Metoda celor mai mici pătrate pentru problema de programare pătratică

Istrati Daniela, Moraru Vasile, Zaporojan Sergiu
Universitatea Tehnică a Moldovei
Chişinău, Republica Moldova
daniela.istrati@ia.utm.md. vasile.moraru@ia.utm.md, zaporojan_s@yahoo.com

***Rezumat*—În această lucrare problema de programare pătratică cu restricții inegalități liniare în urma unor transformări convenabile a condițiilor de optimalitate Karush-Kuhn-Tucker se reduce la rezolvarea unui sistem de ecuații, care la rândul său se poate rezolva cu ajutorul metodei celor mai mici pătrate.**

***Termeni cheie*—optimizare pătratică, condițiile Karush-Kuhn-Tucker, metoda celor mai mici pătrate.**

I. Introducere

Considerăm problema de programare pătratică

$$\left.\begin{array}{c} f(x) = \dfrac{1}{2} x^T Q x + q^T x \to \min \\ \textit{referitor la}: \\ Ax \ge b \end{array}\right\} \qquad (1)$$

unde $x = (x_1, x_2, \ldots, x_n)^T \in \Re^n$, $q = (q_1, q_2, \ldots, q_n)^T \in \Re^n$, $b = (b_1, b_2, \ldots, b_m)^T \in \Re^m$, $Q = \left(q_{ij}\right)_{i,j=1}^{n} \in \Re^{n\times n}$, $Q^T = Q$, $A = \left(a_{ij}\right)_{i=1,m}^{j=1,n} \in \Re^{m\times n}$. Simbolul "$T$" arată operaţia de transpunere

Problemele de optimizare pătratică sunt des întâlnite în diferite aplicaţii reale: suportul maşinilor vectoriale, teoria grafurilor, analiza structurală, VLSI design ş.a. O bibliografie completă referitoare la problemele de tipul (1) poate fi găsită în [1], lucrare care conţine peste 1000 (o mie) de referinţe!

Asociem problemei (1) funcția Lagrange:

$$L(x,\lambda) = f(x) + \lambda^T (Ax - b)$$

Fie $x^* \in \Re^n$ un punct de minim (local). Atunci, cum este bine cunoscut (vezi de exemplu [2]), există multiplicatorii Lagrange

$$\lambda^* = \left(\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*\right)^T$$

astfel încât

$$\nabla_x L\left(x^*, \lambda^*\right) = \nabla f\left(x^*\right) - A^T \lambda^* = 0\,, \qquad (2)$$

$$D(x^*)\lambda^* = 0, \qquad (3)$$

$$\lambda^* \ge 0, Ax^* - b \ge 0. \qquad (4)$$

unde $D(x^*)$ este matricea diagonală cu elementele

$$d_{ii} = \sum_{j=1}^{n} a_{ij} x_j^* - b_i\,,$$

Notăm

$$I(x^*) = \left\{ i : \sum_{j=1}^{n} (a_{ij} x_j^* - b_i) = 0 \right\}.$$

Se presupune că în soluția optimă $(x^*, \lambda^*)^T$ sunt satisfăcute relațiile:

1. $\lambda_i^* > 0$ pentru $\forall i \in I(x^*)$,
2. $z^T \nabla_x L\left(x^*, \lambda^*\right) z > 0$ pentru orice

$$z \in T(x^*) = \left\{ z \in \Re^n : z^T \left(\sum_{j=1}^{n} (a_{ij} x_j^* - b_i) \right) = 0, i \in I(x^*) \right\},$$

Condițiile (3) și (4), numite condiții de complementaritate, ridică mari probleme în rezolvarea directă a sistemului de ecuații și inecuații (2)-(4).

Cu ajutorul funcțiilor de complementaritate [3], [4] relațiile (3) și (4) pot fi reduse la un sistem alcătuit doar din ecuații, astfel încât să putem aplica metodele clasice de rezolvare. Cu părere de rău majoritatea funcțiilor de complementaritate cunoscute ne conduc la sisteme cu proprietatea vădită de singularitate.

În lucrarea de față vom aplica procedeul de transformare a relațiilor (3) și (4), utilizând funcțiile propuse în [5],[6].

II. Reformularea condițiilor Karush-Kuhn-Tucker

Definim funcțiile

$$u(y) = \frac{1}{2}\left(y^3 + y^2|y|\right),$$

$$v(y) = y^2|y| - u(y).$$

Se constată cu ușurință că

1. $u(y) \ge 0, v(y) \ge 0, \forall y \in \Re,$
2. $u(y) = 0, \forall y \le 0,$
3. $v(y) = 0, \forall y \ge 0,$
4. $u(y) \times v(y) = 0, \forall y \in \Re.$

Relațiile (3) și (4) pot fi transformate în ecuații prin introducerea unor variabile auxiliare $y_1, y_2, \dots, y_m$ astfel încât

$$u(y_i) - \sum_{j=1}^{n} a_{ij} x_j + b_i = 0,$$
$$v(y_i) - \lambda_i = 0,$$
$$i = 1,2,\dots,m.$$

Obținem că condițiile Karush-Kuhn-Tucker (2)-(4) pot fi rescrise în forma echivalentă

$$F(x,\lambda,y) = 0, \tag{5}$$

unde

$$F(x,\lambda,y) = \begin{pmatrix} \nabla_x L(x,\lambda) \\ U(y) - Ax + b \\ V(y) - \lambda \end{pmatrix},$$

iar

$$U(y) = \begin{pmatrix} u(y_1) \\ u(y_2) \\ \vdots \\ u(y_m) \end{pmatrix}, V(y) = \begin{pmatrix} v(y_1) \\ v(y_2) \\ \vdots \\ v(y_m) \end{pmatrix}.$$

Așa dar, rezolvarea problemei de programare pătratică (1) este redusă la rezolvarea sistemului de ecuații (5).

III. Metoda celor mai mici pătrate

Rezolvarea sistemului de ecuații (5) poate fi efectuată, minimizând funcția

$$\varphi(x,\lambda,y) = \frac{1}{2}\|F(x,\lambda,y)\|^2. \tag{6}$$

Matricea Jacobiană pentru funcția-vector $F(x,\lambda,y)$ este

$$F'(x,\lambda,y) = \begin{pmatrix} Q & -A^T & O_{n\times m} \\ -A^T & O_{m\times m} & U'(y) \\ O_{m\times n} & -I_m & V'(y) \end{pmatrix}$$

Aici s-a notat $O_{n\times m}$, $O_{m\times m}$ și $O_{m\times n}$ matricele nule de dimensiuni $n\times m$, $m\times m$ și respectiv $m\times n$, I_m - matricea unitate de dimensiune $m\times m$, iar prin $U'(y)$ și $V'(y)$ matricele diagonale de dimensiune $m\times m$:

$$U'(y) = \begin{pmatrix} u'(y_1) & 0 & \cdots & 0 \\ 0 & u'(y_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & u'(y_m) \end{pmatrix},$$

$$V'(y) = \begin{pmatrix} v'(y_1) & 0 & \cdots & 0 \\ 0 & v'(y_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & v'(y_m) \end{pmatrix}$$

cu elementele

$$u'(y_i) = \frac{1}{2} y_i^2 sign(y_i) + y_i|y_i| + \frac{3}{2} y_i^2,$$
$$v'(y) = \frac{1}{2} y_i^2 sign(y_i) + y_i|y_i| - \frac{3}{2} y_i^2.$$

Menționăm ca avem garantată relația $\lambda_i^* > 0$ pentru $\forall i \in I(x^*)$. Într-adevăr, dacă $\sum_{j=1}^{n}(a_{ij}x_j^* - b_i) = 0$, atunci $u(y_i^*) = 0$, de unde $\lambda_i^* = v(y_i^*) > 0$.

Luând în considerație acest lucru și presupunerile de mai sus se poate demonstra că matricea Jacobiană $F'(x,\lambda,y)$ este nedegenerată în vecinătatea soluției optime

Gradientul funcției (6) este

$$\nabla\varphi(x,\lambda,y) = F'(x,\lambda,y)F((x,\lambda,y)),$$

Așa cum $\det(F'(x,\lambda,y)) \neq 0$ avem că în vecinătatea soluției optime $\nabla\varphi(x,\lambda,y) = 0$ dacă și numai dacă $F((x,\lambda,y)) = 0$

Prin urmare rezolvarea problemei (1) poate fi redusă la minimizarea funcției (6).

IV. Concluzii

Pentru a depăși lipsa derivatelor funcțiilor de complementaritate [3],[4] s-au propus utilizarea funcțiilor *u* si *v*, funcții continui împreuna cu derivatele sale. Aceasta permite minimizarea funcției (6) cu ajutorul metodelor de gradient [2],[7].

Bibliografie

[1] N.I.M. Gould, Ph.L Toint. A Quadratic Programming Bibliography. Numerical Analysis Group Internal Report 200-1 Rutherford Appleton Laboratory, Chilton, England, 2001. http://www.optimization-online.org/DB_4TML/2001/02/285.html.

[2] M.S. Bazaraa, H. D. Sheraly, C.M. Shetty Nonlinear programming. Theory and algorithms. 3 rd. edition, Wiley Interscience. A John Wiley and Sons, Inc. Publication, 2006.

[3] M.C.Ferris, Ch. Kanzow, Complementary and related problems: A survey. 1998.

[4] C. Kanzow, Some equation – based methods for nonlinear complementarity problem. Optimization Method and Software, 3 (1), (1994), pp. 327-340.

[5] O. O. Stein, Lifting mathematical programs with complementary constraints. Mathematical Programming, 29, (2010), DOI: 10.1007/s10107-010-0345-y.

[6] V. Moraru, A Smooth Newton Method for Nonlinear Programming Problems with Inequality Constraints. Computer Science Journal of Moldova, vol. 20, no. 3(57), 2011, p.333-355. http://www.math.md/files/csjm/v19-n3/v19-n3-(pp333-355).pdf

[7] A.F. Izmailov, M.V. Solodov, Numerical Methods of Optimization. Fizmatlit/Nauka, Moscow, Russia, Second Edition – 2008. (In Russian).