ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—июнь 2006. Серия 2. Том 13, № 1. 3–9

УДК 519.853.4

ГЛОБАЛЬНАЯ ОПТИМИЗАЦИЯ И МЕТОДЫ НАХОЖДЕНИЯ ВСЕХ КОРНЕЙ СИСТЕМ НЕЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

В. П. Булатов, Т. И. Белых

Предлагается численный метод нахождения всех вещественных корней системы нелинейных алгебраических уравнений, основанный на некоторой редукции исходной задачи в эквивалентную ей вспомогательную задачу. Изложенные в работе конструкции построения отсечений текущего корня системы нелинейных алгебраических уравнений аналогичны правилам построения отсечений в целочисленном программировании и могут также применяться в итерационных процессах глобальной оптимизации с использованием необходимых условий оптимальности.

Введение

Пусть $g(x) = (g_1(x), \ldots, g_n(x))$, где $g_i(x)$, $1 \leqslant i \leqslant n$, — выпуклые, непрерывно дифференцируемые функции, R^0 — выпуклое замкнутое ограниченное множество из $\mathbb{R}^n$. В [1, 2] рассмотрена задача поиска всех таких точек $\overline{x} \in R^0$, что

$$g(\overline{x}) = 0. \tag{1}$$

Корни $\overline{x} \in R^0$ системы (1) будем называть *различными*, если $\mathrm{rang}(\nabla g(\overline{x})) = n$, и кратности k, $1 \leqslant k < n$, если $\mathrm{rang}(\nabla g(\overline{x})) = n - k$.

Мы предполагаем, что число корней системы (1) на R^0 конечно и система имеет различные корни. С помощью приводимой ниже итерационной процедуры находится текущий корень системы (1). Затем этот корень отсекается с помощью дополнительного полупространства, причём на каждом шаге итерационного процесса остальные корни системы не отсекаются. Приводимые в работе конструкции построения отсечений аналогичны правилам построения правильных отсечений в целочисленном программировании [1]; они могут применяться в итерационных процессах глобальной оптимизации с использованием необходимых условий оптимальности.

© 2006 Булатов В. П., Белых Т. И.

1. Описание вычислительной процедуры поиска всех решений. Некоторые вспомогательные редукции

Поиск текущего корня системы (1) основан на использовании следующей вспомогательной конструкции [3]: найти

$$\min\{\varphi(x) \mid \ x \in R\}, \tag{2}$$

где $R = \{x \in E^n \mid g_i(x) = 0,\ x \in L,\ 1 \leqslant i \leqslant m,\ m \leqslant n\}$,

$$L = \{x \mid \ f_j(x) \leqslant 0,\ 1 \leqslant j \leqslant l\}, \tag{3}$$

где $\varphi(x) = \|x\|$, $g_i(x)$, $f_j(x)$ — гладкие функции.

Задаче (2)–(3) сопоставим следующую: найти

$$\min\{\psi(x) = ||x|| + N\sum_{i=1}^{m} g_i(x) \mid x \in R^0\}, \tag{4}$$

где

$$R^0 = \{x \mid g_i(x) \geq 0,\ f_j(x) \leqslant 0,\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant l\}, N > 0. \tag{5}$$

Введём обозначение $\Phi(x) = \begin{pmatrix} g(x) \\ f(x) \end{pmatrix}$, где $g(x) \in \mathbb{R}^m$, $f(x) \in \mathbb{R}^l$, $\Phi \in \mathbb{R}^{m+l}$.

Пусть R^0 удовлетворяет условиям регулярности и X есть множество точек минимума функции $\varphi(x) = ||x||$ на R^0, R_λ — множество оптимумов в двойственной задаче. Тогда справедливо следующее обобщение теоремы из [3].

Теорема 1. *Пусть существует $\varepsilon > 0$ такое, что $||\nabla\Phi(x)^T\lambda|| \geqslant \varepsilon||\lambda||$ при любых $x \in X$ и $\lambda \in R_\lambda$, и $||\nabla\varphi(x)|| \leqslant c < \infty$. Тогда найдётся N^* такое, что при любом $N > N^*$ каждая точка x минимума функции $\psi(x)$ на R^0 удовлетворяет равенству $g(x) = 0$.*

В [1, 3] приводятся две редукции решения задачи (1) к вспомогательным задачам, аналогичным (4), (5). Допустим, что $m = n$. Пусть построена последовательность $\{x^j\}$ такая, что $g(x^j) = 0$ при каждом $j = 1, \ldots, k$. В точке x^k определим конус

$$L^k = \{x \mid \ \nabla g(x^k)^T x \leqslant \nabla g(x^k)^T x^k\}.$$

Тогда уравнения рёбер конуса L^k имеют вид:

$$x^j = x^k - \lambda^j s^{k_j},\ \lambda^j > 0, \tag{6}$$

где s^{k_j} — столбцы матрицы, обратной к $\nabla g(x^k)$ (по условию $|\nabla g(x^k)| \neq 0$).

Из уравнений $\sum\limits_{i=1}^{n} g_i(x^k - \lambda^j s^{k_j}) = 0$ находим $\lambda^{k_j} > 0$ и соответствующие им точки $x^{k_j} = x^k - \lambda^{k_j} s^{k_j}$ пересечения лучей (6) с поверхностью $\sum\limits_{i=1}^{n} g_i(x) = 0$.

Определим отсекающее x^k полупространство $\{x \mid \beta^{k^T} x \leqslant t_k\}$ и следующее решение x^{k+1} системы нелинейных уравнений из решения задачи: найти $\min\{\psi(x) \mid x \in R^{k+1}\}$, где $R^{k+1} = \{x \mid x \in R^k,\ \beta^{k^T} x \leqslant t_k\}$.

Пусть S^k — выпуклое замкнутое множество такое, что $S^k \supset R = \{x \mid g_i(x) = 0,\ 1 \leqslant i \leqslant n\}$, причём все решения системы (1) являются граничными точками множества S^k. Пусть далее существуют $\lambda^{k_j}(0 < \lambda^{k_j} < \infty)$ и соответствующие им точки $x^{k_j} = x^k - \lambda^{k_j} s^{k_j}$ пересечения лучей (6) с границей множества S^k. Через точки x^{k_j} проведём плоскость $\{x \mid \beta^{k^T} x = t_k\}$ и определим отсекающее x^k полупространство $\{x \mid \beta^{k^T} x \leqslant t_k\}$. Тогда следующее решение x^{k+1} системы нелинейных алгебраических уравнений найдем из решения задачи: найти $\min\{\psi(x) \mid x \in R^{k+1}\}$, где $R^{k+1} = \{x \mid x \in R^k,\ \beta^{k^T} x \leqslant t_k\}$. Из геометрических соображений следует

Лемма. *Пересечение* $L^k \cap \{x \mid \beta^{k^T} x > t_k\}$ *не содержит решений системы нелинейных уравнений* (1), *отличных от* x^k.

Приведём пример конструктивного построения множеств S^k: $S = S^k(\lambda) = \{x \mid \sum\limits_{i=1}^{n} \lambda_i g_i(x) \leqslant 0$ при всех $\lambda_i > 0\}$ или, как в описанном выше алгоритме, $S = S^k(1) = \{x \mid \sum\limits_{i=1}^{n} g_i(x) \leqslant 0\}$.

Теперь рассмотрим задачу булевого программирования: найти

$$\min\{\varphi(x) \mid x \in R,\ x_i = 0 \vee 1,\ 1 \leqslant i \leqslant n\},\ R \subset \mathbb{R}^n.$$

Определим $e = (1, ..., 1) \in \mathbb{R}^n$, $R^0 = \{x \in \mathbb{R}^n,\ 0 \leqslant x_i \leqslant 1,\ 1 \leqslant i \leqslant n\}$, $g_i(x) = x_i - x_i^2$. Очевидно, ограничения $g_i(x) = x_i - x_i^2 = 0$ эквивалентны условию $x_i = 0 \vee 1,\ 1 \leqslant i \leqslant n,\ x \in R^0$.

Исходной задаче поставим в соответствие следующую задачу математического программирования: найти $\min\{\varphi(x) + N \sum\limits_{i=1}^{n} g_i(x) \mid x \in R,$ $g_i(x) \geqslant 0,\ 1 \leqslant i \leqslant n\},\ N > 0$.

Если $\varphi(x)$ — вогнутая функция и R задано системой линейных неравенств: $R = \{x \mid Ax \leqslant b\}$, то последнюю задачу можно переписать в

виде: найти
$$\min\{\varphi(x) + Nx^T(e-x) \mid x \in R \cap R^0\}.$$

Эта задача есть задача минимизации вогнутой функции на выпуклом многограннике. Известно, что существует такое N^*, что при любом $N > N^*$ множество глобальных минимумов в последней задаче совпадает с множеством оптимальных целочисленных решений исходной задачи.

Пусть требуется найти
$$\min\{\varphi(x) \mid f_i(x) = 0,\ i = \overline{1,m},\ m < n\}. \tag{7}$$

Допустим, что необходимые условия оптимальности в задаче (7) определяют систему нелинейных алгебраических уравнений с выпуклыми функциями в левых частях: $g(x,\lambda) = 0$, здесь λ — вектор двойственных переменных. Задача (7) эквивалентна поиску минимума функции $\varphi(x)$ на решениях системы $g(x,\lambda) = 0$ с выпуклыми функциями в левых частях. В случае вогнутой функции $\varphi(x)$ аналогичный итерационный процесс решения (7) приводится в [1]. Ниже даются примеры задач, сводимых к задаче (1).

1) Пусть задана система уравнений вида
$$x^T A^j x + {b^j}^T x + c_j = 0,\ 1 \leqslant j \leqslant n. \tag{8}$$

Здесь матрицы A^j, вообще говоря, не определены. Система (8) эквивалентна системе
$$\{x^j A^{1j} x + x^T A^{2j} x + (b^{1j})^T x + (b^{2j})^T x + c_j^1 + c_j^2 = 0\},\ 1 \leqslant j \leqslant n,$$

где A^{1j} неотрицательно, а A^{2j} неположительно определенные матрицы. Вводя новые переменные y_j, получим систему
$$x^T A^{1j} x + (b^{1j})^T x + c_j^1 + y_j = 0,$$
$$-x^j A^{2j} x + (b^{2j})^T x + c_j^2 + y_j = 0$$

$2n$ уравнений с выпуклыми функциями в левых частях.

2) Пусть дана система уравнений
$$\varphi_{1j}(x) + \varphi_{2j}(x) = 0,\ 1 \leqslant j \leqslant n, \tag{10}$$

где $\varphi_{1j}(x)$ — выпуклые, а $\varphi_{2j}(x)$ — вогнутые функции. Очевидно, система (10) эквивалентна системе
$$\varphi_{1j}(x) + y_j = 0,\ \ -\varphi_{2j}(x) + y_j = 0$$

$2n$ уравнений с выпуклыми функциями в левых частях.

2. Численные методы глобальной оптимизации на основе необходимых условий оптимальности

Применим теперь изложенную технику для решения более общих задач математического программирования: найти

$$\max\{\varphi(x) \mid \ q_j(x) \leqslant 0,\ 1 \leqslant j \leqslant m\}. \tag{11}$$

Здесь $\varphi(x)$ — выпуклая гладкая функция, $q_j(x),\ 1 \leqslant j \leqslant m,$ — гладкие функции, максимум понимается в глобальном смысле.

Запишем в форме равенства $q_j(x)+y_j^2 = 0,\ 1 \leqslant j \leqslant m$, ограничения в задаче (11), сопоставим ей функцию Лагранжа и выпишем необходимые условия экстремума. Тогда эквивалентная (11) задача будет иметь вид: найти

$$\max\{\varphi(x) \mid \ q_j(x) + y_j^2 = 0, \\ \nabla\varphi(x) + \lambda^T \nabla q(x) = 0;\ \lambda_j y_j = 0,\ 1 \leqslant j \leqslant m\}. \tag{12}$$

В обозначениях

$$\frac{\partial\varphi}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial q_j}{\partial x_i} = \psi_i,\ 1 \leqslant i \leqslant n;\ q_j(x) + y_j^2 = \psi_j;\ j = n+1, ..., n+m;$$

$$2\lambda_j y_j = \psi_j;\ j = n+m+1, ..., n+2m$$

получим

$$\max\{\varphi(x) \mid \ \psi_j(x, y, \lambda) = 0; j = 1, ..., n+2m\}. \tag{13}$$

Допустимое множество последней задачи задано замкнутой системой нелинейных алгебраических уравнений. Среди решений системы нас интересуют лишь решения, принадлежащие выпуклому замкнутому компакту R^0.

Предположим, что множество решений системы $\psi_j(x, y, \lambda) = 0$ конечно на R^0, т. е. рассматриваются лишь изолированные корни. Тогда исходная задача (11) эквивалентна задаче поиска максимума выпуклой функции $\varphi(x)$ на конечном множестве $R = \{x \mid \ x \in R^0,\ \psi_j(x, y, \lambda) = 0,\ j = 1, ..., n+2m\}$.

Данная задача исследовалась ранее в работах [1, 3], в которых предложены различные редукции этой задачи и методы отсечения.

Задаче (13) поставим в соответствие следующую вспомогательную задачу поиска локального максимума выпуклой функции аналогично тому,

как это сделано ранее: найти

$$\max\{\varphi(x) + N \sum_{j=1}^{n+2m} \psi_j(x, y, \lambda) \mid \ x \in R\}, \tag{14}$$

$$R = \{(x, y, \lambda) \mid \ \psi_j(x, y, \lambda) \leqslant 0, \ j = 1, ..., n + 2m; \ x \in R^0\}, \ N > 0. \tag{15}$$

Согласно теореме 1 при любом достаточно большом N каждое локальное оптимальное решение (x^*, y^*, λ^*) задачи (14), (15) удовлетворяет условию $\psi_j(x^*, y^*, \lambda^*) = 0$ при любом $j = 1, ..., n + 2m$.

Допустим, что задача (14), (15) не вырождена, т. е. $|\nabla\psi_j(x^*, y^*, \lambda^*)| \neq 0$ на любом её решении, и кроме того, вектор-функция $\psi_j(x, y, \lambda)$ выпукла по совокупности своих аргументов. Нетрудно видеть, что в этом случае поиск локального решения в задаче (14), (15) сводится к последовательной максимизации на выпуклом множестве линейных функций [1—3]. Перейдем теперь к применению описанного выше метода для решения задачи (13).

Итак, пусть уже найдена точка $z^k = \{x^k, y^k, \lambda^k\}$, являющаяся решением следующей задачи математического программирования типа (14), (15): найти

$$\max\{\varphi(x) + N \sum_{j=1}^{n+2m} \psi_j(x, y, \lambda) \mid \ (x, y, \lambda) \in R^k\},$$

$$R^k = \{(x, y, \lambda) \mid \ \psi_j(x, y, \lambda) \leqslant 0, \ j = 1, ..., n + 2m; \ x \in R^0\},$$
$$s^{iT} z \leqslant t_i, \ i = 1, ..., k.$$

Пусть $\beta_k = \max\{\varphi(x^k), \beta_{k-1}\}$. В точке z^k определим конус

$$L^k = \{z \mid \ \nabla\psi(x^k, y^k, \lambda^k)^T z \leqslant \nabla\psi(x^k, y^k, \lambda^k)^T z^k\}. \tag{16}$$

В силу предположения матрица $\nabla\psi(x^k, y^k, \lambda^k)$ невырожденная. Тогда уравнения рёбер конуса (16) будут иметь вид

$$z^j = z^k - \lambda^j s^{kj}, \ j = 1, ..., n + 2m; \ z^k = \{x^k, y^k, \lambda^k\}, \ \lambda^j > 0, \tag{17}$$

где s^{kj} — столбцы матрицы, обратной к $\nabla\psi(z^k)$.

Из уравнений $\psi(z^k - \lambda^j s^{kj}) = \beta_k$, где $\psi(z) = \varphi(x) + N \sum\limits_{j=1}^{n+2m} \psi_j(z)$, определяем параметры λ^{kj} и соответствующие им точки z^{kj} пересечения

лучей из (17) с границей множества $\{z \mid \ \psi(z) \leqslant \beta_k\}$. В силу невырожденности матрицы $\nabla\psi(x^k, y^k, \lambda^k)$ при j, $1 \leqslant j \leqslant n+2m$, имеем $\lambda^{kj} > 0$. Через точки z^{kj} проведём плоскость и построим полупространство $\{z \mid \ (s^{k+1})^T z \leqslant t_{k+1}\}$, не содержащее z^k. Очередное приближение z^{k+1} получим из решения задачи: найти

$$\max\{\psi(z) \mid \ z \in R^k, \ (s^{k+1})^T z \leqslant t_{k+1}\},$$

причём $\psi_j(z^{k+1}) = 0$ при любом $j = 1, ..., n+2m$. Через конечное число шагов получим решение задачи (13).

Замечание. Приводимое в работе предположение о выпуклости функции $\psi_j(x, y, \lambda)$ по совокупности переменных не столь обременительно, как кажется на первый взгляд. Если $\psi_j(x)$ и $q_i(x)$ в задаче заданы произвольными «отрезками» степенного ряда, то введением новых переменных за счёт увеличения размерности можно построить эквивалентную задачу типа (13) с выпуклыми функциями в левых частях системы ограничений.

ЛИТЕРАТУРА

1. **Булатов В. П.** Методы погружений в задачах оптимизации. Новосибирск: Наука, 1997.
2. **Булатов В. П.** Глобальная оптимизация и численные методы поиска всех решений систем нелинейных алгебраических уравнений // Журнал вычислит. матем. и матем. физики. 2000. Т. 40, № 3. С. 348–355.
3. **Булатов В. П., Алексеева Т. А.** Численные методы поиска всех решений систем нелинейных алгебраических уравнений // Методы оптимизации и их приложения. Труды XI Междунар. Байкальской школы-семинара. Т. 1. Иркутск, 1998. С. 21–32.

Адреса авторов:

В. П. Булатов

Институт систем энергетики им. Л.А.Мелентьева СО РАН,
ул.Лермонтова, 130,
664033 Иркутск, Россия.
E-mail: bulatov@isem.sei.irk.ru

Т. И. Белых

Байкальский государственный университет экономики и права,
ул. Карла Маркса, 16,
664001 Иркутск, Россия.

Статья поступила
26 октября 2004 г.

Переработанный вариант —
20 мая 2006 г.