УДК 513.25/26

МЕТОДЫ МИНИМИЗАЦИИ КВАЗИВЫПУКЛОЙ ФУНКЦИИ НА ВЫПУКЛОМ МНОГОГРАННИКЕ

Г.Ш.Рубинштейн, В.И.Шмырев

Излагаемые в этой работе методы решения задачи минимизации квазивыпуклой функции на выпуклом многограннике основаны на решении последовательности аналогичных задач для аффинных носителей граней исходного многогранника. Попытка построения методов такого рода была предпринята еще в работе [1]. Однако предложенный в этой работе комбинаторный алгоритм оказался малоэффективным, так как требует перебора слишком большого числа граней многогранника.

В рассматриваемых здесь методах, существенно использующих идеи двойственности из общей теории математического программирования, удаётся получить достаточно экономные схемы направленного перебора граней, близкие к разработанным для задач линейного и квадратичного программирования.

§ 1. Некоторые сведения о квазивыпуклых функциях и выпуклых многогранниках в R^n

Определенная на выпуклом множестве $X \subset R^n$ вещественнозначная функция f называется квазивыпуклой, если при любых $x, y \in X$ и $t \in (0,1)$

$$f(tx+(1-t)y) \le \max\{f(x), f(y)\}, \qquad (1)$$

причем в (1) имеет место строгое неравенство при $f(x) \ne f(y)$.

Если строгое неравенство в (I) имеет место при любых $x \neq y$, то функция f называется строго квазивыпуклой.

Нетрудно проверить, что функция f на выпуклом множестве X тогда и только тогда является квазивыпуклой, если

a) лебеговы множества этой функции

$$L_c = \{x \in X \mid f(x) \leq c\}$$

выпуклы при любом вещественном c;

b) каково бы ни было выпуклое подмножество X' выпуклого множества X, каждый локальный минимум функции f на X' является глобальным минимумом этой функции на X'.

Функция f на выпуклом множестве X называется квазивогнутой (строго квазивогнутой), если функция $-f$ на X является квазивыпуклой (соответственно строго квазивыпуклой).

Из сказанного ясно, что точки минимума квазивыпуклой функции (точки максимума квазивогнутой функции) на выпуклом множестве образуют выпуклое подмножество. При этом для строго квазивыпуклой и строго квазивогнутой функций указанные подмножества содержат не более одного элемента.

Перейдем теперь к определению и перечислению некоторых свойств выпуклых многогранников.

В каждом замкнутом выпуклом множестве $M \subset R^n$ можно ввести отношение эквивалентности, полагая $x \overset{M}{\sim} y$, если при некотором $t > 0$

$$x + t(x-y) \in M, \quad y + t(y-x) \in M.$$

Определяемые этим отношением классы эквивалентности в M называются открытыми гранями. Нетрудно проверить, что каждая открытая грань G является выпуклым множеством, открытым относительно своей аффинной оболочки L. При этом если $G \neq L$, то граница G в L представима в виде объединения открытых граней множества M, имеющих меньшую размерность, чем G.

Выпуклые многогранники в R^n можно определить как замкнутые выпуклые множества, имеющие конечное число открытых граней.

Из этого определения легко следует, что пересечение двух, следовательно, и любого конечного числа выпуклых многогранников будет выпуклым многогранником. Действительно, если M_1 и M_2 - выпуклые многогранники, то множество $M = M_1 \cap M_2$, очевидно, выпуклое и замкнутое. Далее, число открытых граней

в M не превосходит произведения соответствующих чисел для исходных многогранников M_1 и M_2, так как точки x и y тогда и только тогда принадлежат различным открытым граням выпуклого множества M, когда эти точки принадлежат различным открытым граням в M_1 или в M_2.

Простейшими примерами выпуклых многогранников являются непустые аффинные многообразия, имеющие по одной открытой грани, и так называемые полуаффинные многообразия (в частности, замкнутые полупространства), имеющие по две открытые грани.

Из сказанного выше следует, что каково бы ни было конечное число замкнутых полупространств в R^n, их пересечение будет выпуклым многогранником. Одним из важных результатов теории выпуклых многогранников является справедливость обратного утверждения: каждый выпуклый многогранник в R^n представим в виде пересечения конечного числа замкнутых полупространств, т.е. является совокупностью решений некоторой конечной системы линейных неравенств *).

Допустим теперь, что в R^n определена некоторая квазивыпуклая функция f и задан непустой выпуклый многогранник M.

Если функция f достигает минимума на M в некоторой точке x°, то эта точка, как нетрудно проверить, доставляет минимум функции f также на аффинном носителе L открытой грани $G \subset M$, содержащей точку x°.

Покажем, что и в том случае, когда функция f не достигает минимума на многограннике M, в последнем имеется такая открытая грань G, что для нее и её аффинного носителя L справедливо соотношение

$$\inf_{x\in G} f(x) = \inf_{x\in L} f(x) = \inf_{x\in M} f(x), \tag{2}$$

причем функция f не достигает минимума на L. Для доказательства этого факта, очевидно, достаточно убедиться в справедливости следующего вспомогательного утверждения.

ЛЕММА I. Каковы бы ни были определенная в R^n квазивыпуклая функция f и непустой выпуклый многогранник M, найдутся открытая

---

*) Во многих работах приведенное свойство выпуклых многогранников принимается в качестве их определения.

грань G многогранника M и $\varepsilon>0$ такие, что для этой грани и её аффинного носителя L справедливо соотношение (2) и, кроме того, эта грань содержит непустое выпуклое множество

$$P=\{y\in L \mid f(y)\leq \inf_{x\in M} f(x)+\varepsilon\}. \qquad (3)$$

ДОКАЗАТЕЛЬСТВО. Так как многогранник M представим в виде конечного объединения своих открытых граней, то среди них найдется такая открытая грань G, что

$$\inf_{x\in G} f(x)=\inf_{x\in M} f(x), \qquad (4)$$

но при некотором $\varepsilon>0$ для любой открытой грани G', имеющей меньшую размерность, чем G,

$$\inf_{x\in G'} f(x)>\inf_{x\in M} f(x)+\varepsilon . \qquad (5)$$

Аффинный носитель указанной открытой грани G обозначим через L и рассмотрим выпуклое множество (3). Ввиду (4) и (5), оно имеет общие точки с открытой гранью G и не пересекается с границей этой грани в L, так как последняя представима в виде объединения открытых граней G' меньшей размерности. Но тогда $P\subset G$ и лемма доказана.

Из приведенных фактов следует, что задача минимизации квазивыпуклой функции f на выпуклом многограннике $M\subset R^n$ может быть сведена к минимизации этой функции на аффинных носителях открытых граней многогранника M. Однако, как отмечалось во введении, полный перебор указанных аффинных многообразий уже в самых простых случаях практически не осуществим. Предлагаемые в этой работе методы решения рассматриваемой задачи реализуют различные схемы направленного перебора.

§ 2. Основная задача

В настоящей работе в качестве основной рассматривается следующая экстремальная задача.

Основная задача. Определить вектор

$$x=(x_1, x_2,\dots, x_n), \qquad (6)$$

доставляющий минимум квазивыпуклой функции f на выпуклом многограннике M, определяемом условиями:

$$x_j \geq 0, \quad j \in N^+, \tag{7}$$

$$\sum_{j \in N} x_j A^j = B, \tag{8}$$

где $N = \{1, 2, \ldots n\}$, N^+ – фиксированное подмножество множества N, а $A^j = (a_{1j}, a_{2j}, \ldots, a_{mj})$ и $B = (b_1, b_2, \ldots, b_m)$ – заданные m-мерные векторы.

Элементы выпуклого многогранника M, как это принято в математическом программировании, мы будем называть допустимыми векторами, а искомый допустимый вектор – оптимальным.

Каждый допустимый вектор (6) определяет разбиение множества N на два подмножества

$$J^0(x) = \{j \in N^+ \mid x_j = 0\}, \quad J(x) = N \setminus J^0(x). \tag{9}$$

При этом открытая грань многогранника M, содержащая допустимый вектор x^*, и её аффинный носитель допускают следующие описания:

$$G(J) = \{x \in M \mid J(x) = J\}, \tag{10}$$

$$L(J) = \{x \in R^n \mid \sum_{j \in N} x_j A^j = B, \ x_j = 0, \ j \in N \setminus J\}, \tag{11}$$

где $J = J(x^*)$. Однако ниже будут рассматриваться аффинные многообразия (11), отвечающие произвольным множествам J, удовлетворяющим соотношению

$$N \setminus N^+ \subset J \subset N. \tag{12}$$

При изложении численных методов решения поставленной задачи мы будем предполагать, что минимизируемая в ней квазивыпуклая функция f определена во всем R^n и удовлетворяет условиям:

А. Функция f непрерывно дифференцируема в R^n, причем её градиент

$$grad \ f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

обращается в нуль только в точках минимума f на R^n.

В. Каждое непустое аффинное многообразие $L \subset R^n$, на котором функция f не достигает минимума, содержит некоторый луч

$$\Lambda = \{x \in R^n \mid x = x^* + \lambda g, \ \lambda \geq 0\}, \tag{13}$$

на котором эта функция строго убывает;

c. Если функция f строго убывает на некотором луче (I3), то она строго убывает также на любом параллельном луче

$$\Lambda' = \{x \in R^n \mid x = x^{**} + \lambda q, \lambda > 0\} . \tag{I4}$$

Сделанные предположения обеспечивают справедливость следующих утверждений.

ТЕОРЕМА I. Допустимый вектор x^* в поставленной основной задаче тогда и только тогда является оптимальным, если при некотором векторе $y^* = (y_1^*, y_2^*, \dots, y_m^*)$ выполняются соотношения:

$$(A^j, y^*) = \frac{\partial f(x^*)}{\partial x_j}, \quad j \in J(x) , \tag{I5}$$

$$(A^j, y^*) < \frac{\partial f(x^*)}{\partial x_j}, \quad j \in J(x) , \tag{I6}$$

в которых через (A^j, y^*) обозначается скалярное произведение соответствующих векторов.

ТЕОРЕМА 2. Для поставленной основной задачи имеет место один из следующих трех взаимоисключающих случаев:

α) в задаче нет допустимых векторов, т.е. условия (7) и (8) несовместимы;

β) в задаче имеется оптимальный вектор;

γ) определяемый условиями (7) и (8) выпуклый многогранник M содержит некоторый луч (I3), на котором функция f строго убывает.

ДОКАЗАТЕЛЬСТВО теоремы I. В случае, когда вектор $x^* \in M$ доставляет минимум функции f на всем R^n, утверждения теоремы очевидно. При этом вектор x^* является оптимальным, и в качестве искомого может быть принят вектор $y^* = 0$.

Допустим теперь, что вектор $x^* \in M$ не доставляет минимума

функции f на R^n, т.е. выпуклое множество

$$Q=\{x\in R^n \mid f(x)<f(x^*)\} \tag{17}$$

не является пустым. Ввиду условия A в этом случае $grad\, f(x^*)\neq 0$, и линейное уравнение

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j < \sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j^* \tag{18}$$

задает единственную гиперплоскость, которая проходит через точку x° и выделяет множество Q. При этом последнее расположено в открытом полупространстве

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j < \sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j^\circ. \tag{19}$$

Если при некотором $y^*\in R^m$ выполняются соотношения (15) и (16), то тогда

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j^* = \sum_{j\in N}(A^j,y^*)x_j^\circ = \Big(\sum_{j\in N}x_j^* A^j, y\Big)=(B,y^*),$$

а при любом $x\in M$

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j \geqslant \sum_{j\in N}(A^j,y^\circ)x_j = \Big(\sum_{j\in N}x_j A^j,y^*\Big)=(B,y^\circ).$$

Это означает, что $M\cap Q=\varnothing$ и, следовательно, вектор x° является оптимальным.

Наоборот, если x^* – оптимальный вектор, то $M\cap Q=\varnothing$ и, следовательно, найдется гиперплоскость, разделяющая M и Q. Эта гиперплоскость, очевидно, проходит через точку x° и поэтому совпадает с гиперплоскостью (18). При этом множество Q, как отмечалось, расположено в полупространстве (19). Поэтому точки $x\in M$ удовлетворяют противоположному неравенству

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j \geqslant \sum_{j\in N}\frac{\partial f(x^\circ)}{\partial x_j}x_j^\circ.$$

Это означает, что вектор x° является оптимальным в задаче минимизации линейной функции

$$\sum_{j\in N}\frac{\partial f(x^*)}{\partial x_j}x_j$$

при ограничениях (7) и (8). Но тогда на основании признака

оптимальности для задач линейного программирования найдется вектор $y^0 \in R^m$, удовлетворяющий условиям (I5) и (I6), что и требовалось показать.

ЗАМЕЧАНИЕ. Попутно показано, что если определенная в R^n квазивыпуклая функция f удовлетворяет условию A, то при любых $x^0, x' \in R^n$ таких, что $f(x^0) > f(x')$, производная функция f в точке x^0 по направлению $g = x' - x^0$ меньше нуля.

Действительно, в рассматриваемом случае $grad\, f(x^0) \neq 0$ и точка x' принадлежит множеству Q, определяемому согласно (I7). Но тогда точка x' удовлетворяет неравенству (I9), т.е.

$$\frac{\partial f(x^0)}{\partial g} = \sum_{j \in N} \frac{\partial f(x^0)}{\partial x_j} (x_j' - x_j^0) < 0.$$

ДОКАЗАТЕЛЬСТВО теоремы 2. Требуется показать, что если многогранник M не является пустым и на нем функция f не достигает минимума, то имеет место случай γ. На основании леммы I (см. § I) при сделанных предположениях некоторая открытая грань G многогранника M и её аффинный носитель удовлетворяют соотношению (2). Кроме того, при некотором $\varepsilon > 0$ множество P, определяемое согласно (3), целиком содержится в G. Это означает, что при отсутствии оптимального вектора в рассматриваемой задаче функция f не может достигать минимума на L. Но тогда в силу условия B в L найдется луч (I3), на котором функция f строго убывает, и в качестве искомого может быть принят параллельный луч (I4) с вершиной в произвольной точке $x^{**} \in P$. Действительно, в силу условия C этот луч целиком содержится в $P \subset G \subset M$ и на нем функция f строго убывает.

§ 3. Метод последовательного улучшения допустимого вектора

Излагаемый здесь метод позволяет свести решение интересующей нас основной задачи к решению конечного числа задач, состоящих в минимизации исходной квазивыпуклой функции f на некоторых аффинных носителях открытых граней многогранника M. При этом, благодаря направленности перебора, в процессе, как правило, участвует лишь незначительная часть указанных граней. С

таких общих позиций рассматриваемый метод сходен с предложенным в [2]. Однако в этих методах используются различные идеи построения схемы направленного перебора открытых граней.

При изложении метода будет предполагаться, что для любого непустого аффинного многообразия $L \subset R^n$ мы умеем находить точку минимума функции f, если такая существует, или луч $\Lambda \subset L$, на котором функция f строго убывает, – в противном случае.

Для простоты изложения мы будем считать также выполненным следующее условие, которое, как и в задачах линейного программирования, гарантирует отсутствие ситуаций вырождения процесса.

Условие (1). Для любого допустимого вектора x среди m-мерных векторов

$$A^j,\ j \in J(x),$$

где множество $J(x)$ определяется согласно (9), имеется m линейно независимых.

Отметим, что для любого допустимого вектора x фигурирующая в теореме I система линейных уравнений (15) при выполнении условия (1) имеет, очевидно, не более одного решения.

Перейдем теперь к описанию интересующего нас метода.

Прежде всего необходимо выяснить, имеются ли в рассматриваемой задаче допустимые векторы, и, если $M \neq \varnothing$, найти некоторый допустимый вектор x^*. Для этого можно воспользоваться приёмами, развитыми в линейном программировании.

На K-м шаге процесса имеется допустимый вектор $x^K = (x_1^K, x_2^K, \ldots, x_n^K)$ и множество J_K такое, что

$$J(x^K) \subset J_K \subset N.$$

Для $K=1$ принимаем $x^1 = x^*$, $J_1 = J(x^*)$.

Рассматривается открытая грань $G(J_K)$ и её аффинный носитель $L(J_K)$, определяемые согласно (10) и (11).

В силу сделанного предположения для непустого аффинного многообразия $L(J_K)$, содержащего точку x^K, мы можем найти точку минимума функции f, если такая существует, или луч

$$\Lambda = \{x \in R^n \mid x = x^K + tg,\ t > 0\} \subset L(J_K), \qquad (20)$$

на котором функция f строго убывает. При этом мы будем различать три случая:

а) исходная точка x^K доставляет минимум функции f на $L(J_K)$;

b) найдена точка $\bar{x}^{\kappa}$, доставляющая минимум функции f на $L(\mathcal{J}_{\kappa})$, причем $f(\bar{x}^{\kappa}) < f(x^{\kappa})$;

c) найден луч (20), на котором функция f строго убывает.

Пусть имеет место первый из указанных случаев. Тогда находим единственное решение

$$y=(y_1, y_2, \dots, y_m)$$

системы линейных уравнений [ж)]

$$(A^j, y)=\frac{\partial f(x^{\kappa})}{\partial x_j}, \; j\in \mathcal{J}_{\kappa}, \tag{21}$$

и проверяем для него выполнение неравенств

$$(A^j, y)\leq \frac{\partial f(x^{\kappa})}{\partial x_j}, \quad j\in N\setminus \mathcal{J}_{\kappa}.$$

Если все эти неравенства выполняются, то на основании теоремы I имеющийся допустимый вектор x^{κ} является оптимальным (процесс окончен). В противном случае находим $j_0\in N\setminus \mathcal{J}_{\kappa}$ такое, что

$$z_{j_0}=\frac{\partial f(x^{\kappa})}{\partial x_{j_0}}-(A^{j_0}, y)<0 , \tag{22}$$

и, полагая $x^{\kappa+1}=x^{\kappa}$, $\mathcal{J}_{\kappa+1}=\mathcal{J}(x^{\kappa})\cup\{j_0\}$, переходим к следующему шагу.

Пусть имеет место случай b . Тогда, полагая $g=\bar{x}^{\kappa}-x^{\kappa}$, определим максимальное $\tau\in[0,1]$, при котором

$$x_j^{\kappa}+\tau g_j\geq 0, \quad j\in N^{+}\cap \mathcal{J}_{\kappa}. \tag{23}$$

Далее, принимаем $x^{\kappa+1}=x^{\kappa}+\tau g$, $\mathcal{J}_{\kappa+1}=\mathcal{J}(x^{\kappa+1})$ и переходим к следующему шагу. При этом если $\tau=1$, то на $(\kappa+1)$-м шаге, очевидно, будет иметь место случай a .

Пусть теперь имеет место случай c. Если при этом неравенства (23) выполняются при любом $\tau\geq 0$, то для рассматриваемой задачи, очевидно, реализуется случай γ из теоремы 2 и, следовательно, оптимального вектора не существует (процесс окончен). Если же это не так, то находим максимальное τ, при котором выполняются неравенства (23), и переходим к следующему шагу, принимая, как и в случае b $x^{\kappa+1}=x^{\kappa}+\tau g$, $\mathcal{J}_{\kappa+1}=\mathcal{J}(x^{\kappa+1})$.

ж) Эта система разрешима, так как x^{κ} является точкой минимума функции f на $L(\mathcal{J}_{\kappa})$. С другой стороны, в силу предположения (!), как отмечалось, даже некоторая подсистема рассматриваемой системы не может иметь более одного решения.

ТЕОРЕМА 3. При сделанных предположениях описанный процесс последовательного улучшения допустимого вектора конечен.

ДОКАЗАТЕЛЬСТВО. Если в рассматриваемой задаче нет допустимых векторов, то это устанавливается в предварительной части процесса. Поэтому достаточно рассмотреть случай, когда $M \neq \varnothing$.

Прежде всего заметим, что если на k-м шаге процесса имеет место случай b со значением $\tau < 1$ или случай c, не приводящий к окончанию процесса, то на следующем шаге процесса множество $J_{k+1} = J(x^{k+1})$ является собственным подмножеством множества J_k. Поэтому указанные ситуации в процессе могут повториться последовательно лишь конечное число раз. Это означает, что через определенное число шагов мы либо убедимся в отсутствии оптимального вектора в рассматриваемой задаче (будет найден луч $\Lambda \subset M$, на котором функция f строго убывает), либо же придем к случаю a.

С другой стороны, в случае a имеем

$$f(x^k) = \min_{x \in L(J_k)} f(x) ,$$

и, следовательно, величина $f(x^k)$ однозначно определяется множеством J_k.

Поэтому для доказательства конечности процесса достаточно показать, что если на каких-то двух шагах k и ℓ в процессе реализуется случай a, причем $k < \ell$, то $f(x^\ell) < f(x^k)$ и поэтому $J_k \neq J_\ell$.

Покажем прежде всего, что в указанной ситуации на $(k+1)$-м шаге процесса, на котором $x^{k+1} = x^k$, $J_{k+1} = J(x^k) \cup \{j_0\}$, будем иметь

$$\inf_{x \in L(J_{k+1})} f(x) < f(x^{k+1}) , \tag{24}$$

т.е. реализуется случай b или случай c. Для этого вычислим производную функции f в точке x^{k+1} по направлению h такому, что прямая $\{x^{k+1} + t h\}_{t \in (-\infty, +\infty)}$ содержится в $L(J_{k+1})$. Ясно, что последнее имеет место тогда и только тогда, если вектор

$$h = (h_1, h_2, \ldots, h_n) \tag{25}$$

удовлетворяет условиям:

$$h_j = 0, \quad j \in N \setminus J_{k+1},$$
$$\sum_{j \in N} h_j A^j = 0. \tag{26}$$

Из (21), (22) и (26) имеем:

$$\frac{\partial f(x^{k+1})}{\partial h} = \sum_{j \in N} \frac{\partial f(x^{k+1})}{\partial x_j} h_j = \frac{\partial f(x^{k+1})}{\partial x_{j_0}} h_{j_0} + \sum_{j \in J(x^k)} \frac{\partial f(x^{k+1})}{\partial x_j} h_j =$$

$$= h_{j_0}\left((A^{j_0}, y) + z_{j_0}\right) + \sum_{j \in J(x^k)} (A^j, y) h_j = h_{j_0} z_{j_0} + \left(\sum_{j \in N} h_j A^j, y\right) = h_{j_0} z_{j_0},$$

где $z_{j_0} < 0$. Далее, из $J(x^{k+1}) \subset J_{k+1}$ и условия (!) следует существование таких решений (25) системы (26), у которых $h_{j_0} > 0$. Но тогда имеет место неравенство (24) и, следовательно, на $(k+1)$ - м шаге процесса реализуется случай b или случай c.

Далее, в силу замечания к теореме I при реализации указанных случаев производная функции f в точке x^{k+1} по направлению g меньше нуля, т.е.

$$\frac{\partial f(x^{k+1})}{\partial g} = g_{j_0} z_{j_0} < 0,$$

и потому $g_{j_0} > 0$. Но тогда при достаточно малых $\tau > 0$ векторы $x = x^{k+1} + \tau g$ являются допустимыми, и для вектора x^{k+2} имеем:

$$f(x^{k+2}) < f(x^{k+1}).$$

Наконец, так как по условию $\ell > k$ и на ℓ -м шаге реализуется случай , то $\ell \geq k+2$ и, ввиду невозрастания функции f от шага к шагу, $f(x^\ell) \leq f(x^{k+2})$. Таким образом,

$$f(x^\ell) \leq f(x^{k+2}) < f(x^{k+1}) = f(x^k),$$

что и требовалось показать.

§ 4. Двойственная задача

Помимо методов описанного типа в математическом программировании систематически используются так называемые двойственные методы. Последние базируются на рассмотрении наряду с исследуемой задачей другой, тесно связанной с ней, экстремальной

задачи. Эту вспомогательную задачу принято называть двойственной, а исходную экстремальную задачу – прямой.

Для построения двойственных задач разработан ряд универсальных схем. Одной из таких схем (см.[3]) мы здесь и воспользуемся. При этом будем предполагать, что для исследуемой основной задачи, которую мы будем теперь называть прямой, имеет место неравенство

$$\inf_{x \in M} f(x) > \inf_{x \in R^n} f(x) \quad . \tag{27}$$

Это не исключает случая, когда $M \neq \phi$, так как инфимум функции на пустом множестве естественно считать равным $+\infty$.

Нас будут интересовать некоторые из замкнутых полупространств в R^n. Замкнутое полупространство $F \subset R^n$ будем считать элементом множества $\mathcal{F}_1$, если это полупространство содержит многогранник M, и элементом множества $\mathcal{F}_2$, если функция f на F достигает минимума, причем последний не совпадает с инфимумом f на R^n.

Заметим, что для любого замкнутого полупространства $F \in \mathcal{F}_2$ точки минимума функции f принадлежат граничной гиперплоскости этого полупространства, так как квазивыпуклая функция, определенная во всем R^n, не может иметь локальных минимумов, отличных от её глобального минимума. Далее, если точка $x^\circ \in R^n$ доставляет минимум функции f на полупространстве $F \in \mathcal{F}_2$, то последнее при сделанных предположениях относительно функции f (см. условие А в § 2) задается неравенством

$$\sum_{j \in N} \frac{\partial f(x^\circ)}{\partial x_j} x_j \geqslant \sum_{j \in N} \frac{\partial f(x^\circ)}{\partial x_j} x_j^\circ . \tag{28}$$

Наоборот, если замкнутое полупространство $F \subset R^n$ задается неравенством (28), то $F \in \mathcal{F}_2$ и функция f достигает минимума на F в точке x°.

Определим теперь на множестве $\mathcal{F}_2$ функцию Φ, полагая

$$\Phi(F) = \min_{x \in F} f(x), \tag{29}$$

и в качестве вспомогательной рассмотрим следующую экстремальную задачу.

ДВОЙСТВЕННАЯ ЗАДАЧА. В множестве $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$ допустимых полупространств найти оптимальное полупространство F, для которого достигает максимума величина (29).

Прежде всего отметим, что при любых допустимых элементах $x \in M$ и $F \in \mathcal{F}$ имеет место очевидное неравенство

$$f(x) \geqslant \Phi(F).$$

Поэтому для оптимальности допустимых элементов x° и $F_\circ$ достаточно, чтобы для них имело место равенство

$$f(x^\circ) = \Phi(F_\circ). \tag{30}$$

Более тесную связь между прямой и двойственной задачей устанавливает следующая

ТЕОРЕМА 4. Если в прямой и двойственной задачах существуют допустимые элементы, т.е. $M \neq \varnothing$ и $\mathcal{F} \neq \varnothing$, то в обеих задачах существуют оптимальные элементы x° и $F_\circ$, причем для них имеет место равенство (30). Если же по крайней мере в одной из рассматриваемых задач нет допустимых элементов[ж)], то ни в одной из этих задач нет оптимальных элементов.

ДОКАЗАТЕЛЬСТВО. В силу теоремы 2 (см. § I) для прямой задачи имеет место один из следующих трех взаимоисключающих случаев:

α) в задаче нет допустимых элементов, т.е. $M = \varnothing$;

β) в задаче имеется оптимальный элемент x°;

γ) выпуклый многогранник M содержит некоторый луч (13), на котором функция f строго убывает.

Предположим вначале, что имеет место случай β. Тогда по известной теореме отделимости найдется замкнутое полупространство $F_\circ \subset R^n$, содержащее многогранник M и не имеющее общих точек с непустым[жж)] открытым выпуклым множеством

$$Q = \{x \in R^n \mid f(x) < f(x^\circ)\}.$$

ж) Нетрудно проверить, что для рассматриваемых задач соотношения $M = \varnothing$ и $\mathcal{F} = \varnothing$ могут выполняться одновременно лишь в случае, когда $f(x) \equiv const$.

жж) Непустота этого множества следует из предположения (27).

Но тогда точка $x^\circ \in M \subset F_0$ доставляет минимум функции f на замкнутом полупространстве F_0. Следовательно, последнее является допустимым, и для него имеет место равенство (30). А это означает, что F_0 – оптимальный элемент в двойственной задаче. Таким образом, в рассматриваемом случае в обеих задачах существуют оптимальные элементы x° и F_0, и для них имеет место равенство (30).

Пусть теперь имеет место случай γ. При этом, ввиду предположений В и С функция f не достигает минимума ни на каком замкнутом выпуклом множестве, содержащем многогранник M. Следовательно, в двойственной задаче не существует допустимых элементов. Таким образом, в этом случае $\mathcal{F} = \phi$, и ни в одной из рассматриваемых задач нет оптимальных элементов.

Остается проверить, что и в случае α в двойственной задаче не существует оптимального элемента. Для этого достаточно заметить, что если $M = \phi$ и полупространство $F \in \mathcal{F}$, то любое замкнутое полупространство $F' \subset F$, не совпадающее с F, также является допустимым в двойственной задаче и для него имеет место неравенство

$$\Phi(F') > \Phi(F).$$

Это завершает доказательство теоремы.

Непосредственными следствиями доказанной теоремы являются теорема существования и признак оптимальности для рассматриваемых задач.

ТЕОРЕМА 5. *Следующие утверждения относительно прямой и двойственной задачи равносильны:*

1°. *В обеих задачах существуют допустимые элементы.*

2°. *В прямой задаче существует оптимальный элемент.*

3°. *В двойственной задаче существует оптимальный элемент.*

ТЕОРЕМА 6. *Для оптимальности допустимого элемента x° в прямой задаче (допустимого элемента F_0 в двойственной задаче) необходимо и до-*

статочно существование элемента $F_0 \in \mathcal{F}$ (соответственно элемента $x^* \in M$) такого, что имеет место равенство (30).

Отметим, что если квазивыпуклая функция f не является строго квазивыпуклой, то в прямой задаче нельзя гарантировать единственность оптимального вектора. Однако при сделанных предположениях относительно функции f в двойственной задаче существует не более одного оптимального полупространства. Последнее задается неравенством (28), где x^* - любой из оптимальных векторов прямой задачи.

В излагаемых ниже численных методах будет использоваться некоторая модификация рассмотренной двойственной задачи, к построению которой мы и переходим.

Каждому $(m+n)$-мерному вектору

$$W=(y_1, y_2, \dots, y_m, z_1, z_2, \dots, z_n) \qquad (31)$$

сопоставим в R^n замкнутое полупространство $F(W)$, определяемое неравенством

$$\sum_{j \in N} [(A^j, y) + z_j] x_j \geqslant (B, y), \qquad (32)$$

где вектор $y=(y_1, y_2, \dots, y_m)$ представляет из себя усечение соответствующего вектора W.

Нетрудно проверить, что если последние компоненты вектора (31) удовлетворяют условиям

$$z_j \geqslant 0, \quad j \in N^+, \quad z_j = 0, \quad j \in N \setminus N^+,$$

то отвечающее этому вектору полупространство $F(W)$ содержит многогранник M, т.е. принадлежит определенному выше множеству $\mathcal{F}_1$. Действительно, в рассматриваемом случае для любого вектора $x \in M$ имеем

$$\sum_{j \in N} [(A^j, y) + z_j] x_j = (\sum_{j \in N} x_j A^j, y) + \sum_{j \in N^+} z_j x_j \geqslant (B, y).$$

Множество указанных векторов w мы обозначим через W_1, а через W_2 обозначим множество векторов w, для которых $F(W) \in \mathcal{F}_2$. Тогда при любом векторе $w \in W_1 \cap W_2$ замкнутое полупространство $F(w)$, очевидно, является допустимым в двойственной задаче. С другой стороны, справедлива следующая

ЛЕММА 2. **Для любого полупространства $F \in \mathcal{F}$ найдется вектор $w \in W_1 \cap W_2$ та-**

кой, что $F(w)\subset F$.

ДОКАЗАТЕЛЬСТВО. Пусть полупространство $F\in\mathcal{F}$ задается неравенством

$$\sum_{j\in N} p_j x_j \geqslant p_0 \ .$$

Тогда это неравенство является следствием системы (7) и (8), определяющей многогранник M. Поэтому в силу известной теоремы Фаркаша найдется вектор $w\in W_1$ такой, что

$$p_0 \leqslant (B,y), \quad p_j = (A^j, y) + z_j\ , \quad j\in N.$$

Но тогда полупространство $F(w)$ содержится в исходном полупространстве F. Остается проверить, что $w\in W_2$. Если бы это было не так, то нашелся бы луч $\Lambda\subset F(w)\subset F$, на котором функция f строго убывает. Но это невозможно, так как функция f достигает минимума на F.

СЛЕДСТВИЕ 1. Каково бы ни было полупространство $F\in\mathcal{F}$, найдется вектор $w\in W_1\cap W_2$ такой, что $\Phi(F(w))\geqslant\Phi(F)$.

СЛЕДСТВИЕ 2. Если граничная гиперплоскость H полупространства $F\in\mathcal{F}$ имеет общие точки с многогранником M, то при некотором $w\in W_1\cap W_2$ полупространство $F(w)$ совпадает с F. В частности, оптимальное полупространство в двойственной задаче (если оно существует) представимо в виде $F=F(w)$, где $w\in W_1\cap W_2$.

Из приведенных фактов ясно, что рассматриваемая выше двойственная задача эквивалентна следующей.

МОДИФИЦИРОВАННАЯ ДВОЙСТВЕННАЯ ЗАДАЧА. В множестве $W=W_1\cap W_2$ допустимых векторов w найти оптимальный, для которого достигает максимума величина

$$\Psi(w) = \Phi(F(w)) = \min_{x\in F(w)} f(x) \ .$$

Заметим, что векторы w' и w'' из R^{m+n}, отличающиеся положительным множителем, определяют одно и то же замкнутое полупространство $F\subset R^n$. Поэтому если один из этих векторов допустимый, то допустимым является также второй и при этом

$\mathcal{Y}(w) = \mathcal{Y}(w')$. Однако из совпадения полупространств $F(w')$ и $F(w')$, вообще говоря, не следует коллинеарности векторов w' и w'. В частности, это относится и к оптимальному полупространству. Несмотря на отмеченную выше единственность последнего оно может определяться различными не коллинеарными векторами w' и w' из W. Таким образом, в модифицированной двойственной задаче нельзя гарантировать единственность оптимального вектора, если даже ограничиться рассмотрением нормированных допустимых векторов.

Допустим теперь, что $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ – оптимальный вектор в прямой задаче. Тогда единственное оптимальное полупространство F_0 в двойственной задаче, как отмечалось, задается неравенством (28). Следовательно, в модифицированной двойственной задаче найдется оптимальный вектор $w^* = (y_1^*, y_2^*, \dots, y_m^*, z_1^*, z_2^*, \dots, z_n^*)$ такой, что

$$(A^j, y^*) + z_j^* = \frac{\partial f(x^*)}{\partial x_j}, \quad j \in N, \tag{33}$$

$$\sum_{j \in N} \frac{\partial f(x^*)}{\partial x_j} x_j = (B, y^*). \tag{34}$$

Отсюда, учитывая допустимость векторов x^* и w^*, получаем

$$(B, y^*) = \sum_{j \in N} \frac{\partial f(x^*)}{\partial x_j} x_j^* = \sum_{j \in N} [(A^j, y^*) + z_j^*] x_j^* =$$

$$= \left(\sum_{j \in N} x_j^* A^j, y^*\right) + \sum_{j \in N^*} z_j^* x_j^* = (B, y^*) + \sum_{j \in N^*} z_j^* x_j^*,$$

т.е.

$$\sum_{j \in N^*} z_j^* x_j^* = 0 \, . \tag{35}$$

Наоборот, пусть для допустимых векторов x^* и w^* выполняются соотношения (33) и (35). Тогда

$$\sum_{j \in N} \frac{\partial f(x^*)}{\partial x_j} x_j^* = \sum_{j \in N} [(A^j, y^*) + z_j^*] x_j^* = (B, y^*) + \sum_{j \in N^*} z_j^* x_j^* = (B, y^*).$$

Поэтому неравенство

$$\sum_{j \in N} [(A^j, y^*) + z_j^*] x_j \geq (B, y^*),$$

задающее полупространство $F(w^*) \in \mathcal{F}$, может быть переписано в виде (28). Но тогда, как отмечалось в начале параграфа, точка x^* доставляет минимум функции f на $F(w)$ и,

следовательно,

$$f(x^*) = \psi(w^*).$$

А это означает, что допустимые векторы x^* и w^* являются оптимальными.

Из сказанного ясно, что в новых переменных признак оптимальности из теоремы 6 может быть переформулирован в виде, по существу совпадающем с теоремой I из § 2.

ТЕОРЕМА 7. *Для оптимальности допустимого вектора x^* в прямой задаче необходимо и достаточно существование допустимого вектора w^* в модифицированной двойственной задаче такого, что выполняются соотношения (33) и (35). При этом вектор w^* является оптимальным в модифицированной двойственной задаче.*

В заключение параграфа приведем еще одно вспомогательное утверждение, которое будет использоваться при доказательстве конечности излагаемых ниже методов.

ЛЕММА 3. *Если полупространства $F(w')$ и $F(w'')$, отвечающие векторам w', $w''\in W_2$, имеют непустое пересечение, то отрезок $[w', w'']$ целиком содержится в W_1 и функция f на этом отрезке является квазивогнутой. Если же $w'\in W_2$, а w'' таков, что соответствующее неравенство (32) несовместно, то луч*

$$\Lambda = \{w\in R^{n+m} \mid w = w' + t w'',\ t > 0\}$$

содержится в множестве W_2 и на этом луче функция ψ строго возрастает.

ДОКАЗАТЕЛЬСТВО. Пусть векторы w', $w''\in W_2$ таковы, что $F(w')\cap F(w'')\neq\varnothing$. Тогда при любом $t\in[0,1]$ полупространство $F(w)$, где $w = tw' + (1-t)w''$, очевидно, содержится в множестве $F(w')\cup F(w'')$. Поэтому

$$\inf_{x\in F(w)} f(x) \geq \inf_{x\in F(w')\cup F(w'')} f(x) = \min\{\varphi(w'),\ \varphi(w'')\} > \inf_{x\in R^n} f(x).$$

Покажем, что функция f достигает минимума на $F(w)$. Если бы это было не так, то при сделанных предположениях относительно функции f эта функция убывала бы на некотором луче $\Lambda \subset F(w)$, который принадлежит также одному из исходных полупространств $F(w')$, $F(w'')$. Но это невозможно при w', $w''\in W_2$. Следовательно, $w \in W_2$, и справедливо неравенство

$$\varphi(w) \geq \min\{\varphi(w'),\ \varphi(w'')\} .$$

Для доказательства квазивогнутости функции φ на $[w', w'']$ остается проверить, что при $\varphi(w') < \varphi(w'')$ и $t \in (0,1)$ имеет место строгое неравенство

$$\varphi(w) > \varphi(w'). \tag{36}$$

В указанном случае каждая точка $x \in F(w)$, как легко видеть, принадлежит множеству $F(w'')$ или множеству $F(w')\setminus H$, где через H обозначена граничная гиперплоскость полупространства $F(w')$. Если $x \in F(w'')$, то

$$f(x) \geq \varphi(w'') > \varphi(w').$$

Если же $x \in F(w)\setminus H$, то

$$f(x) > \varphi(w') , \tag{37}$$

так как точки минимума функции f на $F(w')$ принадлежат граничной гиперплоскости H. Таким образом, неравенство (37) имеет место для всех точек $x \in F(w)$ и, в частности, для точки минимума функции f на $F(w)$, т.е. справедливо неравенство (36).

Допустим теперь, что $w' \in W_2$, а w'' таково, что соответствующее неравенство (32) несовместно. Тогда при любых $t_2 > t_1 \geq 0$ полупространство $F(w' + t_2 w'')$, очевидно, является собственной частью полупространства $F(w' + t_1 w'')$, и эти полупространства содержатся в полупространстве $F(w')$. Но тогда $w' + t w'' \in W_2$, $w' + t w'' \in W_2$, и имеет место неравенство

$$\varphi(w' + t_2 w'') > \varphi(w' + t_1 w'') .$$

Это завершает доказательство леммы.

§ 5. Двойственный метод последовательного улучшения

В предыдущем параграфе было показано (см. теорему 4), что если в модифицированной двойственной задаче имеются допустимые векторы, то для рассматриваемых задач имеет место один из следующих двух случаев:

α) В прямой задаче нет допустимых векторов, и поэтому ни в одной из рассматриваемых задач нет оптимальных векторов;

β) В обеих задачах существуют оптимальные векторы.

В излагаемом здесь методе, исходя из некоторого вектора $w^0 \in W$, последовательно строятся векторы $w^1, w^2, \ldots$ из W такие, что $\psi(w^{s+1}) > \psi(w^s)$. При этом если для рассматриваемых прямой и модифицированной двойственной задач имеет место случай β, то на некотором шаге процесса определяются соответствующие оптимальные векторы. В противном случае на некотором шаге процесса выясняется, что для рассматриваемых задач имеет место случай α.

При изложении метода предполагается, что для любого аффинного многообразия L, на котором функция f достигает минимума, мы умеем находить точку, реализующую этот минимум.

Каждому вектору (3I) из W сопоставим множество

$$T(w) = \{j \in N \mid z_j = 0\}, \tag{38}$$

а каждому J, удовлетворяющему соотношению (I2), аффинное многообразие $L(J)$, определяемое согласно (II).

Заметим, что при сделанных предположениях относительно функции f эта функция достигает минимума на любом аффинном многообразии $L(T(w))$, где $w \in W$. Это связано с тем, что $L(T(w)) \in F(w)$.

Для простоты изложения мы будем дополнительно предполагать выполненным следующее условие, гарантирующее отсутствие вырождения процесса.

Условие (!!). Для любого вектора $w \in W$ функция f на $L(T(w))$ достигает минимума в единственной точке *).

---

*) Если функция f является строго квазивыпуклой, то выполнение этого условия очевидно. В случае линейной функции f это условие сводится к известному условию, обеспечивающему отсутствие вырождения в двойственных методах линейного программирования.

Перейдем теперь к изложению интересующего нас метода.

На κ -ом шаге процесса имеется допустимый вектор w^κ и множество J_κ такое, что

$$N \setminus N^+ \subset J_\kappa \subset T(w^\kappa).$$

Для $\kappa = 1$ принимаем $w^1 = w^*$, $J_1 = T(w^*)$, где w^* – исходный допустимый вектор модифицированной двойственной задачи.

Рассматривается система линейных уравнений

$$\sum_{j \in N} x_j A^j = B, \quad x_j = 0, \quad j \in N \setminus J_\kappa, \qquad (39)$$

определяющая аффинное многообразие $L(J_\kappa)$.

I. Если система (39) совместна, т.е. $L(J_\kappa) \neq \emptyset$, то функция f достигает минимума на $L(J_\kappa)$, так как $w^\kappa \in W$ и

$$L(J_\kappa) \subset F(w^\kappa).$$

Находим точку $\bar{x}^\kappa$, доставляющую минимум функции f на $L(J_\kappa)$, а также вектор

$$\bar{w}^\kappa = (\bar{y}_1^\kappa, \bar{y}_2^\kappa, \dots \bar{y}_m^\kappa, \bar{z}_1^\kappa, \bar{z}_2^\kappa, \dots, \bar{z}_n^\kappa) \qquad (40)$$

из условия

$$(A^j, \bar{y}^\kappa) + \bar{z}_j^\kappa = \frac{\partial f(\bar{x}^\kappa)}{\partial x_j}, \quad j \in N,$$

$$\bar{z}_j^\kappa = 0, \quad j \in J_\kappa.$$

Совместность системы (41) следует из того, что функция f в точке $\bar{x}^\kappa$ достигает минимума на $L(J_\kappa)$.

Из (39) и (41) для векторов $\bar{x}^\kappa$ и $\bar{w}^\kappa$ имеем:

$$\sum_{j \in N} [(A^j, \bar{y}^\kappa) + \bar{z}_j^\kappa] \bar{x}_j^\kappa = \Big(\sum_{j \in N} \bar{x}_j^\kappa A^j, \bar{y}^\kappa\Big) + \sum_{j \in N} \bar{z}_j^\kappa \bar{x}_j^\kappa = (B, \bar{y}^\kappa). \quad (42)$$

Следовательно, вектор $\bar{x}^\kappa \in L(J_\kappa) \subset F(w^\kappa)$ принадлежит также полупространству $F(\bar{w}^\kappa)$. Далее, ввиду (41) и (42), неравенство

$$\sum_{j \in N} [(A^j, \bar{y}^\kappa) + \bar{z}_j^\kappa] x_j \geq (B, \bar{y}^\kappa), \qquad (43)$$

определяющее полупространство $F(\bar{w}^\kappa)$, может быть переписано в виде

$$\sum_{j \in N} \frac{\partial f(\bar{x}^\kappa)}{\partial x_j} x_j \geq \sum_{j \in N} \frac{\partial f(\bar{x}^\kappa)}{\partial x_j} \bar{x}_j^\kappa.$$

Но тогда, как отмечалось в начале § 4, точка x^κ доставляет минимум функции f на $F(\bar{w}^\kappa)$, и потому $\bar{w}^\kappa \in W_2$ и

$$\varphi(\bar{w}^\kappa) = f(\bar{x}^\kappa) \geq \varphi(w^\kappa).$$

Для найденных векторов $\bar{w}^\kappa$ и $\bar{x}^\kappa$ проверяем неравенства:

$$\bar{z}_j^\kappa \geq 0, \quad j \in N \setminus J_\kappa, \tag{44}$$

$$\bar{x}_j^\kappa \geq 0, \quad j \in N^+ \cap J_\kappa. \tag{45}$$

$\bar{a}$) Если условия (44) и (45) выполняются, то $\bar{x}^\kappa \in M$, $\bar{w}^\kappa \in W$ и для этих векторов имеет место равенство $\varphi(\bar{w}^\kappa) = f(\bar{x}^\kappa)$. Следовательно, эти векторы являются оптимальными в прямой и модифицированной двойственной задачах (процесс окончен).

$\bar{b}$) Если условие (44) выполняется, но при некотором $j_0 \in N^+ \cap J_\kappa$ имеет место неравенство $\bar{x}_{j_0}^\kappa < 0$, то переходим к следующему шагу, полагая

$$w^{\kappa+1} = \bar{w}^\kappa, \qquad J_{\kappa+1} = T(\bar{w}^\kappa) \setminus \{j_0\}. \tag{46}$$

$\bar{c}$) Если по крайней мере одно из неравенств (44) нарушается, то находим максимальное $t \in [0,1]$, при котором

$$z_j^\kappa + t(\bar{z}_j^\kappa - z_j^\kappa) \geq 0, \quad j \in N \setminus J_\kappa, \tag{47}$$

и переходим к следующему шагу, принимая

$$w^{\kappa+1} = w^\kappa + t(\bar{w}^\kappa - w^\kappa), \quad J_{\kappa+1} = T(w^{\kappa+1}).$$

Допустимость вектора $w^{\kappa+1}$ следует из (47) и леммы 3 предыдущего параграфа.

II. Пусть теперь система (39) несовместна, т.е. $L(J_\kappa) = \varnothing$. Находим вектор $\bar{y}^\kappa = (\bar{y}_1^\kappa, \bar{y}_2^\kappa, \dots, \bar{y}_m^\kappa)$ из условий

$$(B, \bar{y}^\kappa) = 1, \qquad (A^j, \bar{y}^\kappa) = 0, \quad j \in J_\kappa. \tag{48}$$

Совместность системы (48) является следствием несовместности системы (39).

Далее, рассматриваем вектор (40), первые компоненты которого совпадают с компонентами вектора $\bar{y}_\kappa$, а последние определяются по формулам

$$\bar{z}_j^\kappa = -(A^j, \bar{y}^\kappa), \quad j \in N.$$

Отвечающее этому вектору неравенство (43), как легко видеть, несовместно. Поэтому в силу леммы 3 луч

$$\Lambda = \{w \in R^{n+m} \mid w = w^\kappa + t\bar{w}^\kappa, \; t \geq 0\}$$

содержится в множестве W_2 и функция φ на нем строго возрастает.

Для построенного вектора w^{κ} проверяем условия (44).

а) Если условия (44) выполняются, то $\Lambda \subset W$ и для многогранника M имеем:

$$M \subset \bigcap_{w \in \Lambda} F(w) = \emptyset ,$$

т.е. $M = \emptyset$. Это означает, что в рассматриваемом случае ни в прямой, ни в модифицированной двойственной задаче нет оптимальных векторов (процесс окончен).

b) Если по крайней мере одно из неравенств (44) нарушается, то находим максимальное $t > 0$, при котором

$$z_j^{\kappa} + t \bar{z}_j^{\kappa} \geq 0, \quad j \in N \setminus J_{\kappa} ,$$

и переходим к следующему шагу, принимая

$$w^{\kappa+1} = w^{\kappa} + t \bar{w}^{\kappa}, \quad J_{\kappa+1} = T(w^{\kappa+1}).$$

ЛЕММА 4. Если описанный процесс на $(\kappa+1)$ - м шаге не оканчивается, то для получаемого в конце этого шага вектора $w^{\kappa+2} \in W$ имеет место неравенство

$$\varphi(w^{\kappa+1}) < \varphi(w^{\kappa+2}). \tag{49}$$

ДОКАЗАТЕЛЬСТВО. Так как по условию процесс на $(\kappa+1)$ -ом шаге не оканчивается, то на этом шаге реализуется один из случаев Ib, Ic, IIb.

В случае Ib $w^{\kappa+2} = \bar{w}^{\kappa+1}$ и при этом

$$\varphi(\bar{w}^{\kappa+1}) = f(\bar{x}^{\kappa+1}) > \varphi(w^{\kappa+1}), \tag{50}$$

откуда следует (49).

В случае Ic $w^{\kappa+2} = w^{\kappa+1} + t(\bar{w}^{\kappa+1} - w^{\kappa+1})$, где $t \in (0,1)$, а $\bar{w}^{\kappa+1}$ удовлетворяет соотношению (50). Но тогда неравенство (49) следует из леммы 3.

Наконец, в случае IIb $w^{\kappa+2} = w^{\kappa+1} + t \bar{w}^{\kappa+1}$, где $t > 0$, а векторы $w' = w^{\kappa+1}$ и $w'' = \bar{w}^{\kappa+1}$ удовлетворяют условиям второй части леммы 3. Поэтому и в этом случае неравенство (49) следует из леммы 3.

ЛЕММА 5. Если в условиях предыдущей леммы на κ - м шаге процесса имел

место случай Iь, то для вектора $w^{\kappa+2} \in W$ выполняется строгое неравенство

$$\varphi(w^{\kappa+1}) < \varphi(w^{\kappa+2}) . \qquad (51)$$

ДОКАЗАТЕЛЬСТВО. Здесь в дополнение к предыдущей лемме требуется показать, что при реализации на $(\kappa+1)$ -ом шаге случаев Iь и Iб для вектора $\bar{w}^{\kappa+1}$ имеет место неравенство

$$\varphi(\bar{w}^{\kappa+1}) > \varphi(w^{\kappa+1}). \qquad (52)$$

Кроме того, если реализуется случай Iс или случай IIь, то определяемая в этих случаях неотрицательная величина t отлична от нуля.

Так как по условию на κ -ом шаге имел место случай Iь, то $w^{\kappa+1}$ и $J_{\kappa+1}$ определяются согласно (46), где $j_0 \in N^+ \cap J_\kappa$ таково, что $\bar{x}^\kappa_{j_0} < 0$. При этом

$$\varphi(w^{\kappa+1}) = \varphi(\bar{w}^\kappa) = \min_{x \in L(J_\kappa)} f(x) = f(\bar{x}^\kappa). \qquad (53)$$

Допустим, что $L(J_{\kappa+1}) \neq \varnothing$, т.е. на $(\kappa+1)$ -ом шаге реализуется случай Iь или Iб. Тогда

$$\varphi(\bar{w}^{\kappa+1}) = \min_{x \in L(J_\kappa)} f(x) = f(\bar{x}^{\kappa+1}). \qquad (54)$$

Кроме того, как отмечалось при изложении пункта I рассматриваемого метода, точка $x^\kappa \in L(J_\kappa) \subset L(T(\bar{w}^\kappa)) \subset F(\bar{w}^\kappa)$ доставляет минимум функции f на $F(\bar{w}^\kappa)$, следовательно, и на $L(T(\bar{w}^\kappa))$. Далее, точка $\bar{x}^{\kappa+1}$, доставляющая минимум функции f на $L(J_{\kappa+1})$, также принадлежит $L(T(\bar{w}^\kappa))$ и не совпадает с точкой $\bar{x}^\kappa$, расположенной вне аффинного многообразия $L(J_{\kappa+1})$. А тогда, учитывая, что $\bar{w}^\kappa \in W$, на основании условия (11) имеем

$$f(\bar{x}^{\kappa+1}) > f(\bar{x}^\kappa). \qquad (55)$$

Из (53), (54), (55) следует требуемое неравенство (52).

Покажем теперь, что при реализации на $(\kappa+1)$ -ом шаге случая Iб максимальное $t \in [0,1)$, при котором

$$z_j^{\kappa+1} + t(\bar{z}_j^{\kappa+1} - x_j^{\kappa+1}) \geq 0, \quad j \in N \setminus J_{\kappa+1} = (N \setminus T(w^{\kappa+1})) \cup \{j_0\},$$

отлично от нуля. Так как при этом

$$z_{j_0}^{\kappa+1} = 0, \quad z_j^{\kappa+1} > 0, \quad j \in N \setminus T(w^{\kappa+1}),$$
$$\bar{z}_j^{\kappa+1} = 0, \quad j \in T(w^{\kappa+1}) \setminus \{j_0\},$$

то достаточно проверить, что

$$\bar{z}_{j_0}^{\kappa+1} > 0. \tag{56}$$

Для этого вычислим производную функции f в точке $\bar{x}^{\kappa+1}$ по направлению $g = \bar{x}^{\kappa} - \bar{x}^{\kappa+1}$:

$$\frac{\partial f(\bar{x}^{\kappa+1})}{\partial g} = \sum_{j\in N} \frac{\partial f(\bar{x}^{\kappa+1})}{\partial x_j}(\bar{x}_j^{\kappa} - \bar{x}_j^{\kappa+1}) = \frac{\partial f(\bar{x}^{\kappa+1})}{\partial x_{j_0}}\bar{x}_{j_0}^{\kappa} +$$

$$+\sum_{j\in J_{\kappa+1}} \frac{\partial f(\bar{x}^{\kappa+1})}{\partial x_j}(\bar{x}_j^{\kappa} - \bar{x}_j^{\kappa+1}) = [(A^{j_0}, \bar{y}^{\kappa+1}) + \bar{z}_{j_0}^{\kappa+1}]\bar{x}_{j_0}^{\kappa+1} + \sum_{j\in J_{\kappa+1}}(A^j, \bar{y}^{\kappa+1})(\bar{x}_j^{\kappa} - \bar{x}_j^{\kappa+1}) =$$

$$= \bar{z}_{j_0}^{\kappa+1}\bar{x}_{j_0}^{\kappa} + \Big(\sum_{j\in T(w^{\kappa+1})} \bar{x}_j^{\kappa} A^j, \bar{y}^{\kappa+1}\Big) - \Big(\sum_{j\in J_{\kappa+1}} \bar{x}_j^{\kappa+1} A^j, \bar{y}^{\kappa+1}\Big) =$$

$$= \bar{z}_{j_0}^{\kappa+1}\bar{x}_{j_0}^{\kappa} + (B, \bar{y}^{\kappa+1}) - (B, \bar{y}^{\kappa+1}) = \bar{z}_{j_0}^{\kappa+1}\bar{x}_{j_0}^{\kappa+1}.$$

По замечанию к теореме I из § 2

$$\frac{\partial f(\bar{x}^{\kappa+1})}{\partial g} < 0 \quad ,$$

и так как $\bar{x}_{j_0}^{\kappa} < 0$, то имеет место (56).

Для завершения доказательства леммы остается проверить, что при реализации на $(\kappa+1)$ -ом шаге случая IIb максимальное $t \in [0, +\infty)$, при котором

$$z_j^{\kappa+1} + t\,\bar{z}_j^{\kappa+1} \geq 0, \quad j \in N \setminus J_{\kappa+1},$$

также отлично от нуля. Для этого, как и выше, достаточно показать, что для рассматриваемого здесь вектора $\bar{w}^{\kappa+1}$ справедливо неравенство (56).

В рассматриваемом случае

$$(B, \bar{y}^{\kappa+1}) = 1, \quad (A^j, \bar{y}^{\kappa+1}) = 0, \quad j \in J_{\kappa+1} \supset J_\kappa \setminus \{j_0\},$$

$$\bar{z}_j^{\kappa+1} = -(A^j, \bar{y}^{\kappa+1}), \quad j \in N.$$

Но тогда, так как точка $\bar{x}^{\kappa}$ принадлежит аффинному многообразию $L(J_\kappa)$, имеем:

$$1 = (B, \bar{y}^{\kappa+1}) = \Big(\sum_{j\in J_\kappa} \bar{x}_j^{\kappa} A^j, \bar{y}^{\kappa+1}\Big) = \sum_{j\in J_\kappa\setminus\{j_0\}} x_j^{\kappa}(A^j, \bar{y}^{\kappa+1}) +$$

$$+ x_{j_0}^{\kappa}(A^{j_0}, \bar{y}^{\kappa+1}) = -\bar{x}_{j_0}^{\kappa}\bar{z}_{j_0}^{\kappa+1},$$

откуда, ввиду $\bar{x}_{j_0}^{\kappa} < 0$, следует требуемое неравенство (56).

ЛЕММА 6. Если при $\kappa \neq \ell$ на κ -ом и ℓ -ом шагах процесса реализуется случай Ib, то $J_\kappa \neq J_\ell$.

ДОКАЗАТЕЛЬСТВО. Пусть $k<\ell$. Тогда в силу двух предыдущих лемм

$$\psi(w^{k+1})<\psi(w^{\ell+1}).$$

С другой стороны,

$$\psi(w^{k+1})=\min_{x\in L(J_k)} f(x),$$
$$\psi(w^{\ell+1})=\min_{x\in L(J_\ell)} f(x)\ .$$

Поэтому множества J_k и J_ℓ не совпадают, что и требовалось показать.

ТЕОРЕМА 9. Описанный двойственный метод последовательного улучшения при сделанных предположениях конечен.

ДОКАЗАТЕЛЬСТВО. Так как при реализации случаев Iб и IIб множество J_k заменяется более широким множеством J_{k+1} , то эти случаи в процессе могут повториться последовательно лишь конечное число раз (не превосходящее числа элементов в множестве N^+). Это означает, что при продолжении процесса через определенное число шагов должен встречаться случай Iв. Однако этот случай ввиду леммы 6, может реализовываться в процессе лишь конечное число раз, и теорема доказана.

§ 6. Комбинированные методы

В этом параграфе излагается схема комбинированного использования описанных выше прямого и двойственного методов последовательного улучшения. При этом на каждом шаге процесса имеются допустимый вектор прямой задачи и допустимый вектор модифицированной двойственной задачи, которые от шага к шагу улучшаются. Благодаря этому в ходе процесса мы имеем возможность судить о близости текущих допустимых векторов к оптимальным, точнее, о близости значений соответствующих функций на этих векторах к их значениям на оптимальных векторах. Кроме того, при использовании излагаемых здесь методов направленный перебор аффинных носителей граней исходного многогранника, как правило, существенно сокращается.

Для простоты изложения будем предполагать выполненными приведенные выше условия (I) и (II), обеспечивающие отсутствие ситуаций вырождения в прямом и двойственном методах.

Для начала процесса необходимо иметь векторы $x^* \in M$ и $w^* \in W$, удовлетворяющие условию

$$\mathcal{J}(x^*) \subset T(w^*), \tag{57}$$

где множества $\mathcal{J}(x^*)$ и $T(w^*)$ определяются согласно (9) и (38).

Наличие допустимых векторов в прямой и модифицированной двойственной задачах, как было показано в теореме 4, обеспечивает существование в этих задачах оптимальных векторов x° и w°, причем

$$f(x^\circ) = \mathcal{Y}(w^\circ).$$

На κ-ом шаге процесса помимо допустимых векторов x^κ и w^κ таких, что $\mathcal{J}(x^\kappa) \subset T(w^\kappa)$, имеется еще множество $\mathcal{J}_\kappa$, удовлетворяющее соотношению

$$\mathcal{J}(x^\kappa) \subset \mathcal{J}_\kappa \subset T(w^\kappa) . \tag{58}$$

На первом шаге принимается $x^1 = x^*$, $w^1 = w^*$, а в качестве $\mathcal{J}_1$, вообще говоря,- произвольное множество, удовлетворяющее соотношению (58) при $\kappa = 1$.

Заметим, что ввиду левой части соотношения (58) имеющаяся точка $x^\kappa \in M$ принадлежит аффинному многообразию $L(\mathcal{J}_\kappa)$, определяемому согласно (II). Далее, ввиду правой части указанного соотношения $L(\mathcal{J}_\kappa) \subset F(w^\kappa)$ и так как $w^\kappa \in W$, то функция f достигает минимума на $L(\mathcal{J}_\kappa)$.

Находим точку $\bar{x}^\kappa$, доставляющую минимум функции f на $L(\mathcal{J}_\kappa)$, и вектор (40) из условия (4I). При этом, как легко видеть,

$$\sum_{j \in N} \bar{\pi}_j^\kappa z_j^\kappa = 0 . \tag{59}$$

Для найденных векторов $\bar{x}^\kappa$ и $\bar{w}^\kappa$ проверяем выполнение неравенств (44) и (45). При этом возможны следующие случаи:

a) векторы $\bar{x}^\kappa$ и $\bar{w}^\kappa$ удовлетворяют условиям (44) и (45);

b) вектор $\bar{w}^\kappa$ удовлетворяет условию (44), но по крайней мере одно из неравенств (45) для вектора $\bar{x}^\kappa$ нарушается;

c) вектор $\bar{x}^\kappa$ удовлетворяет условию (45), но по крайней мере одно из неравенств (44) для вектора $\bar{w}^\kappa$ нарушается;

d) векторы $\bar{x}^k$ и $\bar{w}^k$ не удовлетворяют как условию (44), так и условию (45).

Если имеет место случай a, то $\bar{x}^k \in M$, $\bar{w}^k \in W$ и, ввиду (41) и (59), эти векторы, по теореме 7 являются оптимальными в прямой и модифицированной двойственной задачах (процесс окончен).

В случае b находим максимальное $\tau \in [0,1)$, при котором

$$x_j^k + \tau(\bar{x}_j^k - x_j^k) \geq 0, \quad j \in N^+ \cap J_k,$$

и переходим к следующему шагу, принимая

$$x^{k+1} = x^k + \tau(\bar{x}^k - x^k), \quad w^{k+1} = \bar{w}^k; \quad J_{k+1} = J(x^{k+1}).$$

Ясно, что при этом

$$x^{k+1} \in M, \quad w^{k+1} \in W, \qquad J(x^{k+1}) \subset J_{k+1} \subset T(w^{k+1}). \tag{60}$$

В случае c находим максимальное $t \in [0,1)$, удовлетворяющее условию (47), и переходим к следующему шагу, принимая

$$x^{k+1} = \bar{x}^k; \quad w^{k+1} = w^k + t(\bar{w}^k - w^k), \quad J_{k+1} = T(w^{k+1}).$$

Допустимость получаемого при этом вектора w^{k+1} была показана в предыдущем параграфе при описании двойственного метода последовательного улучшения, а выполнение первого и последнего из соотношений (60) очевидно.

Наконец, если имеет место случай d, то при этом возможны различные варианты реализации процесса.

Первый вариант. Поступаем, как и в случае b. Однако так как теперь уже $\bar{w}^k \bar{\in} W$, то в качестве w^{k+1} принимаем имеющийся допустимый вектор x^k.

Второй вариант. Поступаем, как и в случае c. Однако так как теперь уже $\bar{x}^k \in M$, то в качестве x^{k+1} принимаем имеющийся допустимый вектор x^k.

Третий вариант. Как и в случаях $\bar{b}$ и $\bar{c}$, находим величины $\tau, t \in [0,1)$ и переходим к следующему шагу, принимая

$$x^{k+1} = x^k + \tau(\bar{x}^k - x^k), \quad w^{k+1} = w^k + t(\bar{w}^k - w^k), \quad J_{k+1} = T(w^{k+1}) \setminus (J_k \setminus J(x^{k+1}))$$

В зависимости от используемого варианта в случае d описанная схема приводит к трем различным комбинированным методам. При использовании первого варианта получается метод, аналогичный прямому методу последовательного улучшения. Однако при нарушении условия (44) множество J_k пополняется уже не произвольным j_0, для которого $\bar{z}_{j_0}^k < 0$, а совокупностью таких j_0,

что

$$\frac{z_{j_0}^*}{z_j^* - \bar{z}_{j_0}^*} = \min_{\bar{z}_j^* < 0} \frac{z_j^*}{z_j^* - \bar{z}_j^*} .$$

При использовании второго варианта получаем метод, аналогичный двойственному методу последовательного улучшения. Отличие этого метода от описанного в предыдущем параграфе состоит в том, что при нарушении условия (45) из множества J_* исключается теперь не произвольное j_0, для которого $\bar{x}_{j_0}^* < 0$, а совокупность таких j_0, что

$$\frac{x_{j_0}^*}{x_{j_0}^* - \bar{x}_{j_0}^*} = \min_{\bar{x}_j^* < 0} \frac{x_j^*}{x_j^* - \bar{x}_j^*} .$$

Вместе с тем наиболее целесообразной, по-видимому, является реализация описанной схемы с использованием третьего варианта в случае d.

Следует отметить, что в линейном программировании непосредственных аналогов изложенных здесь комбинированных методов последовательного улучшения не существует. Это объясняется тем, что в случае задач линейного программирования из условия (57) для допустимых векторов x^* и w^* следует оптимальность этих векторов в прямой и модифицированной двойственной задачах.

§ 7. Некоторые дополнительные замечания

1^0. При изложении методов предполагались выполненными условия, обеспечивающие отсутствие вырождения процессов. Предварительная проверка этих условий трудно осуществима. С другой стороны, выполнение указанных условий гарантирует отсутствие вырождения лишь при идеально точных вычислениях. В связи с этим при разработке машинных алгоритмов описанные методы нуждаются в некоторой модификации, в принципе мало отличающейся от используемой при написании машинных программ в линейном программировании.

2^0. Эффективность реализации описанных методов во многом определяется тем, насколько хорошими алгоритмами мы располагаем для решения на каждом шаге задачи минимизации рассматриваемой квазивыпуклой функции на аффинном многообразии $L(J_*)$.

Указанная вспомогательная задача легко решается в случае, когда минимизируемая квазивыпуклая функция является квадратичной, и, следовательно, для этого класса задач рассмотренные методы допускают эффективные реализации. При этом прямой метод, как и метод из [2], приводит к конечному алгоритму, типа известных (см. [4], [5]). Двойственный и комбинированные методы приводят к новым алгоритмам, которым предполагается посвятить специальную статью.

В общем случае трудоемкость минимизации квазивыпуклой функции на аффинных многообразиях $L(J_\kappa)$ существенно зависит от размерности этих аффинных многообразий. Поэтому представляется целесообразным выделить классы задач, в которых можно априори указать достаточно хорошие верхние оценки для размерностей, встречающихся в процессе аффинных многообразий.

Простейший класс такого рода образует задачи, в которых минимизируемая функция зависит от небольшого числа аффинных функций. Например, пусть на многограннике $M \subset R^n$, определяемом условиями

$$x_j \geqslant 0, \quad j=1,2,\dots,n_1,$$
$$\sum_{j=1}^{n_1} a_{ij} x_j = b_i, \quad i=1,2,\dots,m_1,$$
$$\sum_{j=1}^{n_1} c_{sj} x_j + d_s = x_{n_1+s}, \quad s=1,2,\dots,z \; (z=m-m_1=n-n_1),$$

требуется минимизировать квазивыпуклую функцию

$$f(x_{n_1+1}, x_{n_1+2}, \dots, x_{n_1+z}).$$

Существенным для нас является то обстоятельство, что в этих задачах минимизируемая функция f зависит лишь от незначительной части общего числа переменных, причем на остальные переменные наложено условие неотрицательности. Таким образом, из множества изученных задач выделяются такие, в которых

$$\frac{\partial f(x)}{\partial x_j} \equiv 0, \quad j \in \dot{N} \setminus S \subset N^+, \qquad (61)$$

причем число элементов в множестве S, которое мы обозначим через z, сравнительно невелико.

Прежде всего заметим, что при построении исходного допустимого вектора x^* в прямом методе последовательного улучшения решается вспомогательная задача линейного программирования. Получаемое при этом множество $J_1 = J(x^*)$, которое определяется

согласно (9), обладает тем свойством, что векторы

$$A^j, \quad j \in J(x^*) \cap N^*,$$

линейно независимы.

Таким образом, в рассматриваемом случае на первом шаге процесса множество $J_1 = J(x^*)$ содержит не более $m+z$ элементов, и, следовательно, размерность аффинного многообразия $L(J_1)$ не превосходит z. Покажем, что эта верхняя оценка сохраняется в течение всего процесса. Для этого достаточно проверить, что на всех шагах процесса векторы

$$A^j, \quad j \in J_k \cap (N \setminus S) \tag{62}$$

остаются линейно независимыми. Линейная независимость этих векторов, очевидно, может нарушаться только при расширении множества J_k, т.е. при реализации на k-ом шаге процесса случая а. Допустим, что это произойдет. Тогда вектор A^{j_0}, отвечающий добавленному индексу $j_0 \in N \setminus S$, представим в виде

$$A^{j_0} = \sum_{j \in P} \lambda_j A^j, \tag{63}$$

где $P = (J_{k+1} \setminus \{j_0\}) \cap (N \setminus S) \subset J_k \cap (N \setminus S)$. При этом для получаемого на k-ом шаге вектора y на основании (21), (22), (61) имеем

$$(A^j, y) = \frac{\partial f(x^k)}{\partial x_j} = 0, \quad j \in P, \tag{64}$$

$$(A^{j_0}, y) > \frac{\partial f(x^k)}{\partial x_{j_0}} = 0. \tag{65}$$

Однако соотношения (64) и (65) противоречат допущению (63) и, следовательно, при реализации прямого метода размерность встречающихся в процессе аффинных многообразий $L(J_k)$ не превосходит z.

Нетрудно проверить, что при реализации двойственного и комбинированных методов векторы (62) также остаются линейно-независимыми в течение всего процесса. А это, как отмечалось, означает, что множества J_k содержат не более $m+z$ элементов. Однако если бы при реализации двойственного или комбинированного метода на некотором шаге процесса множество J_k содержало $m+z$ элементов, то из линейной независимости векторов (62) следовало бы, что $S \subset J_k$ и векторы $A^j (j \in J_k \setminus S)$ образуют базис в R^m. Но тогда при любых $\bar{x}_j$ $(j \in S)$ нашелся бы век-

тор $x \in L(J_k)$ такой, что $x_j = \bar{x}_j (j \in S)$. Следовательно, инфимум функции f на $L(J_k)$, и тем более на полупространстве $F(w^*) \supset L(J_k)$, совпадал бы с инфимумом f на R^n, что противоречит допустимости вектора w^*. Таким образом, при реализации двойственного и комбинированного методов встречающиеся в процессе множества J_k содержат не более $m+z-1$ элементов и размерность соответствующих аффинных многообразий не превосходит $z-1$.

В заключение этого пункта заметим, что для некоторых конкретных задач, удовлетворяющих условию (61), разыскание точки $\bar{x}^*$, доставляющей минимум функции f на аффинном многообразии $L(J_k)$ размерности $s \leq z-1$ удается свести к разысканию максимума некоторой квазивогнутой функции на аффинном многообразии размерности $z-s-1$. Это означает, что при $z \leq 4$ получается весьма эффективная реализация методов, требующая, в худшем случае, на каждом шаге процесса минимизации выпуклой функции или максимизации вогнутой функции одной вещественной переменной. Указанный прием, базирующийся на рассмотренных идеях двойственности, применительно к одной конкретной задаче описан в статье [6] настоящего сборника.

3^0. Примерами функций, определенных в R^n и удовлетворяющих требованиям А, В и С из § 2 могут служить функции

$$f_1(x) = \sum_{j \in S} e^{x_j},$$

$$f_2(x) = \sum_{j \in S} |x_j|^{\alpha},$$

где S – некоторое подмножество множества N, а $\alpha > 1$. Заметим, что вторая из этих функций достигает минимума на любом замкнутом множестве и, в частности, на любом аффинном многообразии. Поэтому при применении прямого метода последовательного улучшения к задаче минимизации функции f_2 на выпуклом многограннике M в процессе не встречается случай с (см. описание прямого метода в § 3), что несколько упрощает реализацию метода.

Интересно отметить, что для задач минимизации указанных функций на многограннике $M \subset R^n$, определяемом условиями

$$x_j \geq 0, \quad j \in N \setminus S, \tag{66}$$

$$\sum_{j \in N} x_j A^j = B, \tag{67}$$

удается в явном виде записать двойственные задачи, точнее, задачи, эквивалентные соответствующим модифицированным двойственным задачам.

Первая из этих задач сводится к максимизации функции

$$\varphi_1(y) = \frac{e^{(B,y)}}{\prod\limits_{j\in S} (A^j, y)^{(A^j,y)}}$$

на множестве векторов $y \in R^m$, удовлетворяющих неравенствам

$$(A^j, y) \leq 0, \quad j \in N \setminus S, \tag{68}$$

$$(A^j, y) > 0, \quad j \in S, \tag{69}$$

и условию нормировки

$$\sum_{j\in S} (A^j, y) = 1.$$

Вторая двойственная задача состоит в максимизации функции

$$\varphi_2(y) = \Big(\sum_{j\in S} |(A^j, y)|^{\alpha'}\Big)^{\frac{\alpha}{\alpha'}}, \quad \frac{1}{\alpha'} + \frac{1}{\alpha} = 1,$$

на множестве векторов $y \in R^m$, удовлетворяющих неравенствам (68) и условию нормировки

$$(B, y) = 1. \tag{70}$$

4^0. Важно отметить, что описанные методы в ряде случаев применимы без существенных изменений к задачам, в которых минимизируемая функция определена не на всем R^n, а лишь на некотором выпуклом множестве $G \subset R^n$. При этом на каждом шаге процесса необходимо минимизировать функцию f уже не на аффинном многообразии $L(J_k)$, а на пересечении этого аффинного многообразия с областью задания функции. Эта вспомогательная задача при произвольной выпуклой области $G \subset R^n$ является задачей квазивыпуклого программирования общего типа, и поэтому говорить об эффективных методах её решения не приходится.

В приводимых ниже примерах многогранник $M \subset R^n$ задается соотношениями (66) и (67), а минимизируемая функция f удовлетворяет условию (61) и определена в одной из следующих двух областей:

$$G_1 = \{x \in R^n \,|\, x_j > 0, \quad j \in S\},$$

$$G_2 = \{x \in R^n \,|\, x_j \geq 0, \quad j \in S\}.$$

ПРИМЕР I. Минимизировать функцию

$$f_3(x) = -\sum_{j\in S} \ln x_j$$

в области $M \cap G_1$.

ПРИМЕР 2. Минимизировать функцию

$$f_4 = \sum_{j\in S} x_j^{\alpha}, \quad \alpha < 0,$$

в области $M \cap G_1$.

ПРИМЕР 3. Минимизировать функцию

$$f_5 = -\sum_{j\in S} x_j^{\alpha}, \quad 0 < \alpha < 1,$$

в области $M \cap G_2$.

Для всех приведенных задач, как и в предыдущем пункте, удается в явном виде записать двойственные задачи. Приведем их формулировки в порядке нумерации примеров.

ЗАДАЧА I. Среди векторов $y \in R^m$, удовлетворяющих неравенствам (68), неравенствам

$$(A^j, y) < 0, \quad j \in S, \tag{71}$$

и условию нормировки

$$(B, y) = -\tau,$$

определить вектор, доставляющий максимум функции

$$\varphi_3(y) = \ln\left[\prod_{j\in S} (-A^j, y)\right].$$

ЗАДАЧА 2. Определить вектор $y \in R^m$, доставляющий максимум функции

$$\varphi_4(y) = \left[\sum_{j\in S} (-A^j, y)^{\alpha'}\right]^{-\frac{\alpha}{\alpha'}}, \quad \frac{1}{\alpha} + \frac{1}{\alpha'} = 1$$

при условиях (68), (71) и условии нормировки

$$(B, y) = -1. \tag{72}$$

ЗАДАЧА 3. Определить вектор $y \in R^m$, доставляющий максимум функции $\varphi_5(y) = -\varphi_4(y)$ при условиях (68), (71) и (72).

В первых двух примерах встречающиеся в процессе аффинные многообразия $L(J_k)$, имеющие общие точки с множеством G_1, обладают следующим свойством. Если минимизируемая функция на $L(J_k) \cap G_1$ не достигает минимума, то найдется луч $\Lambda \subset L(J_k) \cap G_1$,

на котором эта функция строго убывает. При этом рассматриваемая функция строго убывает также на любом параллельном луче с вершиной в G_1. Благодаря этому описанные методы для указанных задач реализуются без каких-либо изменений, если только позаботиться о том, чтобы на первом шаге аффинное многообразие $L(J_1)$ имело общие точки с G_1.

Аналогичное утверждение справедливо и для третьего примера, если многогранник M имеет общие точки с G_1. Если же $M \cap G_1 = \varnothing$, но $M \cap G_2 \neq \varnothing$, то прежде чем применять какой-либо из описанных методов, необходимо решить вспомогательную задачу, состоящую в определении максимального подмножества S' множества S, при котором множество $M \cap G_2$ пересекается с множеством

$$G_1' = \{x \in R^n \mid x_j > 0, \quad j \in S'\} .$$

После этого, полагая

$$x_j = 0, \qquad j \in S \setminus S',$$

приходим к рассмотренному выше случаю. Заметим, что приведенная вспомогательная задача легко решается с помощью методов линейного программирования.

Литература

1. Theil H. and Vande. Quadratic programming as an extension of conventional quadratic maximization. Man. Sci. 7(1961).

2. В.А.Булавский, Г.Ш.Рубинштейн. О решении задач выпуклого программирования с линейными ограничениями методом последовательного улучшения допустимого вектора. ДАН СССР, т.150, № 2 (1963), 231-234.

3. Г.Ш.Рубинштейн. Несколько примеров двойственных экстремальных задач. Сб. "Математическое программирование", М. 1966, 9-39.

4. Beale E.M.L. On quadratic programming. Nav. Res. Log. Qu.6 (1959).

5. Ph.Wolfe. The simplex method for quadratic programming, Econometrica 27, No. 3 (1959).

6. В.И.Шмырёв. Об эффективных алгоритмах для одного класса задач нелинейного программирования. Наст. сб., стр. 118-133.

Поступила в редакцию
15.II. 1971 г.