УДК 512.25/26

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

А.А.Каплан

В настоящей статье дается систематизированное изложение методов выпуклого программирования. В основу классификации положены различия способов исследования экстремальных задач. Такой подход, принятый во многих вопросах вычислительной математики, не является новым и в отношении математического программирования. Достаточно сослаться на хорошо известные у нас работы Зойтендейка [7], Левитина и Поляка [16], где с позиций теоретических концепций, предлагаемых авторами, рассмотрены отдельные классы вычислительных методов [*]. Представляется, что отсутствие сколь-нибудь полной систематизации такого рода препятствует направленному развитию методов выпуклого программирования и является причиной возникновения довольно большого числа алгоритмов, близких друг другу по существу и в то же время весьма различных по авторской интерпретации.

Обзору вычислительных методов предшествует изложение некоторых теоретических результатов.

§ 1. Определение и простейшие свойства вогнутых функций

Функция f, заданная на выпуклом множестве X m-мерного евклидова пространства R^m, называется <u>вогнутой</u>, если для лю-

---

[*] Принятая в настоящей статье классификация близка к намеченной в [16].

бых двух точек x^1 и x^2 из X и любого λ, $0 \le \lambda \le 1$,

$$f(\lambda x^1 + (1-\lambda)x^2) \ge \lambda f(x^1) + (1-\lambda) f(x^2). \qquad (1)$$

Если в соотношении (1) для любых различных x^1 и x^2 из X и любого λ, $0<\lambda<1$, имеет место строгое неравенство, функция f называется <u>строго вогнутой.</u> Заменяя (1) неравенством

$$f(\lambda x^1 + (1-\lambda)x^2) \le \lambda f(x^1) + (1-\lambda) f(x^2),$$

мы получим определения <u>выпуклой</u> и <u>строго выпуклой функции</u>. Отсюда ясно, что если функция f вогнута (строго вогнута), то $-f$ -выпукла (строго выпукла). Данное замечание позволяет ограничиться изучением свойств вогнутых функций.

Приведенные определения предполагают, что вместе с каждыми двумя точками x^1 и x^2 множеству X принадлежит весь отрезок $\lambda x^1 + (1-\lambda)x^2$, то есть требование выпуклости X является существенным.

Заметим, что функция f, заданная на выпуклом множестве X, является вогнутой в том и только том случае, когда выпукло множество $(m+1)$ -мерных векторов

$$A = \{a = (\alpha, x_1, x_2, \dots, x_m) : x = (x_1, x_2, \dots, x_m) \in X, \alpha \le f(x)\} \qquad (2)$$

Следовательно, выпуклыми при любом c являются множества $X_c = \{x \in X \ f(x) \ge c\}$.

Вогнутая функция непрерывна в каждой относительно внутренней точке *) множества, на котором она определена. В граничных точках области задания вогнутая функция может и не быть непрерывной. Так, функция $f(x) = \begin{cases} x+1 & \text{при } x>0 \\ 0 & \text{при } x=0 \end{cases}$ вогнута на множестве $[0,\infty)$, но в точке $x=0$ она разрывна. Ясно, что введенные выше множества X_c будут замкнутыми, если существует множество Y_c, открытое относительно аффинной оболочки X и такое, что $X_c \subset Y_c \subset X$.

Если функция f непрерывно дифференцируема и X - открытое множество, в определении вогнутости и строгой вогнутости соотношение (1) может быть заменено соотношением

---

*) Точка x называется относительно внутренней точкой множества X, если она является внутренней для X в его аффинной оболочке.

$$f(x^2) \le f(x^1) + \nabla f(x^1)(x^2 - x^1), \tag{3}$$

где $\nabla f(x^1)$ – градиент функции f, вычисленный в точке x^1.

Определенная в R^m вогнутая функция φ называется <u>суперлинейной относительно фиксированной точки x°</u>, если

$$\varphi(x^\circ + t(x - x^\circ)) = \varphi(x^\circ) + t(\varphi(x) - \varphi(x^\circ)),$$

каковы бы ни были вектор x и неотрицательное число t.

Пусть μ – вогнутая функция, заданная на открытом выпуклом множестве $G \subset R^m$, x° – фиксированная точка из G. Легко видеть, что функция

$$\mu_{x^\circ}(x) = \mu(x^\circ) + \lim_{t \to +0} \frac{\mu(x^\circ + t(x - x^\circ)) - \mu(x^\circ)}{t} \tag{4}$$

является суперлинейной относительно x° и $\mu_{x^\circ}(x) \ge \mu(x)$ при любом $x \in G$ (выражение $\frac{\mu(x^\circ + t(x - x^\circ)) - \mu(x^\circ)}{t}$ ограничено снизу и монотонно убывает при убывании t до 0, поэтому существование предела гарантировано). Функцию μ_{x° мы будем называть <u>опорной к μ в точке x°</u>. Если функция μ дифференцируема в точке x°, то μ_{x° линейна. По определению, $\mu_{x^\circ x^1 \ldots x^t}(x) = (\mu_{x^\circ x^1 \ldots x^{t-1}})_{x^t}(x)$. Очевидно, $\mu_{x^\circ x^1 \ldots x^t}$ суперлинейна относительно любой точки из аффинной оболочки точек $x^\circ, x^1, \ldots, x^t$. С введением опорной функции неравенство (3) сохраняет свой смысл и при отсутствии дифференцируемости, если под $f(x^1) + \nabla f(x^1)(x - x^1)$ понимать вычисленную в точке x функцию, опорную к f в точке x^1.

Принято говорить, что функция f достигает <u>локального максимума</u> на заданном множестве Q в точке $x^\circ \in Q$, если для любого $x \in Q$ из некоторой окрестности x° имеет место $f(x) \le f(x^\circ)$. Если при этом f – вогнутая функция, а Q – выпуклое множество, то $f(x^\circ) = \max_{x \in Q} f(x)$, то есть x° доставляет глобальный максимум. Множество $\{z \in Q : f(z) = \max_{x \in Q} f(x)\}$, как нетрудно показать, является выпуклым и, если f – строго вогнутая функция, состоит из одной точки.

<u>Конусом возможных направлений</u> выпуклого множества Q в точке $\xi^\circ \in Q$ будем, следуя Зойтендейку, называть конус

$$K_{\xi^0}(Q) = \bigcup_{x\in Q} \mathcal{L}(\xi^0, x),$$

где

$$\mathcal{L}(\xi^0, x) = \{z = \xi^0 + t(x - \xi^0) : t > 0\}.$$

Множество $Q_{\xi^0} = \overline{K_{\xi^0}(Q)}$ (черта означает замыкание) назовем <u>предельным конусом</u> выпуклого множества Q в точке ξ^0. По определению, $Q_{\xi^0 \xi^1 \ldots \xi^t} = (Q_{\xi^0 \xi^1 \ldots \xi^{t-1}})_{\xi^t}$ при любых

$$\xi^0 \in Q, \ \xi^1 \in Q_{\xi^0}, \ldots, \xi^t \in Q_{\xi^0 \xi^1 \ldots \xi^{t-1}}.$$

§ 2. Основные результаты теории выпуклого программирования

Пусть $f, g^1, \ldots, g^n$ – вогнутые функции, определенные на некотором открытом выпуклом множестве $G \subset R^m$, содержащем заданное выпуклое множество X^0.

<u>Основная задача</u> выпуклого программирования состоит в максимизации ж) функции f в области $\Omega = \bigcap_{j=0}^{n} X^j$, где $X^j = \{x \in G : g^j(x) \geq 0\}$ при $j = 1,2,\ldots, n$.

Отметим сразу, что добавлением одной переменной и одного ограничения можно перейти к эквивалентной задаче максимизации линейной функции

$$\psi(x_1, x_2, \ldots, x_{m+1}) = x_{m+1}$$

в выпуклой области $\tilde{\Omega} = \bigcap_{j=0}^{n+1} \tilde{X}^j$, где $\tilde{X}^0 = X^0 \times R^1$,

$$\tilde{X}^j = \{(x_1, x_2, \ldots, x_{m+1}) \in G \times R^1 : g^j(x_1, x_2, \ldots, x_m) \geq 0\}$$

при $j = 1, 2, \ldots, n$,

$$\tilde{X}^{n+1} = \{(x_1, x_2, \ldots, x_{m+1}) \in G \times R^1 : f(x_1, x_2, \ldots, x_m) - x_{m+1} \geq 0\}$$

Во многих работах основная задача формулируется в терминах выпуклых функций с заменой слова "максимизация" словом "минимизация" и неравенств $g^j(x) \geq 0$ неравенствами $g^j(x) \leq 0$.

Наиболее глубокие теоретические результаты связаны с исследованием наряду с основной задачей выпуклого программирования

ж) То есть ищется точка $x^0 \in \Omega$ такая, что $f(x^0) = \max_{x \in \Omega} f(x)$.

следующей в определенном смысле ей __двойственной задачи__.

Пусть Y° - множество векторов $y=(y_1, y_2, \dots, y_n)$ из R^n с неотрицательными компонентами ($y \geq 0$). Требуется минимизировать выпуклую функцию

$$\varphi(y) = \sup_{x \in X^\circ} L(x, y),$$

где

$$L(x, y) = f(x) + \sum_{j=1}^{n} y_j g^j(x), \tag{5}$$

на выпуклом множестве $\Omega^* = \{ y \in Y^\circ : \varphi(y) < +\infty \}$

Векторы $x \in \Omega$ и $y \in \Omega^*$ называются __допустимыми,__ а искомые векторы - __оптимальными.__

Для любых допустимых x и y справедливо соотношение

$$f(x) \leq L(x, y) \leq \varphi(y)$$

Следовательно, для оптимальности x° и y° достаточно, чтобы $x^\circ \in \Omega$, $y^\circ \in \Omega^*$ и $f(x^\circ) = \varphi(y^\circ)$. Последнее эквивалентно требованию, чтобы $x^\circ \in X^\circ$, $y^\circ \in Y^\circ$ и при любых $x \in X^\circ$, $y \in Y^\circ$ имело место соотношение

$$L(x, y^\circ) \leq L(x^\circ, y^\circ) \leq L(x^\circ, y) \tag{6}$$

При этом говорят, что точка (x°, y°) является __седловой__ для $L(x, y)$ - __функции Лагранжа__ исходной задачи. Заметим, что если (x°, y°) - седловая точка функции $L(x, y)$, то x° будет решением основной задачи, даже если не предполагать вогнутости функций $f, g^1, \dots, g^n$ и выпуклости множества X°. Однако наличие решения в задаче выпуклого программирования не означает существования седловой точки функции Лагранжа. В этом легко убедиться на простом примере [27], когда $m=1$, $n=1$, $f(x) = x$, $g^1(x) = -x^2$.

Вопрос о существовании седловой точки функции Лагранжа при наличии решения в основной задаче *) представляет собой центральный пункт теории выпуклого программирования. Точнее, вопрос состоит в том, при каких дополнительных условиях относительно функций $g^1, g^2, \dots, g^n$ и множества X° при произвольной функции f справедливо следующее утверждение:

---

*) То есть вопрос об эквивалентности задачи выпуклого программирования и соответствующей ей задачи о седловой точке.

(!) Для оптимальности вектора $x^0 \in X^0$ в основной задаче необходимо, чтобы существовал вектор y^0 такой, что (x^0, y^0) - седловая точка функции (5).

Необходимые и достаточные условия такого рода в классе дифференцируемых функций были получены Куном и Таккером [17]. Однако непосредственная их проверка, как правило, представляется весьма затруднительной. В настоящее время известен ряд более простых для проверки достаточных признаков (Слейтер [24], Удзава [27], Рубинштейн [23]), не предполагающих дифференцируемости функций f и $g^1, g^2, \ldots, g^n$. В практическом плане (то есть насколько это требуется для построения численных методов) они в значительной степени решают указанный выше вопрос.

При изложении численных методов мы будем всегда считать выполненным следующее достаточное условие [23], которое при сделанном в формулировке задачи предположении, что G - открытое множество, является наиболее общим. Именно:

(*) Для справедливости утверждения (!) достаточно, чтобы существовала точка $x^* \in X^0$, в которой $g^j(x^*) > 0$, $j = 1, 2, \ldots, n$, причем равенство $g^j(x^*) = 0$ допускается лишь для функций вида $g^j(x) = \sum_{i=1}^{m} a_{ij} x_i - b_j$.

Приведенный выше результат позволяет заменить непосредственное отыскание решения задачи выпуклого программирования отысканием седловой точки соответствующей функции Лагранжа. Методы, предназначенные для отыскания седловой точки функции Лагранжа, составляют <u>первую группу</u> в данной классификации.

Ниже формулируются приведенные в [14] требования, необходимые и достаточные для справедливости утверждения (!) без каких-либо дополнительных условий относительно характера функций f и $g^1, g^2, \ldots, g^n$.

УТВЕРЖДЕНИЕ I. Если в основной задаче выпуклого программирования $\sup_{x \in G} g^j(x) > 0$, $j = 1, 2, \ldots, n$, то для справедливости утверждения (!) при произвольной вогнутой функции необходимо и достаточно, чтобы

$$\bigcap_{j=0}^{n} X^j_{x^0 x^1 \ldots x^t} = \Omega_{x^0 x^1 \ldots x^t},$$

каковы бы ни были векторы $x^0 \in \Omega$, $x^k \in \Omega_{x^0 x^1 \ldots x^{k-1}}$; $k = 1, 2, \ldots, t$.

УТВЕРЖДЕНИЕ 2. Для справедливости утверждения (I) при произвольной вогнутой функции f необходимо и достаточно, чтобы для любых векторов $x^0 \in \Omega$, $x^k \in \Omega_{x^0 x^1 \dots x^{k-1}}$, $k = 1, 2, \dots, \tau$, имело место равенство

$$\Omega_{x^0 x^1 \dots x^\tau} = \{x \in X_{x^0 x^1 \dots x^\tau} : g^j_{x^0 x^1 \dots x^\tau}(x) \geq 0, j \in \mathcal{J}^\tau\},$$

где

$$\mathcal{J}^\tau = \{j : g^j_{x^0 x^1 \dots x^\tau}(x^\tau) = 0\}.$$

Ко <u>второй группе</u> отнесены методы выпуклого программирования, основанные на критерии локального экстремума.

Ввиду того, что, как уже отмечалось, любой локальный максимум в задаче выпуклого программирования (в рассматриваемой постановке) представляет собой решение задачи, исследование допустимого решения на оптимальность сводится к проверке условий локального экстремума. Точка $x^0 \in \Omega$ будет решением основной задачи в том и только том случае, когда пересечение конуса Ω_{x^0} (или $K_{x^0}(\Omega)$) с конусом $K_{x^0}(V)$, где $V = \{x \in G : f(x) > f(x^0)\} \cup \cup \{x^0\}$, совпадает с x^0. При выполнении приведенного выше условия (ж) вместо Ω_{x^0} можно рассматривать $\bigcap_{j=0}^{n} X^j_{x^0}$. Если к тому же функции $f, g^1, g^2, \dots, g^n$ дифференцируемы, а $X^0 = R^m$ (эти требования мы будем также считать выполненными всегда, когда речь будет идти о численных методах), указанное достаточное условие оказывается легко проверяемым. А именно, имеет место следующее утверждение:

(II) Для оптимальности вектора $x^0 = (x^0_1, x^0_2, \dots, x^0_m) \in \Omega$ необходимо и достаточно, чтобы нашлись неотрицательные числа y_j $(j \in \mathcal{J})$ такие, что

$$-\nabla f(x^0) = \sum_{j \in \mathcal{J}} y_j \nabla g^j(x^0), \tag{7}$$

где $\mathcal{J} = \{j : g^j(x^0) = 0\}$.

Приведенное условие близко к условию (I). Из (7) следует, что точка $(x^0, \bar{y})$, где $\bar{y}$ - n -мерный вектор с компонентами

$$\bar{y}_j = \begin{cases} y_j, & \text{если } j \in \mathcal{J}, \\ 0, & \text{если } j \in \{1, 2, \dots, n\} \setminus \mathcal{J}, \end{cases}$$

является седловой точкой функции Лагранжа исходной задачи.

Указанный критерий позволяет определять в произвольной допустимой точке $\bar{x} \in \Omega$, не являющейся решением задачи, направление возрастания целевой функции, принадлежащее конусу $\Omega_{\bar{x}} = \bigcap_{j=1} X_{\bar{x}}^j$ и такое, что сдвиг по этому направлению приводит нас в точку $\bar{\bar{x}}$ с лучшим значением целевой функции, причем

а) $\bar{\bar{x}} \in \Omega$ (методы возможных направлений);

б) $\bar{\bar{x}}$ – граничная точка, полученная проектированием $\nabla f(\bar{x})$ на $\Omega_{\bar{x}}$, и, кроме того, в некоторой точке $\bar{\bar{\bar{x}}} \in \Omega$, в определенном смысле близкой к $\bar{\bar{x}}$, имеем $f(\bar{\bar{\bar{x}}}) > f(\bar{x})$ (метод проекций градиента).

Заметим, что случаи а) и б) не являются взаимно исключающими.

Во многих численных методах решение исходной задачи получается как предел последовательности решений более простых вспомогательных экстремальных задач (в определенном смысле это относится и к методам второй группы, однако для них такая интерпретация не представляется естественной). Будем различать два типа таких методов.

<u>К третьей группе</u> мы будем относить те методы, в которых каждая вспомогательная задача состоит в отыскании безусловного экстремума некоторой функции. Естественно, в этом случае вспомогательные функции должны быть построены таким образом, чтобы при их максимизации было нежелательным сколь-нибудь существенное нарушение ограничений исходной задачи. Это достигается введением "штрафа" за нарушение ограничений. В большинстве таких методов речь идет о максимизации вспомогательных функций вида $f(x) + \Phi_K(x)$, где $\Phi_K(x) = \psi_K(g^1(x), g^2(x), \ldots, g^n(x)) < 0$ при всех x, причем $\lim_{K \to \infty} \Phi_K(x) = 0$, если $g^j(x) > 0$, $j = 1, 2, \ldots, n$, и $\lim_{K \to \infty} \Phi_K(x) = -\infty$, если хотя бы одно $g^j(x)$ меньше нуля. Отнесенные к третьей группе методы различаются, главным образом, способами введения штрафа.

<u>К четвертой группе</u> отнесем методы, в которых вспомогательные задачи являются задачами линейного программирования или задачами, специфика которых позволяет применять для их решения методы линейного программирования. При этом в основу построения вспомогательных задач положено то обязательство, что выпуклое ограниченное множество может быть как угодно хорошо аппроксимировано с помощью выпуклого многогранника, содержащего это множество или содержащегося в нем. Следовательно, рассматривая

задачу с линейной целевой функцией (как уже отмечалось, к такому виду может быть сведена любая задача выпуклого программирования), мы можем получить ее решение как предел последовательности решений задач максимизации линейной функции на многограннике, если будем аппроксимировать допустимое множество Ω последовательностью вложенных многогранников, содержащих Ω, или последовательностью расширяющихся многогранников, содержащихся в Ω. Методы четвертой группы различаются, в первую очередь, способами эффективного построения указанной последовательности многогранников.

Из методов, появившихся в последнее время, в самостоятельную, пятую в данной классификации группу следует выделить так называемые методы центров.

Эти методы остроумно эксплуатируют весьма простую идею, в которой в свое время, вероятно, успел разочароваться почти каждый, кто работает в данной области.

Идея состоит в следующем. Пусть нам известна допустимая точка x^1. Начиная из x^1, найдем точку x^2, в определенном смысле центральную для множества $\{x \in \Omega : f(x) \geq f(x^1)\}$. Повторим процесс, заменяя везде x^1 на x^2, и т.д.. Ясно, что такой подход оставался бесперспективным до тех пор, пока не был найден способ отыскания центральных точек, достаточно эффективный при имеющем здесь место усечении исходного множества.

Методы центров и состоят в последовательном построении вспомогательных экстремальных задач, с помощью которых определяется сходящаяся к решению исходной задачи последовательность центральных точек. При этом структура вспомогательных задач такова, что, как и в методах штрафов, для их решения без существенных изменений применимы алгоритмы отыскания безусловного экстремума.

Рассматривая методы второй, третьей, четвертой и пятой групп, мы будем считать, что допустимое множество Ω ограничено. Ясно, что с практической точки зрения это предположение не является обременительным.

Все перечисленные методы, естественно, пригодны для решения задач максимизации вогнутой функции с линейными ограничениями. Данную классификацию можно было бы продолжить, рассматривая специальные алгоритмы для задач этого типа.

При изложении численных методов выпуклого программирования предполагается, что читатель знаком с линейным программированием и методами отыскания абсолютного максимума вогнутой функции.

К сожалению, в литературе, посвященной выпуклому программированию, нет достаточных данных о численных экспериментах, по которым можно было бы судить о сравнительной эффективности различных методов. Поэтому при изложении пришлось ограничиться их качественным анализом.

В настоящей статье не рассматривается ряд весьма интересных в теоретическом отношении релаксационных методов ([1],[4],[6]). Следует, однако, отметить, что практическое использование этих методов осложнено отсутствием достаточно обоснованных рекомендаций по выбору множителей релаксации (это обстоятельство особенно важно ввиду немонотонности релаксационных процессов).

§ 3. Методы первой группы

На протяжении этого параграфа будем дополнительно предполагать, что функции $f, g^1, \ldots, g^n$ дважды непрерывно дифференцируемы, причем f - строго вогнута.

Изложению методов, основанных на непосредственном отыскании седловой точки, предпошлем описание некоторого градиентного процесса, построенного, главным образом, из геометрических соображений [3].

Как известно, градиент $\nabla f(x)$ определяет в каждой точке направление наискорейшего возрастания функции f, так что система дифференциальных уравнений

$$\frac{dx}{dt} = \nabla f(x), \quad x(0) = \xi \tag{8}$$

описывает траекторию из точки ξ до точки, в которой достигается абсолютный максимум f. Этот максимум может быть найден как решение системы (8) с любым начальным ξ при $t \to \infty$.

Для того, чтобы указанный метод соответствовал задаче с ограничениями, в (8) следует ввести условия, удерживающие x в области допустимых значений, причем удобно, чтобы эти условия влияли на решение системы лишь вне области допустимых значений.

Новая система имеет вид

$$\frac{dx}{dt} = \nabla f(x) + \sum_{j=1}^{n} \delta_j(x) \nabla g^j(x), \tag{9}$$

где

$$\delta_j(x) = \begin{cases} 0 & \text{при } g^j(x) \geq 0, \\ L & \text{при } g^j(x) < 0, \end{cases} \tag{10}$$

а L выбирается достаточно большим, так что в точках $x \in \Omega$, близких к границе Ω, вектор $\nabla f(x) + \sum_{j=1}^{n} \delta_j(x) \nabla g^j(x)$ направлен внутрь Ω.

На рис. I изображена качественная картина решения системы (9)-(10) конечно-разностным методом. Внутри области система описывает поиск абсолютного максимума f. Приблизившись к границе допустимой области, мы переходим её на K-том шаге. Далее, в соответствии с рис.I, так как нарушено неравенство $g^1(x) \geq 0$, решается система

$$\frac{dx}{dt} = \nabla f(x) + L \nabla g^1(x), \quad x(t_K) = x^K.$$

Возвратившись в допустимую область, мы как бы продолжаем поиск абсолютного максимума, и все повторяется.

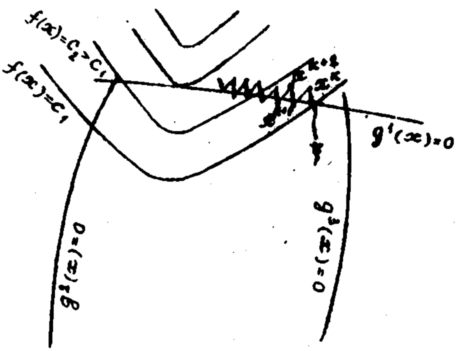

Рис. I

Рассмотренный метод лишь иллюстрирует градиентный процесс для случая задачи с ограничениями. Естественно, что выбранный весьма произвольный способ управления процессом не гарантирует сходимости к решению исходной задачи.

Эрроу и Гурвиц предложили ([32], [25]) градиентный метод, в котором управление процессом осуществляется в самой системе дифференциальных уравнений. Этот метод состоит в отыскании предела (при $t \to \infty$) решения системы дифференциальных уравнений

$$\frac{dx}{dt} = \nabla f(x) + \sum_{j=1}^{n} y_j \nabla g^j(x) \; ; \tag{II}$$

$$\frac{dy}{dt} = \begin{cases} -g^j(x), \text{ если } y_j > 0 \text{ или } g^j(x) < 0; \\ 0 \quad \text{- в противном случае.} \end{cases} \tag{I2}$$

В [25] показано, что при любых начальных $x(0), y(0)$ ($y(0) \geq 0$) существует единственное решение данной системы $x(t)$, $y(t)$, которое непрерывно зависит от $x(0)$, $y(0)$, причем $\lim_{t\to\infty} x(t) = \bar{x}$, где $\bar{x}$ - решение исходной задачи. Более сильное утверждение, что, кроме того, $y(t) \to \bar{y}$, причем ($\bar{x}$, $\bar{y}$) - некоторая седловая точка функции

$$L(x,y) = f(x) + \sum_{j=1}^{n} y_j g^j(x),$$

принадлежит, по-видимому, Карлину [15].

Несложный анализ показывает, что характер изменения y_j, определяемый уравнениями (II)-(I2), в общем соответствует тем требованиям, которые формулировались из геометрических соображений. Хотя здесь y_j не обязано равняться 0 при $g^j(x) \geq 0$, однако y_j как функция t, оставаясь неотрицательной, убывает при $g^j(x) \geq 0$ и возрастает при $g^j(x) < 0$, т.е. влияние слагаемого $y_j g^j(x)$ уменьшается при входе в допустимую область и увеличивается при нарушении соответствующего ограничения.

Следует отметить, что метод Эрроу-Гурвица мало применяется на практике. Это, в первую очередь, связано с серьёзными трудностями, возникающими при численном решении системы нелинейных дифференциальных уравнений даже в случае небольшого числа уравнений. Некоторые особенности численного решения системы (II)-(I2) рассмотрены на примере, приведенном в работе (I0).

Интересная итеративная схема реализации метода Эрроу-Гурвица исследована в работе Удзавы [26]. Рассматриваемая им система уравнений

$$x(t+1) = x(t) + \rho\left\{\nabla f(x(t)) + \sum_{j=1}^{n} y_j(t) \nabla g^j(x(t))\right\}, \tag{I3}$$

$$y_j(t+1)=\max\{0, y_j(t)-\rho g^j(x(t))\}, j=1,2,\dots,n, \qquad (14)$$

как нетрудно видеть, является разностным аналогом системы (11)-(12); с другой стороны, при малых $\rho>0$ она может интерпретироваться как некоторый итеративный процесс решения алгебраической системы

$$\nabla f(x)+\sum_{j=1}^{n} y_j \nabla g^j(x)=0,$$

$$y_j=\max\{0, y_j-\rho g^j(x)\}, j=1,2,\dots,n.$$

Пусть $\{x^\circ\}\times Y$ - множество всех седловых точек функции (в принятых предположениях $\{x^\circ\}\times Y$ -компакт). Введем в рассмотрение функцию

$$V(x,y)=\min_{y^\circ\in Y}\{|x-x^\circ|^2+|y-y^\circ|^2\} \qquad (15)$$

О характере последовательных приближений, построенных по формулам (13)-(14) при достаточно малом $\rho>0$, позволяет судить следующее утверждение ([26], доказательство см. в [34]).

Для любой начальной точки $x(0), y(0)$, где $y(0)\geq 0$, и любого $\varepsilon>0$ существует такое число $\rho_0>0$, что при решении системы (13)-(14) с $0<\rho\leq\rho_0$ найдется целое t^* со следующими свойствами:

$$V[x(t+1), y(t+1)]\leq V[x(t), y(t)] \quad \text{при } 0\leq t\leq t^*;$$

$$V[x(t), y(t)]\leq\varepsilon \quad \text{при } t\geq t^*.$$

Указанное утверждение носит характер теоремы существования. Оно не дает способа определения ρ_0. Проверить пригодность выбранного значения ρ эмпирически, основываясь хотя бы на монотонности $V[x(t), y(t)]$, также не представляется возможным, ибо в V входят неизвестные векторы x°, y°. Однако в последнее время появилась информация о том, что данный метод дает на практике хорошие результаты, если менять ρ, исходя из сравнения последовательных приближений, как это часто делается при решении нелинейных систем.

Известен ряд модификаций приведенных градиентного и итерационного методов. В частности, если характер функций f, g^j,

$g^1, \ldots, g^n$ позволяет установить, что $\max_x L(x,y)$ конечен для любого $y \geq 0$, то x_i можно рассматривать как функцию y и, решая относительно x систему $\nabla f(x) + \sum y_i \nabla g^i(x) = 0$ процесс, аналогичный (13)-(14), удается проводить лишь по y.

§ 4. Методы второй группы

Обозначим через J множество индексов, соответствующих линейным ограничениям задачи, $\bar{J} = \{1, 2, \ldots, n\} \setminus J$. Будем предполагать, что нам известна некоторая допустимая точка $\bar{x}$.

Методы возможных направлений [7], [9]

Как уже отмечалось, для того, чтобы точка $\bar{x} \in \Omega$ была решением исходной задачи, необходимо и достаточно, чтобы пересечение конусов $K_{\bar{x}}(\Omega)$ и $K_{\bar{x}}(V(\bar{x}))$, где

$$V(\bar{x}) = \{x \in G : f(x) > f(\bar{x})\} \cup \{\bar{x}\},$$

совпадало с $\bar{x}$. Если в точке $\bar{x}$

$$\Gamma(\bar{x}) = K_{\bar{x}}(\Omega) \cap K_{\bar{x}}(V(\bar{x})) \neq \{\bar{x}\},$$

то любой луч $\bar{x} + \lambda s \subset \Gamma(\bar{x})$ определяет в $\bar{x}$ <u>подходящее возможное направление</u>.

Зойтендейк [7] охарактеризовал методы возможных направлений следующими условиями:

1. Исходная точка $\bar{x}$ принадлежит допустимой области;

2. В полученной на k-том шаге точке $x^k \in \Omega$ определяется подходящее возможное направление s^k. Ищется точка $x^{k+1} = x^k + \lambda_k s^k \in \Omega$, в которой $f(x^{k+1}) > f(x^k)$;

3. Способ выбора направлений s^k и чисел λ_k обеспечивает сходимость последовательности $\{f(x^k)\}$ к $\max_{x \in \Omega} f(x)$.

При такой слишком общей и потому недостаточно содержательной характеристике к указанному классу можно отнести совершенно различные как по геометрической интерпретации, так и по способам доказательства сходимости методы, - по сути дела - все методы, в которых процесс идет по допустимым точкам с монотонным возрастанием значений целевой функции.

Для того, чтобы избежать такой общности, мы будем относить к методам возможных направлений лишь те методы, в которых подходящее возможное направление определяется путем непосредствен-

ного анализа конуса $\Gamma(x^*)$, т.е. информация о допустимом множестве используется только при построении указанного конуса.

Опишем идею отыскания подходящего возможного направления. Пусть $x \in \Omega$ – фиксированная точка,

$$N(x) = \{j \in J : g^j(x) = 0\}, \quad T(x) = \{j \in \bar{J} : g^j(x) = 0\},$$

$\alpha_j \ (j \in T(x))$ – положительные числа; σ – переменная величина. Определим

$$S(x) = \{(s,\sigma) \mid \nabla g^j(x) s - \alpha_j \sigma \geq 0 (j \in T(x)); \nabla g^j(x) s \geq 0 (j \in N(x))\}$$

Для выбора искомого направления в принципе достаточно решить следующую задачу:

Максимизировать σ

при условиях:

а) $(s,\sigma) \in S(x)$,

б) $\nabla f(x) s - \sigma \geq 0$,

в) требование нормализации.

Нормализация вводится с тем, чтобы задача имела ограниченное решение. Она необходима, так как условия а) и б) однородные.

Простое условие нормализации

$$-1 \leq s_i \leq 1, \quad i = 1, 2, \ldots, m, \tag{16}$$

введено в [7], [9]. В [7] также изучены четыре других условия и сформулированы общие требования, характеризующие нормализацию. Для отыскания подходящего возможного направления с учетом (16) приходится решать задачу линейного программирования.

Интересным является условие нормализации

$$-\sum_{i=1}^{m} s_i^2 \geq -1. \tag{17}$$

При его использовании следует ожидать более быстрой сходимости метода, чем в случае нормализации (16). Если обратиться к примеру, когда все ограничения исходной задачи линейные, легко видеть, что при нормализации (17) вектор направления s , отвечающий максимальному значению σ , будет составлять минимальный угол с градиентом целевой функции, вычисленным в точке x . Однако для определения вектора s здесь приходится решать задачу, в которой помимо линейных ограничений имеется

одно квадратичное.

Отметим, что в обоих случаях подходящее возможное направление получается в процессе решения вспомогательной задачи, как только будет найден допустимый вектор (s , σ) со значением $\sigma > 0$ *). Однако, если выбрать направление, соответствующее слишком малому значению σ , может оказаться, что в этом направлении в допустимой области возможен лишь малый сдвиг.

Длина шага определяется так, чтобы точка $x+\lambda s$ доставляла максимум функции f на пересечении луча $\{x+\tau s : \tau > 0\}$ с областью Ω . При этом обычно, прежде всего, находится точка $x+\tau' s$, доставляющая максимум функции f на указанном луче.

Описанный выше метод позволяет определить последовательность точек x^k такую, что соответствующая последовательность $\{f(x^k)\}$ сходится. Однако, если не принять мер, запрещающих малые шаги вдали от решения задачи (так называемое зигзагообразное движение), может оказаться, что $\lim_{k\to\infty} f(x^k) < \max_{x\in\Omega} f(x)$. Для устранения этого дефекта разработаны специальные приемы, уточняющие способ отыскания подходящего направления.

Первый прием, описанный в [7] (см. также [9]) состоит в следующем.

Пусть $x\in\Omega$ - фиксированная точка, $\varepsilon > 0$ - параметр,

$$N(x,\varepsilon)=\{j\in J : 0<g^j(x)<\varepsilon\},\ T(x,\varepsilon)=\{j\in T : 0<g^j(x)\leq\varepsilon\},$$

α_j $(j\in T(x,\varepsilon))$ - положительные числа. Рассмотрим

$$S(x,\varepsilon)=\{(s,\sigma) \mid \nabla g^j(x)s-\alpha_j\sigma\geq 0\,(j\in T(x,\varepsilon)),\ \nabla g^j(x)s>0\,(j\in N(x,\varepsilon))\}.$$

Для определения подходящего направления в сформулированной выше задаче теперь конус $S(x)$ заменяется конусом $S(x,\varepsilon)$. Кроме того, при решении задачи, выбирается подходящее направление, соответствующее значению $\sigma\geq\varepsilon$. Если оказывается, что $\max\sigma<\varepsilon$, вместо ε берется $\varepsilon/2$ и т.д. После перехода в следующую точку, в качестве нового значения параметра ε берем значение, которое этот параметр имел в конце предыдущего шага.

Опишем теперь второй прием (из [7]).

Пусть на K-том шаге решена задача выбора направления (с конусом $S(x^{k-1})$) и определена точка x^k . Если для

---

*) Если $\max\sigma = 0$, то x , как легко видеть, является решением исходной задачи.

$j \in T(x^k) + N(x^k)$ такого, что $j \notin T(x^{k-1}) + N(x^{k-1})$, мы ранее имели $j \in T(x^l) + N(x^l)$ при некотором $l \leq k-2$ (т.е. мы вернулись на гиперповерхность $g^j(x) = 0$, с которой ранее сошли), то в последующих задачах выбора направления добавляется дополнительное ограничение $\nabla g^j(x^k) - \lambda_j \sigma \geq 0$, препятствующее выходу на гиперповерхность $g^j(x) = 0$ при дальнейших шагах.

Эти дополнительные ограничения сохраняются до тех пор, пока при некотором k' не встретится один из следующих случаев: либо $\lambda_{k'} < \tau'_{k'}$, либо максимальное значение σ на k'-ом шаге меньше наперед заданного $\delta > 0$.

Если имеет место $\lambda_{k'} < \tau'_{k'}$, дополнительные ограничения опускаются, начиная с $(k'+1)$-го шага.

Если максимальное значение σ меньше δ, нужно заново решить задачу отыскания вектора $s^{k'}$, отбросив дополнительные ограничения. При этом δ заменяется на $\delta/2$.

Заметим, что учитывать первый из указанных случаев не обязательно. Однако при его учете как бы уменьшается опасность того, что дополнительное ограничение сохраняется, хотя мы уже далеко ушли от соответствующей гиперповерхности.

Сходимость последовательности $\{f(x^k)\}$ к $\max_{x \in \Omega} f(x)$ при использовании рассмотренных приемов показана в [7]. Заметим, что дробление ε и δ (до нуля) осуществляется автоматически в процессе решения по мере приближения к множеству оптимальных решений исходной задачи.

Второй из описанных приемов, по-видимому, более эффективный. При его применении мы скорее покидаем зону, далекую от оптимального решения.

Методы проекций градиента

Мы ограничимся изложением основной идеи метода проекций градиента, предложенного Розеном [22].

Будем предполагать, что целевая функция f основной задачи линейна и нам известна граничная допустимая точка x^k. Для ограничений, существенных[ж)] в точке x^k, определяются гиперплоскости, касательные к соответствующим гиперповерхностям $g^j(x) = g^j(x^k)$ в x^k. Градиент функции f проектируется на ортогональное дополнение подпространства, порождаемого нормалями к

ж) Т.е. для таких, что $|g^j(x^k)| < \delta$ при заранее выбранном $\delta > 0$

указанным касательным гиперплоскостям. Далее осуществляется сдвиг в направлении полученной проекции градиента, вообще говоря, выводящий из допустимой области. Для возвращения в Ω (точнее, в δ-окрестность границы множества Ω) в [22] предложен специфический итеративный процесс, приводящий нас в точку x^{k+1}. Получена оценка, позволяющая выбрать длину шага в направлении проекции градиента таким образом, чтобы при возвращении в Ω имело место $f(x^{k+1}) > f(x^k)$ и

$$\lim_{k\to\infty} f(x^k) = \max_{x\in\Omega} f(x) .$$

Преимущество данного метода по сравнению с методами возможных направлений заключается в том, что определение направлений сдвига не требует решения экстремальной задачи.

Следует иметь в виду, что если в точке x^k нормали к указанным касательным плоскостям линейно независимы, проектирование $\nabla f(x^k)$ сводится к вычислению вектора

$$z^k = \left(E - Q_k (Q_k' Q_k)^{-1} Q_k'\right) \nabla f(x^k),$$

где Q_k - матрица, столбцами которой являются указанные векторы - единичные нормали, штрих означает транспонирование, E - единичная матрица.

При переходе к точке x^{k+1} матрица (Q_{k+1}' Q_{k+1} $)^{-1}$, вообще говоря, не может быть вычислена, исходя из (Q_k', Q_k $)^{-1}$ (так как одновременно изменяются градиенты функций $g^i(x)$ и совокупность существенных ограничений), - это серьезный недостаток процесса. Второй недостаток в том, что для обеспечения сходимости метода приходится, как следует из имеющихся в [22] оценок, двигаться весьма малым шагом.

Отметим, что в [22] особо приведена схема метода проекций градиента для задачи максимизации вогнутой функции при линейных ограничениях. В этом случае указанные недостатки исчезают и метод представляется весьма перспективным.

§ 5. Методы третьей группы

Идея введения штрафа при решении основной задачи выпуклого программирования наиболее просто проявляется в следующем случае.

Пусть (x°, y°) – седловая точка функции $L(x,y)$, $g_-^j(x) = \min[0, g^j(x)]$, A – фиксированная величина, $A > \max_{j=1,2,\dots,n} y_j^\circ$. Тогда, как легко видеть,

$$f(x) + A\sum_{j=1}^{n} g_-^j(x) \le f(x) + \sum_{j=1}^{n} y_j^\circ g_-^j(x) \le f(x) + \sum_{j=1}^{n} y_j^\circ g^j(x) \le f(x^\circ),$$

причем

$$f(x^\circ) + A\sum_{j=1}^{n} g_-^j(x^\circ) = f(x^\circ)$$

и

$$f(\bar{x}) + A\sum_{j=1}^{n} g_-^j(\bar{x}) \le f(\bar{x}) \qquad , \text{ если } f(\bar{x}) \le f(x^\circ) .$$

Следовательно, если для задачи выпуклого программирования выполнено условие $(*)$, то всегда найдется константа a такая, что при любом $A > a$ точка безусловного максимума функции $f(x) + A\sum_{j=1}^{n} g_-^j(x)$ является решением указанной задачи.

Функция $A\sum_{j=1}^{n} g_-^j(x)$ естественным образом интерпретируется как штраф за нарушение ограничений, причем, как ясно из сказанного выше, при $A > a$ нарушение ограничений становится невыгодным (не может быть компенсировано возрастанием функции f)*).

Указанное благоприятное обстоятельство является специфическим свойством лишь некоторых функций штрафа**). При других функциях штрафа подобной константы a , как правило, не существует. Рассмотрим простой пример:

Максимизировать

$$f(x) = x$$

при ограничении

$$g(x) = -x \ge 0 .$$

Если здесь в качестве функции штрафа взять функцию $-A[g_-(x)]^2 = -A\{\min[0, -x]\}^2$, близкую по смыслу штрафа к $Ag_-(x)$,

---

*) Интересно в связи с этим вернуться к рассмотрению методов первой группы. Система (II)-(I2) может при некоторых уточнениях рассматриваться как система дифференциальных уравнений, описывающая поиск максимума функции $f(x) + A\sum_{j=1}^{n} g_-^j(x)$.

**) К сожалению, это обстоятельство обычно имеет место лишь при нарушении дифференцируемости функций штрафа в граничных точках множества Ω и исчезает при попытках сглаживания этих функций.

то при любом $A>0$ максимум $x-A\{\min[0,-x]\}^2$ достигается в точке $x=\frac{1}{2A}$, в то время как решением исходной задачи является $x=0$.

В данном примере введение указанной функции при большом A делает невыгодным значительное нарушение ограничения, однако выгодно нарушить ограничение в пределах $-\frac{1}{2A}\le g(x)<0$. Если же рассмотреть последовательность функций $x-A_k\{\min[0,-x]\}^2$ при условии, что $\lim_{k\to\infty} A_k=+\infty$, то, очевидно, мы получим последовательность точек максимума, сходящуюся к искомому решению задачи.

Отвлекаясь от конкретного вида штрафных функций, метод штрафов можно описать следующим образом:

Начиная процесс из некоторой, как правило, произвольной точки $\bar{x}$, на k-том шаге, мы ищем максимум функции $f(x)+\Phi_k(x)$, получаемой добавлением к целевой функции f исходной задачи штрафа Φ_k за нарушение ограничений. Точка x^k, в которой достигается экстремум функции $f(x)+\Phi_k(x)$, является начальной при отыскании на следующем шаге безусловного максимума функции $f(x)+\Phi_{k+1}(x)$.

Как уже отмечалось, будем предполагать, что

а) функции $\Phi_k(x)=\varphi_k(g^1(x), g^2(x), \ldots, g^n(x))$ вогнуты по x;

б) $\Phi_k(x)\le 0$ при всех x;

в) $\lim_{k\to\infty}\Phi_k(x)=0$, если $x\in\Omega$;

г) $\lim_{k\to\infty}\Phi_k(x)=-\infty$, если $x\notin\Omega$. Уже эти весьма естественные требования, как легко видеть, гарантируют, что любая предельная точка последовательности $\{x^k\}$ будет принадлежать Ω и доставлять максимум функции f на Ω.

Примерами штрафных функций такого типа являются (см., напр., [18]) следующие функции:

$$\Phi_k(x)=\sum_{j=1}^{n} A_k^j g^j(x), \qquad (18)$$

$$\Phi_k(x)=-\sum_{j=1}^{n} A_k^j [g^j(x)]^2, \qquad (19)$$

$$\Phi_k(x)=-\sum_{j=1}^{n}(-g^j(x)+1)^{A_k^j}+n, \qquad (20)$$

$$\Phi_k(x) = -\sum_{j=1}^{n}\left[-(g^j(x)-1)_-\right]^{A_k^j}. \tag{21}$$

Во всех этих формулах A_k^j – неотрицательные числа, причем в (20) и (21) они являются целыми; в (21), кроме того, $A_k^j \geq 2$.

Следует отметить, что так как $(-g_-^j(x)+1)^{A_k^j} - 1 =$ $= A_k^j(-g_-^j(x)) + c(x)$, где $c(x) \geq 0$, то $-(-g_-^j(x)+1)^{A_k^j}+1 =$ $= A_k^j\, g_-^j(x) - c(x) \leq A_k^j\, g_-^j(x)$. Отсюда ясно, что при $A_k^j \geq y_j^\circ$, $j=1,2,\ldots,n$, точка максимума $f(x) - \sum_{j=1}^{n}(-g_-^j(x)+1)^{A_k^j}$ доставляет решение исходной задачи. Функции (19) и (21) являются дифференцируемыми. Все перечисленные функции обладают указанными выше свойствами, если в каждом случае потребовать, чтобы $\lim_{k\to\infty} A_k^j = +\infty\ (j=1,2,\ldots,n)$.

Следует отметить полезные результаты [5], связанные с оценкой быстроты сходимости метода штрафов в случае штрафной функции (19). Так, если предположить, что f –линейная функция и вспомогательные задачи решаются точно, имеет место следующая оценка *)

$$f(x^\circ) \leq \max(f(x)+\Phi_k(x)) \leq f(x^\circ) + \sum_{j=1}^{n}\frac{(y_j^\circ)^2}{4A_k^j}$$

Если $A_k^j = A_k\ (j=1,2,\ldots,n)$, для величины

$$d(x^k) = \max\{|f(x^k)-f(x^\circ)|, |g_-^1(x^k)|,\ldots,|g_-^n(x^k)|\},$$

в определенном смысле характеризующей близость точки x^k, в которой достигается максимум функции $f(x) - A_k\sum_{j=1}^{n}[g_-^j(x)]^2$, до множества оптимальных решений исходной задачи, имеем

$$d(x^k) \leq \frac{1}{A_k}\max\left\{\frac{\sqrt{2}+1}{2}\|y^\circ\|, \frac{2+\sqrt{2}}{2}n\|y^\circ\|^2\right\},$$

где

$$\|y^\circ\| = \left\{\sum_{j=1}^{n}(y_j^\circ)^2\right\}^{1/2}.$$

Специфические штрафные функции, вогнутые лишь на множестве Ω, построены, исходя из несколько иных принципов, для случая, когда известна точка x^*, в которой $g^j(x^*) > 0$, $j=1,2,\ldots,n$

---

*) Здесь по-прежнему $(x^\circ, y_1^\circ, \ldots, y_n^\circ)$ – седловая точка $L(x,y)$.

Эти функции Φ_κ таковы, что при максимизации $f(x)+\Phi_\kappa(x)$ некоторый процесс отыскания безусловного экстремума может быть проведен строго внутри Ω.

Так, в работах [28] и [30] рассмотрены штрафные функции:

$$\Phi_\kappa(x) = -\alpha_\kappa \sum_{j=1}^{n} \frac{1}{g^j(x)}, \tag{22}$$

$$\Phi_\kappa(x) = \alpha_\kappa \sum_{j=1}^{n} \ln g^j(x), - \tag{23}$$

обе при условии, что $\alpha_\kappa > 0$.

В обоих случаях, полагая $\alpha_\kappa > \alpha_{\kappa+1} > 0$, $\lim_{\kappa\to\infty} \alpha_\kappa = 0$, получаем последовательность точек x^κ (доставляющих максимум во вспомогательных задачах), любая предельная точка которой является решением исходной задачи. В работе [28] при дополнительном предположении, что функции $f, g^1, \dots, g^n$ непрерывно дифференцируемы[ж)], а функция $P(x,\alpha) = f(x) - \alpha \sum_{j=1}^{n} \frac{1}{g^j(x)}$ строго вогнута по x при любом $\alpha > 0$, показано, что последовательность $\{x^\kappa\}$ сходится, причем имеет место следующее неравенство, позволяющее оценить близость по функционалу приближенного решения к точному:

$$f(x^\kappa) \le \max_{x\in\Omega} f(x) \le f(x^\kappa) + \alpha_\kappa \sum_{j=1}^{n} \frac{1}{g^j(x^\kappa)}.$$

Переходя к вопросу о решении вспомогательных задач, здесь полезно иметь в виду следующее. Ясно, что максимум функции $f(x) - \alpha_\kappa \sum_{j=1}^{n} \frac{1}{g^j(x)}$ на Ω достигается во внутренней точке множества Ω, сама функция вогнута на Ω. Далее, если $\bar{x}$ внутренняя точка Ω, при фиксированном κ непременно найдется $\varepsilon > 0$ такое, что $\bar{x}$ является также внутренней точкой множества $\Omega^\varepsilon = \{x : g^j(x) \ge \varepsilon, \quad j = 1, 2, \dots, n\}$ и, кроме того,

$$f(x) - \alpha_\kappa \sum_{j=1}^{n} \frac{1}{g^j(x)} < f(\bar{x}) - \alpha_\kappa \sum_{j=1}^{n} \frac{1}{g^j(\bar{x})},$$

ж) В [28] предполагается, что функции $f, g^1, \dots, g^n$ дважды непрерывно дифференцируемы, однако это требование избыточно. Доказательства используют лишь непрерывность первых производных.

если только $x \in \Omega \setminus \Omega^{\varepsilon}$.

Поэтому для отыскания максимума $f(x) - \alpha_k \sum_{j=1}^{n} \frac{1}{g^j(x)}$ может быть применен любой метод градиентного типа, предназначенный для отыскания экстремума без ограничений.[*] Придется лишь следить за длиной градиентного шага, чтобы не выйти из Ω (для этого предварительно определять точку пересечения вектора-градиента с границей множества Ω). При анализе сходимости процесса максимизации функции $f(x) - \alpha_k \sum_{j=1}^{n} \frac{1}{g^j(x)}$, начинающегося из точки $\tilde{x}$, можно воспользоваться известными теоремами, считая, что $\varphi_k(x)$ равна $-\alpha_k \sum_{j=1}^{n} \frac{1}{g^j(x)}$ на Ω^{ε} и продолжена с сохранением вогнутости и дифференцируемости на $R^m \setminus \Omega^{\varepsilon}$. Приближения градиентного процесса не зависят в данном случае от того, как задана φ_k вне Ω^{ε}. Аналогичные рассуждения проходят и в случае штрафной функции (23).

Остановимся теперь на некоторых деталях реализации метода штрафов, связанных в основном с отысканием безусловного максимума вспомогательных функций (для краткости ограничимся случаем, когда $A_k^j = A_k$, $j = 1, 2, \dots, n$).

1) Если обратиться к теоремам, дающим оценку быстроты сходимости различных процессов градиентного типа, нетрудно видеть, что в данном случае скорость сходимости будет, как правило, тем меньше, чем больше величина A_k в случае функций штрафа (18) - (21) и чем меньше величина α_k в случае функций (22) -(23)[**]. По этой причине даже при использовании функций штрафа (18) и (20) не следует пытаться в начале процесса выбирать

---

[*] Приводимые ниже замечания основаны на рассмотрении широко известного метода наискорейшего спуска и почти целиком переносятся на его модификации. Ссылки на ряд относительно новых работ, посвященных методам градиентного типа, читатель найдет в статье [21].

[**] Пример, иллюстрирующий это обстоятельство, приведен в [18]. Однако сделанное на основании этого примера заключение о том, что метод штрафов целесообразно применять лишь тогда, когда другие методы неприменимы, либо с целью получения начального приближения для других методов, представляется все же преждевременным, ибо, с одной стороны, замедление сходимости по мере приближения к решению обычно имеет место во всех методах выпуклого программирования, с другой стороны, часто удается достичь приемлемой скорости сходимости за счет удачного выбора функции штрафа.

большое A_1 (хотя, как уже отмечалось, выбирая в (18) и (20) A_1^i достаточно большими, мы сразу получаем решение исходной задачи). Не следует также быстро увеличивать A_k (уменьшать α_k).

2) По мере приближения к границе допустимой области при решении вспомогательных задач возникают серьезные трудности чисто вычислительного характера. Именно, требуется очень высокая точность промежуточных вычислений при определении градиента функции штрафа. Указанное обстоятельство дает основания переходить в этом случае при решении вспомогательных задач к методу покоординатного спуска (II). Сходимость процесса покоординатного спуска при использовании штрафных функций (22) - (23) получается с учетом сделанного выше замечания о продолжении функций $\varphi_k(x)$).

§ 6. Методы четвертой группы

Внутри этой группы мы будем различать так называемые методы отсечения, основанные на идее аппроксимации допустимого множества последовательностью выпуклых многогранников, и методы, основанные на аппроксимации вогнутых функций кусочно-линейными вогнутыми функциями.

<u>Методы отсечения.</u>

Будем считать, что целевая функция f исходной задачи линейна[ж)] и найдена система линейных неравенств, задающая ограниченный многогранник $\mathfrak{M}_0$, который содержит допустимое множество Ω этой задачи.

Идея отсечения принадлежит Келли [16] и состоит в построении последовательности многогранников

$$\mathfrak{M}_1 \supset \mathfrak{M}_2 \supset \dots$$

таких, что

$$\mathfrak{M}_0 \supset \mathfrak{M}_k \supset \Omega \quad \text{при любом} \quad k = 1, 2, \dots$$

ж) Как показано в § 2, к такому виду может быть приведена любая задача выпуклого программирования. Так как допустимое множество исходной задачи ограничено, добавлением еще одного неравенства $x_{m+1} \geq a$, где a -достаточно большое по абсолютной величине отрицательное число, можно добиться, сохраняя эквивалентность, чтобы допустимое множество преобразованной задачи также было ограниченным.

Многогранник $\mathfrak{M}_k$ получается в результате добавления линейных ограничений, отсекающих точку $y^{k-1} \in \mathfrak{M}_{k-1}$, в которой

$$f(y^{k-1}) = \max_{x \in \mathfrak{M}_{k-1}} f(x)$$

(для определения точки y^{k-1} требуется решить задачу линейного программирования). При этом отсечения должны строиться таким образом, чтобы $\lim_{k \to \infty} \inf_{x \in \Omega} \| y^k - x \| = 0$.

Последнее означает, что любая сходящаяся подпоследовательность $\{y^{k_i}\}$ будет иметь пределом некоторую точку максимума функции f на Ω .

В методе Келли отсечения строятся следующим образом.

Пусть $J_k = \{ j : g^j(y^{k-1}) < 0, \ j = 1, 2, \ldots, n \}$ (если $J_k = \emptyset$, то y^{k-1} является решением исходной задачи). Многогранник $\mathfrak{M}_k$ получается из $\mathfrak{M}_{k-1}$ добавлением ограничений

$$g^j(y^{k-1}) + \nabla g^j(y^{k-1})(x - y^{k-1}) \geq 0, \ j \in J_k .$$

Очевидно, что точка $x = y^{k-1}$ не удовлетворяет вновь введенным ограничениям и, тем самым, отсекается часть многогранника $\mathfrak{M}_{k-1}$, содержащая y^{k-1} . С другой стороны, в силу вогнутости $g^j(x)$

$$g^j(y^{k-1}) + \nabla g^j(y^{k-1})(x - y^{k-1}) \geq g^j(x),$$

т.е. в любой точке $x \in \Omega$

$$g^j(y^{k-1}) + \nabla g^j(y^{k-1})(x - y^{k-1}) \geq 0, \ j \in J_k .$$

Следовательно, $\mathfrak{M}_{k-1} \supset \mathfrak{M}_k \supset \Omega$.

Метод Келли дает оценку для максимума сверху, так что его одновременное использование с каким-либо методом второй группы позволяет оценивать близость полученного приближенного решения к точному (по значению функционала).

Следует особо отметить, что в данном случае не требуется, чтобы ограничения исходной задачи удовлетворяли условию (ж).

Другой способ отсечения предложен в работах [2], [12], [8]. Изложим этот способ в предположении, что множество Ω телесно и известна его внутренняя точка x^* .

На k-том шаге метода определяется точка $y^{k-1} \in \mathfrak{M}_{k-1}$, в которой $f(y^{k-1}) = \max_{x \in \mathfrak{M}_{k-1}} f(x)$, и крайняя точка x^{k-1} пересечения луча $\{ x = x^* + \lambda (y^{k-1} - x^*) : \lambda > 0 \}$ с множеством Ω . Затем строится любой ненулевой линейный функционал l_{k-1} такой, что

$\eta_{k-1}(x - x^{k-1}) \geq 0$ для $x \in \Omega$. Многогранник $\mathcal{M}_k$ получается из $\mathcal{M}_{k-1}$ добавлением одного ограничения $\eta_{k-1}(x) - \eta_{k-1}(x^{k-1}) \geq 0$, отсекающего точку y^{k-1}.

Доказано, что при обоих способах отсечения, если $k \to \infty$, то $\lim \inf\limits_{x \in \Omega} \| y^k - x \| = 0$, т.е. имеет место указанная выше сходимость. При втором методе на каждом шаге имеются двухсторонние оценки для решения исходной задачи

$$f(y^k) \geq \max_{x \in \Omega} f(x) \geq f(x^k),$$

причем

$$\lim_{k \to \infty} \| y^k - x^k \| = 0.$$

Отметим также, что этот метод в отличие от других методов, приведенных в данной статье, без каких-либо изменений применим при любом способе задания выпуклого множества Ω, в частности, когда функции $g^1, g^2, \ldots, g^n$ квазивогнутые*). Способ задания Ω может проявиться разве лишь при определении крайней точки пересечения луча $\{x = x^* + \lambda (y^{k-1} - x^*) : \lambda > 0\}$ с множеством Ω. Отметим, что, какова бы ни была дифференцируемая функция g^j, обращающаяся в ноль в точке $\bar{x} \in \Omega$, неравенство $\nabla g^j(\bar{x})(x - \bar{x}) \geq 0$ справедливо для всех точек $x \in \Omega$, так что, если на k-том шаге процесса $g^j(x^k) = 0$, $\nabla g^j(x^k) \neq 0$, в качестве η_k можно взять $\nabla g^j(x^k)$.

Естественно, что число шагов, требующихся для получения приближенного решения задачи, особенно в первом методе, существенно зависит от выбора начального многогранника $\mathcal{M}_0$. Если мы не имеем достаточной информации, чтобы построить начальный многогранник, сколько-нибудь удовлетворительно аппроксимирующий множество Ω, можно пойти на риск, что $\mathcal{M}_0$ отрежет часть Ω. Если окажется, что

$$\max_{x \in \mathcal{M}_0 \cap \Omega} f(x) < \max_{x \in \Omega} f(x)$$

(изменение $\mathcal{M}_0$ потребуется только в этом случае), то этот факт будет нетрудно обнаружить в процессе решения. Одновремен-

---

*) Заданная на выпуклом множестве $X \subset R^m$ функция f называется квазивогнутой, если при любом c выпуклы множества $X_c = \{x \in X : f(x) \geq c\}$.

но выяснится, в каком направлении следует расширить исходный многогранник $\mathcal{M}_0$. При расширении многогранника добавленные на предыдущих шагах ограничения целесообразно сохранить, так как они и в этом случае не могут отсечь точек множества Ω. Ясно, что при построении $\mathcal{M}_0$ выгодно использовать линейные ограничения исходной задачи, если таковые имеются.

В процессе реализации указанных методов на каждом шаге приходится решать задачу линейного программирования, которая получается из задачи, решенной на предыдущем шаге, добавлением нескольких (во втором методе - одного) ограничений. В связи с этим для решения задач линейного программирования целесообразно применять так называемые двойственные алгоритмы. В качестве начальн о базиса (двойственной задачи) следует выбирать оптимальный базис, полученный на предыдущем шаге. Если число добавленных на предыдущем шаге ограничений мало, такой подход, как правило, позволяет найти решение рассматриваемой задачи линейного программирования за небольшое число шагов.

Сравнивая приведенные выше алгоритмы, следует отметить, что во втором методе ограничения создают более плотную аппроксимацию допустимого множества, число ограничений растет медленнее, чем в методе Келли. Кроме того, как уже отмечалось, на каждом шаге второго метода мы получаем двухсторонние оценки решения.

С другой стороны, метод Келли, вероятно, более эффективен в том случае, если отыскание внутренней точки $x^* \in \Omega$ требует решения вспомогательной задачи общего вида (см. § 8).

В работе [13] предложена модификация, позволяющая использовать второй способ отсечения, когда допустимое множество не удовлетворяет условию (*). В качестве x^* в этом случае берется относительно внутренняя точка множества Ω. Следует отметить, что без существенных изменений рассмотренный метод применим в практически важном случае, когда множество Ω имеет внутреннюю точку относительно аффинного многообразия, определяемого некоторой частью линейных ограничений исходной задачи.

<u>Метод линеаризации</u> ([20], [31])

Пусть $x^1, x^2, \ldots, x^T$ - набор m-мерных векторов. Каждая точка x, принадлежащая выпуклой оболочке этого набора, представима в виде

$$x = \sum_{t=1}^{T} \lambda_t x^t, \qquad (24)$$

где

$$\lambda_t \geq 0,\ t=1,2,\ldots,T,\ \sum_{t=1}^{T}\lambda_t = 1. \qquad (25)$$

В качестве линеаризации вогнутой функции $\eta(x)$ примем

$$\bar\eta(x) = \sum_{t=1}^{T}\lambda_t\,\eta(x^t). \qquad (26)$$

Заменяя функции исходной задачи f, $g^1,\ldots,g^n$ функциями $\bar f$, $\bar g^1,\ldots,\bar g^n$, мы получаем следующую задачу линейного программирования относительно λ_1, $\lambda_2,\ldots,\lambda_T$:

$$\begin{aligned}
&\sum_{t=1}^{T}\lambda_t\,g^j(x^t) \geq 0,\ j=1,2,\ldots,n;\\
&\lambda_t \geq 0,\ t=1,2,\ldots,T;\ \sum_{t=1}^{T}\lambda_t = 1; \qquad (27)\\
&\sum_{t=1}^{T}\lambda_t\,f(x^t) - \max.
\end{aligned}$$

Исследуем соотношение между решениями построенной и исходной задачи.

Пусть $\bar\lambda_1$, $\bar\lambda_2,\ldots,\bar\lambda_T$ – решение задачи (27). Тогда вектор $\bar x = \sum_{t=1}^{T}\bar\lambda_t x^t$ удовлетворяет ограничениям $g^j(x)\geq 0,\ j=1,2,\ldots,n$, если только точки x^1, $x^2,\ldots,x^T$ принадлежат множеству Ω.

Действительно, из вогнутости g^j следует, что

$$g^j(\bar x) = g^j\Big(\sum_{t=1}^{T}\bar\lambda_t x^t\Big) = g^j\Big[\bar\lambda_1 x^1 + (1-\bar\lambda_1)\sum_{t=2}^{T}\frac{\bar\lambda_t}{1-\bar\lambda_1}x^t\Big] \geq$$

$$\bar\lambda_1 g^j(x^1) + (1-\bar\lambda_1)\,g^j\Big[\sum_{t=2}^{T}\frac{\bar\lambda_t}{1-\bar\lambda_1}x^t\Big].$$

Как легко видеть, $\sum_{t=2}^{T}\frac{\bar\lambda_t}{1-\bar\lambda_1} = 1$, и несложная индукция с учетом того, что $g^j(x^t)\geq 0$ при $t = 1,2,\ldots,T$ дает

$$g^j(\bar x) \geq \sum_{t=1}^{T}\bar\lambda_t\,g^j(x^t) \geq 0.$$

Из вогнутости f точно так же имеем $\sum_{t=1}^{T}\bar\lambda_t f(x^t) \leq f(\bar x)$.

Тем самым точка $\bar x$, соответствующая решению задачи (27), принадлежит множеству Ω, причем максимальное значение целевой функции в (27) не превосходит $f(\bar x)$.

Решение задачи (27) рассматривается как приближенное для

исходной задачи. Основная трудность состоит в выборе системы векторов $x^1, x^2, \ldots, x^T$ для получения хорошего приближения. Ясно, что построение равномерной сетки на Ω даже при небольшом числе переменных $x_1, x_2, \ldots, x_m$, как правило, будет приводить к огромному числу λ_t. Однако при наличии хорошего начального приближения в исходной задаче, линеаризацию можно проводить лишь в окрестности решения, что позволяет обойтись вполне обозримым количеством переменных.

Сравнительно небольшое число переменных при линеаризации получается, когда функции f, g^j, $j = 1,2,\ldots, n$, имеют вид:

$$f(x) = \sum_{i=1}^{m} f_i(x_i),$$

$$g^j(x) = \sum_{i=1}^{m} g_i^j(x_i).$$

В этом случае линеаризация проводится отдельно по каждой переменной ($x_i = \sum_{t=1}^{T_i} \lambda_{it} x_i^t$), и соответствующая задача линейного программирования может быть записана следующим образом:

$$\begin{gathered}\sum_{i=1}^{m}\sum_{t=1}^{T_i} \lambda_{it}\, g_i^j(x_i^t) \geq 0;\\ \lambda_{it} \geq 0,\quad i = 1,2,\ldots,m;\quad \sum_{t=1}^{T_i} \lambda_{it} = 1;\\ \sum_{i=1}^{m}\sum_{t=1}^{T_i} \lambda_{it}\, f(x_i^t) - \max.\end{gathered} \tag{28}$$

Сходимость последовательности решений линеаризованных задач к решению исходной задачи при дроблении шага показана в [20]. Естественно, для уточнения полученного при выбранной линеаризации решения дробление шага следует проводить лишь в окрестности этого решения.

Заметим, что аппроксимации в (28) с Tm переменными λ_{it} в (27) соответствует аппроксимация с T^m переменными, ибо в (27) приходится рассматривать выпуклую оболочку векторов ($x_1^{t_1}$, $x_2^{t_2}, \ldots, x_m^{t_m}$) с различными наборами $\{t_1, t_2, \ldots, t_m\}$. (При сравнении предполагается, что все t_j меняются от 1 до T).

В некоторых работах при построении задачи (28) дополнительно предполагается, что при каждом i не более двух значе-

ний λ_{it_1} и λ_{it_2} отличны от нуля, и если $\lambda_{it_1}>0$, $\lambda_{it_2}>0$, то $|t_2-t_1|=1$. Указанное условие введено с тем, чтобы гарантировать единственность представления $x_i=\sum_{t=1}^{T_i}\lambda_{it}x_i^t$, однако в данном случае оно излишне [*] и обременительно, так как должно учитываться при изменении базиса в процессе решения задачи линейного программирования.

§ 7. Методы пятой группы

Как пояснялось в § 2, в основу методов центров положена идея эффективного отыскания центральной (в определенном смысле) точки выпуклого телесного множества.

Ниже формулируются приведенные в [33] весьма общие условия, которым должна удовлетворять вспомогательная функция с тем, чтобы она могла быть использована для управления поиском центральных точек при описанном в § 2 процессе усечения допустимой области.

Именно, пусть $v(x,\lambda)$ - заданная на множестве

$$\{(x,\lambda)|\, x\in\Omega;\ f(x)-\lambda\geq 0;\ 0\leq\lambda\leq\max_{x\in\Omega} f(x)\}$$

непрерывная функция, удовлетворяющая следующим условиям:

1) $v(x,\lambda)=0$, если $\min\{g^1(x), g^2(x),\dots,g^n(x), f(x)-\lambda\}=0$;

2) $v(x,\lambda)>0$, если $\min\{g^1(x), g^2(x),\dots,g^n(x), f(x)-\lambda\}>0$;

3) Если $\lambda'\leq\lambda''$, то $v(x,\lambda')\geq v(x,\lambda'')$ для всех $x\in\Omega$ таких, что $f(x)\geq\lambda''$;

4) Если $\bar\lambda\leq\max_{x\in\Omega} f(x)$ - фиксированное число, $f(x(\lambda))-\lambda\to 0$ при $\lambda\to\bar\lambda-0$, то $v(x(\lambda),\lambda)\to 0$.

Здесь $x(\lambda)$ - произвольная фиксированная вектор-функция параметра λ такая, что $x(\lambda)\in\Omega_\lambda=\{x\in\Omega: f(x)\geq\lambda\}$ при каждом $\lambda\leq\max_{x\in\Omega} f(x)$.

Выберем в качестве x^1 внутреннюю точку множества Ω и, полагая $\lambda=f(x^1)$, определим точку x^2, в которой $v(x^2,\lambda)=\max_{x\in\Omega_\lambda} v(x,\lambda)$. Повторяя затем процесс с x^2 вместо x^1 и

---

*) Единственность этого представления важна при решении нелинейных задач общего вида.

т.д., мы найдем последовательность $\{x^k\}$ внутренних точек множества Ω, которая, как можно показать, сходится к решению исходной задачи.

Само понятие центральной точки, тем самым, определяется заданием функции $\vartheta(x,\lambda)$. Оно становится более ясным на примере тех функций $\vartheta(x,\lambda)$, которые реально используются при вычислениях.

Естественное требование, предъявляемое к этим функциям, состоит в том, что вспомогательные задачи, решаемые на каждом шаге, должны быть как можно проще.

В работе [33] рассмотрены две функции

$$\vartheta(x,\lambda)=\min\{f(x)-\lambda, g^1(x),\ldots,g^n(x)\}, \tag{29}$$

$$\vartheta(x,\lambda)=(f(x)-\lambda)\times\prod_{j=1}^{n} g^j(x). \tag{30}$$

При использовании функции (29) в процессе ее максимизации (по x при заданном λ) нет необходимости заботиться о выполнении ограничений: они, как легко видеть, будут выполняться автоматически при любом монотонном процессе. Использование второй функции делает учет ограничений существенным. Однако, как и в случае метода штрафов о функцией (22), здесь с незначительной модификацией может быть использован любой процесс градиентного типа, предназначенный для отыскания безусловного экстремума. Сходимость процесса следует из рассуждений, аналогичных приведенным в § 5. Отметим, что функция (30) не является вогнутой, однако при любом фиксированном λ она квазивогнута*) на множестве Ω_λ, так что и в этом случае в качестве x^k на ($k-1$)-ом шаге можно брать любую точку, доставляющую локальный максимум функции $\vartheta(x,\lambda)$ при фиксированном $\lambda=f(x^{k-1})$, - одновременно она будет точкой глобального максимума $\vartheta(x,\lambda)$.

Представляется весьма перспективным использовать для определения центральной точки области Ω_λ следующую функцию (см. [29])

$$\psi(x,\lambda)=\frac{1}{\lambda-f(x)}-\sum_{j=1}^{n}\frac{1}{g^j(x)} \tag{31}$$

*) К доказательству этого утверждения мы вернемся ниже.

Хотя эта функция не удовлетворяет требованиям, предъявляемым к функциям $\vartheta(x,\lambda)$, однако, как нетрудно показать, функция $-\frac{1}{\Psi(x,\lambda)}$ удовлетворяет этим требованиям. Так как точка $x^*(\lambda)$, в которой $-\frac{1}{\Psi(x,\lambda)}$ достигает максимума на Ω_λ при фиксированном λ, совпадает с точкой максимума функции $\Psi(x,\lambda)$ на Ω_λ, задача отыскания центра сводится к максимизации $\Psi(x,\lambda)$. При этом, начиная на K-том шаге из точки x^K, необходимо учитывать ограничение $f(x) \geq f(x^K)$.

Заметим, что при максимизации функции (31) с помощью какого-либо метода градиентного типа, остаются в силе все те соображения относительно сходимости градиентного процесса и относительно учета ограничений, которые были приведены при исследовании метода штрафов с функцией штрафа (22). Так что, сдвинувшись по направлению $\nabla f(x^K)$ из точки x^K внутрь Ω_λ, мы, по существу, оказываемся в той же ситуации, которая описана в § 5.

В заключение покажем квазивогнутость функции (30). Пусть $\eta^1(x), \eta^2(x), \dots, \eta^k(x)$ – вогнутые функции, заданные на открытом выпуклом множестве M, причем $\eta^j(x) > 0$ на M при всех j. Ввиду положительности η^j можем написать $\prod\limits_{k=1}^{} \eta^k(x) = e^{\sum\limits_{k=1} \ln \eta^k(x)}$

Функция $\ln \eta^k(x)$ вогнута, ибо для любых x', x'' на M на основании вогнутости η^k и неравенства $\frac{\eta^k(x') + \eta^k(x'')}{2} \geq \sqrt{\eta^k(x')\eta^k(x'')}$ имеет место $\ln \eta^k\left(\frac{x'+x''}{2}\right) \geq$

$\geq \ln\left(\frac{1}{2}\eta^k(x') + \frac{1}{2}\eta^k(x'')\right) \geq \frac{1}{2}\ln \eta^k(x') + \frac{1}{2}\ln \eta^k(x'')$.

Следовательно, вогнута и функция $y(x) = \sum\limits_{k=1} \ln \eta^k(x)$. Далее, $e^{y\left(\frac{x'+x''}{2}\right)} \geq e^{\frac{1}{2}y(x') + \frac{1}{2}y(x'')} \geq \min\left[e^{y(x')}, e^{y(x'')}\right]$, откуда сразу следует, что множество $\{x \in T : e^{y(x)} \geq c\}$ выпукло при любом c.

Квазивогнутость функции (30) на множестве Ω_λ получается из указанных рассуждений с учетом непрерывности $\vartheta(x,\lambda)$ на Ω_λ.

Заметим, что точка максимума функции (30) на Ω_λ совпадает с точкой максимума функции $\ln(f(x)-\lambda) + \sum\limits_{j=1}^{n} \ln g^j(x)$. Следовательно, если при реализации выбранного метода отыскания безусловного экстремума существенна вогнутость функции, вместо $\prod\limits_{j=1}^{n} g^j(x) \times (f(x)-\lambda)$ можно взять $\ln(f(x)-\lambda) + \sum\limits_{j=1}^{n} \ln g^j(x)$.

§ 8. Об отыскании исходной допустимой точки

При изложении методов второй, четвертой и пятой групп, а также метода штрафов со штрафными функциями (22) и (23) предполагалось, что заранее известна исходная точка $x^* \in \Omega$, которая в некоторых методах должна быть внутренней или относительно внутренней, в других – произвольной. В настоящем параграфе мы остановимся на двух способах отыскания таких точек.

В первом способе точка x^* получается при отыскании абсолютного максимума функции

$$F(x) = -\sum_{j=1}^{n} \{\max_{x} [0, -g^j(x)]\}^2.$$

Легко видеть, что в условиях исходной задачи функция F вогнута и дифференцируема, $F(x) = 0$ при $x \in \Omega$, $F(x) < 0$ при $x \notin \Omega$. Для отыскания максимума F может быть применен, например, один из методов градиентного типа с произвольным начальным $\bar{x}$.

Опишем теперь второй способ, предполагая, что множество Ω телесно. Вместе с условием (*) это означает, что в произвольной внутренней точке $\bar{x} \in \Omega$ должно быть $g^j(\bar{x}) > 0$, $j = 1, 2, \ldots, n$.

Пусть $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_m)$ – фиксированная точка, $\min_j g^j(\bar{x}) < 0$. Выберем $A \geq -\min g^j(\bar{x})$. Вспомогательная задача для отыскания внутренней точки множества Ω состоит в максимизации функции $\varphi(x) = x_{m+1}$ при ограничениях

$$g^j(x_1, x_2, \ldots, x_m) - x_{m+1} \geq 0, \; j = 1, 2, \ldots, n,$$

$$-A \leq x_{m+1} \leq A.$$

Добавление последнего условия позволяет добиться ограниченности допустимого множества вспомогательной задачи (в предположении, что Ω ограничено). В качестве исходной допустимой точки можно взять $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m, \bar{x}_{m+1})$, где $-A \leq \bar{x}_{m+1} \leq \min_j g^j(\bar{x})$.

Теперь для решения вспомогательной задачи может быть использован тот же метод, который предполагается применить затем при решении исходной задачи. Так как в точке максимума $\varphi(x) > 0$, положительное значение максимизируемой функции получается в конечное число шагов, этому значению будет соответствовать внутренняя точка множества Ω.

Если множество Ω не является телесным, вспомогательная задача для определения допустимой точки выглядит следующим образом:

Максимизировать

$$\varphi(x) = x_{m+1}$$

при ограничениях

$$g^j(x_1, x_2, \ldots, x_m) \geq 0, \quad j \in J_1;$$

$$g^j(x_1, x_2, \ldots, x_m) - x_{m+1} \geq 0, \quad j \in \{1, \ldots n\} \setminus J_1,$$

где множество индексов J_1 соответствует линейным ограничениям. В качестве исходного набора ($\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$) берется какое-либо решение линейной системы $g^j(x_1, x_2, \ldots, x_m) \geq 0$, $j \in J_1$, а $\bar{x}_{m+1}$ выбирается затем таким образом, чтобы удовлетворялись нелинейные ограничения. Ограниченность допустимого множества вспомогательной задачи достигается, как и в предыдущем случае. Модификация этого подхода для отыскания относительно внутренней точки описана в [13].

Сопоставляя приведенные способы, следует отметить, что, несмотря на сравнительную простоту, первый из них используется, видимо, значительно реже. Это связано с двумя обстоятельствами. Во-первых, при определении абсолютного максимума функции F обычно получается последовательность точек $\{x^k\} \notin \Omega$, так что для отыскания внутренней точки множества Ω приходится затем использовать какой-то дополнительный прием. Во-вторых, что более существенно (особенно с точки зрения реализации метода решения на ЭВМ), удобно, чтобы процесс решения вспомогательной и исходной задач был единообразным. Это достигается при втором способе и не имеет места при первом.

Литература

1. Л.М.Брэгман. Релаксационный метод отыскания общей точки выпуклых множеств и его применение к решению задач выпуклого программирования. Ж. вычисл.матем. и матем. физ., 1967, т.7, 620-631.
2. А.Вейнотт (A.Veinott). The supporting huperplane method for unimodal programming. Oper. Res., 1967, v.15, 147-152.

3. Ф.Вулф. Новые методы в нелинейном программировании. В сб. "Применение математики в экономических исследованиях", т.3, Мысль, 1965, 312-332.

4. И.И.Еремин. Релаксационный метод решения систем неравенств с выпуклыми функциями в левых частях. Докл. АН СССР, 1965, 160, № 5, 994-996.

5. И.И.Еремин. О методе штрафов в выпуклом программировании. В сб. "Математические методы в некоторых задачах оптимального планирования", (ред. И.И.Еремин), 1967, 43-51.

6. И.И.Еремин, Вл.Д.Мазуров. Итерационный метод решения задачи выпуклого программирования. Мат.зап. Уральского Госуниверситета, 1966, 5, № 3, 40-48.

7. Г.Зойтендейк. Методы возможных направлений, ИЛ, 1963.

8. Г.Зойтендейк (G.Zoutendijk). Nonlinear programming. A numerical survey. SIAM J. Control, 1966, v.4, 194-210.

9. .И.Зуховицкий, Р.А.Поляк, М.Е.Примак. Алгоритм для решения задачи выпуклого программирования. Докл. АН СССР, 1963, 153, № 5, 991-994.

10. А.А.Каплан. О некоторых методах решения задач нелинейного программирования. В сб. "Математические модели и методы оптимального планирования" (ред. Л.В.Канторович), Наука, 1966, 36-53.

11. А.А.Каплан. Некоторые свойства оператора покоординатного спуска. Оптимальное планирование, 1967, № 7, 25-31.

12. А.А.Каплан. Об отыскании экстремума линейной функции на выпуклом множестве. Докл. АН СССР, 1968, 178, №6, 1245-1247.

13. А.А.Каплан. К вопросу о реализации метода отсечения для решения задач выпуклого программирования. Оптимальное планирование, 1969, № 14, 43-49.

14. А.А.Каплан, Г.Ш.Рубинштейн. К теореме Куна-Таккера. Докл. АН СССР, 1969, 188, № 5, 993-996.

15. С.Карлин. Математические методы в теории игр, программировании и экономике. Мир, 1964.

16. Д.Келли (J.E.Kelley). The cutting plane method for solving convex programs. J. Soc. Ind. Appl. Math., 1960, 8, N. 4, 703-712.

17. Г.Кун, А.Таккер (H.Kuhn, A.Tucker). Nonlinear programming. Proc. 2-th Berkeley Symp. on Math. Stat. and Prob. (ed J. von Neumann) Univ. of California Press, 1951, 481-492.

18. Е.С.Левитин, Б.Т.Поляк. Методы минимизации при наличии ограничений. Ж. вычисл.матем. и матем. физ., 1966, 6, № 5, 787-823.

19. Вл.Д.Мазуров. Некоторые итерационные методы решения задач выпуклого программирования. В сб. "Математические методы в некоторых задачах оптимального планирования"(ред. И.И. Еремин), 1967, 53-72.

20. К.Миллер (C.Miller). The simplex method for local separable programming., In. "Recent advances in mathematical programming" (ed. R.Grave, P.Wolfe), 1963.

21. Б.Т.Поляк. Градиентные методы минимизации функционалов. Ж. вычисл. матем. и матем.физ., 1963, 3, № 4, 643-654.

22. Д. Розен (J.Rosen). The gradient projection Method for nonlinear programming. Part I. J. Soc. Ind. Appl. Math., 1960, 8, N 2, 181-217; Part II. J. Soc. Ind. Appl. Math., 1961, 9, N 3, 514-532.

23. Г.Ш.Рубинштейн. Несколько примеров двойственных экстремальных задач. В сб. "Математическое программирование" (ред. Л.В.Канторович). Наука, 1966, 9-59.

24. М.Слейтер (M.Slater). Lagrange multipliers revisited: A contribution to nonlinear programming. Cowles Comiss. Discuss. Paper, Math., 403, Nov. 1950.

25. Х.Удзава. Градиентный метод для вогнутого программирования. В сб. "К.Эрроу, Л.Гурвиц, Х.Удзава. Исследования по линейному и нелинейному программированию", ИЛ, 1962, 189-197.

26. Х.Удзава. Итерационные методы вогнутого программирования. В сб. "К.Эрроу, Л.Гурвиц, Х.Удзава. Исследования по линейному и нелинейному программированию". ИЛ, 1962, 228-245.

27. Х.Удзава. Теорема Куна-Таккера о вогнутом программировании. В сб. "К.Эрроу, Л.Гурвиц, Х.Удзава. Исследования по линейному и нелинейному программированию", ИЛ, 1962, 57-64.

28. А.Фиакко, Г.Мак Кормик (A.Fiacco, G.Mc Cormik). The sequential unconstrained minimization technique for nonlinear programming, a primal-dual method. Manag. Sci.,1964, 10, N 2, 360-366.

29. А.Фиакко, Г.Мак Кормик (A.Fiacco, G. Mc Cormik). The sequential unconstrained minimisation technique without parameters. J. Oper. Res., 1967, v. 15, 820-827.

30. Р.Фриш (R.Frisch). The logarithmic potential method for solving linear programming problems. Memorandum Univ. of Inst. of Econ., Oslo, 1955.

31. Д.Хедли. Нелинейное и динамическое программирование. Мир, 1967, 116-169.

32. К.Эрроу, Л.Гурвиц. Градиентный метод для вогнутого программирования. В сб. "К.Эрроу, Л.Гурвиц, Х.Удзава. Исследования по линейному и нелинейному программированию", ИЛ,1962, 175-188.

33. П.Юа (P.Huard) Resolution of mathematical programming with nonlinear constraints by the Method of centres. In "Nonlinear programming"(ed. J.Abodi), Amsterdam,1967, 207-209.

34. Д.Зангвилл (J.Zangwill). Convergence conditions for nonlinear programming algorithms. Man. Sci., 1969, 16, 1-13

Поступила в редакцию
5.IX. 1970 г.