В. Г. Жадан, Метод модифицированной функции Лагранжа для задач многокритериальной оптимизации, *Ж. вычисл. матем. и матем. физ.*, 1988, том 28, номер 11, 1603–1618

ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 28, 1988 № 11

УДК 519.85

МЕТОД МОДИФИЦИРОВАННОЙ ФУНКЦИИ ЛАГРАНЖА ДЛЯ ЗАДАЧ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

ЖАДАН В. Г.

(Москва)

Для задачи многокритериальной минимизации с ограничениями вводится функция Лагранжа, содержащая оценки эффективных значений критериев. На ее основе строится метод решения задачи, являющийся обобщением известного в нелинейном программировании метода модифицированной функции Лагранжа. Дается обоснование метода, и описываются ситуации, в которых применение метода может оказаться полезным при решении задач многокритериальной минимизации.

§ 1. Введение

При решении задач многокритериальной оптимизации (м.к.о.) с выпуклыми целевыми функциями один из основных подходов состоит в использовании линейной функции для скаляризации векторного критерия. Однако такой прием совершенно неприемлем в случае невыпуклых критериев. Ниже для решения общей невыпуклой задачи м.к.о. с ограничениями предлагается использовать модифицированную функцию Лагранжа (м.ф.Л.), в которой к обычной функции Лагранжа, осуществляющей линейную свертку критериев и ограничений, добавляются дополнительные члены, являющиеся штрафами за нарушение условий оптимальности.

М.ф.Л. и методы, основанные на их использовании, широко применяются в нелинейном программировании (н.л.п.) (см., например, [1]–[6]). Для задач м.к.о. они рассматривались в [7]. В предлагаемой ниже м.ф.Л., в отличие от [7], улучшение поведения функции Лагранжа в окрестности решения осуществляется не только за счет ограничений, но и за счет критериев. Это потребовало введения в функцию Лагранжа, помимо двойственных переменных, дополнительно оценок эффективных значений критериев. Такое расширение функции Лагранжа позволяет, в частности, получить условия оптимальности в удобной симметричной форме. Предлагается метод решения задачи м.к.о., в котором за один просчет отыскивается одно эффективное решение. Для точного указания того эффективного решения, которое желательно получить, в методе, аналогично [8], предусмотрена возможность «пронизывать» пространство критериев из заданной целевой точки лучами и двигаться вдоль них.

Рассмотрим задачу м.к.о., которую условно обозначим

$$(1.1)\qquad \min_{x\in X} f(x),\qquad X=\{x\in R^n\,|\,g(x)\leqslant 0\},$$

где R^n есть n-мерное евклидово пространство, $f(x)$ и $g(x)$ – непрерывные вектор-функции, осуществляющие, соответственно, отображения $f\colon R^n\to R^r$, $g\colon R^n\to R^m$, $r\geqslant 1$, $m\geqslant 0$.

Под решением задачи (1.1) понимается множество оптимальных по Слейтеру точек в пространстве R^n, определяемое следующим образом:

$$(1.2)\qquad X_*=\{x_*\in X\,|\,\max_{1\leqslant i\leqslant r}[f^i(x)-f^i(x_*)]\geqslant 0\ \ \forall x\in X\}.$$

Всюду ниже предполагается, что множество X_* не пусто и внешне устойчиво.

Помимо X_* определим также множество строго оптимальных решений:

$$(1.3)\qquad X_*^c=\{x_*\in X\,|\,\max_{1\le i\le r}[f^i(x)-f^i(x_*)]>0\ \ \forall x\in X,\ x\ne x_*\}.$$

Если неравенство в определении (1.3) выполняется только для x из некоторой окрестности точки x_*, то x_* называется строгим локально-оптимальным решением.

Ниже символы $\operatorname{cl} B$, $\operatorname{int} B$ и $\operatorname{fr} B$ обозначают, соответственно, замыкание множества B, его внутренность и границу. Символ $f(B)$ используется для обозначения образа множества $B\subseteq R^n$ при отображении f. Множество $F=f(X)$ является множеством достижимых оценок, множества $F_*=f(X_*)$ и $F_*^c=f(X_*^c)$ — соответственно, множествами оптимальных и строго оптимальных оценок.

§ 2. Условия оптимальности

Пусть вектор $y\in R^r$ и $l=r+m$ — общее количество критериев и ограничений в задаче (1.1). Пусть, кроме того, R_+^i — неотрицательный ортант R^i, R_-^i — неположительный ортант R^i, $F_+=F+R_+^r$, S^r — симплекс в пространстве R^r, определяемый следующим образом:

$$S^r=\left\{z\in R_+^r\,\middle|\,\sum_{i=1}^{r} z^i=1\right\}.$$

Через $h(x, y)$ обозначим r-мерную вектор-функцию, компонентами которой являются функции $h^i(x, y)=f^i(x)-y^i$, $1\le i\le r$.

Составим для задачи (1.1) функцию Лагранжа:

$$(2.1)\qquad L(x, u, v, y)=\langle u, h(x, y)\rangle+\langle v, g(x)\rangle,$$

где $u\in S^r$, $v\in R_+^m$, угловые скобки обозначают скалярное произведение. Используя функцию (2.1), находим условия оптимальности для задачи (1.1). С этой целью рассмотрим задачу отыскания минимакса функции (2.1) при фиксированном векторе $y\in R^r$:

$$(2.2)\qquad L^*(y)=\min_{x\in R^n}\max_{u\in S^r}\max_{v\in R_+^m} L(x,u,v,y).$$

Обозначим

$$\varphi(x,y)=\max_{u\in S^r}\langle u,h(x,y)\rangle,\qquad \psi(x)=\sup_{v\in R_+^m}\langle v,g(x)\rangle,$$

$$\omega(x, y)=\varphi(x, y)+\psi(x),$$

$$\Omega(x, y)=\{[u, v]\in S^r\times R_+^m\,|\,L(x, u, v, y)=\omega(x, y)\}.$$

Точка $[x_*, u_*, v_*]\in R^{n+l}$ является решением задачи (2.2), если $[u_*, v_*]\in\Omega(x_*, y_*)$ и, кроме того,

$$(2.3)\qquad x_*\in\operatorname{Arg}\min_{x\in R^n}\omega(x,y).$$

Задачу нахождения множества $\Omega(x, y)$ и построения функции $\omega(x, y)$ назовем внутренней подзадачей задачи (2.2), задачу отыскания точки x_*, удовлетворяющей условию (2.3),— внешней подзадачей задачи (2.2). Заметим, что $\omega(x, y)=+\infty$, если $x\notin X$.

Лемма 1. *Пусть точка* $[x_*, u_*, v_*]$ *является решением задачи* (2.2) *при некотором* $y=y_*\in R^r$. *Тогда равенство*

$$L(x_*, u_*, v_*, y_*)=0 \tag{2.4}$$

возможно в том и только том случае, когда $y_*\in \mathrm{fr}\, F_+$.

Доказательство. Предположим сначала, что $y_*\in \mathrm{fr}\, F_+$. Из данного условия вытекает, что для любого $x\in X$ выполнено неравенство $\varphi(x, y_*)\geqslant 0$. Кроме того, поскольку всегда $\psi(x)\geqslant 0$, причем $\psi(x)=+\infty$, когда $x\notin X$, то $\omega(x, y_*)\geqslant 0$ для всех $x\in R^n$. Убедимся теперь, что найдется точка $x_0\in X$, для которой $\omega(x_0, y_*)=0$. Так как, по предположению, множество X_* не пусто и внешне устойчиво, то для любой точки $x\in X\backslash X_*$ можно указать точку $x_1\in X_*$ такую, что $f(x_1)<f(x)$. Отсюда вытекает, что $F_+=F_*+R_+{}^r$. Кроме того, в силу непрерывности функции $f(x)$ и замкнутости множества X, множество F_* замкнуто. Поэтому $y_*\in F_+$ и, следовательно, найдется точка $x_0\in X$ такая, что $f(x_0)\leqslant y_*$. В этом случае обязательно $\omega(x_0, y_*)=\varphi(x_0, y_*)\leqslant 0$. Но для всех $x\in R^n$, как было установлено, $\omega(x, y_*)\geqslant 0$, поэтому $\omega(x_0, y_*)=0$. Сопоставляя данное равенство с предыдущим неравенством, приходим к выводу, что имеет место (2.4). Таким образом, необходимость условия (2.4) доказана.

Покажем теперь достаточность условия (2.4) для того, чтобы $y_*\in \mathrm{fr}\, F_+$. Предположим, что это не так. Тогда либо $y_*\notin F_+$, либо $y_*\in \mathrm{int}\, F_+$. В первом случае $\omega(x, y_*)=\varphi(x, y_*)>0$ для любого $x\in X$ и, стало быть, равенство (2.4) ни при одном $x\in R^n$ не может иметь места. Во втором случае в силу того, что множество F_+ представимо в виде $F_+=F_*+R_+{}^r$, всегда можно указать точку $x'\in X_*$ такую, что $f(x')<y_*$. Поэтому $\omega(x', y_*)=\varphi(x', y_*)<0$. Отсюда вытекает, что $L(x_*, u_*, v_*, y_*)\leqslant\omega(x', y_*)<0$, и, следовательно, равенство (2.4) снова нарушено. Лемма доказана.

Теорема 1. *Для того чтобы точка* $x_*\in R^n$ *была решением задачи* (1.1), *необходимо и достаточно, чтобы существовали* $u_*\in S^r$, $v_*\in R_+{}^m$ *и* $y_*\in R^r$ *такие, что точка* $[x_*, u_*, v_*]$ *является решением задачи* (2.2) *при* $y=y_*$ *и имеет место равенство* (2.4).

Доказательство. Достаточность. Точка x_* допустима, поскольку она является решением внешней подзадачи задачи (2.2). Осталось показать, что $x_*\in X_*$. Если $x_*\notin X_*$, то найдется $x'\in X$ такое, что $f(x')<f(x_*)$. Так как выполнено (2.4), то, в силу утверждения леммы 1, обязательно $y_*\in \mathrm{fr}\, F_+$ и $f(x_*)\leqslant y_*$. Поэтому $f(x')<y_*$. Отсюда следует, что $\omega(x', y_*)=\varphi(x', y_*)<\varphi(x_*, y_*)=\omega(x_*, y_*)=0$ и, значит, точка x_* не является решением внешней подзадачи задачи (2.4). Полученное противоречие доказывает достаточность.

Необходимость. Возьмем $y_*=f(x_*)$. Тогда $h(x_*, y_*)=0$ и $\omega(x_*, y_*)=0$, причем существуют такие векторы $u_*\in S^r$ и $v_*\in R_+{}^m$, что $L(x_*, u_*, v_*, y_*)=\omega(x_*, y_*)$. Покажем, что $\omega(x, y_*)\geqslant 0$ для любых других $x\in R^n$. В самом деле, если $x\notin X$, то $\omega(x, y_*)=+\infty$. В случае же, когда $x\in X$, в силу определения (1.2), $\varphi(x, y_*)\geqslant 0$ и, следовательно, $\omega(x, y_*)\geqslant 0$. Таким образом, точка $[x_*, u_*, v_*]$ является решением задачи (2.2) при $y=y_*$ и имеет место (2.4). Теорема доказана.

Замечание. Из доказательства теоремы видно, что если $x_*\in X_*$, то в качестве вектора y_*, фигурирующего в формулировке необходимых условий, можно взять $y_*=f_*=f(x_*)$. Но результат теоремы в этой части не изменится, если в качестве y_* взять любой вектор из множества $Y(x_*)=(f_*+R_+{}^r)\cap \mathrm{fr}\, F_+$. В самом деле, пусть $y_*\in Y(x_*)$ и $y_*\neq f_*$. Тогда $L(x, u, v, y_*)\leqslant L(x, u, v, f_*)$ для любых $x\in R^n$, $u\in S^r$ и $v\in R_+{}^m$.

Отсюда следует, что $\omega(x, y_*) \leqslant \omega(x, f_*)$. Кроме того, поскольку $f_*^i = y_*^i$ хотя бы для одного индекса i, то $\omega(x_*, y_*) = \omega(x_*, f_*)$. Таким образом, если x_* не есть точка минимума функции $\omega(x, y_*)$, то можно указать точку $x' \in X$ такую, что $\omega(x', y_*) < 0$. Но тогда $f(x') < y_*$. Поэтому $y_* \notin \mathrm{fr}\, F_+$, что противоречит определению множества $Y(x_*)$.

Теорема 1 — это, по-существу, условия оптимальности Ю. Б. Гермейера, выраженные посредством функции Лагранжа (2.1). Если $r=1$, то (1.1) — задача н.л.п. Для нее множество S^1 допустимых значений вектора u вырождается в единственную точку, так что внутренняя подзадача максимизации по этой переменной в (2.2) пропадает. Уровень y перестает влиять как на решение внутренней, так и на решение внешних подзадач (2.2), и его можно опустить. Утверждение теоремы в данном случае переходит в хорошо известное свойство эквивалентности задачи н.л.п. задаче отыскания минимакса функции Лагранжа (см. [4]).

Предполагая далее, что все функции $f^i(x)$ и $g^j(x)$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant m$, непрерывно дифференцируемы, получаем необходимые условия оптимальности первого порядка с использованием функции (2.1). Пусть $f_x(x)$ и $g_x(x)$ — матрицы первых производных $f(x)$ и $g(x)$ размером, соответственно, $r \times n$ и $m \times n$; L_x, L_u и L_v — градиенты функции L, вычисленные, соответственно, по переменным x, u и v. Определим множества

$$J_1(x, y) = \{1 \leqslant i \leqslant r \mid f^i(x) = y^i\}, \qquad J_2(x) = \{1 \leqslant j \leqslant m \mid g^j(x) = 0\}.$$

Для произвольной точки $x_* \in X_*$ рассмотрим линейную задачу м.к.о.: найти

$$\min_{d \in D_1(x_*)} f_x(x_*) d, \tag{2.5}$$

$$D_1(x_*) = \{d \in R^n \mid \langle g_x^j(x_*), d \rangle \leqslant 0,\ j \in J_2(x_*)\}.$$

Множеством ее достижимых оценок является образ множества $D_1(x_*)$ при отображении $f_x(x_*) d$. Обозначим его $F^1(x_*)$. Кроме того, положим $F_+^1(x_*) = F^1(x_*) + R_+^r$, $Y^1(x_*) = f(x_*) + [R_+^r \cap \mathrm{fr}\, F_+^1(x_*)]$. Символом $P_B(b)$ будем обозначать проекцию вектора b на множество B.

Теорема 2. *Пусть $x_* \in X_*$ и в задаче* (2.5) *ограничения удовлетворяют условию Слейтера. Тогда найдутся $u_* \in S^r$ и $v_* \in R_+^m$ такие, что для любого $y_* \in Y^1(x_*)$ в точке $z_* = [x_*, u_*, v_*, y_*]$ имеют место равенства*

$$L_x(z_*) = 0, \tag{2.6}$$

$$P_{S^r}(c_1 L_u(z_*) + u_*) = u_*, \qquad P_{R_+^m}(c_2 L_v(z_*) + v_*) = v_*, \tag{2.7}$$

$$L(z_*) = 0, \tag{2.8}$$

где c_1 и c_2 — произвольные положительные константы.

Доказательство проводится примерно по той же схеме, что и доказательство теоремы Да Канха — Полака из [9].

В дальнейшем точки z_*, в которых выполнено (2.6) — (2.8), называются точками Куна — Таккера.

Сформулируем с помощью функции (2.1) достаточные условия оптимальности второго порядка. Предположим, что функции $f^i(x)$ и $g^j(x)$, $1 \leqslant i \leqslant r$, $1 \leqslant j \leqslant m$, дважды непрерывно дифференцируемы на R^n и L_{xx} — матрица вторых производных функции L по первому аргументу.

Определение 1. В точке $z_*=[x_*, u_*, v_*, y_*]$ выполнены достаточные условия второго порядка, если имеет место (2.6)–(2.8) и

$$\langle d, L_{xx}(z_*)d\rangle>0 \tag{2.9}$$

для любого ненулевого $d\in C(z_*)$, где

$$C(z_*)=\{d\in D_1(x_*)\cap D_2(x_*, y_*)|\langle f_x^i(x_*), d\rangle=0,$$

$$i\in I_1(u_*), \ \langle g_x^j(x_*), d\rangle=0, \ j\in I_2(v_*)\},$$

$$D_2(x, y)=\{d\in R^n|\langle f_x^i(x), d\rangle\leqslant 0|i\in J_1(x, y)\},$$

$$I_1(u)=\{i|u^i>0\}, \qquad I_2(v)=\{j|v^j>0\}.$$

Теорема 3. *Пусть в точке* $[x_*, u_*, v_*, y_*]$ *выполнены достаточные условия второго порядка. Тогда* x_* *является строгим локальным решением задачи* (1.1).

Доказательство теоремы аналогично доказательству теоремы Мак-Кормика для задачи н.л.п. [10]. Отметим также, что условие (2.9) более жесткое, чем достаточные условия второго порядка из [11].

§ 3. Вспомогательная функция

Обратимся к минимаксной задаче (2.2). Функция $L(x, u, v, y)$ линейна по u и v. Поэтому для того, чтобы точка $[u, v]$ была решением внутренней подзадачи задачи (2.2) при фиксированных x и y, необходимо (см., например, [12]), чтобы

$$L_u(x, u, v, y)\in K^0(u|S^r), \qquad L_v(x, u, v, y)\in K^0(v|R_+^m), \tag{3.1}$$

где $K(q|Q)$ – конус допустимых направлений относительно множества Q в точке $q\in Q$, $K^0(q|Q)$ – поляра конуса $K(q|Q)$. Согласно определению, $K^0(q|Q)=\{z_*|\langle z, z_*\rangle\leqslant 0 \ \forall z\in K(q|Q)\}$. Таким образом, решение задачи (2.2) является одновременно решением следующей оптимизационной задачи: найти

$$\min_{x\in R^n} L(x, u, v, y) \tag{3.2}$$

при ограничениях (3.1).

Задача (3.2), (3.1) – это задача параметрического н. л. п., и для ее решения может быть применен подход, основанный на использовании внешних штрафных функций.

Пусть Q – выпуклое замкнутое множество и точка $q\in Q$. Возьмем произвольную выпуклую функцию $F(p)$ и составим функцию

$$M(p, q|Q)=[F(p)+\delta(p|Q-q)]^*, \tag{3.3}$$

где $\delta(p|Q)$ – индикаторная функция множества Q; индекс звездочка означает сопряжение. Символами $M^*(p, q|Q)$ и $(M0^+)(p, q|Q)$ будем обозначать, соответственно, сопряженную и рецессивную функции функции $M(p, q|Q)$, причем везде считается, что эти операции относятся только к первому аргументу.

Лемма 2. *Пусть* $F(p)$ – *гладкая выпуклая кофинитная функция, достигающая своего минимума на* R^l *в нуле. Тогда функция* $M(p, q|Q)$ *выпукла и замкнута по* p, *достигает своего минимума по* p *при фиксированном* $q\in Q$ *в точках множества* $K^0(q|Q)$ *и* $(M0^+)(p, q|Q)=\delta^*(p|Q-q)$.

Доказательство. Выпуклость и замкнутость функции $M(p, q|Q)$ следуют из основных свойств сопряженной функции. Далее, по теореме 27.1 из [12], множество минимумов функции $M(p, q|Q)$ по p совпадает с субдифференциалом функции $M^*(p, q|Q)$ по первому аргументу в нуле. С учетом утверждения теоремы 25.6 (из [12]) получаем

$$\partial_p M^*(0, q|Q)=\partial F(0)+\partial\delta(0|Q-q)=\partial\delta(0|Q-q)=K^0(q|Q).$$

Наконец, согласно теореме 13.3 из [12], рецессивная функция $(M0^+)(p, q|Q)$ является опорной функцией эффективного множества функции $M^*(p, q|Q)$. Но это множество, очевидно, есть $Q-q$. Отсюда заключаем, что $(M0^+)(p, q|Q)=\delta^*(p|Q-q)$. Лемма доказана.

Построим по формуле (3.3) функцию $M(p, q|Q)$, используя в качестве $F(p)$ функцию $F(p)=\|p\|^2/2$, где $\|p\|$ – евклидова норма вектора p. По теореме 16.4 из [12],

$$[F(p)+\delta(p|Q-q)]^*=F^*(p)\Box\delta^*(p|Q-q).$$

Здесь $\Box$ – знак операции инфимальной конволюции. Имеем $F^*(p)=F(p)$. Помимо того, $\delta^*(p|Q-q)=\delta^*(p|Q)-\langle p, q\rangle$. Тогда, согласно определению операции инфимальной конволюции,

$$M(p,q|Q)=\inf_{p_1+p_2=p}\left[\frac{1}{2}\|p_1\|^2+\sup_{q_1\in Q}\langle q_1,p_2\rangle-\langle q,p_2\rangle\right]= \tag{3.4}$$

$$=\inf_{p_2\in R^l}\sup_{q_1\in Q}\left[\frac{1}{2}\|p-p_2\|^2+\langle p_2,q_1-q\rangle\right].$$

Рассмотрим функцию

$$M_1(p,q|Q)=\sup_{q_1\in Q}\inf_{p_2\in R^l}\left[\frac{1}{2}\|p-p_2\|^2+\langle p_2,q_1-q\rangle\right]. \tag{3.5}$$

Минимум по p_2 для любых $q_1\in Q$ достигается в точке $p_2=p+q-q_1$. Подставляя это выражение в (3.5), получаем

$$M_1(p,q|Q)=\sup_{q_1\in Q}\left[\frac{1}{2}\|q_1-q\|^2+\langle q_1-q,p+q-q_1\rangle\right]= \tag{3.6}$$

$$=\sup_{q_1\in Q}\frac{1}{2}(\|p\|^2-\|p+q-q_1\|^2)=\frac{1}{2}(\|p\|^2-\inf_{q_1\in Q}\|p+q-q_1\|^2).$$

В силу замкнутости и выпуклости множества Q, проекция $P_Q(p+q)$ существует и единственна. Отсюда следует, что множество, на котором достигается супремум по q_1 в (3.6), всегда не пусто и компактно. Поэтому, по обобщенной теореме фон-Неймана (см. [13]), функции (3.4) и (3.5) совпадают. Таким образом,

$$M(p, q|Q)=[\|p\|^2-\|p+q-P_Q(p+q)\|^2]/2. \tag{3.7}$$

Уточним теперь, какой вид принимает функция (3.7) для двух важных для дальнейшего множеств

Для $Q=R_+^m$. Согласно (3.7),

$$M(p, q|R_+^m)=[\|p\|^2-\|p+q-(p+q)_+\|^2]/2=$$

$$=[\|p\|^2-\|(p+q)_-\|^2]/2.$$

Здесь через b_+ и b_- обозначены векторы, j-е компоненты которых определяются через b^j по формулам $b_+^j=\max[0, b^j]$, $b_-^j=\min[0, b^j]$.

Для $Q=S^r$. Чтобы отыскать $P_{S^r}(b)$, упорядочим все компоненты вектора b в порядке их убывания: $b^{i_1}\geqslant\ldots\geqslant b^{i_r}$, и определим величины

$$\mu^k=k^{-1}\left(kb^{i_k}-\sum_{j=1}^{k}b^{i_j}+1\right),\qquad k=1,2,\ldots,r.$$

Нетрудно видеть, что $\mu^1=1$ и если $\mu^k>0$, то $\mu^k>\mu^{k+1}$. Действительно, при $\mu^{k+1}\leqslant 0$ это неравенство очевидно. В противном случае

$$\mu^k\geqslant k^{-1}\left(kb^{i_{k+1}}-\sum_{j=1}^{k}b^{i_j}+1\right)=k^{-1}\left[(k+1)b^{i_{k+1}}-\sum_{j=1}^{k+1}b^{i_j}+1\right]>$$

$$>(k+1)^{-1}\left[(k+1)b^{i_{k+1}}-\sum_{j=1}^{k+1}b^{i_j}+1\right]=\mu^{k+1}.$$

Пусть $k_*(b)$ – максимальный индекс из $1, 2, \ldots, r$ такой, что $\mu^{k_*(b)}>0$. Обозначим $J(b)=\{i_1,\ldots,i_{k_*(b)}\}$ и положим

$$\lambda(b)=k_*^{-1}(b)\left(\sum_{i\in J(b)}b^i-1\right).$$

Тогда

$$P_{S^r}(b)=\begin{cases}b^i-\lambda(b), & i\in J(b),\\ 0, & i\notin J(b).\end{cases}\tag{3.8}$$

Подставляя это выражение в (3.7), получаем

$$M(p,q\,|\,S^r)=\frac{1}{2}\left\{\|p\|^2-\sum_{i\notin J(p+q)}(p^i+q^i)^2-\left[\sum_{i\in J(p+q)}(p^i+q^i)-1\right]^2\right\}.$$

Из формулы (3.7) вытекает, что функция $M(p,q|Q)$ непрерывно дифференцируема по первому аргументу. В самом деле, в силу единственности точки, в которой строго выпуклая функция $\|p+q-q_1\|^2$ достигает своего минимума на выпуклом множестве Q, функция

$$\chi(p,q)=\min_{q_1\in Q}\|p+q-q_1\|^2,$$

согласно [14], дифференцируема по первому аргументу по всем направлениям $d\in R^l$ и имеет место формула

$$\partial_p\chi(p,q)/\partial d=\langle p+q-P_Q(p+q),d\rangle.$$

Отсюда следует, что функция $\chi(p,q)$ обладает частными производными по p. Но поскольку функция $\chi(p,q)$ выпукла, то существования частных производных достаточно для того, чтобы $\chi(p,q)$ была дифференцируемой. Таким образом, функция $M(p,q|Q)$ также дифференцируема и ее градиент равен

$$M_p(p,q|Q)=P_Q(p+q)-q.\tag{3.9}$$

Ниже I_s и E_s – матрицы размеров $s\times s$, I_s – единичная матрица, E_s – матрица, у которой все элементы равны единице.

Лемма 3. *Пусть точки* $p_*\in R^r$ *и* $q_*\in S^r$ *таковы, что* $p_*=[0,\ldots,0,p_*^{s+1},\ldots,p_*^r]$, $q_*=[q_*^1,\ldots,q_*^s,0,\ldots,0]$, *где* $p_*^i\leqslant 0$, $s<i\leqslant r$, $q_*^j>0$, $1\leqslant j\leqslant s$,

$1\leqslant s\leqslant r$. *Тогда* $M_p(p_*, q_*|S^r)=0$. *Если, помимо того,* $p_*^i<0$, $s<i\leqslant r$, *то функция* $M(p, q|S^r)$ *дважды дифференцируема по* p *в точке* $[p_*, q_*]$ *и ее матрица вторых производных имеет вид*

$$(3.10)\qquad M_{pp}(p_*, q_*|S^r)=\left\|\begin{array}{c:c} I_s & 0 \\ \hdashline 0 & 0 \end{array}\right\|-\frac{1}{s}\left\|\begin{array}{c:c} E_s & 0 \\ \hdashline 0 & 0 \end{array}\right\|.$$

Доказательство. Как уже отмечалось, функция $M(p, q|Q)$ непрерывно дифференцируема по p и имеет место формула (3.9).

Вычислим $P_{S^r}(p_*+q_*)$. В силу сделанных предположений, $p_*^i+q_*^i=$ $=q_*^i>0$, $1\leqslant i\leqslant s$, и $p_*^j+q_*^j=p_*^j\leqslant 0$, $s<j\leqslant r$. Пусть, для определенности, $q_*^1\geqslant\ldots\geqslant q_*^s$ и $p_*^{s+1}\geqslant\ldots\geqslant p_*^r$. Тогда

$$(3.11)\qquad \mu^i=i^{-1}\left[iq_*^i+\left(1-\sum_{j=1}^{j}q_*^j\right)\right]>0,\qquad 1\leqslant i\leqslant s,$$

$$(3.12)\qquad \mu^i=i^{-1}\left[ip_*^i-\sum_{j=s+1}^{s}p_*^i\right]\leqslant 0,\qquad s<i\leqslant r.$$

Отсюда приходим к выводу, что $J(p_*+q_*)=\{1, 2, \ldots, s\}$ и, следовательно,

$$\lambda(p_*+q_*)=s^{-1}\left[\sum_{i=1}^{s}(p_*^i+q_*^i)-1\right]=\sum_{i=1}^{0}p_*^i/s=0.$$

Поэтому, согласно (3.8), $P_{S^r}(p_*+q_*)=q_*$ и, стало быть, $M_p(p_*, q_*|S^r)=0$.

Пусть теперь $p_*^i<0$, $s<i\leqslant r$. Так как в этом случае неравенства (3.12) строгие, то найдется достаточно малая окрестность точки p_* такая, что для всех p из этой окрестности строгие неравенства (3.11) и (3.12) сохраняются. Поэтому множество $J(p+q_*)$ для этих p также будет состоять из первых s индексов и справедлива формула

$$M_p^i(p, q_*|S^r)=\begin{cases} p^i-\lambda(p+q_*), & 1\leqslant i\leqslant s,\\ 0, & s<i\leqslant r,\end{cases}$$

$$\lambda(p+q_*)=s^{-1}\sum_{i=1}^{0}p^i.$$

Таким образом, вектор-функция $M_p(p, q_*|S^r)$ непрерывно дифференцируема по p в окрестности точки p_* и ее матрица Якоби $M_{pp}(p_*, q_*|S^r)$ имеет вид (3.10). Лемма доказана.

Аналогично доказывается

Лемма. 4. *Пусть точки* $p_*\in R^m$ *и* $q_*\in R_+^m$ *таковы, что* $p_*=[0, \ldots, 0, p_*^{t+1}, \ldots, p_*^m]$, $q_*=[q_*^1, \ldots, q_*^t, 0, \ldots, 0]$, *где* $p_*^i\leqslant 0$, $t<i\leqslant m$, *и* $q_*^j>0$, $1\leqslant j\leqslant t$, $1\leqslant t\leqslant m$. *Тогда* $M_p(p_*, q_*|R_+^m)=0$. *Если, помимо того,* $p_*^i<0$, $t<i\leqslant m$, *то функция* $M(p, q|R_+^m)$ *дважды дифференцируема по* p *в точке* $[p_*, q_*]$ *и ее матрица вторых производных имеет вид*

$$M_{pp}(p_*, q_*|R_+^m)=\left\|\begin{array}{c:c} I_t & 0 \\ \hdashline 0 & 0 \end{array}\right\|.$$

Используя вид матрицы $M_{pp}(p_*, q_*|S^r)$, получаем, что для любого вектора $d\in R^r$ выполнено

$$\langle d, M_{pp}(p_*, q_*|S^r)d\rangle=\sum_{i=1}^{s}(d^i)^2-s^{-1}\left(\sum_{i=1}^{s}d^i\right)^2.$$

Отсюда на основании неравенства между квадратичным и арифметическим средним приходим к выдову, что матрица $M_{pp}(p_*, q_*|S^r)$ неотрицательно определена, причем равенство $\langle d, M_{pp}(p_*, q_*|S^r)d\rangle=0$ возможно только в том случае, когда первые s компонент вектора d равны между собой.

§ 4. Численный метод

Ниже для сокращения записи переменные x, u, v и y объединим одним символом, положив $z=[x, u, v, y]$. Символ z_* будем использовать для обозначения точки $z_*=(x_*, u_*, v_*, y_*]$. Составим м. ф. Л.:

$$H(z, \alpha)=L(z)+(M\alpha)(L_u(z), u|S^r)+(M\alpha)(L_v(z), v|R_+^m). \tag{4.1}$$

Здесь $(M\alpha)$ — функция (3.7), умноженная справа на константу α, причем, как и прежде, считается, что эта операция относится только к первому аргументу. Согласно определению из [12], $(M\alpha)(p, q|Q)=\alpha M(p/\alpha, q|Q)$. Если (1.1) — задача н. л. п., функция (4.1) переходит в хорошо известную м. ф. Л., исследованную многими авторами (см., например, [2] — [4]).

Обозначим через H_x градиент функции H по x, через H_{xx} — матрицу ее вторых производных. Имеет место

Лемма 5. *Пусть z_* — точка Куна — Таккера. Тогда для любого $\alpha>0$*

$$H(z_*, \alpha)=0, \qquad H_x(z_*, \alpha)=0, \tag{4.2}$$

$$(M\alpha)_p(L_u(z_*), u_*|S^r)=0, \qquad (M\alpha)_p(L_v(z_*), v_*|R_+^m)=0. \tag{4.3}$$

Добавление к функции Лагранжа штрафных членов значительно улучшает ее поведение, а именно: можно показать, что при некоторых дополнительных предположениях вблизи точек Куна — Таккера всегда существуют локальные решения задачи безусловной минимизации

$$\min_{x\in R^n} H(x, u, v, y, \alpha). \tag{4.4}$$

Определение 2. В точке z_* выполнено условие дополняющей нежесткости, если $h^i(x_*, y_*)u_*^i=0$, $g^j(x_*)v_*^j=0$, $1\leqslant i\leqslant r$, $1\leqslant j\leqslant m$. Если, кроме того, для всех $1\leqslant i\leqslant r$ и $1\leqslant j\leqslant m$ равенства $h^i(x_*, y_*)=0$ и $g^j(x_*)=0$ возможны в том и только том случае, когда $u_*^i>0$, $v_*^j>0$, то в точке z_* выполнено условие строгой дополняющей нежесткости.

Лемма 6. *Пусть в точке z_* выполнены достаточные условия второго порядка и условия строгой дополняющей нежесткости. Тогда можно указать константу $\alpha_*>0$ такую, что при $0<\alpha<\alpha_*$ для всех $[u, v, y]$ из некоторой окрестности $\Delta(u_*, v_*, y_*)$ точки $[u_*, v_*, y_*]$, $u\in S^r$, $v\in R_+^m$, существует изолированное локальное решение $x(u, v, y, \alpha)$ задачи (4.4), удовлетворяющее условию $x(u_*, v_*, y_*, \alpha)=x_*$.*

Доказательство. Согласно утверждению леммы 5, $H_x(z_*, \alpha)=0$ для любого $\alpha>0$. Пусть, для определенности, первые s компонент вектора u_* и первые t компонент вектора v_*, и только они, больше нуля. Из предположения о строгой дополняющей нежесткости вытекает, что точка $[p_*, q_*]=[L_u(z_*), u_*]$ удовлетворяет условиям леммы 3, причем $p_*^i<0$, $s<i\leqslant r$. Точно так же точка $[p_*, q_*]=[L_v(z_*), v_*]$ удовлетворяет условиям леммы 4, поэтому функция $H(z, \alpha)$ дважды дифференцируема по x в точке z_* и

$$H_{xx}(z_*, \alpha)=L_{xx}(z_*)+f_x^T(x_*)(M\alpha)_{pp}(L_u(z_*), u_*|S^r)f_x(x_*)+$$
$$+g_x^T(x_*)(M\alpha)_{pp}(L_v(z_*), v_*|R_+^m)g_x(x_*).$$

Матрицы $(M\alpha)_{pp}(L_u(z_*),\ u_*|S^r)$ и $(M\alpha)_{pp}(L_v(z_*),\ v_*|R_+^m)$ неотрицательно определены, причем $\langle d,\ g_x^T(x_*)(M\alpha)_{pp}(L_v(z_*),\ v_*|R_+^m)g_x(x_*)d\rangle=0$ в том и только том случае, когда $\langle g_x^i(x_*),\ d\rangle=0$, $1\leqslant i\leqslant t$. Кроме того, как отмечалось выше, равенство $\langle d,\ f_x^T(x_*)(M\alpha)_{pp}(L_u(z_*),\ u_*|S^r)f_x(x_*)d\rangle=0$ имеет место, когда величины $\langle f_x^i(x_*),\ d\rangle$, $1\leqslant i\leqslant s$, равны между собой и, в частности, совпадают с нулем. Отсюда, согласно (2.6), (2.9) и лемме Финслера (см. [4]), матрица $H_{xx}(z_*,\ \alpha)$ положительно определена при α достаточно малых, $\alpha<\alpha_*$. В силу непрерывности она будет оставаться положительно-определенной и вблизи точки z_*. Тогда, по теореме о неявной функции, можно указать окрестность $\Delta(u_*,\ v_*,\ y_*)$ точки $[u_*,\ v_*,\ y_*]$ такую, что на ней существует непрерывно дифференцируемая функция $x(u,\ v,\ y,\ \alpha)$, определяемая уравнением $H_x(z,\ \alpha)=0$ и являющаяся решением задачи (4.4). Лемма доказана.

Пусть задан вектор начальных уровней $y_0\in R^r$ и выбрано произвольное направление $e\in\operatorname{int}R_+^r$. Положим $y(y_0,\ e,\ \sigma)=y_0+\sigma e$, $\sigma\in R^1$, и построим итеративный процесс для отыскания точек z_*, удовлетворяющих условиям (2.6)–(2.8):

(4.5а) $$y_k=y(y_0,e,\sigma_k),\qquad x_k\in\operatorname{Arg}\min_{x\in R^n}H(x,u_k,v_k,y_k,\alpha),$$

(4.5б) $$u_{k+1}=P_{S^r}(L_u(z_k)/\alpha+u_k),\qquad v_{k+1}=P_{R_+^m}(L_v(z_k)/\alpha+v_k),$$

(4.5в) $$\sigma_{k+1}=\sigma_k+H(z_k,\ \alpha)/\langle u_k,\ e\rangle.$$

Здесь $\sigma_0=0$, $u_0\in S^r$, $v_0\in R_+^m$, $z_k=[x_k,\ u_k,\ v_k,\ y_k]$.

Положим $w=[u,\ v,\ \sigma]$, $x(w,\ \alpha)=x(u,\ v,\ y(y_0,\ e,\ \sigma),\ \alpha)$, где $x(u,\ v,\ y,\ \alpha)$ — непрерывно дифференцируемая вектор-функция, определяемая из решения задачи минимизации (4.4). Процесс (4.5) — это по-существу метод простой итерации, примененный для решения системы уравнений

(4.6а) $$(M\alpha)_p(L_u(x(w,\ \alpha),\ u,\ v,\ y(y_0,\ e,\ \sigma)),\ u|S^r)=0,$$

(4.6б) $$(M\alpha)_p(L_v(x(w,\ \alpha),\ u,\ v,\ y(y_0,\ e,\ \sigma)),\ v|R_+^m)=0,$$

(4.6в) $$H(x(w,\ \alpha),\ u,\ v,\ y(y_0,\ e,\ \sigma),\ \alpha)/\langle u,\ e\rangle=0.$$

Обозначим

$$G(x,\ w,\ \alpha)=\left\|\begin{matrix}P_{S^r}(L_u(x,\ u,\ v,\ y(y_0,\ e,\ \sigma))/\alpha+u)\\ P_{R_+^m}(L_v(x,\ u,\ v,\ y(y_0,\ e,\ \sigma))/\alpha+v)\\ \sigma+H(x,\ u,\ v,\ y(y_0,\ e,\ \sigma),\ \alpha)/\langle u,\ e\rangle\end{matrix}\right\|$$

и определим отображение

(4.7) $$W(w,\ \alpha)=G(x(w,\ \alpha),\ w,\ \alpha).$$

Пусть $y_*(y_0,\ e)$ — точка, лежащая на прямой $\gamma(y_0,\ e)=\{y\in R^r|y=y_0+\sigma e,\ \sigma\in R^1\}$, такая, что $y_*(y_0,\ e)\in\operatorname{fr}F_+$. Понятно, что для каждых $y_0\in R^r$ и $e\in\operatorname{int}R_+^r$ существует только одна такая точка. Величину σ, для которого $y(y_0,\ e,\ \sigma)=y_*(y_0,\ e)$, обозначим через $\sigma(y_0,\ e)$. Если точка w_* — решение системы (4.6) и $y_*=y(y_0,\ e,\ \sigma_*)$, $x_*=x(w_*,\ \alpha)$, то

(4.8а) $$(M\alpha)_p(L_u(z_*),\ u_*|S^r)=0,\qquad (M\alpha)_p(L_v(z_*),\ v_*|R_+^m)=0,$$

(4.8б) $$H_x(z_*,\ \alpha)=0,\qquad H(z_*,\ \alpha)/\langle u_*,\ e\rangle=0.$$

Равенства (4.8а) означают, что имеет место (2.7). Кроме того, из них и из второго равенства (4.8б) следует, что в z_* выполнено (2.8). Наконец, из (4.4) и (4.8) вытекает (2.6). Таким образом, z_* — точка Куна — Таккера. Если в точке z_* имеет место условие строгой дополняющей нежесткости, то, в силу утверждений лемм 3 и 4, отображение (4.7) дифференцируемо в точке w_*.

Пусть $N(y_*)=\{y\in R^r|y=y_*+\sigma e,\ e\in \operatorname{int} R_+^r,\ \sigma\in R^1\}$ — конус в пространстве R^r. Всюду ниже под окрестностью точки $[u_*, v_*, y_*]$ будем понимать множество вида $\Delta(u_*, v_*, y_*)=\Delta_{S^r}(u_*)\times\Delta_{R_+^m}(v_*)\times\Delta_{N(y_*)}(y_*)$, где $\Delta_Q(q)=$ $=\Delta(q)\cap Q$, $\Delta(q)$ — некоторая окрестность точки q, содержащая ее внутри себя.

Определение 3. Метод (4.5) локально сходится к точке $[u_*, v_*, y_*]$, если найдется такая ее окрестность $\Delta(u_*, v_*, y_*)$, что для любых u_0, v_0 и y_0 из этой окрестности и для соответствующего $e\in\operatorname{int} R_+^r$, для которого $y_*\in$ $\in\gamma(y_0, e)$, последовательность $\{[u_k, v_k, y_k]\}$, порождаемая (4.5), сходится к $[u_*, v_*, y_*]$.

Нетрудно видеть, что для сходимости процесса (4.5) необходимо и достаточно, чтобы соответствующая последовательность $\{[u_k, v_k, \sigma_k]\}$ сходилась к точке $[u_*, v_*, \sigma_*]$, где $\sigma_*=\sigma(y_0, e)$.

Через $\Phi(x, y)$ обозначим l-мерную вектор-функцию, первыми r компонентами которой являются функции $h^i(x, y)$, $1\leqslant i\leqslant r$, а последующими m компонентами — функции $g^j(x)$, $1\leqslant j\leqslant m$; через $J_0(x, y)$ обозначим индексное множество $\{1\leqslant i\leqslant l|\Phi^i(x, y)=0\}$. Символом $\Phi_x(x, y)$ будем обозначать матрицу первых производных функции Φ относительно x размером $l\times n$.

Определение 4. В точке $[x_*, y_*]$ выполнено условие регулярности, если $\Phi_x(x_*, y_*)d\neq0$ для любого ненулевого вектора $d\in R^l$, у которого сумма r первых компонент равна нулю и такого, что $d^i=0$, $i\notin J_0(x_*, y_*)$.

Теорема 4. *Пусть в точке z_* выполнены достаточные условия второго порядка и условия строгой дополняющей нежесткости. Пусть, кроме того, в $[x_*, y_*]$ выполнено условие регулярности и матрица $L_{xx}(z_*)$ не вырождена. Тогда можно указать $\alpha_*>0$ такое, что для всех $0<\alpha<\alpha_*$ процесс (4.5) локально сходится к точке $[u_*, v_*, y_*]$ со скоростью геометрической прогрессии.*

Доказательство. Для произвольного $y_0\in\Delta_{N(y_*)}(y_*)$ и соответствующего $e\in\operatorname{int} R_+^m$ такого, что $y_*\in\gamma(y_0, e)$, рассмотрим отображение $W(w, \alpha)$. Пусть $W_w(w_*, \alpha)$ — матрица его первых производных, вычисленная в точке w_*. Тогда

$$W_w(w_*,\alpha)=G_w(x_*,w_*,\alpha)+G_x(x_*,w_*,\alpha)\,\frac{dx(w_*,\alpha)}{dw},$$

где $x_*=x(w_*, \alpha)$. В силу утверждения леммы 6, если α выбрано достаточно малым ($\alpha<\alpha_1$), то непрерывно дифференцируемая функция $x(u, v, y, \alpha)$ определена на некоторой окрестности точки $[u_*, v_*, y_*]$. Естественно, что тогда функция $x(w, \alpha)$ будет обладать этим свойством в окрестности точки w_*, где $\sigma_*=\sigma(y_0, e)$. Подставляя $x(w, \alpha)$ в равенство $H_x(z, \alpha)=0$, следующее из необходимых условий экстремума для задачи (4.4), и дифференцируя получившееся тождество по w, находим

$$\frac{dx}{dw}(w_*,\alpha)=-H_{xx}^{-1}(z_*,\alpha)\,H_{xw}(z_*,\alpha).$$

Не умаляя общности, ниже будем считать точку z_* такой, что первые s компонент вектора $L_u(z_*)$ и первые t компонент вектора $L_v(z_*)$ обращаются в нуль. Из условия строгой дополняющей нежесткости тогда следует, что первые s компонент вектора u_* и первые t компонент вектора v_*, и только они, отличны от нуля.

На основании утверждения лемм 3 и 4, для z из окрестности точки z_* выполняется

$$P_{S^{r}}{}^{i}\left(\frac{L_u(z)}{\alpha}+u\right)=\begin{cases}\dfrac{L_u{}^i(z)}{\alpha}+u^i-s^{-1}\left\{\sum\limits_{i=1}^{s}\left[\dfrac{L_u{}^i(z)}{\alpha}+u^i\right]-1\right\}, & 1\leqslant i\leqslant s,\\ 0, & s<i\leqslant r,\end{cases}$$

$$P^{j}_{R_+{}^m}(L_v(z)/\alpha+v)=\begin{cases}L_v{}^j(z)/\alpha+v^j, & 1\leqslant j\leqslant t,\\ 0, & t<j\leqslant m,\end{cases}$$

$$H(z,\alpha)=L(z)+\frac{\alpha}{2}\left\{\sum_{i=1}^{r}\left[\frac{L_u{}^i(z)}{\alpha}\right]^2-s^{-1}\left\{\sum_{i=1}^{s}\left[\frac{L_u{}^i(z)}{\alpha}+u^i\right]-1\right\}^2-\right.$$

$$\left.-\sum_{i=s+1}^{r}\left[\frac{L_u{}^i(z)}{\alpha}+u^i\right]^2+\sum_{j=1}^{m}\left[\frac{L_v{}^i(z)}{\alpha}\right]^2-\sum_{j=t=1}^{m}\left[\frac{L_v{}^j(z)}{\alpha}+v^j\right]^2\right\}.$$

Дифференцируя получившиеся выражения, находим, что матрица $W_w(w_*,\alpha)$ представима в виде

$$W_w(w_*,\alpha)=\left\|\begin{array}{c:c}A_{11} & A_{12}\\ \hdashline A_{21} & A_{22}\end{array}\right\|,$$

где A_{11} — матрица размера $l\times l$, A_{12} и A_{21} — соответственно, l-мерные вектор-столбец и вектор-строка, A_{22} — элемент:

$$A_{11}=\Gamma-\alpha^{-1}\Gamma\Phi_x H_{xx}{}^{-1}\Phi_x{}^{\mathrm{T}}\Gamma,$$

$$A_{12}=-\alpha^{-1}(\Gamma-\alpha^{-1}\Phi_x H_{xx}{}^{-1}\Phi_x{}^{\mathrm{T}}\Gamma)\bar{e},$$

$$A_{21}=[\Phi^{\mathrm{T}}(x_*,y_*)\Gamma-H_x{}^{\mathrm{T}}H_{xx}{}^{-1}H_{xw}]\langle e,u_*\rangle^{-1}-$$
$$-[H(z_*,\alpha)\langle e,u_*\rangle^{-2}]\bar{e}^{\mathrm{T}}=0,$$

$$A_{22}=\left\{1-\left[\sum_{i=1}^{s}u_*{}^ie^i+\alpha^{-1}\sum_{i=1}^{s}\Phi^i(x_*,y_*)(\Gamma\bar{e})^i\right]\right\}\langle e,u_*\rangle^{-1}=0;$$

здесь Γ — блочно-диагональная матрица размера $l\times l$:

$$\Gamma=\left\|\begin{array}{c:c:c:c}I_s-s^{-1}E_s & 0 & \multicolumn{2}{c}{0}\\ \hdashline 0 & 0 & & \\ \hdashline \multicolumn{2}{c:}{} & I_t & 0\\ \hdashline \multicolumn{2}{c:}{0} & 0 & 0\end{array}\right\|,$$

$\bar{e}=[e,0]$ есть l-мерный вектор. Для сокращения записи здесь и ниже аргументы матриц опускаются, при этом считается, что все они вычислены в точке z_*. Нетрудно видеть, что матрица Γ идемпотентна, т. е. $\Gamma^2=\Gamma$.

Так как $A_{21}=0$, то множество собственных значений матрицы $W_w(w_*,\alpha)$ состоит из одного собственного значения, соответствующего элементу A_{22} и равного нулю. и из l собственных значений матрицы A_{11}.

В силу симметричности A_{11}, все ее собственные значения вещественны. Уточним вид A_{11}. Имеем

$$H_{xx}=L_{xx}+\alpha^{-1}\Phi_x^{\mathrm{T}}\Gamma\Phi_x=L_{xx}+\alpha^{-1}\Phi_x^{\mathrm{T}}\Gamma^2\Phi_x.$$

С помощью формулы Шермана – Моррисона – Вудбери (см. [15]) найдем выражение для обратной матрицы:

$$H_{xx}^{-1}=L_{xx}^{-1}-\alpha^{-1}L_{xx}^{-1}\Phi_x^{\mathrm{T}}\Gamma(I_l+\alpha^{-1}\Gamma\Phi_xL_{xx}^{-1}\Phi_x^{\mathrm{T}}\Gamma)\Gamma\Phi_xL_{xx}^{-1}.$$

Подставляя его в матрицу A_{11}, получаем $A_{11}=\Gamma-\alpha^{-1}Z+\alpha^{-2}(I_l+\alpha^{-1}Z)^{-1}Z$, где $Z=\Gamma\Phi_xL_{xx}^{-1}\Phi_x^{\mathrm{T}}\Gamma$. Матрица $I_l+\alpha^{-1}Z$ для достаточно малых α будет неособой. Кроме того, в силу идемпотентности матрицы Γ выполняется $Z=\Gamma Z=Z\Gamma=\Gamma Z\Gamma$, поэтому

$$A_{11}=\Gamma[I_l-\alpha^{-1}Z-\alpha^{-2}Z(I_l+\alpha^{-1}Z)^{-1}Z]\Gamma=\Gamma(I_l+\alpha^{-1}Z)\Gamma.$$

Собственные значения ν матрицы A_{11} находятся из характеристического уравнения $\det(A_{11}-\nu I_l)=0$. Умножая матрицу $A_{11}-\nu I_l$ справа на неособую симметрическую матрицу $I_l+\alpha^{-1}Z$ и учитывая перестановочность матрицы Γ с любой симметрической матрицей, получаем

$$(I_l+\alpha^{-1}Z)(A_{11}-\nu I_l)=\Gamma(I_l+\alpha^{-1}Z)(I_l+\alpha^{-1}Z)^{-1}\Gamma-\nu(I_l+\alpha^{-1}Z)=$$
$$=\Gamma-\nu(I_l+\alpha^{-1}Z)=\Gamma(I_l-\nu\alpha^{-1}Z_1)\Gamma-\nu I_l,$$

где $Z_1=\Phi_xL_{xx}^{-1}\Phi_x^{\mathrm{T}}$. Таким образом, характеристическое уравнение может быть преобразовано к виду

$$\det[\Gamma(I_l-\nu\alpha^{-1}Z_1)\Gamma-\nu I_l]=0. \tag{4.9}$$

В силу симметричности, матрица Γ ортогонально подобна вещественной диагональной матрице Λ, диагональными элементами которой являются собственные значения матрицы Γ. Пусть U – ортогональная матрица такая, что $\Lambda=U^{\mathrm{T}}\Gamma U$. Так как $|\det U|=1$, то из (4.9) следует

$$\det(UU^{\mathrm{T}}\Gamma UU^{\mathrm{T}}(I_l-\nu\alpha^{-1}Z_1)UU^{\mathrm{T}}\Gamma UU^{\mathrm{T}}-\nu UU^{\mathrm{T}})=$$
$$=\det[\Lambda(I_l-\nu\alpha^{-1}Z_2)\Lambda^{\mathrm{T}}-\nu I_l]=0,$$

где $Z_2=U^{\mathrm{T}}Z_1U$.

Найдем собственные значения матрицы Γ. Из ее вида следует, что заведомо $l-(t+s)$ собственных ее значений равны нулю, а t собственных значений – единице. Оценим собственные значения, соответствующие верхней левой подматрице $I_s-s^{-1}E_s$. Они удовлетворяют уравнению $\det(I_s-s^{-1}E_s-\beta I_s)=0$, из которого вытекает, что собственные значения β^i матрицы $I_s-s^{-1}E_s$ связаны с собственными значениями θ^i матрицы E_s соотношениями $\theta^i=s(1-\beta^i)$, $1\leqslant i\leqslant s$. Чтобы найти θ^i, возьмем первый столбец матрицы $E_s-\theta I_s$ и вычтем его из всех других столбцов, а затем прибавим к первой строке сумму всех последующих строк. Определитель матрицы от этого не изменится, но сама матрица преобразуется к нижней треугольной, у которой определитель равен произведению диагональных элементов. Таким образом, $\det(E_s-\theta I_s)=(s-\theta)\theta^{s-1}=0$, и, следовательно, среди s собственных значений матрицы E_s лишь одно отлично от нуля и равно s. Вспоминая связь между числами θ^i и β^i, приходим к выводу, что $s+t-1$ собственных значений Γ равны единице, а остальные числа совпадают с нулем. Поэтому у диагональной матрицы Λ лишь $s+t-1$ диагональных элементов отличны от нуля и, следовательно, среди собственных значений ν^i по крайней мере $l-(s+t-1)$ чисел равны нулю. Отбросим их. Оставшиеся числа могут быть найдены из уравнения $\det[I_{s+t-1}(1-\nu)-\nu\alpha^{-1}Z_3]=0$, где Z_3 – квадратная матрица Z_2, у которой удалены строки и столбцы, соответствующие ее нулевым диагональным элементам.

Пусть λ^i – собственные значения матрицы Z_3. Поскольку выполнены условия регулярности, то обе матрицы, на которые умножается в Z_3 матрица L_{xx}^{-1}, имеют полный ранг. Но сама матрица L_{xx}^{-1} не вырождена, поэтому Z_3 — также невырожденная матрица и все ее собственные значения отличны от нуля. Числа ν^i связаны с числами λ^i соотношениями $\nu^i=1-\lambda^i\times\times(1+\lambda^i)^{-1}$, $1\leqslant i\leqslant s+t-1$. Если все λ^i положительны, то $|\nu^i|<1$ для любых $\alpha>0$. В противном случае это неравенство имеет место, если $\alpha<\alpha_2$, где $\alpha_2=|\lambda_*|/2$. Здесь λ_* – минимальное по модулю отрицательное λ^i. Таким образом, если взять $\alpha<\alpha_*=\min\{\alpha_1, \alpha_2\}$, то спектральный радиус матрицы $W_w(w_*, \alpha)$ меньше 1 и, следовательно, по теореме Островского (см. [15]), процесс (4.5) локально сходится относительно переменных w к точке w_* со скоростью геометрической прогрессии. В силу непрерывности соответствующая последовательность $\{x_k\}$ сходится к точке x_*. Теорема доказана.

Обратим внимание, что, используя другие функции, удовлетворяющие условиям леммы 2, можно по формуле (4.1) получить целое семейство различных м.ф.Л. Соответствующие этим функциям итеративные процессы строятся аналогично (4.5).

§ 5. Заключение

Рассмотрим вопрос о достоинствах и недостатках метода и те ситуации, в которых он может оказаться полезным для решения задач м.к.о.

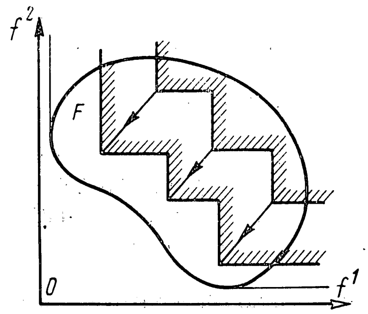
В отличие, например, от [7], метод применим для решения задач м.к.о. без всякого предположения о выпуклости функций, определяющих эти задачи. Он сохраняет все те хорошие вычислительные свойства, которые присущи методу м.ф.Л. при решении задач н.л.п., а именно: линейную скорость сходимости по двойственным переменным и существование точных локальных минимумов у вспомогательной функции вблизи решений задачи. Процесс (4.5) обобщает традиционную для задач н.л.п. схему и применим как для решения задач н.л.п., так и для задач м.к.о. с произвольным количеством критериев и ограничений, в том числе и при отсутствии последних.

Из недостатков метода основными являются его локальный характер и то, что в результате одного просчета он находит только одну оптимальную оценку.

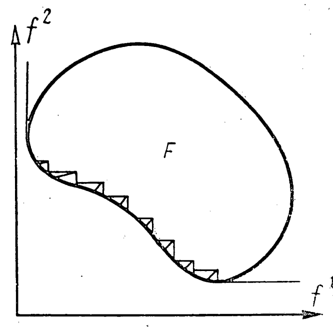
Сравнение достоинств и недостатков метода выявляет те ситуации при решении задач м.к.о., в которых применение метода наиболее целесообразно:

1) в глобальных методах аппроксимации множества Парето (см. фиг. 1) – в качестве вспомогательных процедур. Дело в том, что эффективность многих глобальных методов покрытия, например методов перебора на неравномерной сетке [16], в сильной степени зависит от того, насколько текущее рекордное значение близко к оптимальному. Локальные методы позволяют существенно ускорить процесс отыскания хороших рекордных оценок;

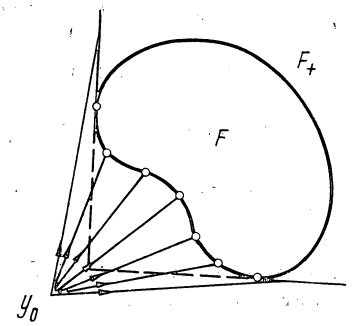
2) в методах построения ε-сети множества Парето на завершающей стадии расчетов – для получения ближайших точных решений (фиг. 2);



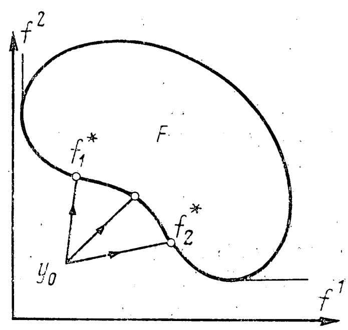
Фиг. 1



Фиг. 2



Фиг. 3



Фиг. 4

3) самостоятельное использование метода для построения аппроксимации множества Парето, особенно в задачах с большим числом переменных, когда методы перебора становятся неприемлемыми. В этом случае можно выбрать начальный вектор y_0 так, чтобы он лежал строго «юго-западнее» идеальной точки, и из него провести лучи по направлению к множеству F_+. Разные лучи будут приводить к разным оптимальным решениям (фиг. 3). Таким образом, в отличие от многих известных подходов к решению задачи м.к.о., когда она сводится с помощью линейной или какой-либо другой свертки к задаче н.л.п., параметризация в методе осуществляется не за счет выбора «весовых» коэффициентов, а за счет изменения направлений;

4) при построении сечений множества Парето и отыскании «срединных» решений. В последнем случае, если известны две точки, принадлежащие множеству Парето, и требуется найти точку, лежащую между ними, достаточно соединить их произвольными, принадлежащими положительному ортанту лучами, исходящими из одной и той же точки, и, проведя луч между ними в этой плоскости, найти новую оптимальную точку (фиг. 4). Подобным образом можно найти произвольное число промежуточных решений.

Автор выражает признательность В. В. Подиновскому за обсуждение работы и ценные замечания.

Литература

1. *Поляк Б. Т., Третьяков Н. В.* Метод штрафных оценок для задач на условный экстремум // Ж. вычисл. матем. и матем. физ. 1973. Т. 13. № 1. С. 34–46.
2. *Бертсекас Д.* Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987.

3. *Гольштейн Е. Г., Третьяков Н. В.* Слабые модифицированные функции Лагранжа и связанные с ними теоремы двойственности // Экономика и матем. методы. 1979. Т. 15. № 6. С. 1180–1193.
4. *Евтушенко Ю. Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.
5. *Жадан В. Г.* Модифицированные функции Лагранжа в нелинейном программировании // Ж. вычисл. матем. и матем. физ. 1982. Т. 22. № 2. С. 296–308.
6. *Wierzbicky A. P.* A penalty function shifting method in constrained static optimization and its convergence properties // Arch. Automat. i Telemech. 1971. V. 16. № 4. P. 395–416.
7. *Абасов Т. М.* О модифицированных функциях Лагранжа в многокритериальных задачах // Принятие решений в условиях многокритериальности и неопределенности. Тезисы докл. на IV Всес. семинаре по иссл. операций и системному анализу. М.– Батуми, 1983. С. 3.
8. *Жадан В. Г.* Метод параметризации целевых функций в условной многокритериальной оптимизации // Ж. вычисл. матем. и матем. физ. 1986. Т. 26. № 2. С. 177–189.
9. *Подиновский В. В., Ногин В. Д.* Парето-оптимальные решения многокритериальных задач. М.: Наука, 1982.
10. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной минимизации. М.: Мир, 1972.
11. *Гороховик В. В.* Условия слабой эффективности в конечномерных задачах векторной оптимизации. Минск: Ин-т матем. АН БССР, 1976.
12. *Рокафеллар Р.* Выпуклый анализ. М.: Мир, 1973.
13. *Гольштейн Е. Г.* Выпуклое программирование. Элементы теории. М.: Наука, 1970.
14. *Демьянов В. Ф., Васильев Л. В.* Недифференцируемая оптимизация. М.: Наука, 1981.
15. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975.
16. *Евтушенко Ю. Г., Потапов М. А.* Численные методы решения многокритериальных задач // Кибернетика и вычисл. техн. М.: Наука, 1987. Вып. 3. С. 209–218.

Поступила в редакцию 9.VII.1987
Переработанный вариант 25.V.1988