УДК 519.85

Обоснование алгоритмов внутренних точек для задач оптимизации с нелинейными ограничениями*

В.И. Зоркальцев, С.М. Пержабинский

Институт систем энергетики им. Л.А. Мелентьева Сибирского отделения Российской академии наук, ул. Лермонтова, 130, Иркутск, 664033

E-mails: zork@isem.sei.irk.ru (Зоркальцев В.И.), smper@isem.sei.irk.ru (Пержабинский С.М.)

Зоркальцев В.И., Пержабинский С.М. Обоснование алгоритмов внутренних точек для задач оптимизации с нелинейными ограничениями // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2013. — Т. 16, № 1. — С. 27–38.

Рассматривается семейство алгоритмов внутренних точек. Алгоритмы предназначены для решения задач математического программирования с нелинейными ограничениями-неравенствами. При поиске направления улучшения решения используются изменяющиеся по итерациям взвешенные евклидовы нормы. Представлены результаты теоретического обоснования алгоритмов при некоторых предположениях (в том числе о невырожденности задачи).

Ключевые слова: *метод внутренних точек, взвешенная евклидова норма, линеаризация.*

Zorkaltsev V.I., Perzhabinsky S.M. Theoretical justification of interior point algorithms for solving optimization problems with nonlinear constraints // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2013. — Vol. 16, № 1. — P. 27–38.

A family of interior point algorithms is considered. These algorithms can be used for solving mathematical programming problems with nonlinear inequality constraints. The weighted Euclidean rates are applied to find a descent direction for improving a solution. These rates are varying in iterations. Theoretical justification of the algorithms with some assumptions (such as non-degeneracy of a problem) is presented.

Key words: *interior point method, weighted Euclidean rate, linearization.*

Введение

Известно, что алгоритмы внутренних точек являются высокоэффективными процедурами решения задач математического программирования. Из множества алгоритмов внутренних точек в особый класс выделяются алгоритмы, в которых поиск направления улучшения решения основывается на идее стимулирования движения вдоль границ области допустимых по ограничениям-неравенствам решений. Это обуславливает оригинальность и эффективность этих методов, но в то же время затрудняет их теоретическое обоснование. Пионерные разработки таких алгоритмов были осуществлены в СССР в 60–70-х гг. прошлого века С.М. Анцызом, И.И. Дикиным [1, 2], Ю.Г. Евтушенко и В.Г. Жаданом [3, 4], В.И. Зоркальцевым [2, 5, 6]. Алгоритмы внутренних точек обсуждаемого типа успешно используются с 70-х гг. прошлого века при реализации ряда моделей энергетики в Институте систем энергетики им. Л.А. Мелентьева СО РАН. При этом для моделей с нелинейными ограничениями применяются процедуры итеративной линеаризации.

*Работа выполнена при поддержке РФФИ (проект № 09-01-00306а).

1. Постановка задачи

Рассмотрим задачу нелинейного программирования

$$f_0(x) \to \min, \quad x \in X, \tag{1}$$

где

$$X = \{x \in R^n : \underline{x} \le x \le \overline{x}, \quad f_i(x) \le 0, \ i = 1, \ldots, m\}. \tag{2}$$

Заданы векторы $\underline{x}$, $\overline{x}$ из R^n (причем $\underline{x}_j < \overline{x}_j$, $j = 1, \ldots, n$) и сепарабельные выпуклые дважды дифференцируемые функции $f_i(x)$, $i = 0, \ldots, m$.

Векторы из X будем называть допустимыми решениями. Подмножество векторов из X, удовлетворяющих ограничениям-неравенствам в строгой форме, обозначим

$$\mathrm{int}X = \left\{ x \in R^n : \ f_i(x) < 0, \ i = 1, \ldots, m, \quad \underline{x} < x < \overline{x} \right\}.$$

Будем считать, что $\mathrm{int}X \neq \emptyset$ и задан вектор x^0 из $\mathrm{int}X$. Следовательно, задача (1), (2) удовлетворяет условию Слейтера. Согласно условиям оптимальности Куна–Таккера, для того, чтобы вектор $x^* \in X$ был оптимальным решением задачи (1), (2), необходимо и достаточно существования векторов $u^* \in R^m$, $\alpha^* \in R^n$, $\beta^* \in R^n$, при которых

$$\nabla f_0(x^*) + \sum_{i=1}^{m} u_i^* \nabla f_i(x^*) + \alpha^* - \beta^* = 0, \tag{3}$$

$$u_i^* f_i(x^*) = 0, \qquad i = 1, \ldots, m, \tag{4}$$

$$\beta_j^*(x_j^* - \underline{x}_j) = 0, \quad j = 1, \ldots, n, \tag{5}$$

$$\alpha_j^*(\overline{x}_j - x_j^*) = 0, \quad j = 1, \ldots, n, \tag{6}$$

$$u^* \ge 0, \quad \alpha^* \ge 0, \quad \beta^* \ge 0.$$

Допустимое решение $x \in X$ назовем стационарным, если для него выполняются соотношения (3)–(6) при некоторых векторах u, α, β. Если такие векторы единственные, то решение x будем называть невырожденным. Задачу (1), (2) назовем невырожденной, если у нее любое стационарное решение не вырождено.

Заметим, что оптимальное решение является всегда стационарным, но не любое стационарное решение является оптимальным.

2. Описание семейства алгоритмов внутренних точек

В рассматриваемом вычислительном процессе последовательность векторов x^k из $\mathrm{int}X$ вырабатывается по правилу

$$x^{k+1} = x^k + \lambda_k \Delta x^k. \tag{7}$$

Здесь $k = 0, 1, 2, \ldots$ — номер итерации, Δx^k — вектор спуска, λ_k — шаг корректировки.

Основной проблемой в излагаемом вычислительном процессе является определение направления улучшения решения. Для ее разрешения сначала в точке x^k вычисляется градиент целевой функции

$$c^k = \nabla f_0(x^k).$$

Если $c^k = 0$, то дальнейшие вычисления прекращаются, так как x^k — точка минимума функции $f_0(x)$. Далее будем считать, что $c^k \neq 0$.

После этого находится матрица A_k размера $m \times n$, строками которой являются градиенты функций $f_i(x)$, $i = 1, \ldots, m$. Затем вычисляется диагональная матрица B_k размера $n \times n$ с положительными элементами на главной диагонали.

В работе [8] в качестве B_k использовалась взвешенная сумма матриц вторых производных функций $f_i(x)$, $i = 0, \ldots, m$:

$$B_k = \nabla^2 f_0(x^k) + \sum_{i=1}^{m} v_i^{k-1} \nabla^2 f_i(x^k). \tag{8}$$

Здесь v_i^{k-1} — приближения на итерации $k-1$ к множителям Лагранжа нелинейных ограничений задачи (1), (2). На начальной итерации $v_i^0 = 1$, $i = 1, \ldots, m$. Экспериментальные исследования [8] такого алгоритма внутренних точек показали его вычислительную эффективность.

Поскольку функции $f_i(x)$, $i = 0, \ldots, m$, сепарабельные, то матрица B_k является диагональной. В случае несепарабельных функций $f_i(x)$ возможна модификация правила (8), в которой для всех i, $i = 0, \ldots, m$, вместо матрицы $\nabla^2 f_i(x^k)$ используется диагональная матрица того же размера с элементами $\partial^2 f_i(x^k)/\partial x_j^2$, $j = 1, \ldots, n$, размещенными на главной диагонали.

Затем определяем векторы весовых коэффициентов $d^k \in R^n$, $h^k \in R^m$. Их компоненты должны удовлетворять неравенствам:

$$\underline{\sigma}\left(y_j^k\right) \le d_j^k \le \overline{\sigma}\left(y_j^k\right), \qquad j = 1, \ldots, n, \tag{9}$$

$$\underline{\sigma}\left(-f_i^k\right) \le h_i^k \le \overline{\sigma}\left(-f_i^k\right), \quad i = 1, \ldots, m. \tag{10}$$

Здесь

$$y_j^k = \min\left\{\overline{x}_j - x_j^k,\ x_j^k - \underline{x}_j\right\}, \quad j = 1, \ldots, n,$$

$$f_i^k = f_i\left(x^k\right), \qquad i = 1, \ldots, m,$$

$\overline{\sigma}$, $\underline{\sigma}$ — некоторые непрерывные неубывающие функции положительного аргумента, удовлетворяющие двум условиям:

$$0 < \underline{\sigma}(t) \le \overline{\sigma}(t), \quad \forall t > 0, \tag{11}$$

$$\overline{\sigma}(t) \le Mt \tag{12}$$

при некоторых $\varepsilon > 0$, $M > 0$ для всех $t \in (0,\ \varepsilon]$.

Исследуемый вычислительный процесс представляет семейство алгоритмов, поскольку могут использоваться различные правила вычисления весовых коэффициентов. В частности, можно пользоваться правилом

$$d_j^k = \left(y_j^k\right)^2, \quad h_i^k = \left(f_i^k\right)^2, \quad j = 1, \ldots, n,\ i = 1, \ldots, m.$$

Другими эффективными, как показали результаты некоторых экспериментальных и теоретических исследований [6], вариантами правил нахождения весовых коэффициентов могут служить следующие:

$$d_j^k = \frac{y_j^k}{\min\{\varepsilon, \alpha_j^{k-1}, \beta_j^{k-1}\}}, \quad h_i^k = \frac{-f_i^k}{\min\{\varepsilon, u_i^{k-1}\}}, \quad j = 1, \ldots, n, \; i = 1, \ldots, m,$$

при заданном $\varepsilon > 0$. Здесь u_i^{k-1}, α_j^{k-1}, β_j^{k-1} — приближения на итерации $k-1$ к множителям Лагранжа ограничений, $f_i(x) \le 0, \; i = 1, \ldots, m, \; x \le \overline{x}, \; \underline{x} \le x$.

Обозначим

$$D_k = \operatorname{diag}(d^k), \quad H_k = \operatorname{diag}(h^k)$$

диагональные матрицы размера $n \times n$ и $m \times m$ с компонентами векторов d^k и h^k на главной диагонали.

Интервальный и аксиоматический подходы к определению весовых коэффициентов были введены в [5] применительно к алгоритмам внутренних точек для решения задач линейного программирования. Такой подход обобщает имеющиеся способы нахождения весовых коэффициентов и позволяет создавать новые. При этом не обязательно иметь конкретные выражения для функций $\overline{\sigma}$, $\underline{\sigma}$. Достаточно доказать существование таких функций и указанных их свойств.

2.1. Вспомогательная задача поиска направления улучшения решения

Вектор Δx^k определим как результат решения вспомогательной задачи минимизации квадратичной выпуклой функции от векторов $\Delta x \in R^n$ и $z \in R^m$:

$$(c^k)^\top \Delta x + \frac{1}{2} \Delta x^\top B_k \Delta x + \frac{1}{2} \Delta x^\top D_k^{-1} \Delta x + \frac{1}{2} z^\top H_k^{-1} z \to \min \tag{13}$$

при линейных ограничениях-равенствах

$$A_k \Delta x - z = 0. \tag{14}$$

В целевой функции задачи (13), (14) составляющие $\Delta x^\top D_k^{-1} \Delta x$, $z^\top H_k^{-1} z$ стимулируют при выборе Δx^k к движению "вдоль" границ множества векторов, удовлетворяющих неравенствам $\underline{x} \le x \le \overline{x}$, $f_i(x) \le 0$, $i = 1, \ldots, m$. Составляющая $\Delta x^\top B_k \Delta x$ является квадратом взвешенной евклидовой нормы. Если матрица B_k вычисляется по правилу (8), то это слагаемое представляет вторые производные в аппроксимации функций $f_i(x)$, $i = 0, \ldots, m$.

Обозначим вектор множителей Лагранжа ограничений (14) как u^k. Определим вектор приближений к множителям Лагранжа нелинейных ограничений исходной задачи

$$v_i^k = \max\{0, \; u_i^k\}, \quad i = 1, \ldots, m. \tag{15}$$

Множество номеров i, при которых $u_i^k < 0$, обозначим I_k.

2.2. Два варианта вычисления направления корректировки

Вспомогательная задача (13), (14) сводится к решению системы линейных уравнений с симметрической положительно определенной матрицей. Возможны два варианта вычислений векторов Δx^k, z^k, u^k.

Один из них основывается на решении системы линейных уравнений с матрицей размера $n \times n$:

$$\left(B_k + D_k^{-1} + A_k^\top H_k^{-1} A_k\right) \Delta x = -c^k. \tag{16}$$

Найдя в результате решения системы (16) вектор Δx^k, определяем векторы

$$z^k = A_k \Delta x^k, \qquad u^k = H_k^{-1} z^k. \tag{17}$$

Система (16) следует из задачи безусловной минимизации

$$\big(c^k\big)^\top \Delta x + \frac{1}{2}\Delta x^\top B_k \Delta x + \frac{1}{2}\Delta x^\top D_k^{-1}\Delta x + \frac{1}{2}\Delta x^\top A_k^\top H_k^{-1} A_k \Delta x \to \min,$$

к которой приходим в результате подстановки вектора z (из условия (14)) в целевую функцию (13).

Второй вариант вычислений основан на решении системы размера $m \times m$:

$$\Big(A_k\big(B_k + D_k^{-1}\big)^{-1} A_k^\top + H_k\Big) u = -A_k\big(B_k + D_k^{-1}\big)^{-1} c^k. \tag{18}$$

Вычислив u^k, находим вектор

$$\Delta x^k = -\big(B_k + D_k^{-1}\big)^{-1}\big(A_k^\top u^k + c^k\big) \tag{19}$$

и по формуле (14) — вектор z^k. Система (18) — результат решения задачи (13), (14) методом неопределенных множителей Лагранжа. Второй из двух вариантов вычислений предпочтительней в том случае, когда $m < n$.

Векторы $\alpha^k \in R^n$, $\beta^k \in R^n$, являющиеся приближениями на итерации k к множителям Лагранжа ограничений $x \le \overline{x}$, $\underline{x} \le x$, определим следующим образом:

$$\alpha_j^k = \max\big\{\, 0,\, -q_j^k \big\}, \quad \beta_j^k = \max\big\{\, 0,\ q_j^k \big\}, \quad j = 1, \ldots, n. \tag{20}$$

Вектор q^k вычисляется по правилу

$$q^k = -\big(B_k + D_k^{-1}\big)\Delta x^k,$$

что, согласно (19), равносильно следующей формуле:

$$q^k = A_k^\top u^k + c^k.$$

2.3. Нахождение шага корректировки

Шаг корректировки вычисляется следующим образом:

1) $\lambda_1^k = \gamma \max\{\lambda : x^k + \lambda \Delta x^k \in X,\ \lambda \ge 0\}$, γ — заданный параметр из интервала $\big(0, 1\big)$;
2) если $I_k \ne \emptyset$, то $\lambda_2^k = \max\big\{\lambda :\ f_i\big(x^k + \lambda\Delta x^k\big) \le f_i\big(x^k\big),\ i \in I_k,\ 0 \le \lambda \le \lambda_1^k\big\}$;
3) если $I_k = \emptyset$, то $\lambda_2^k = \lambda_1^k$;
4) $\lambda_k = \operatorname{argmin}\big\{f_0\big(x^k + \lambda \Delta x^k\big) :\ 0 \le \lambda \le \lambda_2^k\big\}$.

Правила нахождения шага λ_k гарантируют, что вектор x^{k+1} будет находиться в $\operatorname{int} X$.

2.4. Критерий останова

Вычисления заканчиваются при выполнении с заданной точностью $\varepsilon > 0$ для $x^k \in X$, $v^k \ge 0$, $\alpha^k \ge 0$, $\beta^k \ge 0$, ограничения на множители Лагранжа

$$\big\| c^k + A^\top v^k + \alpha^k - \beta^k \big\| \le \varepsilon$$

и условий дополняющей нежесткости (4)–(6):

$$\begin{aligned} -v_i^k f_i\bigl(x^k\bigr) &\le \varepsilon, \quad i = 1,\ldots,m,\\ \beta_j^k\bigl(x_j^k - \underline{x}_j\bigr) &\le \varepsilon, \quad j = 1,\ldots,n,\\ \alpha_j^k\bigl(\bar{x}_j - x_j^k\bigr) &\le \varepsilon, \quad j = 1,\ldots,n. \end{aligned}$$

Здесь $\|r\| = \bigl(\sum_{i=1}^n r_i\bigr)^{1/2}$ — евклидова (невзвешенная) норма вектора $r \in R^n$, векторы v^k, α^k, β^k находятся по правилам (15), (20).

2.5. Свойства вырабатываемых приближений

В силу (16) и положительной определенности матрицы $B_k + D_k^{-1} + A_k^\top H_k^{-1} A_k$ для всех $k = 0, 1, 2, \ldots$ имеет место неравенство

$$\bigl(c^k\bigr)^\top \Delta x^k < 0. \tag{21}$$

Из выполнения неравенства (21) и правил определения шага λ_k следует, что

$$f_0\bigl(x^{k+1}\bigr) < f_0\bigl(x^k\bigr). \tag{22}$$

Поскольку целевая функция задачи (1), (2) ограничена снизу на множестве X, то в силу (22) значения целевой функции сходятся к некоторой величине

$$F = \lim_{k\to\infty} f_0\bigl(x^k\bigr). \tag{23}$$

3. Евклидовы проекции начала координат на линейное многообразие

В применяемой здесь технике доказательства используется установленный в [5] факт ограниченности множества евклидовых проекций начала координат на линейное многообразие. Сначала поясним этот факт.

Пусть L — линейное многообразие в R^n. Вектор $s \in L$ назовем особым вектором многообразия L, если не существует вектора $y \in L$ такого, что $J\bigl(y\bigr) \subset J\bigl(s\bigr)$, $J\bigl(y\bigr) \ne J\bigl(s\bigr)$, где $J\bigl(y\bigr) = \{j : \ y_j \ne 0\}$ — множество номеров ненулевых компонент вектора y. Обозначим выпуклую оболочку особых векторов линейного многообразия L как $Q\bigl(L\bigr)$. Поскольку число особых векторов у линейного многообразия L конечно, то их выпуклая оболочка $Q\bigl(L\bigr)$ будет ограниченным множеством.

Обозначим

$$r\bigl(p\bigr) = \operatorname{argmin}\left\{\sum_{j=1}^n p_j\bigl(r_j\bigr)^2 : \ r \in L\right\}$$

евклидову проекцию начала координат на линейное многообразие L при использовании евклидовой нормы

$$\|\, r\,\|_p = \left(\sum_{j=1}^n p_j\bigl(r_j\bigr)^2\right)^{1/2}$$

с заданными весовыми коэффициентами $p_j > 0$, $j = 1,\ldots,n$. В [5] доказано следующее утверждение.

Лемма 1. *Любая евклидова проекция начала координат R^n на линейное многообразие $L \subset R^n$ находится в выпуклой оболочке особых векторов этого многообразия, $r(p) \in Q(L)$ при любых $p_j > 0$, $j = 1, \ldots, n$.*

Этот факт служит основой теоретического обоснования алгоритмов внутренних точек для решения задач линейного программирования [5, 6]. Усложняющим обстоятельством для нелинейных задач является изменение линейных многообразий, рассматриваемых при доказательстве, в зависимости от итеративно меняющейся точки, в которой осуществляется линеаризация. При этом меняются также особые векторы.

Выделим два линейных многообразия, соответствующих вектору x такому, что $\underline{x} \le x \le \bar{x}$. Линейное многообразие $L_1(x)$ состоит из векторов $\Delta x \in R^n$, $z \in R^m$, удовлетворяющих условиям

$$A(x)\Delta x - z = 0, \tag{24}$$

$$c^\top \Delta x = -1. \tag{25}$$

Здесь $A(x)$ — матрица размера $m \times n$, строками которой являются градиенты функций $f_i(x)$, $i = 1, \ldots, m$.

Линейное многообразие $L_2(x)$ состоит из векторов $q \in R^n$, $u \in R^m$, удовлетворяющих условию

$$q - \big(A(x)\big)^\top u = c. \tag{26}$$

4. Теоретическое обоснование алгоритмов внутренних точек

Для доказательства сходимости последовательности приближений, вырабатываемых вычислительным процессом (7), к оптимальному решению задачи (1), (2) введем некоторые предположения.

1. Будем считать, что целевая функция линейна: при заданном $c \in R^n$, $c \neq 0$,

$$f_0(x) = c^\top x.$$

Отметим, что это предположение вводится для удобства доказательства и не является существенным.

2. С изменением вектора x могут меняться значения особых векторов линейных многообразий $L_1(x)$, $L_2(x)$. Предположим, что объединения по x, удовлетворяющим условию $\underline{x} \le x \le \bar{x}$, множеств особых векторов многообразий $L_1(x)$ и объединения по x, удовлетворяющим $\underline{x} \le x \le \bar{x}$, множеств особых векторов многообразий $L_2(x)$ будут ограниченными множествами.

3. Будем считать, что величина шага λ_k на всех итерациях ограничена снизу некоторым положительным числом τ, т. е. существует $\tau > 0$ такое, что

$$\lambda_k \ge \tau, \quad k = 0, 1, 2, \ldots\ . \tag{27}$$

В случае линейных ограничений это условие обязательно выполняется, поэтому нет необходимости вводить его особо. При наличии нелинейных ограничений выполнение этого условия — открытый вопрос. Поэтому оно, к сожалению авторов, вводится в условия теоремы.

Для доказательства сходимости векторов x^k к оптимальному решению задачи (1), (2) потребуется следующее вспомогательное утверждение, доказываемое стандартным образом.

Лемма 2. *Если положительный ряд* $\sum_{k=0}^{\infty}\alpha_k$ *сходится, последовательность чисел* β_k, $k=0,1,2,\ldots,$ *— ограниченная, то ряд* $\sum_{k=0}^{\infty}\alpha_k\beta_k$ *сходится.*

С учетом вышеперечисленных фактов перейдем к доказательству следующей теоремы.

Теорема. *Пусть*

а) *задача* (1), (2) *невырождена;*

б) *целевая функция линейна;*

в) *объединения по всем* x, *удовлетворяющим условию* $\underline{x}\le x\le\bar{x}$, *множеств особых векторов линейных многообразий* $L_1(x)$, $L_2(x)$ *ограничены;*

г) *существует* $\tau>0$ *такое, что* $\lambda_k\ge\tau$, $k=0,1,2,\ldots$.

Тогда существуют векторы $\tilde{x}\in X$, $\tilde{u}\in R^m$, $\tilde{\alpha}\in R^n$, $\tilde{\beta}\in R^n$ *такие, что*

$$x^k\to\tilde{x},\quad u^k\to\tilde{u},\quad \alpha^k\to\tilde{\alpha},\quad \beta^k\to\tilde{\beta}\quad \text{при } k\to\infty.$$

Вектор $\tilde{x}$ *является оптимальным решением, а векторы* $\tilde{u}$, $\tilde{\alpha}$, $\tilde{\beta}$ *— множителями Лагранжа ограничений задачи* (1), (2).

Доказательство. Будем нумеровать отдельные фрагменты рассуждений.

1. Вычислительный процесс (7) равносилен следующему:

$$x^{k+1}=x^k+\tilde{\lambda}_k\Delta\tilde{x}^k, \tag{28}$$

где

$$\Delta\tilde{x}^k=\frac{-1}{c^\top\Delta x^k}\Delta x^k,\quad \tilde{\lambda}_k=-\big(c^\top\Delta x^k\big)\lambda_k. \tag{29}$$

Положим

$$\tilde{z}^k=\frac{-1}{c^\top\Delta x^k}\,z^k.$$

Из вспомогательной задачи (13), (14) вытекает, что векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$ являются решением задачи

$$\frac{1}{2}\Delta x^\top\big(B_k+D_k^{-1}\big)\Delta x+\frac{1}{2}z^\top H_k^{-1}z\to\min$$

при ограничениях

$$A_k\Delta x-z=0,\qquad c^\top\Delta x=-1.$$

То есть векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$ являются евклидовыми проекциями начала координат на линейное многообразие $L_1(x^k)$.

Векторы q^k, u^k определяются как результат решения следующей задачи нахождения евклидовой проекции начала координат на линейное многообразие $L_2(x^k)$:

$$\frac{1}{2}q^\top\Big(B_k+D_k^{-1}\Big)^{-1}q+\frac{1}{2}u^\top H_k u\to\min,$$
$$q-A_k^\top u=c.$$

Из леммы 1 и условия ограниченности объединений по x^k, $k=0,1,2,\ldots$, множеств особых векторов линейных многообразий $L_1(x)$, $L_2(x)$ следует, что векторы $\Delta\tilde{x}^k$, $\tilde{z}^k$, q^k, u^k будут образовывать ограниченные последовательности при некоторых $M_1>0$, $M_2>0$ для всех $k=0,1,2,\ldots$:

$$\| \Delta \tilde{x}^k \| \le M_1, \quad \| \tilde{z}^k \| \le M_1, \tag{30}$$

$$\| q^k \| \le M_2, \qquad \| u^k \| \le M_2. \tag{31}$$

В силу (28), (29) для $k = 0, 1, 2, \ldots$:

$$c^\top x^k - c^\top x^{k+1} = \tilde{\lambda}_k, \quad \tilde{\lambda}_k > 0.$$

Из (23) следует, что положительный ряд $\sum_{k=0}^{\infty} \tilde{\lambda}_k$ является сходящимся

$$\sum_{k=0}^{\infty} \tilde{\lambda}_k = c^\top x^0 - F. \tag{32}$$

Из (28), (30), (32) и леммы 2, примененной к процессу (28), получаем, что при некотором $\tilde{x} \in R^n$:

$$x^k \to \tilde{x} \quad \text{при } k \to \infty.$$

При этом $\tilde{x} \in X$, так как на всех итерациях $x^k \in X$ и множество X замкнуто.

Отметим, что в полученном обосновании сходимости векторов x^k условия на выбор весовых коэффициентов не использовались в полной мере. Использовался только факт положительности значений весовых коэффициентов d_j^k, h_i^k, $j = 1, \ldots, n$, $i = 1, \ldots, m$.

2. Из (21) и (23) следует

$$c^\top x^{k+1} - c^\top x^k \to 0 \quad \text{при} \quad k \to \infty.$$

В силу (7):

$$c^\top x^{k+1} - c^\top x^k = \lambda_k \big(c^\top \Delta x^k \big).$$

Поэтому

$$\lambda_k \big(c^\top \Delta x^k \big) \to 0 \quad \text{при} \quad k \to \infty.$$

Из (27) получаем, что

$$c^\top \Delta x^k \to 0 \quad \text{при} \quad k \to \infty. \tag{33}$$

Производная в точке оптимума минимизируемой в (13) функции должна быть равна нулю по любому направлению из R^n, в том числе и по направлению Δx^k:

$$-c^\top \Delta x^k = \big(\Delta x^k\big)^\top \big(B_k + D_k^{-1}\big) \Delta x^k + \big(\Delta x^k\big)^\top A_k^\top H_k^{-1} A_k \Delta x^k.$$

Из (14) и (33) следует, что

$$\big(\Delta x^k\big)^\top \big(B_k + D_k^{-1}\big) \Delta x^k + \big(z^k\big)^\top H_k^{-1} z^k \to 0 \quad \text{при} \quad k \to \infty. \tag{34}$$

Согласно (17), (19), (26), (34) при $k \to \infty$:

$$\big(q^k\big)^\top \big(B_k + D_k^{-1}\big)^{-1} q^k \to 0, \tag{35}$$

$$\big(u^k\big)^\top H_k u^k \to 0. \tag{36}$$

Из (31) следует, что предельные при $k \to \infty$ значения последовательностей векторов u^k, q^k образуют ограниченные множества, которые обозначим $Lim\, u^k$, $Lim\, q^k$.

В силу (9)–(12) последовательности векторов d^k, h^k являются ограниченными и, следовательно, имеют ограниченные множества предельных при $k \to \infty$ значений, которые обозначим $Lim\, d^k$, $Lim\, h^k$.

Для любого $\tilde{u} \in Lim\, u^k$ существует, согласно (36), вектор $\tilde{h} \in Lim\, h^k$ такой, что

$$\text{если } \tilde{u}_i \neq 0, \text{ то } \tilde{h}_i = 0, \quad \text{если } \tilde{h}_i \neq 0, \text{ то } \tilde{u}_i = 0, \quad i = 1, \ldots, m.$$

Аналогично, в силу (35), для любого $\tilde{q} \in Lim\, q^k$ существует вектор $\tilde{d} \in Lim\, d^k$ такой, что

$$\text{если } \tilde{q}_j \neq 0, \text{ то } \tilde{d}_j = 0, \quad \text{если } \tilde{d}_j \neq 0, \text{ то } \tilde{q}_j = 0, \quad j = 1, \ldots, n.$$

Учитывая условия на выбор весовых коэффициентов (9)–(12), получаем, что для вектора $\tilde{x} \in X$ при любых векторах $\tilde{u} \in Lim\, u^k$, $\tilde{q} \in Lim\, q^k$ выполняются условия стационарности.

Поскольку задача (1), (2) невырожденная, то для стационарного решения $\tilde{x} \in X$ векторы q и u, удовлетворяющие условиям (35) и (36) соответственно, единственны. Поэтому последовательность векторов u^k будет сходиться к одному вектору

$$\tilde{u} = \lim_{k \to \infty} u^k.$$

Соответственно, последовательность векторов q^k будет сходиться к единственному вектору

$$\tilde{q} = \lim_{k \to \infty} q^k.$$

3. Предположим, что у вектора $\tilde{u}$ есть отрицательные компоненты. Если $\tilde{u}_i < 0$ для некоторого i, то, начиная с некоторой итерации k_0, на всех последующих $u_i^k < 0$. Согласно (17), $z_i^k = h_i^k u_i^k$ и $h_i^k > 0$, то на всех последующих итерациях $z_i^k < 0$. Из (14) следует, что

$$\Big(\nabla f_i\big(x^k\big)\Big)^{\top} \Delta x^k < 0. \tag{37}$$

В силу (37) и условий на выбор шага λ_k для всех $k \geq k_0$ имеет место неравенство

$$f_i\big(x^k + \lambda_k \Delta x^k\big) < f_i\big(x^k\big).$$

Получаем, что значение $f_i\big(x^k\big)$ для данного i будет убывать по итерациям, поэтому $\tilde{h}_i > 0$. Неравенства $\tilde{h}_i > 0$ и $\tilde{u}_i < 0$ противоречат условию (36). Тем самым доказано, что $\tilde{u} \geq 0$.

Если $\tilde{u}_i > 0$ для некоторого i, то, начиная с некоторой итерации k_0, для всех $k \geq k_0$ значение $f_i\big(x^k\big)$ для данного i будет возрастать. Следовательно, значение $f_i\big(\tilde{x}\big)$ либо меньше нуля, либо равно нулю. В силу (36), $\tilde{h}_i = 0$. Отсюда, согласно (10), следует, что $\underline{\sigma}\big(-f_i(\tilde{x})\big) = 0$. Это означает, что $f_i\big(\tilde{x}\big) = 0$. Аналогично, если $f_i\big(\tilde{x}\big) < 0$ для некоторого i, то $\tilde{h}_i > 0$. Следовательно, $\tilde{u}_i = 0$. Тем самым доказано выполнение условий дополняющей нежесткости (3) для векторов $\tilde{x}$, $\tilde{u}$.

Неотрицательность компонент векторов $\tilde{\alpha}$, $\tilde{\beta}$ следует из правила (20). Покажем, что для векторов $\tilde{x}$, $\tilde{\alpha}$, $\tilde{\beta}$ выполняются условия дополняющей нежесткости (4), (5).

Множества номеров положительных, отрицательных и нулевых компонент $\tilde{q}$ обозначим Q_+, Q_-, Q_0 соответственно.

Пусть $j \in Q_+$, тогда $\tilde{q}_j > 0$ и $\tilde{\alpha}_j = 0$, $\tilde{\beta}_j = \tilde{q}_j$. Поскольку $\tilde{q}_j > 0$, то, начиная с некоторой итерации $\hat{k}$, на всех последующих $q_j^k > 0$. Согласно (19):

$$\Delta x_j^k = -\frac{d_j^k}{b_j^k d_j^k + 1}\, q_j^k. \tag{38}$$

Поскольку $q_j^k > 0$, $d_j^k > 0$, $b_j^k \geq 0$, то, в силу (38), $\Delta x_j^k < 0$ для всех $k \geq \hat{k}$. Из (7) и неравенства $\lambda_k > 0$ получаем, что величина x_j^k для данного j будет убывать по итерациям. Поэтому

$$\bar{x}_j - \tilde{x}_j > 0. \tag{39}$$

Из (35) и положительности q_j^k следует, что для данного номера j последовательность величин d_j^k сходится к нулю при $k \to \infty$. Это означает, что в силу определения весовых коэффициентов d_j^k и (39), $\left(x_j^k - \underline{x}_j\right) \to 0$ при $k \to \infty$. Отсюда следует, что условия (4), (5) выполняются для $j \in Q_+$.

Пусть $j \in Q_-$, тогда $\tilde{q}_j < 0$ и $\tilde{\alpha}_j = \tilde{q}_j$, $\tilde{\beta}_j = 0$. Если $\tilde{q}_j < 0$, то, начиная с некоторой итерации $\hat{k}$, на всех последующих $q_j^k < 0$. Поскольку $q_j^k < 0$, $d_j^k > 0$, $b_j^k \geq 0$, то, согласно (38), $\Delta x_j^k > 0$ для всех $k \geq \hat{k}$. Из (7) и неравенства $\lambda_k > 0$ получаем, что величина x_j^k для данного j будет возрастать по итерациям. Поэтому

$$\tilde{x}_j - \underline{x}_j > 0. \tag{40}$$

Из (35) и отрицательности q_j^k следует, что для данного номера j последовательность величин d_j^k сходится к нулю при $k \to \infty$. Это означает, что в силу определения весовых коэффициентов d_j^k и (40), $\left(\bar{x}_j - x_j^k\right) \to 0$ при $k \to \infty$. Отсюда следует, что для $j \in Q_-$ выполняются условия (4), (5).

Пусть $j \in Q_0$, тогда $\tilde{q}_j = 0$ и, следовательно, $\tilde{\alpha}_j = 0$, $\tilde{\beta}_j = 0$. Это означает, что для $j \in Q_0$ имеют место (4), (5).

Итак, доказано, что для векторов $\tilde{x}$, $\tilde{u}$, $\tilde{\alpha}$, $\tilde{\beta}$ выполняются условия дополняющей нежесткости (3)–(5). □

5. Обсуждение результатов

Представлено описание нового семейства алгоритмов внутренних точек для решения задач оптимизации с нелинейными ограничениями. Основной вычислительной проблемой в алгоритмах является нахождение на каждой итерации направления корректировки решения. Эта задача представляется в виде задачи безусловной минимизации квадратичной выпуклой функции от n или m переменных. Такая альтернатива может быть полезна в вычислительном отношении.

Осуществлено теоретическое обоснование сходимости последовательности приближений, вырабатываемых алгоритмами внутренних точек из описанного класса, к стационарному и оптимальному решению задачи (1), (2). Весовые коэффициенты в этих алгоритмах удовлетворяют условиям (9)–(12). Отметим, что для доказательства сходимости последовательности приближений, вырабатываемых вычислительным процессом (7), имеющиеся стандартные техники доказательства, описанные, например, у У.И. Зангвилла [7], не годятся. Это связано с тем, что при поиске направления улучшения решения используются изменяющиеся по итерациям нормы. Доказательство теоремы является некоторым переложением техники доказательства сходимости, предложенной в [5], для теоретического обоснования алгоритмов внутренних точек в линейном программировании.

Литература

1. **Дикин И.И.** Итеративное решение задач линейного и квадратичного программирования // Докл. АН СССР. — 1967. — Т. 174. — С. 747–748.
2. **Дикин И.И., Зоркальцев В.И.** Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). — Новосибирск: Наука, 1980.
3. **Евтушенко Ю.Г., Жадан В.Г.** Численные методы решения некоторых задач исследования операций // Журн. вычисл. матем. и мат. физики. — 1973. — Т. 13, № 3. — С. 583–597.
4. **Евтушенко Ю.Г., Жадан В.Г.** Барьерно-проективные методы решения задач нелинейного программирования // Журн. вычисл. матем. и мат. физики. — 1994. — Т. 34, № 5. — С. 669–684.
5. **Зоркальцев В.И.** Относительно внутренняя точка оптимальных решений. — Сыктывкар: Изд-во Коми филиала АН СССР, 1984.
6. **Зоркальцев В.И.** Метод относительно внутренних точек. — Сыктывкар: Изд-во Коми филиала АН СССР, 1986.
7. **Зангвилл У.И.** Нелинейное программирование. — М.: Советское радио, 1973.
8. **Пержабинский С.М.** Алгоритм внутренних точек, использующий квадратичные аппроксимации // Современные технологии. Системный анализ. Моделирование. — Иркутск: Изд-во ИрГУПС, 2008. — Т. 18, № 3. — С. 97–101.

Поступила в редакцию 30 июня 2011 г.