

PROGRAMACION LINEAL EN VALORES ABSOLUTOS APLICADA A LA RESOLUCION NUMERICA DE ECUACIONES DIFERENCIALES E INTEGRALES POR METODOS ANALOGOS A LOS DE RITZ-GALERKIN (*)

Por J. M. ANTON CORRALES

Ing. de Caminos, Canales y Puertos

Muchos problemas numéricos conducen a minimizar una suma o integral de cuadrados. Se propone un método alternativo usando una suma de valores absolutos, indicando cómo se resuelven por programación lineal en el caso lineal. Se indica cómo puede ser aplicado lo que precede a los métodos en diferencias finitas, a los de Ritz-Galerkin, al cálculo de errores y a la aproximación de curvas entre otros. El método es más interesante que el de mínimos cuadrados cuando hay que poner restricciones lineales.

1. INTRODUCCION

Varios problemas físicos o matemáticos muy frecuentes pueden resolverse aproximadamente minimizando una suma de cuadrados o la integral de cuadrados de funciones o funcionales reales. Entre ellos están los de mínimos cuadrados, los de Ritz-Galerkin, etc. (referencias 3, 5).

En muchos de estos problemas se puede hallar otra solución aproximada minimizando una suma de valores absolutos o una integral de valor absoluto que se reduce a suma para cálculo aproximado. Ello conduce a programas de programación en valores absolutos, muy fáciles de transformar si son lineales en programas lineales de resolución fácil con ordenador. Esta solución es del mismo orden de complicación que la de las ecuaciones a las que conduce el método de mínimos cuadrados.

En este artículo se indica cómo se pueden resolver bastantes de esos problemas por programación lineal. Aunque ha sido publicado algo sobre programación en valores absolutos (problemas $\text{Min } \sum_i C_i |x_i|$ con C_i positivos) en los Proceedings of the Operational Research Society, el autor de este trabajo no ha encontrado publicado este tema, aunque es probable que lo haya sido, vista su sencillez e interés.

2. PROGRAMACION LINEAL EN VALORES ABSOLUTOS

Denominamos programa al problema consistente en hallar el mínimo de una función $f(y^1, \dots, y^n) \equiv f(Y)$ con unas restricciones $g^1(Y) \geq 0, \dots, g^m(Y) \geq 0$ algún signo $\geq$ puede ser $=$). Como el mínimo de f es el máximo de $-f$, consideramos sólo mínimos sin pérdida de generalidad.

Se demuestra que si las funciones g y $-f$ son cóncavas (**) existe un conjunto convexo solución del problema, y suele este conjunto reducirse a un punto (2).

(*) Se admiten comentarios sobre el presente artículo, que pueden remitirse a la Redacción de esta Revista hasta el 31 de marzo de 1973.

(**) Si la restricción correspondiente a g es un signo $=$, g ha de ser lineal en general para que el conjunto sea convexo.

El cálculo del óptimo es fácil si las funciones g y f son lineales, y más complicado en caso contrario (6). En el primer caso existen programas (4) para su resolución para casi todo ordenador científico, por métodos derivados del de Dantzig o simples. Con muy pocas variables puede buscarse la solución operando a mano; en ordenador se puede llegar a cientos de variables y de restricciones si se dispone por lo menos de dos cintas magnéticas.

Denominamos programa en valores absolutos a uno del tipo:

$$\text{Min} \sum_{j=1}^N |f_j| + f^{N+1}$$

con restricciones $g^k \geq 0$ o $g^k = 0$.

Si existe solución, es la misma que la del problema:

$$\text{Min} \left(\sum_{j=1}^N (B_j - f_j) + f^{N+1} \right)$$

con las restricciones para todo j :

$$B_j \geq 2f_j, \quad B_j \geq 0,$$

además de las $g^k \geq 0$ o $g^k = 0$.

Ello es debido a que en el óptimo del segundo programa, las B_j deben tomar el mayor valor compatible con las restricciones, que es $2f_j$ si $f_j \geq 0$, o si $f_j \leq 0$. En ambos casos aparece $|f_j|$ en el mínimo. Es fácil, pues, demostrar que si uno de los problemas tiene solución, el otro la tiene y es para los mismos y^i y con el mismo valor del programa.

Si las f y g no son lineales el problema es de programación no lineal. No entramos en su resolución, pero puede haber alguna ventaja en utilizar el segundo programa para poder utilizar métodos de los gradientes: el gradiente de $\sum_j |f_j|$ es $\sum_j \pm \text{grad}(f_j)$ según el signo de f_j y el problema del signo está resuelto en el segundo programa.

En el caso de ser las f y las g lineales, el problema del tipo:

$$\text{Min} \left(\sum_{j=1}^N \left| a_{j0} + \sum_{i=1}^n a_{ji} y^i \right| + a^0_0 + \sum_{i=1}^n a^0_i y^i \right)$$

con m restricciones:

$$g^k = \sum_{i=1}^n b^k_i y^i - b^k_0 \geq 0$$

(o con $= 0$ en vez de ≥ 0) equivale al:

$$\text{Min} \left(\sum_{j=1}^N \left(B_j - a_{j0} - \sum_{i=1}^n a_{ji} y_i \right) + a_{j0} + \sum_{i=1}^n a_{ji} y_i \right)$$

con las restricciones $g_i^k \geq 0$ o bien $g^k = 0$ anteriores y las nuevas:

$$B_j \geq 2 \left(a_{j0} + \sum_{i=1}^n a_{ji} y_i \right),$$

con las $B_j \geq 0$.

Las y_i pueden ser cualesquiera, o no negativas. En el primer caso puede ser necesario hacer $y_i = \alpha'_i - \alpha''_i$, donde α' y α'' son no negativas, para introducir los datos en ordenador.

Se observa que si $a_{j0} \geq 0$ (lo que siempre se puede lograr cambiando de signo dentro del (| |)), las B_j pueden formar parte de la primera solución básica, sin necesidad de variables de penalización. Por ello, la primera matriz puede ser de rango $(n + N + m + (\text{variables de holgura para las } g^m \geq 0)) \times (N + m)$, aunque más automáticamente pueda ser de orden superior $(n + 2N + m + \text{variables de holgura}) \times (1 + N + m)$. Son $n + N$ variables con $N + m$ restricciones.

Si las a_{ji} son nulas, el problema está siempre acotado y tiene solución.

En cuanto a longitud de cálculo, se recalcula la matriz en cada paso, siendo el número de pasos limitado. Es del mismo orden de complicación que el de resolver N' ecuaciones con N' incógnitas, que requiere calcular $N' (1 + N') N'$ elementos de matriz aproximadamente.

En los apartados siguientes se plantean dos programas de este tipo.

3. APLICACION A LOS METODOS EN DIFERENCIAS FINITAS

Dada la aplicación $h(y)$ de una función definida en (a, b) , y $y(x)$, en el mismo espacio de funciones definidas en (a, b) , se aproximan $y(x)$ por $(y_0, y_1, \dots, y_n) = Y$ en n puntos consecutivos $x_0, x_1, \dots, x_n$, $h(y)$ por $H(Y) = (h_0, h_1, \dots, h_N)$ (*) en N puntos consecutivos $x_0, x_1, \dots, x_N$. La distancia entre dos puntos consecutivos debe tender a 0 si n o N tienden a infinito.

El operador $H(Y)$ es consistente con el $h(y)$ si

$$\sum_{j=0}^N |h(y)_{x=x_j} - h_j(Y)|,$$

donde

$$Y = (y_1, \dots, y_n) \text{ e } y_i = y(x_i),$$

tiende a 0 si N y n tienden a infinito para toda función admisible $y(x)$.

(*) $H(Y) = (h_0(y_0, y_1, \dots, y_n), h_1(y_0, y_1, \dots, y_n), \dots, h_N(y_0, y_1, \dots, y_n))$.

Por ejemplo, $h(y) = \int_a^b K(x, \xi) y(\xi) d\xi$ se reemplaza por:

$$h_j(Y) = \sum_{i=0}^n K(x_j, \xi_i) y_i D_i.$$

Con

$$\begin{aligned} \xi_i &= a + i l, & x_j &= a + j l', \\ l &= (b - a)/n, & l' &= (b - a)/N, \end{aligned}$$

y los D_i forman la matriz $(D_i) = h(1/2, 1, 1, \dots, 1, 1/2)$ en el método trapezoidal de integración $(D_i) = (h/3)(1, 4, 2, 4, 2, \dots, 2, 4, 1)$, en el de Simpson, mejor si y sólo si la función $y(x)$ es relativamente lisa.

Para resolver un problema planteado con $h(y)$ se puede resolver uno en $H(Y)$ análogo, siempre y cuando h y H sean consistentes y las soluciones de ambos problemas dependan continuamente de los datos. Se dice entonces que se emplea un método en diferencias finitas.

Así, las ecuaciones $h(y) = 0$ equivalen a minimizar una norma $|h(y)|$ y se resuelven aproximadamente minimizando una norma de $H(Y)$. Si esa norma es

$$\sum_{j=0}^N |h_j(Y)|^2,$$

tenemos que resolver numéricamente un problema en mínimos cuadrados, y si es

$$\sum_{j=0}^N |h_j(Y)|,$$

un problema en valores absolutos que es lineal si las h_j lo son en Y . Si $n = N$ es un método de colocación se resuelven n ecuaciones con n incógnitas $h_j(Y) = 0$, si $n > N$ la solución no es única, si $n < N$ es un problema de programación en valores absolutos. Puede aplicarse el método a ejemplos de la referencia 3.

4. METODOS DE RITZ-GALERKIN

La solución aproximada de un problema variacional $\min(F(y))$ o de una ecuación $h(y) = 0$ se busca a menudo (referencias 3 y 5) con ventaja buscando una solución de la forma:

$$y = \sum_{i=1}^n C_i \cdot w_i(x).$$

Si el problema es lineal, hay ventaja en restar a $y(x)$ previamente una $w_0(x)$ que cumpla las condiciones de borde y suponer que las $w_i(x)$ cumplen condiciones de borde homogéneas.

Reemplazando $h_j(Y)$ en el apartado 3 por:

$$h_j \left(\sum_{i=1}^n C_i \cdot w_i(x) \right)$$

función de $C_1, \dots, C_n$, encontramos que el problema se aproxima minimizando

$$\sum_{j=0}^N D_j \left| h_j \left(\sum_{i=1}^n C_i w_i(x) \right) \right|,$$

que es un problema en valores absolutos, lineal si lo son las h_j . Las condiciones de borde pueden darnos restricciones.

Así, tomamos de la referencia 5 el ejemplo:

$$h(x) = -y(x) + x + \int_0^1 K(x, \xi) y(\xi) d\xi = 0$$

con:

$$K(x, \xi) = x(1 - \xi) \text{ si } x \leq \xi, \\ \xi(1 - x) \text{ si } x \geq \xi.$$

Se demuestra que esta ecuación equivale al problema de contorno $y'' + y = 0$; $y(0) = 0$, $y(1) = 1$, de solución $\sin x / \sin 1$.

Sea $x_i = i/4$ (i de 0 a 5).

Aproximamos la solución por el polinomio $A_1 + A_2 x + A_3 x^2$. Obtenemos el sistema en valores absolutos.

Quedan las ecuaciones:

$$\sum_{j=0}^4 D_j \left| x_j - (A_1 + A_2 x_j + A_3 x_j^2) + \sum_{i=0}^4 K(x_j, \xi_i) (A_1 + A_2 \xi_i + A_3 \xi_i^2) D_i \right|,$$

o sea:

$$\begin{aligned} & \text{Mín } \frac{1}{8} (|A_1| + |A_1 + A_2 + A_3 - 1|) \\ & + 2 \left| -0,90625 A_1 - 0,2109375 A_2 - 0,0419921875 A_3 + \frac{1}{4} \right| \\ & + 2 \left| -0,8751 A_1 - 0,4375 A_2 - 0,2135416 A_3 + \frac{1}{2} \right| \\ & + 2 \left| 0,90625 A_1 - 0,6953125 A_2 - 0,5263671875 A_3 + \frac{3}{4} \right| \end{aligned}$$

El problema se transforma en uno de programación lineal por el método del apartado 1:

$$\text{Min } (B_1 + B_2 + B_3 + B_4 + B_5 - 3,6875 A_1 - 1,84375 A_2 - 1,2819010416 A_3 + 2)$$

$$B_1 \geq A_1,$$

$$B_5 \geq A_1 + A_2 + A_3 - 1,$$

$$B_2 \geq 1,8125 A_1 + 0,421875 A_2 + 0,083984375 A_3 - 0,5,$$

$$B_3 \geq 1,75 A_1 + 0,875 A_2 + 0,42708333 A_3 - 1,$$

$$B_4 \geq 1,8125 A_1 + 1,390625 A_2 + 1,052734375 A_3 - 1,5,$$

de resultados:

$$A_1 = 0,$$

$$A_2 = 1,27907,$$

$$A_3 = -0,27907.$$

Que nos da la aproximación:

$$1,27907 x - 0,27907 x^2.$$

Que comparamos con la solución exacta (para el mismo ejemplo con otros métodos, ver referencia en pág. 458), resultando en este caso una aproximación algo menor que por mínimos cuadrados (ver tabla al final).

5. CALCULO DE ERRORES

Si $y_1, \dots, y_n$ se conocen aproximadamente por verificar las relaciones $f_j(y_1, \dots, y_n) = 0$, aproximadamente, asignándose el peso D_j a la relación $f_j = 0$, es costumbre minimizar

$$\sum_{j=1}^N D_j (f_j)^2,$$

lo que conduce a las ecuaciones

$$\sum_{j=1}^N D_j f_j (\partial f_j / \partial y_i) = 0.$$

El problema está determinado, en general, si $N \geq n$, y las ecuaciones son lineales si las f_j lo son.

Podemos también minimizar $\sum_j D_j |f_j|$, teniendo un programa en valores absolutos, lineal si los f_j lo son.

Prescindimos aquí del estudio de la resolución de programas de programación no lineal.

6. PROBLEMAS EN MAS DIMENSIONES

Muchos de los problemas resueltos en la referencia 3 por los métodos de Ritz, Galerkin o Trefftz (ecuaciones integrales, diferenciales, cálculo de variaciones), minimizando una expresión $\sum |f_j|^2$ (con $\sum$ múltiples a menudo), pueden ser resueltos minimizando $\sum |f_j|$ como en los apartados 4 ó 5.

7. APROXIMACION DE CURVA

Para aproximar $\sum_{i=1}^n C_i \cdot W_i(x)$ a una función $f(x)$, se puede minimizar en $N + 1$ puntos ($N + 1 \geq n$) la suma ponderada por D_j (j de 0 a $N + 1$):

$$\sum_{j=0}^{N+1} D_j \left| \sum_{i=1}^n C_i \cdot w_i(x_j) - f(x_j) \right|$$

con las notaciones del apartado 4, quedando un programa lineal en valores absolutos que se resuelve según el apartado 1, pudiendo imponerse restricciones a la aproximación (paso de la curva por un punto; limitación derivadas, intervalo de paso).

Un método análogo ha sido utilizado por el Dr. Ing. de Caminos J. Puy, por indicación nuestra para aproximar un trazado optimizado de carreteras. El interés del método es el poder poner restricciones lineales, en este caso muy útiles.

8. PROGRAMACION LINEAL APLICADA AL CALCULO PLASTICO

Una aplicación de los métodos del apartado 1 al cálculo plástico elemental, fue publicada por el autor (referencia 1), sirviendo para calcular la seguridad de una estructura o para hacer un dimensionamiento óptimo (con ciertas hipótesis). Es de observar que algunas de esas ideas ya habían sido publicados.

9. CONCLUSION

En muchos métodos de cálculo aproximado, en los que tradicionalmente se puede calcular el mínimo de una suma de cuadrados, se puede calcular el mínimo de una suma de valores absolutos para aproximar la solución.

Se indica el método de obtener programas sin valores absolutos de aplicación por programación lineal en los casos lineales.

Las mayores aplicaciones son los métodos de Ritz-Galerkin, la aproximación de curvas y el cálculo plástico elemental.

La complicación de cálculo en ordenador es del mismo orden que con los mínimos cuadrados.

x	sen x/sen 1	Aproximada
0	0	0
0,1	0,1186	0,1251
0,2	0,2361	0,2446
0,3	0,3512	0,3586
0,4	0,4628	0,4670
0,5	0,5697	0,5698
0,6	0,6710	0,6670
0,7	0,7656	0,7586
0,8	0,8526	0,8446
0,9	0,9309	0,9251
1,0	1,0000	1,0000

BIBLIOGRAFIA

1. ANTON CORRALES, J. M.: "Programación lineal aplicada al cálculo plástico elemental de estructuras lineales". Informes de la Construcción, del Instituto Torroja, junio 1970. Madrid.
2. BERGE, C., y A. GHOUILA-HENRI: "Programmes, jeux et réseaux de transport". Dunod, 1962. París.
3. COLLATZ: "Nümerishe Behandlung von Differentialgleichungen". Springer Verlag, New York (también en inglés).
4. GASS: "Programación lineal". C.E.C.S.A., México. Original Mac Graw Hill, en inglés, 1958-1964.
5. HILDEBRAND, F. B.: "Methods of applied mathematics". Prentice Hall Inc., 1952 (reeditado con título y contenido parecido, después, en otra editorial).
6. HUÁRD, P.: "Tour d'horizon: programmation non linéaire". Revue française d'informatique et de recherche opérationnelle, 1971, R-4, pp. 8 a 48.