ДОКЛАДЫ АКАДЕМИИ НАУК, 2006, том 408, № 4, с. 439–442

МАТЕМАТИКА

УДК 519.852.2

2-ФАКТОР-МЕТОД МОДИФИЦИРОВАННЫХ ФУНКЦИЙ ЛАГРАНЖА ДЛЯ РЕШЕНИЯ ВЫРОЖДЕННЫХ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ

©2006 г. О.А. Брежнева*, член-корреспондент РАН Ю. Г. Евтушенко, А.А. Третьяков**,*****

* Miami University, Ohio, USA

** Вычислительный центр им. А.А. Дородницына Российской Академии наук, Москва

*** Академия Подляска, Сиедлице, Польша;
Институт системных исследований Польской академии наук, Варшава, Польша

Поступило 28.02.2006 г.

С помощью специальной модификации функции Лагранжа задача нелинейного программирования сводится к решению системы нелинейных уравнений, в общем случае вырожденной. С использованием p-фактор-преобразования эта система редуцируется к новой невырожденной системе, имеющей то же самое решение и к которой становятся применимы методы ньютоновского типа. Таким образом, предлагаемый метод представляет собой синтез варианта метода модифицированных функций Лагранжа (МФЛ), предложенного Ю.Г. Евтушенко в [3], и 2-фактор-метода А.А. Третьякова, описание которого можно найти, например, в работе [1].

Нелинейные системы традиционно разделяются на два класса: регулярные и нерегулярные. Регулярными называются системы, к которым применима классическая теорема о неявной функции, и нерегулярными — те, к которым теорема о неявной функции неприменима. В этом сообщении мы имеем дело с нерегулярными системами.

Рассмотрим задачу нелинейного программирования

$$\min_{x\in X} f(x). \tag{1}$$

Здесь допустимое множество $X = \{x \in \mathbb{R}^n \mid g(x) \le 0_m\}$, где 0_m — нулевой вектор из $\mathbb{R}^m$, $(g(x))^\top = (g_1(x), g_2(x), \dots, g_m(x))$ — вектор-функция (строка), функции $f(x)$ и $g_j(x)$ отображают $\mathbb{R}^n$ в пространство $\mathbb{R}$.

Функция Лагранжа для задачи (1) имеет вид $\mathcal{L}(x, v) = f(x) + v^\top g(x)$, где $v \in \mathbb{R}^m_+$ — вектор множителей Лагранжа. Предполагая, что функции $f(x)$ и $g(x)$ дважды непрерывно дифференцируемы, определим градиент и матрицу Гессе для функции Лагранжа:

$$\begin{aligned}\nabla_x \mathcal{L}(x, v) &= \nabla f(x) + \sum_{i=1}^m v_i \nabla g_i(x),\\ \nabla_{xx} \mathcal{L}(x, v) &= \nabla^2 f(x) + \sum_{i=1}^m v_i \nabla^2 g_i(x).\end{aligned}$$

Относительно задачи (1) считаем, что множество ее решений $X^* \subset \mathbb{R}^n$ непусто. Всюду ниже предполагаем, что выполнено условие регулярности ограничений (УРО) или, иными словами, градиенты активных ограничений $\nabla g_i(x^*)$ линейно-независимы. Это условие гарантирует, что каждому $x^* \in X^*$ соответствует единственный вектор множителей Лагранжа $v^* \in V^*$, для которого выполнено соотношение $\nabla_x \mathcal{L}(x^*, v^*) = 0_n$ и $v_i^* = 0$, если $g_i(x^*) > 0$, $i = 1, 2, \ldots, m$.

Рассмотрим предложенный в [3] нестандартный вариант МФЛ, в котором модифицированная функция Лагранжа имеет вид

$$\mathcal{L}_E(x, \lambda) = f(x) + \frac{1}{2} \sum_{i=1}^{m} \lambda_i^2 g_i(x), \tag{2}$$

где $\lambda^\top = (\lambda_1, \lambda_2, \ldots, \lambda_m)$.

Очевидно, что i-я компонента вектора множителей Лагранжа v выражается через i-ю компоненту нового вектора λ по формуле $v_i = \dfrac{(\lambda_i)^2}{2}$. Таким образом, использование вектора λ автоматически обеспечивает неотрицательность соответствующего вектора множителей Лагранжа v.

Если $x^* \in X^*$, то ему соответствует вектор λ^* с компонентами $\lambda_i^* = \pm\sqrt{2v_i^*}$. Объединим векторы x и λ одним символом $w \in \mathbb{R}^{n+m}$. Аналогично пару $[x^*, \lambda^*]$ будем обозначать w^*, поэтому $\mathcal{L}_E(x, \lambda) = \mathcal{L}_E(w)$. Согласно теореме Куна–Таккера, вектор w^* удовлетворяет системе

$$G(w) = \begin{bmatrix} \nabla f(x) + \frac{1}{2} \sum\limits_{i=1}^{m} \lambda_i^2 \nabla g_i(x) \\ D(\lambda) g(x) \end{bmatrix} = 0_{m+n}. \tag{3}$$

Здесь $D(\lambda)$ — диагональная матрица, размерность которой определяется размерностью вектора λ, у нее i-й диагональный элемент есть λ_i. Отметим, что система уравнений (3) может иметь, вообще говоря, бесконечное множество решений даже в окрестности точки w^*. Пусть $\nabla g(x)^\top$ — матрица Якоби отображения $g(x)$. Для системы (3) матрица Якоби имеет вид

$$G'(w) = \left[\begin{array}{c|c} \nabla^2 f(x) + \frac{1}{2} \sum\limits_{i=1}^{m} \lambda_i^2 \nabla^2 g_i(x) & \nabla g(x) D(\lambda) \\ \hline D(\lambda) \nabla g^\top(x) & D(g(x)) \end{array}\right]. \tag{4}$$

Для пары $[x^*, \lambda^*]$ определим множество активных ограничений $I(x^*)$, множество слабо активных ограничений $I_0(x^*)$ и множество строго активных ограничений $I_+(x^*)$ следующим образом:

$$\begin{aligned} I(x^*) &= \{j = 1, 2, \ldots, m \mid g_j(x^*) = 0\}, \\ I_0(x^*) &= \{j = 1, 2, \ldots, m \mid \lambda_j^* = 0, \;\; g_j(x^*) = 0\}, \\ I_+(x^*) &= \{j = 1, 2, \ldots, m \mid \lambda_j^* \neq 0, \;\; g_j(x^*) = 0\}. \end{aligned}$$

При обосновании и анализе метода МФЛ обычно помимо УРО вводятся следующие условия:

1) условие строгой дополняющей нежесткости (УСДН), т.е. $\lambda_i^* g_i(x^*) = 0$, $i = 1, 2, \ldots, m$, и если $g_i(x^*) = 0$, то $\lambda_i^* \neq 0$ для всех $i = 1, 2, \ldots, m$;

2) достаточные условия оптимальности 2-го порядка: существует число $\nu > 0$ такое, что

$$z^\top \nabla_{xx}^2 \mathcal{L}_E(x^*, \lambda^*) z \geq \nu \|z\|^2 \tag{5}$$

для всех $z \in \mathbb{R}^n$, для которых $\nabla g_j(x^*)^\top z \leq 0$, $j \in I(x^*)$.

Предположим, что в точке x^* не выполнено УСДН. Тогда для некоторого индекса i выполнены оба равенства $\lambda_i^* = 0$ и $g_i(x^*) = 0$, поэтому множество $I_0(x^*)$ непусто. В этом случае матрица (4) становится вырожденной в точке w^* и, следовательно, методы типа Ньютона для решения системы уравнений (3) неприменимы. Главная цель сообщения — показать, что в этой ситуации эффективным будет применение теории p-регулярности, основные положения которой даны в [4, 5].

Рассмотрим систему нелинейных уравнений (3). Пусть отображение G нерегулярно в точке w^*, или, другими словами, матрица Якоби (4) вырождена и $\operatorname{rank}(G'(w^*)) = r < n+m$. В этом случае w^* называется вырожденным решением системы (3).

Условие вырожденности матрицы $G'(w^*)$ означает, что существует по крайней мере один ненулевой вектор h такой, что

$$G'(w^*)h = 0_{m+n}. \tag{6}$$

Очевидно, что для такого вектора h решение системы (3) будет также решением следующей модифицированной системы:

$$\Phi(w) = G(w) + G'(w)h = 0_{m+n}. \tag{7}$$

Если матрица $G'(w^*)$ вырождена, то ниже, в лемме 2, показано, что матрица $\Phi'(w^*) = G'(w^*) + G''(w^*)h$ не вырождена и, следовательно, решение w^* системы (7) локально единственно. Свойство невырожденности матрицы $\Phi'(w^*)$ положено в основу конструкции 2-фактор-метода решения вырожденных систем нелинейных уравнений.

Рассмотрим 2-фактор-оператор: $G'(w) + G''(w)h$, $h \in \mathbb{R}^{n+m}$, $\|h\| \neq 0$, где вектор h удовлетворяет условию

$$\operatorname{rank}\left[G'(w^*) + G''(w^*)h\right] = n + m. \tag{8}$$

Конкретный вид вектора h формируется с учетом специфики системы (3). Отметим, что 2-фактор-оператор может быть определен разными способами (см., например, [4, 5]). В данной работе мы используем наиболее удобную форму.

Определение 1. *Отображение G называется* **2-регулярным в точке w^* на некотором элементе** $h \in \mathbb{R}^{n+m}$, *если выполнено условие* (8).

Рассмотрим итерационный процесс решения системы (3), называемый 2-фактор-методом:

$$w^{k+1} = w^k - [G'(w^k) + G''(w^k)h]^{-1}[G(w^k) + G'(w^k)h], \qquad k = 0, 1, \ldots, \tag{9}$$

где w^0 — начальное приближение из достаточно малой окрестности точки w^*.

Теорема 1. *Пусть w^* — решение системы* (3), *$U_\varepsilon(w^*)$ — некоторая достаточно малая окрестность w^*, отображение $G \in C^3(\mathbb{R}^{n+m} \to \mathbb{R}^{n+m})$ и 2-регулярно в точке w^* на некотором ненулевом элементе $h \in \mathbb{R}^{n+m}$, удовлетворяющем* (6).

Тогда метод (9) *сходится к точке w^* и имеет место оценка скорости сходимости*

$$\|w^{k+1} - w^*\| \leq \alpha \|w^k - w^*\|^2, \tag{10}$$

где $\alpha > 0$ — независимая константа и $w^0 \in U_\varepsilon(w^*)$.

Из приводимой ниже леммы 2 следует, что при выполнении достаточных условий оптимальности для задачи (1) отображение G, определенное в (3), является 2-регулярным в точке w^* на некотором элементе h. Следовательно, для решения системы $G(w) = 0_{m+n}$ можно применять 2-фактор-метод. В силу теоремы 1 метод обладает квадратичной скоростью сходимости.

Допускаем, что УСДН в точке x^* может быть не выполнено, и без ограничения общности считаем, что множество $I_0(x^*)$ состоит из первых s индексов, т.е. $I_0(x^*) = \{1, 2, \ldots, s\}$.

Для численного определения множества $I_0(x^*)$ можно использовать процедуру идентификации нулевых элементов [2]. Кроме того, предположим, что $I_+ = \{s+1, s+2, \dots, p\}$, и обозначим $\ell = m - p$. Так как $\lambda_j^* = 0$ и $g_j(x^*) = 0$ для всех $j = 1, 2, \dots, s$, то строки и столбцы матрицы $G'(w^*)$ с $(n+1)$-го номера по $(n+s)$-й состоят из одних нулей. Определим вектор $h \in \mathbb{R}^{n+m}$ как $h^\top = (0_n^\top, 1_s^\top, 0_{m-s}^\top)$, где $1_s^\top$ — это s-мерный вектор-строка, у которого все компоненты равны 1. Рассмотрим отображение

$$\Phi(w) = G(w) + G'(w)h. \tag{11}$$

Лемма 1. *Пусть V — матрица размерности $n \times n$ и Q — матрица размерности $n \times p$, такие, что столбцы матрицы Q линейно-независимы и $\langle Vx, x\rangle > 0$ для всех $x \in \{\ker Q^\top\} \backslash \{0\}$. Предположим также, что G_N — диагональная матрица полного ранга размерности $\ell \times \ell$.*

Тогда матрица

$$\bar{A} = \begin{pmatrix} V & Q & 0 \\ Q^\top & 0 & 0 \\ 0 & 0 & G_N \end{pmatrix} \tag{12}$$

не вырождена.

Лемма 2. *Пусть f, $g_i \in C^3(\mathbb{R}^n)$, $i = 1, 2, \dots, m$, отображение Φ задается формулой* (11), *выполнены УРО и достаточные условия оптимальности 2-го порядка* (5).

Тогда 2-фактор-оператор $\Phi'(w) = G'(w) + G''(w)h$ не вырожден в точке w^*.

Доказательство вытекает из леммы 1, если в ней положить $V = \nabla^2_{xx}\mathcal{L}_E(x^*, \lambda^*)$, $G_N = D(g_N(x^*))$, где $g_N(x) = (g_{p+1}(x), g_{p+2}(x), \dots, g_m(x))^\top$, и

$$Q = [\nabla g_1(x^*), \dots, \nabla g_s(x^*), \lambda^*_{s+1}\nabla g_{s+1}(x^*), \dots, \lambda^*_p \nabla g_p(x^*)].$$

Тогда $\Phi'(w^*) = \bar{A}$.

Из леммы 2 следует, что для решения системы (3) можно использовать 2-фактор-метод (9). Применяя теорему 1 к решению исходной задачи (1), приходим к следующему результату.

Теорема 2. *Предположим, что точка x^* — решение задачи* (1). *Пусть f, $g_i \in C^3(\mathbb{R}^n)$, $i = 1, 2, \dots, m$, выполнены УРО и достаточные условия оптимальности 2-го порядка* (5).

Тогда существует достаточно малая окрестность $U_\varepsilon(w^*)$ точки Куна–Таккера $w^* = [x^*, \lambda^*]$, такая, что для метода (9) *имеет место оценка* (10).

Отметим, что условие вырожденности матрицы $G'(w^*)$ дает возможность построить целый класс 2-фактор-методов, используя условие $PG'(w^*)h = 0$, где h — ненулевой вектор из пространства $\mathbb{R}^{n+m}$ и P — ортопроектор на ортогональное дополнение к образу $G'(w^*)$. В этом случае условие 2-регулярности имеет вид $\operatorname{rank}[G'(w^*) + PG''(w^*)h] = n + m$ и схема 2-фактор-метода записывается следующим образом:

$$w^{k+1} = w^k - [G'(w^k) + PG''(w^k)h]^{-1} \times [G(w^k) + PG'(w^k)h]. \tag{13}$$

Для метода (13) будет также справедлива теорема о сходимости и скорости сходимости. Отличие состоит лишь в том, что здесь не требуется вычислять вектор $h \in \operatorname{Ker} G'(w^*)$. Однако этот случай требует задания дополнительной матрицы P.

Проиллюстрируем применение описанного метода на следующем примере.

Пример. Рассмотрим задачу

$$\min_{x \in \mathbb{R}^2} (x_1^2 + x_2^2 + 4x_1x_2) \tag{14}$$

при ограничениях $x_1 \geq 0$, $x_2 \geq 0$.

Легко проверить, что точка $x^* = (0,0)^\top$ — решение этой задачи с соответствующим множителем Лагранжа $v^* = (0,0)^\top$. В этом примере множество $I_0(x^*) = \{1,2\}$ и модифицированная функция Лагранжа имеет вид $\mathcal{L}_E(x,\lambda) = x_1^2 + x_2^2 + 4x_1x_2 - \frac{1}{2}\lambda_1^2 x_1 - \frac{1}{2}\lambda_2^2 x_2$. Пусть $h = (0,0,1,1)^\top$, тогда система (3) запишется следующим образом:

$$G(w) = \begin{bmatrix} 2x_1 + 4x_2 - \frac{1}{2}\lambda_1^2 \\ 2x_2 + 4x_1 - \frac{1}{2}\lambda_2^2 \\ -\lambda_1 x_1 \\ -\lambda_2 x_2 \end{bmatrix} = 0_4. \tag{15}$$

Матрица Якоби для системы (15) имеет вид

$$G'(w) = \begin{bmatrix} 2 & 4 & -\lambda_1 & 0 \\ 4 & 2 & 0 & -\lambda_2 \\ -\lambda_1 & 0 & -x_1 & 0 \\ 0 & -\lambda_2 & 0 & -x_2 \end{bmatrix}.$$

Эта матрица вырождена в точке $w^* = (0,0,0,0)^\top$. Отображение G является 2-регулярным в точке w^* на введенном элементе h, и схема 2-фактор-метода запишется как

$$\begin{bmatrix} 2 & 4 & -\lambda_1 - 1 & 0 \\ 4 & 2 & 0 & -\lambda_2 - 1 \\ -\lambda_1 - 1 & 0 & -x_1 & 0 \\ 0 & -\lambda_2 - 1 & 0 & -x_2 \end{bmatrix} \begin{bmatrix} x_1^{k+1} - x_1 \\ x_2^{k+1} - x_2 \\ \lambda_1^{k+1} - \lambda_1 \\ \lambda_2^{k+1} - \lambda_2 \end{bmatrix} = - \begin{bmatrix} 2x_1 + 4x_2 - \frac{1}{2}\lambda_1^2 - \lambda_1 \\ 2x_2 + 4x_1 - \frac{1}{2}\lambda_2^2 - \lambda_2 \\ -\lambda_1 x_1 - x_1 \\ -\lambda_2 x_2 - x_2 \end{bmatrix},$$

где $k = 0, 1, \ldots$, $(x_1, x_2, \lambda_1, \lambda_2)^\top = (x_1^k, x_2^k, \lambda_1^k, \lambda_2^k)^\top$. ■

В данном примере система (7) имеет неединственное решение, поэтому отсутствует глобальная сходимость предложенного метода. Здесь можно использовать другую версию метода, решая линейную относительно w систему $\Psi(w) = G'(w)h = 0_4$, где $h \in \operatorname{Ker} G'(w^*)$. В силу 2-регулярности отображения G на элементе h в точке w^* матрица $\Psi'(w^*)$ не вырождена, и, следовательно, эта задача имеет единственное решение $w^* = 0_{n+m}$.

Работа выполнена на основе совместных исследований, проводимых в Вычислительном центре Российской Академии наук (Москва, Россия) по программе поддержки ведущих научных школ РФ (грант НШ-2240.2006.1), Академии Подляска (Сиедлице, Польша) и Институте системных исследований Польской академии наук (Варшава, Польша).

СПИСОК ЛИТЕРАТУРЫ

1. *Белаш К.Н., Третьяков А.А.* // ЖВМиМФ. 1988. Т. 28. С. 1097–1102.

2. *Брежнева О.А., Третьяков А.А.* Методы решения существенно нелинейных задач. М.: ВЦ РАН, 2000.

3. *Evtushenko Yu.G.* // J. Optim. Theory and Appl. 1977. V. 21. №2. P. 121–135.

4. *Измаилов А.Ф., Третьяков А.А.* Фактор-анализ нелинейных отображений. М.: Наука, 1994.

5. *Tret'yakov A.A., Marsden J.E.* // Communs Pure and Appl. Anal. 2003. V. 2. P. 425–445.