ДАН, 2012, т. 442, No.2, с. 156–161

МАТЕМАТИКА

УДК 519.626

Элементарное доказательство конструктивного варианта теоремы о касательных направлениях и теоремы о неявной функции

Академик Ю.Г. Евтушенко[1], А.А. Третьяков[1, 2]

[1] *119991 Москва, ул. Вавилова, 40, ВЦ РАН;*
[2] *System Res. Inst., Polish Acad. Sie, Newelska 6, 01-447 Warsaw, and University of Podlasie, 08-110 Siedlce, Poland*
e-mail: evt@ccas.ru, tret@ap.siedlce.pl

Дана формула, определяющая семейство приближенных решений системы нелинейных уравнений, и с ее помощью элементарно доказывается теорема о касательных направлениях и теорема о неявной функции. В предложенном доказательстве используется факт равенства нулю градиента отображения типа квадрата нормы в его стационарной точке, а также хрестоматийный результат безусловной оптимизации — теорема Ферма о необходимых условиях экстремума — и теорема Вейерштрасса о достижении непрерывной функцией минимума на компакте. Эти результаты доказываются для конечномерного случая и распространяются на банахово пространство. Приводится формула аппроксимации неявной функции с погрешностью третьего порядка малости. Построены итерационные методы решения систем уравнений и задач нелинейного программирования с ограничениями типа равенства, обладающие локальной кубической скоростью сходимости. В качестве приложения дается элементарное доказательство теоремы Лагранжа о необходимых условиях экстремума в задаче НЛП с ограничениями-равенствами.

1. Пусть задана вектор-функция $F(x) : \mathbb{R}^n \to \mathbb{R}^m$, $F^\top(x) = [f^1(x), f^2(x), \ldots, f^m(x)]$, где $f^i(x) : \mathbb{R}^n \to \mathbb{R}^1$, $i = \overline{1, m}$, $m \le n$, $F(x) \in C^3(\mathbb{R}^n)$. Введем допустимое множество

$$X = \{x \in \mathbb{R}^n \mid F(x) = 0_m\}. \tag{1}$$

В точке $x \in \mathbb{R}^n$ определим прямоугольную $m \times n$ матрицу Якоби $F_x(x)$, ее ядро и образ

$$\begin{array}{lcl} \operatorname{Ker} F_x(x) & = & \{h \in \mathbb{R}^n \mid F_x(x)h = 0_m\}, \\ \operatorname{Im} F_x(x) & = & \{\eta \in \mathbb{R}^m \mid \eta = F_x(x)y, \quad y \in \mathbb{R}^n\} = (\operatorname{Ker} F_x(x))^\perp . \end{array} \tag{2}$$

Ниже понадобится следующая

Лемма 1. *Пусть $\varphi(x) = \|F(x)\|^2$, где $F(x) \in C^1(\mathbb{R}^n)$, $F : \mathbb{R}^n \to \mathbb{R}^m$. Тогда множеству X принадлежит любая стационарная точка x_* функции $\varphi(x)$, в которой $\varphi_x(x_*) = 0_n$ и строки матрицы $F_x(x_*)$ линейно независимы.*

Доказательство следует из формулы $\varphi_x(x_*) = 2F_x^\top(x_*)F(x_*) = 0_n$ и из того, что здесь столбцы матрицы $F_x^\top(x_*)$ линейно независимы. Поэтому равенство нулю может иметь место только если $F(x_*) = 0_m$, т.е. при $x_* \in X$. $\square$

Пусть ранг матрицы $F_x(x)$ равен m. Тогда существует невырожденная $m \times m$ матрица Грамма $G(x) = F_x(x)F_x^\top(x)$ и матрица $F_x^+(x) = F_x^\top(x)G^{-1}(x)$, псевдообратная (правая обратная) к $F_x(x)$, т.е. $F_x(x)F_x^+(x) = I_m$. Определим два n-мерных вектора

$$N(x) = F_x^+(x)F(x), \qquad p(x) = -\frac{1}{2}\, F_x^+(x)h^\top F_{xx}(x)h. \tag{3}$$

Здесь $h \in \mathbb{R}^n$, $h^\top F_{xx}(x_*)h$ — m-мерный вектор, i-я компонента которого есть $h^\top f^i_{xx}(x_*)h$.

Теорема 1. *Пусть точка $x_* \in X$ такова, что столбцы матрицы $F_x^\top(x_*)$ линейно независимы, $h \in \operatorname{Ker} F_x(x_*)$. Тогда в $\mathbb{R}^n$ существует семейство дуг*

$$x(\alpha) = x_* + \alpha h + \alpha^2 p(x_*), \tag{4}$$

исходящих из x_* и с точностью до величин второго порядка малости принадлежащих множеству X.

Доказательство. Матрица Грамма $G(x_*)$ не вырождена и, следовательно, при любых $\eta \in \mathbb{R}^m$, $\eta \neq 0$, имеет место неравенство

$$\inf_\eta \|G(x_*)\eta\|/\|\eta\| \geq \gamma,$$

где $\gamma = \|G^{-1}(x_*)\|$ — нижняя грань матрицы $G(x_*)$, $\gamma > 0$.

Рассмотрим семейство дуг

$$z(\alpha,\eta) = x(\alpha) + \alpha^3 F_x^\top(x_*)\eta, \qquad \eta \in \mathbb{R}^m.$$

Используя дифференцируемость $F(x)$ и формулу (4), получим

$$F(z(\alpha,\eta)) = F(x_*) + \alpha \left[F_x(x_*)(h + \alpha p(x_*) + \alpha^2 F_x^\top(x_*)\eta) + \frac{\alpha}{2}\, h^\top F_{xx}(x_*)h\right] + O(\alpha^3),$$

где $O(\alpha^3)$ — m-мерный вектор, все компоненты которого имеют третий порядок малости. Учтем определения (2) и (3), тогда

$$F(z(\alpha,\eta)) = \alpha^3 G(x_*)\eta + O(\alpha^3).$$

Возьмем нормы левой и правой частей, оценим их, положив $\|\eta\| = \rho$, получим

$$\|F(z(\alpha,\eta))\| \geq \alpha^3 \|G(x_*)\eta\| - \|O(\alpha^3)\| \geq \gamma\rho\alpha^3 - \|O(\alpha^3)\|.$$

Можно выбрать ρ достаточно большим, таким, что $\gamma\rho\alpha^3 \geq 2\|O(\alpha^3)\|$. При этом ρ не зависит от α и

$$\|F(z(\alpha,\eta))\| \geq \frac{1}{2}\,\gamma\rho\alpha^3.$$

Обозначив $\Psi(\eta) = \|F(z(\alpha,\eta))\|^2$, при $\rho > 0$ имеем

$$\Psi(\eta) \geq \frac{1}{4}\,\gamma^2\rho^2\alpha^6 > \Psi(0_m) = O(\alpha^6).$$

Согласно теореме Вейерштрасса функция $\Psi(\eta)$ достигает минимума внутри шара радиуса ρ при некотором $\bar\eta(\alpha)$, и по теореме Ферма

$$\Psi_\eta(\bar\eta(\alpha)) = 2BF\left(x(\alpha) + \alpha^3 F_x^\top(x_*)\bar\eta(\alpha)\right) = 0_m, \tag{5}$$

где $\|\bar\eta(\alpha)\| < \rho$ и квадратная матрица $B = F_x\left(x(\alpha) + \alpha^3 F_x^\top(x_*)\bar\eta(\alpha)\right) F_x^\top(x_*)$.

Из невырожденности матрицы Грамма следует невырожденность матрицы B при достаточно малых α, поэтому из (5) и леммы 1 следует, что

$$F\left(x(\alpha) + \alpha^3 F_x^\top(x_*)\bar\eta(\alpha)\right) = 0_m.$$

Отсюда заключаем, что существует точка $x(\alpha)+\alpha^3 F_x^\top(x_*)\bar\eta(\alpha)$, принадлежащая множеству X и такая, что евклидово расстояние $\mathrm{dist}\,(x(\alpha), X)$ от точки x до множества X не превышает $\|x(\alpha)-z(\alpha,\bar\eta(\alpha))\|\le c\,\alpha^3$. Теорема доказана. $\square$

Приведенное доказательство использует тот факт, что из равенства нулю производной квадрата вещественной функции в точке, где ее производная не равна нулю, следует, что сама функция в этой точке равна нулю. Если ядро матрицы $F_x(x_*)$ тривиально, что имеет место, например, при $m=n$, то семейство дуг $x(\alpha)$ вырождается в точку x_*. Близкий подход был использован в книге [8].

2. Формулы (3) и (4) дают возможность построить итерационный метод решения системы $F(x)=0_m$, в которой $m\le n$. Чтобы сохранить аппроксимацию отображения $F(x)$ с погрешностью $O(\|x-x_*\|^3)$ в формуле (4), в качестве векторов $x(\alpha)$, x_*, αh, $\alpha^2 p(x_*)$ возьмем соответственно векторы x_{i+1}, x_i, $-N(x_i)$, $-\dfrac{1}{2}\,F_x^+(x_i)N^\top(x_i)F_{xx}(x_i)N(x_i)$. В результате получим итерационный метод

$$x_{i+1}=x_i-N(x_i)-\frac{1}{2}\,F_x^+(x_i)N^\top(x_i)F_{xx}(x_i)N(x_i). \tag{6}$$

Теорема 2. *Пусть $F\in C^2(\mathbb{R}^n)$ и существует точка $x_*\in X$, в которой ранг матрицы $F_x(x_*)$ равен m. Тогда существует окрестность $U(x_*)$ точки x_*, такая что для любой точки $x_i\in U(x_*)$ имеет место локальная сходимость итерационного процесса* (6) *с кубической скоростью, при этом*

$$\mathrm{dist}\,(x_{i+1},X)\le c\ \mathrm{dist}\,(x_i,X)^3,$$

где $c>0$ — независимая константа.

Доказательство. Рассмотрим функцию

$$\Psi(\eta)=\|F(x_i-N(x_i)-\frac{1}{2}\,F_x^+(x_i)N^\top(x_i)F_{xx}(x_i)N(x_i)+F_x^\top(x_i)\eta)\|^2.$$

Аналогично тому, как это делалось в теореме 1, доказывается существование точки $\bar\eta$, такой что

$$\Psi_\eta(\bar\eta)=0_m \qquad \text{и} \qquad \|\bar\eta\|\le c\,\|N(x_i)\|^3.$$

Следовательно, вектор

$$x_i-N(x_i)-\frac{1}{2}\,F_x^+(x_i)N^\top(x_i)F_{xx}(x_i)N(x_i)+F_x^\top(x_i)\bar\eta$$

принадлежит множеству X, т.е.

$$\mathrm{dist}\,(x_{i+1},X)\le c\,\|N(x_i)\|^3\le\bar c\|x_i-z_i\|^3=\bar c(\mathrm{dist}\,(x_i,X))^3,$$

где z_i — проекция точки x_i на X. $\square$

При этом последовательность $\{x_k\}$ сходится к некоторой точке $\bar x\in X$ таким образом, что справедлива следующая оценка скорости сходимости:

$$\|x_{i+1}-\bar x\|\le c\,\|x_i-\bar x\|^3.$$

3. Рассмотрим частный случай системы (1), когда $F : \mathbb{R}^n \to \mathbb{R}^n$. Предположим, что квадратная матрица $F_x(x)$ не вырождена. Тогда расчетные формулы упрощаются, в частности $F_x^+(x) = F_x^{-1}(x)$, $N(x) = F_x^{-1}(x)F(x)$, итерационный метод (6) принимает вид

$$x_{i+1} = x_i - F_x^{-1}(x_i)\left(F(x_i) + \frac{1}{2}\, N^\top(x_i)F_{xx}(x_i)N(x_i)\right). \tag{7}$$

Теорема 3. *Пусть $F \in C^2(\mathbb{R}^n)$, существует точка $x_* \in X$, где матрица F_x не вырождена. Тогда итерационный метод* (7) *локально сходится к x_* с кубической скоростью*

$$\|x_{i+1} - x_*\| \le c\,\|x_i - x_*\|^3, \tag{8}$$

где c — некоторая константа.

Действительно, из (7) следует

$$x_{i+1} - x_* = F_x^{-1}(x_i)\left\{F_x(x_i)(x_i - x_*) - F(x_i) - \frac{1}{2}\, N^\top(x_i)F_{xx}(x_i)N(x_i)\right\}. \tag{9}$$

Предположим, что точка x_i принадлежит достаточно малой окрестности x_*. Воспользуемся формулой Тейлора, получим

$$F(x_*) = F(x_i) + F_x(x_i)(x_* - x_i) + \frac{1}{2}\,(x_* - x_i)^\top F_{xx}(x_i)(x_* - x_i) + O(\|x_* - x_i\|^3) = 0_n,$$
$$N(x_i) = F_x^{-1}(x_i)F(x_i) = x_i - x_* + O(\|x_i - x_*\|^2).$$

Из этих формул находим

$$F(x_i)(x_i - x_*) - F(x_i) = \frac{1}{2}\, N(x_i)^\top F_{xx}(x_i)N(x_i) + O(\|x_i - x_*\|^3).$$

Подставляя это выражение в (9), приходим к оценке (8).

Если в (7) отбросить члены второго порядка малости (т.е. порядка $\|F(x_i)\|^2$), то (7) переходит в обычный метод Ньютона:

$$x_{i+1} - x_i = -F_x^{-1}(x_i)F(x_i) = -N(x_i). \tag{10}$$

В частном случае, когда $n = m = 1$, из (7) следует

$$x_{i+1} = x_i - \frac{f(x_i)}{f_x(x_i)} - \frac{f^2(x_i)f_{xx}(x_i)}{2f_x^3(x_i)}.$$

Этот метод отыскания корней уравнения $f(x) = 0$ был предложен в 1838 году П.Л. Чебышевым [1]. Итерационные методы, обладающие более высокой, чем квадратичная, скоростью сходимости, изучались в работах [2]–[5].

4. Перейдем к классической теореме о неявной функции для уравнения $F(x, y) = 0_m$, где $F^\top(x, y) = [f_1(x, y), \ldots, f_m(x, y)]$, $F : X \times Y \to \mathbb{R}^m$, X, Y — B-пространства. Вначале рассмотрим случай $Y = \mathbb{R}^n$, а затем покажем, как можно общий случай B-пространства Y свести к конечномерному $Y = \mathbb{R}^n$.

Введем новую матрицу Грамма $Q(x, y)$, матрицу $F_y^+(x, y)$, псевдообратную к $F_y(x, y)$, n-мерные вектор-функции $M(x, y)$ и $y(x)$:

$$Q(x, y) = F_y(x, y)F_y^\top(x, y), \qquad F_y^+(x, y) = F_y^\top(x, y)\,(Q(x, y))^{-1},$$
$$M(x, y) = F_y^+(x_*, y)F(x, y), \qquad y(x, \eta) = y_* - M(x, y_*) + F_y^\top(x_*, y_*)\eta.$$

Через $U_\varepsilon(x_*)$ и $U_\delta(y_*)$ обозначим окрестности точек x_* и y_* с диаметрами ε и δ соответственно.

Теорема 4 (о неявной функции). *Пусть $Y = \mathbb{R}^n$, $F(x,y) \in C^2(X\times Y)$, $F(x_*,y_*) = 0_m$, ранг матрицы $F_y(x_*,y_*)$ равен m. Тогда для любого $x \in U_\varepsilon(x_*)$ существует дифференцируемая функция $y(x) \in U_\delta(y_*)$, такая что*

$$y(x_*) = y_*, \quad F(x, y(x)) = 0_m, \quad \|y(x) - y_*\| \le c\|F(x,y_*)\|, \quad y_x(x_*) = -F_y^+(x_*,y_*)F_x(x_*,y_*), \tag{11}$$

где $c > 0$ — независимая константа, а числа $\varepsilon > 0$ и $\delta > 0$ достаточно малые.

Доказательство проводится аналогично доказательству предыдущей теоремы.

Для любых $\eta \in \mathbb{R}^m$, $\eta \ne 0$ выполнено неравенство $\inf \|Q(x_*,y_*)\eta\|/\|\eta\| \ge \gamma > 0$. Считаем $\|F(x,y_*)\|$ и $\|\eta\|$ величинами первого порядка малости. Для $\varphi(\eta) = \|F(x, y(x,\eta))\|^2$ имеем: $\varphi(\eta) \ge \dfrac{1}{2}\gamma^2\|\eta\|^2$. Число $\varphi(0)$ есть величина четвертого порядка малости. Минимум функции φ по η достигается в некоторой внутренней точке $\bar\eta(x)$, причем аналогом оценки $\|\bar\eta(\alpha)\| < \rho$ в теореме 1 в данном случае будет

$$\|\bar\eta(x)\| < c\,\|F(x,y_*)\|\,\|x - x_*\|.$$

Отсюда имеем

$$\varphi_\eta(\bar\eta(x)) = 0_m, \quad F_y(x_*, y(x))F_y^\top(x_*,y_*)F(x_*, y_* - M(x,y_*) + F_y^\top(x_*,y_*)\bar\eta(x)) = 0_m. \tag{12}$$

Если $\tilde y$ достаточно близко к y_*, то из невырожденности $Q(x_*,y_*)$ следует невырожденность матрицы $Q(x_*,\tilde y)$, поэтому из (12) получим, что

$$F(x_*, y_* - M(x,y_*) + F_y^\top(x_*,y_*)\bar\eta(x)) = 0_m.$$

Обозначая $y(x) = y(x, \bar\eta(x))$, завершаем доказательство. $\square$

Неявная функция $y(x)$, удовлетворяющая условию $F(x,y(x)) = 0_m$, может быть определена приближенно по формуле, аналогичной (6):

$$\begin{aligned} \bar y(x) \;&=\; y_* - M(x,y_*) + F_y^+(x_*,y_*)\Big\{F_{xy}(x_*,y_*)M(x,y_*)(x - x_*) - \\ &-\; \frac{1}{2}M^\top(x,y_*)F_{yy}(x_*,y_*)M(x,y_*)\Big\}, \end{aligned} \tag{13}$$

где $F_{xy}(x_*,y_*)M(x,y_*)(x-x_*)$, $M^\top(x,y_*)F_{yy}(x_*,y_*)M(x,y_*)$ — векторы из $\mathbb{R}^m$, i-е компоненты которых суть соответственно $f^i_{xy}(x_*,y_*)M(x,y_*)(x-x_*)$ и $M^\top(x,y_*)f^i_{yy}(x_*,y_*)M(x,y_*)$. Отсюда заключаем, что $\|\bar y(x) - y(x)\| \le c\,\|x - x_*\|^3$, где $c > 0$ — независимая константа.

5. Покажем, как данные результаты переносятся на случай банаховых пространств. Для этого понадобится следующая лемма, доказательство которой можно найти в [5].

Лемма 2. *Пусть X — банахово пространство и X^* — его сопряженное. Тогда для любой линейно независимой системы элементов $\{\xi_i\}$, где $\xi_i \in X^*$, $i = \overline{1,m}$, существует m линейно независимых элементов $\{\eta_i\}$, $\eta_i \in X$, $i = \overline{1,m}$, таких, что матрица Грамма*

$$\begin{pmatrix} \langle\xi_1,\eta_1\rangle & \dots & \langle\xi_1,\eta_m\rangle \\ \dots & \dots & \dots \\ \langle\xi_m,\eta_1\rangle & \dots & \langle\xi_m,\eta_m\rangle \end{pmatrix}$$

не вырождена.

Приведенные доказательства теорем 1 и 3 не изменятся, если взять $X = \mathbb{R}^m$ в теореме 1 и $Y = \mathbb{R}^m$ — в теореме 3. При этом роль $\{\xi_i\}$ будут играть соответственно $f_x^i(x_*)$ либо $f_y^i(x_*, y_*)$, $i = \overline{1, m}$.

6. Классический принцип Лагранжа о необходимых условиях минимума в задаче нелинейного программирования

$$\min_{x \in X} \varphi(x), \qquad X = \{x \in \mathbb{R}^n : g(x) = 0_m\}, \qquad \varphi(x) : \mathbb{R}^n \to \mathbb{R}^1. \tag{14}$$

просто следует из теоремы 1. Действительно, пусть x_* — решение задачи (14) и выполнено условие регулярности ограничений; тогда

$$\nabla\varphi(x_*) \in \operatorname{Im}\{g_x^\top(x_*)\} = \mathcal{L}; \tag{15}$$

здесь вектор-столбец $\nabla\varphi(x)$ есть градиент функции $\varphi(x)$ в точке x.

Предположим противное, тогда $\varphi_x^\top(x_*) = h + \bar{h}$, где $h \in \operatorname{Ker} g_x^\top(x_*) = \mathcal{L}^\perp$, $\bar{h} \in \mathcal{L}$. Согласно теореме 1 точка $x_* + \alpha h + \alpha^2 \bar{h} + O(\alpha^3)$ принадлежит X. Ненулевой вектор h из ядра выберем так, чтобы выполнялось строгое неравенство $\langle \nabla\varphi(x_*), h\rangle < 0$. Тогда

$$\varphi(x_* + \alpha h + \alpha^2\bar{h} + O(\alpha^3)) = \varphi(x_*) + \alpha\langle\nabla\varphi(x_*), h\rangle + O(\alpha^2) < \varphi(x_*).$$

Полученное противоречие доказывает утверждение (15).

7. Используем формулы (9), (13) для получения приближенного решения задач (14). Составим для нее функцию Лагранжа и внешнюю штрафную функцию:

$$L(x, \lambda) = \varphi(x) + \lambda^\top g(x), \qquad P(x, \varepsilon) = \varepsilon\varphi(x) + \frac{1}{2}\|g(x)\|^2, \qquad \lambda \in \mathbb{R}^m, \quad \varepsilon \in \mathbb{R}. \tag{16}$$

Необходимые условия минимума функций L и P по x имеют вид

$$\begin{array}{rcl} \nabla L(x, \lambda) & = & \nabla\varphi(x) + g_x^\top(x)\lambda = 0_n, \\ \nabla P(x, \varepsilon) & = & \varepsilon\nabla\varphi(x) + g_x^\top(x) g(x) = 0_n. \end{array}$$

Эти системы совпадают, если $g(x) = \varepsilon\lambda$.

Предполагаем, что (14) и двойственная к ней задача имеют решения x_* и λ_*. Введем две $(n+m)$-мерные вектор-функции

$$F(x, \lambda) = \begin{pmatrix} \nabla\varphi(x) + g_x^\top(x)\lambda \\ g(x) \end{pmatrix}, \qquad \Phi(\varepsilon, x, \lambda) = \begin{pmatrix} \nabla\varphi(x) + g_x^\top(x)\lambda \\ g(x) - \varepsilon\lambda \end{pmatrix}.$$

Пусть x_* и λ_* — невырожденные решения следующей системы из $(n+m)$ уравнений с $(n+m)$ неизвестными:

$$F(x, \lambda) = 0_{n+m}. \tag{17}$$

Дифференцируя F по x и λ, определим квадратную матрицу Якоби размерности $(n+m) \times (n+m)$:

$$F''(x, \lambda) = \begin{bmatrix} L_{xx}(x, \lambda) & g_x^\top(x) \\ g_x(x) & 0_{mm} \end{bmatrix}.$$

Требуемое в теореме 3 условие невырожденности матрицы $F''(x_*, \lambda_*)$ будет выполнено, если ранг матрицы $g_x(x_*)$ равен m и матрица $L_{xx}(x_*, \lambda_*)$ положительно определена на ядре матрицы $g_x(x_*)$.

Первый метод решения задачи (14) состоит в решении системы (17). Для этого воспользуемся формулой (7), считая, что в ней вектор x состоит из двух компонент x и λ, участвующих в формулах (16).

Второй метод решения задачи (14) состоит в параметрическом представлении $x(\varepsilon)$ и $\lambda(\varepsilon)$ решений системы из $(n+m)$ уравнений с таким же числом неизвестных

$$\Phi(\varepsilon, x(\varepsilon), \lambda(\varepsilon)) = 0_{n+m}. \tag{18}$$

Здесь решения x_* и λ_* получаются асимптотически в пределе при $\varepsilon \to 0$.

В работах по методу штрафных функций показано, что $\|x(\varepsilon) - x_*\| \le c\,\varepsilon$, если ε достаточно мало. Возьмем в формуле (13) в качестве x параметр ε, а в качестве функции $y(x)$ — пару $[x(\varepsilon), \lambda(\varepsilon)]$. Полагаем $x_* = \varepsilon$, $y_*^\top = [x_*, \lambda_*] = [x(\varepsilon), \lambda(\varepsilon)]$. Тогда роль псевдообратной матрицы F_y^+ будет играть матрица $\left(\Phi'_{(x,\lambda)}\right)^{-1}$, поэтому

$$M(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) = \left\{\Phi'_{(x,\lambda)}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon)))\right\}^{-1} \Phi(0, (x(\varepsilon), \lambda(\varepsilon))).$$

Из формулы (13) следует

$$\begin{aligned}
(\bar{x}(0), \bar{\lambda}(0)) \;&=\; (x(\varepsilon), \lambda(\varepsilon)) - M(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) - \left\{\Phi'_{(x,\lambda)}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon)))\right\}^{-1} \times \\
\times\; &\left[\varepsilon \frac{\partial}{\partial \varepsilon}\, \Phi'_{(x,\lambda)}(\varepsilon, (x(\varepsilon)\lambda(\varepsilon))) M(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) + \right. \\
+\; &\left. \frac{1}{2}\, M^\top(\varepsilon, (x(\varepsilon), \lambda(\varepsilon)))\Phi''_{(x,\lambda)}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) M(\varepsilon(x(\varepsilon), \lambda(\varepsilon)))\right],
\end{aligned} \tag{19}$$

так что $\|\bar{x}(0) - x_*\| \le c\,\varepsilon^3$. Таким образом, имея точку $x(\varepsilon)$, мы, используя формулу (19), сразу получаем точку $\bar{x}(0)$, которая аппроксимирует решение x_* с точностью ε^2.

8. **Пример.** Рассмотрим простейший вариант задачи (14). Пусть $n = m = 1$, $\varphi(x) = x$, $g(x) = x + x^2$. Тогда $x_* = 0$, $\lambda_* = -1$, $L(x, \lambda) = x + \lambda(x + x^2)$, $P(x, \varepsilon) = \varepsilon x + \dfrac{1}{2}\,(x + x^2)^2$. Задача сводится к решению системы (17), в которой $n = m = 2$, $F_x^+ = F_x^{-1}$, а в качестве вектора x берется пара x и λ, определяемая в (16). Поэтому

$$F = \begin{pmatrix} 1 + \lambda + 2\lambda x \\ x + x^2 \end{pmatrix}, \quad F_{x\lambda} = \begin{pmatrix} 2\lambda & 1 + 2x \\ 1 + 2x & 0 \end{pmatrix}, \quad F_{x\lambda}^{-1} = \frac{1}{(1+2x)^2} \begin{pmatrix} 0 & 1 + 2x \\ 1 + 2x & -2\lambda \end{pmatrix}.$$

Считаем $x_i - x_*$, $\lambda_i - \lambda_*$ величинами первого порядка малости (порядка δ). Тогда можно записать $x_i - x_* = x_i \simeq O(\delta)$, $\lambda_i - \lambda_* = \lambda_i + 1 \simeq O(\delta)$.

Воспользуемся стандартным методом Ньютона (10), получим

$$x_{i+1} = (x_i)^2 \simeq O(\delta^2), \qquad \lambda_{i+1} + 1 = 2x_i(1 + \lambda_i) \simeq O(\delta^2). \tag{20}$$

Более точный результат следует из метода (9):

$$x_{i+1} \simeq O(\delta^3), \qquad \lambda_{i+1} + 1 \simeq O(\delta^3).$$

Согласно методу Ньютона (10) x_{i+1} и λ_{i+1} определены с погрешностями соответственно $(x_i)^2$ и $2x_i(1+\lambda_i)$, имеющими второй порядок малости. В случае использования метода (9), напротив, получены точные решения $x_{i+1} = 0$, $\lambda_{i+1} = -1$ с гарантией, что погрешность, если она есть, имеет третий порядок малости.

К приведенному примеру применим параметризацию. В этом случае

$$\Phi(\varepsilon, x, \lambda) = \begin{pmatrix} 1+\lambda+2x\lambda \\ x+x^2-\varepsilon\lambda \end{pmatrix}, \qquad \Phi'_{(x,\lambda)}(\varepsilon, x, \lambda) = \begin{pmatrix} 2\lambda & 1+2x \\ 1+2x & -\varepsilon \end{pmatrix},$$

$$\left[\Phi'_{(x,\lambda)}(\varepsilon, x, \lambda)\right]^{-1} = \frac{1}{2\varepsilon\lambda+(1+2x)^2}\begin{pmatrix} \varepsilon & 1+2x \\ 1+2x & -2\lambda \end{pmatrix}.$$

С точностью до величин второго порядка малости имеем:

$$x(\varepsilon) = -\varepsilon - 3\varepsilon^2, \qquad \lambda(\varepsilon) = -1-2\varepsilon-10\varepsilon^2.$$

Поэтому

$$\Phi^{-1}_{(x,\lambda)}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon)))\Phi(0, (x(\varepsilon), \lambda(\varepsilon))) = \frac{1}{6\varepsilon-1}\begin{pmatrix} \varepsilon \\ 2\varepsilon+8\varepsilon^2 \end{pmatrix},$$

$$\Phi''_{(x,\lambda)}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) = \begin{pmatrix} \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \\ \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix}, \qquad \frac{\partial}{\partial\varepsilon}\Phi'_{(x,\lambda),\varepsilon}(\varepsilon, (x(\varepsilon), \lambda(\varepsilon))) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Отсюда следует, что разности $|\bar{x}(0) - x_*|$ и $|\bar{\lambda}(0) - \lambda_*|$ имеют третий порядок малости.

9. Приведенные в работе быстросходящиеся методы могут быть использованы для решения задач ЛП с ограничениям типа равенства. Однако если воспользоваться специальным вариантом метода проекции [6, 7], то исходная задача линейного программирования редуцируется к задачам безусловной минимизации дифференцируемых нелинейных функций, что можно осуществить с помощью изложенных подходов.

Работа выполнена при финансовой поддержке РФФИ (грант 11-01-00786-а) и Программы поддержки ведущих научных школ НШ-4096.2010.1.

Литература

1. *Чебышев П.Л.* Собрание сочинений. Т. V. М.-Л.: Изд-во АН СССР, 1951. С. 7–25.
2. *Третьяков А.А.* Нелинейные методы решения экстремальных задач. Дисс. на соиск. уч. степ. канд. физ.-мат. наук. Москва, 1983.
3. *Третьяков А.А.* Две схемы нелинейного метода оптимизации в экстремальных задачах // Журн. вычисл. мат. и матем. физ. 1984. Т. 24. № 7. С. 986–992.
4. *Денисов А.А., Карманов В.Г., Третьяков А.А.* Ускоренный метод Ньютона для решения функциональных уравнений // ДАН СССР, 1985. Т. 281. № 6. С. 1293–1297.
5. *Третьяков А.А.* Элементарное доказательство классических необходимых условий оптимальности для задачи математического программирования с ограничениями-неравенствами // ДАН. 2009. Т. 425. № 4. С. 459–461.
6. *Голиков А.И., Евтушенко Ю.Г.* Нахождение проекции заданной точки на множество решений задач линейного программирования / Тр. ин-та матем. и мех. УрО РАН. 2008. Т. 14. № 2. С. 33-47.
7. *Гаранжа В.А., Голиков А.И., Евтушенко Ю.Г., Нгуен М.Х.* Параллельная реализация метода Ньютона для решения больших задач линейного программирования //Журн. вычисл. мат. и матем. физ. 2009. Т. 49. № 8. С. 1369-1384.
8. *Блисс Г.А.* Лекции по вариационному вычислению. М.: Изд-во иностранной литературы, 1950.