Math-Net.Ru
Общероссийский математический портал

Ю. Г. Евтушенко, Два численных метода решения задач нелинейного программирования, *Докл. АН СССР*, 1974, том 215, номер 1, 38–40

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
http://www.mathnet.ru/rus/agreement

Параметры загрузки:
IP: 18.217.12.218
5 ноября 2024 г., 00:19:40

Доклады Академии наук СССР
1974. Том 215, № 1

УДК 518:519.3:62-50 · **МАТЕМАТИКА**

Ю. Г. ЕВТУШЕНКО

ДВА ЧИСЛЕННЫХ МЕТОДА РЕШЕНИЯ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

(Представлено академиком А. А. Дородницыным 28 VI 1973)

1. Рассмотрим задачи нахождения

$$\min_{x\in X_1} F(x), \quad X_1=\{x\in E_n\colon G(x)\leqslant 0,\ 0\leqslant x\}, \tag{1}$$

$$\min_{x\in X_2} F(x), \quad X_2=\{x\in E_n\colon g(x)=0,\ h(x)\leqslant 0,\ 0\leqslant x\}; \tag{2}$$

здесь F, G, g, h — непрерывно дифференцируемые функции n-мерного вектора x, определенные на открытых покрытиях множеств X_1, X_2; $G=\{G^1, G^2,\ldots, G^m\}\in E_m$, $g(x)\in E_s$, $h(x)\in E_k$, E_i есть i-мерное эвклидово пространство.

Через X_1^0, X_2^0 обозначим множества строго внутренних точек множеств X_1, X_2 соответственно:

$$X_1^0=\{x\colon G(x)<0,\ 0<x\},\quad X_2^0=\{x\colon g(x)=0,\ h(x)<0,\ 0<x\}.$$

Пусть G_x — матрица $n\times m$, (i, j)-элемент которой равен $\partial G^j/\partial x^i$, F_x — матрица-столбец $n\times 1$, i-й элемент которой есть $\partial F/\partial x^i$. Символ $D(z)$ обозначает диагональную матрицу, у которой i-й диагональный элемент есть z^i, размерность D совпадает с размерностью вектора z. Штрих обозначает транспонирование матрицы.

2. Для численного решения задачи (1) предлагается искать предельные при $t\to\infty$ точки решения задачи Коши для системы

$$dx/dt=-D(x)[F_x+G_xw],\quad x(0)=x_0\in X_1^0; \tag{3}$$

здесь вектор $w\in E_m$ определяется из решения линейной системы

$$Aw+G_x'D(x)F_x=0,\quad A=G_x'(D)(x)G_x-D(G). \tag{4}$$

Считаем, что множества X_1^0, X_2^0 не пустые и либо множества X_1, X_2 компактные, либо множества $Z_i=\{x\colon F(x)\leqslant F(x_0),\ x\in X_i\}$, $i=1, 2$, ограниченные.

Обозначим через $e_i\in E_n$ единичный вектор-столбец, у которого равна единице i-я координата. Предполагаем, что функции, определяющие X_1, удовлетворяют усиленным условиям регулярности, т. е. для всех $x\in X_1$ векторы G_x^i такие, что $G^i(x)=0$ и e_i такие, что $x^i=0$ линейно независимые.

Лемма 1. *Для всех $x\in X_1$ определитель системы* (4) *отличен от нуля.*

Предполагаем, что правые части системы (3) удовлетворяют условиям существования и единственности решений для всех $x\in X_1$.

Вычислим первые производные функции $G(x)$ и $F(x)$ в силу системы (3), получим

$$\dot G=-D(G)w,\quad \dot F=-\|D(x^{1/2})(G_xw+F_x)\|^2-\|D((-G)^{-1/2})w\|^2;$$

здесь $\|z\|^2=z'z$, $x^{1/2}=(\sqrt{x^1}, \sqrt{x^2},\ldots, \sqrt{x^n})\in E_n$, $(-G)^{1/2}=(\sqrt{-G^1},\ldots,\sqrt{-G^m})$.

Из этих формул и вида системы (3) следует, что множества X_1 и Z_1 являются положительно инвариантными по отношению к системе (3), т. е. если $x_0 \in X_1$, то $x(x_0, t) \in X_1$, $x(x_0, t) \in Z_1$ для всех $0 \leqslant t < \infty$.

Теорема 1. *Пусть $F(x)$, $G(x)$ — строго выпуклые, непрерывно дифференцируемые на X_1 функции, выполнены усиленные условия регулярности; тогда при $t \to \infty$ решения* (3) *сходятся к единственному решению задачи* (1) *при любых $x_0 \in X_1^0$, функция $F(x(t))$ монотонно убывает вдоль траекторий системы* (3).

Множество ω-предельных точек $\omega(x_0)$ не пусто для всех $x_0 \in X_1^0$ в силу ограниченности либо множества X_1, либо Z_1. Для каждого $x_0 \in X_1^0$ существует единственная точка $x_* \in \omega(x_0)$, $F_x(x_*) = 0$. Действительно, в противном случае на односвязном множестве $\omega(x_0)$ функция $F(x)$ принимала бы одно и то же значение в нескольких точках, что невозможно в силу строгой выпуклости $F(x)$. Если $x_* \in \omega(x_0)$, $x_* \in X_1^0$, то $\hat{F}_x(x_*) = 0$, правые части системы (3) обращаются в нуль; в этом случае в точке x_* выполнены достаточные условия Куна — Таккера минимума $F(x)$ в задаче (1). Из анализа системы (3) следует, что аналогичная ситуация имеет место, когда $x_* \in X_1 \setminus X_1^0$.

3. Метод (3) можно использовать для решения задачи (2). Вид системы (3) для нее останется прежним. Вектор-функцию следует теперь считать состоящей из двух векторов $G = \{g, h\}$, начальная точка x_0 должна принадлежать X_2^0.

В некоторых частных случаях (3) переходит в методы, предложенные в ([1], [2]). В частности (см. ([1])), когда $X = \{x: \sum_{i=1}^{n} x^i = 1,\ x \geqslant 0\}$, система (3) имеет вид

$$\dot{x}^i = x^i \left[\sum_{j=1}^{n} x^j F_{x^j} - F_{x^i}\right], \quad i = 1, 2, \ldots, n;$$

если $X = \{x: a \leqslant x \leqslant b\}$, то

$$\dot{x}^i = (x^i - a^i)(b^i - x^i) F_{x^i}, \quad i = 1, 2, \ldots, n.$$

4. Приведем дискретный вариант метода:

$$x_{s+1} = x_s - \alpha B(x_s) F_x(x_s), \quad B(x) = D(x) - D(x) G_x A^{-1}(x) G_x' D(x), \tag{5}$$

где $0 < \alpha$. Симметрическая матрица $B(x)$ неотрицательно определенная для всех $x \in X$, поэтому $\max\limits_{x \in X_1,\ z \in E_n} z'B(x)z/z'z = \lambda > 0$. Считаем, что градиент $F(x)$ удовлетворяет условию Липшица для всех x_1, x_2 из X_1:

$$\|F_x(x_1) - F_x(x_2)\| \leqslant M \|x_1 - x_2\|. \tag{6}$$

Обозначим $c = \max\limits_{x \in X_1} \|w(x)\|$, $a = \max\limits_{x \in X_1} \|F_x + G_x w\|$. Если множество X_1 не ограничено, то максимум берется на множестве Z_1.

Теорема 2. *Если выполнены условия теоремы* 1 *и условие* (6), *тогда при $0 < \alpha < \min[1/a, 1/c, 2\lambda]$ и любых $x_0 \in X_1^0$, метод* (5) *сходится к единственному решению задачи* (1), *причем $x_s \in X_1$, $F(x_{s+1}) \leqslant F(x_s)$ для* $s = 0, 1, 2, \ldots$

Возможны варианты метода, аналогичные методу наискорейшего спуска, когда $\alpha(s)$ выбирается на каждом шаге из условия минимизации $F(x_{s+1})$ при условии $x_{s+1} \in X_1^0$. Для решения задачи (2) метод (5) можно применить лишь в том случае, если функции $g(x)$ линейные.

5. Учет ограничений типа неравенства можно проводить, используя модификацию метода штрафных функций ([3]). Составим функцию

$$H(x, w, t) = F(x) + w'g(x) + f(t) \sum_{i=1}^{k} \varphi(h^i(x));$$

здесь $w \in E_s$, $f(t)$ — положительная, дифференцируемая, строго монотонно убывающая функция t, $\lim\limits_{t\to\infty} f(t)=0$. Функция $\varphi(z)\in E_1$ выпуклая, возрастающая, дифференцируемая, положительная при $z<0$, $\lim\limits_{z\to -0}\varphi(z)=\infty$. В качестве этих функций можно взять, например, следующие:

$$f(t)=e^{-t},\quad f=(1+t)^{-1},\quad \varphi(z)=-z^{-1},\quad \varphi(z)=-z^{-2}.$$

Для решения задачи (2) воспользуемся системой

$$\dot{x}=-D(x)H_x,\quad H_x=F_x+g_xw+f(t)\sum_{i=1}^{k}\varphi'(h^i)h_x^i. \tag{7}$$

Вектор $w(t)$ выбираем таким образом, чтобы вектор-функция $g(x)$ была интегралом системы (7)

$$\dot{g}=-g_x'D(x)H_x=0. \tag{8}$$

Считаем, что векторное соотношение $g(x)=0$ определяет $(n-s)$-мерное гладкое многообразие в X_1; векторы g_x^i, h_x^i такие, что $h^i=0$, и e^i такие, что $x^i=0$, линейно независимые для всех $x\in X_2$. В этом случае система (8) однозначно разрешима относительно w. Дифференцируя H в силу (7), имеем

$$\dot{H}=-\|D(x^{1/2})H_x\|^2+\dot{f}\sum_{i=1}^{k}\varphi(h^i(x))<0.$$

Если $h(x(t))\to 0$ при $t\to t_*<\infty$, то $H\to\infty$, что невозможно в силу полученного неравенства, поэтому траектории (7) продолжимы при $t\to\infty$ и множество X_2^0 инвариантно относительно (7).

Если множество X_2 выпуклое (в этом случае функции $g(x)$ линейные), выполнены усиленные условия регулярности, тогда справедлива теорема, аналогичная теореме 1:

Теорема 3. *Пусть функции F, h и множество X_2 строго выпуклые; тогда решения* (3), (7) *при $t\to\infty$ сходятся к единственному решению задачи* (2) *при любых $x_0\in X_2^0$.*

Если функции $g(x)$ не линейные, то методы (3) и (7) сходятся к точкам, в которых выполнены необходимые условия Куна — Таккера в задаче (2).

В тех случаях, когда отсутствует ограничение $x\geqslant 0$, в методах (3) и (7) следует вместо матрицы $D(x)$ писать единичную матрицу. Если существует ограничение $x\geqslant a$, то вместо $D(x)$ следует взять $D(x-a)$.

Вычислительный центр
Академии наук СССР
Москва

Поступило
14 VI 1973

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] *Ю. Г. Евтушенко, В. Г. Жадан*, Журн. вычислит. матем. и матем. физ., т. 13, № 3 (1973). [2] *И. И. Дикин*, Управляемые системы, Новосибирск. в. 9, 1971, стр. 59. [3] *А. Фиакко, Г. Мак-Кормик*, Нелинейное программирование, М., 1972.