder para $\overline{\text{Im}(A)} = N(A')^\perp$. Aquí, $A' : F \rightarrow U$ es el adjunto de A (i.e. $\langle Au, v \rangle_F = \langle u, A'v \rangle_U$, para toda $u \in U$ y toda $v \in F$).

No es difícil ver que

$$(2.3.2) \quad \begin{aligned} A'Av_i &= s_i^2 v_i \\ AA'w_i &= s_i^2 w_i \end{aligned} \quad , i=1,2,\dots$$

Más aún, se tienen los siguientes desarrollos espectrales

$$(2.3.3) \quad \begin{aligned} A'A &= \sum_{i=1}^{\infty} s_i^2 \langle \cdot, v_i \rangle_U v_i \\ AA' &= \sum_{i=1}^{\infty} s_i^2 \langle \cdot, w_i \rangle_F w_i, \end{aligned}$$

(*) . Por comodidad, nosotros consideramos sólo el caso real.

de $A'A$ y AA' , respectivamente. Esto es, $\{v_i\}$ y $\{w_i\}$ son los sistemas de vectores propios de $A'A$ y AA' respectivamente, con los correspondientes mismos valores propios.

El siguiente resultado da una representación explícita de u^+ , $D(A^+)$ y de $A^+ : D(A^+) \subset F \rightarrow U$.

Teorema 2.3.1 (Picard). Sea $A : U \rightarrow F$ lineal compacto y con desarrollo singular (2.3.1). Dado $f \in F$, la ecuación $Au = f$ tiene solución generalizada (M-P) u^+ si y sólo si

$$(2.3.4) \quad \sum_{i=1}^{\infty} \frac{|\langle f, w_i \rangle_F|^2}{s_i^2} < \infty.$$

Si tal es el caso, entonces,

$$(2.3.5) \quad u^+ = \sum_{i=1}^{\infty} \frac{\langle f, w_i \rangle_F}{s_i} v_i.$$

Demostración. Por el teorema de la proyección ortogonal, dado $f \in F$, se tiene que $f = Pf + h$, con $Pf \in \overline{\text{Im}(A)}$ y $h \in \text{Im}(A)^\perp$. Y como $\{w_i\}$ es una base de Schauder de $\overline{\text{Im}(A)}$, se sigue que

$$Pf = \sum_{i=1}^{\infty} \langle Pf, w_i \rangle_F w_i = \sum_{i=1}^{\infty} \langle f, w_i \rangle_F w_i.$$

Suficiencia: Por demostrar que $Pf \in \text{Im}(A)$ (véase la Proposición 1.3.1). Como $\sum_{i=1}^{\infty} \frac{|\langle f, w_i \rangle_F|^2}{s_i^2} < \infty$, se tiene que

$u_0 = \sum_{i=1}^{\infty} \frac{\langle f, w_i \rangle_F}{s_i} v_i \in U$. Y ya que $Au_0 = \sum_{i=1}^{\infty} \langle f, w_i \rangle_F w_i$, se tiene que $Pf \in \text{Im}(A)$.

Necesidad: Si u^+ existe entonces $Au^+ = Pf^{(*)}$ con $u^+ \in N(A)^1$. Como $\{v_i\}$ es una base de Schauder de $N(A)^1$, se tiene que $u^+ = \sum_{i=1}^{\infty} \alpha_i v_i$ ($\sum_{i=1}^{\infty} \alpha_i^2 = \|u^+\|_U^2 < \infty$). De la relación, $Au^+ = (-Pf) = \sum_{i=1}^{\infty} \langle f, w_i \rangle_P w_i$ se sigue que $\sum_{i=1}^{\infty} \alpha_i s_i v_i = \sum_{i=1}^{\infty} \langle f, w_i \rangle_P w_i$; de donde $\alpha_i = \frac{\langle f, w_i \rangle_P}{s_i}$, $i=1, 2, \dots$ (i.e. $u^+ = \sum_{i=1}^{\infty} \frac{\langle f, w_i \rangle_P}{s_i} v_i$), y $\sum_{i=1}^{\infty} \frac{|\langle f, w_i \rangle_P|^2}{s_i^2} = \sum_{i=1}^{\infty} \alpha_i^2 < \infty$. \square

Como consecuencia,

$$(2.3.6) \quad D(A^+) = \text{Im}(A) \circ \overline{\text{Im}(A)}$$

$$= \{f \in F \mid \sum_{i=1}^{\infty} \frac{|\langle f, w_i \rangle_P|^2}{s_i^2} < \infty\}, \text{ y}$$

$$(2.3.6') \quad u^+ = A^+ f = \sum_{i=1}^{\infty} \frac{\langle f, w_i \rangle_P}{s_i} v_i.$$

La interpretación de este teorema es muy importante, pues nos dice que para que $Au = f$ tenga solución generalizada u^+ , f debe ser suficientemente suave, en el sentido que sus coeficientes de Fourier $\langle f, w_i \rangle_P$ tiendan, con respecto a s_i , suficientemente rápido a cero de manera que (2.3.4) tenga lugar. Esto es, que los coeficientes de Fourier $\langle f, w_i \rangle_P$ de f sean prácti-

(*) . Por el teorema de la proyección, para toda $u \in U$, $\|f - Au\|_F^2 = \|(Pf + h) - Au\|_F^2 = \|Pf - Au\|_F^2 + \|h\|_F^2 \geq \|h\|_F^2$. Si u^0 es una solución del problema $\text{Min } \|f - Au\|_F^2$ (en U), entonces $\|f - Au^0\|_F^2 = \|h\|_F^2$, lo que implica que $\|Pf - Au^0\|_F^2 = 0$, o bien que $Au^0 = Pf$.

camente cero con respecto a s_1 , en las direcciones de las "altas frecuencias" s_1^{-1} .

Con base en el siguiente resultado, se obtiene (considerando a (2.2.9) y (2.2.13)), que el problema $Au=f$ es regularizable a la Tifonov, siempre que u^+ exista (i.e. que $f \in D(A^+)$).

Teorema 2.3.2. Si $f \in D(A^+)$ (i.e. si u^+ existe) entonces $R_\gamma f \rightarrow u^+$ (en U) cuando $\gamma \rightarrow 0^+$.

Demostración: En términos de (2.3.1) se tiene que

$$R_\gamma f = \sum_{i=1}^{\infty} \frac{s_i \langle f, w_i \rangle F}{\gamma + s_i^2} v_i,$$

y por (2.3.6') que

$$u^+ = A^+ f = \sum_{i=1}^{\infty} \frac{\langle f, w_i \rangle F}{s_i} v_i.$$

Luego,

$$R_\gamma f - u^+ = - \sum_{i=1}^{\infty} \frac{\gamma}{(s_i^2 + \gamma)} \frac{\langle f, w_i \rangle F}{s_i} v_i,$$

de donde

$$\begin{aligned} \|R_\gamma f - u^+\|_U^2 &= \sum_{i=1}^{\infty} \left(\frac{\gamma}{s_i^2 + \gamma} \right)^2 \frac{|\langle f, w_i \rangle F|^2}{s_i^2} \\ &< \sum_{i=1}^{\infty} \frac{|\langle f, w_i \rangle F|^2}{s_i^2} < +\infty. \end{aligned}$$

Así, por el teorema de la convergencia acotada, $\|R_\gamma f - u^+\|_U \rightarrow 0$ cuando $\gamma \rightarrow 0^+$. \square

En otras palabras, este teorema dice que R_γ se aproxima a A^+ cuando $\gamma \rightarrow 0^+$, en el sentido que $R_\gamma f \rightarrow A^+ f$ (en U), para toda $f \in D(A^+)$. Pero, desafortunadamente no da información sobre su razón de convergencia. De hecho, ésta puede resultar extremadamente lenta (con ello también la de $R_{\gamma(\delta)} f_\delta$ a u^+), como puede verse en Franklin [30].

Sin embargo, como veremos en el capítulo siguiente, bajo ciertos razonables supuestos a priori sobre la suavidad de la solución u^+ más exigentes que $Pf \in \text{Im}(A)$ (equivalentemente (2.3.4)), es posible determinar la razón de convergencia de $R_\gamma f$ a $u^+ = A^+ f$.

Por último, debe decirse que la condición que $f \in D(A^+)$ de este último teorema es también una condición necesaria. Pues se puede demostrar, si $f \notin D(A^+)$, que existe $\{\gamma_n\}$ tendiendo a cero de manera que $\{\|R_{\gamma_n} f\|_U\}$ no es acotada (Groetsch [39]).

CAPITULO 3

PRINCIPIO DE SEUDO-OPTIMALIDAD PARA LA ELECCION DEL PARAMETRO DE REGULARIZACION

El objetivo de este capítulo es el estudio desde los puntos de vista métrico y estadístico de una alternativa al problema (de capital importancia práctica) sobre la propuesta de un *criterio* (o *principio*) que llamaremos de *Seudo-optimalidad*^(*) para la *elección* del parámetro de regularización, esto es, para la construcción de la función $\gamma = \gamma(\delta)$ de manera - que $R_{\gamma}(\delta) \rightarrow u^+$ cuando $\delta \rightarrow 0^+$ con aceptable razón de convergencia con respecto a δ .

§3.1. Selección del Parametro de Regularización.

Hasta el momento, se ha dejado de lado el *problema de la selección del parametro de regularización*, el cual, como ya se

(*) Por razones que, como se verá, son transparentes.

ha mencionado, resulta ser una de las más serias dificultades, sobre todo prácticas, del método de regularización. Y aún cuando el objetivo del presente capítulo es el estudio de una alternativa a tal problema, cabe decir que en la actualidad se cuenta con varios principios para la construcción (ninguna explícita) del criterio $\gamma = \gamma(\delta)$ de elección del parámetro de regularización. A continuación se describen muy brevemente los más sobresalientes, sobre todo por su efectividad en la práctica.

El primero que se estudió y desarrolló, principalmente por la escuela soviética, se le conoce por el Principio de Discrepancia (Morozov ([62], [63] y [65]) y Groetsch [39]), y consiste, conocida $\delta > 0$, en tomar $\gamma = \gamma_d(\delta)$ de manera que

$$\rho(\gamma) (\equiv \|Au_Y^S - f_\delta\|_F) = c\delta \quad (0 < c < 1)$$

donde

$$u_Y^\delta = R_Y f_\delta.$$

Esta idea surge de manera muy natural a partir del problema variacional

$$\text{Min} \|u\|_U^2$$

$$\text{s. a. } \|Au - f_\delta\|_F \leq c\delta.$$

Bajo condiciones adecuadas se demuestra que: (1). $\gamma = \gamma_d(\delta)$ existe y es única; y (2). $u_{Y_d}^\delta \rightarrow u^0$ (en U) cuando $\delta \rightarrow 0^+$, u^0 la solución de $Au = f$.

Es importante decir que salvo el resultado siguiente (Groetsch [39]): "Si $u^0 = A^0 h$ ($h \in F$) entonces $\|u_Y^\delta - u^0\|_U \leq 2 \sqrt{\lambda} \|\delta\|$ ($\gamma = \gamma_\delta(\delta)$)"; poco se sabe con respecto a la razón de convergencia de u_Y^δ ($\gamma = \gamma_\delta(\delta)$) a u^0 . Y cabe agregar que para la aplicación de este principio es indispensable el conocimiento de δ , lo cual es un inconveniente debido a que con frecuencia dicho conocimiento no está disponible.

El Principio de Quasi-óptimalidad desarrollado principalmente por Glasko y Tijonov ([34], [35] y [32]), permite la elección del parámetro $\gamma = \gamma(\delta)$ sin presuponer el conocimiento de δ explícitamente. Este principio consiste en tomar a $\gamma = \gamma_{q-op}(\delta)$ como el primer punto de mínimo local de la función de *quasi-óptimalidad*

$$\psi(\gamma) = \| \gamma \frac{du_Y^\delta}{d\gamma} \|_U,$$

la cual surge de simplificar el cálculo de hallar $\gamma = \gamma_{op}(\delta)$, que corresponde a determinar en donde la función de optimalidad

$$\psi_{op}(\gamma) = \max(\|u_Y^\delta - u^0\|_U \mid u_Y^\delta = R_\gamma f_\delta, \|f_\delta - f\|_F < \delta)$$

alcanza su mínimo. En términos probabilísticos se demuestra: (1) que $\gamma = \gamma_{q-op}(\delta)$ existe; y (2) que $u_Y^\delta \rightarrow u^0$ ($\gamma = \gamma_{q-op}(\delta)$) cuando $\delta \rightarrow 0^+$, con probabilidad uno. Sin embargo, no se conocen resultados sobre su razón de convergencia con respecto a $\delta \rightarrow 0^+$.

El Principio de la Validación Cruzada Generalizada de G. Wahba [100] es uno de los principios competitivos al anterior, en el sentido de que tampoco requiere del conocimiento explícito de δ . Para indicar en que consiste la idea de Wahba, considérese el problema

$$Au=f \quad (A[u](x) = \int_0^1 k(x,t)u(t) dt),$$

donde $A:U \rightarrow F$, $U=H_R$ un espacio de Hilbert con kernel reproductor $R(x,t)$ (*) o bien $U=L_2[0,1]$, $F=L_2[0,1]$, en el cual f es dada por una tabla de sus valores $T_{\xi\delta}$ conteniendo errores. Esto es

$$T_{\xi\delta} : y_i = f(x_i) + c_i; \quad 0 < x_1 < x_2 < \dots < x_n;$$

$$E\{c_i\} = 0, \quad i=1(1)n; \quad \text{y} \quad E\{c_i c_j\} = \delta^2 \delta_{ij}.$$

Suponiendo que el kernel

$$Q(t,x) = \begin{cases} \int_0^1 k(t,\xi)k(x,\xi)d\xi, & \text{si } U=L_2[0,1] \\ \int_0^1 \int_0^1 k(t,\xi)k(x,\xi)R(\xi,\zeta)d\xi d\zeta, & \text{si } U=H_R \end{cases}$$

es continuo y que f es una función "muy suave" (i.e. que $f \in D((AA')^+)$, Groetsch [39]), considérese la funcional regularizante

(*) Véase a Mashed-Wahba ([68], [69]) y Mashed [67] sobre el Método de Regularización y Espacios de Hilbert $k.R.$

$$\Phi_{\gamma}[u; y] = \frac{1}{n} \|(Au)_n - y\|_2^2 + \gamma \|u\|_U^2$$

$$((Au)_n = (A[u](x_1), \dots, A[u](x_n))^T; y = (y_1, \dots, y_n)^T).$$

Sea $u_{n,\gamma}$ (en U) donde $\Phi_{\gamma}[u; y]$ alcanza su mínimo. Para elegir a γ de forma que se dé un buen balance entre la medida de suavidad $\|u_{n,\gamma}\|_U^2$ y la medida de infidelidad a los datos $\frac{1}{n} \|(Au)_n - y\|_2^2$ de $u_{n,\gamma}$, se considera a $u_{n,\gamma}^{(k)}$ donde

$$\frac{1}{n} \sum_{\substack{j=1 \\ j \neq k}}^n (A[u](x_j) - y_j)^2 + \gamma \|u\|_U^2$$

se minimiza, para $k=1, 2, \dots, n$. Si γ^0 es una buena elección entonces, en promedio, $A[u_{n,\gamma}^{(k)}](x_k) \approx y_k$. Para $\gamma > 0$, el error cuadrático medio y pesado de predicción se mide por la función de *Validación Cruzada Generalizada* (VCG)

$$V(\gamma) = \frac{1}{n} \sum_{k=1}^n (A[u_{n,\gamma}^{(k)}](x_k) - y_k)^2 w_k(\gamma),$$

en donde los pesos $w_k(\gamma)$ se escogen de manera que, bajo ciertas hipótesis modelo:

(a) la función de VCG se puede escribir como

$$V(\lambda) = \frac{\frac{1}{n} y^T (Q_n + n\gamma I)^{-2} y}{\left[\frac{1}{n} \text{Traza} (Q_n + n\gamma I)^{-1} \right]^2},$$

siendo $Q_n = (q_{ij})$, $q_{ij} = Q(x_i, x_j)$, $x_i = \frac{1}{n}$, $i=1, 2, \dots, n$;

(b) $\gamma_{VCG} \approx \gamma^*$, para n grande,

siendo γ_{VCG} donde $EV(\gamma)$ alcanza su mínimo y γ^* donde $ET(\gamma)$ alcanza su mínimo, siendo $T(\gamma) = \frac{1}{n} \sum (Au)_n - \gamma \mathbb{N}_i^2$ el error cuadrático medio de predicción; y

(c). $E\|A^+f - u_{n, \gamma_{VCG}}\|_U^2 \rightarrow 0$ cuando $n \rightarrow \infty$

La demostración que da Wahba para (b) requiere el mantener fija a δ^2 ; por ello, no se puede hablar de su validez para $\delta \rightarrow 0^+$ (Wahba [100], p. 662). Así, no se sabe (por lo menos hasta lo que el que suscribe conoce) que el principio de VCG define un algoritmo de regularización a la Tijonov. Esto es, que para n grande,

$$E\|u_{n, \gamma_{VCG}} - u^+\|_U^2 \rightarrow 0 \text{ cuando } \delta \rightarrow 0^+;$$

ni menos aún se conoce su razón de convergencia con respecto a $\delta \rightarrow 0^+$.

Nuestra alternativa, Principio de Seudo-Optimalidad, el cual estudiamos en lo que sigue del presente capítulo, tiene la propiedad de definir un algoritmo de regularización a la Tijonov con orden de convergencia tipo Holder con respecto a δ (i.e. $O(\delta^q)$, $q > 0$). Y aún cuando este estudio es hecho en sentido a priori, da las bases para llevarlo a la práctica, como se hará ver en el próximo capítulo, sin requerir del conocimiento explícito de δ^2 y u^+ . En un lenguaje estadístico, respecto al modelo lineal (formal) $Au + \xi = f$, ξ un vector aleatorio de la

familia $\{\xi\}_\delta = \{\xi \in \mathbb{F} | \xi \text{ aleatorio, } \mathbb{E}[\xi] = 0 \text{ y } \text{vt}[\xi] < \delta^2\}^{(*)}$,
 y bajo el supuesto (a priori) que $u^+ = (A'A)^{-1} h$ (h en U , $0 < \nu < 1$),
 la idea, muy sencilla por cierto, consiste en tomar $\gamma = \gamma_{80}(\delta)$
 de manera que

$$v(\gamma; \delta) = \tau \beta^2(\gamma; \|h\|_{\mathbb{F}, \nu}), \quad \gamma > 0,$$

($\tau > 0$ dada)

donde $v(\gamma; \delta)$ y $\beta^2(\gamma; \|h\|_{\mathbb{F}, \nu})$ son mayoraciones inmejorables de
 la variabilidad total y sesgo cuadrado de $\hat{u}^+ = R_\gamma f$, el estima
 dor Gauss-Markov regularizado de u ; estableciendo que

$$\|\hat{u}_\gamma^+ - u\|_{\mathbb{E}} = O(\delta^{2\nu/(2\nu+1)}), \quad \gamma = \gamma_{80}(\delta),$$

donde $\|\hat{u}_\gamma^+ - u\|_{\mathbb{E}}^2 = \sup\{\mathbb{E}\|\hat{u}_\gamma^+ - u\|_{\mathbb{U}}^2 | \xi \in \{\xi\}_\delta\}$.

siendo $\hat{u}_\gamma^+ = R_\gamma f$ para $Au + \xi = f$.

§3.2. Principio de Pseudo-Optimalidad: I. Versión Métrica.

Para el propósito central de esta sección, se hace necesari
 o recordar que^(**)

$$(3.2.1) \quad \begin{aligned} \|R_\gamma f_\delta - u^+\|_{\mathbb{U}} &\leq \|R_\gamma(f_\delta - f)\|_{\mathbb{U}} + \|R_\gamma f - u^+\|_{\mathbb{U}} \\ &\leq \|R_\gamma\| \delta + \|R_\gamma f - u^+\|_{\mathbb{U}}. \end{aligned}$$

(*) $\text{vt}[\xi] = \mathbb{E}\|\xi - \mathbb{E}[\xi]\|_2^2$; $\text{vt} \equiv$ varianza total.

(**) Secs. §2.2 y §2.3

Desigualdad que es válida para toda familia regularizante $\{R_Y\}_{Y>0}$ con $R_Y: F \rightarrow U$ continuos. La idea del *Principio de Pseudo-optimalidad*, con base en esta fundamental desigualdad es la siguiente: "tomar $Y = Y_{\text{so}}(\delta)$ como la solución de la ecuación

$$\|R_Y\| \delta = \tau \|R_Y f - u^+\|_U \quad (\tau > 0 \text{ dada}).$$

Pero debido a que el cálculo de $\|R_Y\|$ y $\|R_Y f - u^+\|$ es en general muy difícil de realizar, es natural buscar "buenas" mayoraciones $\mu(Y)$ y $\alpha(Y; u^+)$ de $\|R_Y\|$ y $\|R_Y f - u^+\|_U$, respectivamente.

Para la familia regularizante de Tijonov $\{R_Y\}_{Y>0}$ con $R_Y = (A^*A + YI)^{-1} A^*$ del problema $Au = f$ (A lineal compacto) se tiene el siguiente

Lema 3.2.1. Para $Y > 0$, se tiene que

$$(3.2.1) \quad \|R_Y\| < \mu(Y) = \frac{1}{2\sqrt{Y}}.$$

Demostración: Del desarrollo singular (2.3.1) de A , se obtiene que

$$R_Y = \sum_{i=1}^{\infty} \frac{s_i \langle \cdot, w_i \rangle F v_i}{s_i^2 + Y}$$

y con ello que

$$\|R_Y\| = \max \left\{ \frac{s_1}{s_1^2 + Y} \right\}.$$

Ahora, como la función $\varphi(s) = \frac{s}{s^2 + \gamma}$, $s > 0$, alcanza su máximo $g(s^*) = \frac{1}{2\sqrt{\gamma}}$ en $s^* = \sqrt{\gamma}$, el lema queda demostrado. #

Nota 3.2.1. La cota (3.2.2) es inmejorable. #

Con respecto al término $\|R_\gamma f - u^+\|_U$, el asunto es más complicado; pues como ya hemos mencionado (§2.3), con la sola condición de la existencia de u^+ , es sólo posible demostrar que $R_\gamma f \rightarrow u^+$ (en U). Para lograr una estimación del orden de convergencia de $\|R_\gamma f - u^+\|_U$ con respecto a γ cuando $\gamma \rightarrow 0^+$, se hace necesario exigir a u^+ propiedades de "suavidad" más fuertes que las requeridas por el teorema de Picard. (+)

Teorema 3.2.1 Si $Pf = A(A'A)^v h^{(*)}$ para alguna $h \in U$, $0 < v \leq 1$, entonces

$$(3.2.3) \quad \|R_\gamma f - u^+\|_U \leq \alpha(\gamma; \|h\|_U, v) \equiv c(v) \|h\|_U \gamma^v$$

donde $c(v) = (1-v)^{1-v} v^v$.

Demostración: Como

$$R_\gamma f = R_\gamma Pf = R_\gamma A(A'A)^v h = (A'A + \gamma I)^{-1} (A'A)^{v+1} h,$$

se tiene que

$$R_\gamma f - u^+ = (A'A)^v \{ (A'A + \gamma I)^{-1} A'A - I \} h,$$

de donde

(+). Morozov es el primero en dar resultados en esta dirección [62].

(*) O equivalentemente que $u^+ \in \text{Im}((A'A)^v)$.

$$(3.2.3') \quad \|R_Y f - u^+\|_U \leq \| (A'A)^V \{ (A'A + \gamma I)^{-1} A'A - I \} \| \|h\|_U$$

Por otro lado se tiene que

$$(3.2.3'') \quad \| (A'A)^V \{ (A'A + \gamma I)^{-1} A'A - I \} \| = \max_x \left\{ \frac{s_1^{2\nu} \gamma}{s_1^2 + \gamma} \right\} \\ \leq c(\nu) \gamma^\nu,$$

pues $\varphi(s) = s^{2\nu} \gamma / (s^2 + \gamma)$ alcanza su máximo en $s^0 = (\nu \gamma / (1 - \nu))^{1/2}$, dando $\varphi(s^0) = c(\nu) \gamma^\nu$.

De (3.2.3') y (3.2.3'') se obtiene (3.2.3). #

Nota 3.2.2 La mayoración (3.2.3) de $\|R_Y f - u^+\|_U$ es inmejorable. #

Es importante comentar el significado de la condición $Pf = A(A'A)^V h$, la cual se puede describir en forma equivalente: como

$$(3.2.4) \quad \sum_{i=1}^m \frac{|\langle f, w_i \rangle_F|^2}{s_i^{2+4\nu}} < \infty.$$

La cual dice que f debe ser muy "suave", en el sentido de que f se halla prácticamente concentrada en las direcciones w_1 (en F) de bajas "frecuencias" $1/s_1$. Dicha situación viene a ser razonable para efectos de la ciencia aplicada y la tecnología (Lee et. al. [56] p. 66).

A continuación se describen las gráficas de las mayoraciones $\alpha(\gamma; \|h\|_U, \nu)$ y $\mu(\gamma; \delta) = \mu(\gamma) \cdot \delta$ con respecto a γ .

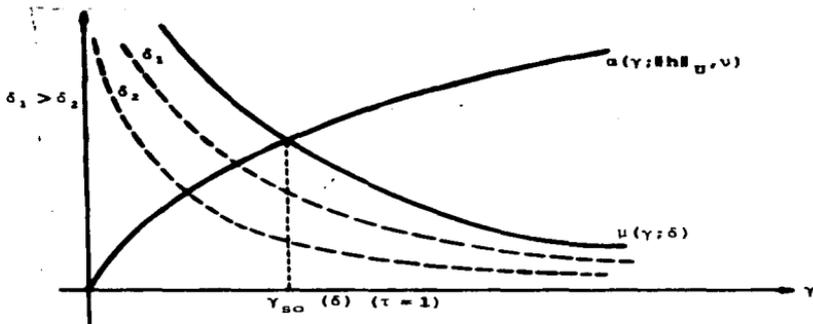


Figura 2. Gráficas típicas de las funciones mayorantes $\alpha(\gamma; \|h\|_{U, V})$ y $\mu(\gamma; \delta)$.

Principio de Seudo-optimalidad. Para $\tau > 0$ dada, tómesese $\gamma = \gamma_{SO}(\delta)$ como la solución de la ecuación

$$(3.2.5) \quad \mu(\gamma)\delta = \tau \alpha(\gamma; \|h\|_{U, V}), \quad \gamma > 0. \quad \#$$

La figura anterior muestra la existencia única de $\gamma = \gamma_{SO}(\delta)$.

En la idea inicial del Principio de Seudo-optimalidad, $\tau=1$. La conveniencia de este nuevo enunciado lo explica el hecho de que en el orden de convergencia de $\|R_{\gamma} \xi_{\delta} - u^+\|_U(\gamma = \gamma_{SO}(\delta))$ con respecto a $\delta \rightarrow 0^+$ el valor de $\tau=1$ no es esencial, como puede verse del resultado siguiente. Por lo demás, tal flexibilidad puede tener sus ventajas prácticas.

Teorema 3.2.2. Sea $A: U \rightarrow F$ lineal y compacto. Si $Pf = A(A'A)^v h$ ($0 < v < 1$) para alguna $h \in U$ y se elige $\gamma = \gamma_{20}(\delta)$ entonces para la familia regularizante $\{R_\gamma\}_{\gamma > 0}$ con $R_\gamma = (A'A + \gamma I)^{-1}A'$ del problema $Au = f$ se tiene que

$$(3.2.6) \quad \|R_{\gamma_{20}} f_\delta - u^+\|_U \leq k(\gamma, v, \|h\|_U) \delta^{2v/(2v+1)}$$

donde

$$k(\gamma, v, \|h\|_U) = \frac{\tau+1}{2\tau} (2\tau c(v) \|h\|_U)^{1/(2v+1)}$$

Demostración: De la ecuación (3.2.5) se tiene que $\gamma_{20}(\delta) = (2\tau c(v) \|h\|_U)^{-2/(2v+1)} \delta^{2/(2v+1)}$,

de donde

$$\begin{aligned} \|R_{\gamma_{20}} f_\delta - u^+\|_U &\leq \mu(\gamma_{20}) \cdot \delta + \alpha(\gamma_{20}; \|h\|_U, v) \\ &= (\tau+1)\alpha(\gamma_{20}; \|h\|_U, v) \\ &= k(\tau, v, \|h\|_U) \delta^{2v/(2v+1)}. \quad \square \end{aligned}$$

Para $v=1$, se tiene que

$$(3.2.6^*) \quad \|R_{\gamma_{20}} f_\delta - u^+\|_U \leq (2\tau \|h\|_U)^{1/3} \cdot \delta^{2/3}.$$

Ahora, bajo la hipótesis $Pf = A(A'A)^v h$ ($0 < v < 1$) para alguna $h \in U$, se puede ver que la función mayorante $\psi(\gamma; \delta, v, \|h\|_U) \equiv \mu(\gamma) \cdot \delta + \alpha(\gamma; \|h\|_U, v)$, para $\gamma > 0$, de $\|R_\gamma f_\delta - u^+\|_U$, alcanza su valor mínimo

$$\begin{aligned} \psi_{\min} &= B\delta^{2v/(5-2v)} + C\delta^{(4-2v)/(5-2v)} \\ &\leq (B+C) \delta^{2v/(5-2v)}, \quad 0 < v < 1, \end{aligned}$$

$$(B = ((4v)^{-2v} c(v) \text{sh} \delta_U)^{1/(5-2v)}, C = \frac{1}{2} (4vc(v) \text{sh} \delta_U)^{1/(5-2v)},$$

en

$$\gamma_{op}(\delta) = (4vc(v) \text{sh} \delta_U)^{-2/(5-2v)} \delta^{2/(5-2v)}.$$

En particular para $v=1$, se tiene que

$$(3.2.6'') \quad \|\mathbb{R}_\gamma \mathbb{Z}_\delta - u^+\|_U \leq (2^{-1/3} + 2^{-4/3}) \text{sh} \delta_U^{1/3} - \delta^{2/3}$$

que corresponde a (3.2.6') con $\tau=27/32$.

En consecuencia, para $v=1$, la convergencia asintótica de $\|\mathbb{R}_\gamma \mathbb{Z}_\delta - u^+\|_U \rightarrow 0$ con respecto a $\delta \rightarrow 0^+$, para γ -pseudo- δ ptima y γ - δ ptima son esencialmente la misma.

§3.3. Regularización de la Inversa Generalizada (Moore-Penrose) de una Matriz.

Como vimos en §1.2 (ejemplo 6), el cálculo de la INVERSA GENERALIZADA Moore-Penrose (M-P) A^+ de una matriz A , es un problema M-Pdo. El principal propósito de esta sección es el cálculo "estable" de A^+ con respecto a perturbaciones de A , vía el método de regularización de Tijonov con la selección del parámetro por el Principio de Pseudo-optimalidad.

Sea $M_{m \times n}$ el espacio vectorial de todas las matrices reales A de dimensiones $m \times n$ ($m \geq n$). Como es bien sabido, $M_{m \times n}$ con el producto interior

$$\langle A, B \rangle = \text{traza } A'B \quad (= \text{tr}(A'B)),$$

es un espacio euclidiano, cuya norma inducida es la norma de Frobenius: $\|A\| = (\langle A, A \rangle)^{1/2}$.

Nuestro problema consiste en hallar una solución aproximada A_{sp}^+ que dependa continuamente de A , a partir de A_0 con $\|A_0 - A\| \leq \delta \|A\|$, para el sistema matricial

$$(3.3.1) \quad \begin{aligned} AX &= P \\ XA &= Q, \end{aligned}$$

en donde $P: \mathbb{R}^m \rightarrow \text{Im}(A)$ y $Q: \mathbb{R}^n \rightarrow \text{N}(A)^\perp$ son las proyecciones ortogonales sobre $\text{Im}(A)$ y $\text{N}(A)^\perp$ respectivamente, sistema matricial que determina a A^+ completamente.

Para tal objetivo, considérese a la *funcional regularizante* siguiente para el problema (3.3.1)

$$(3.3.2) \quad \Phi_\gamma[X; A_0] = \|A_0 X - P\|^2 + \gamma \|X\|^2, \quad \gamma > 0,$$

siendo A_0 tal que $\|A_0 - A\| \leq \delta \|A\|$, para alguna $\delta > 0$ dada.

Por ser $\Phi_\gamma[\cdot; A_0]$ continua no-negativa y M_{\max} un espacio vectorial normado localmente compacto, existe $R_\gamma(A_0)$ en M_{\max} tal que

$$\Phi_\gamma[R_\gamma(A_0); A_0] \leq \Phi_\gamma[X; A_0], \quad \text{para toda } X \text{ en } M_{\max}.$$

Más adn, de la identidad

$$\begin{aligned} \|A_0 X - A_0 Y\|^2 + \gamma \|X - Y\|^2 &= \frac{1}{2} \Phi_\gamma[X; A_0] \\ &+ \frac{1}{2} \Phi_\gamma[Y; A_0] - \Phi_\gamma\left[\frac{1}{2}(X+Y); A_0\right], \end{aligned}$$

se sigue que $R_Y(A_g)$ es única.

Para el cálculo de $R_Y(A_g)$, considérese la primera variación de (3.3.2)

$$\delta \Phi_Y[X; A_g](H) = 2 \langle A_g^* (A_g X - I) + \gamma X, H \rangle.$$

Ahora, $\delta \Phi_Y[X; A_g](H) \equiv 0$ para toda H en $M_{n \times n}$ si y sólo si X es la solución del sistema

$$(\gamma I + A_g^* A_g) X = A_g^*.$$

Por ser $\gamma I + A_g^* A_g$ positiva definida para $\gamma > 0$ se tiene que

$$(3.3.3) \quad R_Y(A_g) = (\gamma I + A_g^* A_g)^{-1} A_g^*.$$

Es directo ver que $R_Y: M_{n \times n} \rightarrow M_{n \times n}$ dado por (3.3.3) es continuo para toda $\gamma > 0$ dada. Por lo tanto, $\{R_Y\}_{\gamma > 0}$ con R_Y dado por (3.3.3) es una familia candidata para ser regularizante del problema (3.3.1).

Lema 3.3.1. Se tiene que

$$(3.3.4) \quad \|R_Y(A) - A^+\| \leq a(\gamma), \quad \gamma > 0,$$

donde $a(\gamma)$ viene dada por

$$(3.3.4') \quad a(\gamma) = \|(\gamma I + A^* A)^{-1} A^* A^+\|;$$

más aún, $a(\gamma)$ tiene las propiedades siguientes

(i) $a(\gamma)$ es monótona creciente,

(ii) $a(\gamma) \rightarrow 0$ cuando $\gamma \rightarrow 0^+$ (i.e. $A_\gamma^+ \rightarrow A^+$ cuando $\gamma \rightarrow 0^+$), y

(iii) $a(\gamma) \rightarrow \infty$ cuando $\gamma \rightarrow -$.

Demostración: De la descomposición singular de A :

$$A = U \Sigma V', \quad \Sigma = \left[\begin{array}{ccc|ccc} s_1 & & & & & \\ & s_2 & & & & \\ & & \ddots & & & \\ & & & s_r & & \\ \hline & & & & & \\ & & & & & \\ & & & & & \end{array} \right]; s_1 > s_2 > \dots > s_r > 0;$$

se sigue que

$$A^+ = V \Sigma^+ U', \quad \Sigma^+ = \left[\begin{array}{ccc|ccc} s_1^{-1} & & & & & \\ & s_2^{-1} & & & & \\ & & \ddots & & & \\ & & & s_r^{-1} & & \\ \hline & & & & & \\ & & & & & \\ & & & & & \end{array} \right];$$

y que

$$A_\gamma^+ = V \Sigma_\gamma^+ U', \quad \Sigma_\gamma^+ = (\gamma I + \Sigma^+ \Sigma)^{-1} \Sigma^+;$$

$$\Sigma_\gamma^+ = \left[\begin{array}{ccc|ccc} \frac{s_1}{\gamma + s_1^2} & & & & & \\ & \frac{s_2}{\gamma + s_2^2} & & & & \\ & & \ddots & & & \\ & & & \frac{s_r}{\gamma + s_r^2} & & \\ \hline & & & & & \\ & & & & & \\ & & & & & \end{array} \right]$$

Luego,

$$\begin{aligned}
 \|A_Y^* - A^*\|^2 &= I - V(\Sigma^* - \Sigma_Y^*) U^*{}^2 \\
 &= \frac{\Sigma}{\Sigma_1} \left(\frac{1}{\Sigma_1} - \frac{\sigma}{\gamma + \sigma^2} \right)^2 = \frac{\Sigma}{\Sigma_1} \left(\frac{\gamma}{\sigma_1(\gamma + \sigma_1^2)} \right)^2 \\
 &< \gamma^2 \frac{1}{\sigma_1^2} \frac{\Sigma}{\Sigma_1} \left(\frac{1}{\gamma + \sigma^2} \right)^2 = \gamma^2 \|A^*\|_1^2 \|(\gamma I + A^*A)^{-1}\|^2 \\
 &< \|A^*\|^2 \|(\gamma I + A^*A)^{-1}\|^2 \gamma^2 = (a(\gamma))^2.
 \end{aligned}$$

Las propiedades (i), (ii) y (iii) se siguen de manera directa de (3.3.4'). \square

Lema 3.3.2. Para toda A_δ dada con $\|A_\delta - A\| \leq \delta \|A\|$, se cumple que

$$(3.3.5) \quad \|R_Y(A_\delta) - R_Y(A)\| \leq V(\gamma; \delta),$$

donde $V(\gamma; \delta)$ viene dada por

$$(3.3.5') \quad V(\gamma; \delta) = \|(\gamma I + A^*A)^{-1}\| \|A\| \left(1 + \frac{b(\delta)}{\gamma}\|A\|^2\right) \delta,$$

con $b(\delta) = \sqrt{n}(1+\delta)(2+\delta)$. Además, $V(\gamma; \delta)$ tiene las propiedades siguientes:

- (i) $V(\gamma; \delta)$ respecto a γ es monótona decreciente para $\gamma > 0$.
- (ii) $V(\gamma; \delta) \rightarrow 0$ cuando $\gamma \rightarrow \infty$, y
- (iii) $V(\gamma; \delta) \rightarrow \infty$ cuando $\gamma \rightarrow 0^+$.

Demostración: Se tiene que

$$\begin{aligned} \|R_Y(A) - R_Y(A_\delta)\| &\leq \|(\gamma I + A'A)^{-1}\| \|A - A_\delta\| \\ &\quad + \|(\gamma I + A'A)^{-1} - (\gamma I + A_\delta'A_\delta)^{-1}\| \|A_\delta\|; \\ \|(\gamma I + A'A)^{-1} - (\gamma I + A_\delta'A_\delta)^{-1}\| &\leq \|(\gamma I + A'A)^{-1}\| \|(\gamma I + A_\delta'A_\delta)^{-1}\| \|A_\delta'A_\delta - A'A\|; \text{ y} \\ \|A_\delta'A_\delta - A'A\| &\leq (\|A\| + \|A_\delta\|) \|A_\delta - A\|. \end{aligned}$$

Por otro lado, de la descomposición singular de A_δ , se obtiene que

$$\|(\gamma I + A_\delta'A_\delta)^{-1}\| \leq \frac{\sqrt{n}}{\gamma}.$$

El resultado se sigue directamente de estas desigualdades. \square

De estos dos lemas se tiene que $R_Y(A_\delta) \rightarrow A^+$, si γ y $\delta \rightarrow 0^+$ de manera que $\delta/\gamma \rightarrow 0^+$. Por lo tanto, $\{R_Y\}_{\gamma>0}$ con R_Y dado por (3.3.3) es una familia regularizante para el problema (3.3.1). Esto es, para el cálculo aproximado y estable respecto a A de A^+ .

A continuación se ilustran las gráficas de las funciones mayorantes $a(\gamma)$ y $V(\gamma;\delta)$ de $\|R_Y(A) - A^+\|$ y $\|R_Y(A_\delta) - R_Y(A)\|$ respectivamente.

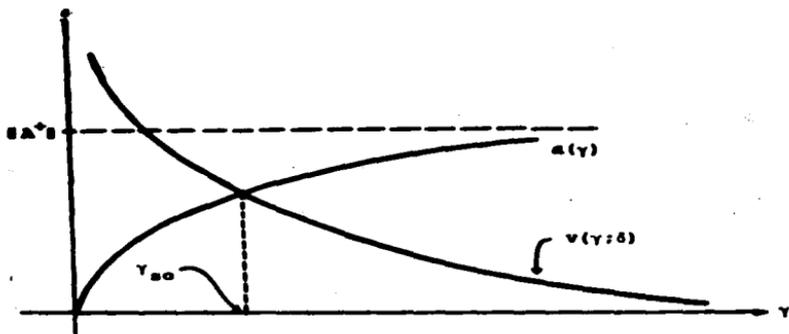


Figura 3. Gráficas típicas de las funciones mayores $a(y)$ y $V(y;\delta)$.

Ahora, la elección del parámetro de regularización por el Principio de Seudo-optimalidad se obtiene resolviendo la ecuación

$$V(y;\delta) = \tau a(y), \quad \tau > 0 \text{ dado,}$$

que por (3.3.4') y (3.3.5') da lugar a la relación

$$EAE \left(1 + \frac{b(\delta)}{Y} EAE^2\right) \delta = \tau EA^+ EY,$$

de la cual se obtiene la ecuación cuadrática

$$\tau EA^+ E Y^2 - EAE \delta Y - b(\delta) EAE^2 \delta = 0,$$

cuya raíz positiva viene dada por

$$(3.3.6) \quad \gamma_{so}(\delta) = \delta \text{AN} (\delta + \sqrt{4\tau b(\delta)K(A)\delta + \delta^2}) / (2\tau \text{A}^+ \text{I})$$

donde $K(A) = \text{AN} \text{A}^+ \text{I}$, es el condicional numérico de A.

Luego entonces

$$\begin{aligned} \|\mathbb{R}_{\gamma_{so}}(A_\delta) - A^+\| &\leq V(\gamma_{so}, \delta) + a(\gamma_{so}) \\ &= (\tau + 1)a(\gamma_{so}) = (\tau + 1)\text{A}^+ \text{I} (\gamma_{so} \text{I} + \text{A}'\text{A})^{-1} \text{I} \gamma_{so}(\delta) \\ &\leq (\tau + 1)\text{A}^+ \text{I}^2 \gamma_{so}(\delta), \end{aligned}$$

ya que

$$(3.3.6') \quad \text{I} (\gamma_{so} \text{I} + \text{A}'\text{A})^{-1} \text{I} \leq \text{I} \text{A}^+ \text{I}^2.$$

Así, echando mano de (3.3.6), se tiene el siguiente

Teorema 3.3.1 Si se aplica a la familia regularizante $\{\mathbb{R}_\gamma\}_{\gamma > 0}$ con \mathbb{R}_γ dado por (3.3.3) el Principio de Seudo-optimalidad para la selección del parámetro $\gamma = \gamma_{so}(\delta)$, entonces

$$(3.3.7) \quad \frac{\|\mathbb{R}_{\gamma_{so}}(A_\delta) - A^+\|}{\text{I} \text{A}^+ \text{I}} \leq \frac{\tau + 1}{2\tau} K(A) (\sqrt{4\tau b(\delta)K(A)\delta + \delta^2} + \delta),$$

para toda A_δ con $\|A_\delta - \text{AN}\| \leq \delta \text{AN}$.

Con este resultado queda demostrado que la familia $\{\mathbb{R}_\gamma\}_{\gamma > 0}$ con \mathbb{R}_γ dado por (3.3.3) con el Principio de Seudo-optimalidad define un algoritmo de regularización para el problema (3.3.1); esto es, para el cálculo aproximado de A^+ "estable" a perturbaciones (relativas) en A.

La importancia de este resultado radica en que no se requiere de la muy restrictiva y poco realista hipótesis de sólo aceptar perturbaciones "pequeñas" δA de A con $\text{rgo}(A + \delta A) = \text{rgo}(A)$ (Medin [10], y Stewart [5]).

Para δ 's pequeñas, (3.3.7) se puede describir, en forma simplificada, como sigue

$$(3.3.8) \quad \frac{\|R_{Y_{SO}}(A_\delta) - A^+\|}{\|A^+\|} < \frac{1+\epsilon}{2\epsilon} \sqrt{8\tau[K(A)]} \cdot \delta^{1/2} + O(\delta),$$

o bien como

$$\frac{\|R_{Y_{SO}}(A_\delta) - A^+\|}{\|A^+\|} \leq O(\delta^{1/2}), \quad (\|A_\delta - A\| < \delta N \epsilon).$$

Técnicamente, esto es muy satisfactorio, pues dice que el Principio de Seudo-optimalidad para el caso bajo discusión, define un algoritmo de regularización *Bien-Comportado* a la F. John ([50], [51]). Sin embargo, debe notarse que la constante $[K(A)]^{3/2}$ de $O(\delta^{1/2})$ es "muy grande" para A *mal-comportada* (i.e. cuando $K(A)\# \approx 1$, $\#$ la precisión de la computadora anfitrión). Así pues, nuestro resultado es un primer importante logro en la dirección del tan difícil problema de regularizar el mapeo inversión generalizado (M-P): $A \xrightarrow{\psi^+} A^+$; bajo hipótesis realistas y aceptables.

Conviene hacer notar que el coeficiente $[K(A)]^{3/2}$ de $O(\delta^{1/2})$ viene a consecuencia de la burda sobrestimación (3.3.6') para $\|Y_{SO} I + A^* A\|^{-1} N$. Luego es de esperar, bajo hipótesis razonables

adicionales, el logro de un mejor orden de convergencia para $\|R_{Y_{so}}(A_\delta) - A^+ \delta / \delta A^+\|$ con respecto a $\delta \rightarrow 0^+$. El problema en esta dirección queda aún abierto.

Por último, es importante anotar que el Principio de Quasi-óptimalidad de Glasko-Tijonov ([34], [35] y [32]), no es aplicable en el presente caso. En efecto, para $B(A_\delta)$, de $B_Y^+ (=R_Y(B))$: $(\gamma I + B^*B)^{-1}B^*$, se sigue que

$$\frac{d}{d\gamma} B_Y^+ = -\gamma(\gamma I + B^*B)^{-2}B^*,$$

y con ello que

$$(3.3.9) \quad \psi_{q-op}(\gamma) = \| \gamma \frac{dB_Y^+}{d\gamma} \| = \gamma^2 \| (\gamma I + B^*B)^{-2}B^* \|^2.$$

Ahora, de la descomposición singular de B:

$$B = U \Sigma_B V', \quad \Sigma_B = \begin{bmatrix} \tilde{\sigma}_1 & & & \\ & \ddots & & \\ & & \tilde{\sigma}_p & \\ & & & \text{---+---} \end{bmatrix}, \quad \tilde{\sigma}_1 > \tilde{\sigma}_2 > \dots > \tilde{\sigma}_p > 0,$$

$$U'U = I_m \quad \text{y} \quad V'V = I_n,$$

se tiene que

$$(\gamma I + B^*B)^{-2}B^* = V((\gamma I + \Sigma_B^* \Sigma_B)^{-2} \Sigma_B^*) U',$$

por lo que

$$\psi_{q-op}(\gamma) = \left(\sum_{i=1}^p \frac{\gamma^2 \tilde{\sigma}_i^2}{(\gamma + \tilde{\sigma}_i^2)^2} \right)^{1/2}.$$

De donde se sigue que $\phi_{q-\text{op}}^*(\gamma) > 0$ para toda $\gamma > 0$. Y por lo tanto, $\phi_{q-\text{op}}^*(\gamma)$ no tiene un solo mínimo local para $\gamma > 0$, o en otras palabras, $\gamma = \gamma_{q-\text{op}}(\delta)$ no existe.

Con este ejemplo se muestra que el Principio de Pseudo-optimalidad es una alternativa a los otros prestigiosos Principios para la construcción del criterio $\gamma = \gamma(\delta)$, digna de tomarse en cuenta.

§3.4. Principio de Pseudo-Optimalidad. II: Versión Estadística.

En esta sección se plantea el Principio de Pseudo-Optimalidad (P.S.O.) en términos estadísticos y se establece que éste define, para $A: U \rightarrow F$ lineal compacto sobre espacios de Hilbert U y F , un algoritmo de regularización estadístico a la Tijonov para el problema

$$(3.4.1) \quad Au = g; \quad g = f + \xi; \quad Au^{\dagger} = f;$$

donde ξ es un vector aleatorio^(*) en F (no observable) que satisface:

$$(3.4.1') \quad \text{a) } E(\xi) = 0$$

$$\text{b) } \text{v.t. } (\xi) = \sum_{i=1}^{\infty} \sigma_i^2 < \delta^2, \text{ con } \sigma_i^2 = E(\xi_i^2), \xi_i = \langle \xi, w_i \rangle_F, \text{ siendo } \{w_i\}_{i=1}^{\infty} \text{ cualquier conjunto ortonormal a lo más numerable en } F. \text{ (**)}$$

(*) En el apéndice de este capítulo se ve el sentido de vector aleatorio ξ sobre un espacio de Hilbert F , útil a nuestros propósitos.

(**) Aquí, $\{w_i\}_{i=1}^{\infty}$ puede ser de los vectores singulares izquierdos de A .

A todos los vectores aleatorios ξ que cumplen con (3.4.1') los denotaremos por $\{\xi\}_g$.

El problema es "estimar" u^+ de manera estable a perturbaciones en el lado derecho de (3.4.1).

Si $g \in D(A^+)$ entonces el "estimador Gauss-Markov" es

$$\hat{u}^+ = A^+ g;$$

pues si $g \in D(A^+)$ entonces \hat{u}^+ es la solución de menor norma del problema:

$$(3.4.2) \quad \min_{u \in U} \|\xi\|_Y^2$$

Es muy importante observar que (3.4.2), a diferencia del caso de dimensión finita, no siempre tiene solución. Más aún, $\hat{u}^+ = A^+ g$ por ser A^+ no acotado, no depende continuamente de g . Mas grave aún, para un pequeño cambio en g , \hat{u}^+ puede dejar de existir.

Consideremos, ahora, el "estimador Gauss-Markov regularizado":

$$(3.4.3) \quad \begin{aligned} \hat{u}_Y^+ &= R_Y(g; \xi) (=R_Y(g)) \\ &= (A'A + \gamma I)^{-1} A'g, \end{aligned}$$

que resulta ser la solución del problema

$$(3.4.3') \quad \min_{u \in U} J_Y[u; g, \xi]$$

donde

$$J_{\gamma}(u; \eta, \xi) = \|\xi\|_{\mathcal{D}}^2 + \gamma \|u\|_{\mathcal{U}}^2$$

como ya se ha visto antes.

Claramente el estimador GM regularizado siempre existe y es estable a perturbaciones en η .

Para presentar la formulación "estadística" del P.S.O., se hace necesario introducir un poco de terminología.

Como se sabe, el *error cuadrático medio* de \hat{u}_{γ}^{+}

$$(3.4.4) \quad \text{ecm}(\gamma; u, \delta) \equiv \mathbb{E}(\|\hat{u}_{\gamma}^{+} - u^{+}\|_{\mathcal{U}}^2),$$

para $\xi \in \{\xi\}_{\delta}$, se puede escribir como la suma de las funciones de *varianza total*

$$(3.4.4') \quad \text{vt}(\gamma; \delta) = \mathbb{E}(\|u_{\gamma}^{+} - \mathbb{E}(\hat{u}_{\gamma}^{+})\|_{\mathcal{U}}^2),$$

y de *sesgo cuadrado*

$$(3.4.4'') \quad \text{b}^2(\gamma; u^{+}) = \|\mathbb{E}(\hat{u}_{\gamma}^{+}) - u^{+}\|_{\mathcal{U}}^2.$$

Lema 3.4.1. Para $\xi \in \{\xi\}_{\delta}$, en términos de la descomposición singular de $A = \sum_{i=1}^{\infty} d_i \langle \cdot, v_i \rangle_{\mathcal{U}} w_i$, se tiene que

$$(3.4.5) \quad \text{vt}(\gamma; \delta) = \sum_{i=1}^{\infty} \frac{d_i^2 \sigma_i^2}{(d_i^2 + \gamma)^2},$$

y que

$$(3.4.5') \quad b^2(\gamma; u^+) = \frac{\sigma^2}{\gamma} \frac{\sum_{i=1}^k \sigma_i^2 \gamma^2}{(d_1^2 + \gamma)^2}$$

en donde $\sigma_i = \langle u^+, v_i \rangle_U$, $i=1, 2, \dots$

Demostración: Calculando

$$\begin{aligned} \mathbb{E}(\hat{U}_Y^+) &= \mathbb{E}(R_Y(g)) = \mathbb{E}(R_Y(Au^+ + \xi)) \\ &= \mathbb{E}(R_Y A(u^+)) + R_Y(\mathbb{E}(\xi)) \\ &= (R_Y A)(u^+), \end{aligned}$$

se sigue que

$$\begin{aligned} \hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+) &= R_Y(g) - R_Y A(u^+) \\ &= R_Y(g - A(u^+)) = R_Y(\xi), \end{aligned}$$

de donde

$$\begin{aligned} \mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+)) \mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+)) &= \mathbb{E}(R_Y(\xi) \mathbb{E}(\xi)) \\ &= \frac{\sigma^2}{\gamma} \frac{\sum_{i=1}^k d_i^2 \xi_i^2}{(d_1^2 + \gamma)^2}; \quad \xi_i = \langle \xi, w_i \rangle, i=1, 2, \dots \end{aligned}$$

Y como $\xi \in (\xi)_0$, se obtiene que

$$\begin{aligned} \mathbb{E}(\mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+)) \mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+))) &= \mathbb{E}\left(\frac{\sigma^2}{\gamma} \frac{\sum_{i=1}^k d_i^2 \xi_i^2}{(d_1^2 + \gamma)^2}\right) \\ &= \frac{\sigma^2}{\gamma} \frac{\sum_{i=1}^k d_i^2 \mathbb{E}(\xi_i^2)}{(d_1^2 + \gamma)^2}, \end{aligned}$$

o sea que

$$vt(\gamma; \delta) = \mathbb{E}(\mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+)) \mathbb{E}(\hat{U}_Y^+ - \mathbb{E}(\hat{U}_Y^+))) = \frac{\sigma^2}{\gamma} \frac{\sum_{i=1}^k d_i^2 \sigma_i^2}{(d_1^2 + \gamma)^2},$$

lo que da (3.4.5).

Ahora, veamos que

$$\begin{aligned} \mathbb{E}(\hat{u}_Y^+) - u^+ &= (R_Y A - I) (u^+) \\ &= \sum_1 \left(\frac{d_1^2}{d_1^2 + \gamma} - 1 \right) \langle u, v_1 \rangle_U w_1. \end{aligned}$$

luego

$$b^2(\gamma; u^+) = \mathbb{E}(\hat{u}_Y^+) - u^+ = \sum_1 \frac{\alpha_1^2 \gamma^2}{(d_1^2 + \gamma)^2}.$$

siendo $\alpha_1 = \langle u, v_1 \rangle_U$, $i=1, 2, \dots$ que es (3.4.5'). #

Lema 3.4.2. (Morozov) Sea $vt(\gamma) = \sup_{\xi \in \{\xi\}_\delta} vt(\gamma; \xi)$.

Se tiene que

$$vt(\gamma) \leq v(\gamma)$$

donde

$$(3.4.6) \quad v(\gamma) = \frac{\delta^2}{4\gamma}.$$

Demostración. Sea $g(\lambda) = \lambda/(\lambda + \gamma)^2$, se tiene que $g(\lambda)$ alcanza su máximo en $\lambda^0 = \gamma$, dando $g(\lambda^0) = \frac{1}{4\gamma}$. Luego,

$$vt(\gamma; \xi) \leq \sum_1 \frac{\alpha_i^2}{4\gamma} = \frac{1}{4\gamma} \sum_1 \alpha_i^2 \leq \frac{\delta^2}{4\gamma}.$$

para toda $\xi \in \{\xi\}_\delta$. Por tanto

$$vt(\gamma) \leq \frac{\delta^2}{4\gamma}. \quad \#$$

Nota 3.4.1. La desigualdad (3.4.6) es inmejorable. En efecto, si $\xi = \xi_1 w_1$, $\mathbb{E}(\xi_1^2) = \delta^2$ y $d_1^2 = \gamma$ entonces $vt(\gamma; \xi) = \delta^2/4\gamma$. #

Lema 3.4.3 (Morozov) Si $u^{\dagger} = (A'A)^{\nu} [h]^{(*)}$, para alguna $h \in U$, $0 < \nu < 1$, entonces se tiene que

$$(3.4.7) \quad b^2(\gamma) \leq \hat{b}^2(\gamma)$$

donde

$$\hat{b}^2(\gamma) = (\|h\|_U C(\nu))^2 \gamma^{2\nu}$$

con

$$C(\nu) = \nu^{\nu} (1-\nu)^{1-\nu}$$

Demostración: De que $u^{\dagger} = (A'A)^{\nu} [h]$, se sigue que $\alpha_i^2 = d_i^{4\nu} h_i^2$, $i=1, \dots, m$. Así,

$$b^2(\gamma; \xi) = \sum_1^m \frac{\alpha_i^2 \gamma^2}{(d_i^2 + \gamma)^2} = \sum_1^m \frac{d_i^{4\nu} h_i^2 \gamma^2}{(d_i^2 + \gamma)^2}.$$

Como ya se vio en la demostración del teorema 3.2.1 $g(d) =$

$$\frac{d^{2\nu}}{d^2 + \gamma} \text{ alcanza su máximo en } d^0 = \left(\frac{\nu}{1-\nu}\right)^{1/2} \gamma^{1/2} \text{ dando } g(d^0) = c(\gamma) \gamma^{\nu}.$$

Luego,

$$b^2(\gamma) \leq \sum_1^m (g(d^0))^2 h_i^2 \leq (d(d^0))^2 \sum_1^m h_i^2 = \hat{b}^2(\gamma). \quad \square$$

Nota 3.4.2. La cota (3.4.7) es inmejorable; para verlo, sean $u^{\dagger} = \alpha_i v_i$, $\alpha_i^2 = d_i^{4\nu} h_i^2$, $h = h_i v_i$ y $\gamma = \frac{(1-\nu)d_i^2}{\nu}$. \square

El comportamiento de las funciones $vt(\gamma; \delta)$, $v(\gamma)$, $b^2(\gamma)$ y $\hat{b}^2(\gamma)$ es como se ilustra a continuación.

(*) . Con relación a esta condición, véase el comentario que sigue a la nota 3.2.2.

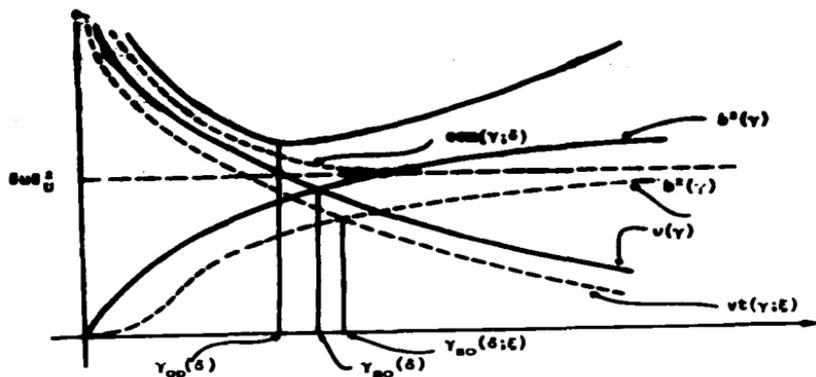


Figura 4. Ilustración geométrica del Principio de Seudo-Optimalidad Estadístico (con $\tau = 1$)

Principio Estadístico de Seudo-Optimalidad. Tómese $\gamma = \gamma_{so}(\delta; \xi, \tau)$ (o bien $\gamma = \gamma_{so}(\delta, \tau)$) como la solución de la ecuación

$$(3.4.8) \quad vt(\gamma; \xi) = \tau^2 b^2(\gamma) \text{ (o bien } v(\gamma) = \tau^2 b^2(\gamma) \text{), } \tau > 0 \text{ dado.}$$

Es inmediato de ver que la ecuación (3.4.8) siempre tiene una única solución. En la fig. 4 antecedente se ilustra el caso cuando la función $ecm(\gamma; \xi)$ no tiene mínimo, por ello, $\gamma_{op}(\delta; \xi)$ en general no existe. Es de notar, sin embargo, que la función mayorante $\psi(\gamma) = v(\gamma) + b^2(\gamma)$ del $ecm(\gamma; \xi)$, para toda $\xi \in \{\xi\}_\delta$ siempre tiene mínimo. Esto es, $\gamma = \gamma_{op}(\delta)$ siempre existe.

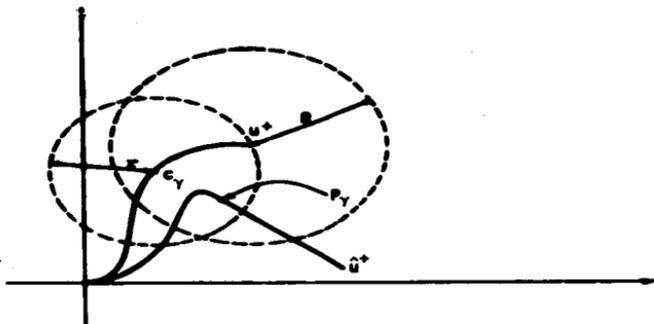


Figura 5. Ilustración geométrica del significado estadístico del Principio de Pseudo-Optimalidad con $r=1$. Aquí, $C_Y = E(\hat{U}_Y^+)$, $P_Y = \hat{U}_Y^+$, $r^2 = vt(y; \epsilon)$ (o bien $v(y)$) y $Y = Y_{SO}(\delta; \epsilon)$ (o bien $Y = Y_{SO}(\delta)$).

La relación (3.4.8) con $r=1$ tiene un significado estadístico muy claro, el cual consiste en determinar \hat{U}_Y^+ con $Y = Y_{SO}(\delta; \epsilon)$ (o bien $Y = Y_{SO}(\delta)$), de manera que la bola con centro en $E(\hat{U}_Y^+)$ y radio cuadrado igual a $vt(y; \epsilon)$ (o bien $v(y)$); contenga a u^+ con "alta probabilidad" (*). La figura de arriba ilustra esta situación.

(*) Para la distribución normal $P\{|x - \mu| \leq c\} \approx .66$.

Es importante anotar que para cualquier $R > 0$ dado y de existir \hat{u}^+ , \hat{u}^+ cae fuera de la bola con centro en u^+ y de radio R con "probabilidad significativa" (*).

El resultado siguiente (el principal de este trabajo) dice que el Principio de Seudo-Optimalidad estadístico define un algoritmo de regularización a la Tijonov, y colateralmente también dice que nuestro problema (3.4.1), bajo discusión, es uno *Sémprecomportado* a la F. John [51].

Teorema 3.4.1 Bajo los supuestos (3.4.1'). Si $u = (A'A)^{-1}h$ ($0 < \gamma \leq 1$) para alguna $h \in U$ entonces la elección del parámetro de regularización $\gamma = \gamma_{so}(\delta, \tau)$ por el Principio de Seudo-Optimalidad estadístico define un algoritmo de regularización con convergencia asintótica.

$$(3.4.9) \quad \sup_{\xi \in (\xi)_{\delta}} \| \hat{u}_{\gamma}^+ - u \|_{\Sigma} \leq n(v) \| h \|_{U, \tau^2} \delta^{2\nu / (2\nu + 1)}$$

en donde

$$\| \hat{u}_{\gamma}^+ - u \|_{\Sigma} \equiv [\mathbb{E} (\| \hat{u}_{\gamma_{so}}^+ - u \|_{\Sigma}^2)]^{1/2}, \text{ y}$$

$$n(v) \| h \|_{U, \tau^2} = \left(\frac{1 + \tau^2}{4\tau^2} \right)^{1/2} (2\tau c(v) \| h \|_{U})^{1/(2\nu + 1)}$$

(*) Pensando en una variable aleatoria con distribución normal, para $R^2 = vt(\gamma; \xi)$, la "probabilidad significativa" es .33 aproximadamente.

Demostración: De la ecuación básica (3.4.8), se obtiene que

$$(3.4.10) \quad Y_{so}(\delta) = [1/(2\tau c(v) \|h\|_U)]^{2/(2\nu+1)} \cdot \delta^{2/(2\nu+1)}$$

De aquí que, para $Y = Y_{so}(\delta)$:

$$\begin{aligned} \| \hat{U}_Y^+ - u^+ \|_E^2 &\leq \varphi(Y) = \nu(Y) + b^2(Y) \\ &= (1 + \tau^2) b^2(Y). \end{aligned}$$

Y por lo tanto, para toda $\xi \in \{\xi\}_\delta$, se tiene que

$$\| \hat{U}_Y^+ - u^+ \|_E \leq \eta(\nu; \|h\|_U, \tau^2) \delta^{2\nu/(2\nu+1)}. \quad \#$$

Nota 3.4.3. Con las mismas hipótesis del teorema 3.4.1, se puede verificar, para toda $\xi \in \{\xi\}_\delta$, que

$$\| \hat{U}_{Y_{op}}^+(\delta) - u^+ \|_E \leq \xi(\nu; \|h\|_U) \delta^{2\nu/(2\nu+1)}$$

en donde

$$(3.4.11) \quad \begin{aligned} \xi(\nu; \|h\|_U) &= (1 + 2\nu)^{1/2} (8\nu)^{-\nu/(2\nu+1)} \\ &\quad \cdot [c(v) \|h\|_U]^{1/(2\nu+1)}, \end{aligned}$$

siendo $Y_{op}(\delta)$ el punto de mínimo de $\varphi(Y) = \nu(Y) + b^2(Y)$ (Véase a Morozov [65], sec. 25). Esto justifica el nombre (de Pseudo-optimalidad) de nuestro principio, en cuanto al orden asintótico de convergencia cuando $\delta \rightarrow 0^+$. Más aún, para el Principio

de Seudo-optimalidad, se tiene que

$$(3.4.12) \quad n(v; \mathbb{H}_U, \tau^2) = (1 + \tau^2)^{1/2} (4\tau^2)^{-v/(2v+1)} \\ \cdot [c(v) \mathbb{H}_U]^{1/(2v+1)}$$

De donde es directo ver que para $\tau^2 = 2v$:

$$f(v; \mathbb{H}_U) = n(v; \mathbb{H}_U, \tau^2).$$

Así, el Principio de Optimalidad es un caso especial del Principio de Seudo-optimalidad. $\#$

APENDICE. Vectores Aleatorios en Espacios de Hilbert.

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad sobre una σ -álgebra \mathcal{A} de subconjuntos de Ω , con P una medida de probabilidad continua σ -aditiva. Y sea $\xi: \Omega \rightarrow F$ una variable aleatoria con valores $\xi = \xi(\omega)$, ξ vector aleatorio, en un espacio de Hilbert (real) F de manera que $\xi_w^{-1}(\xi_w > a) \in \mathcal{A}$, siendo $\xi_w = \langle \xi, w \rangle_F$, para toda $w \in F$ dada.

Por la *distribución* del vector aleatorio ξ en dirección de w ($w \in F$) se entenderá la medida μ_w sobre $(\mathbb{R}, \mathcal{A}_{\mathbb{R}})$ ^(*), definida por

$$\mu_w^{\xi}(C) = P\{\omega \in \Omega \mid \xi_w(\omega) \in C\}$$

(*) $\mathcal{A}_{\mathbb{R}}$ es la menor σ -álgebra de subconjuntos de \mathbb{R} que contiene a todos los subconjuntos borelianos de \mathbb{R} .

para todo $C \in A_{\mathbb{R}}$ dado. Y por la función de distribución del vector aleatorio ξ en dirección de $w \in F$ se entenderá la función

$$G_w^\xi(x) = \mu_w^\xi(C_x), \quad C_x = (-\infty, x].$$

De las propiedades de continuidad de μ_w^ξ , heredadas de la función de probabilidad P , se sigue que $G_w^\xi(x)$ es una función monótona no-decreciente, continua por la derecha y tal que

$$\lim_{x \rightarrow -\infty} G_w^\xi(x) = 0 \quad \text{y} \quad \lim_{x \rightarrow +\infty} G_w^\xi(x) = 1.$$

En estos términos, el k-ésimo momento del vector aleatorio ξ en dirección de w , se define por

$$\mathbb{E}\{(\xi_w)^k\} = \int_{-\infty}^{+\infty} (\xi_w)^k dG_w^\xi,$$

siempre que dicha integral exista.

Por $\mathbb{E}\{\xi\} = 0$ se entenderá que $\mathbb{E}\{\xi_w\} = 0$ para toda $w \in F$ dada. Por la varianza total (v.t.) de un vector aleatorio ξ con $\mathbb{E}\{\xi\} = 0$ con respecto a un sistema ortonormal (a lo más numerable) $\beta = \{e_i\}_{i=1}^{\infty}$ de F se entenderá

$$\text{v.t.}(\xi)_\beta = \sum_{i=1}^{\infty} \mathbb{E}\{(\xi_{e_i})^2\},$$

siempre que $\sum_{i=1}^{\infty} \mathbb{E}\{(\xi_{e_i})^2\}$ exista. Y se dirá que un vector aleatorio ξ (sobre F) tiene varianza total v.t. $\{\xi\}$ a lo más δ^2 , si v.t. $\{\xi\}_\beta \leq \delta^2$ para todo sistema ortonormal (a lo más numerable) β de F .

Es importante observar que las condiciones (3.4.1') se satisfacen para todo vector aleatorio ξ con $\mathbb{E}\{\xi\} = 0$ y v.t. $\{\xi\} \leq \delta^2$.

CAPITULO 4

ESTIMACION GAUSS-MARKOV REGULARIZADA CON ELECCION SEUDO-OPTIMA DEL PARAMETRO DE REGULARIZACION.

El propósito central de este capítulo consiste en mostrar la viabilidad práctica del *Principio de Seudo-optimalidad* para la elección del parámetro de regularización en la resolución numérica de la ecuación integral de Fredholm de primer tipo, viendo su estrecha relación con la estimación (sesgada) Gauss-Markov regularizada (o "ridge") en la regresión lineal.

§4.1. Resolución Numérica de la Ecuación Integral de Fredholm de Primer Tipo, y Colinealidad en la Regresión Lineal.

4.1.1. Colinealidad en la Regresión Lineal.

Uno de los problemas centrales de la *regresión lineal* consiste en hallar una "buena estimación" $\hat{\beta}$ para β en el modelo lineal estandard

$$(4.1.1) \quad \underline{y} = \underline{X} \underline{\beta} + \underline{\epsilon}, \quad \underline{\epsilon} \sim (0, \delta^2 \underline{I}),$$

en donde \mathbf{X} es una matriz de $n \times p$ ($n \geq p$) real no-estocástica, \mathbf{y} es el vector (real) de observaciones, $\underline{\epsilon}$ es un vector aleatorio (de errores) no-observables de media igual a $\underline{0}$ y matriz de covarianzas ($\text{Cov}(\underline{\epsilon}) = \delta^2 \mathbf{I}$) y $\underline{\beta}$ es el vector (real) de parámetros a estimar.

Cuando \mathbf{X} es de rango máximo (i.e. cuando $\text{rgo}(\mathbf{X}) = p$), la estimación más frecuentemente usada, debido principalmente a sus propiedades de ser insesgada y de mínima varianza⁽¹⁾ (entre todas las lineales sobre \mathbf{y}), es la estimación de Gauss-Markov (G M)

$$(4.1.2) \quad \hat{\underline{\beta}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y}, \quad (2)$$

para la cual es bien sabido que

$$(4.1.2') \quad \hat{\underline{\beta}} \sim (\underline{\beta}, \delta^2 (\mathbf{X}' \mathbf{X})^{-1}),$$

y que es mini-max con respecto a cualquier función de pérdida cuadrática

$$(4.1.3) \quad \mathcal{L}_C(\hat{\underline{\beta}}) = (\hat{\underline{\beta}} - \underline{\beta})' \mathbf{C} (\hat{\underline{\beta}} - \underline{\beta}),$$

(1). A estas propiedades hay que agregar las de máxima verocimilitud, consistencia y eficiencia, si $\underline{\epsilon} \sim N(\underline{0}, \delta^2 \mathbf{I})$ (i.e. si adicionalmente $\underline{\epsilon}$ tiene distribución normal).

(2). En lo sucesivo, \mathbf{T}' denotará a la matriz transpuesta de \mathbf{T} , y \mathbf{T}^+ a su matriz inversa generalizada Moore-Penrose (Véase la sec. 2 del Cap. 1).

en donde C es una matriz de $n \times n$ real simétrica y positiva definida. Así se tiene, con respecto a la función de pérdida (4.1.3), que el error cuadrático medio (ecm_c) de $\hat{\beta}$ está dado por

$$ecm_c(\hat{\beta}) = E\{L_c(\hat{\beta})\} = \text{tr}\{C \text{Cov}(\hat{\beta})\} = \delta^2 \text{tr}\{C(K'X)^{-1}\},$$

el que en términos de la descomposición en valores singulares de $X = UAV'$ se puede describir como

$$(4.1.4) \quad ecm_c(\hat{\beta}) = vt_c(\hat{\beta}) = \delta^2 \sum_{i=1}^p \frac{v_i'}{s_i^2} C v_i / s_i^2,$$

en donde v_i es la i -ésima columna de V y s_i es el correspondiente valor singular de X . Luego entonces

$$(4.1.4') \quad \begin{aligned} ecm_c(\hat{\beta}) &= vt_c(\hat{\beta}) \geq \alpha \frac{\delta^2}{s_p^2} \\ &= \alpha \delta^2 \frac{[K_2(X)]^2}{\|X\|_2^2}, \end{aligned}$$

con

$$0 < \alpha = \min\{v' C v \mid \|v\|_2 = 1\}.$$

Así, si X tiene vectores columna numéricamente hablando "débilmente independiente"⁽³⁾ (i.e. si X es cercana a una matriz de rango deficiente), o sea *mal-comportada* en términos del Análisis Numérico, o bien bajo presencia de *colinealidad* en lenguaje de la Estadística,⁽⁴⁾ entonces (como ya se dijo en la sec.1.2

(3). Desde un punto de vista del Análisis Numérico, en Barrera et. al. [8] se discute el término de vectores "débilmente independientes".

(4). Para una revisión sobre la colinealidad en la regresión lineal consultese, entre otros, a Bas [11] y a Stewart [85'].

de este trabajo), el cálculo numérico de $\hat{\beta}$ resulta ser generalmente insatisfactorio, si no es que hasta inservible, lo cual se suele manifestar con la aparición de componentes de $\hat{\beta}$ muy grandes y/o de signos inaceptables.

En lenguaje de la Estadística la explicación se sigue de (4.1.4) y (4.1.4'), en donde se ve claramente el gran efecto que pueden tener los valores singulares "pequeños" en el $\text{scm}_c(\hat{\beta})$ como en la C-varianza total $\text{vt}_c(\hat{\beta})$, sin importar β .

Ahora, con respecto al objetivo central de este trabajo sobre la resolución numérica de la ecuación integral de Fredholm

$$(4.1.5) \quad \int_a^b k(x,t) u(t) dt = f(x), \quad c \leq x < d;$$

la importancia del fenómeno de colinealidad en la regresión lineal radica (como ya se dijo en el aparato 8 de la Sec. 1.2 de este trabajo) en el hecho de que su discretización directa da lugar al modelo lineal (4.1.1):

$$\mathbb{X} \underline{\beta} + \underline{\epsilon} = \underline{y}, \quad \underline{\epsilon} \sim (0, \delta^2 \mathbb{I}),$$

con

$$\mathbb{X} = (k(x_i, t_j) \sigma_j), \quad \underline{y}' = (y_1, \dots, y_n), \quad y_i = f(x_i) + \epsilon_i,$$

bajo presencia de colinealidad, como reflejo o herencia de ser la ecuación (4.1.5) un problema *Mal-Plantado* a la Hadamard, cuando a $f(x)$ se le conoce por medio de una tabla de sus valores (observados)

$$T_{\mathbb{X}} = \{(x_i, y_i) | y_i = f(x_i) + \epsilon_i, i = 1, 2, \dots, n\}$$

y al término integral se le discretiza por medio de una regla de cuadratura del tipo siguiente

$$\int_a^b g(t) dt \approx \sum_{j=1}^n g(t_j) \sigma_j .$$

Regresando al contexto general de la regresión lineal se tiene que debido a las propiedades de la estimación $\hat{\beta}$ de GM para β , si se quiere hallar una estimación $\tilde{\beta}$ para β con $\text{cov}_{\tilde{\beta}}(\tilde{\beta})$ y $\text{vt}_{\tilde{\beta}}(\tilde{\beta})$ menores que los de $\hat{\beta}$, entonces se está obligando a considerar, entre otras, a las estimaciones sesgadas.

En la sección siguiente se discute la estimación GM regularizada, también conocida por la estimación "ridge" o HK de Hoerl y Kennard ([46], [47]), en la literatura de la Estadística.

Para tal efecto es conveniente ver antes, como una motivación a la estimación GM regularizada desde nuestra perspectiva del método de regularización de Tijonov, la formulación discreta de tal método o bien sus modos de discretización.

4.1.2. Convergencia Discreta del Método de Regularización.

Como ya se dijo en la sección 2.2. de este trabajo, el cálculo numérico de la solución regularizada $u_{\gamma}(t)$ para el problema integral

$$(4.1.6) \quad \int_a^b k(x,t) u(t) dt = f_{\delta}(x), \quad c < x < d,$$

con $f_{\delta}(x) = f(x) + \xi(x)$, siendo $\xi(x)$ un vector aleatorio en $L_2(c,d)$ de media cero y de $\text{v.t.}[\xi] < \delta^2$, con respecto a la funcional de Tijonov

$$(4.1.6') \quad \Phi_\gamma\{u; f_0\} = \|f_0\| - \lambda \|u\|_{L_2}^2 + \gamma S\{u\}, \quad \gamma > 0,$$

en donde

$$\lambda\{u\}(x) = \int_a^b k(x,t) u(t) dt$$

$$y \quad S\{u\} = \int_a^b (|u^{(m)}(t)|^2 + \alpha |u(t)|^2) dt^{(*)}, \quad \alpha > 0 \text{ dada.}$$

se puede obtener resolviendo numéricamente la ecuación integro-diferencial de Euler

$$(4.1.7) \quad \int_a^b \tilde{k}(x,t) u(t) dt + \gamma \{(-1)^m u^{(2m)}(x) + \alpha u(x)\} \\ = \int_c^d k(s,x) f_0(s) ds,$$

siendo

$$\tilde{k}(x,t) = \int_c^d k(s,x) k(s,t) ds,$$

con las condiciones de frontera naturales

$$(4.1.7') \quad u^{(\ell)}(a) = u^{(\ell)}(b) = 0, \quad \ell = m, m+1, \dots, 2m-1.$$

En efecto, con referencia a las particiones $a = t_0 < t_1 < \dots < t_p = b$ (uniforme de paso h) para $[a,b]$ y $c = x_1 < x_2 < \dots < x_n$ (con paso máximo k) para $[c,d]$, la aplicación de las reglas de cuadratura

$$(4.1.8) \quad \int_a^b \psi(t) dt = \sum_{j=1}^p \psi(t_j) \sigma_j \quad y \quad \int_c^d \psi(s) ds = \sum_{i=1}^n \psi(x_i) \rho_i,$$

a los términos integrales, y la discretización por diferencias divididas de los términos diferenciales del problema (4.1.7)

(4.1.7'), dan lugar al sistema lineal algebraico

(*) Con $\alpha = 0$, para $m = 0$.

Si para las reglas de cuadratura (4.1.8) las sumatorias $\sum_{j=1}^p |\sigma_j|$ y $\sum_{i=1}^n |\rho_i|$ son uniformemente acotadas para todas h , $h > 0$, entonces se demuestra que

$$\max_j |u_Y(t_j) - u_j| \rightarrow 0 \text{ cuando } h, k \rightarrow 0^+,$$

como puede verse en Linz ([108], [109]).

En particular, si las cuadraturas (4.1.8) son las del trapecio (las que tienen error de truncamiento de orden $O(h^2)$ y $O(k^2)$ respectivamente), y la discretización por diferencias - divididas de los términos diferenciales en (4.1.7) - (4.1.7') tienen errores de truncamiento de orden no menor a $O(h^2)$, lo que es siempre posible para $m \geq 1$, entonces además de la convergencia se tiene que

$$\max_j |u_Y(t_j) - u_j| \text{ es del orden de } O(h^2 + k^2),$$

para h y k suficientemente pequeñas.

Otro procedimiento de discretización para el cálculo numérico de la solución regularizada $u_Y(t)$ para el problema (4.1.6) bajo discusión consiste, para empezar, en considerar a la funcional regularizante discretizada:

$$(4.1.10) \quad \Phi_Y[\tilde{u}; \chi] = \sum_{i=1}^n \sum_{j=1}^p k(x_i, t_j) u_j \sigma_j - y_i)^2 \rho_i + \gamma (h^{-(2m-1)} \sum_{j=1}^{n-m} |\Delta^{(m)} u_j|^2 + \alpha k \sum_{j=1}^n |u_j|^2).$$

donde $\Delta^{(m)} u_j$ denota a la m -ésima diferencia hacia adelante con pivote en u_j . La cual se obtiene al discretizar la funcional de Tijonov $\Phi_Y[u; f_0]$ dada por (4.1.6') de manera análoga al procedimiento (de Glasko y Tijonov [34]) anteriormente descrito.

Aplicando las técnicas de Cálculo Diferencial para minimizar a $\Phi_Y[\tilde{E}; \underline{y}]$ con respecto a \tilde{E} , se verifica directamente que el mínimo se alcanza en \tilde{E}_Y , la solución del sistema lineal

$$(4.1.11) \quad (\tilde{X}'P\tilde{X} + \gamma\tilde{B})\tilde{E} = \tilde{X}'P\underline{y}$$

con

$$\tilde{X} = K\tilde{S} \quad \text{y} \quad \tilde{B} = h^{-(2m-1)} \Delta^{[2m]} + o(hI),$$

en donde

$$\Delta^{[2m]} = (\Delta^{[m]})_+ \Delta^{[m]},$$

siendo $\Delta^{[m]}$ de $(n-m) \times p$ la matriz de m -ésima diferencias hacia adelante. Por ejemplo, para $m=1$,

$$(m=1) \quad \Delta^{[2m]} = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & & & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix}.$$

Cabe decir que la matriz \tilde{B} que resulta de discretizar la norma de Sobolev

$$S[u] = \int_a^b (|u^{(m)}(t)|^2 + \alpha|u(t)|^2) dt,$$

es siempre simétrica y positiva definida, lo que no ocurre con B^* en (4.1.9) en general; y que los sistemas lineales (4.1.9) y (4.1.11) son también en general diferentes.

Ahora, si $\pi_p: \mathbb{R}^p \rightarrow C^{m-1}[a, b]$ denota al operador de prolongación mediante, por ejemplo, la interpolación spline de orden m , entonces se demuestra que (*)

(4.1.12) $\|u_Y - \pi_p B_{-1}^{-1} c_{m-1}\| \rightarrow 0$ cuando $h, h \rightarrow 0^+$, como puede verse en Vasin [110] (véase también a Goncharakii et. al. [106]).

Una tercera variante de discretización para efectos del cálculo numérico de la solución regularizada $u_Y(t)$ para la ecuación integral (4.1.6), sobre la cual se basa nuestro algoritmo numérico el cual se describe en la sec. 4 de este capítulo, es la siguiente.

Primeramente se discretiza mediante la aplicación de la cuadratura numérica.

$$\int_a^b \varphi(t) dt = \sum_{j=1}^p \varphi(t_j) \sigma_j,$$

nuestra ecuación integral (4.1.6) tal y como se describió en la subsección anterior, obteniéndose el modelo lineal

$$\sum_{k=1}^n \frac{B}{n \times p} + \underline{\varepsilon} = \underline{Y}, \quad \underline{\varepsilon} \sim (0, \delta^2 I),$$

con \underline{Y} , \underline{B} , $\underline{K} = KS$ como en (4.1.9), y $\underline{\varepsilon}$ siendo un vector aleatorio en \mathbb{R}^n .

(*) Recuérdese que $\|g\|_{C^{m-1}} = \max_{0 \leq i < m-1} \max_{a < t < b} |g^{(i)}(t)|$.

A continuación se considera a la funcional regularizante

$$J_{\gamma}[\tilde{u}; y] = \|y - X \tilde{u}\|_2^2 + \gamma \|\tilde{u}\|_B^2 \quad (*), \quad \gamma > 0,$$

en donde

$$(4.1.12') \quad B = h^{-1} \tilde{B} = h^{-2m} \Delta^{[2m]} + \alpha I,$$

siendo \tilde{B} la matriz como en (4.1.11) que se obtiene al discretizar por diferencias divididas la norma de Sobolev

$$S[u] = \int_a^b (|u^{(m)}(t)|^2 + \alpha |u(t)|^2) dt,$$

Finalmente, aplicando las técnicas del Cálculo Diferencial para hallar el mínimo de la funcional $J_{\gamma}[\tilde{u}; y]$ con respecto a \tilde{u} , se obtiene el sistema lineal

$$(4.1.13) \quad (X'X + \gamma B)\tilde{u} = X'y,$$

el cual está estrechamente relacionado con la estimación sesgada conocida por "ridge, en la Regresión Lineal; estrecha relación que aprovecharemos para la realización práctica del *Principio de Seudo-optimalidad* para una buena elección del parámetro de regularización.

Ahora, si en (4.1.13) se toma a $B = C_0 h^{-1} \tilde{B}$ con $C_0 = h/k$, entonces es directo verificar que (4.1.11) y (4.1.13) son exactamente el mismo sistema lineal cuando la cuadratura numérica

(*) En el entendido que $\|\tilde{u}\|_B^2 = \tilde{u}' B \tilde{u}$.

$$\int_0^1 \psi(s) ds \approx \sum_{i=1}^n \psi(x_i) \rho_i$$

es la del Rectángulo sobre una partición uniforme de paso h para $[c,d]$; y que (4.1.11) y (4.1.13) son también, salvo una ligera discrepancia despreciable en el límite, esencialmente el mismo sistema lineal cuando la antes especificada cuadratura es la del Trapecio. En conclusión para esta tercera variante de discretización para efectos del cálculo numérico de la solución u_Y regularizada para la ecuación integral (4.1.6) se tiene que

$$\|u_Y - \pi_p \hat{u}_Y\|_C = O(h^k),$$

cuando $h, k \rightarrow 0^+$ manteniendo constante a la razón h/k .

§4.2. Estimación Gauss-Markov Regularizada.

Una alternativa a la estimación Gauss-Markov clásica $\hat{\beta}$ bajo presencia de colinealidad, la cual usualmente se manifiesta al tomar $\|N\hat{\beta}\|_2$ valores inaceptablemente grandes, la sugiere de manera natural el Método de Regularización de Tikhonov para la resolución de Problemas Mal-Planteados a la Madamard, el cual ya hemos estudiado en el Cap. 2.

Con respecto al modelo lineal general

$$(4.2.1) \quad \sum_{i=1}^n \beta_i x_i = y \quad \text{con } \xi \sim (0, \delta^2 \Sigma),$$

(*) Como ya se anotó antes, $\|N\hat{\beta}\|_2 = \sqrt{\sum_{i=1}^n \beta_i^2}$; análogamente se entien-

la idea consiste en tomar a la siguiente funcional regularizante

$$(4.2.2) \quad J_{\gamma}[\underline{\beta}; \underline{y}] = \|\underline{y} - \underline{X}\underline{\beta}\|_{\Sigma}^2 + \gamma \|\underline{\beta}\|_B^2 \quad (*), \quad \gamma > 0,$$

en donde Σ y B son matrices simétricas y positivas definidas⁽⁵⁾; para luego definir a la *Estimación Gauss-Markov Regularizada* como

$$(4.2.3) \quad \hat{\underline{\beta}}_{\gamma} = \arg \min_{\underline{\beta}} J_{\gamma}[\underline{\beta}; \underline{y}].$$

Así pues, a la estimación GM regularizada se le puede pensar como una estimación con base al criterio de mínimo de cuadrados - generalizados amortiguados.

Ahora, aplicando las técnicas del Cálculo Diferencial, es directo verificar que

$$(4.2.3') \quad \hat{\underline{\beta}}_{\gamma} = (\underline{X}'\Sigma^{-1}\underline{X} + \gamma B)^{-1} \underline{X}'\Sigma^{-1} \underline{y},$$

la cual para $\Sigma = \underline{I}_n$ y $B = \underline{I}_p$ se reduce a la clásica expresión conocida por *estimación "ridge"* debida a Hoerl y Kennard.

Conviene observar que para $\gamma = 0$, $\hat{\underline{\beta}}_{\gamma}$ es la estimación GM generalizada

$$\hat{\underline{\beta}} = (\underline{X}\Sigma^{-1}\underline{X})^{-1} \underline{X}'\Sigma^{-1} \underline{y};$$

(*) Como ya se anotó antes, $\|\underline{z}\|_B^2 = \underline{z}' B \underline{z}$; análogamente se entiende $\|\underline{z}\|_{\Sigma}^2$.

(5) Para \underline{B} pensado como un vector aleatorio con $\text{COV}(\underline{\beta})$, conocida es natural tomar a B de manera que $B^{-1} = \text{COV}(\underline{\beta})$. Por otro lado, con relación a la resolución numérica de la ecuación integral de Fredholm de primer tipo, objetivo central de este trabajo, B se toma como dada por (4.1.12').

la cual es "blue" con respecto al producto interior $\underline{x}' \Sigma^{-1} \underline{x}$,
 R-optimal con respecto a toda función de pérdida cuadrática dada

$$(4.2.4) \quad L_C(\underline{\beta}) = (\underline{\beta} - \underline{\beta})' C(\underline{\beta} - \underline{\beta})$$

como puede verse en Bibby et. al. [104], y que es precisamente la
 estimación GM ordinaria

$$(4.2.4') \quad \hat{\underline{\beta}} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y},$$

para $\underline{\beta}$ es el modelo lineal estandard

$$(4.2.5) \quad \underline{X} \underline{\beta} + \underline{\varepsilon} = \underline{Y} \quad \text{con} \quad \underline{\varepsilon} \sim (0, \delta^2 \underline{I}),$$

si con base a la descomposición de Cholesky $\Sigma = \underline{L}\underline{L}'$ (\underline{L} triangular
 inferior) de Σ , se toma

$$\underline{X} = \underline{L}^{-1} \underline{X}, \quad \underline{\varepsilon} = \underline{L}^{-1} \underline{\varepsilon} \quad \text{y} \quad \underline{Y} = \underline{L}^{-1} \underline{Y}.$$

Es cuyo caso, la estimación GM regularizada (4.2.3') toma la
 forma

$$(4.2.5') \quad \hat{\underline{\beta}}_{\gamma} = (\underline{X}' \underline{X} + \gamma \underline{B})^{-1} \underline{X}' \underline{Y};$$

y aunque esta simplificación siempre es posible, cabe decir que
 bien puede acarrear serias dificultades prácticas desde un punto
 de vista numérico cuando Σ es mal-comportada.

Para lo que sigue de nuestra discusión, el modelo (4.2.1)
 será referido por el modelo θ , y el modelo lineal

$$(4.2.6) \quad \underline{T} \underline{\theta} + \underline{\varepsilon} = \underline{y} \quad \text{con} \quad \underline{\varepsilon} \sim (0, \delta^2 \underline{\Sigma})$$

por el modelo θ ; el cual se obtiene al considerar la descomposición

de Cholesky $B = R'R$ (R triangular superior) de B , al tomar

$$T = X R^{-1},$$

y efectuar el cambio de variables $\hat{\beta} = \hat{\Theta}$. Y cuya estimación GM regularizada $\hat{\Theta}_Y$ para Θ con base a (4.2.3') toma la forma

$$(4.2.7) \quad \hat{\Theta}_Y = (T' \Sigma^{-1} T + \gamma I)^{-1} T' \Sigma^{-1} Y,$$

que bien se puede describir como

$$(4.2.8) \quad \hat{\Theta}_Y = (T' T + \gamma I)^{-1} T' Y,$$

si se le piensa con respecto al modelo lineal estandard

$$(4.2.8') \quad T \Theta + \underline{\varepsilon} = Y \quad \text{con} \quad \underline{\varepsilon} \sim (0, \delta^2 I),$$

el cual se obtiene de (4.2.5), tomando

$$T = L^{-1} T, \quad \underline{\varepsilon} = L^{-1} \underline{\varepsilon} \quad \text{y} \quad Y = L^{-1} Y.$$

Las relaciones siguientes entre estos dos modelos (β y Θ) son de directa verificación:

$$(4.2.9.a) \quad \hat{\beta}_B^2 = \hat{\beta}_\Theta^2 \quad (6),$$

$$(4.2.9.b) \quad \hat{\beta}_C(\hat{\beta}_Y) = \hat{\beta}_C(\hat{\Theta}_Y) \quad \text{donde} \quad G = R^{-1} C R^{-1} \quad (*) \quad \text{y}$$

$$(4.2.9.c) \quad \text{ecm}_C(\hat{\beta}_Y) = \text{ecm}_G(\hat{\Theta}_Y).$$

(6). Aquí, siendo X de rango máximo: $\hat{\beta}_B^2 = \frac{1}{n-p} (Y - X \hat{\beta})' \Sigma^{-1} (Y - X \hat{\beta})$ y $\hat{\beta}_\Theta^2 = \frac{1}{n-p} (Y - T \hat{\Theta})' \Sigma^{-1} (Y - T \hat{\Theta})$.

(7). De aquí en adelante $R^{-1} = (R')^{-1}$.

La importancia de estas relaciones se debe a que los cálculos con respecto al modelo Θ , pensando en el importante caso particular con $\Sigma = I$, son de mucho menor complejidad que los correspondientes con respecto al modelo β . Por otra parte, permiten dar, en el contexto del modelo lineal general (4.2.1), demostración directa del siguiente

Teorema 4.2.1. Existe $\gamma^0 = \gamma^0(\underline{\beta}; \delta^2) > 0$ para la cual

$$\text{ecm}_C(\hat{\underline{\beta}}_\gamma) \leq \text{ecm}_C(\hat{\underline{\beta}}), \text{ si } 0 < \gamma \leq \gamma^0.$$

En efecto, su demostración se sigue de (4.2.9.c) y el teorema 10.2 en Bibby et. al. [104] aplicado al modelo Θ dado por (4.2.7).

Así pues, con base en este teorema es natural proponer el criterio siguiente, el cual llamaremos *Principio de Optimalidad*, para la elección del parámetro de regularización

$$\gamma_{OP} = \arg \min_{\gamma > 0} \text{ecm}_C(\hat{\underline{\beta}}_\gamma),$$

estando claro que

$$\text{ecm}_C(\hat{\underline{\beta}}_\gamma) = \text{ecm}_C(\hat{\underline{\beta}}_\gamma; \underline{\beta}; \delta^2);$$

y con ello que

$$\gamma_{OP} = \gamma_{OP}(\underline{\beta}; \delta^2).$$

Ahora bien, con el propósito de indagar bajo qué condiciones la función en γ :

$$\text{ecm}_C(\gamma) = \text{ecm}_C(\hat{\underline{\beta}}_\gamma; \underline{\beta}; \delta^2)$$

tiene mínimo, considérese a la descomposición en valores singulares de \mathbf{T} :

$$(4.2.10) \quad \mathbf{T}_{n \times p} = \mathbf{U} \Delta \mathbf{V}' \quad \Delta_{n \times p} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\mathbf{D} = \text{diag}(s_1, s_2, \dots, s_r), \quad s_1 > s_2 > \dots > s_r > 0;$$

siendo $r \leq \min\{n, p\}$ el rango de \mathbf{T} ; en términos de la cual se tiene el siguiente

Lema 4.2.1. Si $\mathbf{G}(\gamma) = -\mathbf{R}^{-1} \mathbf{C} \mathbf{R}^{-1} = \mathbf{V} \mathbf{W} \mathbf{V}'$ con $\mathbf{W} = \text{diag}(w_1, w_2, \dots, w_p)$ donde $w_i > 0$ para toda i , entonces

$$(4.2.11) \quad \text{vt}_{\mathbf{C}}(\gamma; \delta^2) = \text{vt}_{\mathbf{C}}(\hat{\beta}_{\gamma}; \delta^2) = \delta^2 \sum_{i=1}^r w_i \frac{s_i^2}{(s_i^2 + \gamma)^2},$$

$$(4.2.12) \quad \text{b}_{\mathbf{C}}^2(\gamma; \underline{\beta}) = \text{b}_{\mathbf{C}}^2(\hat{\beta}_{\gamma}; \delta^2) \\ = \gamma^2 \sum_{i=1}^r w_i \frac{\alpha_i^2}{(s_i^2 + \gamma)^2} + \sum_{i=r+1}^p w_i \alpha_i^2.$$

Y con ello,

$$(4.2.13) \quad \text{ecm}_{\mathbf{C}}(\gamma; \underline{\beta}, \delta^2) = \text{ecm}_{\mathbf{C}}(\hat{\beta}_{\gamma}; \underline{\beta}, \delta^2) \\ = \delta^2 \sum_{i=1}^r w_i \frac{s_i^2}{(s_i^2 + \gamma)^2} + \gamma^2 \sum_{i=1}^r w_i \frac{\alpha_i^2}{(s_i^2 + \gamma)^2} + \sum_{i=r+1}^p w_i \alpha_i^2,$$

en donde $\underline{\alpha} = \mathbf{V}' \underline{\mathbf{0}}$ con $\underline{\mathbf{0}} = \mathbf{R} \underline{\beta}$.

Cuya laboriosa demostración aparece el final de esta sección como un apéndice.

Así, con apoyo en la expresión (4.2.13) y considerando que

$$(4.2.14.i) \quad \underline{\hat{\beta}} \in N(X) \text{ ssi } X\underline{\hat{\beta}} = \underline{0} \text{ ssi } L^{-1}X\underline{\hat{\beta}} = \underline{0} \text{ ssi } (L^{-1}XR^{-1})R\underline{\hat{\beta}} = \underline{0} \\ \text{ssi } T\underline{\hat{\beta}} = \underline{0} \text{ ssi } \underline{\hat{\beta}} \in N(T),$$

$$(4.2.14.ii) \quad \underline{\hat{\beta}} \in N(X)^\perp \text{ ssi } \underline{\hat{\beta}} = X^+X\underline{\hat{\beta}} \text{ ssi } \underline{\hat{\beta}} = (LFR)^+(LFR)\underline{\hat{\beta}} \\ \text{ssi } \underline{\hat{\beta}} = T^+T\underline{\hat{\beta}} \text{ ssi } \underline{\hat{\beta}} \in N(T)^\perp, \quad y$$

$$(4.2.14.iii) \quad \|\underline{\hat{\beta}}\|_C = \underline{\hat{\beta}}'C\underline{\hat{\beta}} = (R\underline{\hat{\beta}})'G(R\underline{\hat{\beta}}) = \|V\underline{\hat{\beta}}\|_W^2 = \|\underline{\hat{\alpha}}\|_W^2;$$

se demuestra, como puede verse en Barrios et. al. [9], el siguiente

Teorema 4.2.2. Sea $\text{rg}(X) > 1$. La función $\text{ecm}_C(\hat{\beta}^2; \underline{\beta}, \cdot)$ como función de y tiene un único mínimo en $0 < y < \infty$ ssi $\underline{\beta} \notin N(X)$.

Las figuras siguientes ilustran este resultado.

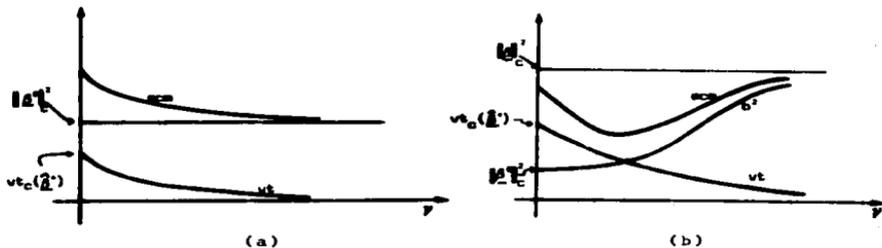


Fig. 1. Dos casos excluyentes: (a) $\underline{\hat{\beta}} \in N(X)$, y (b) $\underline{\hat{\beta}} \notin N(X)$. Aquí, $\underline{\hat{\beta}}^0$ es la proyección de $\underline{\hat{\beta}}$ sobre $N(X)$;

$$\|\underline{\hat{\beta}}^0\|_C^2 = \sum_{i=1}^p w_i \alpha_i^2 \quad \text{y} \quad \|\underline{\hat{\beta}}\|_C^2 = \sum_{i=1}^p w_i \alpha_i^2.$$

Para lo que sigue se dirá que $\underline{\beta}$ es estimable si $\underline{\beta} = \underline{0}$ o $\underline{\beta} \notin N(X)$ (compárese con Albert [102]). En estos términos, se tiene que $\text{ecm}_C(\gamma) = \text{ecm}_C(\hat{\underline{\beta}}_\gamma; \underline{\beta}; \delta^2)$ tiene un único mínimo ssi $\underline{\beta}$ es estimable.

Cabe anotar que cuando $\underline{\beta}$ es estimable, $\hat{\underline{\beta}}^+ = X^+ \underline{y}^{(*)}$ es una estimación insesgada para $\underline{\beta}^{\dagger(7)}$, la proyección ortogonal de $\underline{\beta}$ sobre $N(X)^\perp = R(X')$.

Ahora bien, considerando que cuando $\underline{\beta}$ es estimable, lo que es realmente relevante es la estimación de $\underline{\beta}^{\dagger}$; y ya que $\hat{\underline{\beta}}_\gamma \rightarrow \underline{\beta}^{\dagger}$ cuando $\gamma \rightarrow 0^+$, en lo que resta del presente trabajo, nuestro estudio se llevará a cabo en lo que respecta a las funciones $\text{ecm}_C(\hat{\underline{\beta}}_\gamma; \underline{\beta}; \delta^2)$ y $b_C^2(\hat{\underline{\beta}}_\gamma; \underline{\beta})$, pensando que $\underline{\beta} = \underline{\beta}^{\dagger}$, o equivalentemente que $\underline{\beta} = X^+ X \underline{\beta}$.

Ahora bien, como para fines numéricos da prácticamente casi lo mismo que X sea de rango deficiente o que X sea mal-comportada, se tiene en la mano una buena explicación de por qué el cálculo numérico de $\gamma_{\text{op}}^{(**)}$ es difícil de determinar con una relativa buena precisión, en cuanto al orden de magnitud se refiere. Pues en tal situación, la función $\text{ecm}_C(\gamma) = \text{ecm}_C(\hat{\underline{\beta}}_\gamma)$ presenta usualmente un amplio plano valle, como se ilustra a continuación.

(*) Aquí, X^+ es la inversa generalizada de Moore-Penrose con respecto a R^p con el producto interior usual y R^n con el producto interior $\underline{x}' \Sigma^{-1} \underline{x}$.

(7) Más aún, si $\phi = \underline{a}' \underline{\beta}$ es estimable (i.e. $\underline{a} \notin N(X)$) entonces la "blu" estimación para ϕ está dada por $\hat{\phi} = \underline{a}' \underline{\beta}$ ssi $\underline{\beta} = X^+ \underline{y}$. Compárese con Albert [102]

(**) Para $\underline{\beta}$ en el modelo (4.2.1) estimable.

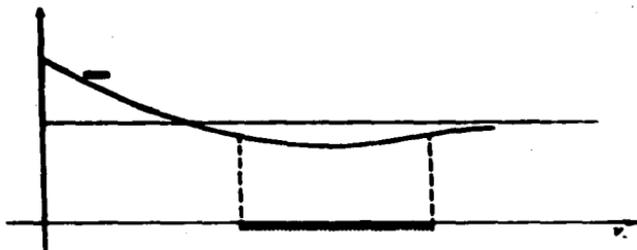


Fig. 2. Gráfica típica del $ecm_C(\gamma)$, cuando X es mal-comportada.

Así pues, para X mal-comportada, el cálculo numérico de γ_{Op} resulta ser un problema, en general, numéricamente mal-comportado.

Concluimos esta sección advirtiendo al lector que en lo que resta de este trabajo, nuestro estudio lo limitaremos alrededor del modelo lineal estándar β dado por (4.2.5), por ser ello suficiente a nuestros propósitos específicos.

APENDICE. Demostración del Lema 4.2.1.

Con respecto al modelo lineal general (4.2.1), la estimación GM regularizada (para β)

$$(4.2.3') \quad \hat{\beta}_Y = X_Y^{-1} X' \Sigma^{-1} Y = \tilde{Z}_Y \hat{\beta}^+$$

en donde

$$\tilde{Z}_Y = X_Y^{-1} X' \Sigma^{-1} X, \quad X_Y = X' \Sigma^{-1} X + \gamma B$$

$$\hat{\beta}^+ = (X' \Sigma^{-1} X + \gamma B)^{-1} X' \Sigma^{-1} Y = A^+ Y;$$

se puede describir con respecto al modelo lineal estandar (4.2.5) como

(4.2.5')

$$\hat{\beta}_Y = \mathbf{X}_Y^{-1} \mathbf{X}' \mathbf{Y} = \mathbf{Z}_Y \hat{\beta}^+$$

en donde

$$\mathbf{Z}_Y = \mathbf{X}_Y^{-1} \mathbf{X}' \mathbf{X}, \quad \mathbf{X}_Y = \mathbf{X}' \mathbf{X} + \gamma \mathbf{I},$$

y

$$\hat{\beta}^+ = (\mathbf{X}' \mathbf{X})^+ \mathbf{X}' \mathbf{Y} = \mathbf{X}^+ \mathbf{Y}.$$

Afirmación 4.2.1. Para el modelo lineal estandar (4.1.5), $\hat{\beta}^+$ es una estimación insesgada para $\beta^1 = \mathbf{X}^+ \mathbf{X} \beta$. Más aún,

$$\hat{\beta}^+ \sim (\beta^1, \delta^2 (\mathbf{X}' \mathbf{X})^+).$$

En efecto, como en el modelo (4.2.5), $\underline{\varepsilon} \sim (\mathbf{0}, \delta^2 \mathbf{I})$,

$$\mathbb{E}(\hat{\beta}^+) = \mathbb{E}(\mathbf{X}^+ (\mathbf{X} \beta + \underline{\varepsilon})) = \mathbf{X}^+ \mathbf{X} \beta = \beta^1,$$

Y como

$$\hat{\beta}^+ = \mathbf{X}^+ (\mathbf{X} \beta + \underline{\varepsilon}) = \mathbb{E}(\hat{\beta}^+) + \mathbf{X}^+ \underline{\varepsilon},$$

se sigue que

$$\text{Cov}(\hat{\beta}^+) = \text{Cov}(\mathbf{X}^+ \underline{\varepsilon}) = \mathbf{X}^+ \text{Cov}(\underline{\varepsilon}) (\mathbf{X}^+)' = \delta^2 (\mathbf{X}' \mathbf{X})^+.$$

Ahora, recordando que bajo la transformación $\hat{\theta} = \mathbf{R} \tilde{\theta}$ la estimación GN regularizada (4.2.3') en su forma (4.2.5'), se convierte en

(4.2.8)

$$\hat{\theta}_Y = \mathbf{T}_Y^{-1} \mathbf{T}' \mathbf{Y} = \mathbf{Z}_Y \hat{\theta}^+$$

con

$$\hat{\theta}^+ = \mathbf{T}^+ \mathbf{Y} \sim (\theta^1, \delta^2 (\mathbf{T}' \mathbf{T})^+),$$

en donde

$$\mathbb{E}_Y = \mathbb{T}_Y^{-1} \mathbb{T}' \mathbb{T} \quad \text{y} \quad \mathbb{T}_Y = \mathbb{T}' \mathbb{T} + \gamma \mathbb{I};$$

se tiene la siguiente

Afirmación 4.2.2. Si $\underline{\theta} = \underline{\theta}^1 + \underline{\theta}^0$ con $\underline{\theta}^0$ siendo la proyección ortogonal de $\underline{\theta}$ sobre $N(\mathbb{T})$ entonces

$$(4.2.15) \quad b_C^2(\gamma; \underline{\beta}) = \|\gamma \mathbb{T}_Y^{-1} \underline{\theta}^1 - \underline{\theta}^0\|_G^2.$$

En efecto, ya que

$$\mathbb{E}\{\hat{\underline{\theta}}_Y\} = \mathbb{E}_Y \underline{\theta}^1,$$

se tiene que

$$\begin{aligned} b_C^2(\gamma; \underline{\beta}) &= \|\mathbb{E}\{\hat{\underline{\beta}}_Y\} - \underline{\beta}\|_C^2 = \|\mathbb{R}^{-1}(\mathbb{E}\{\hat{\underline{\theta}}_Y\} - \underline{\theta})\|_C^2 \\ &= \|\mathbb{E}_Y \underline{\theta}^1 - \underline{\theta}\|_G^2 = \|(\mathbb{E}_Y - \mathbb{I})\underline{\theta}^1 - \underline{\theta}^0\|_G^2 \\ &= \|\gamma \mathbb{T}_Y^{-1} \underline{\theta}^1 - \underline{\theta}^0\|_G^2, \end{aligned}$$

pues

$$\mathbb{E}_Y - \mathbb{I} = \mathbb{T}_Y^{-1} (\mathbb{T}' \mathbb{T} - \mathbb{T}_Y) = -\gamma \mathbb{T}_Y^{-1}.$$

Con base en la descomposición (4.2.10) en valores singulares de \mathbb{T} se demuestra la siguiente

Afirmación 4.2.3. Si $G = (\mathbb{R}^{-1} \mathbb{C} \mathbb{R}^{-1}) = \mathbb{V} \mathbb{W} \mathbb{V}'$ donde $\mathbb{W} = \text{diag}(w_1, \dots, w_p)$ con $w_i > 0$ para toda i , entonces

$$(4.2.12) \quad b_C^2(\gamma; \underline{\beta}) = \gamma^2 \sum_{i=1}^p w_i \frac{a_i^2}{(w_i^2 + \gamma)^2} + \sum_{i=p+1}^P w_i a_i^2,$$

en donde $\underline{\alpha} = V' \underline{\theta}$.

En efecto, de (4.2.15) se sigue que

$$b_C^2(\gamma; \underline{\beta}) = \|\gamma - \mathbb{T}_Y^{-1} \underline{\theta}^1 - \underline{\theta}^0\|_C^2 = \|\gamma - \Delta_Y^{-1} \underline{\alpha}^1 - \underline{\alpha}^0\|_W^2,$$

en donde

$$\Delta_Y^{-1} = (\Delta' \Delta + \gamma I)^{-1}$$

Y como

$$(\underline{\alpha}^0)' W \Delta_Y^{-1} \underline{\alpha}^1 = (\underline{\alpha}^1)' \Delta_Y^{-1} W \underline{\alpha}^0 = 0,$$

debido a que $\underline{\alpha}^0$ y $\underline{\alpha}^1$ son ortogonales y $W \Delta_Y^{-1} = \Delta_Y^{-1} W$ es una matriz diagonal; se sigue que

$$b_C^2(\gamma; \underline{\beta}) = \|\gamma - \Delta_Y^{-1} \underline{\alpha}^1\|_W^2 + \|\underline{\alpha}^0\|_W^2. \quad \theta$$

Afirmación 4.2.4.

$$(4.2.16) \quad vt_C(\gamma; \delta^2) = vt_C(\hat{\underline{\theta}}_Y; \delta^2) = \text{tr}[G \mathbb{T}_Y^{-1} (\mathbb{T}_Y' \mathbb{T}) \mathbb{T}_Y^{-1}].$$

En efecto,

$$\begin{aligned} vt_C(\gamma; \delta^2) &= \mathbb{E}\{\|\hat{\underline{\beta}}_Y - \mathbb{E}\{\hat{\underline{\beta}}_Y\}\|_C^2\} \\ &= \mathbb{E}\{\|\mathbb{R}^{-1}(\hat{\underline{\theta}}_Y - \mathbb{E}\{\hat{\underline{\theta}}_Y\})\|_C^2\} \\ &= \mathbb{E}\{\|\mathbb{S}_Y(\hat{\underline{\theta}}^+ - \underline{\theta}^1)\|_C^2\} = \text{tr}[G \text{Cov}(\mathbb{S}_Y \hat{\underline{\theta}}^+)] \\ &= \text{tr}[G \mathbb{S}_Y \text{Cov}(\hat{\underline{\theta}}^+) \mathbb{S}_Y'] = \delta^2 \text{tr}[G \mathbb{S}_Y (\mathbb{T}' \mathbb{T})^+ \mathbb{S}_Y] \\ &= \delta^2 \text{tr}[G \mathbb{T}_Y^{-1} (\mathbb{T}' \mathbb{T}) \mathbb{T}_Y^{-1}]. \quad \theta \end{aligned}$$

Afirmación 4.2.5. Si $G = R^{-1}CR^{-1} = WVW'$ con $W = \text{diag}(w_1, w_2, \dots, w_p)$, donde $w_i > 0$ para toda i , entonces

$$(4.2.11) \quad vt_C(\gamma; \delta^2) = \delta^2 \sum_{i=1}^p w_i \frac{s_i^2}{(s_i^2 + \gamma)^2}.$$

Efectivamente, en términos de la descomposición (4.2.10) en valores singulares de T , se tiene que

$$G T_Y^{-1} T' T_Y^{-1} = V(W \Delta_Y^{-1} \Delta' \Delta_Y^{-1}) V'$$

en donde $\Delta_Y = \Delta' \Delta + \gamma I$. Y con ello, que

$$\begin{aligned} vt_C(\gamma; \delta^2) &= \delta^2 \text{tr}[G T_Y^{-1} (T' T) T_Y^{-1}] \\ &= \delta^2 \text{tr}[W \Delta_Y^{-1} (\Delta' \Delta) \Delta_Y^{-1}]. \quad \# \end{aligned}$$

Con esto se completa la demostración del lema. #

§4.3. Principio de Seudo-optimalidad: Realización Práctica.

En el capítulo anterior se vió, como una alternativa al criterio natural con respecto al error cuadrático medio (i.e. al Principio de Optimalidad), que la elección del parámetro de regularización por el Principio de Seudo-optimalidad define un algoritmo de regularización a la Tijonov (teorema 3.4.1) para el problema

$$(4.3.1) \quad Au = f_\delta \text{ con } f_\delta = f + \xi,$$

en donde $A:U \rightarrow F$ (con U y F espacios de Hilbert) es lineal compacto y ξ es un vector aleatorio sobre $F^{(*)}$ con $E\{\xi\}=0$ y v.t. $\{\xi\} \leq \delta^2$; el cual comparte con el Principio de Optimalidad las mismas propiedades de convergencia asintótica con respecto a δ .

Específicamente, para el modelo lineal estandar

$$(4.3.2) \quad \mathbb{X} \underline{\beta} + \underline{\epsilon} = \underline{y} \quad \text{con } \underline{\epsilon} \sim (0, \delta^2 I),$$

(el cual se obtiene, como se vió en la sec. 4.1, al discretizar el problema (4.3.1) con A siendo un operador integral de Fredholm de primer tipo) por el lema 4.2.1. (con $\underline{\beta}$ estimable) se sigue que

$$vt_C(\hat{\underline{\beta}}_Y; \underline{\beta}) \leq w^2 \cdot v(\gamma), \quad v(\gamma) = \frac{F}{4\gamma} \delta^2,$$

y

$$b_C^2(\hat{\underline{\beta}}_Y; \underline{\beta}) \leq b^2(\gamma), \quad b^2(\gamma) = \|\underline{h}\|_C^2 \gamma^2$$

$$(w^2 = \max_i \{w_i\}; \quad r = \text{rgo}(\mathbb{X}))$$

siendo \underline{h} tal que $\underline{\beta} = \mathbb{X}' \mathbb{X} \underline{h}$; con base a lo cual el teorema 3.4.1 toma la forma siguiente⁽⁸⁾.

Teorema 4.3.1. Sea $\underline{\beta}$ estimable (i.e. $\underline{\beta} = \mathbb{X}' \mathbb{X} \underline{\beta}$), \underline{h} tal que $\underline{\beta} = \mathbb{X}' \mathbb{X} \underline{h}$ y $\tau^2 > 0$ dada. Si se elige a $\gamma = \gamma_{\text{SC}}(\delta)$ por el Principio de Seudo-optimalidad (i.e. como la solución de: $v(\gamma) = \tau^2 b^2(\gamma)$, $\gamma > 0$) entonces

(*) En el sentido dado en el apéndice de la sec. 3.4 de este trabajo.

(8). Aquí, por estar en dimensión finita, se tiene que $\underline{\beta} = (\mathbb{X}' \mathbb{X})^{\underline{v}} \underline{h}$ se satisface para toda $0 < \underline{v} \leq 1$ (con \underline{h} dependiendo de \underline{v}).

$$(\underline{c}) \sim \left(\frac{0}{\sigma^2 I} \right) \sup (\mathbb{E} \hat{\underline{\beta}}_Y - \underline{\beta} \mathbb{E} \underline{C}^2)^{1/2} < n \delta^{2/3},$$

con

$$n = (1 + \tau^{-2})^{1/2} (\tau \mathbb{E} \mathbb{H} \mathbb{H} \mathbb{C} / 2w)^{1/2}$$

De manera análoga, para γ -óptima se tiene que

$$\text{Mfn } \gamma > 0 \quad (\underline{c}) \sim \left(\frac{0}{\sigma^2 I} \right) \sup (\mathbb{E} \hat{\underline{\beta}}_Y - \underline{\beta} \mathbb{E} \underline{C}^2)^{1/2} < \bar{n} \delta^{2/3},$$

$$\bar{n} = n \text{ con } \tau^2 = 2;$$

lo cual como ya se ha anotado antes justifica el nombre, *Principio de Pseudo-optimalidad*, para nuestro criterio para la elección del parámetro de regularización.

Hasta aquí, se ha estudiado el *Principio de Pseudo-optimalidad* suponiendo a $\underline{\beta}$ (estimable) y δ^2 conocidos. Así pues, para poderlo llevar a la práctica se hace necesario recurrir a la *estimación estadística*, con el objetivo de lograr "buenas" estimaciones para

$$vt_{\underline{C}}(\hat{\underline{\beta}}_Y; \delta^2) \quad \text{y} \quad b_{\underline{C}}^2(\hat{\underline{\beta}}_Y; \underline{\beta})$$

con $\hat{\underline{\beta}}_Y$ siendo la *estimación de GM regularizada*

$$(4.3.3) \quad \hat{\underline{\beta}}_Y = (\mathbb{X}'\mathbb{X} + \gamma \mathbb{B})^{-1} \mathbb{X}' \underline{Y}$$

para $\underline{\beta}$ en el modelo lineal (4.3.2); en el cual, por razones de simplicidad, de aquí en lo sucesivo se supondrá con \mathbb{X} de rango máximo.

Para tal propósito es conveniente recordar (como se vió en la sección anterior) que con respecto al modelo lineal Θ :

$$(4.3.3') \quad \mathbb{T} \underline{\Theta} + \underline{\varepsilon} = \underline{Y} \quad \text{con} \quad \underline{\varepsilon} \sim (\underline{0}, \delta^2 \mathbf{I}),$$

donde $\mathbb{T} = \mathbf{R}\mathbf{R}'^{-1}$, siendo $\mathbf{B} = \mathbf{R}'\mathbf{R}$ la descomposición de Cholesky de \mathbf{B} ; se tiene, para el caso en que $\mathbf{G} (= \mathbf{R}'^{-1}\mathbf{C}\mathbf{R}^{-1}) = \mathbf{V}\mathbf{W}\mathbf{V}'$, que

$$(4.3.4) \quad \begin{aligned} \text{vt}_{\mathbf{C}}(\underline{Y}; \delta^2) &= \text{vt}_{\mathbf{C}}(\hat{\underline{\beta}}_{\underline{Y}}; \delta^2) = \text{vt}_{\mathbf{C}}(\hat{\underline{\Theta}}_{\underline{Y}}; \delta^2) \\ &= \delta^2 \sum_{i=1}^p w_i \frac{s_i^2}{(s_i^2 + \gamma)^2}, \end{aligned}$$

y que

$$(4.3.5) \quad \begin{aligned} b_{\mathbf{C}}^2(\underline{Y}; \underline{\beta}) &= b_{\mathbf{C}}^2(\hat{\underline{\beta}}_{\underline{Y}}; \underline{\beta}) = b_{\mathbf{C}}^2(\hat{\underline{\Theta}}_{\underline{Y}}; \underline{\Theta}) \\ &= \gamma^2 \sum_{i=1}^p w_i \frac{\alpha_i^2}{(s_i^2 + \gamma)^2}, \end{aligned}$$

en donde $\underline{\alpha} = \mathbf{V}'\underline{\Theta}$ ($\underline{\Theta} = \mathbf{R}\underline{\beta}$), siendo

$$(4.3.5') \quad \mathbb{T} = \mathbf{U}\mathbf{\Delta}\mathbf{V}', \quad \mathbf{\Delta} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \mathbf{D} & \cdot \\ \cdot & \cdot & \mathbf{O} \end{bmatrix}, \quad \mathbf{D} = \text{diag}(s_1, \dots, s_p)$$

$$(s_1 > s_2 > \dots > s_p > 0)$$

la descomposición en valores singulares de \mathbb{T} ($= \mathbf{R}\mathbf{R}'^{-1}$).

La idea natural que se ocurre es utilizar las estimaciones insesgadas

$$(4.3.6) \quad \hat{\delta}_{\underline{\beta}}^2 = \hat{\delta}_{\underline{\Theta}}^2 = \frac{1}{n-p} \|\underline{Y} - \mathbb{T}\hat{\underline{\Theta}}\|_2^2 (= \hat{\delta}^2)$$

$$(4.3.7) \quad \hat{v}_C(\gamma) = v t_C(\hat{\beta}_V; \delta^2) - v t_C(\hat{\beta}_V; \delta^2) \\ = \delta^2 \sum_{i=1}^p w_i \frac{a_i^2}{(s_i^2 + \gamma)^2},$$

$$(4.3.8) \quad \hat{b}_C^2(\gamma) = \hat{b}_C^2(\hat{\beta}_V; \hat{\beta}) = \hat{b}_C^2(\hat{\beta}_V; \theta) \\ = \gamma^2 \sum_{i=1}^p w_i \frac{\hat{a}_i - \delta^2/s_i^2}{(s_i^2 + \gamma)^2}$$

para $\delta^2, v t_C(\hat{\beta}_V; \delta^2)$ y $b_C^2(\hat{\beta}_V; \hat{\beta})$, respectivamente. Es directo verificar que (4.3.7) y (4.3.8) se obtienen de las expresiones siguientes

$$\hat{v}_C(\gamma) = v t_C(\hat{\beta}_V; \delta^2) - \text{sen}(v t_C(\hat{\beta}_V; \delta^2)), \text{ y}$$

$$\hat{b}_C^2(\gamma) = \hat{b}_C^2(\hat{\beta}_V; \hat{\beta}) - \text{sen}(b_C^2(\hat{\beta}_V; \hat{\beta}))$$

aplicadas a (4.3.4) y (4.3.5), respectivamente.

Para nuestros propósitos, en especial de cálculo numérico, es conveniente reescribir (4.3.8) como sigue

$$(4.3.8') \quad \hat{b}_C^2(\gamma) = \sum_{i=1}^p w_i \frac{a_i^2 - \delta^2}{s_i^2} \left(\frac{\gamma}{s_i^2 + \gamma} \right)^2,$$

la cual se obtiene directamente de ver que

$$\hat{\underline{a}} = v' \hat{\underline{\theta}} = v' \underline{\theta} + \underline{\gamma} = v' (UAV')^+ \underline{\gamma} = \Delta^+ (U' \underline{\gamma}) \\ = \Delta^+ \underline{a}, \text{ con } \underline{a} = U' \underline{\gamma}.$$

Nótese que en (4.3.8') no se requiere del cálculo explícito de $\hat{\beta}$ (y con ello de $\hat{\beta}$).

Regresando a nuestro punto central sobre la búsqueda de una realización práctica del *Principio de Seudo-optimalidad*, empezemos por ver que dicho principio, con base a las estimaciones insesgadas (4.3.7) y (4.3.8') para $vt_C(\hat{\beta}_Y; \delta^2)$ y $b_C^2(\hat{\beta}_Y; \underline{\beta})$ respectivamente, toma la forma

$$(4.3.9) \quad \hat{v}_C(\gamma) = \tau^2 \hat{b}_C^2(\gamma) \quad \text{con } \gamma > 0 (\tau^2 > 0 \text{ dada}).$$

Por el sorprendente hecho de que $\hat{b}_C^2(\gamma)$ puede tomar valores negativos o ser negativa, a pesar de ser $\hat{b}_C^2(\gamma)$ una estimación insesgada para $b_C^2(\underline{\beta}_Y; \underline{\beta})$ que es siempre no-negativa, resulta que la ecuación de *seudo-optimalidad* (4.3.9) no siempre tiene solución.

Teorema 4.3.2. Si $\|\hat{\beta}\|_C^2 > \hat{v}_C(\hat{\beta})$ entonces la ecuación de *seudo-optimalidad* (4.3.9) tiene solución. Y si $\min\{d_1^2\} > \delta^2$ entonces (4.3.9) tiene una única solución.

La *demonstración* para la primera parte se obtiene de aplicar el tma. del valor intermedio a la función $g(\gamma) = \hat{v}_C(\gamma) - \tau^2 \hat{b}_C^2(\gamma)$, viendo que $g(0^+) = \hat{v}_C(0^+) > 0$ y que

$$\lim_{\gamma \rightarrow \infty} g(\gamma) = -\tau^2 \lim_{\gamma \rightarrow \infty} \hat{b}_C^2(\gamma) = -\tau^2 (\|\hat{\beta}\|_C^2 - \hat{v}_C(\hat{\beta})) < 0.$$

Y la segunda parte se obtiene de ver que $\hat{b}_C^2(\gamma)$ es monótona creciente si $\min\{d_1^2\} > \delta^2$. *

Las figuras que vienen a continuación ilustran esta situación.

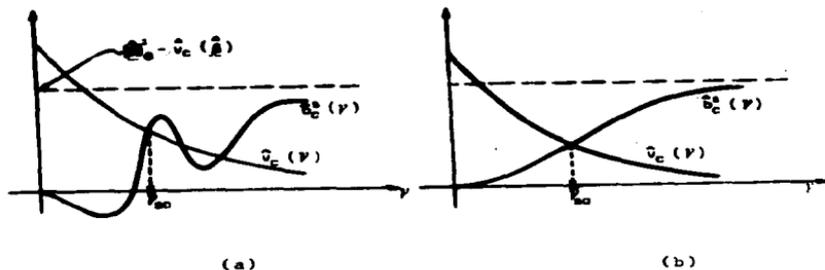


Fig. 3 (a). Existencia de $\tilde{\gamma}_{SO}$ cero de (4.3.9).

(b). Existencia y unicidad de $\tilde{\gamma}_{SO}$ cero de (4.3.9).

Es de interés mencionar que las condiciones de existencia con unicidad de este tma. son exactamente las mismas que se requieren satisfacer para que la estimación insesgada

$$(4.3.10) \quad C(\gamma) = -\hat{V}(\gamma) + \hat{B}_C^2(\gamma),$$

para $\text{ecm}_C(\hat{\beta}_\gamma; \beta, \delta^2)$, tenga un único mínimo. En concreto, si en (4.3.5') se considera a

$$U = [U_1 | U_2] \text{ con } U_1 \text{ de } n \times p,$$

entonces se tiene el siguiente

Teorema 4.3.3. Sea $P_{\text{Im}(\mathbb{B})} \mathbf{Y}$ la proyección ortogonal de \mathbf{Y} sobre $\text{Im}(\mathbb{B})$. Si $\hat{\delta}^2 \text{tr}(W)/\text{tr}(\mathbf{U}) < 1^{(9)}$ entonces la estimación insesgada $C(\gamma)$ para $\text{scm}_C(\hat{\beta}_\gamma; \underline{\beta}, \delta^2)$ tiene mínimo. Y si $\min(d_i^2) > \delta^2$ entonces $C(\gamma)$ tiene un único mínimo.

Su demostración es esencialmente la misma que se da para el tma. 2 en Barrios et. al. [9]. En efecto, para $C(\gamma) = \hat{V}_G(\gamma) + \hat{b}_G^2(\gamma)$ se tiene que

$$C'(\gamma) = 2 \frac{P(\gamma) - Q(\gamma)}{R(\gamma)}$$

en donde

$$P(\gamma) = \prod_{i=1}^p w_i \prod_{j \neq i} (s_j^2 + \gamma)^3 (d_i^2 - \delta^2) \gamma, \quad y$$

$$Q(\gamma) = \prod_{i=1}^p w_i \prod_{j \neq i} (s_j^2 + \gamma) (s_i \delta)^2$$

son polinomios de grados $3p-2$ y $3p-3$ respectivamente, y

$$R(\gamma) = \prod_{i=1}^p (s_i^2 + \gamma)^3 > 0, \quad \text{para toda } \gamma > 0.$$

Por lo que los ceros de $C'(\gamma)$ son los ceros de $g(\gamma) = P(\gamma) - Q(\gamma)$, e inversamente.

Así, si el coeficiente líder de $P(\gamma)$

(9). Si $C = B$ en $\mathcal{L}_C(\hat{\theta})$ entonces $G = VWV'(W=I)$ en $\mathcal{L}_C(\hat{\theta})$ como puede verse en la sec. 4.2. En tal circunstancia, esta condición toma la forma $p\hat{\delta}^2/\text{tr} P_{\text{Im}(\mathbb{B})} \mathbf{Y}^2 < 1$; la cual se puede reescribir como $R^2 > p/n$, siendo R^2 el coeficiente de determinación.

$$|U_1' P_{1m(\beta)} Y_{1m}^2 - \hat{\delta}^2 \text{tr}[W]| = \frac{P}{1} \sum_{i=1}^p w_i (d_i^2 - \hat{\delta}^2) > 0^{(*)},$$

entonces, observando que $g(0^+) = -Q(0^+) < 0$ y que $g(\gamma) \rightarrow \infty$ cuando $\gamma \rightarrow \infty$, y aplicando el tms. del valor intermedio, se tiene que $C(\gamma)$ tiene al menos un cero; siendo el primer cero de $C'(\gamma)$ - necesariamente un punto de mínimo.

Por último, si $\min\{d_i^2\} \geq \hat{\delta}^2$ entonces $P(\gamma)$ y $Q(\gamma)$ son monótonos crecientes; consecuentemente, $C'(\gamma)$ tiene un sólo cero. #

Con relación a este punto es recomendable la lectura de J. Bourges [12].

Observación 4.3.1. En nuestro estudio experimental sobre la resolución numérica de la ecuación integral de Fredholm de primer tipo, objetivo central de este trabajo, el cálculo de \tilde{Y}_{SO} como el de \tilde{Y}_{Op} vía las estimaciones insesgadas (4.3.7), (4.3.8') y (4.3.10) para $v_C(\hat{\beta}_Y; \hat{\beta})$, $b_C^2(\hat{\beta}_Y; \hat{\beta})$ y $ecm_C(\hat{\beta}_Y; \hat{\beta}, \hat{\delta}^2)$ con $C=B$ y $C=X'X$ y en donde $\hat{\beta}_Y = (X'X + \gamma B)^{-1} X'Y$, raramente dió resultados aceptables; pues usualmente $\hat{\beta}_Y \approx \hat{\beta}$ con $\gamma = \tilde{Y}_{SO}(\tilde{Y}_{Op})$ y $\hat{\beta}$ muy distinto de $\hat{\beta}$. #

La causa principal por la cual el cálculo de $\hat{\beta}_Y$ con γ determinado vía (4.3.9) (o (4.3.10)) dió raramente buenos resultados fue que $\hat{b}_C^2(\gamma)$ resultó generalmente negativa.

(*) . Escribiendo $U = [U_1 | U_2]$ con U_1 de $n \times p$, se tiene que $T = U_1 DV$;

$$\text{Así, } \frac{P}{1} \sum_{i=1}^p w_i d_i^2 = \frac{P}{1} \sum_{i=1}^p w_i (u_i' \hat{G}_1)^2 = |DV' \hat{G}_1^2| = |U_1' T \hat{\beta}_Y^2| \\ = |U_1' \hat{\beta}_Y^2|.$$

Es por ello que se ocurre la búsqueda de una estimación para $b_C^2(\gamma; \hat{\beta})$ que conserve las características geométricas de su gráfica ca. La siguiente es la obvia

$$(4.3.11) \quad \tilde{b}_C^2(\gamma) = b_C^2(\gamma; \hat{\beta}) = \frac{p}{1-\alpha} w_1 \left(\frac{d_1}{s_1} \right)^2 \left(\frac{\gamma}{s_1^2 + \gamma} \right)^2 ;$$

en términos de la cual, la ecuación de pseudo-optimalidad viene a ser

$$(4.3.12) \quad \hat{V}_C(\gamma) = \tau^2 \tilde{b}_C^2(\gamma), \quad \gamma > 0 \quad (\tau^2 > 0 \text{ dada}).$$

Ahora, es evidente que la ecuación (4.3.12) siempre tiene una única solución. Sin embargo, es importante decir que bien puede $\gamma = \tilde{\gamma}_{\text{MO}}$ verse fuertemente subestimada, debido a que $\tilde{b}_C^2(\gamma)$ tiene usualmente a sobre-estimar demasiado a $b_C^2(\hat{\beta}_\gamma; \hat{\beta})$, a causa de que $\hat{\beta}$ (en norma) tiende a dispararse (recuérdese la discusión vista en la sec. 4.1. sobre la estimación GM).

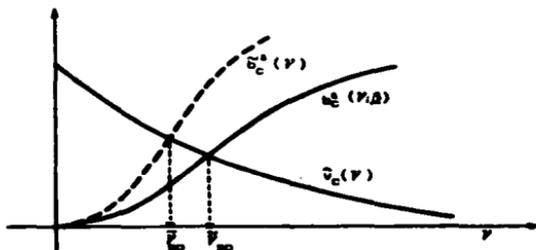


Fig. 4. Principio de Pseudo-optimalidad (versión (4.3.12)): Sub-estimación de $\gamma = \tilde{\gamma}_{\text{MO}}$.

Observación 4.3.2. En nuestro estudio experimental con base a la estimación *sesgada* (4.3.11) para $b_C^2(\gamma; \underline{\beta})$ para la determinación de $\gamma = \tilde{\gamma}_{so}$ por medio de la ecuación (4.3.12) con $C=B$, se observó con frecuencia una excesiva sub-valoración de γ_{so} , siendo $\hat{\beta}_\gamma$ con $\gamma = \tilde{\gamma}_{so}$ una muy mala estimación para $\underline{\beta}$. Sin embargo para $C = \mathbb{E}' \mathbb{E}$ (i.e. para $W = \text{diag}(s_1^2, \dots, s_p^2)$ en (4.3.7) y (4.3.11)), los resultados observados fueron bastante aceptables y competitivos a los obtenidos con $\gamma = \tilde{\gamma}_{Op}$ determinada mediante la minimización de la estimación (sesgada):

$$(4.3.13) \quad \tilde{C}(\gamma) = \hat{V}(\gamma) + \tilde{B}^2(\gamma)$$

para $\text{ecm}_C(\hat{\beta}_\gamma; \underline{\beta}, \delta^2)$; sobre todo para aquellas ecuaciones integrales con núcleo $k(x,t)$ apenas continuo o bien algebraico de clase C^∞ , como bien puede verse en la sección 4.5 para más detalles. #

Por último, diremos que el problema de fondo para todo criterio de selección del parámetro de regularización γ con base (explícita o implícita) en $b_C^2(\gamma; \underline{\beta})$, es un problema muy *mal-comportado*, o bien uno *mal-planteado* a la Hadamard si se piensa en dimensión infinita; que ello depende de $\underline{\beta}$, precisamente lo que se quiere bien "estimar". Esto plantea un *problema nuevo*, el cual se está investigando desde diversos enfoques; se recomienda leer a O' Sullivan [71] para ver uno de los más prometedores. Otro enfoque (tómese como sugerencia, por lo pronto) podría ser (¿Por qué no?). el de hallar una buena regularización para $b_C^2(\gamma; \underline{\beta})$. En fin, este problema es aún un problema abierto.

4.4. El Algoritmo: Descripción.

El objetivo de esta sección es la descripción del algoritmo para la resolución numérica de la ecuación integral de Fredholm de primer tipo

$$(4.4.1) \quad \int_a^b k(x,t) u(t) dt = f(x), \quad c \leq x \leq d,$$

por el método de regularización de Tijonov con la elección del parámetro de regularización por el Principio de Pseudo-optimalidad, objetivo central de este trabajo.

El algoritmo se elaboró en FORTRAN-77 con nombre REGULA.PACK, con apoyo en LINPACK (Sunah J. et al. [15]) reconfigurado a nuestros propósitos específicos como en las rutinas ZERON y FMIN (Forsythe et. al. [28]). El programa inicialmente se preparó en la Burroughs 7800 de la DGCA de la UNAM, luego se implantó en la red de computadoras Vax de la Univ. de Nuevo México en Albuquerque, y finalmente se ha instalado en una PC-Printaform con un disco duro de 20 Mb y coprocesador aritmético 8087.

El algoritmo a grandes rasgos consta de dos fases:

- 1a). *De discretización.* En esta fase como se vió en la sección 4.1:
 - a). Se discretiza a la ecuación integral (4.4.1) obteniéndose la matriz

$$\mathbb{K}_{n \times p} = (k(x_i, t_j) \sigma_j),$$

en donde las σ_j son los pesos de cuadratura empleada (Simpson, Trapecio y Punto Medio, en nuestro programa) sobre una partición uniforme del intervalo $[a,b]$ con espaciamento $h > 0$.

b). se genera al vector de "observaciones"

$$Y = (y_1, y_2, \dots, y_n)', y_i = f(x_i) + \epsilon_i \text{ con } \epsilon_i \sim N(0, \delta^2), \quad (10)$$

siendo la función f evaluada sobre una partición uniforme del intervalo $[c,d]$ con espaciamento $h > 0$. Y,

c). se discretiza a la norma de Sobolev

$$(4.4.2) \quad S[u] = \int_a^b (|u^{(m)}(t)|^2 + \alpha |u(t)|^2) dt$$

con $m=0, 1$ ó 2 opcional (i.e. con convergencia L^2, C o C^1 opcional). Para $m=0$, se toma $\alpha=0$; y para $m=1$ ó 2 se elige a $\alpha > 0$ ⁽¹¹⁾ de manera que siendo lo más pequeña posible se obtenga que la matriz

$$(4.4.3) \quad B = h^{-2m} (\Delta^{[m]})' \Delta^{[m]} + \alpha I = B^{[2m]} + \alpha I,$$

se a simétrica y positiva definida, en donde $\Delta^{[m]}$ de $(n-m) \times p$ es la matriz de diferencias hacia adelante de orden m .

2a). *De estimación Gauss-Markov regularizada.* Esta fase se puede ejecutar de manera independiente a la anterior y consiste en calcular la estimación GM regularizada

(10) Para nuestro estudio experimental se tomó a $\delta = 10^{-5}$.

(11) Nuestra elección fue $\alpha = 2 \times 10^{-7}$.

$$(4.4.4) \quad \hat{\underline{\beta}} = (\mathbf{X}'\mathbf{X} + \gamma\mathbf{B})^{-1} \mathbf{X}' \underline{y}$$

para $\underline{\beta}$ es el modelo lineal.

$$(4.4.5) \quad \mathbf{X}\underline{\beta} + \underline{\varepsilon} = \underline{y} \quad \text{con } \underline{\varepsilon} \sim (0, \delta^2 \mathbf{I}),$$

siendo \mathbf{B} una matriz simétrica positiva definida^(*) dada, y en donde el parámetro de regularización γ es determinado mediante la resolución (con ZEROIN) de la ecuación de pseudo-optimalidad

$$\hat{\mathbf{V}}_C(\gamma) - \tau^2 \tilde{\mathbf{b}}_C^2(\gamma) = 0, \quad (12)$$

con $\hat{\mathbf{V}}_C(\gamma)$ y $\tilde{\mathbf{b}}_C^2(\gamma)$ dadas por (4.3.7) y (4.3.11) respectivamente.

Con fines comparativos el parámetro γ es también determinado mediante la minimización (con FMIN) de la estimación (sesgada) $\tilde{\mathbf{C}}(\gamma) = \hat{\mathbf{V}}_C(\gamma) + \tilde{\mathbf{b}}_C^2(\gamma)$ para el $\text{ecm}_C(\hat{\underline{\beta}}_\gamma; \underline{\beta}, \delta^2)$.

ALGORITMO.

Datos de entrada:

- ax, bx** extremos del intervalo [a,b].
ay, by extremos del intervalo [c,d].
N es el número de "observaciones" $y_i = f(x_i) + \varepsilon_i$,
 igual al número de renglones de \mathbf{X} .
P es el número de nodos de cuadratura, igual al número de columnas de \mathbf{X} .

(*). Véase el pie de nota (5) en la sec. 4.2.

(12). En nuestro estudio experimental se tomó $\tau^2 = 27/32$, y se trabajó con $\mathbf{C} = \mathbf{B}$ y con $\mathbf{C} = \mathbf{X}'\mathbf{X}$.

- NPROB** igual a 1,2,...,11; corresponde a uno de los 11 problemas que se estudiaron experimentalmente, los cuales se describen en la sección siguiente; con **NPROB** = 0, se sale del programa.
- NKER** igual 0 para cuadratura de Simpson, 2 para Trapecio y 3 para Punto Medio.
- ORDER** es el valor (opcional) de $m=0,1 \delta 2$ en (4.4.2), es el orden de la regularización.
- SGMA** es el valor de δ en (4.4.5), la desviación estándar con la cual se generan los números aleatorios ϵ_i , $i=1, \dots, n$; para la construcción del vector de "observaciones" y ($\delta = 10^{-5}$ en nuestro estudio)

FASE I

Se generan las matrices \mathbf{X} de $n \times p$, \mathbf{B} de $p \times p$; y el vector de "observaciones" \mathbf{y} :

- a). $\mathbf{X}_{n \times p} = (k(x_i, t_j) \sigma_j)$; con base a la "function" $f_{ker}(x_i, t_j)$ para $k(x, t)$ y con las σ_j según **NKER**.
- b). $\mathbf{B} \leftarrow \mathbf{I}$ si **ORDER** = 0; $\mathbf{B} \leftarrow \mathbf{B}^{[2m]} + \alpha \mathbf{I}$, si **ORDER** (=m) > 0; con base a la matriz $\mathbf{B}^{[2m]}$ (tridiagonal si **ORDER** = 1, y pentadiagonal si **ORDER** = 2) de diferencias de orden $2m$ descrita en (4.4.3).

- c). $\underset{N \times 1}{Y} + (y_1, \dots, y_M)'$; $y_i = f(x_i) + \epsilon_i$ con base a la "function" $FF(X)$ para $f(x)$ y la generación de M números aleatorios (con base en URAND) $\epsilon_i \sim N(0, \delta^2)$.

FASE II

- 1). Se genera el modelo lineal Θ ($T \underline{Q} + \underline{\epsilon} = \underline{Y}$):

Si $ORDER > 0$:

- a). se calcula con DPBFA-LINPACK la descomposición de Cholesky de B :

$$B = R'R \text{ con } R \text{ triangular superior}$$

- b). se genera a la matrix T .

$$T = X R^{-1}$$

Si $ORDER = 0$:

$$T = X$$

- 2) Se calcula con DQRDC-LINPACK la descomposición QR de T :

$$\underset{N \times P}{T} = \underset{N \times N}{Q} \cdot \left[\begin{array}{c} \tilde{R} \\ \cdot \\ \cdot \\ 0 \end{array} \right] \underset{P}{P}$$

Q ortogonal $N \times N$

\tilde{R} triangular superior $P \times P$

3). Con DQRSL-LINPACK se calcula:

a). $QTY = Q' y$

b). $\hat{Q} = \tilde{R}^{-1} \underline{c}_1$, con \underline{c}_1 siendo el vector cuyas componentes son las primeras p componentes de QTY, y

c). $RDS = y - T\hat{Q}$ el vector residual para \hat{Q} .

4). Se calculan, para el modelo $X\beta + \varepsilon = y$ con $\varepsilon \sim N(0, \delta^2 I)$, las estimaciones insesgadas $\hat{\beta}$ y $\hat{\delta}^2$ para β y δ^2 :

a). $\hat{\beta} = R^{-1} \hat{Q}$, y

b). $\hat{\delta}^2 = \|RDS\|_2^2 / (N-P)$

5). Se calcula con DSVDC-LINPACK la descomposición en valores singulares de \tilde{R} :

$$\tilde{R} = \tilde{U} \Delta V'$$

$$\Delta = \text{diag}(s_1, \dots, s_p),$$

\tilde{U} y V ortogonales

6). Se calcula:

$$\underline{d} = \tilde{U}' \underline{c}_1 \text{ con } \underline{c}_1 \text{ como en el paso 3.b) anterior,}$$

7). Se calcula $\gamma = \gamma_{SO}$ resolviendo con ZERON la ecuación de pseudo-optimalidad:

$$\hat{V}_C(\gamma) - \tau^2 \tilde{B}_C^2(\gamma) = 0, \quad \text{con } \tau^2 = \frac{27}{32},$$

en donde

$$\hat{V}_C(\gamma) = \hat{\sigma}^2 \frac{P}{1-\gamma}, w_i \frac{s_i^2}{(s_i^2 + \gamma)}, \text{ y } \hat{B}_C^2 = \frac{P}{1-\gamma}, w_i \left(\frac{d_i}{s_i}\right)^2 \left(\frac{\gamma}{s_i^2 + \gamma}\right)^2, \quad (13)$$

y con fines comparativos también se calcula α y $\gamma = \gamma_{op}$, hallando con FMN el punto de mínimo de la estimación sesgada.

$$\tilde{C}(\gamma) = \hat{V}_C(\gamma) + \hat{B}_C^2(\gamma)$$

para $ecm_C(\hat{\beta}_\gamma; \underline{\beta}, \delta^2)$.

8). Se calculan

a). $\hat{V}_C(0)$, la estimación insesgada para $ecm_C(\hat{\beta}; \underline{\beta}, \delta^2)$.

b). $\tilde{C}(\tilde{\gamma}_{so})$, la estimación sesgada (4.3.13) para

$ecm_C(\hat{\beta}_\gamma; \underline{\beta}, \delta^2)$ con $\gamma = \tilde{\gamma}_{so}$; y para comparación,

c). $\tilde{C}(\tilde{\gamma}_{op})$, la estimación sesgada (4.3.13) para $ecm_C(\hat{\beta}_\gamma; \underline{\beta}, \delta^2)$

con $\gamma = \tilde{\gamma}_{op}$.

9. Se calcula estimación GM regularizada $\hat{\beta}_\gamma$ con $\gamma = \tilde{\gamma}_{so}$ (y con $\gamma = \tilde{\gamma}_{op}$ para comparación).

a). $\underline{x} = (\gamma I + \Delta^2)^{-1} \Delta \underline{d}$

b). $\hat{\theta}_\gamma = V \underline{x}$

c). $\hat{\beta}_\gamma = R^{-1} \hat{\theta}_\gamma$

(13). Aquí, $w_i = 1$ (para toda i), si $C=B$; y $w_i = s_i^2$ ($i=1,2,\dots,p$) si $C=K'K$ (i.e. $G=T'F$).

Datos de Salida:

```

Problema kernel = ...., f = ...., u = ....
      (función kernel, lado derecho, solución exacta)
nprob = ... n = ... p = ... orden = ... nker = ...
sgma( $\delta$ ) = ...
sggor( $\delta$ ) = ...
valores singulares
... : ... : ... : ... : ...
ecm GM = ...
.....
resultados con función "... (*)
.....
problema kernel = ...., f = ...., u = ....
n prob = ... n = ... p = ... orden = ... discretización = "... (+)
GAM(SOP) = ... (#) ecmgm = ...

BETA EXACTA      BETA(SOP) (z)      ERROR BETA(SOP) (z)      BETGM
...              ...              ...              ...
...              ...              ...              ...
...              ...              ...              ...
.....

```

En la sección siguiente se presentan algunas salidas representativas de las 105 que se realizaron.

-
- (*) "ecm - sesgada" para $\gamma = \tilde{\gamma}_{op}$ con $\tilde{C}(\gamma)$ (véase (4.3.13)) estimación sesgada para ecm_C con $C=B$; "ecm-sesgada-pesada" para $\gamma = \tilde{\gamma}_{op}$ con $\tilde{C}(\gamma)$ estimación sesgada para ecm_C con $C=K'K$; "fso-p-sesgada" para $\gamma = \tilde{\gamma}_{so}$ solución de la ecuación de pseudo-optimalidad (4.3.12) con $C=B$; y "fso-p-sesgada-pesada" para $\gamma = \tilde{\gamma}_{so}$ solución de (4.3.12) con $C=K'K$.
- (+). "simspon", "trapecio", o "punto medio" según NKER.
- (#). O GAM(OP) según función "... en (*).
- (z). O BETA(OP) según función "... en (*).

§4.5. Ejemplos y Resultados Numéricos.

En esta sección se describen brevemente, destacando sus principales características, los *diecisiete* ejemplos de problemas *Mal-plantados* que se trabajaran experimentalmente y - que bien podemos clasificar desde un punto de vista numérico en tres clases o categorías: *blanda, moderada y dura*.

Diremos que el problema

$$(4.5.1) \quad \int_a^b k(x,t)u(t)dt = f(x), \quad c \leq x \leq d,$$

es uno *blando* si su función núcleo $k(x,t)$ es poco suave, por ejemplo, continua y derivable a tramos; como ocurre con el - problema del cálculo numérico de la segunda derivada (Ejemplo 10 de la lista que viene después), en el cual $k(x,t)$ es la - función de Green del problema

$$-y'' = g(x) \quad \text{con} \quad y(0) = y(1) = 0.$$

De esta categoría es también el clásico ejemplo de Phillip (Ejemplo 9).

Diremos que la ecuación integral (4.5.1) es un problema *moderado* si su función núcleo $k(x,t)$ es algebraico y de clase (esencialmente) C^∞ . Entre los ejemplos típicos de esta categoría tenemos a los siguientes

$$k(x,t) = 1/(x+t+k), \quad k > 0 \quad (\text{Ejemplo 7});$$

$$k(x,t) = \sqrt{x^2 + t^2} \quad (\text{Ejemplo 8}); \text{ y}$$

$$k(x,t) = \frac{1}{\sqrt{(x-t)^2 + 1}} \quad (\text{Ejemplo 11})$$

Y diremos que el problema integral (4.5.1) es uno *dato* si su función núcleo $k(x,t)$ es analítica. Como ejemplos típicos están los siguientes

$$k(x,t) = \cos xt \quad (\text{Ejemplo 4}); \text{ y}$$

$$k(x,t) = e^{xt} \quad (\text{Ejemplo 6});$$

los cuales corresponden a las transformaciones inversas "modificadas" coseno y de Laplace, respectivamente.

Por otro lado, bajo ciertas condiciones, Smithies [83'] demuestra que si $\partial_x^r k(x,t)$ es continua a trozos y acotada para alguna $r > 0$, entonces

$$(4.5.2) \quad s_1 = O(i^{-r-\beta}), \quad \beta > 0, \text{ cuando } i \rightarrow \infty.$$

Esto es, a mayor suavidad de la función núcleo $k(x,t)$ de la ecuación integral (4.5.1), se tiene un orden mayor de decaimiento a cero de sus valores singulares s_1 cuando $i \rightarrow \infty$.

Luego es de esperar, si $k(x,t)$ es analítico, que

$$(4.5.3) \quad s_1 = O(a^{-i}), \quad 0 < a < 1,$$

como efectivamente ocurre con el núcleo de Poisson (Sec. 1.2, ejemplo 3)

$$k(x,t) = \frac{R^2 - \rho^2}{2\pi(R^2 + \rho^2 - 2R\rho \cos(t-x))} \\ = \frac{1}{2\pi} \left[1 + 2 \sum_{q=1}^{\infty} \frac{\rho^q}{R^q} \cos q(t-x) \right], \quad a = \rho/R.$$

cuyos valores singulares son $1/2\pi$ y a^q/π .

En estos últimos términos podemos decir que la ecuación integral (4.5.1) es un problema *blando* si sus valores singulares siguen un patrón del tipo (4.5.2) con r un entero positivo "pequeño" (véase el Ejemplo 1), o bien si siguen un patrón del tipo (4.5.3) con (digamos) $a \geq 1/2$ (véase el Ejemplo 2); que es un problema *moderado* si sus valores singulares siguen un patrón del tipo (4.5.3) con (digamos) $1/2 > a > 1/10$; y finalmente, que es un problema *duro* si sus valores singulares siguen un patrón del tipo (4.5.3) con (digamos) $0 < a \leq 1/10$. Así pues, la ecuación (4.5.1) con función núcleo $k(x,t)$ analítico, puede ir de *blando* a *duro* pasando por *moderado*, el problema de extensión armónica es el ejemplo típico (Sec. 1.2., ejemplo 3).

Es ahora oportuno presentar los ejemplos con base a los cuales se realizó nuestro primer estudio experimental del Principio de *Seudo-optimalidad*.

Ejemplo 1 (Jennings et. al. [48]). Este ejemplo consta de cuatro en sí, todos ellos de la categoría de los problemas *blandos*, el cual consiste en generar el modelo lineal

$$\mathbb{X}\underline{\beta} + \underline{\varepsilon} = \underline{y} \quad \text{con} \quad \underline{\varepsilon} \sim N(\underline{0}, \delta^2 \mathbf{I}),$$

de la manera siguiente

- i). Se genera aleatoriamente a la matriz \mathbb{X} de $n \times p$ con valores singulares

$$s_i = i^{-\ell}, \quad i = 1, \dots, p; \quad \ell(\text{opcional}) = 2, 3, 4, 6 \text{ ó } 5. \quad \gamma$$

- ii). Se genera al vector de observaciones

$$\underline{y} = \underline{b} + \underline{\varepsilon} \quad \text{con} \quad \underline{\varepsilon} \sim N(\underline{0}, \delta^2 \mathbb{I})$$

en donde $\underline{b} = \mathbb{X}\underline{\beta}$ con $\underline{\beta} = (1, 2, \dots, p)'$. θ

Ejemplo 2 (Jennings et. al. [48]). Este ejemplo al igual que el antecedente, consta de cuatro en sí, que van de *blando* a *duro*, el cual consiste en generar el modelo lineal

$$\mathbb{X}\underline{\beta} + \underline{\varepsilon} = \underline{y} \quad \text{con} \quad \underline{\varepsilon} \sim N(\underline{0}, \delta^2 \mathbb{I})$$

como se dijo en el ejemplo anterior con la diferencia de que ahora los valores singulares de \mathbb{X} son

$$s_i = a^{-i}, \quad i = 1, \dots, p; \quad a(\text{opcional}) = 2, 5, 7 \text{ ó } 10.$$

($a=2$, *blando*; $a=5$ ó 7 , *moderado*; y $a=10$, *duro*). θ

Ejemplo 3 (Delves L.M. et. al. [105]). Este ejemplo es de los *moderados*, está dado por

$$\int_0^1 k(x,t)u(t) dt = \text{sen } x, \quad 0 < x < 1,$$

con

$$k(x,t) = \begin{cases} \cos(x-t), & t < x \\ 0, & t > x, \end{cases}$$

y cuya solución $u(t) = 1$ satisface las condiciones de frontera naturales de la regularización de primero y segundo orden. #

Ejemplo 4 (Varah J.M. [97]). Este es uno de los ejemplos de la clase de los *datos* que corresponde a la transformada inversa coseno (modificada) y que está dado por

$$\int_0^1 \cos xt u(t) dt = \sin x/x + (\cos x - 1)/x^2, \quad 0 < x < 1,$$

con solución $u(t) = t$, la cual cumple con las condiciones de frontera naturales de la regularización de segundo orden. #

Ejemplo 5 (Bellman R. et. al. [103]). Este ejemplo tiene función núcleo degenerada (véase pie de nota (7) en la pág. 22). por ello es de esperar que en su discretización \mathbb{X} resulte generalmente singular, y está dado por

$$\int_0^1 (x-t)^2 u(t) dt = x^2/2 - 2x/3 + 1/4, \quad 0 < x < 1,$$

con solución $u(t) = t$. #

Ejemplo 6 (Lewis B.A. [107]). Este es otro de los ejemplos de la categoría de los *datos* que corresponde a la transformada (modificada) inversa de Laplace, el cual está dado por

$$\int_0^1 e^{xt} u(t) dt = (e^{x+1} - 1)/(x+1), \quad 0 < x < 1,$$

y cuya solución $u(t) = e^t$ no cumple con las condiciones de frontera naturales de ningún orden de regularización. #

Ejemplo 7. Este ejemplo de la categoría de los *modelados* es de nueva creación y está dado por

$$\int_0^1 (x+t+0.1)^{-1} u(t) dt = f(x), \quad 0 \leq x \leq 2,$$

con

$$f(x) = 5 - 10x + x(10x+1) \ln \left(\frac{10x+1}{10x+1} \right),$$

y cuya solución $u(t) = t(10t+1)$ no cumple las condiciones de frontera naturales de los primeros tres órdenes de regularización. En general, la ecuación integral

$$\int_0^d (x+t+k)^{-1} u(t) dt = f(x), \quad 0 \leq x \leq d,$$

con $d = (n-1)h$, en donde $h = 1/(p-1)$ es el paso de las particiones uniformes para los intervalos $[0,1]$ y $[0,d]$ de p y n nodos respectivamente, da lugar al modelo lineal $\mathbb{E}\underline{u} + \underline{c} = \underline{y}$ con $\mathbb{E} = H \cdot S$ siendo $S = \text{diag}(\sigma_1, \dots, \sigma_p)$ la matriz de pesos de la cuadratura empleada y H la matriz de Hilbert rectangular de $n \times p$. \circ

Ejemplo 8 (Fox - Goodwin [29]). Este ejemplo es otro de la clase de los *modelados*, el cual está dado por

$$\int_0^1 \sqrt{x^2+t^2} u(t) dt = \frac{1}{3} \left[(1+x^2)^{3/2} - x^3 \right],$$

y cuya solución es $u(t) = t$. \circ

Ejemplo 9 (Phillips D.L. [74]). Este ejemplo es uno de los clásicos que sorprendentemente pertenece a la clase de los

problemas blandos (véase la observación 4.3.), el cual está dado por

$$\int_{-6}^6 k(x,t) u(t) dt = f(x), \quad -6 \leq x \leq 6,$$

donde

$$f(x) = (6 - |x|)(1 + \cos(\pi x/3))/2 \\ + \operatorname{sgn}(x) \cdot \frac{9}{2\pi} \sin(\pi x/3), \quad y$$

$$k(x,t) = \begin{cases} 1 + \cos(\pi(x-t)/3), & |x-t| \leq 3 \\ 0, & |x-t| > 3, \end{cases}$$

con solución dada por

$$u(x) = k(0,t) = \begin{cases} 1 + \cos(\pi t/3), & |t| \leq 3 \\ 0, & |t| > 3. \quad \text{§} \end{cases}$$

Ejemplo 10 (Delvas L.M. et. al. [105]). Este ejemplo de la categoría de los blandos corresponde al cálculo numérico de la segunda derivada:

$$\int_0^1 k(x,t) u(t) dt = (x^2 - x)/6, \quad 0 \leq x \leq 1,$$

en donde la función núcleo

$$k(x,t) = \begin{cases} x(t-1), & x \leq t \\ t(x-1), & t \leq x, \end{cases}$$

es la función de Green del problema: $-y'' = g(x)$ con $y(0)=y(1) = 0$. Y cuya solución está dada por $u(t) = t$. §

Ejemplo 11 (Glasko-Tijonov [34]). Este ejemplo de la clase de los *moderados* es clásico en la literatura soviética y está dado por

$$\int_{-1}^1 (\pi((x-t)^2 + 1))^{-1} u(t) dt = f(x), \quad -1 \leq x \leq 1,$$

donde, con solución $u(t) = (t^2 - 1)^2$ dada, $f(x)$ es generada por cuadratura numérica sobre una partición del intervalo $[-1, 1]$ - (en nuestro caso se usó la rutina ADAPT [102']).

En el cuadro 4.5.1 que viene a continuación se dan los porcentajes comparativos de las salidas con resultados aceptables^(*) de un total de 32 corridas por cada criterio de elección del parámetro de regularización (Seudo-optimalidad y *Optimalidad* con base a la función de pérdida $L_c(\hat{\beta}) = (\hat{\beta} - \beta)'C(\hat{\beta} - \beta)$ con $C = B$ y $C = X'X$), y por cada orden de regularización de cero a dos. Se excluyeron 3 corridas en las cuales el proceso se abortó por detectarse singularidad en la descomposición QR o de valores singulares de la matriz $T = XR^{-1}$ del modelo 6.

Cuadro 4.5.1

ORDEN	Y_{so} vs Y_{op} con $C=B$	Y_{so} vs Y_{op} con $C = X'X$
0	21.88 vs 25.08	34.38 vs 28.18
1	21.88 vs 37.58	56.28 vs 65.68
2	21.88 vs 56.28	43.78 vs 65.68

(*) Consideramos como resultados aceptables a todos aquellos que en promedio no dan menos de una cifra exacta.

Observación 4.5.1. En términos generales se observó que los mejores resultados se obtuvieron con el Principio de Optimalidad con respecto a la estimación (seagada) $\tilde{C}(\gamma)$ dada por (4.3.13) para $ocm_c(\gamma)$ con $C = \mathbb{H}'\mathbb{H}$. #

Observación 4.5.2. Para los ejemplos de la clase de los problemas blandos o moderados el Principio de Seudo-optimalidad, con respecto a la ecuación de seudo-optimalidad (4.3.12) con $C = \mathbb{H}'\mathbb{H}$, dió en general buenos resultados (véase las tablas y las figuras que están al final de esta sección). Pero para los ejemplos de la clase de los duros y el ejemplo 5 (de función núcleo degenerada), dicho principio no trabajó en general nada bien (véase la tabla 4.5.3). #

Observación 4.5.3. Con relación a las cuadraturas numéricas que se ensayaron (Trapezio, Simpson y Punto Medio) en la discretización de la ecuación integral de Fredholm (4.5.1), la cuadratura del Trapezio fue la que en general dió los mejores resultados superando claramente a la de Simpson, la cual tiene un orden de truncamiento superior. Al parecer este sorprendente hecho es ya del conocimiento de otros autores. #

Observación 4.5.4. En los ejemplos de la clase de los blandos la resolución numérica dada por la estimación GM clásica resultó generalmente muy aceptable. En tal circunstancia, las resoluciones regularizadas esencialmente coincidieron con la proporcionada por la estimación GM clásica como bien puede apreciarse de las Tablas 4.5.1, 4.5.5, y 4.5.6. La tabla 4.5.5 corresponde al clásico ejemplo de Phillips. #

Info 0

```

.....
PROBLEM
.....
problem osborna - type 0 1
nprob= 11 2 n= 77 p= 15 order= 2 nker= 0
OSMA = 1.000000000000000E+05
MATSCN = 9.279819762989741E-06
Cond(T) = .51111887E+11
diag. values
.14896188E+05 : .33333832E+04 : .62150119E-03 : .16622089E-03 :
.11666118E-03 : .16947017E-04 : .10075790E-04 : .70433649E-05 :
.38463148E-05 : .36640566E-05 : .21390695E-05 : .13649216E-05 :
.11172777E-05 : .88128761E-06 : .68274507E-06 : .00000000E+00 :
.00000000E+00 : .00000000E+00 : .00000000E+00 : .00000000E+00 :
GM-mse = 3.93402886390686
.....

```

```

.....
results with function mse-biased
.....
OPCAM= 11613.99928302644
RESULTS
EXACT BETA          OPBETA          OPBETA ERROR      GMBET
.10000000E+01      .10080810E+01      .80809666E-02      .99222359E+00
.20000000E+01      .20067925E+01      .33962650E-02      .20056096E+01
.30000000E+01      .30055041E+01      .18346978E-02      .29872353E+01
.40000000E+01      .40042157E+01      .10539142E-02      .40100540E+01
.50000000E+01      .50029272E+01      .58544403E-03      .49949918E+01
.60000000E+01      .60016388E+01      .27313059E-03      .60065411E+01
.70000000E+01      .70003503E+01      .50049559E-04      .69935301E+01
.80000000E+01      .79990619E+01      -.11726121E-05      .79979901E+01
.90000000E+01      .89977735E+01      -.24739181E-03      .90149616E+01
.10000000E+02      .99964850E+01      -.35149629E-03      .10004824E+02
.11000000E+02      .10993197E+02      -.43667268E-03      .11004097E+02
.12000000E+02      .11993908E+02      -.50765301E-03      .12009150E+02
.13000000E+02      .12992620E+02      -.56771329E-03      .12979688E+02
.14000000E+02      .13991331E+02      -.61919353E-03      .13974772E+02
.15000000E+02      .14990043E+02      -.66380973E-03      .15001167E+02
.....

```

```

.....
results with function mse-biased-weighted= mse0
.....
OPCAM= 2.546301066249958E-08
RESULTS
EXACT BETA          OPBETA          OPBETA ERROR      GMBET
.10000000E+01      .99951057E+00      -.48943177E-03      .99222359E+00
.20000000E+01      .19995419E+01      -.21905616E-03      .20056096E+01
.30000000E+01      .29996449E+01      -.11837872E-03      .29872353E+01
.40000000E+01      .39997347E+01      -.64328005E-04      .40100540E+01
.50000000E+01      .49997994E+01      -.40126875E-04      .49949918E+01
.60000000E+01      .59998629E+01      -.22853194E-04      .60065411E+01
.70000000E+01      .69999611E+01      -.55543900E-05      .69935301E+01
.80000000E+01      .80000794E+01      .99275395E-05      .79979901E+01
.90000000E+01      .90001593E+01      .17703497E-04      .90149616E+01
.10000000E+02      .10000184E+02      .16367862E-04      .10004824E+02
.11000000E+02      .11000082E+02      .74601382E-05      .11004097E+02
.12000000E+02      .11999882E+02      -.92955173E-05      .12009150E+02
.13000000E+02      .12999578E+02      -.32478338E-04      .12979688E+02
.14000000E+02      .13999185E+02      -.58204282E-04      .13974772E+02
.....

```

```

.....
.15000000E+02 : .14998755E+02 : -.83007514E-04 : .15001167E+02 :
.....
results with function faop-biased= faop
.....
SOPCAM = 1.000000000000000E-25
RESULTS
EXACT BETA          SOPBETA          SOPBETA ERROR      GMBET
.10000000E+01      .99222359E+00      -.77764115E-02      .99222359E+00
.20000000E+01      .20056096E+01      .28048145E-02      .20056096E+01
.30000000E+01      .29872353E+01      -.42549000E-02      .29872353E+01
.40000000E+01      .40100540E+01      .25134966E-02      .40100540E+01
.50000000E+01      .49949918E+01      -.10016434E-02      .49949918E+01
.60000000E+01      .60065411E+01      .10901862E-02      .60065411E+01
.70000000E+01      .69935301E+01      -.92426863E-03      .69935301E+01
.80000000E+01      .79979901E+01      -.25124061E-03      .79979901E+01
.90000000E+01      .90149616E+01      .16623945E-02      .90149616E+01
.10000000E+02      .10004824E+02      .48240393E-03      .10004824E+02
.11000000E+02      .11004097E+02      .37248995E-03      .11004097E+02
.12000000E+02      .12009150E+02      .76253466E-03      .12009150E+02
.13000000E+02      .12979688E+02      -.15624667E-02      .12979688E+02
.14000000E+02      .13974772E+02      -.18020159E-02      .13974772E+02
.15000000E+02      .15001167E+02      .77800704E-04      .15001167E+02
.....

```

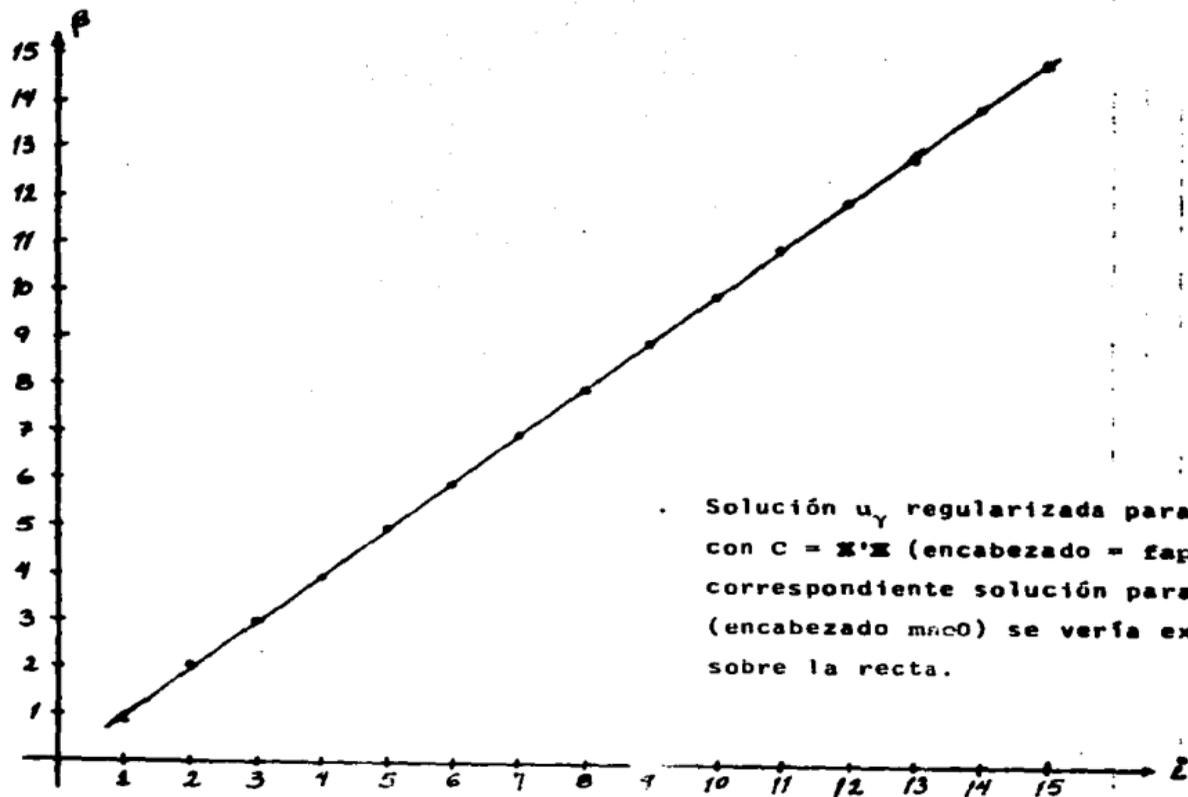
```

.....
results with function faop-biased-weighted= faop0
.....
SOPCAM = 6.280098396185610E-10
RESULTS
EXACT BETA          SOPBETA          SOPBETA ERROR      GMBET
.10000000E+01      .99877294E+00      -.12270605E-02      .99222359E+00
.20000000E+01      .19989076E+01      -.44619652E-03      .20056096E+01
.30000000E+01      .29993633E+01      -.21223194E-03      .29872353E+01
.40000000E+01      .39999733E+01      -.66658474E-05      .40100540E+01
.50000000E+01      .50001087E+01      .21734703E-04      .49949918E+01
.60000000E+01      .59999681E+01      -.53126216E-05      .60065411E+01
.70000000E+01      .70000116E+01      .16629761E-05      .69935301E+01
.80000000E+01      .80004335E+01      .54192527E-04      .79979901E+01
.90000000E+01      .90006415E+01      .71278681E-04      .90149616E+01
.10000000E+02      .10000547E+02      .54736722E-04      .10004824E+02
.11000000E+02      .11000097E+02      .88078682E-05      .11004097E+02
.12000000E+02      .11999327E+02      -.56120879E-04      .12009150E+02
.13000000E+02      .12998479E+02      -.11696440E-03      .12979688E+02
.14000000E+02      .13997770E+02      -.15931490E-03      .13974772E+02
.15000000E+02      .14997279E+02      -.18137447E-03      .15001167E+02
.....

```

S

TABLA 4.5.1. Ejemplo con valores singulares $\lambda_1 = 1^{-2}$. Los resultados para γ_{80} con $C = M^2 X$ (encabezado osm0) son claramente superiores a la estimación GM clásica.



• Solución u_Y regularizada para $y = Y_{so}$ con $C = \Sigma' \Sigma$ (encabezado = $fap0$). La correspondiente solución para $y = Y_{op}$ (encabezado $mac0$) se vería exactamente sobre la recta.

Figura 6. Ilustración gráfica de la Tabla 4.5.1.

PROBLEM

problem kernel= cos(x-y), f=sin(x), u = 1
order= 3 ill= 0 n= 17 p= 15 order= 2 nker= 2
SOPGA = 1.000000000000000E-03
SOPGM = 1.691254386638924E-02
Comdt = -1.4375872E+23

Table with 4 columns: exact values, OPBETA, OPBETA ERROR, GMSET. Rows show numerical results for various parameters.

results with function mae-biased

discretization = trapezoidal
OPGM= 40080.32640165517

Table with 4 columns: EXACT BETA, OPBETA, OPBETA ERROR, GMSET. Rows show numerical results for various parameters.

results with function mae-biased-weighted= mae0

discretization = trapezoidal
OPGM= 40080.32640165517

Table with 4 columns: EXACT BETA, OPBETA, OPBETA ERROR, GMSET. Rows show numerical results for various parameters.

Table with 4 columns: exact values, SOPBETA, SOPBETA ERROR, GMSET. Rows show numerical results for various parameters.

results with function fscop-biased= fap

discretization = trapezoidal
SOPGM = 1.000000000000000E-25

Table with 4 columns: EXACT BETA, SOPBETA, SOPBETA ERROR, GMSET. Rows show numerical results for various parameters.

results with function fscop-biased-weighted= fap0

discretization = trapezoidal
SOPGM = 1.664908569712216E-05

Table with 4 columns: EXACT BETA, SOPBETA, SOPBETA ERROR, GMSET. Rows show numerical results for various parameters.

TABLA 4.5.2. Este ejemplo no trabajó con ORDEN 0; sin embargo, los resultados obtenidos con ORDEN 2 son superiores a los de la estimación GM clásica.

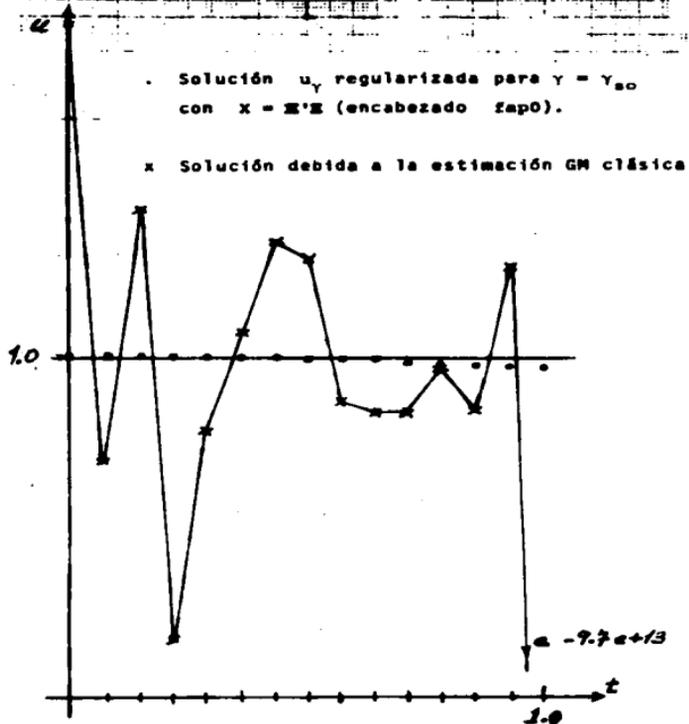


Figura 7. Ilustración Gráfica de la tabla 4.5.2.

```

.....
PROBLEM
.....
problem kernel= exp(x^2) f= (exp(1-x)-1)/(1+x), u = exp(y)
nprob= 6 nll= 0 n= 17 p= 15 order= 2 nker= 0
BGM= 1.000000000000000E-05
MATRGM = 9.322461078496397E-06
Cond(T) = .15495305E+25
sing. values
.97938937E+05 : .77540844E+04 : .19679197E-03 : .15999422E-05
-1.3381300E-07 : .10731504E-09 : .85454192E-12 : .66051228E-14
-6.014630E-16 : .56500298E-17 : .28952713E-17 : .19925099E-17
-6.401630E-18 : .19429700E-18 : .63205231E-19 : .00000000E+00
-0.0000000E+00 : 0.0000000E+00 : .00000000E+00 : .00000000E+00
GM=ms= 3023.658447265625
.....
Results with function ms=biased
.....
discretization = simpson
OPGAM= 97938.93519284059
RESULTS
EXACT BETA OPBETA OPBETA ERROR GMBET
1.0000000E+01 : .85809790E+00 : -.14190210E+00 : .10782982E+10
1.0740414E+01 : .98091995E+00 : -.8670194E-01 : -.26446306E+10
1.1535650E+01 : .11037420E+01 : -.43190455E-01 : -.23204366E+11
1.2389766E+01 : .12265640E+01 : -.10018391E-01 : -.29025328E+11
1.3307122E+01 : .13493861E+01 : -.14033009E-01 : -.39138311E+11
1.4292400E+01 : .14722081E+01 : -.30063606E-01 : -.42747474E+11
1.5350630E+01 : .15950302E+01 : -.39064963E-01 : -.50865039E+11
1.6487213E+01 : .17178522E+01 : -.41930050E-01 : -.58925048E+10
1.7707950E+01 : .18406743E+01 : -.39462125E-01 : -.51837077E+09
1.9019071E+01 : .19634963E+01 : -.32382857E-01 : -.42495322E+10
2.0427271E+01 : .20863184E+01 : -.21339762E-01 : -.27062531E+10
2.1939735E+01 : .22091404E+01 : -.69122924E-02 : -.81127498E+10
2.3564184E+01 : .23319625E+01 : -.10378448E-01 : -.19545775E+10
2.5308910E+01 : .24547845E+01 : -.30071035E-01 : -.13551313E+10
2.7182818E+01 : .25776066E+01 : -.51751535E-01 : .
.....
Results with function ms=biased-weighted= ms=0
.....
discretization = simpson
OPGAM= 4.513049914352759E-12
RESULTS
EXACT BETA OPBETA OPBETA ERROR GMBET
1.0000000E+01 : .10143904E+01 : -.14390428E-01 : .10782982E+10
1.0740414E+01 : .10859205E+01 : -.11060149E-01 : -.26446306E+10
1.1535650E+01 : .11584942E+01 : -.42729966E-02 : -.23204366E+11
1.2389766E+01 : .12352264E+01 : -.30268217E-02 : -.29025328E+11
1.3307122E+01 : .13194623E+01 : -.8450455E-02 : -.39138311E+11
1.4292400E+01 : .14138637E+01 : -.10758399E-01 : -.42747474E+11
1.5350630E+01 : .15203973E+01 : -.95538002E-02 : -.50865039E+11
1.6487213E+01 : .16393262E+01 : -.56988246E-02 : -.81127498E+10
1.7707950E+01 : .17700759E+01 : -.40605441E-03 : -.19545775E+10
1.9019071E+01 : .19107057E+01 : -.46262008E-02 : -.42495322E+10
2.0427271E+01 : .20588466E+01 : .78911962E-02 : -.45112515E+10
2.1939735E+01 : .22118942E+01 : -.81601256E-02 : -.27062531E+10

```

```

.....
.23564184E+01 : .23674386E+01 : .46766502E-02 : -.01127498E+10
-25308910E+01 : .25239623E+01 : -.27378652E-02 : -.19545775E+10
27182818E+01 : .28805115E+01 : .13894924E-01 : .13551313E+10
.....
Results with function feop-biased= feap
.....
discretization = simpson
SOPGAM = 1.000000000000000E-25
RESULTS
EXACT BETA SOPBETA SOPBETA ERROR GMBET
1.0000000E+01 : -.12538439E-05 : -.12539439E+05 : .10782982E+10
1.0740414E+01 : .45237599E+04 : .42109045E+04 : -.26446306E+10
1.1535650E+01 : .69045024E+04 : .59843605E+04 : -.23204366E+11
1.2389766E+01 : -.14621891E+04 : -.11811588E+04 : -.29025328E+11
1.3307122E+01 : -.85677330E+04 : -.64394568E+04 : -.39138311E+11
1.4292400E+01 : -.71422504E+04 : -.49982365E+04 : -.42747474E+11
1.5350630E+01 : .24561635E+04 : .15990408E+04 : -.50865039E+11
1.6487213E+01 : .99958350E+04 : .60617804E+04 : -.89250483E+10
1.7707950E+01 : .69855784E+04 : .39438827E+04 : -.51837077E+09
1.9019071E+01 : -.22580091E+04 : -.11882342E+04 : -.42495322E+10
2.0427271E+01 : .83591769E+04 : .40931654E+04 : -.45112515E+10
2.1939735E+01 : .53472816E+04 : .24382590E+04 : -.27062531E+10
2.3564184E+01 : .45292140E+04 : .19210372E+04 : -.01127498E+10
2.5308910E+01 : .57261001E+04 : .22614838E+04 : -.19545775E+10
2.7182818E+01 : -.11430167E+05 : -.42059234E+04 : .13551313E+10
.....
Results with function feop-biased-weighted= feap0
.....
discretization = simpson
SOPGAM = 1.000000000000000E-25
RESULTS
EXACT BETA SOPBETA SOPBETA ERROR GMBET
1.0000000E+01 : -.12538439E-05 : -.12539439E+05 : .10782982E+10
1.0740414E+01 : .45237599E+04 : .42109045E+04 : -.26446306E+10
1.1535650E+01 : .69045024E+04 : .59843605E+04 : -.23204366E+11
1.2389766E+01 : -.14621891E+04 : -.11811588E+04 : -.29025328E+11
1.3307122E+01 : -.85677330E+04 : -.64394568E+04 : -.39138311E+11
1.4292400E+01 : -.71422504E+04 : -.49982365E+04 : -.42747474E+11
1.5350630E+01 : .24561635E+04 : .15990408E+04 : -.50865039E+11
1.6487213E+01 : .99958350E+04 : .60617804E+04 : -.89250483E+10
1.7707950E+01 : .69855784E+04 : .39438827E+04 : -.51837077E+09
1.9019071E+01 : -.22580091E+04 : -.11882342E+04 : -.42495322E+10
2.0427271E+01 : .83591769E+04 : .40931654E+04 : -.45112515E+10
2.1939735E+01 : .53472816E+04 : .24382590E+04 : -.27062531E+10
2.3564184E+01 : .45292140E+04 : .19210372E+04 : -.01127498E+10
2.5308910E+01 : .57261001E+04 : .22614838E+04 : -.19545775E+10
2.7182818E+01 : -.11430167E+05 : -.42059234E+04 : .13551313E+10
.....

```

5 TABLA 4.5.3. Para este caso γ_{LU} no dá buenos resultados. Siendo este ejemplo de los duros, γ_{UP} con $C = 2 \times 10^4$ dá resultados muy aceptables. Nótese que $\|B_{LU}\| = 10^{+10}$.

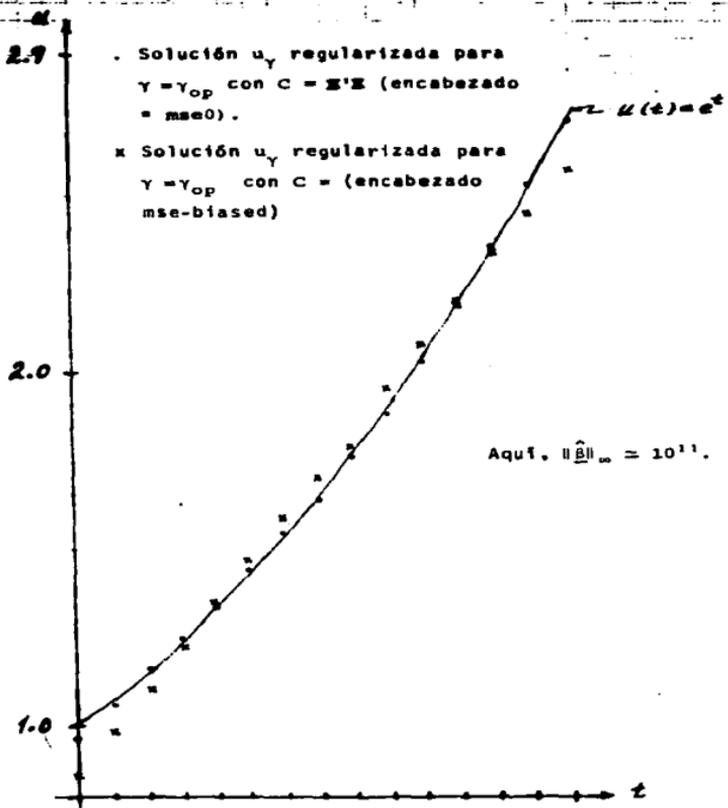


Figura 8. Ilustración gráfica de la tabla 4.5.3.

info = 0

```

.....
PROBLEM
problem kernel=sqrt(x)*y**2 -(10*x**2)**(3/2)-x**3/3, u = y
nprob= 8 ill= 0 n= 17 p= 15 order= 2 niter= 2
SOMA = 1.000000000000000E-05
MATHCON = 1.238100000000000E-05
Cond(T) = .11700710E+23
sing values
.50007290E+05 : .70203726E+04 : .35247563E-03 : .21274402E-04 :
.26711161E-05 : .40613271E-06 : .55268107E-07 : .64304833E-08 :
.63577412E-09 : .5285853E-10 : .3625814E-11 : .19997229E-12 :
.84397676E-14 : .25081523E-15 : .50000000E-17 : .00000000E+00 :
.00000000E+00 : .00000000E+00 : .00000000E+00 : .00000000E+00 :
GM-mse = 2372.0869140625

```

```

.....
results with function mse-biased
discretization = trapezoidal
OPCGM= 50002.2006322727
RESULTS
EXACT BETA      OPBETA      OPBETA ERROR      GMBET
.00000000E+00 : .49929981E-02 : .49929981E-02 : .16962474E-01 :
.71428571E-01 : .75693696E-01 : .42651250E-02 : -.12744587E+02 :
.14285714E+00 : .14639439E+00 : .35372519E-02 : -.34684249E+03 :
.21428571E+00 : .21709509E+00 : .24093788E-02 : .14899889E+05 :
.28571429E+00 : .28779579E+00 : .20815056E-02 : -.23950294E+06 :
.35714286E+00 : .35849649E+00 : .13536325E-02 : .22252528E+07 :
.42857143E+00 : .42919719E+00 : .62575940E-03 : -.13405032E+08 :
.50000000E+00 : .49989789E+00 : -.10211373E-03 : .55077249E+09 :
.57142857E+00 : .57059868E+00 : -.02998665E-03 : -.15811990E+10 :
.64285714E+00 : .64129928E+00 : -.15578600E-02 : .31966588E+09 :
.71428571E+00 : .71199998E+00 : -.22857331E-02 : -.45216429E+09 :
.78571429E+00 : .78270068E+00 : -.30136062E-02 : .43712196E+09 :
.85714286E+00 : .85140138E+00 : -.37814793E-02 : -.27467449E+09 :
.92857143E+00 : .92410208E+00 : -.44693525E-02 : .10093819E+09 :
.10000000E+01 : .99480277E+00 : -.51972256E-02 : -.32879668E+08 :

```

```

.....
results with function mse-biased-weighted- mse0
discretization = trapezoidal
OPCGM= 8.596997303415985E-11
RESULTS
EXACT BETA      OPBETA      OPBETA ERROR      GMBET
.00000000E+00 : .83227550E-02 : .83227550E-02 : .16962474E+01 :
.71428571E-01 : .74118214E-01 : .26696423E-02 : -.12744587E+02 :
.14285714E+00 : .14135004E+00 : -.15071046E-02 : -.34684249E+03 :
.21428571E+00 : .21110255E+00 : -.31831663E-02 : .14899889E+05 :
.28571429E+00 : .28339722E+00 : -.23170635E-02 : -.23950294E+06 :
.35714286E+00 : .35745011E+00 : .30725042E-03 : .22252528E+07 :
.42857143E+00 : .42211329E+00 : .35418830E-02 : -.13405032E+08 :
.50000000E+00 : .50624469E+00 : .62460869E-02 : .55077249E+09 :
.57142857E+00 : .57896236E+00 : .75337851E-02 : -.15811990E+10 :
.64285714E+00 : .64974759E+00 : .64904493E-02 : .31966588E+09 :
.71428571E+00 : .71840951E+00 : .42038949E-02 : -.45216429E+09 :
.78571429E+00 : .78882539E+00 : -.28889298E-03 : .43712196E+09 :

```

```

.85714286E+00 : .85103638E+00 : -.61064767E-02 : -.27467449E+09 :
.92857143E+00 : .91590324E+00 : -.12668192E-01 : -.10093819E+09 :
.10000000E+01 : .98053215E+00 : -.19467854E-01 : -.32879668E+08 :

```

```

.....
results with function feop-biased= feop
discretization = trapezoidal
SOPCGM= 1.000000000000000E-25
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.00000000E+00 : .62563692E-01 : .62563692E-01 : .16962474E+01 :
.71428571E-01 : .13263272E+01 : -.13977558E+01 : -.12744587E+02 :
.14285714E+00 : .28339722E+02 : .28196915E+02 : -.34684249E+03 :
.21428571E+00 : -.26320590E+03 : -.26327378E+03 : .14899889E+05 :
.28571429E+00 : .13023350E+04 : .13020492E+04 : -.23950294E+06 :
.35714286E+00 : -.35459151E+04 : -.35462722E+04 : .22252528E+07 :
.42857143E+00 : .49009724E+04 : .49005438E+04 : -.13405032E+08 :
.50000000E+00 : -.16759842E+04 : -.16764842E+04 : .55077249E+09 :
.57142857E+00 : -.35273518E+04 : -.35279232E+04 : -.15811990E+10 :
.64285714E+00 : .26199505E+04 : .26193077E+04 : .31966588E+09 :
.71428571E+00 : .25449286E+04 : .25442143E+04 : -.45216429E+09 :
.78571429E+00 : -.23840576E+04 : -.23848433E+04 : .43712196E+09 :
.85714286E+00 : -.15868396E+04 : -.15849496E+04 : -.27467449E+09 :
.92857143E+00 : .21906278E+04 : .21896992E+04 : .10093819E+09 :
.10000000E+01 : -.12277758E+04 : -.12287758E+04 : -.32879668E+08 :

```

```

.....
results with function feop-biased-weighted= feop0
discretization = trapezoidal
SOPCGM= 4.948622314040939E-10
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.00000000E+00 : .45953046E-02 : .45953046E-02 : .16962474E+01 :
.71428571E-01 : .73379416E-01 : .19508447E-02 : -.12744587E+02 :
.14285714E+00 : .14287101E+00 : .13868114E-04 : -.34684249E+03 :
.21428571E+00 : .21363603E+00 : -.64968127E-03 : .14899889E+05 :
.28571429E+00 : .28570620E+00 : .23950294E+06 : -.23950294E+06 :
.35714286E+00 : .35866710E+00 : .15242474E-02 : .22252528E+07 :
.42857143E+00 : .43187563E+00 : .33042039E-02 : -.13405032E+08 :
.50000000E+00 : .50466762E+00 : .46676209E-02 : .55077249E+09 :
.57142857E+00 : .57650936E+00 : .50807871E-02 : -.15811990E+10 :
.64285714E+00 : .64708421E+00 : .42270680E-02 : .31966588E+09 :
.71428571E+00 : .71631998E+00 : .20342622E-02 : -.45216429E+09 :
.78571429E+00 : .78346687E+00 : -.13474167E-02 : .43712196E+09 :
.85714286E+00 : .85153628E+00 : .65065775E-02 : -.27467449E+09 :
.92857143E+00 : .91820227E+00 : -.10362159E-01 : .10093819E+09 :
.10000000E+01 : .98472416E+00 : -.15277843E-01 : -.32879668E+08 :

```

TABLA 4.5.4. En este caso los resultados para Y_{00} y Y_{0p} con $C = \Xi'$ son comparables. Aquí, $K_1(\Xi) \approx 2.6e+14$ y $K_2(\Psi) \approx 1.3e+22$. Nótese que $\|\Xi\| = 10e+08$.

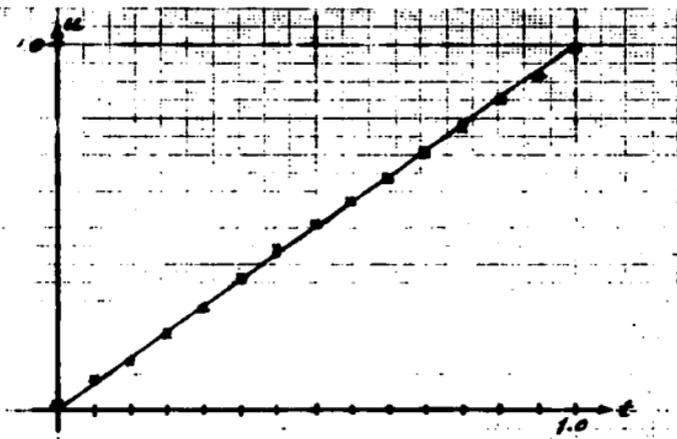


Figura 9. Ilustración gráfica de la tabla 4.5.4.

- x Solución u_Y regularizada para $Y = Y_{so}$ con $C = \Sigma^2 X$ (encabezado = $fap0$).
- Solución u_Y regularizada para $Y = Y_{op}$ con $C = B$ (encabezado mse -biased), la mejor solución en este caso.

```

.....
PROBLEM
problem kernel=leoco (pi * (y - x)/3),f=(6-w/3)*(...
nparm= 9 illi=0 n= 17 p= 13 order= 1 nker= 2
SOPCAM = 2.401773463992222E-05
SOPCAM = 2.401773463992222E-05
Cond(F) = .89534975E+06
sing. values
.13803830E+05 : .22352877E+02 : .98742027E+01 : .53385071E+01
.29752162E+01 : .16016789E+01 : .79298901E+00 : .33507845E+00
.08135111E-01 : .69654955E-01 : .67710511E-01 : .27044123E-01
.24035262E-01 : .18973285E-01 : .14520500E-01 : .00000000E+00
.00000000E-01 : .00000000E+00 : .00000000E+00 : .00000000E+00
GM=msc = .460031766891479
.....
Results with function msc-biased
discretization = trapezoidal
OPCAM= 1.676164433592303E-06
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.00000000E+00 : .54924431E-03 : .54924431E-03 : .54924431E-03
.00000000E+00 : -.87617924E-03 : -.87617924E-03 : -.87617924E-03
.00000000E+00 : .24090460E-02 : .24090460E-02 : .24090460E-02
.00000000E+00 : -.10071987E-01 : -.10071987E-01 : -.10071987E-01
.00000000E-01 : .12516202E+00 : .12516202E+00 : .12516202E+00
.77747907E+00 : .75796890E+00 : .75796890E+00 : .75796890E+00
.16234898E+01 : .15988353E+01 : .15186111E-01 : .75796890E+00
.20000000E+01 : .20527215E+01 : .26360762E-01 : .15987946E+01
.16234898E+01 : .15988353E+01 : .15299311E-01 : .15987946E+01
.77747907E+00 : .75804072E+00 : .19438344E-01 : .15987946E+01
.99031132E-01 : .12531774E+00 : .26286611E-01 : .12531774E+00
.00000000E+00 : -.10137459E-01 : -.10137459E-01 : -.10137459E-01
.00000000E+00 : .20965611E-02 : .20965611E-02 : .20921401E-02
.00000000E+00 : .36769796E-03 : .36769796E-03 : .37370172E-03
.00000000E+00 : -.23754442E-04 : -.23754442E-04 : -.12028870E-04
.....
Results with function msc-biased-weighted=msc0
discretization = trapezoidal
OPCAM= 6.621242514171248E-07
RESULTS
EXACT BETA      OPBETA      OPBETA ERROR      GMBET
.00000000E+00 : .55887838E-03 : .55887838E-03 : .55887838E-03
.00000000E+00 : -.88206936E-03 : -.88206936E-03 : -.88206936E-03
.00000000E+00 : .24078188E-02 : .24078188E-02 : .24078188E-02
.00000000E+00 : -.10061882E-01 : -.10061882E-01 : -.10061882E-01
.99031132E-01 : .12514735E+00 : .12514735E+00 : .12514735E+00
.16234898E+01 : .15988353E+00 : .15492059E-01 : .15987946E+01
.20000000E+01 : .20527507E+01 : .15201288E-01 : .20527507E+01
.16234898E+01 : .15988353E+01 : .26370343E-01 : .15987946E+01
.77747907E+00 : .75805952E+00 : .15315165E-01 : .75805952E+00
.99031132E-01 : .12520348E+00 : .19419549E-01 : .12520348E+00
.....

```

```

.00000000E+00 : -.10127497E-01 : -.10127497E-01 : -.10120988E-01
.00000000E+00 : .20928811E-02 : .20928811E-02 : .20921401E-02
.00000000E+00 : -.37132476E-03 : -.37132476E-03 : -.37370172E-03
.00000000E+00 : -.16646682E-04 : -.16646682E-04 : -.12828870E-04
.....
Results with function fscp-biased=fap
discretization = trapezoidal
SOPCAM= 2.401773463992222E-05
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.00000000E+00 : .34511052E-03 : .34511052E-03 : .34511052E-03
.00000000E+00 : -.75516597E-03 : -.75516597E-03 : -.88593557E-03
.00000000E+00 : .24432301E-02 : .24432301E-02 : .24078188E-02
.00000000E+00 : -.10293269E-01 : -.10293269E-01 : -.10058282E-01
.99031132E-01 : .12547422E+00 : .12547422E+00 : .12513774E+00
.77747907E+00 : .75758722E+00 : .19891849E-01 : .75796890E+00
.16234898E+01 : .15993571E+01 : .148664720E-01 : .15987946E+01
.20000000E+01 : .20521021E+01 : .26051025E-01 : .20527697E+01
.16234898E+01 : .15993571E+01 : .149648072E-01 : .15987946E+01
.77747907E+00 : .75764588E+00 : .19833484E-01 : .75807183E+00
.99031132E-01 : .12562138E+00 : .24650250E-01 : .12562138E+00
.00000000E+00 : -.10355551E-01 : -.10355551E-01 : -.10120988E-01
.00000000E+00 : .21594791E-02 : .21594791E-02 : .20921401E-02
.00000000E+00 : -.29173040E-03 : -.29173040E-03 : .37370172E-03
.00000000E+00 : -.17695579E-03 : -.17695579E-03 : -.12028870E-04
.....
Results with function fscp-biased-weighted=fap0
discretization = trapezoidal
SOPCAM= 3.2150687586655273E-05
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.00000000E+00 : .27466671E-03 : .27466671E-03 : .56518809E-03
.00000000E+00 : -.71485514E-03 : -.71485514E-03 : -.88593557E-03
.00000000E+00 : .24586520E-02 : .24586520E-02 : .24078188E-02
.00000000E+00 : -.10373069E-01 : -.10373069E-01 : -.10058282E-01
.99031132E-01 : .12558285E+00 : .12558285E+00 : .12513774E+00
.77747907E+00 : .75745605E+00 : .20223020E-01 : .75796890E+00
.16234898E+01 : .15995378E+01 : .14753381E-01 : .15987946E+01
.20000000E+01 : .20518866E+01 : .25943322E-01 : .20527697E+01
.16234898E+01 : .15993837E+01 : .14888325E-01 : .15988089E+01
.77747907E+00 : .75751020E+00 : .19968869E-01 : .75807183E+00
.99031132E-01 : .12572714E+00 : .26960065E-01 : .12528415E+00
.00000000E+00 : -.10434186E-01 : -.10434186E-01 : -.10120988E-01
.00000000E+00 : .21840111E-02 : .21840111E-02 : .20921401E-02
.00000000E+00 : -.26581967E-03 : -.26581967E-03 : .37370172E-03
.00000000E+00 : -.23134217E-03 : -.23134217E-03 : -.12028870E-04
.....

```

5

TABLA 4.5.5. Para este caso y_{00} , y_{01} y la estimación GM clásica dan esencialmente los mismos aceptables resultados. Aquí, $K_1(N) = 3.1e+02$ y $K_2(N) = 4.9e+05$.

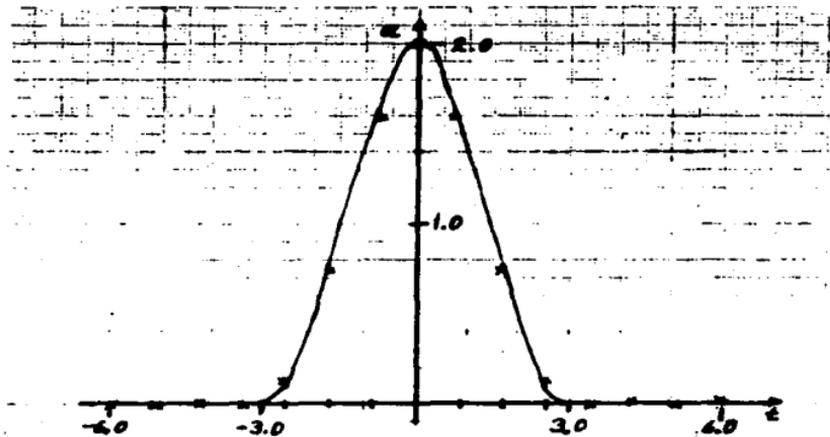


Figura 10. Ilustración gráfica de la tabla 4.5.5.

- x Solución u_y regularizada de orden 1 para $y = y_{20}$ con $C = B$ (encabezado = fap). La cual coincide en tres cifras con la resolución proporcionada por la estimación GM clásica.

Info- 8

```

.....
PROBLEM
problem name: y(1-x) or x(y-1), f=x(x-1)/6, u=y
nprob= 10  ill= 0  n= 17  m= 15  order= 2  niter= 3
SCMA = 1.000000000000000E-05
MATHGM = 3.24399999935580E-06
Cond(T) = .71586302E+10
sig. values
.63739162E+04 : .12797217E+04 : .40779000E-03 : .99167053E-04 :
.56631978E-04 : .27118435E-04 : .13248460E-04 : .71775917E-05 :
.43386073E-05 : .29188826E-05 : .21146069E-05 : .15177078E-05 :
.12360757E-05 : .10240569E-05 : .89078001E-06 : .00000000E+00 :
.00000000E+00 : .00000000E+00 : .00000000E+00 : .00000000E+00 :
GM-ans = 3.05597319874327

```

```

.....
results with function msa-biased
discretization = midpoint
SOPGM = 2.01375999343444E-15
RESULTS
EXACT BETA      OPBETA      OPBETA ERROR      GMBET
.35714286E-01 : .43125762E-01 : .74114767E-02 : .43141485E-01 :
.10714286E+00 : .10853793E+00 : .13950714E-02 : .10852175E+00 :
.17857143E+00 : .19546048E+00 : .16889055E-01 : .19547470E+00 :
.25000000E+00 : .21924164E+00 : -.30750399E-01 : .21924164E+00 :
.32142857E+00 : .31233466E+00 : -.90939106E-02 : .31233180E+00 :
.39285714E+00 : .40858732E+00 : .15710179E-01 : .40858399E+00 :
.46428571E+00 : .46033550E+00 : -.39502147E-02 : .46030738E+00 :
.53571429E+00 : .55612871E+00 : .15779552E-01 : .55612871E+00 :
.60714286E+00 : .62292241E+00 : .20381076E-01 : .62288812E+00 :
.67857143E+00 : .66748163E+00 : .18490200E-01 : .66749545E+00 :
.75000000E+00 : .66477609E+00 : -.65223915E-01 : .66473060E+00 :
.82142857E+00 : .82097861E+00 : .44995862E-03 : .82108012E+00 :
.89285714E+00 : .91413191E+00 : .21274765E-01 : .91392179E+00 :
.96428571E+00 : .97185968E+00 : .73739601E-02 : .97197920E+00 :
.10357143E+01 : .15399899E-02 : -.99851114E+00 : .15517515E-02 :

```

```

.....
results with function msa-biased-weighted= msa0
discretization = midpoint
SOPGM = 2.184111209557753E-15
RESULTS
EXACT BETA      OPBETA      OPBETA ERROR      GMBET
.35714286E-01 : .43124435E-01 : .74101491E-02 : .43141485E-01 :
.10714286E+00 : .10853930E+00 : .13964379E-02 : .10852175E+00 :
.17857143E+00 : .19549288E+00 : .16887854E-01 : .19547470E+00 :
.25000000E+00 : .21925027E+00 : -.30749726E-01 : .21924164E+00 :
.32142857E+00 : .31233490E+00 : -.90939707E-02 : .31233180E+00 :
.39285714E+00 : .40858592E+00 : .15708773E-01 : .40858399E+00 :
.46428571E+00 : .46033787E+00 : -.39478437E-02 : .46030738E+00 :
.53571429E+00 : .55609255E+00 : .15782465E-01 : .55612871E+00 :
.60714286E+00 : .62292530E+00 : .20378265E-01 : .62288812E+00 :
.67857143E+00 : .66749877E+00 : .18887346E-01 : .66749545E+00 :
.75000000E+00 : .66477992E+00 : -.85220077E-01 : .66473060E+00 :
.82142857E+00 : .82097055E+00 : .43852332E-03 : .82108012E+00 :

```

```

.....
.89285714E+00 : .91414955E+00 : .21292503E-01 : .91392179E+00 :
.96428571E+00 : .97185370E+00 : .73489898E-02 : .97197920E+00 :
.10357143E+01 : .15389632E-02 : -.9985140E+00 : .15517515E-02 :
.....
results with function fcap-biased= fap
discretization = midpoint
SOPGM = 1.000000000000000E-25
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.35714286E-01 : .43141482E-01 : .74271961E-02 : .43141485E-01 :
.10714286E+00 : .10852175E+00 : .13788954E-02 : .10852175E+00 :
.17857143E+00 : .19547470E+00 : .16903276E-01 : .19547470E+00 :
.25000000E+00 : .21924164E+00 : -.30758357E-01 : .21924164E+00 :
.32142857E+00 : .31233180E+00 : -.90967673E-02 : .31233180E+00 :
.39285714E+00 : .40858399E+00 : .15726851E-01 : .40858399E+00 :
.46428571E+00 : .46030738E+00 : -.39783319E-02 : .46030738E+00 :
.53571429E+00 : .55612871E+00 : .15745286E-01 : .55612871E+00 :
.60714286E+00 : .62288812E+00 : .20414420E-01 : .62288812E+00 :
.67857143E+00 : .66749545E+00 : .18924026E-01 : .66749545E+00 :
.75000000E+00 : .66473060E+00 : -.85269403E-01 : .66473060E+00 :
.82142857E+00 : .82108012E+00 : .34845480E-03 : .82108012E+00 :
.89285714E+00 : .91392179E+00 : .21064648E-01 : .91392179E+00 :
.96428571E+00 : .97197920E+00 : .76934884E-02 : .97197920E+00 :
.10357143E+01 : .15517515E-02 : -.99850176E+00 : .15517515E-02 :

```

```

.....
results with function fcap-biased-weighted= fap0
discretization = midpoint
SOPGM = 1.000000000000000E-25
RESULTS
EXACT BETA      SOPBETA      SOPBETA ERROR      GMBET
.35714286E-01 : .43141482E-01 : .74271961E-02 : .43141485E-01 :
.10714286E+00 : .10852175E+00 : .13788954E-02 : .10852175E+00 :
.17857143E+00 : .19547470E+00 : .16903276E-01 : .19547470E+00 :
.25000000E+00 : .21924164E+00 : -.30758357E-01 : .21924164E+00 :
.32142857E+00 : .31233180E+00 : -.90967673E-02 : .31233180E+00 :
.39285714E+00 : .40858399E+00 : .15726851E-01 : .40858399E+00 :
.46428571E+00 : .46030738E+00 : -.39783319E-02 : .46030738E+00 :
.53571429E+00 : .55612871E+00 : .15745286E-01 : .55612871E+00 :
.60714286E+00 : .62288812E+00 : .20414420E-01 : .62288812E+00 :
.67857143E+00 : .66749545E+00 : .18924026E-01 : .66749545E+00 :
.75000000E+00 : .66473060E+00 : -.85269403E-01 : .66473060E+00 :
.82142857E+00 : .82108012E+00 : .34845480E-03 : .82108012E+00 :
.89285714E+00 : .91392179E+00 : .21064648E-01 : .91392179E+00 :
.96428571E+00 : .97197920E+00 : .76934884E-02 : .97197920E+00 :
.10357143E+01 : .15517515E-02 : -.99850176E+00 : .15517515E-02 :

```

5

TABLA 4.5.6. Este ejemplo de los *bandas* sólo corrió con la cuadratura de Punto Medio. Los resultados obtenidos son los de la estimación GM clásica. Aquí, $K_1(\mathbb{R}) = 1.6e+02$.

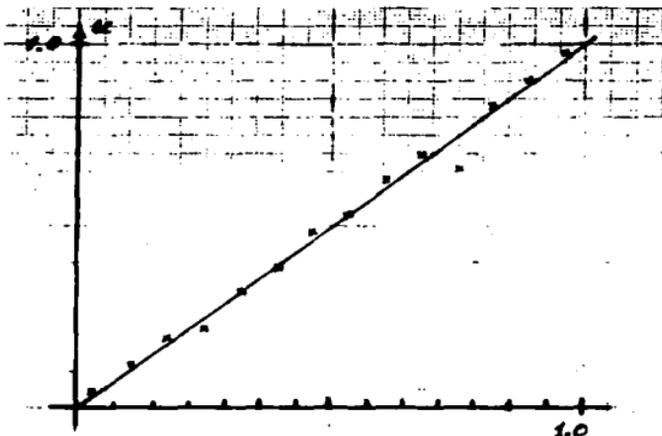


Figura 11. Ilustración gráfica de la tabla 4.5.6.

- x Solución u_Y regularizada para $Y = Y_{80}$ con $C = \mathbb{E}'\mathbb{E}$ (encabezado = $fap0$). La cual, al igual que las otras tres soluciones u_Y regularizadas, esencialmente coincide con la resolución dada por la estimación GM - clásica.

```

.....
PROBLEMA
.....
problem kernel= 1/(1+(x-y)^2),f=17-data,u=(1-y)^2
nprob= 11 n= 30 p= 15 order= 0 nker= 2
sgma= 1.000000000000e-05
SCORR= 9.5702170798299d-06
sing. values
  0.20472897e+01 : 0.58832258e+00 : 0.16989118e+00 : 0.47592619e-01
  0.12862302e-01 : 0.28877631e-02 : 0.11281988e-02 : 0.32978816e-03
  0.88297808e-04 : 0.27298488e-04 : 0.78679127e-05 : 0.19419961e-05
  0.4773797e-06 : 0.10728831e-06 : 0.20274371e-07 : 0.00000000e+00
  0.00000000e+00 : 0.00000000e+00 : 0.00000000e+00 : 0.00000000e+00
SCEN= 231187.25000000

```

```

.....
results with function fexp-biased-weighted=fexp0
.....
problem kernel= 1/(1+(x-y)^2),f=17-data,u=(1-y)^2
nprob= 11 n= 30 p= 15 order= 0 discretization= trapezoidal
sgma= 1.000000000000e-05
GASORP= 1.1037786280103d-06 ecsgas = -147255.31736013

```

```

RESULTADOS
BETA EXACTA      BETASOP      ERROR BETSOP      betgm
0.0000000e+00 : -1.4038451e-01 : -1.4038451e-01 : -28064867e+01
0.70387339e-01 : 0.79051663e-01 : 0.88643242e-02 : 0.80211336e+01
0.23990004e+00 : 0.24292874e+00 : 0.30367024e-02 : -23150645e+02
0.45356102e+00 : 0.44940483e+00 : -41541883e-02 : 0.48451472e+02
0.64638900e+00 : 0.64369716e+00 : -26918410e-02 : -77609896e+02
0.84339858e+00 : 0.84414390e+00 : 0.74531528e-03 : 0.10710841e+03
0.95960017e+00 : 0.96053804e+00 : 0.93587187e-03 : -12377703e+03
0.10000000e+01 : 0.10002847e+01 : 0.28468236e-03 : 0.13028956e+03
0.95960017e+00 : 0.96053159e+00 : 0.12514261e-02 : -21859926e+03
0.84339858e+00 : 0.84437440e+00 : 0.97581984e-03 : 0.99232479e+02
0.64638900e+00 : 0.64371791e+00 : -26710898e-02 : -70103914e+02
0.45356102e+00 : 0.44929290e+00 : -42681126e-02 : 0.43343205e+02
0.23990004e+00 : 0.24263347e+00 : 0.27334316e-02 : -20280512e+02
0.70387339e-01 : 0.78912206e-01 : 0.85248677e-02 : 0.69201795e+01
0.00000000e+00 : -1.3420185e-01 : -1.3420185e-01 : -23610393e+01

```

```

.....
problem kernel= 1/(1+(x-y)^2),f=17-data,u=(1-y)^2
nprob= 11 n= 30 p= 15 order= 1 nker= 2
sgma= 1.000000000000e-05
SCORR= 9.5702170798172d-06
sing. values
  0.46455287e+04 : 0.10220001e+01 : 0.15260363e+00 : 0.29131940e-01
  0.2399209e-02 : 0.14801927e-02 : 0.36679476e-03 : 0.93792397e-04
  0.24870371e-04 : 0.63979958e-05 : 0.16397082e-05 : 0.40532939e-06
  0.95728288e-07 : 0.2056790e-07 : 0.38632893e-08 : 0.00000000e+00
  0.00000000e+00 : 0.00000000e+00 : 0.00000000e+00 : 0.00000000e+00
SCEN= 635484.80000000

```

```

.....
results with function fexp-biased-weighted=fexp0
.....
problem kernel= 1/(1+(x-y)^2),f=17-data,u=(1-y)^2
nprob= 11 n= 30 p= 15 order= 1 discretization= trapezoidal
sgma= 1.000000000000e-05
GASORP= 3.9146881071903d-07 ecsgas = -6067002.4932865

```

```

RESULTADOS
BETA EXACTA      BETASOP      ERROR BETSOP      betgm
0.0000000e+00 : 0.11246104e-02 : 0.11246104e-02 : -28884867e+01
0.70387339e-01 : 0.70878170e-01 : 0.49083173e-03 : 0.80211336e+01
0.23990004e+00 : 0.23810845e+00 : -17915872e-02 : -23150645e+02
0.45356102e+00 : 0.45309438e+00 : -46682548e-03 : 0.48451472e+02
0.64638900e+00 : 0.64793705e+00 : 0.15680413e-02 : -77609896e+02
0.84339858e+00 : 0.8446626e+00 : 0.12678800e-02 : 0.10710841e+03
0.95960017e+00 : 0.95883789e+00 : -76227219e-03 : -12377703e+03
0.10000000e+01 : 0.99818589e+00 : -18141074e-02 : 0.13028956e+03
0.95960017e+00 : 0.95909488e+00 : -50530900e-03 : -11859926e+03
0.84339858e+00 : 0.84492950e+00 : 0.15309116e-02 : 0.99232479e+02
0.64638900e+00 : 0.64796907e+00 : 0.15800640e-02 : -70103914e+02
0.45356102e+00 : 0.45287813e+00 : -68288331e-03 : 0.43343205e+02
0.23990004e+00 : 0.23785781e+00 : -20422315e-02 : -20280512e+02
0.70387339e-01 : 0.70898457e-01 : 0.51111859e-03 : 0.69201795e+01
0.00000000e+00 : 0.14796105e-02 : 0.14796105e-02 : -23610393e+01

```

```

.....
Results with function fexp-biased-weighted=fexp0
.....
problem kernel= 1/(1+(x-y)^2),f=17-data,u=(1-y)^2
nprob= 11 n= 30 p= 15 order= 2 discretization= trapezoidal
sgma= 1.000000000000d-05
GASORP= 1.064516460898d-10 ecsgas = -5451900923.0936

```

```

RESULTADOS
BETA EXACTA      BETASOP      ERROR BETSOP      betgm
0.0000000e+00 : -2.4965318e-01 : -2.4965318e-01 : -28064867e+01
0.70387339e-01 : 0.89717330e-01 : 0.19229992e-01 : 0.80211336e+01
0.23990004e+00 : 0.24012745e+00 : 0.22740478e-03 : -23150645e+02
0.45356102e+00 : 0.4435242e+00 : -10308774e-01 : 0.48451472e+02
0.64638900e+00 : 0.66506376e+00 : -13252426e-02 : -77609896e+02
0.84339858e+00 : 0.84899277e+00 : 0.56008859e-02 : 0.10710841e+03
0.95960017e+00 : 0.96021934e+00 : 0.61918976e-03 : -12377703e+03
0.10000000e+01 : 0.9585440e+00 : -41458018e-02 : 0.13028956e+03
0.95960017e+00 : 0.96042507e+00 : 0.82490800e-03 : -11859926e+03
0.84339858e+00 : 0.84921230e+00 : 0.58127197e-02 : 0.99232479e+02
0.64638900e+00 : 0.66519177e+00 : -11972346e-02 : -70103914e+02
0.45356102e+00 : 0.44318521e+00 : -10375800e-01 : 0.43343205e+02
0.23990004e+00 : 0.23974455e+00 : -15549089e-03 : -20280512e+02
0.70387339e-01 : 0.89492048e-01 : 0.19104710e-01 : 0.69201795e+01
0.00000000e+00 : -2.4173949e-01 : -2.4173949e-01 : -23610393e+01

```

TABLA 4.5.7. Aquí se ilustra que al aumentar el orden de regularización los resultados no necesariamente mejoran. Los resultados para τ_{00} con ORDEN 1 son muy buenos.

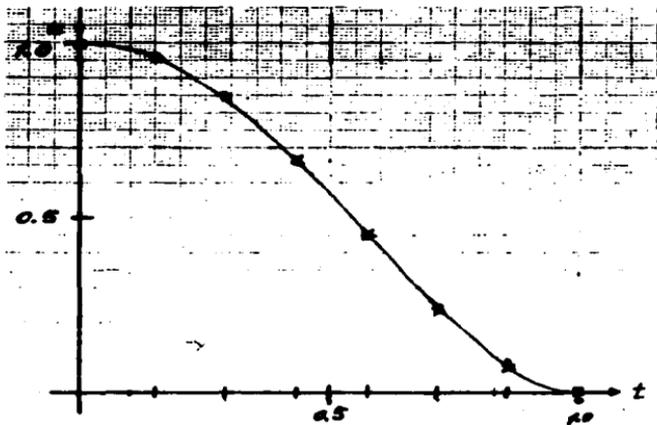


Figura 12. Ilustración gráfica de la tabla 4.5.7.

- x Solución u_y regularizada para $y = y_{SO}$ con $C = \mathbb{R}^2$ y orden 1.
- Solución u_y regularizada para $y = y_{SO}$ con $C = \mathbb{R}^2$ y orden 2.

CONCLUSIONES

El método de Regularización de Tijonov para la resolución de problemas Mal-planteados a la Hadamard presenta varias dificultades prácticas, siendo una de las principales la justificación de buena selección del parámetro de regularización $\gamma = \gamma(\delta)$ acorde al umbral $\delta > 0$ de las perturbaciones en los datos del problema bajo estudio.

A este respecto, la primera principal conclusión del presente trabajo es el haber establecido, bajo condiciones razonables, que para el caso de un operador lineal compacto nuestra nueva propuesta, el Principio de Seudo-optimalidad, define un algoritmo de regularización a la Tijonov (Teoremas 3.2.2 y 3.4.1), el cual resulta además Bien-comportado a la F. John ([50],[51]).

Cabe mencionar que resultados similares a los nuestros han aparecido en la literatura. Por ejemplo, para el problema

$$Au = f_\delta \quad \text{con} \quad \|f_\delta - f\|_F \leq \delta,$$

con $A:U \rightarrow F$ lineal compacto sobre los espacios U y F de Hilbert, Groetsch [39] demuestra que si $A^+f \in \text{Im}((A^+A)^\nu)$ para alguna ν , $0 < \nu \leq 1$, y se toma a $\gamma = K\delta^{2/(2\nu+1)}$ entonces $\|A^+f - u_Y^\delta\| = O(\delta^{2\nu/(2\nu+1)})$, siendo $u_Y^\delta = (A^+A + \gamma I)^{-1} A^+f_\delta$ (véase nuestro teorema 3.2.2.). Más aún, demuestra que la razón de convergencia $O(\delta^{2\nu/(2\nu+1)})$ para $\nu=1$ es la mejor que se puede lograr (véase la discusión que sigue a nuestro teorema 3.2.2). Como puede apreciarse, la elección $\gamma = K\delta^{2/(2\nu+1)}$ carece de fundamento al-

guno que la sustente, a menos que para ello se adopte, por ejemplo, a nuestro Principio de Seudo-optimalidad.

La segunda de nuestras principales conclusiones es relativa a la realización práctica del Principio de Seudo-optimalidad para la resolución numérica de la ecuación integral de Fedholm de primer tipo

$$\int_a^b k(x,t) u(t) dt = f(x), \quad c \leq x \leq d,$$

por el método de Regularización de Tijonov, vía la estimación estadística de las funciones de varianza total y sesgo cuadrado en conexión con la estimación Gauss-Markov regularizada en la regresión lineal, con base a tales estimaciones^(*) se desarrolló el algoritmo automático descrito en la Sec. 4.4., automáticamente en el sentido de que no hace uso del conocimiento explícito de los parámetros β y δ^2 en el modelo lineal estandard

$$X\beta + \varepsilon = y \quad \text{con} \quad \varepsilon \sim (0, \delta^2 I).$$

Ahora bien, es importante mencionar que dicho algoritmo arrojó resultados numéricos generalmente buenos para los ejemplos que hemos clasificado como blandos o moderados. Así pues, con la base de nuestro amplio, aunque aún preliminar estudio experimental, podemos concluir que el Principio de Seudo-optimalidad es digno de ser considerado, tanto por sus atributos teó

(*) Insegada para la varianza total y sesgada para el sesgo cuadrado.

riscos como prácticos, como un principio^(*) prometedor para la elección del parámetro de regularización. El cual a pesar de sus fuertes deficiencias, merece ser estudiado y desarrollado en varias direcciones de perfeccionamiento y extensión.

Para finalizar, no podemos dejar de decir que la principal dificultad práctica de nuestro Principio de Seudo-optimalidad - radica en el problema (abierto) sobre la propuesta de una "buena" estimación para la función de sesgo cuadrado, problema que nosotros reconocemos como uno Mal-planteado de la Hadamard. Y que otros autores (véase a O'Sullivan [71']) también lo reconocen como una seria dificultad en el problema de la elección del parámetro de regularización con apoyo en métodos estadísticos.

(*) De hecho competitivo al Criterio de la Validación Cruzada Generalizada debido a Wahba [37], pues la estimación $\hat{C}(\gamma)$ dada por (4.3.13) para $ECM_C(\gamma)$ con $C = E'E$ es la estimación del error cuadrático medio de predicción $E(T(\hat{\beta}_\gamma))$ donde $T(\gamma) = EEB_\gamma - EEB_\gamma^2 = L_C(\hat{\beta}_\gamma)$, al cual está íntimamente ligada la Validación Cruzada Generalizada.

REFERENCIAS.

- [1] ARSEININ V. Ya., On Ill-Posed Problems, *Russian Math. Surveys* 31, No.6 (1976) 93-107.
- [2] ARSEININ V., TIKNONOV A.N., *Solutions of Ill-Posed Problems*, Wiley (1977).
- [3] AUDLEY D.R., Lee D.A., Ill-Posed and Well-Posed Problems in System Identification, *IEEE Transac. on Autom. Control* 19, No. 6 (1974) 738-743.
- [3'] BACHMAN G., NARICI L., *Functional Analysis*, Academic Press (1966).
- [4] BAKER T.H., Fox L., et. al., Numerical Solution of Fredholm Integral Equations of First Kind, *Computer J.* 7 (1964) 141-148.
- [5] BAKUSHINSKII A.B., A General Method of Constructing Regularizing Algorithms for a Linear Ill-Posed Equation in Hilbert Spaces, *USSR Camp. Math. & Math. Phys.* 7 (1967) 279-286.
- [6] BAKUSHINSKII A.B., A Remark on a Class of Regularizing Algorithms, *USSR Comput Math & Math Phys* 13 (1973) 278-282.
- [7] BARRERA P., LOPEZ J., et. al., La Inversa Generalizada. Un Enfoque Geométrico y un Método Iterativo para Calcularla, *CIMAS Com. Tec.* #7, vol. 1, Ser. A. (1974).
- [8] BARRERA S. P. y LOPEZ E., J., Métodos de Residuos Pesados, Com. Internas (Serie Com. Tec. e Investigación), No. 11, Dpto. de Matemáticas, Fac. de Ciencias-UNAM (1977).
- [9] BARRIOZ Z., Ernesto y LOPEZ E., Jesús, Sobre el Estimador Gauss-Markov Regularizado con Estimación Óptima del Parámetro de Regularización, Reporte presentado en el XIX Congreso Nat. de la S.M.M. (Nov/1986).
- [10] BAZ TELLEZ, Guillermo R., Estimación Sesgada en el Modelo Lineal, Tesis Profesional de Actuaría, Fac. de Ciencias-UNAM (1976).
- [11] BJORCK A & ELDEN L., Methods in Numerical Algebra for Ill-Posed Problems, Rep LITH-MAT-R-33-1979, Dpto. Mathematics, Linköping Univ., Sweden (1979).
- [12] BOURGES, Javier P., On a Criterion for Choosing Biased Estimators in Regression, Com. Tec. (Serie Investigaciones), No. 384, IIMAS-UNAM (Marzo, 1985).

- [13] BOURGIN D.G. & DUFFIN R., The Dirichlet Problem for the Vibrating String Equation, *Bull. Amer. Math. Soc.* 45 (1935) 851-858.
- [14] BOURGIN D.G., The Dirichlet Problem for the Damped Wave Equations, *Duke Math. J.* 7 (1940) 97-120.
- [15] BUNCH J.R., DONGARRA J.J., et. al., *LINPACK: Users' Guide*, SIAM, Philadelphia (1979).
- [16] BURRUS W.R. & RUST B.W., *Mathematical Programming and Numerical Solution of Linear Equations*, ELSEVIER (1972).
- [17] CABAYAN H.S., DESCHAMPS G.A., Antenna Synthesis and Solution of Inverse Problems by Regularization Method, *IEEE Trans. Antenna Propagation* AP-20 (1972) 268-274.
- [18] CARASSO A & STONE A.P. (Ed), *Improperly Posed Boundary Value Problems*, Pitman (1975).
- [19] COURANT E. & HILBERT D., *Methods of Mathematical Physics* (1962), 2 vols. Wiley (1962).
- [20] CRONE L., The Singular Value Decomposition of Matrices and Cheap Numerical Filtering of Systems of Linear Equations, *J. Franklin Inst.* 294 (1972).
- [21] CROUT P.D., Application of Polynomial Approximation to the Solution of Integral Equation Arising in Physical Problems, *J. of Math. & Phys.* 19 (1940) 34-92.
- [22] CULLUM J., Numerical Differentiation and Regularization, *SIAM J. Num. Anal.* 8 (1971) 254-265.
- [23] CULLUM J., Ill-Posed Problems, Regularization and Singular Value Decomposition, Res. Rep. IBM RC 6465, Yorktown Heights, NY 10598 (1977).
- [24] DELVES L.M. & MOHAMED J.L., *Computational Methods for Integral Equations*, Cambridge Univ. Press (1985).
- [25] DMITRIEV V.I., ILINSKII A.S. et. al., the Development of Mathematical Methods for the Study of Direct and Inverse Problems in Electrodynamics, *Russian Math. Surveys* 31, No. 6 (1976) 133-152.
- [26] ELDMEN L., Algorithms for the Regularization of Ill-Conditioned Least Squares Problems, *BIT* 17 (1977) 134-145.

- [27] FLEMING H.E. & MARK D.Q., Indirect Measurements of Atmospheric Temperature Profiles from Satellites, *Non. Weather Rev.* 94, No. 6 (1966) 351-362.
- [28] FORSYTHE G., MALCOLM M.A. et. al., *Computer Methods for Mathematical Computations*, Prentice Hall (1977).
- [29] FOX L. & GOODWIN E.T., The Numerical Solution of Non-singular Linear Integral Equations, *Phil. Trans. Roy. Soc.* 241 (1953) 501-534.
- [30] FRANKLIN J.N., On Tikhonov's Method for Ill-Posed Problems, *Maths. of Comput.* 28, No. 128 (1974) 889-907.
- [31] GANDER W., On the Linear Least Squares Problem with Quadratic Constraint, Tech. R. STANCS-78-697, Comput. Sci. Dpt. Stanford U., Stanford-CA (1978).
- [32] GLASKO V.B., KRIKSIN Ju. A. & Tikhonov A.N., On the Question of Quasi-optimal Choice of Regularized Approximation, *Soviet Math. Dokl.* 20, No. 5 (1979) 1036-1040.
- [33] GLASKO V.B., LITVINENKO O.K. et al., Tikhonov's Regularization Method in Modern Geophysical Prospecting, *Izv. Acad. Sci. USSR Earth Phys.* 13 (1977) 14-24.
- [34] GLASKO V.B. & TIKHONOV A.N., The Approximate Solution of Fredholm Integral Equations of the First Kind, *USSR Comput. Math. & Math. Phys.* 4 (1964) 236-247.
- [35] GLASKO V.B. & TIKHONOV A.N., Use of the Regularization Method in Non-linear Problems, *USSR Comput. Math. & Math. Phys.* 5 (1965) 93-107.
- [36] GLASKO V.B., ZAKHARON M.V. et. al., Application of Regularization Method to Solve an Inverse Prob. of Non-Linear Heat Conduction Theory, *USSR Comput. Math. & Math. Phys.* 15 (1975) 244-248.
- [37] GOLUB G.H., HEATH M., WAHBA G., Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter, *Technometrics* 21, No. 2 (1979) 215-223.
- [37'] GOLUB G.H., VAN LOAN C.F., *Matrix Computations*, Johns Hopkins Univ. Press (1983).
- [38] GORDONOVA V.I. & MOROZOV V.A., Numerical Parameter Selection Algorithms in the Regularization Method, *USSR Comput. Math. & Math. Phys.* 13 (1973).

- [39] GROETSCH C.W., *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman (1984).
- [40] HADAMARD J., *Sur les Problemes Aux Derivées Partielles et Leur Signification Physique*, *Bull. Univ. Princeton* 13(1902) 49-52.
- [41] HADAMARD J., *Lectures on Cauchy Problem in Linear Partial Differential Equations*, Yale U. Press (1923).
- [42] HADAMARD J., *La Theorie des Equations Aux Derivées Partielles*, Editions Scientifiques, Pekin (1964).
- [43] HANSON R.J., *A Numerical Method for Solving Fredholm Integral Equations of the First Kind Using Singular Values*, *SIAM J. Numer. Anal.* 8(1971) 616-622.
- [44] HERRERA HOYOS, Carlos H., *Ridge Regression*, Master of Science Thesis, Dpt. of Statis., Iowa State Univ. (March, 1981).
- [45] HILGERS J.W., *On the Equivalence of Regularization and Certain Reproducing Kernel Hilbert Space Approches for Solving First Kind Problems*, *SIAM J. Num. Anal.* 13 (1976) 172-184.
- [46] HOERL A.L. & KENNARD R.W., *Ridge Regression: Biased Estimation for Non-orthogonal Problems*, *Technometrics* 12 (1970) 55-67.
- [47] HOERL A.E. & KENNARD R.W., *Ridge Regression: Applications to Non-orthogonal Problems*, *Technometrics* 12 (1970) 69-82.
- [48] JENNINGS L.S. & OSBORNE M.R., *Applications of Orthogonal Matrix Transformations to the Solution of Systems of Linear and Nonlinear Equations*, Rep. #37, Computer Center, Australian Nat. Univ. (1970).
- [49] JOHN F., *Numerical Solution of the Equation of Heat Conduction for Proceeding Times*, *Ann. Mat. Pura et Appl.* 40 (1955) 129-142.
- [50] JOHN F., *Continuous Dependence on Data for Solutions of PDE with a Prescribed Bound*, *Comm. Pure & Applied Math.* 13 (1960) 551-585.
- [51] JOHN F., *Remarks on Ill-Posed Problems*, en *Proc. Internat. Conference on Ill-Posed Problems* (Oct-1979).
- [52] KHUDAK Yu. N., *On Regularizing Algorithms*, *USSR Comput & Math Phys* 15, No. 1 (1975) 9-16.
- [53] LATTES R. & LIONS J-L, *Méthode de Quasiréversibilité et Applications*, Dunod (1967).

- [54] LAURENTIEV M.M., On the Cauchy Problem for Laplace Equation. *Dokl. Akad. Nauk. SSSR* 102, No. 2 (1955) 205-206 [en Ruso].
- [55] LAVRENTIEV M.M., *Some Improperly Posed Problems of Mathematical Physics*, Springer (1967).
- [56] LEE D.A., POTTER R.M. et. al., Some Practical Aspects of the Treatment of Ill-Posed Problems by Regularization. Tech R. AR75-0022, Aerospace Research Lab., Wright-Patterson AFB, Ohio 45433 (Feb., 1975).
- [57] LOPEZ E., Je., La Inversa Generalizada, Tesis Prof. en Matemáticas, Fac. de Ciencias-UNAM, México D.F. (1973).
- [58] MARQUARDT D.W. SNEE R., Ridge Regression in Practice. *Amer. Stat.* 12 (1975) 3-19.
- [58'] MIKHLIN S.G., *Mathematical Physics An Advanced Course*, North-Holland Publis. Co., (1970).
- [59] MILLER G.F., Fredholm Equations of the First Kind, Chap. 13 in: DELVES L.M. & Walsh J. (Eds) *Numerical Solution of Integral Equations*, Clarendon Press, Oxford (1974).
- [60] MILLER K., Efficient Numerical Methods for Backward Solution of Parabolic Probs with Variable Coefficients, in: CARASSO A & STONE A.P. (Ed), *Improperly Posed Boundary Value Problems*, Pitman (1975) 54-64.
- [61] MOROZOV V.A., On the Solution of Functional Equations by the Method of Regularization, *Soviet Math. Dokl.* 7, No. 2 (1966) 414-417.
- [62] MOROZOV V.A., Choice of Parameter for the Solution of Functional Eqs by the Regularization Method, *Soviet Math. Dokl.* 8 No. 4 (1967) 1000-1003.
- [63] MOROZOV V.A., The Error Principle in the Solution of Operational Equations by the Regularization Method, *USSR Comput. Math. & Math. Phys.* 8, No. 2 (1968) 63-87.
- [64] MOROZOV A.V., The Solution by Regularization of Incorrectly Posed Problems with Unbounded Operators, *Soviet Math. Dokl.* 6, No. 8 (1970) 1107-1111.
- [65] MOROZOV V.A., *Methods for Solving Incorrectly Posed Problems*, Springer (1984).

- [66] NASHED M.Z., Aspects of Generalized Inverses in Analysis and Regularization, in: NASHED M.Z. (Ed), *Generalized Inverses and Applications*, Academic Press (1976) 193-243.
- [67] NASHED M.Z., Approximate Regularized Solutions to Improperly Posed Linear Integral and Operator Equations, in: COLTON D.L., Gilbert R.P. (Eds), *Constructive and Computational Methods for Differential and Integral Equations*, Lecture Notes in Math. No. 430, Springer (1974).
- [68] NASHED M.Z. & WAHBA G., Regularization and Approximation of Linear Operator Equations in Reproducing Kernel Spaces, *Bull. Amer. Math. Soc.* 80 (1974) 1213-1218.
- [69] NASHED M.Z. & WAHBA G., Generalized Inverses in Reproducing Kernel Spaces, An Approach to Regularization of Linear Operator Equations, *SIAM J. Math. Anal.* 5, No. 6 (1974) 974-978.
- [70] NEDELKOV I.P., Improper Problems in Computational Physics, *Comput. Phy. Comm.* 4 (1972) 157-164.
- [71] OGANESJAN S.M. & STAROSTENKO V.I., A Regularizing Iteration Process Based on the Parametric Functional of A.N. Tihonov, *Soviet Math. Dokl.* 19, No.1 (1978), 39-43.
- [71'] O'SULLIVAN F., A Statistical Perspective on Ill-Posed Inverse Problems, *Statistical Sci.* 1, No.4 (1986) 502-527.
- [72] PAYNE L.E., *Improperly Posed Problems in PDE*, SIAM (1975).
- [73] PENROSE R., A Generalized Inverse for Matrices, *Proc. Camb. Phil. Soc.* 51 (1955) 406-413.
- [74] PHILLIPS D.L. A Technique for the Solution of Certain Integral Equations of the First Kind, *J. of ACM* 9 (1962) 84-97.
- [75] PROVENCHER S.W., Numerical Solution of Linear Integral Equation of the First Kind. Calculation of Molecular Weight Distributions from Sedimentation Equilibrium Data, *J. of Chemical Phys.* 46, No. 8 (1967) 3229-3236.
- [76] PUCCI C., Sui Problemi di Cauchy non Ben Posti, *Atti. Accad. Naz. Lincei* 8, No. 18 (1955) 473-477.
- [77] PUCCI C., Discussione del Problema di Cauchy per le Equazioni di Tipo Ellittico, *Annali di Mat. Pura et Appl.* 46 (1958) 131-153.

- [78] RIESZ F., NAGY B., *Functional Analysis*, Ungar Publ. Co. (1978).
- [79] RUDIN W., *Principles of Mathematical Analysis*, 3rd. Ed., Mc. Graw-Hill (1976).
- [80] SEBER G. A.F., *Linear Regression Analysis*, Wiley (1977).
- [81] SHAW C.B., Improvement of the Resolution of an Instrument by Numerical Solution of an Integral Equation, *J. of Math. Anal. & Applic.* 37 (1972) 83-112.
- [81'] SMITH G., CAMPBELL F., A Critique of Some Ridge Regression Methods (with discussion) *J.A.S.A.* 75 (1980) 74-103.
- [82] SMIRNOV V.I., *A Course of High Mathematics*, v.5, Pergamon (1964).
- [83] SMITHIES F., *Integral Equations*, Cambridge U. Press (1970).
- [83'] SMITHIES F., The Eigen-values and Singular Values of Integral Equations, *Proc. London Math. Soc.* 43 (1937) 255-279.
- [84] STEWART G.W., *Introduction to Matrix Computations*, Academic Press (1973).
- [85] STEWART G.W., On the Perturbation of Pseudo Inverses, Projections and Linear Least Squares, *SIAM Rev.* 19, No. 4 (1977) 643-662.
- [85'] STEWART G.W., Collinearity and Least Squares Regression, *Statistical Sci.* 2, No. 1(1987) 68-100.
- [86] STRAHOV V.N. & VALJASKO G.M., On Methods of Regularizing Linear Ill-Posed Probs, Taking Account of a Priori Information on Noise Properties on the Input Data, *Soviet Math. Dokl.* 17, No. 6 (1976) 734-738.
- [87] TIKHONOV A.N., On the Stability of Inverse Problems, *Dokl. Akademii Nauk SSSR* 39, No. 5 (1943) 195-198 [En Ruso].
- [88] TIKHONOV A.N., Solution of Incorrectly Formulated Problems and the Regularization Method, *Soviet Math. Dokl.* 4 (1963) 1035-1038.
- [89] TIKHONOV A.N., Regularization of Incorrectly Posed Problems, *Soviet Maths.* 4 (1963) 1624-1627.
- [90] TIJONOV A.N. & SAMARSKY A., *Ecuaciones de la Física Matemática*, MIR (1980).

- [91] TURCHIN V.F., KOSLOV V.P. et. al., The Use of the Mathematical Statistics Methods in the Solution of Incorrectly Posed Problems, *Soviet Phys.Usp.* 13 (1971) 681-703.
- [92] TWOMEY S., On the Numerical Solution of Fredholm Integral Equations of the Linear System Produced by Quadrature, *J. of ACM* 10 (1963) 97-101.
- [93] TWOMEY S., *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*, ELSEVIER (1977).
- [94] VABISHCHEVICH P.N., GLASKO V.B. et. al. Solution of the Hadamard Problem by a Tikhonov Regularization Algorithm, *USSR Comput. Math. & Math. Phys.* 19 (1979) 103-112.
- [95] VAINSHTEIN L.A., Filtering of Noise in Numerical Solution of Integral Eqs. of the First Kind, *Soviet Phys.Dokl.* 17 No. 6 (1972) 519-521.
- [96] VARAH J.M., On the Numerical Solution of Ill-Conditioned Linear Systems with Applications to Ill-Posed Problems, *SIAM J. Numer. Anal.* 10, No. 2 (1973) 257-267.
- [97] VARAH J. M., A Practical Examination of Some Numerical Methods for Linear Discrete Ill-Posed Problems, *SIAM Rev.* 21, No. 1 (1979) 100-111.
- [98] VOEVODIN V.V., The Method of Regularization, *USSR Comput. Math. & Math. Phys.* 9, No.3 (1969) 228-232.
- [99] WAHBA G., Converge Rates of Certain Approximate Solutions to Fredholm Integral Equations of the First Kind, *J. of Approx. Thy.* 7 (1973) 167-185.
- [100] WAHBA G., Practical Approximate Solutions to Linear Operator Equations when the Data are Noisy, *SIAM J. Num. Anal.* 14 No. 4 (1977) 651-667.
- [101] Medin P.A, Perturbation Theory for Pseudo-inverses, *SIAM* 13 (1973) 217-232.

REFERENCIAS ADICIONALES.

- [102] ALBERT A., *Regression and the Moore-Penrose Pseudo Inverse*, *Math. in Sci.* & Eng. 94, Academic Press (1972).
- [102'] ALLEN R., PRUESS S. et. al., *Fundamentals of Numerical Computing*, Pre-impreso.
- [103] BELLMAN R. et. al., *Dynamic Programming and Ill-conditioned Linear Systems*, *J. of Math. Anal. & Appl.* 10 (1965) 206-215.
- [104] BIBBY J. y TOUTENBURG H., *Prediction and Improved Estimation in Linear Models*, Wiley (1977).
- [105] DELVES L.M. et. al., *An Augmented Galerkin Method for First Kind Fredholm Equations*, *J. of IMA* 24 (1979) 157-174.
- [106] GONCHARSKII A.V. et. al., *Finite-Difference Approximation of Linear Incorrectly Posed Problems*, *USSR Comput. Math. & Math. Phys.* 14 (1974) 14-23.
- [107] LEWIS B. A., *On the Numerical Solution of Fredholm Integral Equations of the First Kind*, *J. of IMA* 16 (1975) 207-220.
- [108] LINZ P., *A General Theory for the Approximate Solution of Operator Equations of the Second Kind*, *SIAM J. Num. Anal.* 14 n.3 (1977) 543-554.
- [109] LINZ P., *Theoretical Numerical Analysis*, Wiley (1979).
- [109'] TIKHONOV A.N. et. al. (Eds.), *Ill-Posed Problems in the Natural Sciences*, MIR, Moscow (1987).
- [110] VASIN V.V., *Discrete Convergence and Finite-Dimensional Approximation of Regularizing Algorithms*, *USSR Comput. Maths. & Math. Phys.* 19 (1979) 8-19.