БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ   ACADÉMIE BULGARE DES SCIENCES
ИЗВЕСТИЯ НА МАТЕМАТИЧЕСКИЯ ИНСТИТУТ
BULLETIN DE L'INSTITUT DE MATHÉMATIQUES
Том (Tome) VIII

ЕДНА МОДИФИКАЦИЯ НА ГРАДИЕНТНИЯ ПРОЕКТИВЕН МЕТОД В НЕЛИНЕЙНОТО ПРОГРАМИРАНЕ

Веселин Спиридонов

I. Постановка на задачата. Търси се абсолютният максимум на функцията $F(x_1, \dots, x_m)$, дефинирана в изпъкналата, ограничена и затворена област R, определена от системата линейни неравенства

$$(1) \qquad \sum_{j=1}^{m} n_{ij} x_j - b_j \geqq 0, \qquad i = 1, \dots, k;\ k > m.$$

Предполага се, че $F(x_1, \dots, x_m)$ притежава непрекъснати и ограничени втори частни производни по x_i и е вдлъбната функция в областта R.

За решаването на тази задача от нелинейното програмиране J. B. Rosen [1] е предложил така наречения градиентен проективен метод. При неговото използване обаче съществува следното затруднение: налага се на всяка стъпка от итерационния процес да се изчислява спектралната норма на матрица, което е свързано с намирането на собствени значения [2] и [3]. Това усложнение на метода го прави практически неудобен.

Целта на настоящата работа е да се изложи алгоритъм, основан на такава модификация на градиентния проективен метод, при която се избягва намирането на собствени значения на матрица. При използуване на електронна сметачна машина по този начин значително се спестяват машинно време и памет.

II. Означения. Ще спазваме терминологията и означенията, възприети в [1].

Всяка наредена съвкупност от m числа $\{x_1, \dots, x_m\}$ определя точка в m-мерното евклидово пространство E_m, на която съответствува m-мерният вектор-стълб $\mathbf{x} = \{x_1, \dots, x_m\}$. Всяко неравенство от системата (1) определя полупространство в E_m.

Без ограничение на общността може да се счита, че числата n_{ij} са нормирани, така че

$$\sum_{j=1}^{m} (n_{ij})^2 = 1 \qquad (i = 1, 2, \dots, k).$$

Определяме k единични вектори $\mathbf{n}_i\ (i = 1, \dots, k)$ така: $\mathbf{n}_i = \{n_{i1}, \dots, n_{im}\}$; матрица $N_k = [\mathbf{n}_1, \dots, \mathbf{n}_k]$ и k-мерен вектор $\mathbf{b}_k = \{b_1, \dots, b_k\}$.

Тогава системата неравенства (1) може да се запише във вида

$$(2) \qquad \lambda_i(\mathbf{x}) = \mathbf{x}^T \mathbf{n}_i - b_i \geqq 0 \quad (i = 1, 2, \dots, k)$$

или във вида

$$N_k^T\mathbf{x}-\mathbf{b}_k\geqq\mathbf{0}. \tag{3}$$

Всяко уравнение $\lambda_i(\mathbf{x})=0$ $(i=1,\ldots,k)$ определя хиперравнина H_i в E_m. Нека Q е пресечницата (сечението) на q хиперравнини $H_i(i=1,\ldots,q)$. Ще предполагаме, че векторите $\mathbf{n}_i(i=1,\ldots,q)$, а следователно и хиперравнините са линейно независими. От линейната независимост следва, че Q има измерение $m-q$. Да определим матриците $N_q=[\mathbf{n}_1,\ldots,\mathbf{n}_q]$ и

$$P_q=J-N_q(N_q^TN_q)^{-1}N_q^T,$$

където J е единична матрица от ред m. Известно е [1], че P_q е проекционна матрица, която пренася всеки вектор от E_m във вектор от пресечницата Q.

III. Алгоритъм. Нека сме получили точка $\mathbf{x}_\nu$, която се намира на пресечницата Q на хиперравнините H_i $(i=1,\ldots,q)$, $0\leqq q\leqq m$. Следващата точка от итерационния процес $\mathbf{x}_{\nu+1}$ и съответната пресечница, на която тя се намира, се получават посредством следния алгоритъм:

1. Образувай матриците $N_q, (N_q^TN_q)^{-1}$ и P_q. Намери в точката $\mathbf{x}_\nu$ градиента $\mathbf{g}(\mathbf{x}_\nu)$ и проекцията му $P_q\mathbf{g}(\mathbf{x}_\nu)$. Премини към 2.

2. Провери дали е изпълнено условието $P_q\mathbf{g}(\mathbf{x}_\nu)=\mathbf{0}$. Ако ДА — премина към 7. Ако НЕ — продължи с 3.

3. Изчисли единичния вектор

$$\mathbf{z}=\frac{P_q\mathbf{g}(\mathbf{x}_\nu)}{|P_q\mathbf{g}(\mathbf{x}_\nu)|}.$$

Премини към 4.

4. Намери

$$\tau=\min\left\{\frac{\lambda_i(\mathbf{x}_\nu)}{\mathbf{z}^T\mathbf{n}_i}>0\right\},\quad i=q+1,\ldots,k.$$

Отбележи тази хиперравнина H_i, за която се достига минимумът τ Образувай

$$\mathbf{x}'_{\nu+1}=\mathbf{x}_\nu+\tau\mathbf{z}$$

и получи в точката $\mathbf{x}'_{\nu+1}$ градиента $\mathbf{g}(\mathbf{x}'_{\nu+1})$. Премини към 5.

5. Провери дали е изпълнено условието $\mathbf{z}^T\mathbf{g}(\mathbf{x}'_{\nu+1})<0$. Ако ДА — премини към 10. Ако НЕ — продължи с 6.

6. Означи $\mathbf{x}_{\nu+1}=\mathbf{x}'_{\nu+1}$ и считай тази точка за следваща точка от итерационния процес. Замести Q с нова пресечница, получена от пресичането на Q с отбелязаната в 4. хиперравнина. Изтрий признака β, ако има образуван такъв признак. Замести $\mathbf{x}_\nu$ с $\mathbf{x}_{\nu+1}$. Премини към 1.

7. Намери вектора

$$\mathbf{r}=(N_q^TN_q)^{-1}N_q^T\mathbf{g}(\mathbf{x}_\nu).$$

Премини към 8.

8. Провери дали е изпълнено условието $\mathbf{r}\leqq\mathbf{0}$. Ако ДА — премини към 16. Ако НЕ — продължи с 9.

9. Намери най-голямата компонента на вектора $\mathbf{r}$. Нека тя е r_l. Отстрани вектора $\mathbf{n}_l$ от матрицата N_q. В резултат се получава матрица N_{q-1} с размери $m, (q-1)$. Вместо Q постави нова пресечница, получена от пресичането на $q-1$ хиперравнини $H_1, H_2, \ldots, H_{l-1}, H_{l+1}, \ldots, H_q$. Образувай матриците $(N^T_{q-1} N_{q-1})^{-1}$ и P_{q-1}. Намери $P_{q-1}\,\mathbf{g}(\mathbf{x}_\nu)$. Изчисли единичния вектор

$$\mathbf{z} = \frac{P_{q-1}\,\mathbf{g}(\mathbf{x}_\nu)}{|P_{q-1}\,\mathbf{g}(\mathbf{x}_\nu)|}.$$

Премини към 4.

10. Изчисли

$$\varrho = \frac{\mathbf{z}^T\mathbf{g}(\mathbf{x}_\nu)}{\mathbf{z}^T\mathbf{g}(\mathbf{x}_\nu) - \mathbf{z}^T\mathbf{g}(\mathbf{x}'_{\nu+1})}$$

и

$$\mathbf{x}''_{\nu+1} = \varrho\mathbf{x}'_{\nu+1} + (1-\varrho)\mathbf{x}_\nu.$$

Намери $g(\mathbf{x}''_{\nu+1})$ и $\mathbf{z}^T\mathbf{g}(\mathbf{x}''_{\nu+1})$. Избери си подходящо $\varepsilon_1 > 0$. Премини към 11.

11. Провери дали е изпълнено условието $\mathbf{z}^T\mathbf{g}(\mathbf{x}''_{\nu+1}) > 0$. Ако ДА — премини към 17. Ако НЕ — продължи с 12.

12. Изчисли $\delta = |\mathbf{x}_\nu - \mathbf{x}''_{\nu+1}|$. Образувай признак α. Премини към 13.

13. Провери дали е изпълнено условието $\delta < \varepsilon_1$. Ако ДА — премини към 18. Ако НЕ — продължи с 14.

14. Провери дали е изпълнен признакът α. Ако ДА — премини към 22. Ако НЕ — продължи с 15.

15. Замести $\mathbf{x}''_{\nu+1}$ с $\mathbf{x}'''_{\nu+1}$ и $\mathbf{x}_\nu$ с $\mathbf{x}''_{\nu+1}$. Премини към 10.

16. Пресметни $F(\mathbf{x}_\nu)$ и запиши $\mathbf{x}_\nu$ и $F(\mathbf{x}_\nu)$. Спри!

17. Изчисли $\delta = |\mathbf{x}'_{\nu+1} - \mathbf{x}''_{\nu+1}|$. Премини към 13.

18. Означи $\mathbf{x}_{\nu+1} = \mathbf{x}''_{\nu+1}$ и считай тази точка за следващата точка от итерационния процес. Досегашната пресечница остави непроменена. Премини към 19.

19. Провери дали е изпълнен признакът β. Ако ДА — премини към 23. Ако НЕ — продължи с 20.

20. Образувай признак β. Премини към 21.

21. Замести $\mathbf{x}$ с $\mathbf{x}_{\nu+1}$. Премини към 1.

22. Замести $\mathbf{x}''_{\nu+1}$ с $\mathbf{x}'''_{\nu+1}$ и $\mathbf{x}'_{\nu+1}$ с $\mathbf{x}''_{\nu+1}$. Премини към 10.

23. Образувай $|\mathbf{x}_{\nu+1} - \mathbf{x}_\nu|$. Избери си подходящо $\varepsilon_2 > 0$. Премини към 24.

24. Провери дали е изпълнено условието $|\mathbf{x}_{\nu+1} - \mathbf{x}_\nu| \geqq \varepsilon_2$. Ако ДА — премини към 21. Ако НЕ — продължи с 25.

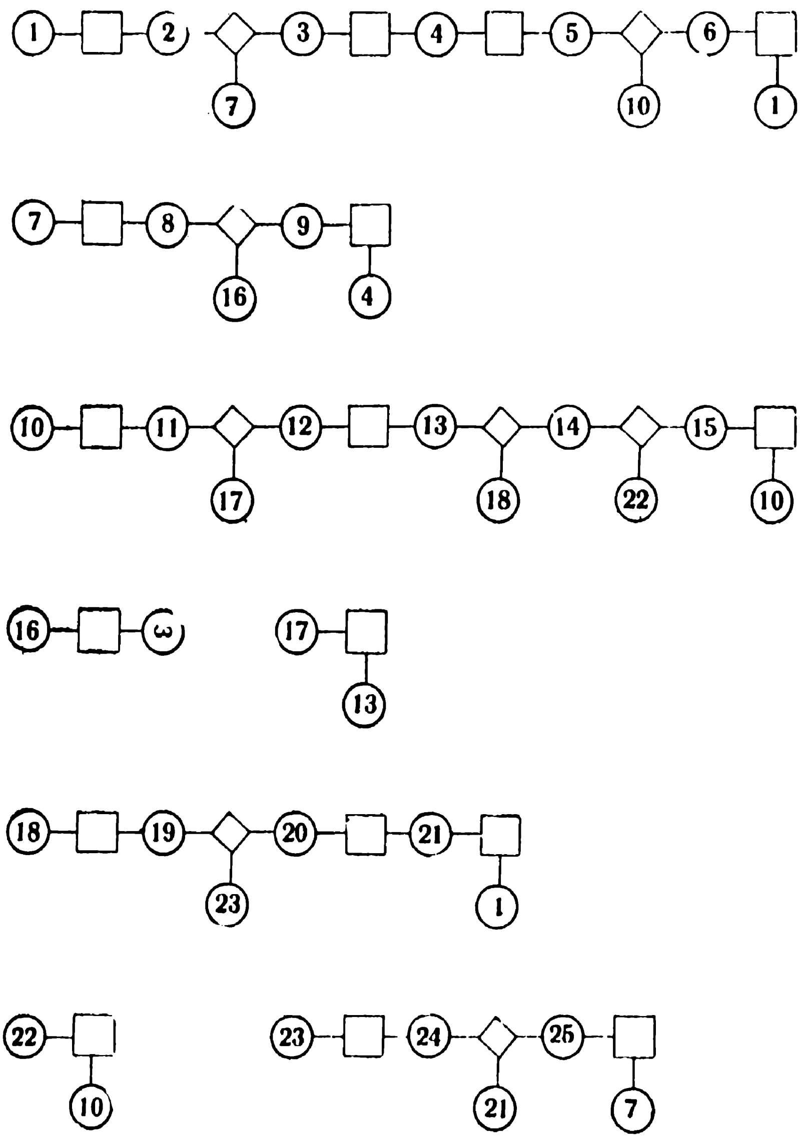
25. Замести $\mathbf{x}_\nu$ с $\mathbf{x}_{\nu+1}$. Премини към 7.

IV. Забележки. а. Алгоритмът е даден подробно и може непосредствено да се използува за програмиране на електронна сметачна машина. Видът, в който е написан, съответствува на начина за записване на блок-схемите, възприет в [4]. Самата блок-схема е дадена на фиг. 1.

б. Отстраняването на изродени случаи се извършва по същия начин, както в [1].

в. В блок 9 $P_{q-1}\mathbf{g}(x_\nu)|\neq 0$ се гарантира от $r_l > 0$.

г. Важен въпрос е намирането на допустима точка, с която да започне алгоритъмът. Ако това не може да се извърши въз основа на кон-



Фиг. 1. Блок-схема на алгоритъма

кретни съображения, трябва предварително да се реши следната задача: да се намери точка $\mathbf{x}_0$, която удовлетворява системата от неравенства (2). За тази цел J. B. Rosen [1] предлага свой метод, който има това преимущество, че използува матриците $N_q, (N_q^T N_q)^{-1}$ и P_q, които се срещат и в алгоритъма. При него обикновено за $\mathbf{x}_0$ се получава контурна точка. За решаването на същата задача може да се приложи и релаксационният метод [5]. Тогава за $\mathbf{x}_0$ обикновено се намира вътрешна точка. По този начин, започвайки с $\mathbf{x}_0$, с помощта на система рекурентни формули от [1] последователните приближения може да се получат, без да се обръща матрицата $(N_q^T N_q)$.

V. Обосноваване на алгоритъма. Ще използуваме следната теорема, доказана в [1]:

Теорема 1. *Ако* $\mathbf{x}_\nu$ *е контурна точка на* R *и лежи точно на* $q (1 \leq q \quad m)$ *хиперравнини, необходимо и достатъчно условие, за да има* $F(\mathbf{x})$ *абсолютен максимум в* $\mathbf{x}_\nu$, е

$$P_q \mathbf{g}(\mathbf{x}_\nu) = 0 \quad \text{и } \mathbf{r} \leq 0.$$

Ако $\mathbf{x}_\nu$ е вътрешна точка в R, необходимото и достатъчно условие, за да има функцията $F(\mathbf{x})$ абсолютен максимум в $\mathbf{x}_\nu$, е $\mathbf{g}(\mathbf{x}_\nu) = \mathbf{0}$.

Трябва да се отбележи, че ако се възприеме за $q = 0, P_0 = J$ и $\mathbf{r} = \mathbf{0}$, втората част от теоремата се явява следствие от първата.

Не представлява затруднение доказателството на следните помощни теореми:

Теорема 2. *Ако* $\mathbf{x}_\nu \in R$, *то и следващата точка, получена чрез алгоритъма,* $\mathbf{x}_{\nu+1} \in R$.

Теорема 3. *Да разгледаме редицата от точки*

$$\mathbf{x}_\nu, \ \mathbf{x}'_{\nu+1}, \ \mathbf{x}''_{\nu+1}, \ldots,$$

която се получава от алгоритъма чрез клона 10 *на конфигурацията* (фиг. 1), *като изключим условното предаване на* 18. *Тази редица е сходяща и за нейната граница* $\mathbf{x}'$ е изпълнено $\mathbf{z}^T \mathbf{g}(\mathbf{x}') = 0$.

Теорема 4. *Ако чрез алгоритъма, от който сме изключили условното предаване на* 25, *получим безкрайна редица от точки*

$$\mathbf{x}_\nu, \ \mathbf{x}_{\nu+1}, \ \mathbf{x}_{\nu+2}, \ldots,$$

които се намират на една и съща пресечница, то редицата е сходяща и за нейната граница $\mathbf{x}''$ *е изпълнено* $P_q \mathbf{g}(\mathbf{x}'') = \mathbf{0}$.

Алгоритъмът се основава на следната

Теорема 5. *Нека* $\mathbf{x}_0 \in R$. *Посредством дадения алгоритъм построяваме редицата от точки*

$$\mathbf{x}_0, \ \mathbf{x}_1, \ldots, \ \mathbf{x}_\nu, \ldots$$

Тогава съществува такъв индекс n, *че в някоя точка от* ε*-околност на* $\mathbf{x}_n$ *(където* ε *евентуално зависи от* ε_1 и ε_2*) функцията* $F(\mathbf{x})$ *има абсолютен максимум. Итерационният процес завършва с изчисляването на* $\mathbf{x}_n$ и $F(\mathbf{x}_n)$.

Доказателство. Нека сме получили точка $\mathbf{x}_\nu$. Съгласно с теорема 2 $\mathbf{x}_\nu \in R$. По-нататък алгоритъмът дава три възможности:

а. От $\mathbf{x}_\nu$ да се премине към $\mathbf{x}_{\nu+1}$ посредством блок 18. Нека все по този начин се получават последователните приближения

$$\mathbf{x}_\nu, \ \mathbf{x}_{\nu+1} \ \mathbf{x}_{\nu+2}, \ldots$$

Според теорема 4 тази редица би била сходяща. Но логическият блок 24 я прекъсва и чрез блок 25 се стига до блок 7. Следователно чрез блок 18 могат да се получат подред само краен брой последователни приближения, след което задължително се появява някой от случаите б. или в.

б. От $\mathbf{x}_\nu$ да се премине към $\mathbf{x}_{\nu+1}$ посредством блок 6. Този случай е разгледан в [1]. След краен брой итерации се появява случай в.

в. Итерационният процес спира в блок 16. В блок 16 се идва само от блок 8 (фиг. 1). В 8 се преминава от 7, а в 7 — от 2 или от 25. Да разгледаме тези два случая.

В първия случай изпълнението на логическите условия в 2 и 8 въз основа на теорема 1 показва, че $F(\mathbf{x})$ има абсолютен максимум в $\mathbf{x}_\nu$. Във втория случай имаме редица от точки

$$\mathbf{x}_k, \mathbf{x}_{k+1}, \ldots, \mathbf{x}_\nu.$$

Съгласно с теорема 4 в някаква ε-околност на $\mathbf{x}_\nu$ ще има точка $\mathbf{x}''$, за която е изпълнено $P_q\mathbf{g}(\mathbf{x}'')=0$. Тъй като $\mathbf{g}(\mathbf{x})$, а следователно и $\mathbf{r}(\mathbf{x})$ са непрекъснати функции на $\mathbf{x}$ (поне когато $\mathbf{x}$ не напуска пресечницата Q), то от $\mathbf{r}(\mathbf{x}_\nu)\leq 0$ следва, че $\mathbf{r}(\mathbf{x}'')\leq 0$. Съгласно с теорема 1 функцията $F(\mathbf{x})$ има абсолютен максимум в $\mathbf{x}''$. Тук влиза и случаят, когато максимумът се достига във вътрешна точка. С това теоремата е доказана.

Постъпила на 5. XI. 1963 г.

ЛИТЕРАТУРА

1. Rosen J. B., The Gradient Projection Method for Nonlinear Programming, Part 1, Linear Constraints, Journ. Soc. Ind. Appl. Math., **8**, 1960, 181—217.
2. Bodewig E., Matrix calculus, Amsterdam, 1959.
3. Фаддеев Д. К. и В. Н. Фаддеева, Вычислительные методы линейной алгебры, Москва, 1960.
4. Димитров Е., Б. Сендов и В. Спиридонов, Аксиоматика и автоматична проверка на блок-схемите, Год. на Соф. унив., Физ.-мат. фак., **56**, кн. 1, 1961—1962, 185—190.
5. Motzkin S. and I. J. Schoenberg, The Relaxation Method for Linear Inequalities. Canadian Journ. Math., **6**, 1954, 393—404.

МОДИФИКАЦИЯ ГРАДИЕНТНОГО ПРОЕКТИВНОГО МЕТОДА В НЕЛИНЕЙНОМ ПРОГРАММИРОВАНИИ

В. Спиридонов

(*Резюме*)

Пусть область R определена системой неравенств (1), а функция $F(x_1, \ldots, x_m)$, определенная в R, имеет непрерывные и ограниченные вторые частные производные по x_i и является вогнутой в R. Требуется найти абсолютный максимум функции F в R. Для этого предлагается алгоритм, основанный на градиентном проективном методе. В сравнении с методом Розена [1], он имеет то преимущество, что собственные значения матриц не вычисляются. Сходимость этого алгоритма доказывается теоремами 1—5.

A MODIFICATION OF THE GRADIENT PROJECTION METHOD IN NONLINEAR PROGRAMMING

V. Spiridonov

(*Summary*)

Let the region R be defined by the system of inequalities (1) and the function $F(x_1, \ldots, x_m)$, defined on R, has continuous and limited partial derivatives of the second order with respect to x_i and be concave on R. The absolute maximum of the function F on R is to be found. With that end in view an algorithm is proposed based on the gradient projection method. Compared with Rosen's method [1] it has the advantage that eigenvalues of matrices are not calculated. The convergence of the algorithm is established by theorems 1—5.