Divulgaciones Matemáticas Vol. 15 No. 2(2007), pp. 161–177

Mínimos Cuadrados Lineales con Restricción Cuadrática

Linear Least Squares with Quadratic Constraints

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 numérico para resolver 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 $d \in \mathbb{R}^p$. Se utiliza una técnica de *regularización* para reformular el problema, y luego resolver una ecuación no-lineal en una variable. Se presentan resultados preliminares y comparaciones con métodos actuales.
Palabras y frases claves: Mínimos cuadrados con restricciones, bidiagonalización, ecuación secular.

Abstract

This work proposes a method to minimize the function $\|Ax-b\|_2$, subject to $\|Cx-d\|_2 \leq \Delta$, with $A \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^m$, and $d \in \mathbb{R}^p$. We use a *regularization* technique to reformulate the problem, in order to solve a non-linear equation. Preliminary results are encouraging. We present a comparison with traditional approaches.
Key words and phrases: Least squares with constraints, bidiagonalization, secular equation.

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

Recibido 2006/03/20. Revisado 2006/09/10. Aceptado 2006/11/08.
MSC (2000): 65F05, 65F22.

la estructura de un sistema físico, partiendo de su comportamiento. Como un ejemplo mencionamos la ecuación integral de primer orden:

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

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:

$$\min_x \|Ax - b\|_2, \tag{2}$$

donde $A \in R^{m\times n}$ 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, 12, 11]).

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 reemplazado por un problema de minimización con restricciones:

$$\min_x \ \|Ax - b\|_2^2 \ \ s.t. \ \ \|Cx - d\|_2^2 \ \leq \Delta, \tag{3}$$

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:

$$\min_x \ \|Ax - b\|_2^2 \ \ s.t. \ \ \|Cx\|_2^2 \ \leq \Delta, \tag{4}$$

que es un caso particular del problema (3).

Este problema ha sido estudiado por otros autores, entre ellos Hansen [9], Gander [4] y Hanke [8]. En este trabajo se propone un método para resolver el problema general (3) basado en el método propuesto por Eldén para resolver (4).

2 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:

$$\min_x \ \|Ax-b\|_2^2 \ \ s.t. \ \ \|Cx\|_2 \ \leq \Delta, \tag{5}$$

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 estudiar el problema:

$$\min_x \ \|Ax-b\|_2^2 \ \ s.t. \ \ \|Cx-d\|_2^2 \ = \Delta^2, \tag{6}$$

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, y denotemos con $rank(A)$ el número de columnas linealmente independientes de la matriz A. Si las siguientes dos condiciones se satisfacen:
Condición (1): mín $\|Cx-d\| < \Delta$
Condición (2): $rank \begin{pmatrix} A \\ C \end{pmatrix} = n$,
entonces el problema (6) tiene una solución global y ésta es única.

Demostración: 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:

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

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. Cuando esta condición 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 $A^TA + \mu C^TC$ es invertible para todo $\mu > 0$.

Definición : El problema (6) está en forma estándar, cuando $d = 0$ y $C = I$.

Las ecuaciones normales para el problema estándar son:

$$(A^T A + \mu I)x = A^T b \tag{7}$$

$$\|x\|_2^2 = \Delta^2 \tag{8}$$

En la actualidad existen métodos eficientes para resolver el problema en forma estándar, ver por ejemplo: [7] y [13].

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

$$f(\mu) = x(\mu)^T x(\mu) - \Delta^2 = 0, \tag{9}$$

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 (7), lo cual es equivalente a resolver el problema irrestricto:

$$\min_x \ \| \begin{pmatrix} A \\ \sqrt{\mu} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \|. \tag{10}$$

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 (10) es equivalente a:

$$\min_x \ \| \begin{pmatrix} Q\hat{R} \\ \sqrt{\mu} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \|$$

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

$$\Leftrightarrow \ \min_x \ \| \begin{pmatrix} \hat{R} \\ \sqrt{\mu} I \end{pmatrix} x - \begin{pmatrix} g_1 \\ 0 \end{pmatrix} \|, \tag{11}$$

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 [6] (pags. 239-242).

La forma natural de resolver (11) es utilizar rotaciones de Givens para sustituir las entradas $\sqrt{\mu}$ debajo de $\hat{R}$, dejando así el siguiente problema:

$$\min_x \ \| \begin{pmatrix} R_\mu \\ 0 \end{pmatrix} x - \ \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \|, \tag{12}$$

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

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, lo cual reduce el costo a $O(n)$ operaciones.

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

Los métodos clásicos resuelven el problema (10) basándose en la descomposición SVD de la matriz A (ver por ejemplo [5] y [7]). Para ello proponen la bidiagonalización de A y algunas transformaciones adicionales para convertirla en diagonal, procedimiento éste que no es necesario, ya que como se verá más adelante, la bidiagonalización de A es suficiente para resolver el problema.

El método que se propone en este trabajo para resolver el problema (6) 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.

2.2 Transformación a forma estándar

Dados los siguientes datos:
$A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^m$, $d \in \mathbb{R}^m$ y $\Delta \in \mathbb{R}$
sobre los cuales se asume que:
i) $\{x / \|Cx - d\| = \Delta\} \neq \oslash$ y $\Delta > \min \|Cx - d\|$
ii) 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 resolver:

$$\min_x \ \|Ax - b\|_2^2 \ \ s.t. \ \ \|Cx - d\|_2^2 \ \leq \Delta, \tag{13}$$

a su forma estándar:

$$\min_x \ \|\tilde{A}\tilde{x} - \tilde{b}\|_2^2 \ \ s.t. \ \ \|\tilde{x}\|_2^2 \ \leq \Delta. \tag{14}$$

Como ya sabemos, el problema (13) es equivalente a :

$$\min_x \ \| \begin{pmatrix} A \\ \sqrt{\mu}C \end{pmatrix} x - \begin{pmatrix} b \\ d \end{pmatrix} \|, \tag{15}$$

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:

$$\min_x \ \| \begin{pmatrix} \tilde{A} \\ \sqrt{\mu}I \end{pmatrix} \tilde{x} - \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \|. \tag{16}$$

caso a: $p > n$
i) $[V, R] = \text{qr}(C)$. Factorización QR de $C : C = V_1 R$

$$C = \left[\underbrace{V_1}_{n} \ \ \underbrace{V_2}_{n-p} \right]_{p\times p} \left(\begin{array}{c} R_{n\times n} \\ 0 \end{array} \right)_{p\times n}$$

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,
$\tilde{A}\tilde{x} - \tilde{b} = AR^{-1}V_1^T[V_1R(x - x_0)] - b + Ax_0 = Ax - Ax_0 - b + Ax_0 = Ax - b$,
$\tilde{x} = C(x - x_0) = V_1R(x - R^{-1}V_1^T d) = V_1Rx - d = Cx - d$.
Por lo tanto,
$\min \|Ax - b\| \quad s.t. \quad \|Cx - d\| \leq \Delta \ \equiv \ \min \|\tilde{A}\tilde{x} - \tilde{b}\| \quad s.t. \quad \|\tilde{x}\| \leq \Delta$.

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

$\tilde{x} = C(x - x_0) = Cx - Cx_0 = Cx - d.$
Por lo tanto,
$\min \|Ax - b\| \quad s.t. \quad \|Cx - d\| \leq \Delta \;\equiv\; \min \|\tilde{A}\tilde{x} - \tilde{b}\| \quad s.t. \quad \|\tilde{x}\| \leq \Delta.$

caso c: $p < n$

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

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

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 $x = V_1 y_1 + V_2 y_2 + x_0$,
$\bar{b} = b - Ax_0$.
Así,
$Ax - b = AV_1 y_1 + AV_2 y_2 - b + Ax_0 = AV_1 y_1 + AV_2 y_2 - \bar{b}$,
$Cx - d = CV_1 y_1 + CV_2 y_2 + Cx_0 - d = R^T V_1^T V_1 y_1 + R^T V_1^T V_2 y_2 + d - d = R^T y_1$
(ya que $V_1^T V_2 = 0$).
Con este primer cambio de variables, el problema (15) es equivalente a:

$$\min_y \; \| \left(\begin{array}{cc} AV_1 & AV_2 \\ \sqrt{\mu} R^T & 0 \end{array} \right) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix} \|. \tag{17}$$

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

$$AV_2 = \left[\underbrace{Q_1}_{n} \quad \underbrace{Q_2}_{n-p} \right]_{p\times p} \left(\begin{array}{c} T_{n\times n} \\ 0 \end{array} \right)_{p\times n}$$

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 (17) es equivalente a:

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

$$\equiv \quad \min_y \; \| \left(\begin{array}{cc} Q1^T AV_1 & Q_1^T AV_2 \\ Q_2^T AV_1 & 0 \\ \sqrt{\mu} R^T & 0 \end{array} \right) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} Q_1^T \bar{b} \\ Q_2^T \bar{b} \\ 0 \end{pmatrix} \|. \tag{19}$$

Ahora bien, definamos

$$r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} Q1^T AV_1 & Q_1^T AV_2 \\ Q_2^T AV_1 & 0 \\ \sqrt{\mu} R^T & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} Q_1^T \bar{b} \\ Q_2^T \bar{b} \\ 0 \end{pmatrix}, \tag{20}$$

como el residual que se minimiza en (19).
Una vez que r_2 y r_3 no dependen de y_2, y y_2 puede ser escogido tal que $r_1 = 0$, el problema se reduce a:

$$\min_{y_1} \| \begin{pmatrix} Q_2^T AV_1 \\ \sqrt{\mu} R^T \end{pmatrix} y_1 - \begin{pmatrix} Q_2^T \bar{b} \\ 0 \end{pmatrix} \| \tag{21}$$

$$y_2 = T^{-1} Q_1^T (\bar{b} - AV_1 y_1)$$

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 (21) es equivalente a:

$$\min_{\tilde{x}} \| \begin{pmatrix} Q_2^T AV_1 R^{-T} \\ \sqrt{\mu} R^T \end{pmatrix} \tilde{x} - \begin{pmatrix} Q_2^T \bar{b} \\ 0 \end{pmatrix} \| \tag{22}$$

$$\equiv \quad \min_{\tilde{x}} \| \begin{pmatrix} \tilde{A} \\ \sqrt{\mu} I \end{pmatrix} \tilde{x} - \begin{pmatrix} \tilde{b} \\ 0 \end{pmatrix} \| \tag{23}$$

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

2.3 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:

$$U^T AV = \begin{pmatrix} B \\ 0 \end{pmatrix} \Longrightarrow A = U \begin{pmatrix} B \\ 0 \end{pmatrix} V^T. \tag{24}$$

Después de esta transformación, hacemos el siguiente cambio de variables:
$xx = V^T x \quad \Rightarrow \quad x = Vxx$,
$g = U^T b = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}$.
Así,

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

y el problema a resolver es:

$$\min_{xx} \| \begin{pmatrix} B \\ \sqrt{\mu}I \end{pmatrix} xx - \begin{pmatrix} g_1 \\ 0 \end{pmatrix} \|_2. \tag{25}$$

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

$$Q_\mu^T \left(\begin{array}{c|c} B & g_1 \\ \sqrt{\mu}I & 0 \end{array} \right) = \left(\begin{array}{c|c} B_\mu & z_1 \\ 0 & z_2 \end{array} \right), \tag{26}$$

y resolver $B_\mu xx = z_1$, para hallar la solución $x(\mu)$ de (7).

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 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, tal como lo sugiere Björck en [2].

2.4 Solución de la ecuación secular

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

$$f(\mu) = x(\mu)^T x(\mu) - \Delta^2 = 0. \tag{27}$$

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 del método de Newton, debemos evaluar $f(\mu)$ y $f'(\mu)$. Hallar $f(\mu)$ requiere de la solución de (25), es decir, hallar Q_μ, resolver $B_\mu xx = z_1$ y devolver los cambios. Nótese que xx satisface :

$$(B^TB + \mu I)xx = B^Tz_1 \quad \Rightarrow \quad xx(\mu) = (B_\mu^TB_\mu)^{-1}B^Tz_1.$$

De esta manera la derivada de $f(\mu)$ en (27) está dada por :

$$\frac{df(\mu)}{d\mu} = -2v^Tv,$$

donde $B_\mu^Tv = xx(\mu)$.

Si aplicamos el método de Newton para resolver (27), 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 difultad Gordonova y Morozov [7] resuelven la ecuación equivalente:

$$f(\mu) = (x^T x)^s - \Delta^{2s} = 0, \qquad para - 1 \leq s < 0. \tag{28}$$

Sorensen [13] usa s = -1 y resuelve $\frac{1}{x^T x} - \frac{1}{\Delta^2} = 0$, 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, la 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:

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

3 Experimentación numérica

Los algoritmos para llevar el problema a forma estándar, para bidiagonalizar la matriz A, y para resolver el problema de los mínimos cuadrados fueron implantados en Matlab 6.5.

En esta sección se presenta una comparación entre el método propuesto (MP) y algunos métodos clásicos implantados en el paquete *Regularization Tools*, desarrollado por Hansen en 1994 [10]. Este paquete consiste en una serie de rutinas, diseñadas para analizar y resolver problemas mal condicionados, que requieren la técnica de regularización, y está a disposición pública en:
`http://www2.imm.dtu.dk/~pch/Regutools/regutools.html`.

Para estas pruebas utilizaremos las rutinas **TIKHONOV**, **TGSVD** y **LSQI**, incluidas en este paquete, las cuales ofrecen diferentes esquemas para regularizar la solución, y están basadas en la descomposición en valores singulares generalizados de (A,C).

TIKHONOV: Resuelve el problema $\quad \min_x\{\|Ax - b\| + \lambda\|Cx\|\}$.

TGSVD: Resuelve el problema $\quad \min_x \|A_k x - b\|$, donde A_k es una aproximación a A con $rank(A_k) = k$.

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 de *Regularization Tools*, usando problemas reales, tomados de la literatura en el área.

En las pruebas (I) se generaron, de manera aleatoria, problemas en forma estándar y problemas en forma general, mientras que 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 [10], entre ellos: **blur**, **deriv2**, **foxgood**, **heat**, **ilaplace**, y otros.

3.1 Resultados

Todas las pruebas fueron realizadas con datos sintéticos, en un microcomputador Pentium IV, a 2.8GHz, 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 en forma 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 parte (a) de esta figura muestra el comportamiento de ambas rutinas, de acuerdo al error relativo en la solución, y la parte (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 la ventaja de la rutina MP sobre la rutina LSQI.

La Figura 2, muestra cómo el comportamiento con respecto a la precisión cambia cuando se resuelven problemas en forma general. Para esta prueba se seleccionó la matriz C como una tridiagonal, simétrica y positivo-definida.

Podemos ver en esta figura, que 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

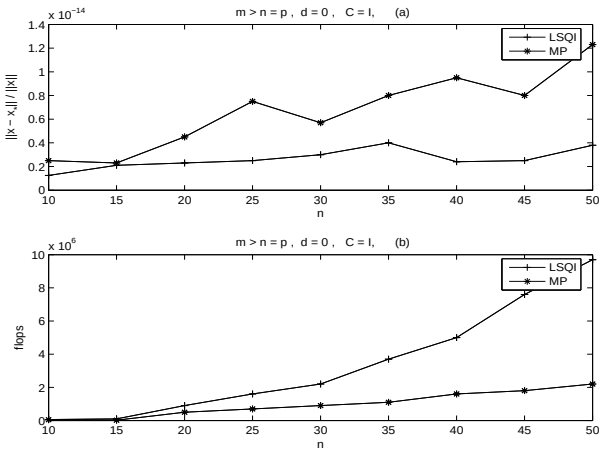

Figura 1: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas en forma estándar.

duplica el número de operaciones de MP.

En la segunda parte de estas pruebas, comparamos el método propuesto (MP) con las otras rutinas de regularización disponibles en *Regularization Tools* [10].

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 la técnica de regularización. Los problemas seleccionados fueron propuestos por Hansen [10] para evaluar métodos de regularización. Los resultados obtenidos con respecto al error relativo en la solución y el número de operaciones en punto flotante, se muestran en la Tabla 1. En la primera columna de la tabla se nombra cada problema junto con su dimensión entre paréntesis.

Tal como se refleja en la Tabla 1, en la mayoría de los casos, la solución hallada por LSQR es completamente errónea. De allí la necesidad de utilizar una técnica de regularización para resolverlos. También podemos ver en esta tabla, que las rutinas TIKHONOV y LSQI resuelven 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. En este trabajo se tomó $k = n$, considerando la descomposición completa en valores singulares de la matriz

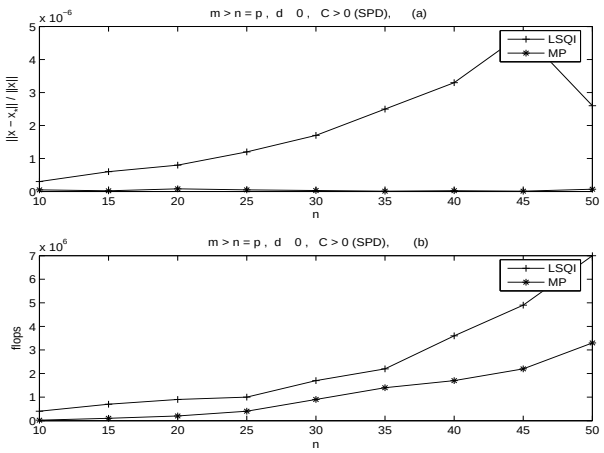

Figura 2: Comparación numérica de los métodos **MP** y **LSQI** sobre problemas en forma general, con C tridiagonal y positivo definida.

A, lo cual no garantiza la calidad del resultado. La escogencia óptima de éste parámetro requiere de un análisis previo para cada problema y el paquete de regularización de Hansen [10] 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).

Un último experimento se diseñó para comparar 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 problemas 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

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)	$1{,}8*10^{-8}$	$1{,}5*10^{-13}$	$8{,}8{*}10^{-3}$	$5{,}8*10^{-8}$	$1{,}2{*}10^{-10}$
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)	$6{,}8*10^{0}$	$1{,}2*10^{-9}$	$8{,}1{*}10^{-1}$	$9{,}5*10^{-9}$	$3{,}6*10^{-9}$

Tabla 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
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

Tabla 2. Número de operaciones

Problema	σ	TIKHONOV	LSQI	MP
deriv2(20)	10^{-8}	$1{,}5*10^{-4}$	$6{,}4{*}10^{-3}$	$1{,}6*10^{-9}$
foxgood(20)	10^{-8}	$2{,}2*10^{-2}$	$,9*10^{-1}$	$3{,}8{*}10^{-10}$
shaw(50)	10^{-8}	$3{,}3*10^{-2}$	$8{,}2{*}10^{-3}$	$1{,}6*10^{-9}$
ilaplace(50)	10^{-16}	$4{,}4*10^{-9}$	$4{,}7{*}10^{-9}$	$2{,}9*10^{-9}$
ilaplace(50)	10^{-8}	,2190	$3{,}6{*}10^{-1}$	$1{,}1*10^{-8}$
ilaplace(50)	10^{-4}	$2{,}4*10^{3}$	,7093	$2{,}7*10^{-5}$
ilaplace(50)	10^{-3}	$1{,}7*10^{4}$	,7102	$4{,}8*10^{-4}$

Tabla 3. Precisión Vs. Perturbación

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 la forma en la que las rutinas TIKHONOV y LSQI van degradando su precisión a medida que se incrementa el orden de la perturbación.

La rutina MP también es afectada por las perturbaciones, sin embargo, siempre logra hallar una aproximación del mismo orden de la perturbación.

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, 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*.

4 Conclusión

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, aun en presencia de perturbaciones. Algunas ventajas de este método con respecto a las rutinas de regularización del paquete *Regularization Tools* son:
1) No está ligado fuertemente al parámetro de regularización Δ, aunque una escogencia apropiada del mismo 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.
Los resultados obtenidos son un estímulo para continuar investigando este método, razón por la cual se propone, como continuación de este trabajo una comparación con métodos actuales de optimización, el estudio de una implantación basada en productos Matriz-vector, a fin de disminuir los requerimientos de memoria y usarlo para problemas de gran tamaño.

5 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. Su culminación fue parcialmente financiada por el CDCH-UCV (PG 03-00-5579-2004).

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] G.H. Golub. Some Modified Eigenvalue Problems. *SIAM Review*, 15:318-324, 1973.

[6] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, 3rd. edition, 1996.

[7] V.I. Gordonova and V.A. Morozov. Numerical Algorithms for Parameter Choice in Regularization Method. *Zh. Vychis1. Mat. Fiz.*, 13:539–545, 1973.

[8] M. Hanke and P.C. Hansen. Regularization methods for large-scale problems. *Geophysical Journal*, 95:135–147, 1988.

[9] P.C Hansen. Numerical tools for analysis and solution of Fredholm integral equations of the first kind. *Inverse Problem*, 8:849–872, 1992.

[10] P.C Hansen. *Regularization Tools: a Matlab package for analisys 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.

[11] P.C Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion SIAM.*, Philadelphia, 1998.

[12] C.L. Lawson, and R.J. Hanson. *Solving Least Square Problems. Classics in Applied Mathematics SIAM.*, Philadelphia, 1995.

[13] D.C. Sorensen. Minimization of a large-scale quadratic fuction subject to spherical constraint. *SIAM J. Optim.*, 7(1):141–161, 1997.