УДК 519.852

Локальное программирование [1)]

Б.Т.Поляк

(117997 Москва, ул. Профсоюзная, 65, ИПУ РАН)

Задачи математического программирования с дополнительным ограничением $\|x-a\| \leq \varepsilon$ назовем локальным программированием. Оказывается, если a является регулярной точкой исходной задачи, а $\varepsilon > 0$ достаточно мало, то эти задачи обладают свойствами выпуклых, хотя целевая функция и ограничения выпуклыми не предполагаются. Эти свойства основываются на общем принципе выпуклости образа малого шара при нелинейном отображении. Для задач локального программирования строится теория двойственности и специальные методы решения, обладающие высокой скоростью сходимости.

1. Введение

Известно, что выпуклость играет центральную роль в теории оптимизации. Для выпуклых задач необходимые условия экстремума , как правило, являются и достаточными; для них удается построить содержательную теорию двойственности; выпуклые задачи допускают применение эффективных глобально сходящихся методов оптимизации [1], [2]. Однако выпуклые задачи являются лишь маленьким островком в океане общих нелинейных задач оптимизации.

Тем интереснее выделить некоторые классы экстремальных задач, которые, будучи невыпуклыми, обладают многими свойствами выпуклых. Первым из таких классов являются задачи с интегральными функционалами [3, 4]. Вторым - задачи с квадратичными ограничениями и квадратичной целевой функцией [5 - 8]. И в том и другом случае удается доказать их выпуклость в пространстве образов. В первом случае это делается на основе теоремы А.А. Ляпунова о значениях векторнозначной меры; во втором - на основе результатов о выпуклости квадратичных отображений.

Ниже описывается еще один класс подобных задач, который мы называем локальным программированием. Здесь ключевую роль играет новый принцип выпуклости образа малого шара при нелинейном отображении. Структура статьи следующая. В п. 2 формулируется упомянутый выше принцип выпуклости. В п. 3 строится теория двойственности для задач локального программирования. Наконец, заключительный п.4 посвящен численным методам локального программирования. Оказывается, здесь возможны специальные методы оптимизации, обладающие быстрой сходимостью.

[1)] Работа поддержана грантами РФФИ 99-01-00340 и 00-15-96018

2. <u>Выпуклость образа малого шара</u>

Пусть X, Y - два гильбертовых пространства, f - нелинейный оператор, действующий из X в Y, имеющий Липшицеву производную в окрестности точки $a \in X$:

(1) $$\| f'(x_1) - f'(x_2) \| \le L \| x_1 - x_2 \|, \quad \forall x_1, x_2 \in B(a,r),$$

где $B(a,r)$ - шар с центром в a радиуса r:

$$B(a,r) = \{ x \in X : \| x - a \| \le r \}.$$

Основное предположение заключается в том, что a - регулярная точка отображения f, т.е. линейный оператор $f'(a)$ является сюръективным: он отображает X на Y ($f'(a) X = Y$). Например, если X, Y конечномерны: $X = R^n$, $Y = R^m$, то это условие означает, что ранг матрицы Якоби $f'(a) = \partial f(a) / \partial x$ равен m . Как хорошо известно, для сюръективного линейного отображения существует такая константа $\nu > 0$, что

(2) $$\| f'(a)^* y \| \ge \nu \| y \| \quad \forall y \in Y.$$

Здесь A^* обозначает сопряженный к A линейный оператор.

<u>Теорема</u> 1. Пусть выполняются условия (1) , (2) и $\varepsilon < \min\{r, \nu/(2L)\}$. Тогда образ шара $B(a,\varepsilon)$ является выпуклым, т.е. множество

(3) $$F = \{ f(x) : \| x - a \| \le \varepsilon \} \subset Y$$

выпукло. Более того, это множество строго выпукло, а его граница порождена граничными точками $B(a, \varepsilon)$: $\partial F \subset \{ f(x) : \| x - a \| = \varepsilon \}$

Для доказательства нам потребуются две леммы.

<u>Лемма</u> 1. Шар в гильбертовом пространстве сильно выпуклый: если $x_1, x_2 \in B(a, \varepsilon)$, то для $x_0 = (x_1 + x_2)/2$ имеем $B(x_0, \rho) \subset B(a, \varepsilon)$ для $\rho = \| x_1 - x_2 \|^2 / (8 \varepsilon)$.

Этот результат хорошо известен, он немедленно следует из равенства параллелограмма.

<u>Лемма</u> 2. Пусть f - нелинейное отображение из X в Y (X, Y - гильбертовы пространства) такое, что для некоторых $L, \rho, \mu > 0$ и $x_0 \in X$, $y_0 \in Y$ имеем

(4) $$\begin{aligned} &\| f'(x_1) - f'(x_2) \| \le L \| x_1 - x_2 \|, \quad \forall x_1, x_2 \in B(x_0, \rho), \\ &\| f'(x)^* y \| \ge \mu \| y \|, \quad \forall x \in B(x_0, \rho), \ y \in Y, \\ &\| f(x_0) - y_0 \| \le \rho \mu . \end{aligned}$$

Тогда уравнение $f(x) = y_0$ имеет решение $x^* \in B(x_0, \rho)$, причем $\| x_0 - x^* \| \le \| f(x_0) - y_0 \| / \mu$.

Это утверждение совпадает со следствием 1 к теореме 1 в [9].

<u>Доказательство теоремы</u>. Пусть x_1, x_2 - произвольные точки в $B(a,\varepsilon) \subset B(a,r)$, $y_i = f(x_i) \subset F$, $i = 1,2$. Обозначим $x_0 = (x_1 + x_2)/2$, $y_0 = (y_1 + y_2)/2$. Чтобы доказать выпуклость F, достаточно найти $x^* \in B(a, \varepsilon)$ такое, что $f(x^*) = y_0$. Мы имеем

$$y_i = f(x_i) = f(x_0) + f'(x_0)(x_i - x_0) + \varepsilon_i , \quad \| \varepsilon_i \| \le \frac{L}{2} \| x_i - x_0 \|^2 = \frac{L}{8} \| x_1 - x_2 \|^2, \ i = 1,2 .$$

Оценка остаточного члена в этой формуле справедлива в силу условия Липшица на $f'(x)$, см. [10] , теорема 3.2.12. Отсюда $y_0 = f(x_0) + \varepsilon_0$, $\varepsilon_0 = (\varepsilon_1 + \varepsilon_2)/2$, $\|\varepsilon_0\| \le \frac{L}{8}\|x_1 - x_2\|^2$. Далее в силу леммы 1 $B(x_0, \rho) \subset B(a, \varepsilon)$ с $\rho = \|x_1 - x_2\|^2/(8\varepsilon)$, а из (1), (2) имеем для любого $x \in B(x_0, \rho)$ и $y \in Y$:

$$\|f'(x)^* y\| \ge \|f'(a)^* y\| - \|(f'(x) - f'(a))^* y\| \ge \nu\|y\| - L\|x - a\|\|y\| \ge (\nu - L\varepsilon)\|y\|.$$

Таким образом, выполнены все условия леммы 2 с $\mu = \nu - L\varepsilon \ge L\varepsilon > 0$, $\rho = \|x_1 - x_2\|^2/8\varepsilon$, поскольку $\|f(x_0) - y_0\| \le \frac{L}{8}\|x_1 - x_2\|^2 \le \rho\,\mu$. Следовательно, уравнение $f(x) = y_0$ имеет решение x^* в $B(x_0, \rho) \subset B(a, \varepsilon)$. Это и доказывает выпуклость F.

Из вышеприведенного рассуждения также следует, что при $x_1 \ne x_2$ уравнение $f(x) = y$ имеет решение и при y, достаточно близких к y_0 ; это подтверждает строгую выпуклость F. Наконец, пусть x_0 - внутренняя точка $B(a, \varepsilon)$, тогда найдется $B(x_0, \rho) \subset B(a, \varepsilon)$ с некоторым $\rho > 0$ и для $\mu = \nu - L\varepsilon > 0$ будут выполнены все условия леммы 2. Поэтому для всякого y достаточно близкого к $f(x_0)$, уравнение $y = f(x)$ будет иметь решение в $B(a, \varepsilon)$. Таким образом, образ внутренних точек $B(a, \varepsilon)$ принадлежит внутренности F, т.е. ∂F порождается только границей $B(a, \varepsilon)$. Это завершает доказательство теоремы.

Вышеприведенное доказательство построено на использовании леммы 2, которая обоснована в [9] с помощью метода Ньютона. Можно опираться и на иные результаты, примыкающие к теореме Люстерника и приведенные в [11-12]. Впрочем сама теорема Люстерника также обычно доказывается с помощью метода Ньютона. Отметим, что нестрогое объяснение результата теоремы весьма просто. Шар $B(a, \varepsilon)$ сильно выпукл, поэтому его образ при линейном отображении $f'(a)$ тоже сильно выпукл. Но свойство сильной выпуклости устойчиво к малым возмущениям, поэтому оно не может потеряться при замене линейного отображения $f'(a)$ на близкое нелинейное $f(x)$. Из этих же соображений видно, что теорема не обобщается на банаховы пространства, в которых нет свойства сильной выпуклости шара. С другой стороны, результат можно обобщить, если вместо шара $B(a, \varepsilon)$ взять любое сильно выпуклое множество. Наконец, предположения о гладкости f нельзя существенно ослабить. Так, А.Иоффе построил пример, в котором f непрерывно дифференцируема, но образ F невыпукл.

Покажем на примере, что константы, входящие в формулировку теоремы , могут быть эффективно вычислены.

<u>Квадратичные отображения</u>. Пусть $X = R^n$, $Y = R^m$, $f(x) = \big(f_1(x), \dots, f_m(x)\big)$, где скалярные функции $f_i(x)$ квадратичны:

(5) $$f_i(x) = \frac{1}{2}(A_i x, x) + (a_i, x) + \alpha_i, \; i = 1, \dots, m .$$

Здесь A_i - симметричные матрицы $n\times n$, $a_i \in R^n$, $\alpha_i \in R^1$, $i=1,\ldots,m$. Возьмем без ограничения общности $a=0$, т.е. $B=\left\{ x: \|x\| \le \varepsilon \right\}$. Составим матрицу A со столбцами a_i, $A=\left(a_1 \middle| a_2 \middle| \ldots \middle| a_m\right)$.

<u>Утверждение</u> 1. Пусть ранг A равен m и $\varepsilon < \gamma/(2L)$, где γ - наименьшее сингулярное число A и $L=\left(\sum_{i=1}^{m} \|A_i\|^2\right)^{1/2}$. Тогда множество

(6) $$F=\left\{ f(x): \|x\| \le \varepsilon \right\}$$

выпукло в R^m.

<u>Доказательство</u>. Очевидно, $f_i'(x) = A_i x + a_i$ и $f'(0)=a_i$. Поэтому $f'(x)$ удовлетворяет условию Липшица с константой $L=\left(\sum L_i^2\right)^{1/2}$, $L_i = \|A_i\|$ на всем R^n; здесь $\|A_i\|$ - операторная норма A_i. С другой стороны, $\|f'(0)^* y\| = \|Ay\| \ge \nu \|y\|$, где ν - наименьшее сингулярное число A: : $\nu = (\min_i \lambda)^{1/2}$, λ_i - собственные числа матрицы $A^T A$. Таким образом, выполнены все условия теоремы.

Приведенный результат резко отличается от известных утверждений об образе произвольного шара при квадратичных отображениях [8] , где выпуклость удается доказать лишь при очень жестких предположениях (например, $m=2$, $n\ge 3$, $a_i \equiv 0$ - теорема Брикмана).

<u>Пример</u> 1. Рассмотрим квадратичное отображение

$$f_1(x) = x_1 x_2 - x_1, \quad f_2(x) = x_1 x_2 + x_2$$

с $m=n=2$. Утверждение 1 гарантирует выпуклость F при $\|x\| \le \varepsilon$, $\varepsilon < \varepsilon^* = 1\big/\left(2\sqrt{2}\right) \approx 0{,}3536$. Можно показать, что в данном случае оценка точная - выпуклость действительно теряется при $\varepsilon > \varepsilon^*$. На рис. 1 показан вид множества F при различных значениях ε.

3. <u>Двойственность в локальном программировании</u>

Наряду с обычной задачей математического программирования

$$\min f_0(x), \quad x\in R^n$$

(7) $$f_i(x) \le 0, \quad i=1,\ldots,l$$

$$f_i(x) = 0, \quad i=l+1,\ldots,m$$

рассмотрим ее "локальный" вариант с одним дополнительным ограничением

$$\min f_0(x),\ x \in R^n$$

(8) $$f_i(x) \le 0,\ i=1,\dots,l$$

$$f_i(x) = 0,\ i=l+1,\dots,m$$

$$\| x-a \| \le \varepsilon \ .$$

Сделаем следующие предположения.

<u>Условие</u> 1. Функции $f_i : R^n \to R^1$, $i = 0,\dots,m$ дифференцируемы и их градиенты удовлетворяют условию Липшица на $B(a, \varepsilon)$.

<u>Условие</u> 2. Точка a является допустимой в задаче (7), причем ограничения - неравенства выполняются как равенства :

$$f_i(a) = 0,\ i=1,\dots,m \ .$$

Это предположение не является ограничительным - если бы a не являлась допустимой, то задача (8) не имела бы решения при малых ε, а если некоторые неравенства в точке a неактивны, то их можно отбросить в задаче (8) при достаточно малых ε.

Следующее условие является ключевым.

<u>Условие</u> 3. Градиенты $f_i'(a)$, $i=0,\dots,m$ являются линейно независимыми. Если ограничения типа неравенств отсутствуют ($l = 0$) , то это условие означает, что a не является стационарной точкой в задаче на условный экстремум. Если же неравенства присутствуют, то это ограничение более жесткое, чем утверждение “ a не является точкой Куна- Таккера в задаче (7)”. Например, если (7) - задача линейного программирования, то условие 3 не выполняется в любой из вершин многогранника ограничений.

Составим функцию Лагранжа

(9) $$L(x,y) = \sum_{i=0}^{m} y_i f_i(x),\ \ y \in R^{m+1}$$

и обозначим $Y_+ = \left\{ y \in R^{m+1} : y_i \ge 0,\ \ i=0,\dots,l \right\}$.

<u>Теорема</u> 2. Пусть выполнены условия 1-3 . Тогда найдется $\varepsilon^* > 0$ такое, что для всех $0 < \varepsilon < \varepsilon^*$ решение x^* задачи (8) существует и выполняется необходимое условие экстремума :

$$\| x^* - a \| = \varepsilon \ ,$$

(10) $$L(x,y^*) \ge L(x^*,y^*) \quad \forall x : \| x-a \| \le \varepsilon \ ,$$

$$y_i^* f_i(x^*) = 0,\ i=1,\dots,l$$

для некоторого $y^* \in Y_+$, $y^* \ne 0$.

<u>Доказательство</u>. Задача (8) эквивалентна оптимизационной задаче в пространстве образов:

$$
(11)\qquad \begin{array}{l} \min\ f_0 \\ f \in F, \\ f_i \le 0, \quad i = 1,\dots,l \\ f_i = 0, \quad i = l+1,\dots,m \end{array} ,
$$

где $f = (f_0, f_1, \dots, f_m) \in R^{m+1}$, а $F = \left\{ f(x) : \left\| x - a \right\| \le \varepsilon \right\}$ - образ шара $B(a, \varepsilon)$ при отображении $f(x) = \left(f_0(x), \dots, f_m(x) \right)$. В силу условий 1- 3 выполнены все предположения Теоремы 1, поэтому F - выпуклое множество для достаточно малых $\varepsilon > 0$. Таким образом, (11) - выпуклая задача с непустым замкнутым ограниченным допустимым множеством. Поэтому у нее существует решение f^*, и по теореме отделимости существует гиперплоскость , задаваемая вектором $y^* \ne 0$, которая является опорной к допустимому множеству в (11) в точке f^*. Иначе говоря , $(y^*, f) \ge 0$ для всех $f \in F$, $f_i \le 0$, $i = 1,\dots,l$, $f_i = 0$, $i = l+1,\dots,m$, $f_0 \ge f_0^*$. Вспоминая определение F, убеждаемся, что это условие эквивалентно (10). Поскольку F строго выпукло (теорема 1), решение f^* лежит на границе F, а граница F порождается границей шара $B(a, \varepsilon)$ (Теорема 1) , поэтому $\left\| x^* - a \right\| = \varepsilon$.

Необходимое условие экстремума в (8) может быть сформулировано и в других формах.

Следствие. Если x^* - решение (8), то найдется $y^* \in Y_+$, $y^* \neq 0$, $y_i^* f_i(x^*) = 0$, $i = 1,\dots,l$ такое, что

(12) $$x^* = a - \varepsilon \frac{L_x(x^*, y^*)}{\| L_x(x^*, y^*) \|} .$$

Действительно, в соответствии с (10) $L(x, y^*)$ достигает минимума на $B(a, \varepsilon)$ в точке x^*, а необходимое условие минимума заключается в том, что градиент $L(x, y^*)$ по x в точке x^* (т.е. $L_x(x^*, y^*)$) является опорным к шару $B(a, \varepsilon)$ в его граничной точке x^*, т.е. $L_x(x^*, y^*) = -\lambda\,(x^* - a)$. Нормирующий множитель $\lambda = \| L_x(x^*, y^*) \| / \varepsilon$ находится из условия $\| x^* - a \| = \varepsilon$.

Как известно, в выпуклом программировании при выполнении условий регулярности (например, условия Слейтера) можно утверждать, что $y_0^* \neq 0$, т.е. можно взять $y_0^* = 1$. Аналогичный результат верен и для локального программирования.

Условие 4. Для любых $\varepsilon > 0$, $\sigma \in R^m$, $\sigma_i = 1$, $i = 1,\dots,l$, $|\sigma_i| = 1$, $i = l+1,\dots,m$ найдется x_σ такое что

(13) $$\sigma_i\, f_i(x_\sigma) < 0, \quad i = 1,\dots,m, \quad \| x_\sigma - a \| \le \varepsilon$$

Если в задаче имеются лишь неравенства $(l = m)$, то это условие является обычным условием Слейтера в задаче (7).

Теорема 3. Пусть выполняются условия 1-4 . Тогда в Теореме 2 можно взять $y_0^* = 1$, а (10) при $y_0^* = 1$ является необходимым и достаточным условием оптимальности в (8).

Доказательство. Пусть x^* - решение задачи (8) , тогда в соответствии с теоремой 2 найдется $y^* \in Y_+$, $y^* \neq 0$ такое что

$$y_0^*(f_0(x) - f_0(x^*)) + \sum_{i=1}^{m} y_i^* f_i(x) \ge 0 \quad \forall x: \| x - a \| \le \varepsilon$$

(здесь мы использовали, что $y_i^* f_i(x^*) = 0$, $i = 1,\dots,m$). Выберем σ такое, что $\sigma_i = sign\, y_i^*$, $i = 1,\dots,m$, и соответствующее x_σ из (13) . Тогда при $y_0^* = 0$ мы получаем из вышеприведенного неравенства $\sum_{i=1}^{m} y_i^* f_i(x_\sigma) \ge 0$, (поскольку $y^* \neq 0$), что противоречит (13) при $x = x_\sigma$. Таким образом $y_0^* > 0$, и мы можем отмасштабировать все y_i^* (поделив на y_0^*), чтобы получить $y_0^* = 1$.

Обратно , пусть выполняется (10) , а x - любая допустимая точка в задаче (8). Тогда из (10) при $y_0^* = 1$ следует $f(x) + \sum_{i=1}^{m} y_i^* f_i(x) \ge f(x^*)$, но поскольку $y_i^* f_i(x) \le 0$ для любой допустимой точки, то $f(x) \ge f(x^*)$, т.е. (10) - достаточное условие экстремума.

Наряду с исходной задачей (8) мы можем рассмотреть двойственную ей

(14) $$\max_{y \in Y_+} \psi(y),$$

где $\psi(y) = \min_{\|x-a\| \le \varepsilon} \left(f(x) + \sum_{i=1}^{m} y_i f_i(x) \right)$. Эта задача является выпуклой (Y_+ выпукло, а $\psi(y)$ как минимум линейных функций вогнута). Теорема 3 может рассматриваться как теорема двойственности - минимум в исходной невыпуклой задаче совпадает с максимумом в двойственной.

4. <u>Численные методы</u>

Мы начнем с простейшей задачи локального программирования, в которой дополнительные ограничения отсутствуют:

(15) $$\min_{\|x-a\| \le \varepsilon} f(x) .$$

Рассмотрим следующий итеративный метод ее решения:

(16) $$x^{k+1} = a - \varepsilon \frac{f'(x^k)}{\|f'(x^k)\|} .$$

Этот метод может рассматриваться, с одной стороны, как метод последовательных приближений для условия экстремума (12) в данном простейшем случае. С другой стороны, он может интерпретироваться как метод условного градиента [13] для решения задачи (15) при весьма специальном способе выбора длины шага. В самом деле, в методе условного градиента ищется минимум z^k линеаризованной функции $(f'(x^k), x)$ на допустимом множестве, а затем делается шаг $x^{k+1} = x^k + \lambda_k (z^k - x^k)$ при некотором $0 \le \lambda_k \le 1$. В нашем случае допустимое множество шар, и $z^k = a - \varepsilon \frac{f'(x^k)}{\|f'(x^k)\|}$. Если взять $\lambda_k \equiv 1$, то $x^{k+1} = z^k$ и мы приходим к методу (16).

Отметим, что (16) выглядит довольно необычно- в нем шаг делается не из текущей точки x^k, как в большинстве итерационных методов, а из одной и той же точки a.

<u>Теорема</u> 4. Пусть $f : R^n \to R^1$, где $f(x)$ дифференцируема на шаре $B(a, \varepsilon)$ и

(17) $$\|f'(x_1) - f'(x_2)\| \le L \|x_1 - x_2\| \quad \forall\ x_1, x_2 \in B(a, \varepsilon) .$$

Пусть, кроме того, $f'(a) \ne 0$ и

(18) $$\varepsilon < \frac{\|f'(a)\|}{2L}.$$

Тогда

а) Решение x^* задачи (15) существует, единственно, лежит на границе шара: $\|x^* - a\| = \varepsilon$ и для него

(19) $$x^* = a - \varepsilon \frac{f'(x^*)}{\|f'(x^*)\|} .$$

б) Для любого $x_0 \in B(a, \varepsilon)$ метод (16) сходится со скоростью геометрической прогрессии:

(20) $$\left\| x^k - x^0 \right\| \le q^k \left\| x^0 - x^* \right\| \le 2\varepsilon q^k, \quad q = \frac{\varepsilon L}{\left\| f'(a) \right\| - \varepsilon L} < 1 \quad .$$

Доказательство. Приведем прямое доказательство, не использующее предыдущие результаты. Для всех $x \in B(a, \varepsilon)$

$$\left\| f'(x) \right\| \ge \left\| f'(a) \right\| - \left\| f'(x) - f'(a) \right\| \ge \left\| f'(a) \right\| - L\varepsilon > L\varepsilon > 0$$

в силу (17), (18) , поэтому $f'(x) \ne 0$ на $B(a, \varepsilon)$ и минимум $f(x)$ не может достигаться внутри $B(a, \varepsilon)$. Поэтому x^* - граничная точка шара $B(a, \varepsilon)$; для нее (19) является необходимым условием экстремума.

Покажем единственность точки минимума. Пусть найдутся две различные точки минимума x_1^*, x_2^* . Для каждой из них выполняется (19), поэтому

(21) $$x_1^* - x_2^* = \varepsilon \left(\frac{f'(x_2^*)}{\left\| f'(x_2^*) \right\|} - \frac{f'(x_1^*)}{\left\| f'(x_1^*) \right\|} \right).$$

Используем следующую простую лемму. Если $\left\| b \right\| \ge \tau$, $\left\| c \right\| \ge \tau$, $\tau > 0$, то

(22) $$\left\| \frac{b}{\left\| b \right\|} - \frac{c}{\left\| c \right\|} \right\| \le \frac{1}{\tau} \left\| b - c \right\|.$$

Действительно, $\tau \dfrac{b}{\left\| b \right\|}$ является проекцией b на шар $B(0, \tau)$, а оператор проектирования - нерастягивающий. Поэтому из (21) получаем для $\tau = \left\| f'(a) \right\| - L\varepsilon$

$$\left\| x_1^* - x_2^* \right\| \le \frac{\varepsilon}{\tau} \left\| f'(x_2^*) - f'(x_1^*) \right\| \le \frac{\varepsilon L}{\tau} \left\| x_1^* - x_2^* \right\| = q \left\| x_1^* - x_2^* \right\|,$$

что невозможно при $x_1^* \ne x_2^*$, поскольку $0 < q < 1$.

Для обоснования скорости сходимости вычтем (19) из (16):

$$x^{k+1} - x^* = \varepsilon \left(\frac{f'(x^*)}{\left\| f'(x^*) \right\|} - \frac{f'(x^k)}{\left\| f'(x^k) \right\|} \right).$$

Поступая как и выше, получаем

$$\left\| x^{k+1} - x^* \right\| \le q \left\| x^k - x^* \right\| ,$$

что и приводит к (20).

Заметим, что знаменатель прогрессии q тем меньше, чем меньше ε : $q = O(\varepsilon)$. Поэтому скорость сходимости для малых ε очень велика. Результат теоремы 5 уточняет утверждение 5 теоремы 6.1 из [13] о линейной скорости сходимости метода условного градиента для минимизации функции с ненулевым градиентом на сильно выпуклом множестве.

Отметим еще, что если метод (16) применяется в ситуации, когда условия теоремы 4 не выполняются, то сходимости, как правило, нет, однако катастрофического "развала" метода не происходит (все итерации остаются в шаре

$B(a, \varepsilon))$. Поэтому поведение точек x^k обычно позволяет легко классифицировать ситуацию.

Таким образом, мы можем считать, что простейшая задача (15) допускает эффективный алгоритм решения. Это позволяет построить методы для решения общей задачи локального программирования. Действительно, пусть выполнены условия теоремы 3. Тогда вместо исходной задачи может быть рассмотрена двойственная (14). Вычисление двойственной функции

(23) $$\psi(y) = \max_{\|x-a\| \le \varepsilon} \left(f(x) + \sum y_i f_i(x) \right)$$

при любом $y \in Y_+$ сводится к решению задачи типа (15) (с $f(x)$, замененной на $f(x) + \sum y_i f_i(x)$). При этом, если $x(y)$ - решение задачи на максимум (23), то субградиент вогнутой функции $\psi(y)$ вычисляется без труда:

(24) $$\left(\partial \psi(y)\right)_i = f_i(x(y)), \ i = 1, \dots, m.$$

Таким образом, для решения двойственной задачи (14) можно применить тот или иной вариант субградиентного метода:

(25) $$\begin{aligned} y_i^{k+1} &= (y_i^k + \gamma_k f_i(x^k))_i^+, \quad i = 1, \dots, l, \\ y_i^{k+1} &= y_i^k + \gamma_k f_i(x^k), \quad i = l+1, \dots, m, \end{aligned}$$

$$x^k = \arg \max_{\| x-a \| \le \varepsilon} \left(f(x) + \sum y_i^k f_i(x) \right),$$

где γ_k - длина шага, $c^+ = \max\{0, c\}$, а для отыскания x^k применяется метод (16). Мы не останавливаемся на деталях (способы выбора γ_k, точность решения вспомогательной задачи для x^k и т.д.), т.к. эти вопросы хорошо изучены в теории субградиентных методов [14], [15].

Метод (16) может быть полезен и в ряде задач, которые не являются непосредственно задачами оптимизации. Пусть, например, нас интересует непосредственное описание множества F - образа (3) шара $B(a, \varepsilon)$ при отображении $f : X \to R^m$. Тогда опорная функция этого множества

$$\psi_F(c) = \max_{f \in F} (f, c), \quad c \in R^m$$

равна

$$\psi_F(c) = \max_{\| x-a \| \le \varepsilon} \sum_{i=1}^{m} c_i f_i(x),$$

и для ее вычисления удобен метод (16). В частности, когда $m=2$ (этот случай часто встречается в приложениях к линейной алгебре и теории устойчивости), вектор $c \in R^2$ может быть описан однопараметрически: $c_1 = \cos t$, $c_2 = \sin t$ и граница $F \subset R^2$ порождается точками $x(t)$, $0 \le t < 2\pi$:

$$x(t) = \arg \max_{\| x-a \| \le \varepsilon} (\cos t \; f_1(x) + \sin t \; f_2(x)),$$

которые удобно искать с помощью алгоритма (16).

5. Заключение.

Приведенная теорема о выпуклости образа малого шара при нелинейном отображении имеет множество приложений в линейной алгебре (спектр матрицы при ее малом возмущении или множество нулей полинома при вариации его коэффициентов), теории робастной устойчивости, оптимальном управлении (достижимое множество управляемой нелинейной системы при ограничении на управление вида $\|u\|_2 \le \varepsilon$ оказывается выпуклым) и ряде других областей. В настоящей работе мы не останавливаемся на этих приложениях, сделав акцент на новом специальном классе задач математического программирования - локальном программировании.

Литература

1. *Рокафеллар Р*. Выпуклый анализ. М.: Мир , 1973.
2. *Иоффе А.Д., Тихомиров В.М.* Теория экстремальных задач. М.: Наука, 1974.
3. *Аркин В.И.* О бесконечномерном аналоге задачи невыпуклого программирования // Кибернетика. 1967. Т. 2. № 1. С. 87-93.
4. *Aumann R.J.. Perles M.A*. A variational problem arising in economics // J. Math. Anal. Appl. 1965. V. 11. P. 488-503.
5. *Якубович В.А.* S - процедура в нелинейной теории управления // Вестник ЛГУ. Сер. матем. 1971. № 1. С. 62-77.

6. *Фрадков А.Л.* Теоремы двойственности в некоторых невыпуклых экстремальных задачах // Сиб. матем. журн. 1973. № 2. С. 357-383.
7. *Матвеев А.С., Якубович В.А.* Невыпуклые задачи глобальной оптимизации в теории управления. // Итоги науки и техники. Сер. "Соврем. матем. и ее приложения". Т. 60. М.: ВИНИТИ. 1998. С. 128-175.
8. *Polyak B.T.* Convexity of quadratic transformations and its use in control and optimization // Journ. Opt. Th. and Appl. 1998. V. 99. №. 3. P. 553-583.
9. *Поляк Б.Т.* Градиентные методы решения уравнений и неравенств // Ж. вычисл. матем. и матем. физ. 1964. Т. 4. № 6. С. 995-1005.
10. *Ортега Д., Рейнболд В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир. 1976.
11. *Дмитрук . А.В., Милютин А.А., Осмоловский Н.М.* Теорема Люстерника и теория экстремума // Усп. матем. наук. 1980. Т. 55. № 6. С. 11-46.
12. *Ioffe A.D.* On the local surjection property // Nonlinear Analysis: Theory, Methods and Appl. 1987. V. 11. № 5. P. 565-592.
13. *Левитин Е.С., Поляк Б.Т.* Методы минимизации при наличии ограничений // Ж. вычисл. матем. и матем. физ. 1966. Т. 6. № 5. С. 787-823.
14. *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка. 1979.
15. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука. 1983.