

**Universidad Central de Venezuela
Facultad de Ciencias
Escuela de Computación**

Lecturas en Ciencias de la Computación
ISSN 1316-6239

**Mínimos cuadrados lineales con
restricción cuadrática**

Z. Castillo

RT 2006-04

Centro de Cálculo Científico y Tecnológico de la UCV
CCCT-UCV
Caracas, Marzo, 2006.

Mínimos Cuadrados Lineales con Restricción Cuadrática

Zenaida Castillo (zenaida@kuaimare.ciens.ucv.ve)

Dpto. de Computación, Facultad de Ciencias
Universidad Central de Venezuela, UCV, Ap. 47002
Caracas 1041-A, Venezuela.

Resumen

En este trabajo se propone un método para solucionar el problema de minimizar el funcional $\|Ax - b\|_2$, sujeto a $\|Cx - d\|_2 \leq \Delta$, donde $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^m$, y $d \in \mathbb{R}^p$. Este problema surge en importantes aplicaciones que generan problemas inversos, los cuales, generalmente son mal condicionados. Se usa una técnica de *regularización* para reformular el problema, y luego resolver una ecuación no-lineal en una variable. Se presentan resultados preliminares y comparativos con métodos actuales.

Palabras y frases claves: Mínimos cuadrados con restricciones, bidiagonalización, ecuación secular.

1 Introducción

La fuente más importante de problemas de mínimos cuadrados con restricciones cuadráticas es sin duda la discretización de problemas inversos mal condicionados. Estos problemas generalmente surgen al tratar de determinar la estructura de un sistema físico, partiendo de su comportamiento. Como un ejemplo mencionamos la ecuación integral de primer orden:

$$g(y) = \int_a^b K(x, y)f(x)dx, \quad (1)$$

donde el operador K es compacto. Este es un problema mal condicionado ya que la función f no depende de manera continua de la función g . El problema que se genera al discretizar esta ecuación, es un problema de mínimos

cuadrados lineales:

$$\min_x \|Ax - b\|_2, \quad (2)$$

donde $A \in \mathbb{R}^{m \times n}$ es la discretización del operador K , y b representa los datos. En este tipo de problemas los valores singulares de A decaen exponencialmente a cero; por esta razón, los métodos tradicionales para resolver problemas de mínimos cuadrados no son útiles (ver [1, 14, 13]).

Introduciendo cierta información sobre la solución, el problema (2) puede ser reemplazado por uno equivalente bien-condicionado, que pueda resolverse con cierta precisión. Esta técnica es conocida como *Regularización*.

De esta manera (2) puede ser reformulado como un problema de minimización con restricciones:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx - d\|_2^2 \leq \Delta, \quad (3)$$

donde $\|Cx - d\|_2^2 \leq \Delta$ representa la información que se tiene acerca de la solución. En [3] Eldén propone un método computacionalmente eficiente para solucionar el problema:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx\|_2^2 \leq \Delta, \quad (4)$$

que es un caso particular del problema (3).

Hasta ahora el método propuesto por Eldén y algunas derivaciones de él, son los más usados para resolver problemas a mediana escala y esparcidos, por lo tanto tiene sentido hacer una extensión de este método para resolver el problema general (3), el cual también puede ser solucionado usando métodos de optimización con restricciones, sin embargo estos métodos pueden requerir muchas iteraciones para lograr una buena aproximación, y por lo tanto resultar ineficientes.

2 Técnicas de Regularización

Una de las técnicas de *regularización* más utilizadas para resolver el problema (2), restringe el espacio de las soluciones imponiendo una cota límite para la cantidad $\|Cx\|$. Esta técnica permite reformular el problema inicial, al problema de hallar x como la solución de:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx\|_2 \leq \Delta, \quad (5)$$

donde el parámetro Δ controla el balance entre el tamaño del residual y la continuidad de la solución.

En los casos de utilidad práctica, la solución se encuentra en la frontera, por lo tanto, tiene sentido plantear el siguiente problema:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx\|_2 = \Delta. \quad (6)$$

Este último problema es un caso particular de:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx - d\|_2^2 = \Delta^2, \quad (7)$$

cuya solución y algunos aspectos teóricos son tratados a continuación.

2.1 Existencia y Unicidad de Solución

Teorema 1.: Sea $F = \{x/\|Cx - d\|_2^2 = \Delta^2\}$ distinto de vacío. Si las siguientes dos condiciones se satisfacen:

Condición (1): $\min \|Cx - d\| < \Delta$,

Condición (2): $\text{rank} \begin{pmatrix} A \\ C \end{pmatrix} = n$,

entonces el problema (7) tiene una solución global y esta es única.

Prueba: Basta con probar que si $\{x_k\}$ es una secuencia de F con $\{Ax_k - b\}$ acotada, entonces $\{x_k\}$ también es acotada.

Si la condición (1) se satisface y $Ax_k - b$ es acotada, lo cual significa que $\|Ax_k - b\| \leq \delta$ para algún δ , entonces para todo x_k en F , existe un número $\epsilon > 0$, $\epsilon = \max(\delta, \Delta)$, tal que:

$$\left\| \begin{pmatrix} A \\ C \end{pmatrix} x_k - \begin{pmatrix} b \\ d \end{pmatrix} \right\| \leq \epsilon.$$

Por lo tanto, si la secuencia $\{Ax_k - b\}$ es acotada, $\{x_k\}$ también lo es.

Una vez que la función objetivo es cuadrática y la secuencia $\{x_k\}$ es acotada, un mínimo global existe y es único.

La condición (2) garantiza la unicidad de la solución, ya que si esto no se cumple, podemos añadir a la solución cualquier elemento de la intersección de los espacios nulos de A y C , y obtener una nueva solución.

Si la condición (2) del Teorema 1 se satisface, entonces la intersección de

los espacios nulos de A y C es el vector cero ($N(A) \cap N(C) = \{0\}$), lo cual implica que $A^T A + \mu C^T C$ es invertible para todo $\mu > 0$.

Definición : El problema (7) está en forma estándar, cuando $d = 0$ y $C = I$. Las ecuaciones normales para el problema estándar son:

$$(A^T A + \mu I)x = A^T b, \quad (8)$$

$$\|x\|_2^2 = \Delta^2. \quad (9)$$

En la actualidad existen métodos eficientes para solucionar el problema en forma estándar, ver por ejemplo: [6], [9] y [19].

El algoritmo que se propone en este trabajo, hace una transformación de (7) a la forma estándar. Luego usa la bidiagonalización propuesta por Eldén en [3], para solucionar el problema estándar, y finalmente regresa los cambios de la transformación.

Una vez que $A^T A + \mu I$ es invertible para todo $\mu > 0$, podemos usar un método iterativo para resolver la ecuación no-lineal

$$f(\mu) = x(\mu)^T x(\mu) - \Delta^2 = 0, \quad (10)$$

donde $x(\mu) = (A^T A + \mu I)^{-1} A^T b$.

Ahora bien, el método iterativo que soluciona esta ecuación no-lineal resuelve, en cada iteración, las ecuaciones normales (8), lo cual es equivalente a resolver el problema irrestricto:

$$\min_x \left\| \begin{pmatrix} A \\ \sqrt{\mu} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\|. \quad (11)$$

Recordemos también, que si A es de orden $m \times n$, con $m > n$, su factorización QR es:

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = Q \hat{R}$$

Así, el problema (11) es equivalente a:

$$\begin{aligned} & \min_x \left\| \begin{pmatrix} Q \hat{R} \\ \sqrt{\mu} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\| \\ \Leftrightarrow & \min_x \left\| U \begin{pmatrix} Q \hat{R} \\ \sqrt{\mu} I \end{pmatrix} x - U \begin{pmatrix} b \\ 0 \end{pmatrix} \right\| \end{aligned}$$

$$\Leftrightarrow \min_x \left\| \begin{pmatrix} \hat{R} \\ \sqrt{\mu}I \end{pmatrix} x - \begin{pmatrix} g_1 \\ 0 \end{pmatrix} \right\|, \quad (12)$$

donde $U = \begin{pmatrix} Q^T & 0 \\ 0 & I \end{pmatrix}$ es ortogonal ($U^T U = I$) y $g_1 = Q^T b$.

Para una explicación detallada de este último procedimiento se recomienda Golub y Van Loan [7] (pags. 239-242).

La forma natural de resolver (12) es utilizar rotaciones de Givens para sustituir las entradas $\sqrt{\mu}$ debajo de \hat{R} , dejando así el siguiente problema:

$$\min_x \left\| \begin{pmatrix} R_\mu \\ 0 \end{pmatrix} - \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|, \quad (13)$$

cuya solución es $x = R_\mu^{-1} h_1$.

Cabe señalar que, si el problema no está en forma estándar, este último procedimiento tiene un costo $O(n^3)$, y además debe hacerse en cada iteración, de allí la importancia de transformar el problema a forma estándar, con lo cual este costo se reduce a $O(n)$ operaciones.

Generalmente C es una matriz bien estructurada y condicionada, es por esta razón que el algoritmo que se propone, está basado en la descomposición QR de C . Nótese además, que si C es una matriz en banda, que es lo que sucede cuando proviene de la discretización de un operador diferencial, esta descomposición puede calcularse a un costo $O(nw)$, donde w es el ancho de la banda.

3 Método propuesto

Los métodos clásicos resuelven el problema (11) basándose en la descomposición SVD de la matriz A (ver por ejemplo [6] y [9]). Estos métodos resuelven el problema usando la bidiagonalización de A , y luego algunas transformaciones para convertirla en diagonal, procedimiento éste que no es necesario, ya que como se verá más adelante, basta con la bidiagonalización de A para resolver el problema.

El método que proponemos en este trabajo para resolver el problema (7) consta básicamente de 4 pasos:

- 1) Transformar el problema a forma estándar.
- 2) Bidiagonalizar la matriz A .
- 3) Resolver la ecuación secular ($f(\mu) = 0$).
- 4) Regresar los cambios de la transformación.

3.1 Transformación a forma estándar

Dados los siguientes datos:

$A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^m$, and $\Delta \in \mathbb{R}$ sobre los cuales se asume que:

- i) $\{x / \|Cx - d\| = \Delta\} \neq \emptyset$ y $\Delta > \min \|Cx - d\|$
- ii) $\text{rank} \begin{pmatrix} A \\ C \end{pmatrix} = n$, es decir $N(A) \cap N(C) = \{0\}$;

a continuación se presentan los cambios de variables necesarios para llevar el problema que se quiere solucionar:

$$\min_x \|Ax - b\|_2^2 \quad s.t. \quad \|Cx - d\|_2^2 \leq \Delta, \quad (14)$$

a su forma estándar:

$$\min_x \|\tilde{A}\tilde{x} - \tilde{b}\|_2^2 \quad s.t. \quad \|\tilde{x}\|_2^2 \leq \Delta. \quad (15)$$

Como ya sabemos el problema (14) es equivalente a :

$$\min_x \left\| \begin{pmatrix} A \\ \sqrt{\mu}C \end{pmatrix} x - \begin{pmatrix} b \\ d \end{pmatrix} \right\|, \quad (16)$$

donde μ es el multiplicador de Lagrange asociado a la restricción cuadrática. Por lo tanto, nuestro objetivo es convertir este último problema a su forma estándar:

$$\min_x \left\| \begin{pmatrix} \tilde{A} \\ \sqrt{\mu}I \end{pmatrix} \tilde{x} - \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \right\|. \quad (17)$$

Caso (a): $p > n$

- i) $[V, R] = \text{qr}(C)$. Factorización QR de C : $C = V_1 R$

$$C = \left[\underbrace{V_1}_n \quad \underbrace{V_2}_{n-p} \right]_{p \times p} \begin{pmatrix} R_{n \times n} \\ 0 \end{pmatrix}_{p \times n}$$

ii) Resolver $Cx_0 = d \equiv Rx_0 = V_1^T d$, (con R no-singular y triangular superior)

iii) Cambio de variables:

$$\tilde{x} = C(x - x_0),$$

$$\tilde{b} = b - Ax_0,$$

$$\tilde{A} = AR^{-1}V_1^T.$$

Así,

$$\tilde{A}\tilde{x} - \tilde{b} = AR^{-1}V_1^T[V_1R(x - x_0)] - b + Ax_0 = Ax - Ax_0 - b + Ax_0 = Ax - b, \text{ y}$$

$$\tilde{x} = C(x - x_0) = V_1R(x - R^{-1}V_1^T d) = V_1Rx - d = Cx - d.$$

Por lo tanto,

$$\text{mín } \|Ax - b\| \quad \text{s.t.} \quad \|Cx - d\| \leq \Delta \equiv \text{mín } \|\tilde{A}\tilde{x} - \tilde{b}\| \quad \text{s.t.} \quad \|\tilde{x}\| \leq \Delta.$$

Caso (b): $p = n$

i) $[V, R] = \text{qr}(C^T)$. Factorización QR de $C : C = R^T V_1^T$.

ii) Resolver $Cx_0 = d \equiv R^T V^T x_0 = d$, donde R es no-singular y triangular superior. Esto puede hacerse en dos pasos: $R^T z = d$; $x_0 = Vz$.

iii) Cambio de variables:

$$\tilde{x} = C(x - x_0),$$

$$\tilde{b} = b - Ax_0,$$

$$\tilde{A} = AVR^{-T}.$$

Así,

$$\tilde{A}\tilde{x} - \tilde{b} = AVR^{-T}(R^T V^T)(x - x_0) - b + Ax_0 = Ax - Ax_0 - b + Ax_0 = Ax - b, \text{ y}$$

$$\tilde{x} = C(x - x_0) = Cx - Cx_0 = Cx - d$$

Por lo tanto,

$$\text{mín } \|Ax - b\| \quad \text{s.t.} \quad \|Cx - d\| \leq \Delta \equiv \text{mín } \|\tilde{A}\tilde{x} - \tilde{b}\| \quad \text{s.t.} \quad \|\tilde{x}\| \leq \Delta.$$

Caso (c): $p < n$

i) $[V, R] = \text{qr}(C^T)$. Primera factorización QR de $C^T : C = R^T V_1^T$.

$$C = \left[\underbrace{R^T}_{p \times p} \quad \underbrace{0}_{p \times n-p} \right]_{p \times n} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}_{n \times n}$$

ii) Resolver $Cx_0 = d \equiv R^T V_1^T x_0 = d$, donde R es no-singular y triangular superior. Esto puede hacerse en dos pasos: $R^T z = d$; $x_0 = Vz$.

iii) Primer cambio de variables:

$\bar{x} = Vy + x_0$, donde $y = [y_1; y_2]^T$ y así $x = V_1y_1 + V_2y_2 + x_0$,
 $\bar{b} = b - Ax_0$.

Así,

$$Ax - b = AV_1y_1 + AV_2y_2 - b + Ax_0 = AV_1y_1 + AV_2y_2 - \bar{b},$$

$$Cx - d = CV_1y_1 + CV_2y_2 + Cx_0 - d = R^T V_1^T V_1 y_1 + R^T V_1^T V_2 y_2 + d - d = R^T y_1$$

(ya que $V_1^T V_2 = 0$).

Con este primer cambio de variables, el problema (16) es equivalente a:

$$\min_y \left\| \begin{pmatrix} AV_1 & AV_2 \\ \sqrt{\mu}R^T & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix} \right\|. \quad (18)$$

iv) Segunda factorización: $AV_2 = Q_1 T : [Q, T] = qr(AV_2)$.

$$AV_2 = \left[\underbrace{Q_1}_n \quad \underbrace{Q_2}_{n-p} \right]_{p \times p} \begin{pmatrix} T_{n \times n} \\ 0 \end{pmatrix}_{p \times n}$$

Nótese que con esta nueva factorización podemos definir una matriz ortogonal:

$$P = \begin{pmatrix} Q_1^T & 0 \\ Q_2^T & 0 \\ 0 & I_p \end{pmatrix},$$

y el problema (18) es equivalente a:

$$\min_y \left\| P \begin{pmatrix} AV_1 & AV_2 \\ \sqrt{\mu}R^T & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - P \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix} \right\| \quad (19)$$

$$\equiv \min_y \left\| \begin{pmatrix} Q_1^T AV_1 & Q_1^T AV_2 \\ Q_2^T AV_1 & 0 \\ \sqrt{\mu}R^T & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} Q_1^T \bar{b} \\ Q_2^T \bar{b} \\ 0 \end{pmatrix} \right\|. \quad (20)$$

Ahora bien, definamos

$$r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} Q_1^T AV_1 & Q_1^T AV_2 \\ Q_2^T AV_1 & 0 \\ \sqrt{\mu}R^T & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} Q_1^T \bar{b} \\ Q_2^T \bar{b} \\ 0 \end{pmatrix}, \quad (21)$$

como el residual que se minimiza en (20).

Una vez que r_2 y r_3 no dependen de y_2 , y y_2 puede ser escogido tal que $r_1 = 0$, nuestro problema se reduce a:

$$\min_{y_1} \left\| \begin{pmatrix} Q_2^T AV_1 \\ \sqrt{\mu}R^T \end{pmatrix} y_1 - \begin{pmatrix} Q_2^T \bar{b} \\ 0 \end{pmatrix} \right\|, \quad (22)$$

con

$$y_2 = T^{-1}Q_1^T(\bar{b} - AV_1y_1).$$

v) Segundo cambio de variables

$$\tilde{x} = R^T y_1,$$

$$\tilde{A} = Q_2^T AV_1 R^{-T},$$

$$\tilde{b} = Q_2^T \bar{b}.$$

Con estos cambios (22) se transforma en:

$$\min_{\tilde{x}} \left\| \begin{pmatrix} Q_2^T AV_1 R^{-T} \\ \sqrt{\mu} R^T \end{pmatrix} \tilde{x} - \begin{pmatrix} Q_2^T \bar{b} \\ 0 \end{pmatrix} \right\| \quad (23)$$

$$\equiv \min_{\tilde{x}} \left\| \begin{pmatrix} \tilde{A} \\ \sqrt{\mu} I \end{pmatrix} \tilde{x} - \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \right\|. \quad (24)$$

Esta última expresión corresponde a la forma estándar del problema original.

3.2 Proceso de Bidiagonalización

Una vez que el problema está en forma estándar, el próximo paso es bidiagonalizar la matriz A del problema estándar, usando transformaciones ortogonales a la derecha y a la izquierda de A tal como lo propone Eldén:

$$U^T AV = \begin{pmatrix} B \\ 0 \end{pmatrix} \implies A = U \begin{pmatrix} B \\ 0 \end{pmatrix} V^T. \quad (25)$$

Después de hacer esta transformación, hacemos el siguiente cambio de variables:

$$xx = V^T x \implies x = V xx,$$

$$g = U^T b = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}.$$

Así,

$$\|Ax - b\| = \|U^T(Ax - b)\| = \|U^T AVxx - U^T b\| = \left\| \begin{pmatrix} B \\ 0 \end{pmatrix} xx - \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\|,$$

y el problema a resolver es:

$$\min_{xx} \left\| \begin{pmatrix} B \\ \sqrt{\mu} I \end{pmatrix} xx - \begin{pmatrix} g_1 \\ 0 \end{pmatrix} \right\|_2. \quad (26)$$

En este punto, debemos encontrar una matriz ortogonal Q_μ , tal que:

$$Q_\mu^T \left(\begin{array}{c|c} B & g_1 \\ \hline \sqrt{\mu} I & 0 \end{array} \right) = \left(\begin{array}{c|c} B_\mu & z_1 \\ \hline 0 & z_2 \end{array} \right) \quad (27)$$

y resolver $B_\mu xx = z_1$, para hallar la solución $x(\mu)$ de las ecuaciones normales (8).

Ahora bien, en vista de que este problema debe ser resuelto en cada iteración para diferentes valores de μ , la bidiagonalización de A es útil si en cada iteración logramos anular las entradas $\sqrt{\mu}$ que se encuentran debajo de B , de manera eficiente. Esto se hace usando rotaciones de Givens sin alterar la forma bidiagonal de B .

3.3 Solución de la ecuación secular

Solucionar el problema estándar, significa básicamente resolver la ecuación secular:

$$f(\mu) = x(\mu)^T x(\mu) - \Delta^2 = 0. \quad (28)$$

En este trabajo, usamos el método de Newton para resolver esta ecuación no-lineal, sin embargo, una vez que las derivadas de $f(\mu)$ son conocidas, también podrían utilizarse otros métodos.

En cada iteración, debemos evaluar $f(\mu)$ y $f'(\mu)$. Hallar $f(\mu)$ requiere de la solución de (26), es decir, hallar Q_μ , resolver $B_\mu xx = z_1$ y devolver los cambios. Nótese que xx satisface :

$$(B^T B + \mu I)xx = B^T z_1 \quad \rightarrow \quad xx(\mu) = (B_\mu^T B_\mu)^{-1} B^T z_1$$

De esta manera la derivada de $f(\mu)$ en (28) esta dada por :

$$\begin{aligned} \frac{df(\mu)}{d\mu} &= 2xx(\mu)^T \frac{dx(\mu)}{d\mu} = -2xx(\mu)(B_\mu^T B_\mu)^{-2} B^T z_1 \\ \Rightarrow \frac{df(\mu)}{d\mu} &= -2xx(\mu)^T (B_\mu^T B_\mu)^{-1} (B_\mu^T B_\mu)^{-1} B^T z_1 = -2xx(\mu)^T (B_\mu^T B_\mu)^{-1} xx(\mu) \\ &\Rightarrow \frac{df(\mu)}{d\mu} = -2v^T v, \end{aligned}$$

donde $B_\mu^T v = xx(\mu)$.

Es de hacer notar, que si aplicamos el método de Newton para resolver (28), las cantidades $\frac{f(\mu_k)}{f'(\mu_k)}$ que se generan en cada iteración, tienden a ser muy pequeñas, y esto dificulta la convergencia del método. Para subsanar esta dificultad Gordonova y Morozov [9] resuelven la ecuación equivalente:

$$f(\mu) = (x^T x)^s - \Delta^{2s} = 0, \quad \text{para } -1 \leq s < 0. \quad (29)$$

Sorensen [19] usa $s = -1$ y soluciona $\frac{1}{x^T x} - \frac{1}{\Delta^2} = 0$, obteniendo muy buenos resultados. Considerando estos resultados y el hecho de que algoritmos numéricos basados en aproximaciones racionales convergen más rápido que aquellos basados en aproximaciones polinomiales, las pruebas presentadas en este trabajo usan $\frac{1}{x^T x} - \frac{1}{\Delta^2}$. En este caso, y siguiendo el mismo esquema anterior, obtenemos que la derivada es:

$$\Rightarrow \frac{df(\mu)}{d\mu} = \frac{-2v^T v}{x(\mu)^T x(\mu)}.$$

3.4 Implementación

En esta sección, se presentan algunos detalles de la implementación del método que se propone, para resolver el problema de los mínimos cuadrados con restricción cuadrática, así como también los resultados de las pruebas realizadas.

Los algoritmos que se muestran a continuación fueron diseñados en Matlab, razón por la cual, algunas instrucciones exhiben este estilo de programación.

Algoritmo Principal

Dados A , C , b , d , Δ , m , n , p

Si $p > n$

1) Factorización QR de C : $[V, R] = qr(C)$;

% $V = [V_1, V_2]$; $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; $C = V_1 R_1$

2) Resolver $R_1 x_o = V_1^T d$; % Sistema Triangular

3) $b \leftarrow b - Ax_o$; % Cambio de variable (b)

4) $A \leftarrow AR_1^{-1}V_1^T$; % Cambio de variable (A)

5) Resolver Problema estándar : $(x, \mu, iter) = std(A, b, delta, n)$;

6) Hallar solución final : $x = R_1^{-1}V_1^T x + x_o$;

fSi

Si $p = n$

1) Factorización QR de C^T : $[V, R] = qr(C^T)$;

% $C = R^T V^T$

- 2) Resolver $R^T V^T x_o = d$; % Dos pasos: $\begin{Bmatrix} (2,1) & R^T w=d \\ (2,2) & x_o=Vw \end{Bmatrix}$
 - 3) $b \leftarrow b - Ax_o$; % Cambio de variable (b)
 - 4) $A \leftarrow AV R^{-T}$; % Cambio de variable (A)
 - 5) Resolver Problema estándar : $(x, \mu, iter) = std(A, b, delta, n)$;
 - 6) Hallar solución final : $x = V R^{-T} x + x_o$;
- fSi

Si $p < n$

- 1) Factorización QR de C^T : $[V, R] = qr(C^T)$;
 % $V = [V_1, V_2]$; $R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; $C = R_1^T V_1^T$
 - 2) Resolver $R_1^T V_1^T x_o = d$; % Dos pasos: $\begin{Bmatrix} (2,1) & R_1^T w=d \\ (2,2) & x_o=V_1 w \end{Bmatrix}$
 - 3) $b \leftarrow b - Ax_o$; % Cambio de variable (b)
 - 4) $A_o = A$; $b_o = b$ % Almacenar variables originales (A, b)
 - 5) Factorizar AV_2 : $[Q, T] = qr(AV_2)$;
 % $Q = [Q_1, Q_2]$; $T = \begin{pmatrix} T_1 \\ 0 \end{pmatrix}$; $AV_2 = Q_1 T_1$
 - 6) $A \leftarrow Q_2^T AV_1 R^{-T}$; % Cambio de variable (A)
 - 7) $b \leftarrow Q_2^T b$; % Nuevo Cambio de variable (b)
 - 8) Resolver Problema estándar : $(x, \mu, iter) = std(A, b, delta, n)$;
 - 9) Hallar solución final :
 - 9.1) Resolver $R_1^T y_1 = x$;
 - 9.2) Resolver $T_1^{-1} y_2 = Q_1^T (b_o - A_o V_1 y_1)$;
 - 9.3) $x = V_1 y_1 + V_2 y_2 + x_o$;
- fSi

A continuación detallamos la rutina "std", la cual soluciona el problema estándar.

$$[x, \mu, iter] = \mathbf{std}(A, b, delta, n)$$

- 1) Inicializar parámetros de trabajo : $\mu = 0$; $delta2 = delta^2$; $tol = 10^{-8}$;
 $maxiter = 50$; $convergencia = 0$; $k = 1$; $x_{k-1} = 0$;
- 2) Bidiagonalizar A : $[U, A, V] = bidiag(A)$;
 % Usando Transformaciones de Householder: $U^T AV = \begin{pmatrix} B \\ 0 \end{pmatrix}$

- 3) $b \leftarrow Q_2^T b$; % Cambio de variable : $b = \begin{pmatrix} g_1 \\ 0 \end{pmatrix}$
- 4) Resolver Ecuación Secular ;
mientras ($k \leq \text{maxiter} \wedge \text{convergencia} = 0$) hacer
- 1) Reducir $\begin{pmatrix} B & | & g_1 \\ \sqrt{\mu}I & | & 0 \end{pmatrix}$ a $\begin{pmatrix} B_\mu & | & z_1 \\ 0 & | & z_2 \end{pmatrix}$; % Usando Rotaciones de Givens
- 1,1) $[B_\mu, z] = \text{giv}(A, b, n, \mu)$; % $B = \begin{pmatrix} B_\mu \\ 0 \end{pmatrix}$; $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$
- 2) Resolver problema reducido $B_\mu x_k = z_1$ % Sistema bidiagonal
- 3) Evaluar función f_μ y su derivada fp_μ
- 3,1) $f = \frac{1}{\text{delta}^2} - \frac{1}{x_k^T x_k}$;
- 3,2) Resolver $B_\mu v = x_k$;
- 3,3) $fp = \frac{-2v^T v}{x_k^T x_k}$;
- 4) si ($\|x_k - x_{k-1}\| < \text{tol} \vee |fp| < \text{tol}$) convergencia = 1; Salir fsi;
- 5) $\mu = \mu - \frac{f}{fp}$; % Paso Newton
- 6) $x_{k-1} = x_k$;
- 7) $k = k + 1$;
- fmientras
- 5) si (convergencia = 1) $x = Vx_k$ fsi ;

A continuación, el algoritmo para la subrutina "giv", utilizada para reducir el sistema (26) a (27).

- $[B_\mu, z] = \mathbf{giv}(A, b, n, \mu)$
- 1) Inicializar parámetros de trabajo : $\text{tao} = 0$; $\text{smu} = \sqrt{\mu}$;
- 2) Transformar las primeras $n - 1$ filas.
Para $i = 1$ hasta $i = n - 1$ hacer
- 1) $B_\mu(i, i) = \sqrt{B(i, i)^2 + \text{smu}^2}$;
- 2) $c = \frac{B(i, i)}{B_\mu(i, i)}$; $s = \frac{\text{smu}}{B_\mu(i, i)}$ % Parámetros de Givens
- 3) $B_\mu(i, i + 1) = c * B(i, i + 1)$; $l = -s * B(i, i + 1)$; % 1a. Rotación
- 4) $z(i) = c * g(i) + s * \text{tao}$; $ro = -s * g(i) + c * \text{tao}$;
- 5) $\text{smu} = \sqrt{\mu + l^2}$; $\text{tao} = \frac{l}{\text{smu} * ro}$; % 2da. Rotación
- fPara
- 3) $B_\mu(n, n) = \sqrt{B(n, n)^2 + \text{smu}^2}$;
- 4) $z(n) = \frac{B(n, n)}{B_\mu(n, n) * z(n)} + \frac{\text{smu}}{B_\mu(q, q) * \text{tao}}$

El resto de las rutinas utilizadas, son primitivas de Matlab.

4 Experimentación Numérica

En 1994, Hansen [12] presenta su paquete de regularización *Regularization Tools*, que constó en una serie de rutinas, diseñadas para analizar y resolver problemas a mediana escala, mal condicionados, que requieren la técnica de regularización.

Estas rutinas fueron codificadas en Matlab y Fortran, y están a la disposición pública en:

<http://www.imm.dtu.dk/documents/users/pch/Regutools/regutools.html>.

Las rutinas **TIKHONOV**, **TGSVD**, y **LSQI**, incluidas en este paquete, usan diferentes esquemas para regularizar la solución, sin embargo todas están basadas en la descomposición en valores singulares generalizados de (A, C) .

TIKHONOV: Resuelve el problema

$$\min_x \{ \|Ax - b\| + \lambda \|Cx\| \}.$$

TGSVD: Resuelve el problema

$$\min_x \|A_k x - b\|.$$

donde A_k es una aproximación a A de rango k .

LSQI: Resuelve el problema

$$\min_x \|Ax - b\| \quad \text{sujeto a} \quad \|L(x - x_0)\| \leq \Delta,$$

donde x_0 es una aproximación a la solución.

En estos problemas, λ , k y Δ juegan el papel de parámetros de regularización.

Nótese que la rutina LSQI usa la misma técnica de regularización que el método propuesto (MP) en este trabajo, y a pesar de tener restricciones sobre la matriz C , ya que exige que p sea menor que n , la forma de resolver el problema, es muy similar a la usada en MP. Por esta razón, las primeras pruebas, fueron realizadas para comparar MP con LSQI.

Los experimentos realizados, pueden dividirse en dos categorías:

I) Pruebas de comparación con la rutina LSQI, usando problemas generados aleatoriamente.

II) Pruebas de comparación con todas las rutinas disponibles, usando problemas reales, tomados de la literatura en el área.

En las pruebas (I) se seleccionaron, de manera aleatoria, problemas de tres tipos:

I.1) Problemas en forma estándar:

$$\min \|Ax - b\| \quad \text{sujeto a} \quad \|x\| \leq \Delta ; \quad d = 0 ; \quad C = I$$

I.2) Problemas de la forma:

$$\min_x \|Ax - b\| \quad \text{sujeto a} \quad \|Cx\| \leq \Delta ; \quad d = 0 ; \quad C \neq I$$

I.3) Problemas en forma general:

$$\min_x \|Ax - b\| \quad \text{sujeto a} \quad \|Cx - d\| \leq \Delta ; \quad d \neq 0 ; \quad C \neq I$$

Para las pruebas (II) se tomaron algunos problemas reales, de mediana escala, los cuales son clásicos en el área de regularización, y están disponibles en el paquete de regularización de Hansen. He aquí la descripción de algunos de ellos:

blur: Este problema proviene de la digitalización de imágenes borrosas debido a la turbulencia atmosférica.

deriv2: Discretización de la ecuación integral de Fredholm de primer orden, donde el *kernel* K está dado por la función de Green para la segunda derivada.

foxgood: Proviene de la discretización artificial de la ecuación integral de Fredholm de primer orden, diseñada para producir un problema severamente mal condicionado.

head: Problema inverso generado de la ecuación integral de calor, al discretizar la integral de Volterra.

ilaplace: Discretización de la transformada inversa de Laplace, mediante una cuadratura de Gauss-Laguerre.

shaw: Restauración de imágenes en $2D$, proveniente de la discretización de una ecuación integral de Fredholm.

wing: Problema proveniente de la discretización de una ecuación integral, con un *kernel* discontinuo.

4.1 Resultados

Todas las pruebas fueron realizadas con datos sintéticos, en un microcomputador Pentium IV, a 2.8 GHz, exigiendo a los métodos, una precisión de 10^{-8} , o un máximo de 50 iteraciones para los métodos iterativos. Estas pruebas fueron diseñadas para medir la precisión de cada rutina, y el número de operaciones requerido para lograr dicha precisión.

En la figura 1 se muestra el resultado de resolver problemas de tipo estándar, para diferentes valores de n . En la misma podemos ver el comportamiento de la rutina LSQI y la del método propuesto MP, a medida que el problema crece. La figura 1.a muestra el comportamiento de ambas rutinas, de acuer-

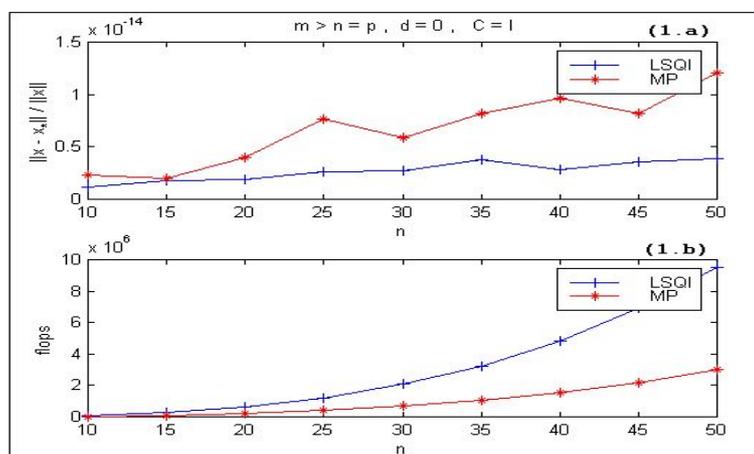


Figura 1: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas en forma estándar.

do al error relativo en la solución, y la figura 1.b de acuerdo al número de operaciones en punto flotante.

Como podemos apreciar, el error relativo en la solución para la rutina MP es mayor, a pesar de mantenerse dentro de la tolerancia exigida. Con respecto al número de operaciones, es notable como la rutina MP aventaja a la rutina LSQI.

La figura 2 por su parte, muestra resultados para problemas del tipo I.2, descrito en la sección 4. Aquí también se puede apreciar la gran ventaja de

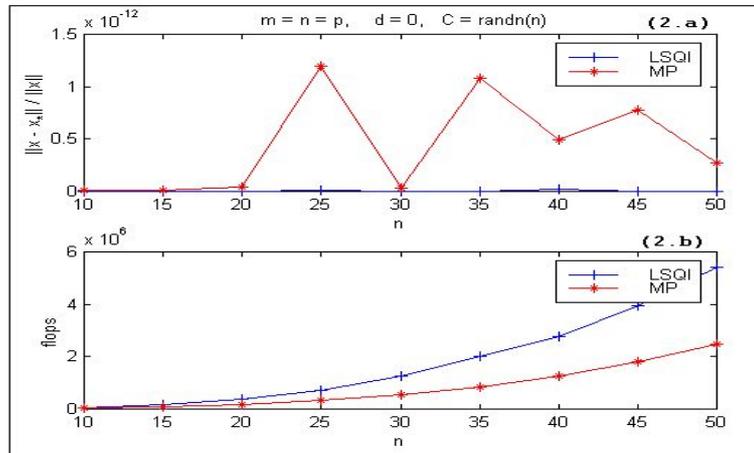


Figura 2: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas en forma general, con C escogida aleatoriamente.

MP sobre LSQI, con respecto al número de operaciones en punto flotante. La diferencia con respecto al error, es poco significativa, ya que ambas medidas se mantienen dentro de la tolerancia.

Por otra parte, la figura 3 muestra, como el comportamiento con respecto a la precisión cambia cuando se resuelven problemas en forma general, del tipo I.3. Para esta prueba se seleccionó la matriz C como una tridiagonal, simétrica, positivo-definida.

Como podemos ver, la rutina LSQI no sólo pierde precisión, sino que además no satisface la tolerancia exigida de 10^{-8} . Con respecto al número de operaciones, la situación sigue siendo la misma, la rutina LSQI duplica el número de operaciones de MP.

En la figura 4, se presentan resultados obtenidos al resolver problemas generales del tipo I.3, con A y C escogidas aleatoriamente. En este caso C no tiene estructura, y pudiera ser positivo-definida, negativo-definida, o indefinida. En esta gráfica, se enfatiza el problema de la rutina LSQI, con respecto a la pérdida de precisión en la solución, a medida que el problema crece.

En la segunda parte de estas pruebas, comparamos el método propuesto

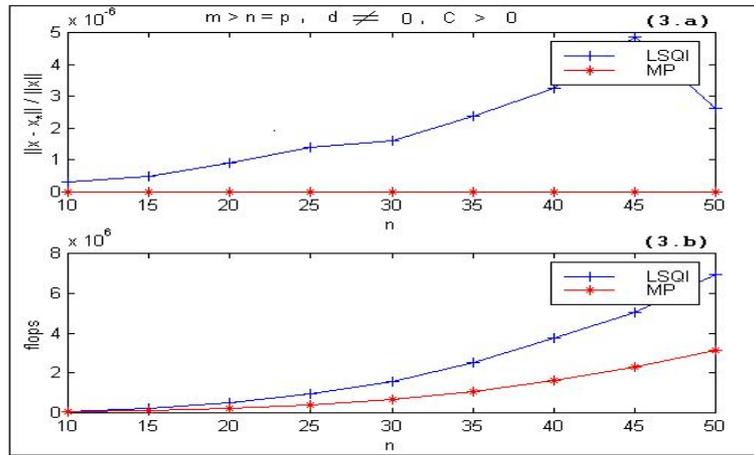


Figura 3: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas generales con C simétrica y positivo definida.

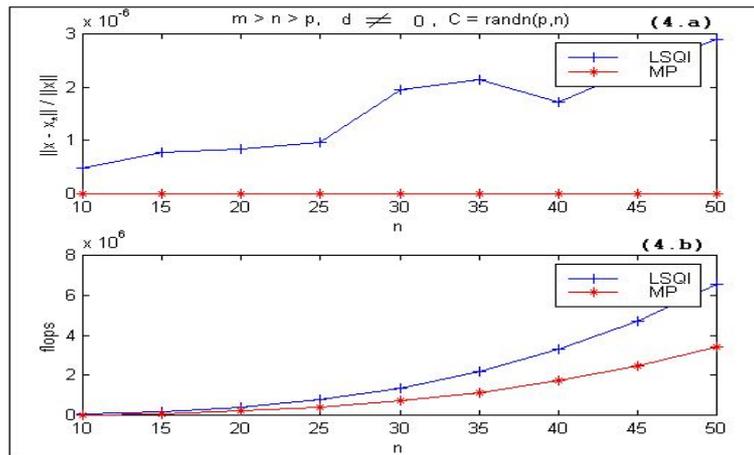


Figura 4: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas en forma general generados aleatoriamente.

(MP) con las otras rutinas de regularización disponibles en *Regularization Tools* [12]. Además de estas rutinas de regularización, hemos añadido la rutina LSQR, la cual resuelve el problema de mínimos cuadrados, usando la descomposición QR de la matriz A , sin utilizar regularización.

Los resultados obtenidos con respecto al error relativo en la solución y el número de operaciones en punto flotante, se muestran en las tablas 1 y 2 respectivamente.

Tal como se refleja en la tabla 1, para la mayoría de los casos, la solu-

Problema	LSQR	TIKHONOV	TGSVD	LSQI	MP
blur(100)	$1,4 * 10^{-15}$	$3,8 * 10^{-14}$	$6,7 * 10^{-1}$	$3,8 * 10^{-14}$	$5,7 * 10^{-16}$
deriv2(10)	$4,9 * 10^{-15}$	$6,4 * 10^{-14}$	$1,5 * 10^{-1}$	$6,4 * 10^{-14}$	$5,3 * 10^{-15}$
foxgood(20)	$6,1 * 10^1$	$3,1 * 10^{-9}$	$8,4 * 10^{-4}$	$3,8 * 10^{-9}$	$2,1 * 10^{-10}$
heat(50)	$1,8 * 10^{-8}$	$1,5 * 10^{-13}$	$8,8 * 10^{-3}$	$5,8 * 10^{-8}$	$1,2 * 10^{-10}$
ilaplace(50,2)	$1,3 * 10^2$	$8,1 * 10^{-10}$	$1,7 * 10^0$	$1,5 * 10^{-9}$	$3,3 * 10^{-10}$
shaw(20)	$3,7 * 10^{-1}$	$1,6 * 10^{-8}$	$1,0 * 10^{-4}$	$2,2 * 10^{-9}$	$4,9 * 10^{-10}$
shaw(500)	$1,5 * 10^6$	$1,5 * 10^{-8}$	$3,7 * 10^1$	$4,4 * 10^{-10}$	$6,7 * 10^{-11}$
wing(15)	$6,8 * 10^0$	$1,2 * 10^{-9}$	$8,1 * 10^{-1}$	$9,5 * 10^{-9}$	$3,6 * 10^{-9}$

Cuadro 1: Error relativo en la solución

Problema	LSQR	TIKHONOV	TGSVD	LSQI	MP
blur(100)	5578230	36774212	36026348	36772611	17594531
deriv2(10)	10991	43241	41575	43080	22013
foxgood(20)	74781	273805	264789	277575	178402
heat(50)	1062951	4381363	4277141	4427049	2523334
ilaplace(50,2)	1062951	4151584	4047108	4162152	2717320
shaw(20)	74781	272771	263681	275533	188659
shaw(500)	10^9	$4,2 * 10^9$	$4,1 * 10^9$	$4,2 * 10^9$	$2,4 * 10^9$
wing(15)	33286	120486	116135	121921	79518

Cuadro 2: Número de operaciones

ción hallada por LSQR es completamente errónea. De allí la necesidad de utilizar una técnica de regularización para resolver estos problemas.

También podemos ver en esta tabla, que las rutinas TIKHONOV y LSQI solucionan los problemas con gran precisión, generando soluciones dentro de la tolerancia exigida.

Los resultados de TGSVD se deben principalmente a que esta rutina es fuertemente dependiente del parámetro de regularización k , y la escogencia de éste parámetro, requiere de un análisis previo de cada problema. El paquete de regularización de Hansen [12] provee herramientas para este tipo de análisis, el cual está fuera del alcance de este trabajo.

Cabe hacer notar que el método propuesto en este trabajo (MP), no sólo resuelve estos problemas mal condicionados con una gran precisión, tal como puede verse en la tabla 1, sino que además mantiene su ventaja con respecto al número de operaciones (ver tabla 2).

De acuerdo a los resultados obtenidos hasta el momento, y para completar esta segunda etapa de pruebas, se diseñó un último experimento, donde comparamos las rutinas TIKHONOV, LSQI y MP en la resolución de problemas mal condicionados y perturbados. En estas pruebas perturbamos el lado derecho b , de tal manera que $b = b + \sigma * e$, con e representando un ruido aleatorio, y σ el orden de la perturbación. Los resultados obtenidos, se muestran en la tabla 3.

Los tres primeros renglones de la tabla 3, presentan resultados sobre proble-

Problema	σ	TIKHONOV	LSQI	MP
deriv2	10^{-8}	$1,5 * 10^{-4}$	$6,4 * 10^{-3}$	$1,6 * 10^{-9}$
foxgood	10^{-8}	$2,2 * 10^{-2}$	$,9 * 10^{-1}$	$3,8 * 10^{-10}$
shaw	10^{-8}	$3,3 * 10^{-2}$	$8,2 * 10^{-3}$	$1,6 * 10^{-9}$
ilaplace	10^{-16}	$4,4 * 10^{-9}$	$4,7 * 10^{-9}$	$2,9 * 10^{-9}$
ilaplace	10^{-8}	,2190	$3,6 * 10^{-1}$	$1,1 * 10^{-9}$
ilaplace	10^{-4}	$2,4 * 10^3$,7093	$2,7 * 10^{-7}$
ilaplace	10^{-3}	$1,7 * 10^4$,7102	$4,8 * 10^{-7}$

Cuadro 3: Precisión Vs. Perturbación

mas donde las componentes del lado derecho se han perturbado en el octavo dígito. Como se puede apreciar, la precisión de las rutinas TIKHONOV y LSQI se afecta notablemente con la perturbación.

Para observar el comportamiento de las rutinas a medida que aumenta la perturbación σ , seleccionamos el problema "ilaplace", con $C = trid(-1, 2, -1)$, que corresponde a la discretización de la segunda derivada. Los resultados se

muestran en los cuatro últimos renglones de la tabla 3.

Es claro que el mal condicionamiento de estos problemas, hace que las rutinas sean sensibles a pequeñas perturbaciones, y puede observarse en la tabla 3, como las rutinas TIKHONOV y LSQI van degradando su precisión a medida que se incrementa el orden de la perturbación. Para perturbaciones de orden mayor a 10^{-4} , estas rutinas producen soluciones totalmente equivocadas.

La rutina MP también es afectada por las perturbaciones, sin embargo, siempre logra hallar una buena aproximación a la solución, aún cuando la perturbación es realmente significativa. Esto se debe principalmente a que este método maneja más información acerca de la solución, que los otros métodos, y a que las transformaciones que hace, están basadas en la matriz C y no en la matriz A , y en este caso C es simétrica, positivo definida y en banda, características que benefician aún más el comportamiento del método.

Es importante señalar acá, que aunque las pruebas no fueron diseñadas para comparar requerimientos de memoria, el método propuesto (MP) tiene un alto requerimiento de memoria, el cual llega a ser de orden $(5n^2)$ en algunos casos. Este número, es por lo menos tres veces mayor que el requerido por las rutinas de *Regularization Tools*.

5 Conclusion

En este trabajo se ha propuesto un método para resolver el problema de mínimos cuadrados con restricción cuadrática, y se han expuesto las bases teóricas sobre las cuales se fundamenta el método.

Las pruebas realizadas son concluyentes, y nos permiten afirmar que el método propuesto es suficientemente robusto, como para alcanzar una buena aproximación a la solución, aún en presencia de perturbaciones. El método efectivamente calcula la solución del problema general, así como también el multiplicador de Lagrange asociado a esta solución.

Una gran variedad de problemas reales, son definidos a través de una ecuación integral con un operador discontinuo, el cual al ser discretizado, produce un problema de mínimos cuadrados lineales, tan mal condicionado, que no puede

ser resuelto por métodos tradicionales y requiere de técnicas como la regularización. El método que se propone, puede ser considerado un híbrido entre iterativo y directo, que usa una técnica de regularización para resolver problemas de mínimos cuadrados, en los cuales la matriz de los coeficientes es muy mal condicionada, y se tiene algún conocimiento de la solución.

Se pueden mencionar algunas ventajas fundamentales de este método con respecto a las rutinas de regularización del paquete *Regularization Tools*:

- 1) No está ligado fuertemente al parámetro de regularización Δ , aunque una buena escogencia de este parámetro garantiza una rápida convergencia.
- 2) Puede resolver problemas perturbados, con bastante precisión .
- 3) No tiene restricciones sobre las dimensiones, estructura, o características de la matriz C .
- 4) No requiere de la descomposición en valores singulares, y por lo tanto, disminuye el costo computacional.

La principal desventaja del método es su requerimiento de memoria, el cual lo restringe a resolver problemas de pequeña y mediana escala.

6 Trabajo futuro

Los resultados obtenidos son un estímulo para continuar investigando este método, razón por la cual se propone, como continuación de este trabajo:

- 1) Estudiar del comportamiento del método frente a perturbaciones de la matriz A .
- 2) Comparar MP con métodos actuales de optimización.
- 3) Hacer pruebas con problemas grandes. Para ello valdría la pena estudiar la posibilidad de obtener un código basado en productos Matriz-vector.

7 Agradecimientos

Este trabajo fue desarrollado mientras el autor se encontraba en la Universidad de Rice (Houston-USA) becado por el Consejo de Desarrollo Científico y Humanístico de la Universidad Central de Venezuela.

Referencias

- [1] Å. Björck. Numerical Methods for Least Squares Problems. SIAM, Philadelphia, 1996.
- [2] Å. Björck, E. Grimme, and P. Van Dooren. An implicit shift bidiagonalization algorithm for ill-posed systems. *BIT*, 34:510–534, 1994.
- [3] L. Eldén. Algorithms for the regularization of ill-conditioned least squares problems. *BIT*, 17:134–145, 1977.
- [4] W. Gander. Least squares with a quadratic constraint. *Numer. Math.*, 36:291–307, 1981.
- [5] D. Gay. Computing optimal locally constrained steps. *SIAM J. Sci. Stat. Comput.*, 2(2):186–197, 1981.
- [6] G.H. Golub. Some Modified Eigenvalue Problems. *SIAM Review*, 15:318–324, 1973.
- [7] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, 3rd. edition, 1996.
- [8] G.H. Golub and U. von Matt. Quadratically constrained least squares and quadratic problems. *Numer. Math.*, 59:561–580, 1991.
- [9] V.I. Gordonova and V.A. Morozov. Numerical Algorithms for Parameter Choice in Regularization Method. *Zh. Vychisl. Mat. Fiz.*, 13:539–545, 1973.
- [10] M. Hanke and P.C. Hansen. Regularization methods for large-scale problems. *Geophysical Journal*, 95:135–147, 1988.
- [11] P.C Hansen. Numerical tools for analysis and solution of Fredholm integral equations of the first kind. *Inverse Problem*, 8:849–872, 1992.
- [12] P.C Hansen. *Regularization Tools: a Matlab package for analysis and solution of discrete ill-posed problems*. *Numer. Algo.*, 6:1–35, 1994.
Software available from:
<http://www.imm.dtu.dk/documents/users/pch/Regutools/regutools.html>.

- [13] P.C Hansen. *Rank-Deficient and Discrete Ill-Posed problems: Numerical Aspects of Linear Inversion SIAM.*, Philadelphia, 1998.
- [14] C.L. Lawson, and R.J. Hanson. *Solving Least Square Problems. Classics in Applied Mathematics SIAM.*, Philadelphia, 1995.
- [15] J.J. Moré and D.C. Sorensen. Computing a trust region step. *SIAM J. Sci. Stat. Compu.*, 4(3):553–572, 1983.
- [16] J.J. Moré and S.J. Wright. *Optimization Software Guide. SIAM.*, Philadelphia, 1993.
- [17] M. Rojas, S.A. Santos and D.C. Sorensen. A matrix-free algorithm for the large-scale trust-region subproblem. Technical Report TR99-19, Department of Computational and Applied Mathematics, Rice University, September 1999.
- [18] D.C. Sorensen. Newton’s method with a model trust region modification. *SIAM J. Numer. Anal.*, 19(2):409–426, 1982.
- [19] D.C. Sorensen. Minimization of a large-scale quadratic fuction subject to spherical constraint. *SIAM J. Optim.*, 7(1):141–161, 1997.
- [20] G.W Stewart. *Introduction to Matrix Computation.* Academic Press, New York, 1973.