

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE CIENCIAS

UN ALGORITMO PARA LA SOLUCION DEL PROBLEMA DE
APROXIMACION UNIFORME EN EL CASO DISCRETO.

T E S I S

QUE PARA OBTENER EL TITULO DE

M A T E M A T I C O

P R E S E N T A

FAUSTO FERNANDO ANGELES ALVAREZ

MEXICO, D.F.

1983



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

P R O L O G O

El presente trabajo surgió como una sugerencia de los Dres. Enrique Chicurel U. y Marco A. Murray-Lasso, de la DEPFI* - UNAM, dada la aplicación que tiene el problema aproximaciones uniformes en el área de la ingeniería de diseño.

La revisión inicial estuvo a cargo del Prof. Jesús López Estrada y la revisión total la hizo el Dr. Pablo Barrera Sánchez, ambos de la Facultad de Ciencias de la UNAM; este último sugirió ideas fundamentales para la elaboración del trabajo.

El desarrollo del programa LAWSON se llevó a cabo en las instalaciones de la DEPFI, durante el tiempo en que el autor prestó ahí sus servicios como ayudante de profesor. El trabajo de mecanografía estuvo a cargo de la Srta. Teresa Quintero. A todos ellos agradece el autor su valiosa colaboración.

* División de Estudios de Posgrado de la Facultad de Ingeniería.

I N D I C E

	pág
PROLOGO	i
1 INTRODUCCION	1
1.1 Introducción al problema de aproximación de funciones discretas.	1
1.2 Problema general de aproximación de funciones discretas.	8
2 EXISTENCIA, CARACTERIZACION Y UNICIDAD DE LA APROXIMACION UNIFORME	12
2.1 Sistemas de Chebyshev.	12
2.2 Existencia de la aproximación uniforme.	13
2.3 Caracterización de la aproximación uniforme.	15
2.4 Unicidad de la aproximación uniforme.	17
3 APROXIMACIONES DE MINIMOS CUADRADOS PESADOS	20
3.1 Propiedades de las aproximaciones de mínimos cuadrados pesados.	20
3.2 Continuidad de las aproximaciones de mínimos cuadrados pesados.	24
3.3 Algoritmo de Lawson.	26
3.4 Criterios para acelerar la convergencia del algoritmo de Lawson.	33
3.5 Alternativas	34
4 REALIZACION DEL ALGORITMO DE LAWSON	36
4.1 Matrices ortogonales y aproximación de mínimos cuadrados.	36

4.2	Descripción del programa LAWSON.	42
5	RESULTADOS Y CONCLUSIONES	47
5.1	Resultados.	47
5.2	Conclusiones.	64
	BIBLIOGRAFIA	66
	APENDICE	67

1. INTRODUCCION

1.1 INTRODUCCION AL PROBLEMA DE APROXIMACION DE FUNCIONES DISCRETAS.

Considérese el siguiente problema:

Determinar los parámetros a_0 y a_1 de una función de la forma $L(a_0, a_1; x) = a_0 + a_1 x$ que se ajuste lo mejor posible a una colección de puntos $\{P_i = (x_i, y_i)\}_{i=1}^m$ dada. (Fig. 1)

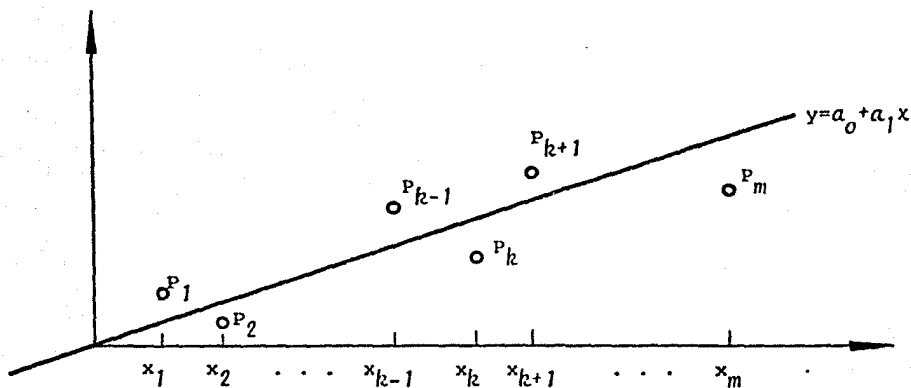


Fig. 1.1. Aproximación por medio de una línea recta.

Entre los criterios más usuales para determinar este ajuste tenemos los siguientes:

$$l_2 : \sum_{i=1}^m r_i^2 = \sum_{i=1}^m (a_0 + a_1 x_i - y_i)^2 \quad (\text{mínimos cuadrados})$$

$$l_{2w} : \sum_{i=1}^m \omega_i r_i^2 = \sum_{i=1}^m \omega_i (a_0 + a_1 x_i - y_i)^2$$

$$\text{con } \omega_i \geq 0 \text{ y } \sum_{i=1}^m \omega_i = 1 \quad (\text{mínimos cuadrados pesados})$$

$$d_{\infty} : \max_{1 \leq i \leq m} |r_i| = \max_{1 \leq i \leq m} |a_0 + a_1 x_i - y_i| \quad (\text{aproximación uniforme de Chebyshev})$$

y en cada caso se trata de hallar a_0 y a_1 de tal forma que minimicen el valor de la expresión correspondiente.

Si hacemos

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}, \quad \bar{y} = \begin{bmatrix} y \\ y \\ \vdots \\ y_m \end{bmatrix}, \quad \bar{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

se puede escribir las expresiones anteriores en términos vectoriales de la manera siguiente

$$d_2(a_0, a_1) = \min_{a_0, a_1} \| X \bar{a} - \bar{y} \|_2^2$$

$$d_{2\omega}(a_0, a_1) = \min_{a_0, a_1} \| X \bar{a} - \bar{y} \|_{2\omega}^2$$

$$d_{\infty}(a_0, a_1) = \min_{a_0, a_1} \| X \bar{a} - \bar{y} \|_{\infty}$$

Lo anterior es un caso particular del problema siguiente:

Determinar los parámetros $A = (a_0, a_1, \dots, a_n)$ de una función de la forma

$$L(A; x) = a_0 \phi_0(x) + a_1 \phi_1(x) + \dots + a_n \phi_n(x),$$

donde $\phi_i(x)$ son las funciones aproximantes (por ejemplo $\phi_i(x) = x^i$), que mejor se ajuste a una colección de puntos $\{P_i = (x_i, y_i)\}$ dada.

Es decir, cada y_i queremos expresarla como

$$y_i = a_0 \phi_0(x_i) + a_1 \phi_1(x_i) + \dots + a_n \phi_n(x_i) \\ i=1, \dots, m$$

Ahora la matriz X del caso anterior toma la forma

$$X = \begin{bmatrix} \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_n(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \dots & \phi_n(x_2) \\ \vdots & \vdots & \dots & \vdots \\ \phi_0(x_m) & \phi_1(x_m) & \dots & \phi_n(x_m) \end{bmatrix}$$

esto es, X es una matriz de $m \times (n+1)$ y $\bar{a} = (a_0, a_1, \dots, a_n)$.

En este trabajo vamos a tratar el caso en el que $m > n+1$.

Definimos una "esfera" en \mathbb{R}^m , respecto a un punto fijo $x_0 \in \mathbb{R}^m$, un radio $\rho \geq 0$ y una medida de distancia $\|\cdot\|$, como

$$\{ x \in \mathbb{R}^m : \|x - x_0\| = \rho \}$$

En la Fig. 1.2 se muestran las esferas en \mathbb{R}^2 para $\|\cdot\|_2$, $\|\cdot\|_{2\omega}$,

$\|\cdot\|_\infty$ con $x_0 = (0,0)$ y $\rho = 1$

El problema de aproximación en \mathbb{R}^m está relacionado con el problema de proyección. La proyección correspondiente a $\|\cdot\|_2$ es la proyección usual.

Una forma de visualizar la proyección correspondiente a $\|\cdot\|_{2\omega}$ y $\|\cdot\|_\infty$ es por medio de las esferas generadas con estas medidas de distancia. Referente a la Fig. 1.3 considérese el problema de encontrar la proyección de P en L con los criterios $\|\cdot\|_2$, $\|\cdot\|_{2\omega}$ y $\|\cdot\|_\infty$.

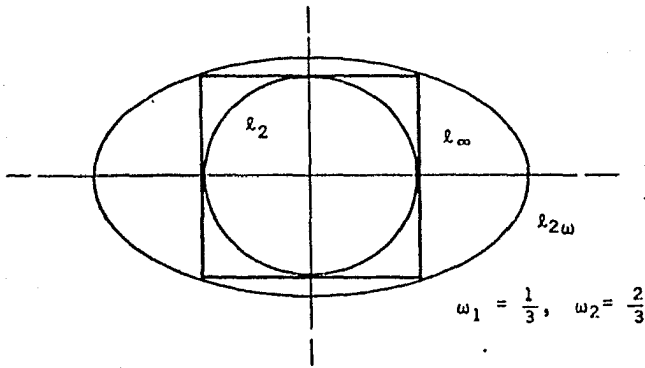


Fig. 1.2 Las esferas unitarias en \mathbb{R}^2 para $\|\cdot\|_2$, $\|\cdot\|_{2\omega}$ y $\|\cdot\|_\infty$

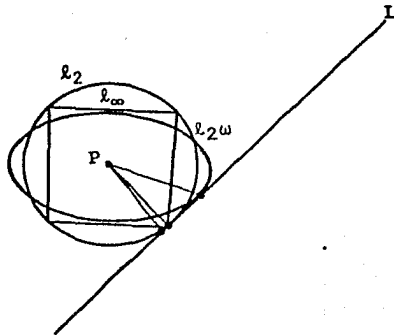


Fig. 1.3 Proyección de un punto en una recta con los criterios $\|\cdot\|_2$, $\|\cdot\|_{2\omega}$ y $\|\cdot\|_\infty$

El punto en el cual cada esfera toca a L sin cruzarla, es la proyección de P en L de acuerdo al criterio correspondiente.

Para introducir la idea de este trabajo considérese el siguiente ejemplo.

- P1. Hallar la distancia mínima entre el punto $P(3,4)$ y la recta $y = .5x$, bajo los dos siguientes criterios (medidas de distancia):

$$l_2 : d_2((3,4), (x, .5x)) = \sqrt{(3-x)^2 + (4-.5x)^2}$$

$$l_\infty : d_\infty((3,4), (x, .5x)) = \max_x \{|3-x|, |4-.5x|\}$$

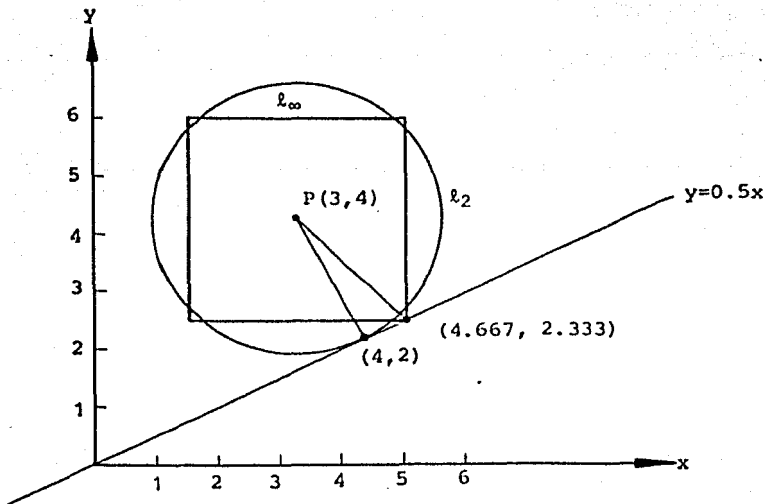


Fig. 1.4 Distancia mínima de un punto a una recta con los criterios l_2 y l_∞ .

Para el primer caso, ℓ_2 , la distancia mínima es $\sqrt{5} \approx 2.24$ y se alcanza en el punto (4,2), el cual es el punto de intersección entre la recta $y=.5x$ y la circunferencia

$$(x-3)^2 + (y-4)^2 = 5$$

Para el segundo caso, ℓ_∞ , la distancia mínima es 1.667 y se alcanza en el punto (4.667, 2.333), el cual es el punto de intersección entre la recta $y=.5x$ y el cuadrado formado por los puntos (x,y) que satisfacen

$$\max_{x,y} \{|x-3|, |y-4|\} = 1.667$$

En este trabajo se presenta un método que tiene como idea principal ir modificando el residual de las aproximaciones de mínimos cuadrados con el fin de minimizar el máximo del valor absoluto de los residuales, esto es, con el fin de obtener la aproximación ℓ_∞ . Estas modificaciones al error se logran con la introducción de pesos adecuados ω_i ($\ell_{2\omega}$).

En seguida se muestra una relación entre las normas ℓ_2 y ℓ_∞ .

Se plantea nuevamente el problema P_1 , pero ahora bajo la siguiente medida de distancia:

$$\ell_{2\omega} : d_{2\omega}((3,4), (x,.5x)) = \sqrt{\omega_1(3-x)^2 + \omega_2(4-0.5x)^2}$$

con $\omega_1, \omega_2 \geq 0$ y $\omega_1 + \omega_2 = 1$.

Los pesos ω_1 y ω_2 los calculamos utilizando el punto de distancia mínima (4,2), obtenido anteriormente para el caso ℓ_2 , de la manera siguiente:

$$\omega_1 = \frac{|4-3|}{|4-3| + |2-4|} = \frac{1}{3}, \quad \omega_2 = \frac{|2-4|}{|4-3| + |2-4|} = \frac{2}{3}$$

$$d_{2\omega}((3,4), (x, .5x)) = \sqrt{\frac{1}{3}(3-x)^2 + \frac{2}{3}(4-.5x)^2}$$

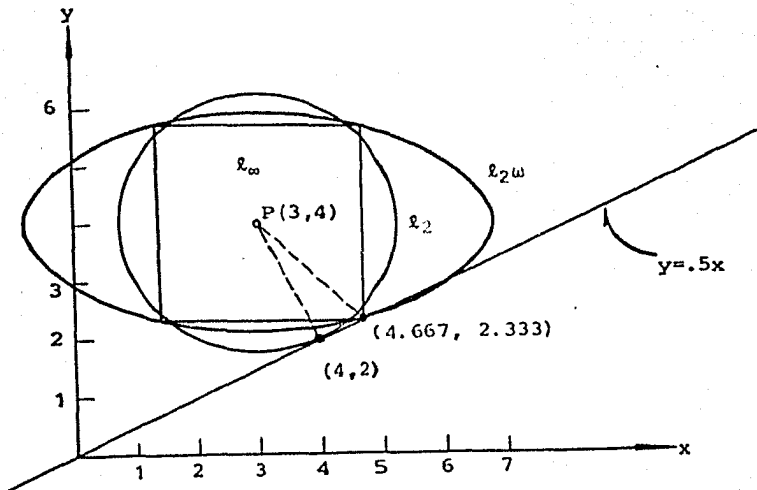


Fig. 1.5 Distancia mínima de un punto a una recta con los criterios l_2 , $l_{2\omega}$ y l_∞ .

En este caso la distancia mínima es 1.667 y se alcanza en el punto $(4.667, 2.333)$, el cual es el punto de tangencia entre la recta $y = 0.5x$ y la elipse

$$\frac{1}{3}(x-3)^2 + \frac{2}{3}(y-4)^2 = (1.667)^2$$

Como se ve, esta solución coincide con la obtenida utilizando el criterio ϵ_∞ .

En general, en el método presentado en este trabajo, se obtiene una sucesión de pesos $\{\omega^k = (\omega_1^k, \omega_2^k, \dots, \omega_m^k)\}$ de tal manera que las soluciones correspondientes $\{A^k = (a_0^k, a_1^k, \dots, a_2^k)\}$, con criterio $\epsilon_{2\omega}$, convergen a la solución correspondiente al criterio ϵ_∞ . En problemas de aproximación en los que m es mucho mayor que n la convergencia no es tan inmediata como en el ejemplo presentado anteriormente.

1.2 PROBLEMA GENERAL DE APROXIMACION DE FUNCIONES DISCRETAS.

El problema general de aproximación de funciones discretas consiste, en abstracto, en lo siguiente:

- i) Dada una función f en un cierto espacio de funciones F , de la cual se conocen sus valores únicamente en un conjunto finito $x = \{x_1, \dots, x_m\}$ de puntos de su dominio.
- ii) Dada una familia de funciones, llamadas funciones aproximantes, A de F ,
- (P) Hallar L^* en A que mejor pueda representar a f .

En las aplicaciones, usualmente la familia de funciones aproximantes A viene representada por un número finito de parámetros a_0, a_1, \dots, a_n ($a_i \in \mathbb{R}$). Esto es,

$$A = \{ L(a_0, a_1, \dots, a_n; x); a_i \in \mathbb{R} \}$$

o bien

$$A = \{ L(A; x) \in F \mid A \in \mathbb{R}^{n+1} \}$$

al definir $A = (a_0, a_1, \dots, a_n)$

Nuestro problema de aproximación se denomina lineal cuando se tiene

$$\{ \phi_j(x) \}_{j=0}^n \in F \text{ y } L(x; A) = a_0 \phi_0(x) + a_1 \phi_1(x) + \dots + a_n \phi_n(x)$$

De aquí, y en todo lo que sigue del presente trabajo, se hablará del problema antes enunciado pensando siempre en el caso lineal.

El problema de aproximación de funciones discretas se presenta cuando los valores $f(x_1), f(x_2), \dots, f(x_m)$ dados de la función f se conocen por medio de una tabla de valores y el interés principal por obtener la función aproximante $L(A, x)$ es con el fin de conocer valores aproximados de f en puntos intermedios a los valores dados. Generalmente las funciones aproximantes $L(A, x)$ son polinomios, funciones trigonométricas, funciones exponenciales o combinaciones de todas éstas. Desde un punto de vista práctico se tienen dos principales tareas en un problema de aproximación de funciones. La primera es cómo determinar el tipo de función aproximante, i.e. la elección de las funciones $\phi_j(x)$, que se va a utilizar, y la segunda es cómo medir la calidad de la aproximación. Para determinar la función aproximante $L(A, x)$ es necesario recurrir a la intuición y a la práctica, ya que no existen métodos directos para hacerlo.

En lo referente a la calidad de la aproximación, requerimos de la introducción de una medida concreta que le dé significado a qué tan bien está representada la función $f(x)$ por la familia

$$A = \{ L(A, x), | A \in R^{n+1} \}$$

La medida de la aproximación entre los valores de la función f conocidos y el elemento aproximante $L(A, x)$ viene dada por la norma del residual

$$r(A) = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_m) \end{bmatrix} - \begin{bmatrix} L(A, x_1) \\ \vdots \\ L(A, x_m) \end{bmatrix}$$

i.e. por

$$||f(x) - L(A, x)||$$

donde $|| \cdot ||$ es una norma de R^m . El problema de aproximación se plantea de la manera siguiente:

Hallar A^* en R^{n+1} de tal forma que satisfaga

$$\|r(A^*)\| \leq \|r(A)\|$$

para toda $A \in R^{n+1}$.

Este punto, visto desde la práctica, nos plantea el problema de proponer una norma específica para F .

Las normas más conocidas son las normas L_p que, para el caso discreto, se definen de la manera siguiente

DEFINICION 1.1 La norma ℓ_p de un vector $\hat{v} = (v_1, \dots, v_m)$ se denota por $\ell_p(\hat{v})$ y se define como

$$\ell_p(\hat{v}) = \left[\sum_{i=1}^m \omega_i |v_i|^p \right]^{1/p}, \quad 1 \leq p < \infty$$

donde las ω_i son pesos que satisfacen $\omega_i \geq 0$, normalmente con $\sum \omega_i = 1$

De esta definición se siguen las aproximaciones ℓ_p que corresponden a minimizar la funcional

$$\phi_p(A) = \left[\sum_{i=1}^m \omega_i (f(x_i) - L(A, x_i))^p \right]^{1/p}$$

Adicionalmente a las normas ℓ_p se tiene otra norma importante que es la norma ℓ_∞ , también conocida como norma uniforme, minimax o de Chebyshev, y que se define como

$$\|\hat{v}\|_\infty = \max |v_i|, \quad i=1, \dots, m$$

Análogamente, tenemos las aproximaciones uniformes que corresponden a minimizar la funcional

$$\phi_\infty(A) = \max_i |f(x_i) - L(A, x_i)|$$

por lo que una aproximación uniforme discreta minimiza el valor máximo de

$$|f(x) - L(A,x)|, \text{ para } x \in X = \{x_1, \dots, x_m\}$$

Dentro de las aproximaciones ℓ_p la aproximación más utilizada en la práctica es la ℓ_2 , conocida como aproximación de mínimos cuadrados. La amplia utilización de esta norma se debe, principalmente, a sus interpretaciones estadísticas.

Otras aproximaciones también utilizadas en la práctica son la ℓ_1 y la ℓ_∞ . Para estas aproximaciones sus cálculos no son tan directos como para la ℓ_2 .

En este trabajo estamos interesados en el problema (P) con respecto a la norma uniforme o ℓ_∞ .

2. EXISTENCIA, CARACTERIZACION Y UNICIDAD DE LA APROXIMACION UNIFORME

2.1 SISTEMAS DE CHEBYSHEV

Un conjunto de funciones continuas reales $\phi_0, \phi_1, \dots, \phi_n$ definidas sobre un conjunto cerrado A en \mathbb{R}^m (espacio euclidiano de dimensión m) es un "sistema de Chebyshev" si satisface las siguientes condiciones

- (a) A contiene al menos $n + 1$ puntos
- (b) Cada polinomio $P(x) = a_0 \phi_0(x) + \dots + a_n \phi_n(x)$, cuyos coeficientes a_i no son simultáneamente iguales a cero, tiene a lo más n ceros distintos en \mathbb{R}^m .

En particular, se sigue de esta definición que las funciones de un sistema de Chebyshev son linealmente independientes.

Ejemplos de sistemas de Chebyshev son las funciones $1, x, x^2, \dots, x^n$ y $1, \cos x, \sin x, \dots, \cos nx, \sin nx$.

La condición (b) se puede expresar como las siguientes condiciones:

- i) Si $\bar{x}_0, \dots, \bar{x}_n$ son puntos distintos de A , entonces el sistema de $n + 1$ ecuaciones con $n + 1$ incógnitas a_0, a_1, \dots, a_n

$$a_0 \phi_0(\bar{x}_k) + a_1 \phi_1(\bar{x}_k) + \dots + a_n \phi_n(\bar{x}_k) = 0$$

$$k = 0, \dots, n$$

tiene únicamente la solución trivial $a_0 = a_1 = \dots = a_n = 0$. Usando propiedades de sistemas de ecuaciones lineales, esto es equivalente a cada una de las siguientes propiedades.

- ii) Si $\bar{x}_0, \dots, \bar{x}_n$ son $n + 1$ puntos distintos de A , el determinante

$$D(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_n) = \begin{vmatrix} \phi_0(\bar{x}_0) & \dots & \phi_n(\bar{x}_0) \\ \dots & \dots & \dots \\ \phi_0(\bar{x}_n) & \dots & \phi_n(\bar{x}_n) \end{vmatrix}$$

es diferente de cero

iii) Si $\bar{x}_0, \dots, \bar{x}_n$ son puntos distintos de A y c_0, \dots, c_n son números reales arbitrarios, entonces el sistema de ecuaciones

$$a_0 \phi_0(\bar{x}_k) + \dots + a_n \phi_n(\bar{x}_k) = c_k, \quad k = 0, \dots, n \quad (2.1)$$

tiene una solución única para a_0, \dots, a_n . Haciendo $P = a_0 \phi_0 + \dots + a_n \phi_n$, la condición (2.1) se puede interpretar como $P(\bar{x}_k) = c_k, k = 0, \dots, n$ y P es un "polinomio interpolante" con valores determinados c_k en los puntos \bar{x}_k .

Si el número de los puntos dados \bar{x}_k y los valores c_k es menor que $n + 1$, entonces también existe un polinomio interpolante pero no es único.

2.2 EXISTENCIA DE LA APROXIMACION UNIFORME

Teorema 2.1 (De existencia) [1]. Dado \bar{y} , vector de dimensión m y $n + 1$ vectores linealmente independientes $\bar{\phi}_0, \bar{\phi}_1, \dots, \bar{\phi}_n$, también de dimensión m , el problema de hallar

$$\min_{a_i} \| \bar{y} - (a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n) \|, \quad a_i - \text{números reales},$$

tienen solución.

Demostración. Considérese la norma del error $d(a_0, a_1, \dots, a_n) = \| \bar{y} - (a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n) \|$ como una función de $n + 1$ variables reales a_0, \dots, a_n . La continuidad de esta función en las variables a_i se sigue de las siguientes desigualdades

$$\begin{aligned} |d(a'_0, \dots, a'_n) - d(a_0, \dots, a_n)| &= | \| \bar{y} - (a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n) \| - \\ &\quad - \| \bar{y} - (a'_0 \bar{\phi}_0 + \dots + a'_n \bar{\phi}_n) \| | \leq \\ &\leq \| (a'_0 - a_0) \bar{\phi}_0 + \dots + (a'_n - a_n) \bar{\phi}_n \| \leq \\ &\leq |a'_0 - a_0| \| \bar{\phi}_0 \| + \dots + |a'_n - a_n| \| \bar{\phi}_n \| \end{aligned}$$

Como las $\bar{\phi}_j$ son fijas, esto implica que la diferencia de las d debe ser pequeña si la diferencia de las a_i lo es.

En forma análoga se demuestra que la función

$$h(a_0, \dots, a_n) = \| a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n \|^2$$

es una función continua de las $\bar{a} = (a_0, \dots, a_n)$

Sea $S = \{ \bar{a} : |a_0|^2 + \dots + |a_n|^2 = 1 \}$

una superficie esférica en \mathbb{R}^{n+1} .

Por ser S un conjunto cerrado y acotado h debe alcanzar su valor mínimo $m > 0$ en S . La posibilidad de que $m = 0$ se descarta, ya que si tenemos, para algunos a_j (no todos iguales a cero), que $\| a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n \| = 0$, esto implicaría $a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n \equiv 0$ (idéntico a cero) y esto contradice la suposición de que las $\bar{\phi}_j$ son linealmente independientes.

Ahora, escribiendo $r = (|a_0|^2 + \dots + |a_n|^2)^{1/2}$, tenemos

$$h(a_0, \dots, a_n) = r \left\| \frac{a_0}{r} \bar{\phi}_0 + \dots + \frac{a_n}{r} \bar{\phi}_n \right\|$$

así que $h(a_0, \dots, a_n) \geq mr$

para todo (a_0, \dots, a_n)

Además, $d = \left\| \bar{y} - (a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n) \right\| \geq$

$$\left\| a_0 \bar{\phi}_0 + \dots + a_n \bar{\phi}_n \right\| - \left\| y \right\| \geq mr - \left\| y \right\|$$

Sean $p = \text{fnfd}(a_0, \dots, a_n)$ y $R = \frac{1 + p + \left\| y \right\|}{m}$

Si $|a_0|^2 + \dots + |a_n|^2 > R^2$, la desigualdad anterior implica

$$d \geq mR - \left\| y \right\| = 1 + p > p$$

Por lo tanto, si A es igual al espacio de todas las \bar{a} y $B = \{ \bar{a} :$

$|a_0|^2 + \dots + |a_n|^2 \leq R^2 \}$, tenemos

$$\text{fnfd}_A(a_0, \dots, a_n) = \text{fnfd}_B(a_0, \dots, a_n) \quad (2.2)$$

Como d es continua, el valor del lado derecho de (2.2) es alcanzado en el conjunto B y esto demuestra el teorema.

2.3 CARACTERIZACION DE LA APROXIMACION UNIFORME.

Teorema 2.2 [2] (Kolmogorov) Un polinomio $P = \sum_{i=0}^n a_i \phi_i(x)$ es una aproximación uniforme o de Chebyshev a una función f definida sobre un conjunto finito $X = \{x_1, \dots, x_n\}$, si y sólo si para cada polinomio $Q(x) = \sum_{i=0}^n c_i \phi_i(x)$

$$\max_{x \in X_0} \{ [f(x) - P(x)] Q(x) \} \geq 0 \quad (2.3)$$

donde $X_0 = \{x \in X : |f(x) - P(x)| = \|f - P\|\}$

Esto significa que la relación $[f(x) - P(x)] \cdot Q(x) < 0$ no puede cumplirse para toda $x \in X_0$.

Demostración. Supóngase primero que $P = \sum_{i=0}^n a_i \phi_i$ es una aproximación polinomial uniforme a f en X y sea $\|f - P\| = \epsilon$. Si la afirmación no es cierta, entonces existe un polinomio $Q(x) = \sum_{i=0}^n c_i \phi_i(x)$ tal que

$$\max_{x \in X_0} \{ [f(x) - P(x)] Q(x) \} = -2\epsilon$$

para alguna $\epsilon > 0$; entonces

$$[f(x) - P(x)] Q(x) < -\epsilon, \text{ para toda } x \in X_0$$

Ahora, formamos el polinomio $P_1 = P - \lambda Q$, con $\lambda > 0$ pequeño. Sea $M = \max_{x \in X_0} |Q(x)|$; entonces, para $x \in X_0$

$$\begin{aligned} [f(x) - P_1(x)]^2 &= [f(x) - P(x) + \lambda Q(x)]^2 = \\ &= [f(x) - P(x)]^2 + 2\lambda (f(x) - P(x))Q(x) + \lambda^2 (Q(x))^2 < \\ &< \epsilon^2 - 2\lambda \epsilon + \lambda^2 M^2 \end{aligned}$$

Si tomamos $\lambda < M^{-2} \epsilon$, entonces $\lambda^2 M^2 < \lambda \epsilon$ y obtenemos

$$[f(x) - P_1(x)]^2 < \epsilon^2 - \lambda \epsilon, \text{ para } x \in X_0 \quad (2.4)$$

Para $x \in X - X_0$, tenemos que $|f(x) - P(x)| < \epsilon$.

De aquí que, para alguna $\delta > 0$, $|f(x) - P(x)| < \epsilon - \delta$ si $x \in X - X_0$. Si tomamos λ tal que $\lambda < (2M)^{-1} \delta$, tenemos

$$|f(x) - P_1(x)| \leq |f(x) - P(x)| + \lambda |Q(x)| \leq \epsilon - \delta + \frac{1}{2} \delta = \epsilon - \frac{1}{2} \delta$$

De (2.4) y esta última desigualdad concluimos que, para valores suficientes pequeños de λ , P_1 aproxima a f mejor que P (contradicción); por lo tanto la condición del teorema es necesaria.

Para demostrar la suficiencia, supóngase que (2.3) se cumple para cada $Q(x) = \sum_{i=0}^n c_i \phi_i(x)$. Tomando un polinomio arbitrario $P_1(x) = \sum_{i=0}^n c_i \phi_i(x)$ vemos que existe un punto $x_0 \in X_0$, tal que para

$$Q = P - P_1$$

$$|f(x_0) - P(x_0)| |Q(x_0)| \geq 0$$

Entonces

$$\begin{aligned} [f(x_0) - P_1(x_0)]^2 &= [f(x_0) - P(x_0)]^2 + 2[f(x_0) - P(x_0)]Q(x_0) + (Q(x_0))^2 \geq \\ &\geq |f(x_0) - P(x_0)|^2 = \|f - P\|^2 \end{aligned}$$

ya que $x \in X_0$. Por lo tanto, P_1 no puede aproximar a f con un error menor que $\|f - P\|$ y P tiene que ser una aproximación uniforme a f .

2.4 UNICIDAD DE LA APROXIMACION UNIFORME.

Lema 2.1. [2], Sea $\phi = \{ \phi_0, \dots, \phi_n \}$ un sistema de Chebyshev de funciones reales sobre un conjunto finito X en R (números reales) que contiene

al menos $n + 2$ puntos y sea P un polinomio $(P = \sum_{i=0}^n a_i \phi_i(x))$ que mejor se aproxima a una función f definida en X en el sentido de la norma uniforme o de Chebyshev. Entonces, el conjunto X_0 de todos los puntos $x \in X$ para los cuales

$$|f(x) - P(x)| = \|f - P\| = E$$

contiene al menos $n + 2$ puntos.

Demostración. Supóngase que $X_0 = \{x_1, \dots, x_s\}$ donde $s \leq n + 1$. Como existen puntos $x \in X$ con $|f(x) - P(x)| \leq E$, debemos tener $E > 0$. Por la condición iii) existe un polinomio $Q = \sum_{i=0}^n c_i \phi_i(x)$ tal que

$$Q(x_k) = - [f(x_k) - P(x_k)], \quad k = 1, \dots, s$$

Entonces

$$\max \{ [f(x) - P(x)] Q(x) \} = \max_{x \in X_0} \{ - (f(x) - P(x))^2 \} = -E^2 < 0$$

lo cual contradice el teorema de caracterización de Kolmogorov.

Teorema 2.3 [2] (Unicidad). Para cada sistema de Chebyshev hay un polinomio único que aproxima uniformemente a cada función f definida sobre $X = \{x_1, \dots, x_m\}$

Demostración. Si X tiene exactamente $n + 1$ puntos entonces, por la condición iii) cada función definida en X es igual a un polinomio interpolante

$P(x) = \sum_{i=0}^n a_i \phi_i(x)$, y este polinomio es único. De aquí que podemos suponer que X consiste de al menos $n + 2$ puntos, por lo tanto el lema 2.2 es aplicable. Supóngase que para una función definida sobre X , existen dos polinomios $P(x) = \sum_{i=0}^n a_i \phi_i(x)$ y $P_1(x) = \sum_{i=0}^n c_i \phi_i(x)$ que aproximan uniformemente a f en X ; por lo tanto

$$\|f - P\| = \|f - P_1\| = E$$

Sea
$$Q(x) = \frac{1}{2} (P(x) + P_1(x)),$$

entonces

$$\|f - Q\| = \left\| \frac{1}{2} (f - P) + \frac{1}{2} (f - P_1) \right\| \leq \frac{1}{2} \|f - P\| + \frac{1}{2} \|f - P_1\| = E$$

Por otro lado, $\|f - Q\| \geq E$, así que Q es también un polinomio que aproxima uniformemente a f en X . Por el lema anterior existen $n + 2$ puntos x para los cuales

$$|f(x) - Q(x)| = \left| f(x) - \frac{1}{2} (P(x) + P_1(x)) \right| = E$$

En cada uno de estos puntos x , para los números

$$\alpha = f(x) - P(x), \alpha_1 = f(x) - P_1(x)$$

se tiene

$$|\alpha + \alpha_1| = 2E, |\alpha| \leq E, |\alpha_1| \leq E$$

y esto es posible sólo si $\alpha = \alpha_1$. Esto es, $P(x) = P_1(x)$ para el menos $n + 2$ puntos de X . Como X es un sistema de Chebyshev, los polinomios $P(x)$ y $P_1(x)$ son idénticos.

3. APROXIMACIONES DE MÍNIMOS CUADRADOS PESADOS

3.1 PROPIEDADES DE LAS APROXIMACIONES DE MÍNIMOS CUADRADOS PESADOS.

En esta sección se tratan propiedades básicas de las aproximaciones ℓ_2 ya que éstas constituyen la base del algoritmo de Lawson descrito posteriormente.

Considérese el siguiente problema:

P3.1 Dada una matriz $X \in \mathbb{R}^{m \times (n+1)}$ y un vector $\bar{y} \in \mathbb{R}^m$, hallar un vector $\bar{a}^* \in \mathbb{R}^{n+1}$ tal que $X\bar{a}^*$ sea el punto del subespacio generado por X más cercano a \bar{y} , de acuerdo a la norma ℓ_2 .

Esto es equivalente a hallar un vector $\bar{a}^* \in \mathbb{R}^{n+1}$ tal que

$$\|X\bar{a}^* - \bar{y}\|_2^2 \leq \|X\bar{a} - \bar{y}\|_2^2$$

para toda $\bar{a} \in \mathbb{R}^{n+1}$.

Si, por ejemplo, $m=3$ y $R(X)=2$ (rango de X), se sigue de consideraciones geométricas (Fig 3.1), que el mínimo se alcanza cuando el vector residual $\bar{r} = X\bar{a} - \bar{y}$ es perpendicular al plano $R(X)$.

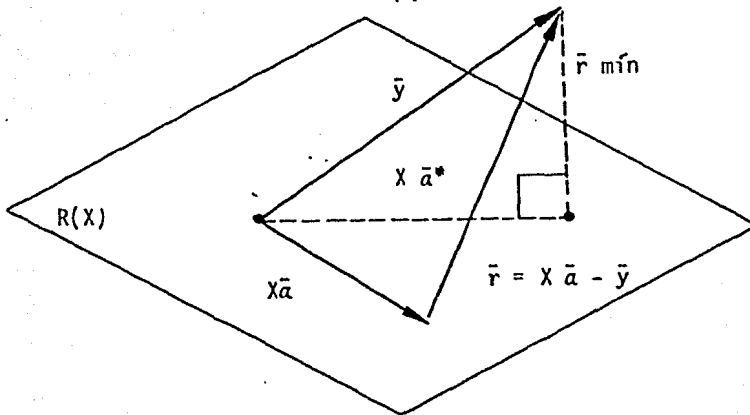


Fig 3.1 Interpretación geométrica del problema P3.1 con $m = 3$ y $R(X) = 2$

Si escribimos $\bar{a} = \bar{a}_1 + \bar{a}_2$ con $\bar{a}_1 \in R(X)$ y \bar{a}_2 en el complemento ortogonal* $R(X)^\perp$ de $R(X)$, entonces \bar{a}_1 es el vector buscado \bar{a}^* y \bar{a}_2 es el vector residual $X\bar{a} - \bar{y}$ con valor mínimo de acuerdo a la norma $\|\cdot\|_2$. Esta última descomposición de un vector tiene la siguiente forma general:

Sean S un subespacio de \mathbb{R}^n y S^\perp su complemento ortogonal. Entonces cualquier vector $\bar{u} \in \mathbb{R}^n$ se puede escribir en forma única como $\bar{u} = \bar{s}_1 + \bar{s}_2$ donde $\bar{s}_1 \in S$ y $\bar{s}_2 \in S^\perp$. El vector \bar{s}_1 es la proyección de \bar{u} en S y el vector \bar{s}_2 es la proyección de \bar{u} en S^\perp .

* Dado un subconjunto S de \mathbb{R}^n , el conjunto de todos los vectores ortogonales a S , se llama "el complemento ortogonal" de S y se denota por S^\perp .

El siguiente teorema establece la existencia de una solución del problema P3.1 y determina cuando es única.

Teorema 3.1 [3]. El problema de minimizar $\| X \bar{a} - \bar{y} \|_2$; respecto a \bar{a} , siempre tiene solución y ésta es única si y sólo si las columnas de X son linealmente independientes.

Demostración. Sean \bar{y}_1 la proyección de \bar{y} en $R(X)$ y \bar{y}_2 la proyección de \bar{y} en $R(X)^\perp$. Entonces $X \bar{a} - \bar{y}_1 \in R(X)$, para toda $\bar{a} \in \mathbb{R}^{n+1}$, y es ortogonal a \bar{y}_2 ; por lo tanto

$$\| X \bar{a} - \bar{y} \|_2^2 = \| X \bar{a} - \bar{y}_1 - \bar{y}_2 \|_2^2 = \| X \bar{a} - \bar{y}_1 \|_2^2 + \| \bar{y}_2 \|_2^2$$

y $\| X \bar{a} - \bar{y} \|_2^2$ será mínimo cuando $\| X \bar{a} - \bar{y}_1 \|_2^2$ lo sea. Pero $\bar{y}_1 \in R(X)$, de aquí que siempre podemos hallar una \bar{a}^* que cumpla $X \bar{a}^* = \bar{y}_1$.

Para esta \bar{a}^* , $\| X \bar{a}^* - \bar{y}_1 \|_2^2 = 0$.

La segunda afirmación es una propiedad de las soluciones de sistemas de ecuaciones algebraicas.

Para el problema de mínimos cuadrados pesados, $\mathcal{L}_{2\omega}$, introducimos las siguientes matrices diagonales $m \times m$

$$W = \text{diag} (\omega_i) = \begin{bmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \omega_m \end{bmatrix}$$

y

$$W' = \text{diag} (\sqrt{\omega_i}), \text{ con } \omega_i \geq 0,$$

por lo que

$$W = W' W'$$

Entonces el problema de aproximación ϵ_{2W} lo podemos plantear de la siguiente manera:

Dadas las matrices $X \in \mathbb{R}^{m \times (n+1)}$ y $W \in \mathbb{R}^{m \times m}$ y un vector $\bar{y} \in \mathbb{R}^m$, hallar un vector $\bar{a}^* \in \mathbb{R}^{n+1}$ tal que $WX\bar{a}^*$ sea el punto del subespacio generado por WX más cercano a $W\bar{y}$, de acuerdo a la norma $\|\cdot\|_2$. Esto es, queremos un vector $\bar{a}^* \in \mathbb{R}^{n+1}$ que minimice $\|W(X\bar{a} - \bar{y})\|_2^2 = \langle W'(X\bar{a} - \bar{y}), W'(X\bar{a} - \bar{y}) \rangle_2 = (X\bar{a} - \bar{y})^T W' W'(X\bar{a} - \bar{y})$ para toda $\bar{a} \in \mathbb{R}^{n+1}$.

Si WX es de rango $n+1$, entonces este problema tiene solución única.

Veremos que esta situación es la que se presenta.

Definición 3.1 Una matriz A $m \times n$, $m \geq n+1$, cumple la condición de Haar si toda submatriz de A de $s \times (n+1)$, con $s \geq n+1$, es de rango $n+1$.

Resultado 3.1. Si $\{\phi_0, \dots, \phi_n\}$ es un sistema de Chebyshev entonces para todo conjunto $\{x_1, x_2, \dots, x_m\}$ la matriz $A=(\phi_j(x_i))$ cumple la condición de Haar.

Demostración. Es inmediata de las propiedades de los sistemas de Chebyshev. (Cap. 2).

Corolario 3.1. Si $\{\phi_0, \phi_1, \dots, \phi_n\}$ es un sistema de Chebyshev y si $\{x_1, \dots, x_m\}$ son m puntos distintos entre sí, entonces el problema

$$\min \|W(X\bar{a} - \bar{y})\|_2^2,$$

donde $W = \text{diag}(\omega_i)$, $\omega_i \geq 0$; $X = (\phi_j(x_i))$, $\bar{y} \in \mathbb{R}^m$, tiene una única solución si W tiene, cuando menos, $n+1$ elementos de la diagonal distintos de cero.

3.2 CONTINUIDAD DE LAS APROXIMACIONES DE MINIMOS CUADRADOS PESADOS.

En esta sección se mencionan dos propiedades de continuidad de las aproximaciones $\mathcal{L}_{2\omega}$. De aquí en adelante, al hacer referencia a las "funciones aproximantes", se entenderá que se está considerando la aproximación a una función f de la que se conocen sus valores $f(x_i) = y_i$, únicamente para los puntos (x_1, \dots, x_m) .

Definimos el siguiente conjunto:

$\Omega = \{\hat{\omega} = (\omega_1, \dots, \omega_m) : \omega_j \geq 0, j=1, \dots, m \text{ y } \sum_{i=1}^m \omega_i = 1\}$
 y sobre este conjunto definimos la función

$$\delta(W) = \sigma(\hat{\omega}) = \min_{\bar{a}} \|W(X\bar{a} - \bar{y})\|_2^2 ;$$

donde $W = \text{diag}(\omega_i)$, X y \bar{y} son como en la sección anterior.

Se considera a la función $\delta(W)$ como una función de los pesos ω_i por la forma de la matriz W .

Esto es, $\sigma(\hat{\omega})$ es una función de

$$\sigma: \Omega \subset \mathbb{R}^m \rightarrow \mathbb{R}$$

Entonces

Teorema 3.2 [4] $\sigma(\hat{\omega})$ es continua en Ω . Se omite la demostración.

A continuación se establece la continuidad de los coeficientes a_0, a_1, \dots, a_n respecto a los pesos $\hat{\omega} = (\omega_1, \dots, \omega_m)$, en las aproximaciones $\mathcal{L}_{2,\omega}$, para un sistema de Chebyshev $\{\phi_0, \dots, \phi_n\}$ dado.

Sea $\Omega' \subset \Omega$ tal que

$$\Omega' = \{\hat{\omega} \in \Omega : \omega_i > 0 \text{ para cuando menos } n+1 \omega_i\}$$

Considérese la siguiente función

$$A : \Omega' \subset \mathbb{R}^m \rightarrow \mathbb{R}^{n+1}$$

que asocia con cada peso $\bar{\omega} \in \Omega'$ los correspondientes coeficientes únicos $\bar{a}^* = (a_0, a_1, \dots, a_n)$ que minimizan

$$\| W X \bar{a} - \bar{y} \|_2^2$$

donde las funciones ϕ_j en $X = (\phi_j(x_i))$ forman un sistema de Chebyshev.

Entonces,

Teorema 3.3 [4] La función A es continua. Se omite la demostración.

3.3 ALGORITMO DE LAWSON

Como se ha visto, el problema de aproximación uniforme discreta consiste en aproximar los valores $f(x_i) = f_i$ de una función f definida sobre el conjunto $\{x_1, \dots, x_m\}$ por medio de funciones aproximantes del tipo

$$L(A; x) = \sum_{i=0}^n a_i \phi_i(x) \quad (3.1)$$

donde $\{\phi_i\}$ es un sistema de Chebyshev (Cap. 2), de tal forma que se minimice

$$\max_A |L(A; x) - f(x)|$$

Sobre todas las funciones $L(A; x)$ del tipo (3.1).

La idea del algoritmo de Lawson para obtener esta aproximación consiste en obtener una sucesión de pesos $\{\omega^k\}$,

$$\omega^k = (\omega_1^k, \dots, \omega_m^k), \text{ con } \sum_{i=1}^m \omega_i^k = 1,$$

de tal forma que su correspondiente sucesión $\{L(A_k^*; x)\}$ de funciones aproximantes bajo la norma $\ell_{2\omega}$, converja a la función aproximante bajo el criterio ℓ_ω (Cap. 1).

Los pesos iniciales ω_i^1 son arbitrarios, por ejemplo $\omega_i^1 = \frac{1}{m}$, para $i = 1, \dots, m$

Los pesos posteriores ω_i^k , $k \geq 2$, tienen como función ir modificando el error obtenido en cada aproximación de tal forma que se les dé mayor peso a aquellos puntos donde el error sea mayor, ya que interesa minimizar dicho error.

El criterio para determinar la convergencia del algoritmo está basado en observar los pesos ω_i^k y desechar aquéllos con valor insignificante (muy cercano a cero) hasta dejar un número de pesos igual a $s+1$, donde s es el orden de la función aproximante (el orden de una función aproximante es igual al número de funciones que constituyen su base), después de lo cual se llega inmediatamente a la solución, como se demuestra más adelante.

Para comprobar la validez de la aproximación se puede recurrir a los teoremas de caracterización del capítulo 2 de este trabajo.

Los pasos del algoritmo de Lawson son:

a) Tómanse pesos iniciales ω_i^1 arbitrarios, por ejemplo

$$\omega_i^1 = 1/m, \text{ para } i=1, \dots, m$$

b) Con los pesos ω_i^k , $i=1, \dots, m$, calcúlese la aproximación, por medio de funciones aproximantes del tipo $L(A; x)$, a $f(x)$ definida en $\{x_1, \dots, x_m\}$ con el criterio de la norma $\ell_{2\omega}$; es decir, calcúlese $L(A_k^*; x)$, $A_k^* \in \mathbb{R}^{n+1}$, tal que

$$\left[\sum_{i=1}^m \omega_i^k (L(A_k^*; x_i) - f_i)^2 \right]^{1/2} \leq \left[\sum_{i=1}^m \omega_i^k (L(A; x_i) - f_i)^2 \right]^{1/2}$$

para toda $A = (a_0, \dots, a_n) \in \mathbb{R}^{n+1}$

c) Si $k=1$ ir a d)

Si

$$\left[\sum_{i=1}^m \omega_i^k (L(A_k^*; x_i) - f_i)^2 \right] = \left[\sum_{i=1}^m \omega_i^{k-1} (L(A_{k-1}^*; x_i) - f_i)^2 \right]$$

para el proceso

d) Calcúlese nuevos pesos ω_j^{k+1} a partir de ω_j^k , haciendo

$$\omega_j^{k+1} = \frac{\omega_j^k |L(A_k^*; x) - f_j|}{\sum_{i=1}^m \omega_i^k |L(A_k^*; x_i) - f_i|}$$

y regrésese a b)

Aunque este método tiene convergencia lenta (convergencia lineal), dicha convergencia se puede acelerar utilizando criterios que indiquen cuáles pesos ω_i^k se deben hacer iguales a cero, hasta quedarse únicamente con los pesos necesarios para obtener la aproximación uniforme deseada. En seguida se demuestra un teorema que justifica este procedimiento.

Se denota para cada peso $\omega \in W$,

$$W = \{ \omega = (\omega_1, \dots, \omega_m) : \sum_{i=1}^m \omega_i = 1, \omega_j \geq 0 \text{ para toda } j \text{ y } > 0$$

para, al menos, $n+1$ valores de j },

su correspondiente aproximación $\ell_{2\omega}$ como $L(A; x)_\omega$ y se define el mapeo

$$F : W \rightarrow W$$

como

$$F(\omega^k) = \omega^{k+1}$$

donde ω^{k+1} se obtiene en la forma que se indica en el algoritmo.

También se define el conjunto

$$E = \{ x_j \in X : |f(x_j) - L(A^*; x_j)| = \| f(x) - L(A^*; x) \|_\infty \}$$

donde $L(A^*; x)$ es la aproximación uniforme a f definida en $X = \{x_1, \dots, x_m\}$; por lo que E contiene los puntos del conjunto X en los cuales se alcanza el máximo error absoluto correspondiente a la aproximación ℓ_∞ (error minimax).

Se definen además los conjuntos de índices

$$J_E = \{j : x_j \in E\}$$

$$J_0 = \{j : x_j \notin E\}$$

Por la caracterización de la aproximación uniforme (Cap. 2) el conjunto E contiene al menos $n+2$ puntos, si las funciones aproximantes son de orden $n+1$.

El siguiente teorema y sus corolarios establecen las condiciones bajo las cuales la convergencia del algoritmo es inmediata.

Teorema 3.4 [5] Sea $\omega = (\omega_1, \dots, \omega_m) \in W$ tal que $\omega_j = 0$ para toda $j \in J_0$ y sean

$$r_j = f(x_j) - L(A; X_j)_\omega$$

$$W' = F(\omega)$$

$$r_j^* = f(x_j) - L(A^*; x_j)$$

si

$$\operatorname{sgn} r_j = \operatorname{sgn} r_j^* ; \quad r^* = (r_1^*, \dots, r_m^*)$$

para toda $j \in J_E$, entonces

$$L(A; x)_{\omega_1} = L(A^*; x)$$

Demostración. Por la propiedad de ortogonalidad de las aproximaciones $\ell_{2\omega}$, $L(A; x)_\omega$ es el único elemento del espacio de funciones

aproximantes L de orden $n+1$, que satisface

$$\sum_{i=1}^m \omega_i (f(x_i) - L(A; x_i)) \cdot L(A; x_i) = 0$$

para toda $L(A; x) \in L$.

Para $j \in J_E$ se tiene

$$|r_j| = \operatorname{sgn}(r_j) \cdot r_j = \operatorname{sgn}(r_j^*) \cdot r_j = \tau^* \cdot r_j / r_j^*,$$

donde

$$\tau^* = \|f(x) - L(A^*; x)\|_\infty$$

y de aquí, por el algoritmo de Lawson,

$$\omega_j^* = \alpha^{-1} \omega_j |r_j| = (\tau^*/\alpha) \cdot \omega_j r_j / r_j^*$$

donde $\alpha = \sum_{i=1}^m \omega_i |r_i|$. De aquí, para cada $L(A; x) \in L$

$$\begin{aligned} \sum_{j=1}^m \omega_j^* (f(x_j) - L(A^*; x_j)) \cdot L(A; x_j) &= \sum_{j \in J_E} \omega_j^* r_j^* L(A; x_j) = \\ &= (\tau^*/\alpha) \sum_{j \in J_E} \omega_j r_j L(A; x_j) = (\tau^*/\alpha) \sum_{j=1}^m \omega_j r_j L(A; x_j) = 0 \end{aligned}$$

lo cual implica que $L(A^*; x) = L(A; x)_{\omega^*}$,

Corolario 3.2 [5] Si para alguna k , $\omega_j^k = 0$ para toda $j \in J_0$

y $\operatorname{sgn} r_j^k = \operatorname{sgn} r_j^*$, para toda $j \in J_E$, entonces

$$L(A; x)_{\omega^{\ell}} = L(A^*; x)$$

para toda $\ell \geq k+1$

Demostración. Es inmediata de la definición del algoritmo

Finalmente se tiene el siguiente corolario que establece la convergencia del algoritmo de Lawson en un número finito de pasos a partir de cierta etapa.

Corolario 3.3 [5] Si para alguna N , $\omega_j^N = 0$ para toda $j \in J_0$, el algoritmo converge en un número finito de pasos.

Demostración. Es claro que para $k \geq N$ $\omega_j^k = 0$ para toda $j \in J_0$. De término $N_1 > N$ tal que $k \geq N_1$ implique

$$\|L(A_k; x) - L(A^*; x)\|_{\infty} < \tau^*$$

Entonces

$$\|r^k - r^*\|_{\infty} = \|(f(x) - L(A_k; x)) - (f(x) - L(A^*; x))\|_{\infty} < \tau^*$$

por lo que, como $|r_j^*| = \tau^*$ para $j \in J_E$,

$$\operatorname{sgn} r_j^k = \operatorname{sgn} r_j^* \quad \text{para } j \in J_E$$

y el corolario anterior nos garantiza que $L(A_k; x) = L(A^*; x)$

para $k \geq N_1 + 1$.

3.4 CRITERIOS PARA ACELERAR LA CONVERGENCIA DEL ALGORITMO DE LAWSON

Como resultado del teorema y los corolarios de la sección anterior se tiene que, para acelerar y alcanzar la convergencia del algoritmo de Lawson en un número finito de pasos, es necesario detectar los pesos ω_j^k tal que $j \in J_0$ y hacerlos iguales a cero. Para esto se utilizan criterios que se describen a continuación.

El primer criterio [6] consiste en hacer $\omega_j^k = 0$ si

$$|f(x_j) - L(A_k; x_j)_{\omega^k}| \leq \frac{(\sum_{i=1}^n \omega_i (r_i^k)^2)}{\max_{x_i \in X} |f(x_i) - L(A_k; x_i)_{\omega^k}|}$$

Esta comparación conviene efectuarla, por lo general, después de cada tres iteraciones.

Aun con este criterio muchas veces no es posible acelerar la convergencia en forma notable, ya que las cantidades $|f(x_i) - L(A; x_i)_{\omega}|$ varían muy poco en cada iteración a medida que se hacen las $\omega_i^k = 0$. Se ha recurrido a un nuevo criterio de aceleración, el cual ha sido utilizado en algunos ejemplos que se presentan en el siguiente capítulo y ha dado buenos resultados. El criterio consiste en hacer $\omega_j^k = 0$ si

$$|f(x_j) - L(A_k; x_j)_{\omega_j^k}| = \min_{x_i \in X'} |f(x_i) - L(A_k; x_i)_{\omega_j^k}|$$

donde

$$X' = \{ x_i \in X: \omega_j^k \neq 0 \}$$

Este criterio se ha aplicado cuando, después de aplicar el criterio anterior, ninguna ω_j^k se hace igual a cero y está relacionado con la caracterización de las aproximaciones uniformes relativa al número de puntos en el que se alcanza el error máximo, el cual es igual a $n+2$ si el orden de la función aproximante es $n+1$.

Por lo que se refiere a la convergencia del algoritmo, se ha demostrado [5] que converge en forma lineal y que el factor de convergencia es a lo más

$$\rho = \max \{ |f(x) - L(A^*; x)| : x \in E \} / \tau^*$$

3.5 ALTERNATIVAS

En [4] se demuestra que la función aproximante $L(A; x)$ obtenida con este algoritmo es válida solamente para un subconjunto de $X = \{ x_1, \dots, x_m \}$; sin embargo, se menciona que en la práctica es raro que esto suceda, ya que normalmente esta función sí es válida sobre todo el conjunto X . En adición se proporciona una alternativa para estos casos raros, la cual consiste en calcular los pesos iniciales ω_j^1 a partir de los últimos pesos obtenidos en un

intento anterior. La forma de calcular los pesos iniciales para estos casos raros está basada en los siguientes teoremas, de los cuales se omite la demostración.

Teorema 3.5 [4] La sucesión $L(A_k; x)$ converge a $L(A_0; x)$ la cual es la aproximación uniforme a $f(x)$ sobre un conjunto $X' \subset X$. La sucesión $\{\sigma^k\}$

$$\sigma^k = \left[\sum_{i=1}^k \omega_i^k (f(x) - L(A_k; x))^2 \right]^{1/2}$$

es monótona creciente y

$$\lim_{k \rightarrow \infty} \sigma^k = \max_{x \in X'} |f(x) - L(A_0; x)| = \sigma^*$$

Teorema 3.6 [4] Si X' es un subconjunto propio de X entonces el algoritmo de Lawson puede reiniciarse con los pesos

$$\bar{\omega}_i^1 = (1-\lambda) \lim_{k \rightarrow \infty} \omega_i^k + \lambda \mu(x), \quad 0 \leq \lambda < 1$$

donde $\mu(x) = 0$ para $x \neq z$ y $\mu(z) = 1$,

donde $z \in X - X'$ y $|f(z) - L(A_0; z)| > \sigma^*$.

Para λ suficientemente pequeña $\bar{\sigma}^1 > \sigma^*$ y después de un número finito de reinicios del algoritmo se obtiene la aproximación uniforme a $f(x)$ sobre X .

4. REALIZACION DEL ALGORITMO DE LAWSON

En este capítulo se describe la técnica utilizada en la realización del algoritmo de Lawson por medio de un programa de computadora. Siendo la parte principal del algoritmo la solución de problemas de aproximación de mínimos cuadrados pesados, se dedica la mayor parte a la descripción de un método eficaz en la solución de estos problemas.

4.1 MATRICES ORTOGONALES Y APROXIMACION DE MINIMOS CUADRADOS

Una matriz cuadrada Q es ortogonal si

$$Q^T Q = Q Q^T = I$$

es decir, la inversa de una matriz ortogonal es su transpuesta.

Una propiedad importante de las matrices ortogonales es que conservan la longitud de la norma ℓ_2 bajo la multiplicación. Esto es, para cualquier vector \bar{y} de dimensión m y cualquier matriz ortogonal Q de $m \times m$,

$$\| Q \bar{y} \|_2 = \| \bar{y} \|_2$$

Esto se sigue de

$$\begin{aligned} \| Q \bar{y} \|_2 &= \langle Q \bar{y}, Q \bar{y} \rangle^{1/2} = (\bar{y}^T Q^T Q \bar{y})^{1/2} = (\bar{y}^T I \bar{y})^{1/2} = (\bar{y}^T \bar{y})^{1/2} = \\ &= \langle \bar{y}, \bar{y} \rangle^{1/2} = \| \bar{y} \|_2 \end{aligned}$$

A continuación se presenta un tipo de transformación que da lugar a una matriz ortogonal; dicha transformación se conoce como "reflexión de Householder" y es de la forma

$$P = I - \frac{1}{\beta} uu^T$$

donde

I es la matriz identidad $m \times m$

u es un vector de dimensión m

distinto del vector nulo

$$\beta = \frac{1}{2} \|u\|_2^2 = \frac{u^T u}{2}$$

Teorema 4.1 [7] Si P es una reflexión de Householder, entonces

$$P = P^{-1} = P^T$$

es decir, P es una matriz ortogonal.

Demostración

$$P^T = (I - \frac{1}{\beta} uu^T)^T = P$$

$$\begin{aligned} PP^T &= (I - \frac{1}{\beta} uu^T) (I - \frac{1}{\beta} uu^T) = \\ &= I - \frac{2}{\beta^2} (\beta - uu^T/2) uu^T = I. \end{aligned}$$

Lo anterior se sigue de que uu^T es una matriz simétrica y de la igualdad $uu^T uu^T = (u^T u) uu^T$.

Si $\bar{a} \neq \bar{0}$ (vector cero), entonces es posible encontrar una reflexión P de Householder que haga igual a cero los componentes de \bar{a} con excepción del primero, esto es, $P\bar{a}$ es un múltiplo de $\bar{e}_1 = (1, 0, \dots, 0)$, lo cual se establece en el siguiente teorema.

Teorema 4.2 [7] dado $\bar{a} \neq \bar{0}$ definimos

$$\alpha = \text{sign}(a_1) \|\bar{a}\|_2$$

$$u = \bar{a} + \alpha \bar{e}_1$$

$$\beta = \alpha u_1$$

(a_1 y u_1 son los primeros componentes de \bar{a} y de u). Entonces P

$P = I - \frac{1}{\beta} u u^T$ es una reflexión de Householder que satisface

$$P\bar{a} = -\alpha e_1$$

Demostración

$$u^T u = (\bar{a} + \alpha \bar{e}_1)^T (\bar{a} + \alpha \bar{e}_1) =$$

$$= \alpha^2 + 2\alpha a_1 + \alpha^2 = 2\alpha (a_1 + \alpha) = 2\alpha u_1 = 2\beta$$

De aquí

$$\beta = \frac{(u^T u)}{2}$$

lo que demuestra que P es una reflexión de Householder.

Ahora, como $u^T \bar{a} = \beta$

$$\begin{aligned} P\bar{a} &= \left(I - \frac{1}{\beta} u u^T \right) \bar{a} = \bar{a} - \frac{1}{\beta} u (u^T \bar{a}) = \\ &= \bar{a} - u = -\alpha e_1 \end{aligned}$$

El siguiente teorema es fundamental en la solución de problemas de aproximación de mínimos cuadrados discretos.

Teorema 4.3 [7] Sea A una matriz $m \times n$ ($m > n$). Entonces existe una matriz ortogonal Q igual al producto de n reflexiones de Householder,

$$Q = P_n \dots P_2 P_1,$$

tal que

$$QA = R = \left[\begin{array}{c} \tilde{R} \\ \mathbf{0} \end{array} \right] \}_{ m}$$

n

donde \tilde{R} es una matriz triangular superior (es decir $r_{ij} = 0$ si $i > j$).

Demostración

La demostración es una descripción del algoritmo para calcular P_1, \dots, P_n

Sean $A_1 = A$ y \bar{a}_1 la primera columna de A_1 . Por el teorema anterior existe una reflexión de Householder P_1 tal que $P_1 \bar{a}_1$ es un múltiplo de \bar{e}_1 . Sean $A_2 = P_1 A_1$ y \bar{a}_2 la segunda columna de A_2 , entonces es posible formar una reflexión de Householder que hace igual a cero los componentes de \bar{a}_2 con excepción de los dos primeros y que no al-

tera a la primera columna de A_2 . Esto último es posible si se toma el vector u de la reflexión de Householder como

$$u = (0, a_{22} + \alpha, a_{32}, \dots, a_{m_2})$$

$$\alpha = \text{sign}(a_{22}) (a_{22}^2 + \dots + a_{m_2}^2)^{1/2}$$

$$\beta = \alpha u_2$$

Ahora, sea $A_3 = P_2 A_2$. Continuando en la misma forma, con el vector u adecuado, después de $n-1$ pasos tenemos una matriz

$$A_n = P_{n-1} \dots P_1 A$$

que es triangular superior excepto por su última columna \bar{a}_n . Sea P_n la reflexión de Householder que hace ceros los últimos $m-n$ componentes de \bar{a}_n y no altera ninguna de las columnas anteriores, esto es

$$u = (0, 0, \dots, a_{nn} + \alpha, a_{(n+1)n}, \dots, a_{mn})$$

Finalmente obtenemos

$$R = A_{n+1} = P_n P_{n-1} \dots P_1 A$$

lo cual termina la demostración.

El uso de este teorema en la solución de problemas de mínimos cuadrados está basado en la propiedad de las transformaciones ortogonales de conservar la norma vectorial euclidiana ℓ_2 . Haciendo

$$c = Qb,$$

entonces

$$\|Ax - b\|_2 = \|Q(Ax - b)\|_2 = \|Rx - c\|_2,$$

donde $R = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix}$, por lo que el problema $Ax = b$ se reduce a uno triangular, que se puede analizar de la manera siguiente:

Sean

\tilde{R} = los primeros n renglones de R

\tilde{c} = los primeros n componentes de c

por lo que \tilde{R} es una matriz cuadrada $n \times n$ triangular superior de rango igual al de A y \tilde{c} es un vector de dimensión n .

Para cualquier vector x

$$\|Rx - c\|_2^2 = \|\tilde{R}x - \tilde{c}\|_2^2 + c_{n+1}^2 + \dots + c_m^2$$

y el problema se reduce a resolver

$$\tilde{R}x = \tilde{c}$$

directamente por sustitución hacia atrás (se supone que rango $(A) = \text{rango}(\tilde{R}) = n$). Esta solución minimiza la suma de los cuadrados de los residuos la cual es igual a $c_{n+1}^2 + \dots + c_m^2$.

En la realización del algoritmo de Lawson por medio de un programa de computadora presentada en este trabajo, se utiliza el método de descomposición $QA = R$, descrito anteriormente, utilizando las subrutinas HECOMP y HOLVE [7]. La subrutina HECOMP sirve para ha-

cer la descomposición $QA = R$ y se utiliza la subrutina HOLVE para calcular Qb y hacer la sustitución hacia atrás para obtener la solución de cada aproximación $\mathcal{L}_{2\omega}$, la cual en cada caso es única ya que la matriz $A = (\phi_j(x_i))$, descrita en el capítulo anterior, es no-singular por ser $\{\phi_j\}$ un sistema de Chebyshev.

4.2 DESCRIPCION DEL PROGRAMA LAWSON

ENTRADA

· Lectura de datos.

M - número de puntos dados

(x_i, y_i) - puntos dados

ϕ_j - funciones aproximantes

$N+1$ - orden de las funciones aproximantes

TOL - tolerancia para aceptar convergencia

ITMAX - número máximo de iteraciones permitidas

CRITER - indica cada cuántas iteraciones se deben aplicar criterios de aceleración de la convergencia

CONT - lleva la cuenta del total de iteraciones

CONT1 - lleva la cuenta del número de iteraciones para aplicar criterios de aceleración.

INDICA - si INDICA = 0 los pesos iniciales se toman como

$$\omega_i^1 = 1/m, \quad i = 1, \dots, m$$

si INDICA = 1 los pesos iniciales los proporciona el usuario.

$W(1)$ - pesos iniciales, si INDICA = 1

PROCESO

1) Inicialización de variables

CONT = 0, CONT1 = 0

Si INDICA = 0 se hacen $\omega(I) = 1/m$, $I = 1, \dots, m$

2) CONT = CONT + 1, CONT1 = CONT1 + 1

Se forma el sistema $WA\bar{x} = W\bar{b}$

$W = \text{diag}(\sqrt{\omega(I)})$

$A = (\phi_j(x_i))$, $i=1, \dots, m$; $j = 1, \dots, n+1$

$\bar{x} = (a_0^*, \dots, a_n^*)$ (valores a calcular)

$\bar{b} = (y_1, \dots, y_m)$

3) Se resuelve el sistema $WA\bar{x} = W\bar{b}$ llamando a las subrutinas:

HECOMP - Aplica reflexiones de Householder a la matriz WA.

Se le pasan los parámetros:

MDIM - dimensión declarada para los renglones de A.

M - número de renglones de A.

N+1 - número de columnas de A.

WA - entrada - matriz a reducir

salida - matriz reducida e información sobre la reducción

U - entrada - soslayar

salida - información sobre la reducción

HOLVE - calcula la solución de mínimos cuadrados de sistemas sobre determinados, encontrando la \bar{x} que minimiza $\|WA\bar{x} - W\bar{b}\|$.

Se le pasan los parámetros:

MDIM, M, N, A, U, resultados de HECOMP

\bar{b} - vector de dimensión m

entrada - $\bar{b} = (y_1, \dots, y_m)$

salida - primeras $n+1$ componentes = la solución

$\bar{x} = (a_0^*, \dots, a_n^*)$; últimas $m - (n+1)$ componentes
= residual transformado.

4) Se calcula

$$\sigma_k^2 = \left[\sum_{i=1}^m \omega_i^k (L(A_k^*; x_i) - y_i)^2 \right]$$

donde $L(A_k^*; x_i) = \sum_{j=1}^{n+1} a_j^* \phi_j(x_i)$ es la aproximación correspondiente a los pesos $\omega_1^k, \dots, \omega_m^k$

5) Para $k > 1$:

$$\text{Si } \left| \sigma_k^2 - \sigma_{k-1}^2 \right| < \text{TOL} ,$$

parar las iteraciones e ir a SALIDA

en caso contrario ir a 6)

Si $k=1$: ir directamente a 6)

6) Si $\text{CONT1} = \text{CRITER}$, hacer $\text{CONT1} = 0$

Si el número de $\omega_i^k \neq 0$ es mayor que $n+1$,

6a) Aplicar el primer criterio de aceleración:

$$\omega_j^k = 0 \quad \text{si}$$

$$\left| y_j - L(A_k^*; x_j) \right| \leq \frac{\left(\sum_{i=1}^m \omega_i (y_i - L(A_k^*; x_i))^2 \right)}{\max_{x_i \in X^*} |y_i - L(A_k^*; x_i)|}$$

para $j = 1, \dots, m$, $X^* = (x_1, \dots, x_m)$

Si ninguna se hizo igual a cero en esta aplicación del criterio, ir a 6b); en caso contrario ir a 7).

6b) Aplicar el segundo criterio de aceleración:

$$\omega_j^k = 0 \text{ si}$$

$$|y_j - L(A_k^*; x_j)| = \min_{x_i \in X'} |y_i - L(A_k^*; x_i)|$$

donde

$$X' = \{x_i \in X : \omega_i^k \neq 0\}$$

ir a 7)

Si al empezar el paso 6, $CONT_1 \neq CRITER$ o el número de

$\omega_i^k \neq 0$ no es mayor que $n+1$, ir directamente a 8).

7) Se actualiza la información manejada de tal forma que los puntos (x_i, y_i) , correspondientes a los pesos ω_i^k que se hicieron iguales a cero, ya no participen en los siguientes cálculos; también se actualiza el valor de m (# de puntos dados), el cual es ahora

$m -$ (# de pesos ω_i^k que se igualaron a cero en la última aplicación de los criterios)

8) Se calculan los nuevos pesos ω_i^{k+1} como

$$\omega_i^{k+1} = \frac{\omega_i^k |L(A_k^*; x_i) - y_i|}{\sum_{j=1}^m \omega_j^k (L(A_k^*; x_j) - y_j)^2}$$

ir a 2):

SALIDA

- Puntos dados (x_i, y_i) , $i=1, \dots, m$
- Valor calculado $L(A^*; x_i)$ en cada punto x_i
- Error de la aproximación en cada punto $y_i - L(A^*; x_i)$
- Error máximo, $\|y - L(A^*; x)\|_\infty$
- Coeficientes a_i^* del polinomio aproximante
- Pesos finales ω_i diferentes de cero
- Número total de iteraciones

5. RESULTADOS Y CONCLUSIONES

5.1 RESULTADOS

En esta sección se presentan resultados numéricos de aproximación uniforme obtenidos con el programa LAWSON, el cual es la realización por medio de un programa de computadora del algoritmo de Lawson. Los ejemplos que se muestran han sido obtenidos de artículos referentes a la aproximación uniforme.

En cada caso se muestran los puntos dados, los valores calculados, el error absoluto en cada punto, el error máximo óptimo, el número de iteraciones con el que se logró la convergencia, los coeficientes del polinomio aproximante, así como los pesos finales diferentes de cero que corresponden a los puntos en donde se alcanza el error máximo.

En el capítulo 3 se menciona que la aproximación uniforme $L(A^*; x)$ obtenida con el algoritmo de Lawson es válida sólo en un subconjunto de X ; a pesar de esto, se ha observado [6] que en la práctica la aproximación obtenida sí es válida en todo el conjunto X , ocurriendo lo contrario muy rara vez. Realmente esto último sucede cuando se hace un peso $\omega_i^k = 0$ bajo condiciones diferentes a las especificadas en los criterios de aceleración de la convergencia. Por ejemplo, cuando una aproximación $L(A; x)$ interpola a alguno de los puntos dados. Uno de estos casos se muestra en el ejemplo 4.

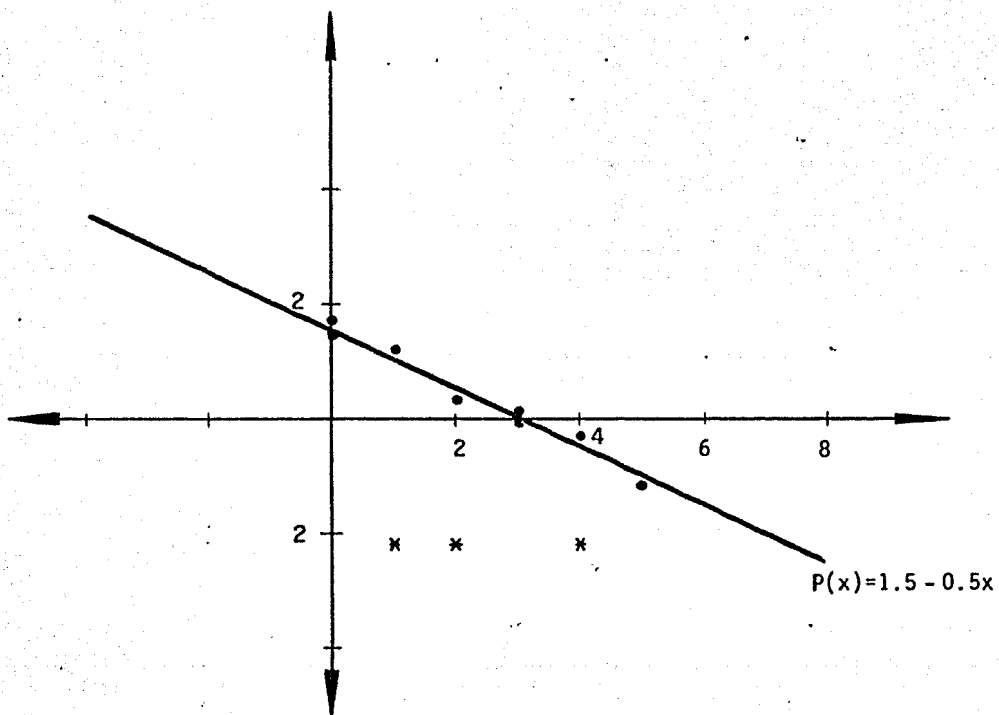
Ejemplo 5.1 [8] Este ejemplo corresponde a una aproximación polinomial de orden 2 con $\phi_0(x) = 1$ y $\phi_1(x) = x$. La aproximación es

sobre cinco puntos dados (x_i, y_i) . Como se observa (Fig 5.1) el error máximo se alcanza en tres de los seis puntos dados. En la Tabla I se muestran los resultados obtenidos sin utilizar criterios de aceleración, mientras que en la Tabla II se muestran los resultados obtenidos con los dos criterios de aceleración mencionados anteriormente (Cap. 3). En el primer caso se paró la iteración cuando

$$\left| \frac{\phi^k - \text{error máx}}{\text{error máx}} \right|^* < .001 . \text{ Como se observa en la Tabla I, para}$$

el primer caso se requieren 29 iteraciones, mientras que para el segundo, únicamente 6 (Tabla II).

$$^* \sigma^k = \left[\sum_{i=1}^m \omega_i^k (f_i - L(A_k; X_i))^2 \right]^{1/2}$$



- * ERROR MAXIMO
- PUNTOS DADOS

FIG 5.1

EU D
NNING 7598

*** R E S U L T A D O S ***

X	F(X)	L(A,X)	F(X)-L(A,X)
0.000000	1.520000	1.500000	0.020000
1.000000	1.025000	1.000000	0.025000
2.000000	0.475000	0.500000	-0.025000
3.000000	0.010000	-0.000000	0.010000
4.000000	-0.475000	-0.500000	0.025000
5.000000	-1.005000	-1.000000	-0.005000

ERROR MAXIMO = 0.025000 ITERACIONES = 29

*** C O E F I C I E N T E S ***

A(0) = 1.500000
A(1) = -0.500000

*** PESOS FINALES DIFERENTES DE CERO ***

W(2) = .333333E+00
W(3) = .500000E+00
W(5) = .166667E+00

=47.8 PT=0.3 IO=0.1

*** RESULTADOS ***

X	F(X)	L(A,X)	F(X)-L(A,X)
0.000000	1.520000	1.500000	0.020000
1.000000	1.025000	1.000000	0.025000
2.000000	0.475000	0.500000	-0.025000
3.000000	0.010000	0.000000	0.010000
4.000000	-0.475000	-0.500000	0.025000
5.000000	-1.005000	-1.000000	-0.005000

ERROR MAXIMO = 0.025000 ITERACIONES = 6

*** COEFICIENTES ***

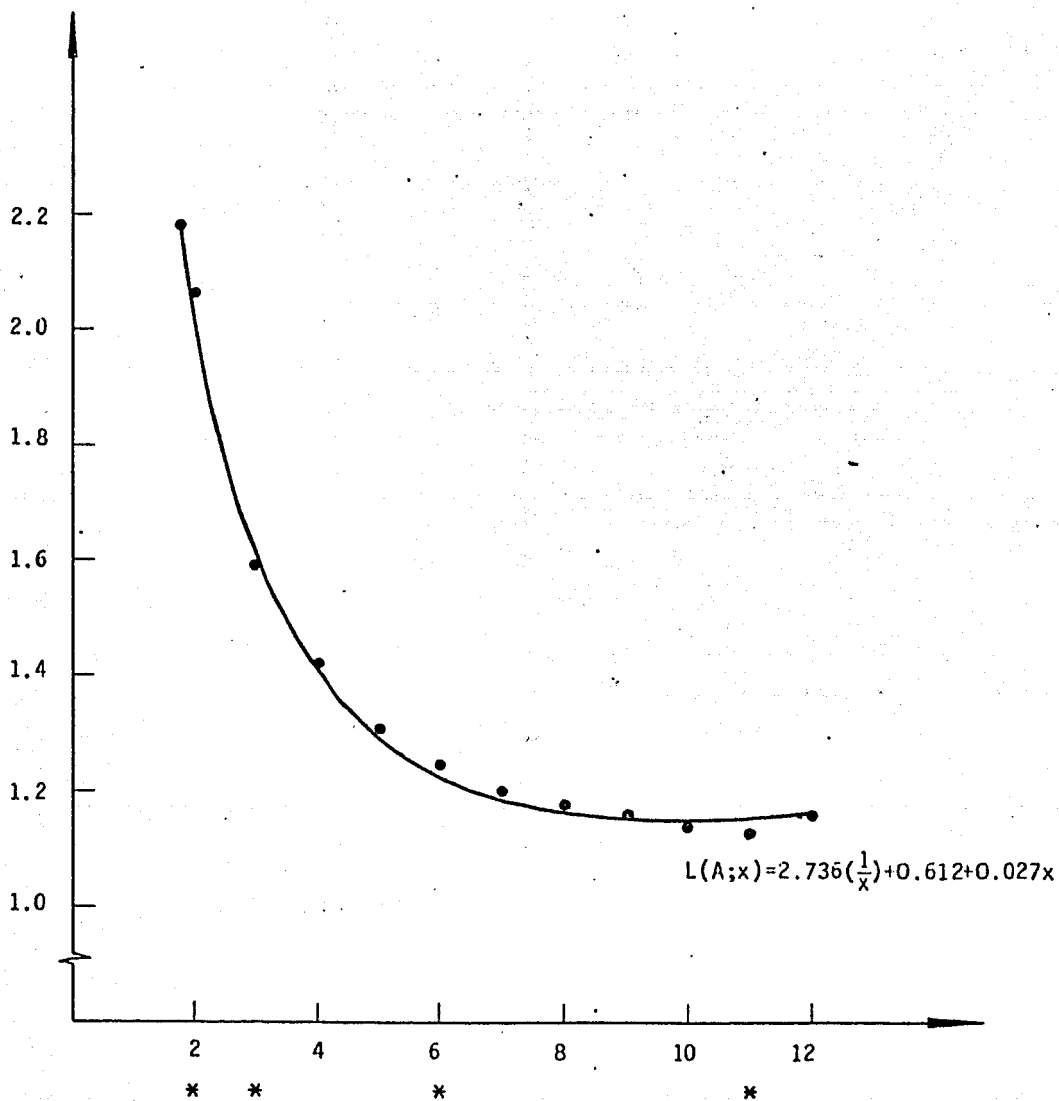
A(0) = 1.500000
 A(1) = -0.500000

*** PESOS FINALES DIFERENTES DE CERO ***

W(2) = .333333E+00
 W(3) = .500000E+00
 W(5) = .166667E+00

ET 37.3 PT=0.2 IO=0.2

Ejemplo 5.2 [9] Este ejemplo (Fig 5.2) corresponde a una aproximación utilizando funciones aproximantes de orden 3; las funciones ϕ_j son $\phi_0(x) = 1/x$, $\phi_1(x) = 1$ y $\phi_2(x) = x$. La aproximación es en este caso sobre diez puntos dados y el error máximo se alcanza en cuatro de los puntos dados, como se muestra en la Tabla III.



- * ERROR MAXIMO
- PUNTOS DADOS

Fig 5.2

*** R E S U L T A D O S ***

X	F(X)	L(A;X)	F(X)-L(A;X)
2.000000	2.057500	2.033392	0.024108
3.000000	1.580000	1.604108	-0.024108
4.000000	1.403700	1.402858	0.000842
5.000000	1.310500	1.292822	0.017678
6.000000	1.252500	1.228392	0.024108
7.000000	1.212900	1.190023	0.022877
8.000000	1.184000	1.167943	0.016057
9.000000	1.162100	1.156721	0.005379
10.000000	1.144800	1.153101	-0.008301
11.000000	1.130900	1.155008	-0.024108

ERROR MAXIMO = 0.024108 ITERACIONES = 6

*** C O E F I C I E N T E S ***

A(0) = 2.736407
A(1) = 0.611621
A(2) = 0.026784

*** PESOS FINALES DIFERENTES DE CERO ***

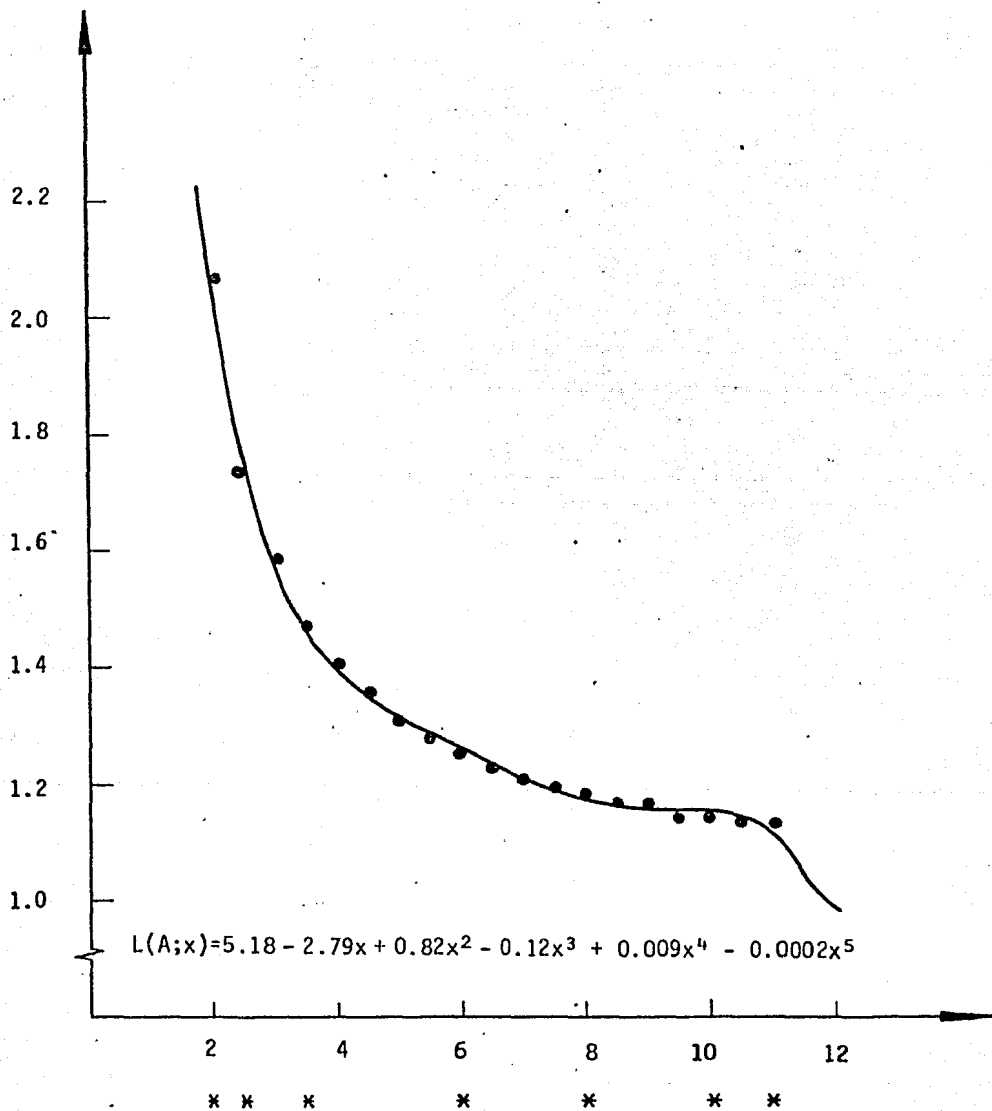
W(1) = .178571E+00
W(2) = .401786E+00
W(5) = .321429E+00
W(10) = .982143E-01

T=51.0 PT=0.3 IO=0.2

Ejemplo 5.3 En este caso (Fig 5.3) se presentan los resultados obtenidos bajo las siguientes condiciones:

- i) Corrida del programa utilizando sólo un criterio de aceleración, el de Rice [6]. Los resultados se muestran en la Tabla IV.
- ii) Corrida del programa utilizando los dos criterios de aceleración mencionados en el capítulo 3. Los resultados se muestran en la Tabla V.

El número de iteraciones requeridas para cada caso fue de 50 y 14, respectivamente. En este ejemplo la función aproximante fue un polinomio de grado 5 sobre 19 puntos dados.



- * ERROR MAXIMO
- PUNTOS DADOS

FIG 5.3

NNING 7725

*** RESULTADOS ***

X	F(X)	L(A,X)	F(X)-L(A,X)
2.000000	2.057500	2.044340	0.013160
2.500000	1.746000	1.759160	-0.013160
3.000000	1.580000	1.576147	0.003853
3.500000	1.475700	1.462540	0.013160
4.000000	1.403700	1.393112	0.010588
4.500000	1.351000	1.349263	0.001737
5.000000	1.310500	1.318120	-0.007620
5.500000	1.278500	1.291630	-0.013130
6.000000	1.252500	1.265660	-0.013160
6.500000	1.231000	1.239090	-0.008090
7.000000	1.212900	1.212913	-0.000013
7.500000	1.197400	1.189329	0.008071
8.000000	1.184000	1.170840	0.013160
8.500000	1.172400	1.159351	0.013049
9.000000	1.162100	1.155264	0.006836
9.500000	1.153000	1.156572	-0.003572
10.000000	1.144800	1.157960	-0.013160
10.500000	1.137500	1.149898	-0.012398
11.000000	1.130900	1.117740	0.013160

ERROR MAXIMO = 0.013160 ITERACIONES = 50

*** COEFICIENTES ***

A(0) = 5.176176
 A(1) = -2.787944
 A(2) = 0.819002
 A(3) = -0.120303
 A(4) = 0.008636
 A(5) = -0.000241

*** PESOS FINALES DIFERENTES DE CERO ***

W(1) = .144791E+00
 W(2) = .305821E+00
 W(4) = .228102E+00
 W(9) = .134035E+00
 W(13) = .105303E+00
 W(17) = .601441E-01
 W(19) = .218039E-01

PT=124.3 PT=1.0 IO=0.1

*** RESULTADOS ***

	X	F(X)	L(A;X)	F(X)-L(A;X)
1	2.000000	2.057500	2.044340	0.013160
2	2.500000	1.746000	1.759160	-0.013160
3	3.000000	1.580000	1.576147	0.003853
4	3.500000	1.475700	1.462540	0.013160
5	4.000000	1.403700	1.393112	0.010588
6	4.500000	1.351000	1.349263	0.001737
7	5.000000	1.310500	1.318120	-0.007620
8	5.500000	1.278500	1.291630	-0.013130
9	6.000000	1.252500	1.265660	-0.013160
10	6.500000	1.231000	1.239090	-0.008090
11	7.000000	1.212900	1.212913	-0.000013
12	7.500000	1.197400	1.189329	0.008071
13	8.000000	1.184000	1.170840	0.013160
14	8.500000	1.172400	1.159351	0.013049
15	9.000000	1.162100	1.155264	0.006836
16	9.500000	1.153000	1.156572	-0.003572
17	10.000000	1.144800	1.157960	-0.013160
18	10.500000	1.137500	1.149898	-0.012398
19	11.000000	1.130900	1.117740	0.013160

ERROR MAXIMO = 0.013160 ITERACIONES = 14

*** COEFICIENTES ***

- A(0) = 5.176176
- A(1) = -2.787944
- A(2) = 0.819002
- A(3) = -0.120303
- A(4) = 0.008636
- A(5) = -0.000241

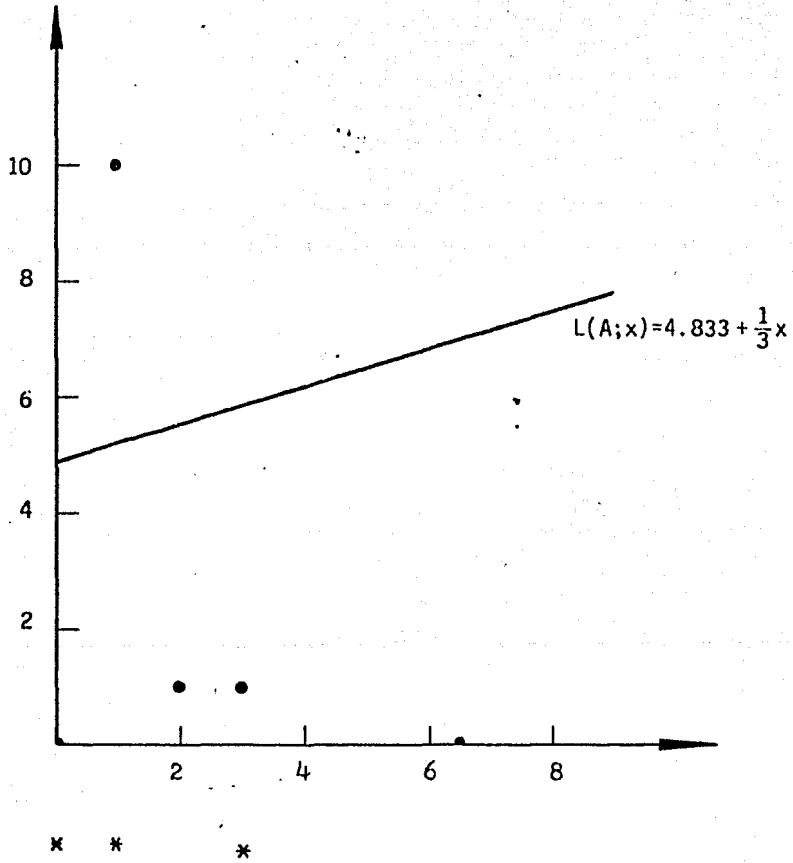
*** PESOS FINALES DIFERENTES DE CERO ***

- W(1) = .144791E+00
- W(2) = .305821E+00
- W(4) = .228102E+00
- W(9) = .134035E+00
- W(13) = .105303E+00
- W(17) = .601441E-01
- W(19) = .218039E-01

ET 1:17.9 PT=0.5 IO=0.2

Ejemplo 5.4 [4] Este ejemplo corresponde a uno de los casos raros en los que la aproximación que se obtiene en un primer intento no corresponde a la aproximación uniforme sobre todo el conjunto X , sino que es válida en un subconjunto de X . Esta aproximación se obtiene utilizando un polinomio de grado 1 sobre cinco puntos dados.

En el primer intento (Fig 5.4) la aproximación obtenida es válida sólo en el subconjunto $\{ (0,0), (1,10), (2,1), (3,1) \}$, obtenida utilizando como pesos iniciales $\omega_i = 1/m$, para $i = 1, \dots, m$. Sin embargo, en un segundo intento (Fig 5.5) se obtiene la función aproximante que es válida en todo el conjunto X , utilizando como pesos iniciales $\omega_1 = 0.25$, $\omega_2 = 0.375$, $\omega_3 = 0$, $\omega_4 = 0.125$ y $\omega_5 = 0.25$, obtenidos aplicando el teorema 3.6 .



- * ERROR MAXIMO
- PUNTOS DADOS

FIG 5.4

RUNNING 4702

*** R E S U L T A D O S ***

	X	F(X)	L(A;X)	F(X)-L(A;X)
1	0.000000	0.000000	4.833333	-4.833333
2	1.000000	10.000000	5.166667	4.833333
3	2.000000	1.000000	5.500000	-4.500000
4	3.000000	1.000000	5.833333	-4.833333
5	6.500000	0.000000	7.000000	-7.000000

ERROR MAXIMO = 7.000000 ITERACIONES = 5

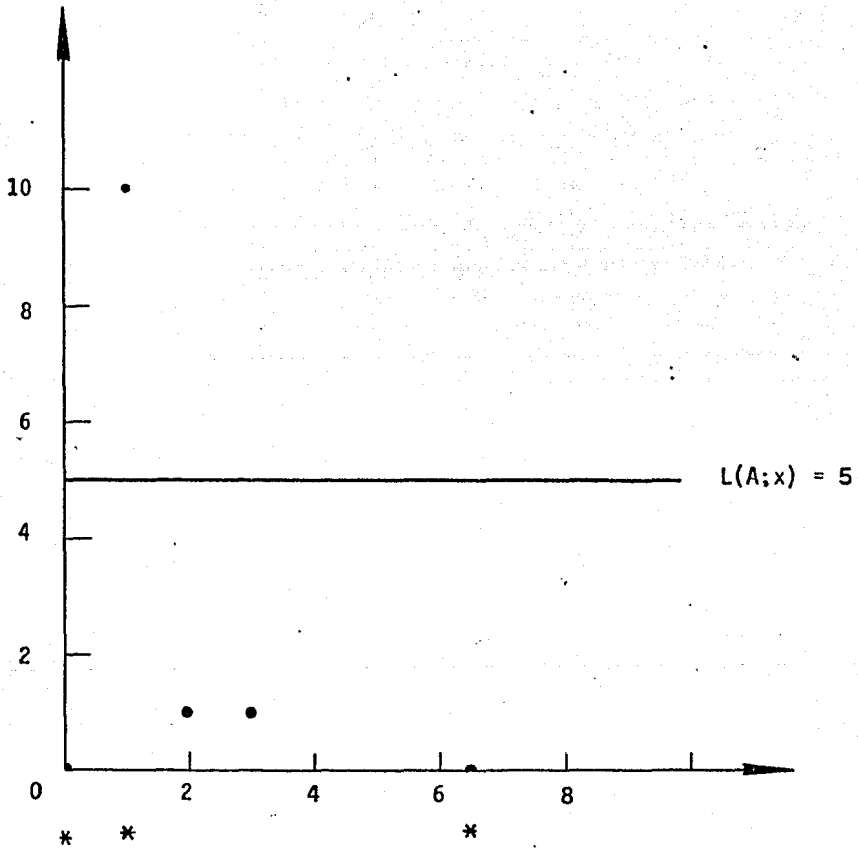
*** C O E F I C I E N T E S ***

A(0) = 4.833333
 A(1) = 0.333333

*** PESOS FINALES DIFERENTES DE CERO ***

W(1) = .333333E+00
 W(2) = .500000E+00
 W(4) = .166667E+00

ET 35.5 PT=0.2 IO=0.2



- * ERROR MAXIMO
- PUNTOS DADOS

FIG 5.5

UNING 5304

*** R E S U L T A D O S ***

X	F(X)	L(A,X)	F(X)-L(A,X)
0.000000	0.000000	5.000000	-5.000000
1.000000	10.000000	5.000000	5.000000
2.000000	1.000000	5.000000	-4.000000
3.000000	1.000000	5.000000	-4.000000
6.500000	0.000000	5.000000	-5.000000

ERROR MAXIMO = 5.000000 ITERACIONES = 5

*** C O E F I C I E N T E S ***

A(0) = 5.000000
 A(1) = -0.000000

*** PESOS FINALES DIFERENTES DE CERO ***

W(1) = .423077E+00
 W(2) = .500000E+00
 W(5) = .769231E-01

T=5.9 PT=0.2 IO=0.2

5.2

CONCLUSIONES

Como se ha mencionado en los capítulos anteriores, la principal ventaja del algoritmo de Lawson es su fácil realización por medio de un programa de computadora, por ser básicamente aproximaciones de mínimos cuadrados pesados.

Aunque en la teoría no está garantizada la convergencia a una función aproximante válida en todo el conjunto X , en la práctica esto último es lo que normalmente sucede, lo cual hace atractivo el método. Finalmente, en el apéndice se muestra el listado del programa LAWSON y el diagrama de flujo correspondiente, en la Fig 5.6 .

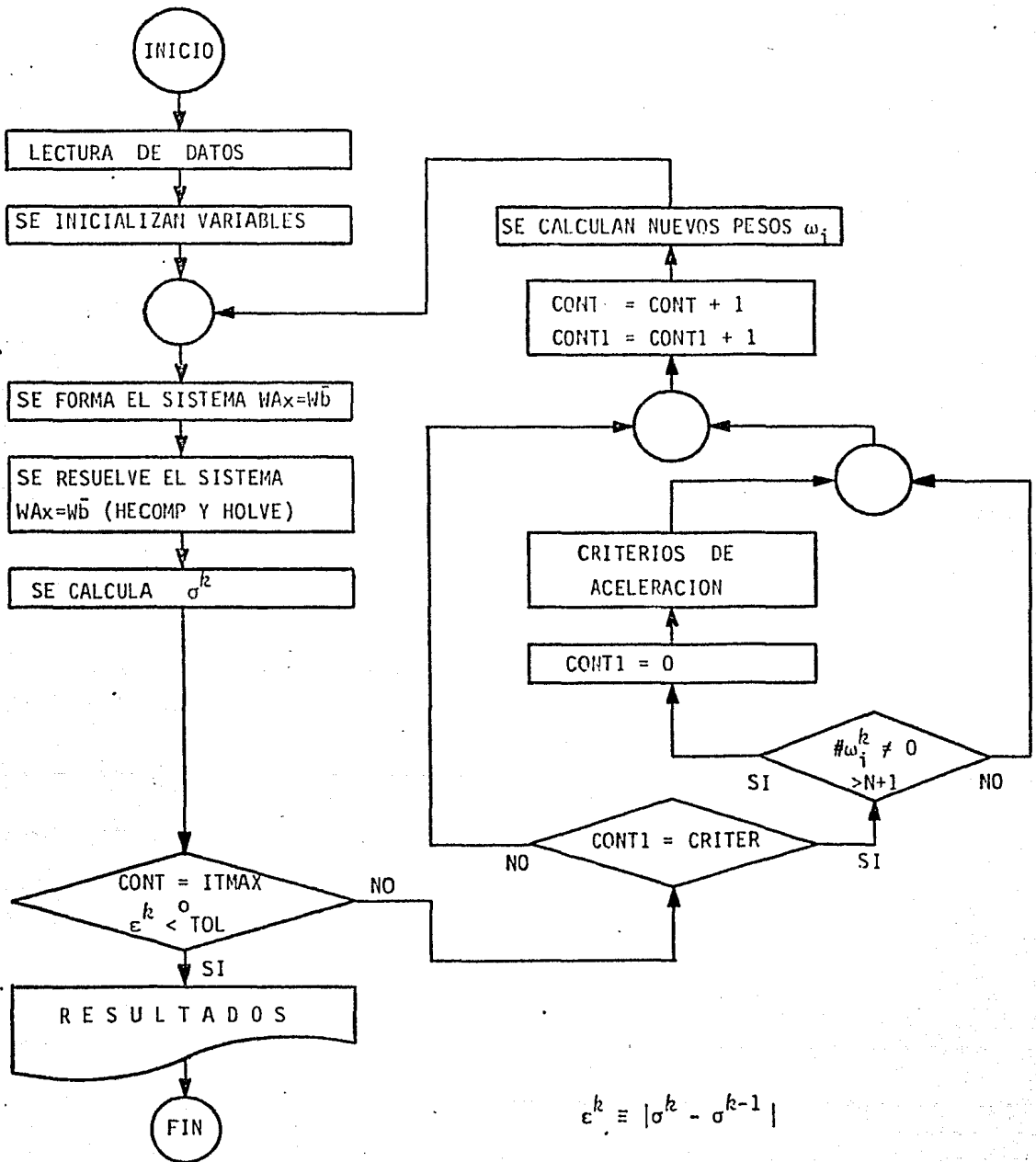


FIG 5.6 Diagrama de flujo del programa LAWSON

BIBLIOGRAFIA

- [1] P.J. Davis, Interpolation and Approximation, Dover Publications, New York, 1963
- [2] G.G. Lorentz, The Approximations of Functions, New York, Holt, Rinehart and Winston, 1966
- [3] G.W. Stewart, Introduction to Matrix Computations, Academic Press, 1973
- [4] C.L. Lawson, Contribution to the Theory of Linear Least Maximum Approximations, Tesis Doctoral, UCLA, 1961
- [5] A.K. Cline, "Rate of Convergence of Lawson's Algorithm", Math. Comp., Vol. 26, pp. 167-176, 1972
- [6] J.R. Rice & K.H. Usow, "The Lawson Algorithm, and Extensions", Math. Comp., Vol. 22, pp. 118-127, 1968
- [7] C.B. Moler, Matrix Eigenvalue and Least Squares Computations, Stanford University, 1973
- [8] I. Barrodale & A. Young, "Algoritms for Best l_1 and l_∞ Linear Approximations on a Discrete Set", Numerische Mathematik 8, pp. 295-306, 1966
- [9] M.A. Murray Lasso, Programa Mejorado para Aproximar Funciones Discretas Minimizando Desviaciones Maximas por Programacion Lineal, Reporte DEPMI-UNAM, Mexico, 1980.

A P E N D I C E

LISTADO DE COMPUTADORA DEL PROGRAMA LAWSON


```

L
100 $RESET FREE
200 FILE 5=DAT005, UNIT=DISK,RECORD=14,BLOCKING=42
300 FILE 6=SAL,UNIT=REMOTE,RECORDSIZE=135
400 C *****
500 C *
600 C *
700 C *
900 C *
905 C *
1000 C *
1100 C *
1200 C *
1300 C *
1400 C *
1500 C *
1600 C *
1700 C *
1800 C *
1900 C *
1905 C *
1910 C *
2000 C *****
2010 C
2100 C
2105 C
2200 C
2300 C
2350 C
2355 C
2360 C
2365 C
2400 C
2500 C
2600 C
2700 C
2800 C
2802 C
2804 C
2806 C
2808 C
2810 C
2812 C
2814 C
2816 C
2818 C
2820 C
2822 C
2824 C
2826 C
2828 C
2900 C
3000 C
3100 C
3102 C
3104 C
3106 C
3110 C
3120 C
3130 C
3140 C
3200 C
3300 C

```

L A W S O N

ESTE PROGRAMA CALCULA LA APROXIMACION UNIFORME A UNA
 FUNCION DISCRETA COMO UN LIMITE DE APROXIMACIONES DE MI -
 NIMOS CUADRADOS PESADOS (ALGORITMO DE LAWSON).

CADA APROXIMACION DE MINIMOS CUADRADOS PESADOS SE CALCULA
 RESOLVIENDO UN SISTEMA LINEAL SOBREDETERMINADO $AX = B$ EN
 EL SENTIDO DE MINIMOS CUADRADOS POR MEDIO DE REFLEXIONES
 DE HOUSEHOLDER (SUBROUTINAS "HECOMP" Y "HOLVE").

ENERO, 1983 FERNANDO ANGELES

DOUBLE PRECISION A(30,10), U(30), B(30), X(30), S(30)
 DOUBLE PRECISION W(30), Y(30), SIGMK1
 DOUBLE PRECISION X1(30), Y1(30), C(30), UMIN, UMAX, SIGMAK
 DOUBLE PRECISION VLAMB0, SUMA1, XC, TOL
 INTEGER V1(30), CRITER, CONT, CONT1

*** E N T R A D A ***

N = ORDEN DE LA FUNCION APROXIMANTE

M = NUMERO DE PUNTOS DADOS

TOL = TOLERANCIA PARA ACEPTAR CONVERGENCIA
 (ERROR ABSOLUTO)

ITMAX = NUMERO MAXIMO DE ITERACIONES PERMITIDAS

CRITER = INDICA CADA CUANTAS ITERACIONES SE
 DEBEN APLICAR CRITERIOS DE ACELERACION

INDICA - SI INDICA = 0 LOS PESOS INICIALES SON
 IGUALES A 1.0
 INDICA = 1 LOS PESOS SE PROPORCIONAN
 EXTERNAMENTE

READ(5,/)M, N, TOL, ITMAX, CRITER, INDICA
 READ(5,/)(X(I), I=1,M)
 READ(5,/)(Y(I), I=1,M)

INICIALIZACION DE VARIABLES

SIGMAK = 0.0
 SIGMK1 = 0.0
 CONT1 = 0
 CONT = 0
 NO = M
 DO 10 I=1,M

```

3500      Y1(I) = Y(I)
3600 10 CONTINUE
3610      DO 18 I=1,M
3620      V1(I) = I
3630 18 CONTINUE
3700      NP1 = N
3705      NP2 = N + 1
3800      N1 = N - 1
3900 C
4000 C      LOS PESOS INICIALES SON IGUALES A 1.0 SI INDICA = 0
4100 C
4105      IF(INDICA .EQ. 1) GO TO 15
4200      DO 40 I=1,M
4300      W(I) = 1.0
4400 40 CONTINUE
4500      GO TO 45
4502 C
4504 C      SE LEEN LOS PESOS INICIALES SI INDICA = 1
4506 C
4510      15 READ(5,/) (W(I), I=1,M)
4700 45 MM = M
4705      CONT1 = CONT1 + 1
4710      CONT = CONT + 1
4800 C
4900 C      SE FORMA LA MATRIZ 'WA' DEL SISTEMA WAX = WB
5000 C
5100      DO 50 I=1,M
5200      A(I,1) = DSQRT(W(I))
5300      DO 50 J=2,NP1
5400      A(I,J) = X(I) * A(I,J-1)
5500 50 CONTINUE
5600 C
5700 C      SE FORMA EL VECTOR 'WB' DEL SISTEMA WAX = WB
5800 C
5900      69 DO 70 I=1,M
6000      B(I) = DSQRT(W(I)) * Y(I)
6100 70 CONTINUE
6200 C
6300 C      SE APLICAN REFLEXIONES DE HOUSEHOLDER A LA MATRIZ 'WA'
6400 C
6500      CALL HECOMP(30,M,NP1,A,U)
6600 C
6700 C      SE APLICAN REFLEXIONES DE HOUSEHOLDER A EL VECTOR 'WB'
6800 C      Y SE CALCULA LA SOLUCION 'X'
6805      CALL HOLVE(30,M,NP1,A,U,B)
6900 C
7000 C      CALCULO DE ABS(L(A,X) - F(X)), L(A,X) FUNCION AFROXIMANTE
7200 C
7300      DO 105 I=1,M
7400      S(I) = B(NP1)
7500      DO 105 J=1,N1
7600      S(I) = S(I) * X(I) + B(NP1-J)
7700 105 CONTINUE
7800      DO 110 I=1,M
7900      S(I) = DABS(Y(I) - S(I))
8000 110 CONTINUE
8100      SUMA1 = 0.0
8200      SUMA2 = 0.0
8210      SIGMAK1 = SIGMAK
8220      SIGMAK = 0.0
8230 C
8240 C      SE CALCULA SIGMAK**2
8250 C
8260      DO 135 I=1,M
8270      SIGMAK = SIGMAK + W(I) * (S(I) ** 2.0)

```

```

8280 135 CONTINUE
8285 IF(SIGMAK .LE. .000001 .OR. SIGMAK .LE. TOL) GO TO 878
002 IF(CONT .EQ. 1) GO TO 17
004 IF(DABS(SIGMAK - SIGMK1) .LE. TOL) GO TO 878 70.
8410 C
8420 C CRITERIO PARA TERMINAR DE ACUERDO AL NUMERO DE ITERACIONES
8430 C
8440 17 IF(CONT .GE. ITMAX) GO TO 878
8500 C
8600 C CRITERIO PARA APLICAR ACELERACION DE CONVERGENCIA
8700 C
8800 IF(CONT1 .LE. CRITER) GO TO 145
8905 CONT1 = 0
9000 C
9400 C SE CALCULA EL VALOR MINIMO DE ABS(L(A,X) - F(X))
9500 C
9600 VMIN = S(1)
9700 K = 2
9800 II = 1
9900 112 IF(VMIN .LE. S(K)) GO TO 114
0000 II = K
0100 VMIN = S(K)
0200 114 K = K + 1
0300 IF(K .LE. M) GO TO 112
10310 C
10320 C SE CALCULA EL VALOR MAXIMO DE ABS(L(A,X) - F(X))
0330 C
0400 VMAX = S(1)
10500 K = 2
0600 120 IF(VMAX .GT. S(K)) GO TO 125
0700 VMAX = S(K)
10800 125 K = K + 1
0810 IF(K .LE. M) GO TO 120
1600 VLAMB = SIGMAK / VMAX
11700 C
11800 C SE APLICA EL PRIMER CRITERIO PARA ACELERAR LA CONVERGENCIA
11900 C
11910 IF(M .LE. NP2) GO TO 145
12000 DO 140 I=1,MM
2100 IF(S(I) .LE. VLAMB) W(I) = 0.0
2200 IF(W(I) .EQ. 0.0) M = M - 1
12300 140 CONTINUE
2400 C
2500 C SI NINGUNA W(I)=0.0 DESPUES DE APLICAR EL PRIMER CRITERIO
12600 C SE APLICA UN SEGUNDO CRITERIO
12700 C
2800 IF(M .NE. MM) GO TO 136
2810 IF(M .LE. NP2) GO TO 145
12900 W(II) = 0.0
3000 M = M - 1
3100 136 CONT1 = 0
15500 139 K = 1
3600 NCONT = 0
3700 C
15800 C SE ACTUALIZAN LOS INDICES
15900 C
4000 141 IF(W(K) .GT. 0.0) GO TO 143
4100 NCONT = NCONT + 1
16200 NCONT1 = MM - NCONT
4300 IF(K .GT. NCONT1) GO TO 145
4400 DO 142 I=K,NCONT1
16500 W(I) = W(I+1)
4600 X(I) = X(I+1)
4700 Y(I) = Y(I+1)
4800 S(I) = S(I+1)
16810 V(I) = V(I+1)

```

```

17000
17100
17200 143 K = K + 1
17300     IF (K .EQ. M+1) GO TO 145
17400     GO TO 141
17500 C
17600 C     CALCULO DE LOS NUEVOS PESOS W(I)
17700 C
17800 145 DO 152 I=1,M
17900     SUMA1 = SUMA1 + W(I) * S(I)
18000 152 CONTINUE
18100     DO 155 I=1,M
18200     W(I) = (W(I) * S(I)) / SUMA1
18300 155 CONTINUE
18305     GO TO 45
18405 C
18410 C     IMPRESION DE RESULTADOS
18415 C
18420 878 WRITE(6,200)
18423     WRITE(6,300)
18425     DO 138 I=1,M0
18430     S(I) = B(NP1)
18435     DO 137 J=1,N1
18440     S(I) = S(I) * X1(I) + B(NP1 - J)
18445 137 CONTINUE
18450     ERROR = Y1(I) - S(I)
18455     WRITE(6,500) I,X1(I), Y1(I), S(I), ERROR
18460 138 CONTINUE
18465     WRITE(6,600)UMAX,CONT
18470     WRITE(6,400)
18475     DO 144 I=1,NP1
18480     IN = I - 1
18485     WRITE(6,700) IN, B(I)
18490 144 CONTINUE
18495     WRITE(6,800)
18500     DO 160 I = 1,NP2
18505     III = V1(I)
18510     WRITE(6,900) III, W(I)
18515 160 CONTINUE
18570     RETURN
18575 C
18580 C     *** FORMATOS DE ESCRITURA ***
18585 C
18590 150 FORMAT(/,10F12.6)
18600 200 FORMAT(20X,"*** R E S U L T A D O S ***",//)
18700 300 FORMAT(8X," X ", " F(X) ", " L(A,X) ",
18800     *" F(X)-L(A;X) ",//)
18900 400 FORMAT(/,10X,"*** C O E F I C I E N T E S ***",//)
19000 500 FORMAT(2X,I2,4F13.6//)
19100 600 FORMAT(/,9X,"ERROR MAXIMO =",F12.6,5X,"ITERACIONES =",I3//)
19200 700 FORMAT(/,15X," A(",I2,") =",F12.6)
19205 800 FORMAT(/,5X,"*** PESOS FINALES DIFERENTES DE CERO ***",//)
19210 900 FORMAT(/,15X," W(",I2,") =",E12.6)
19300     END
19310 C
19320 C     *** S U B R U T I N A S ***
19330 C
19400     SUBROUTINE HECOMP(MDIM,M,N,A,U)
19500     INTEGER MDIM,M,N
19600     REAL*8 A(MDIM,N),U(M)
19700     REAL*8 ALPHA,BETA,GAMMA,SQRT
19800 C
19900 C     HOUSEHOLDER REDUCTION OF RECTANGULAR MATRIX TO UPPER
20000 C     TRIANGULAR FORM, USE WITH HOLVE FOR LEAST-SQUARE
20100 C     SOLUTIONS OF OVERDETERMINED SYSTEMS.

```

```

20300 C MDIM= DECLARED ROW DIMENSION OF A
20400 C M = NUMBER OF ROWS OF A
20500 C N = NUMBER OF COLUMNS OF A
20600 C A = M-BY-N MATRIX WITH M.>.N
20700 C INPUT :
20800 C MATRIX TO BE REDUCED
20900 C OUTPUT:
21000 C REDUCED MATRIX AND INFORMATION ABOUT REDUCTION
21100 C U = M-VECTOR
21200 C INPUT :
21300 C IGNORED
21400 C OUTPUT:
21500 C INFORMATION ABOUT REDUCTION
21600 C
21700 C FIND REFLECTION WHICH ZEROES A(I,K), I= K+1,.....,M
21800 C
21900 DO 6 K= 1,N
22000 ALPHA= 0.0
22100 DO 1 I= K,M
22200 U(I)= A(I,K)
22300 ALPHA= ALPHA+U(I)*U(I)
22400 1 CONTINUE
22500 ALPHA= DSQRT(ALPHA)
22600 IF(U(K).LT.0.0) ALPHA= -ALPHA
22700 U(K)= U(K)+ALPHA
22800 BETA= ALPHA*U(K)
22900 A(K,K)= -ALPHA
23000 IF(BETA.EQ.0.0.OR.K.EQ.N) GO TO 6
23100 C
23200 C APPLY REFLECTION TO REMAINING COLUMNS OF A
23300 KP1= K+1
23400 DO 4 J= KP1,N
23500 GAMMA= 0.0
23600 DO 2 I= K,M
23700 GAMMA= GAMMA+U(I)*A(I,J)
23800 2 CONTINUE
23900 GAMMA= GAMMA/BETA
24000 DO 3 I= K,M
24100 A(I,J)= A(I,J)-GAMMA*U(I)
24200 3 CONTINUE
24300 4 CONTINUE
24400 6 CONTINUE
24500 RETURN
24600 C
24700 C TRIANGULAR RESULT STORED IN A(I,J), I.LE.J
24800 C VECTORS DEFINING REFLECTIONS STORED IN U AND REST OF A
24900 END
25000 SUBROUTINE HOLVE(MDIM,M,N,A,U,B)
25100 INTEGER MDIM,M,N
25200 REAL*8 A(MDIM,N),U(M),B(M)
25300 REAL*8 BETA,GAMMA,T
25400 C
25500 C
25600 C LEAST-SQUARE SOLUTION OF OVERDETERMINED SYSTEMS
25700 C FIND X THAT MINIMIZES NORM(A*X-B)
25800 C
25900 C MDIM,M,N,A,U, RESULTS FROM HECOMP
26000 C B= M-VECTOR
26100 C INPUT :
26200 C RIGHT HAND SIDE
26300 C OUTPUT: FIRST N COMPONENTS = THE SOLUTION, X
26400 C LAST M-N COMPONENTS= TRANSFORMED RESIDUAL
26500 C DIVISION BY ZERO IMPLIES A NOT OF FULL RANK
26600 C
26700 C APPLY REFLECTIONS TO B

```

```
26800 C
26900 DO 3 K= 1,N
27000 T= A(K,K)
27100 BETA= -U(K)*A(K,K)
27200 A(K,K)= U(K)
27300 GAMMA= 0.0
27400 DO 1 I=K,M
27500 GAMMA= GAMMA+A(I,K)*B(I)
27600 1 CONTINUE
27700 GAMMA= GAMMA/BETA
27800 DO 2 I= K,M
27900 B(I)= B(I)-GAMMA*A(I,K)
28000 2 CONTINUE
28100 A(K,K)= T
28200 3 CONTINUE
28300 C
28400 C BACK SUBSTITUTION
28500 DO 5 KB= 1,N
28600 K= N+1-KB
28700 B(K)= B(K)/A(K,K)
28800 IF(K.EQ.1) GO TO 5
28900 KM1= K-1
29000 DO 4 I= 1,KM1
29100 B(I)= B(I)-A(I,K)*B(K)
29200 4 CONTINUE
29300 5 CONTINUE
29400 RETURN
29500 END
```