Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Челябинский государственный университет»

В. И. Ухоботов, И. В. Изместьев

ИЗБРАННЫЕ ГЛАВЫ ТЕОРИИ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Учебное пособие

Челябинск
Издательство Челябинского государственного университета
2020

УДК 517(075.8)
ББК В161.6я7
У895

Печатается по решению редакционно-издательского совета
Челябинского государственного университета

Р е ц е н з е н т ы:
кафедра математического анализа
и методики преподавания математики
Южно-Уральского государственного университета;
В. М. Ситников, кандидат физико-математических наук,
доцент кафедры гуманитарных, естественнонаучных и математических
дисциплин Уральского социально-экономического института (филиала)
Академии труда и социальных отношений

Ухоботов, В. И.

У895 Избранные главы теории экстремальных задач : учеб. пособие / В. И. Ухоботов, И. В. Изместьев. — Челябинск : Изд-во Челяб. гос. ун-та, 2020. — 112 с.

ISBN 978-5-7271-1652-4

Пособие посвящено методам решения оптимизационных задач в конечномерных пространствах. Подробно рассмотрена теория решения задач на условный экстремум: гладкие задачи, задачи выпуклого, квадратичного и линейного программирования. Особый акцент делается на применение численных методов. Изложение теоретического материала дополнено примерами решения прикладных задач.

Пособие предназначено для бакалавров и магистрантов, обучающихся по направлениям «Прикладная математика и информатика», «Математика и компьютерные науки», «Фундаментальная информатика и информационные технологии».

УДК 517.51(075.8)
ББК В161.6я73-1

ISBN 978-5-7271-1652-4

Оглавление

Введение

Учебное пособие написано на основе лекций по курсу «Дискретные модели», читаемых на математическом факультете Челябинского государственного университета. Подробно рассмотрена теория решения задач на условный экстремум. Особый акцент делается на применение численных методов. Изложение теоретического материала дополнено примерами решения прикладных задач. Предполагается, что читатели знакомы с курсами математического анализа и линейной алгебры.

В первой и второй главах приводятся базовые понятия и утверждения, связанные с выпуклыми функциями и выпуклыми множествами.

Третья и четвёртая главы посвящены численным методам решения задач безусловной оптимизации: градиентному методу, методу Ньютона и методу сопряжённых градиентов.

Главы с пятой по десятую посвящены теории решения задач на условный экстремум: гладкие задачи, задачи выпуклого, квадратичного и линейного программирования.

В одиннадцатой главе изложен численный метод поиска минимума дифференцируемой функции при линейных ограничениях.

Двенадцатая глава посвящена проблеме моментов и её приложениям в теории управления движением.

Нумерация теорем, лемм, определений в каждой главе самостоятельная. Нумерация формул двойная.

1. Выпуклые множества и их свойства

Рассмотрим линейное пространство R^n, элементами которого являются вектор-строки $x = (x_1,\ldots,x_n)$. Скалярное произведение двух векторов x и y из R^n задаётся формулой $\langle x, y\rangle = x_1 y_1 + \ldots + x_n y_n$, а евклидова норма вектора x определяется формулой $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2} = \sqrt{< x, x >}$.

Определение 1. Множество $X \subset R^n$ называется выпуклым, если вместе с любыми двумя своими точками она содержит весь соединяющий их отрезок, то есть

$$x \in X \text{ и } y \in X,\ 0 \le \lambda \le 1 \Rightarrow x_\lambda = \lambda x + (1-\lambda)\, y \in X.$$

Задача 1. Доказать, что выпуклыми являются следующие множества:
$S = \{x \in R^n : \|x\| \le 1\}$ — шар единичного радиуса с центром в начале координат;

$X = \{x \in R^n : \langle c, x\rangle = a\}$ — гиперплоскость;

$X = \{x \in R^n : \langle c, x\rangle \le a\}$ — замкнутое полупространство;

$X = \{x \in R^n : \langle c, x\rangle \ge a\}$ — замкнутое полупространство;

$X = \{x \in R^n : \langle c, x\rangle < a\}$ — открытое полупространство;

$X = \{x \in R^n : \langle c, x\rangle > a\}$ — открытое полупространство.

Здесь вектор $c \in R^n$ и число a заданы.

Задача 2. Доказать, что непустое пересечение выпуклых множеств является выпуклым множеством.

Для множества $A \subset R^n$ и для матрицы $G = \{g_{ij}\}$ размерности $m \times n$ определим множество

$$G(A) = \{y \in R^m : y = Ga,\ \forall a \in A\}. \qquad (1.1)$$

Утверждение 1. Если множество $A \subset R^n$ является выпуклым, то для любой матрицы G размерности $m \times n$ выпуклым является множество (1.1).

Доказательство. Пусть точки $y_i = G(A)$, $i = 1,2$. Тогда $y_i = Ga_i$ при некоторых $a_i \in A$. Возьмём число $0 \leq \lambda \leq 1$. Тогда

$$\lambda y_1 + (1 - \lambda)y_2 = \lambda Ga_1 + (1 - \lambda)Ga_2 =$$
$$= G(\lambda a_1 + (1 - \lambda)a_2),\ \lambda a_1 + (1 - \lambda)a_2 \in A.$$

Следовательно, $\lambda y_1 + (1 - \lambda)y_2 \in G(A)$.

Определение 2. Многогранным множеством назовём множество, каждая точка которого является решением линейных неравенств

$$\langle c_i, x\rangle \leq a_i,\ i = 1,\ldots,m.$$

Из решения задач 1 и 2 следует, что, если многогранное множество не является пустым, то оно выпуклое множество.

Определение 3. Говорят, что точка x является выпуклой комбинацией точек $x_1, x_2, \ldots, x_k$, если найдутся числа $\lambda_i \geq 0$, $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$ такие, что $x = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k$.

Утверждение 2. Пусть множество $X \subset R^n$ является выпуклым. Тогда, если точки $x_1, x_2, \ldots, x_k$ принадлежат множеству X, то ему принадлежит любая выпуклая комбинация этих точек

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k \in X,\ \lambda_i > 0,\ \lambda_1 + \lambda_2 + \ldots + \lambda_k = 1.$$

Доказательство. При $k = 2$ требуемое включение следует из определения выпуклого множества. Пусть оно справедливо при числе k. Докажем его для числа $k + 1$. Обозначим $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_k$. Нужно рассмотреть случай, когда $0 < \lambda < 1$. Тогда $\lambda_{k+1} = 1 - \lambda \in (0,1)$ и

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_k x_k + \lambda_{k+1} x_{k+1} = \lambda\left(\frac{\lambda_1}{\lambda} x_1 + \frac{\lambda_2}{\lambda} x_2 + \ldots + \frac{\lambda_k}{\lambda} x_k\right) +$$

$$+ (1 - \lambda)\, x_{k+1} \in X.$$

Определение 4. Пересечение всех выпуклых множеств, содержащих данное множество $X \subset R^n$, называется выпуклой оболочкой множества X и обозначается $co\, X$.

Утверждение 3. Имеет место равенство

$$co\, X = \{x \in R^n \colon x = \sum_{i=1}^{k} \lambda_i x_i,\ \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1,\ x_i \in X\}.$$

Доказательство. Множество, стоящее в правой части доказываемого равенства, обозначим Y. Тогда $X \subset Y$. Множество Y является выпуклым. Если некоторое выпуклое множество Z содержит в себе множество X, то оно содержится в множестве Y. Следовательно, множество Y является пересечением всех выпуклых множеств, содержащих в себе множество X.

Теорема 1 (*Каратеодори*). Пусть множество $X \subset R^n$. Тогда каждая точка $x \in co\, X$ является выпуклой комбинацией не более чем $n + 1$ точек из множества X.

Доказательство. Из предыдущей теоремы следует, что

$$x = \sum_{i=1}^{k} \lambda_i x_i,\ \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1,\ x_i \in X.$$

Пусть $k > n + 1$. Тогда вектора $(x_1;1)$, $(x_2;\ 1)$, …, $(x_k;1)$ в пространстве R^{n+1} являются линейно зависимыми. Следовательно, существуют не все равные нулю числа α_1, …, α_k такие, что

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0,\ \alpha_1 + \alpha_2 + \ldots + \alpha_k = 0. \tag{1.2}$$

Из последнего равенства (1.2) следует, что среди чисел α_i имеются положительные. Пусть

$$0 < \alpha_1, \ldots, 0 < \alpha_p,\ \alpha_{p+1} \leq 0, \ldots, \alpha_k \leq 0.$$

Считаем, что (иначе перенумеруем)

$$\frac{\lambda_1}{\alpha_1} \leq \frac{\lambda_2}{\alpha_2} \leq \ldots \leq \frac{\lambda_p}{\alpha_p}.$$

Обозначим

$$\lambda_1^* = \lambda_1 - \alpha_1 \frac{\lambda_1}{\alpha_1} = 0,\ \lambda_2^* = \lambda_2 - \alpha_2 \frac{\lambda_1}{\alpha_1} \geq 0,\ \ldots,\ \lambda_p^* = \lambda_p - \alpha_p \frac{\lambda_1}{\alpha_1} \geq 0,$$

$$\lambda_{p+1}^* = \lambda_{p+1} - \alpha_{p+1} \frac{\lambda_1}{\alpha_1} \geq 0,\ \ldots,\ \lambda_k^* = \lambda_k - \alpha_k \frac{\lambda_1}{\alpha_1} \geq 0.$$

Тогда

$$\sum_{i=1}^{k} \lambda_i^* = \sum_{i=1}^{k} \lambda_i - \frac{\lambda_1}{\alpha_1} \sum_{i=1}^{k} \alpha_i = 1,\ \sum_{i=1}^{k} \lambda_i^* x_i = \sum_{i=1}^{k} \lambda_i x_i - \sum_{i=1}^{k} \alpha_i x_i = x.$$

Так как число $\lambda_1^* = 0$, то вектор x представим в виде выпуклой комбинации $k - 1$ точки из множества X. Продолжим этот процесс дальше. Придём к тому, что вектор x представим в виде выпуклой комбинации не более чем $n + 1$ точек из множества X.

Отметим, что *компактом* в R^n является замкнутое и ограниченное множество.

Утверждение 4. Если множество X является компактом, то компактом является его выпуклая оболочка $co\ X$.

Доказательство. Обозначим набор $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{n+1}) \in R^{n+1}$. Рассмотрим функцию $f: R^{n+1} \times R^n \times R^n \times \ldots \times R^n \to R^n$, равную

$$f(\lambda, x_1, x_2, ..., x_{n+1}) = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_{n+1} x_{n+1}.$$

Введём в рассмотрение множества

$$\Lambda = \{\lambda \in R^{n+1}: \lambda_1 + \lambda_2 + \ldots + \lambda_{n+1} = 1; \lambda_i \geq 0,\ i = 1, ..., n+1\}$$

и $A = \Lambda \times X \times X \times \ldots \times X$, в котором множество X берётся $n+1$ раз. Теорема Каратеодори утверждает, что $co\ X = f(A)$. Поскольку функция f является непрерывной, а множество A является компактом, то образ $f(A)$ также является компактом.

Утверждение 5. Для любого множества $A \subset R^n$ и для любой матрицы G размерности $m \times n$ выполнено равенство

$$coG(A) = G(coA).$$

Доказательство. Пусть точка $x \in coG(A)$. Тогда, согласно утверждению 2, $x = \sum_{i=1}^{k} \lambda_i\ x_i, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1, x_i \in G(A)$. Отсюда и из формулы (1.1) получим, что $x = \sum_{i=1}^{k} \lambda_i\ Ga_i, a_i \in A$. Следовательно,

$$x = G(\sum_{i=1}^{k} \lambda_i a_i) \in G(coA).$$

Пусть точка $x \in G(coA)$. Тогда $x = Ga$ при некотором $a \in coA$. Отсюда и из утверждения 2 следует, что

$$a = \sum_{i=1}^{k} \lambda_i a_i,\ \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1,\ a_i \in A.$$

Следовательно, $x = G(\sum_{i=1}^{k} \lambda_i\ a_i) = \sum_{i=1}^{k} \lambda_i Ga_i \in coG(A)$.

Определение 5. Выпуклым многогранником в R^n называется выпуклая оболочка конечного числа точек.

Пусть заданы точки $x_1, x_2, \ldots, x_k$ в R^n. Тогда, согласно формуле (1.1), выпуклый многогранник, определяемый этими точками, равен

$$X = \{x \in R^n: x = \sum_{i=1}^{k} \lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1\} = co\{x_1, x_2, ..., x_k\}.$$

Топологические свойства выпуклых множеств

Определение 6. Точка x называется предельной точкой множества X, если существует последовательность точек $x_k \in X$ такая, что $x_k \to x$. Совокупность всех предельных точек множества X называется его замыканием и обозначается $\overline{X}$.

Утверждение 6. Замыкание $\overline{X}$ выпуклого множества X является выпуклым множеством.

Доказательство. Пусть $x \in \overline{X}$ и $y \in \overline{X}$. Тогда существуют последовательности точек $x_k \in X$, $y_k \in X$ такие, что $x_k \to x$, $y_k \to y$. При любом $0 \le \lambda \le 1$ выполнено включение $\lambda y_k + (1 - \lambda) x_k \in X$. Следовательно, предельная точка $\lambda y + (1 - \lambda)x \in \overline{X}$.

Определение 7. Точка x называется внутренней точкой множества X, если существует такое число $\varepsilon > 0$, что $x + \varepsilon s \in X$ для любого вектора $s \in R^n$, $\|s\| \le 1$.

Совокупность всех внутренних точек множества X называется его внутренностью и обозначается $int\, X$.

Теорема 2. Пусть множество X является выпуклым, точка $z \in int\, X$, а точка $x \in \overline{X}$. Тогда для любого числа $0 \le \lambda < 1$ точка

$$y = (1 - \lambda)z + \lambda x \in intX. \tag{1.3}$$

Доказательство. Очевидно, что нужно рассматривать случай $\lambda > 0$. Из включения $z \in intX$ следует, что существует числа $\varepsilon > 0$ такое, что

$$z + \varepsilon s_* \in X \text{ при любом } s_* \in R^n,\ \|s_*\| \le 1. \tag{1.4}$$

Если покажем, что

$$z + \delta s \in X \text{ при любом } s \in R^n,\ \|s\| \le 1, \text{ где } \delta = \frac{(1-\lambda)}{2}\varepsilon, \tag{1.5}$$

то это будет означать выполнение включения (1.3).

Поскольку $x \in \overline{X}$, то найдётся вектор $s^* \in R^n$, $\|s^*\| \le 1$ такой, что

$$x + \beta s^* \in X, \text{ где } \beta = \frac{(1-\lambda)}{2\lambda}\varepsilon. \tag{1.6}$$

Возьмём любой вектор $s \in R^n$, $\|s\| \le 1$. Тогда

$$y + \delta s = |\text{ согласно } (1.3)\,| = (1 - \lambda)\, z + \lambda x + \delta\, s =$$

$$= |\text{ согласно обозначению } (1.5)\,| = (1 - \lambda)\,(z + \frac{1}{2}\,\varepsilon\, s) + \lambda\, x =$$

$$= |\ \text{согласно обозначению (1.6)}\ | =$$

$$= (1 - \lambda)\,(z + \frac{1}{2}\,\varepsilon\, s + \frac{1}{2}\,\varepsilon\,(-s^*)) + \lambda\,(x + \beta s^*).$$

Обозначим $s_* = \frac{1}{2}\,s + \frac{1}{2}\,(-s^*)$. Тогда $||\,s_*||\ \leq 1$ и

$$y + \delta\, s = (1 - \lambda)\,(z + \frac{1}{2}\,\varepsilon\, s_*) + \lambda\,(x + \beta s^*).$$

Используя включения (1.4), (1.6) и выпуклость множества X, получим требуемое включение (1.5).

Следствие 1. Пусть множество X является выпуклым и $int\, X \neq \varnothing$. Тогда $int\, X$ является выпуклым множеством.

Утверждение 7. Пусть множество X является выпуклым и $int\, X \neq \varnothing$. Тогда $int\overline{X} = int\, X$.

Доказательство. Поскольку $\overline{X} \supset X$, то $int\overline{X} \supset int\, X$. Покажем обратное включение. Возьмём любой вектор $y \in int\,\overline{X}$. Зафиксируем вектор $z \in int\, X$. Считаем, что $z \neq y$. Рассмотрим точку

$$x = (1 - v)z + vy = y + (1 - v)\,(z - y). \tag{1.7}$$

Отсюда и из того, что $y \in int\overline{X}$ следует, что при всех достаточно близких к единице числах v точка $x \in \overline{X}$. Рассмотрим такое число $v > 1$. Тогда из формулы (7) следует, что $y = (1 - \lambda)\, z + \lambda x$ при $\lambda = v^{-1}$. Отсюда и из теоремы 2 получим, что $y \in int\, X$.

Теорема 3. Пусть множество X является выпуклым. Тогда либо множество X имеет внутреннюю точку, либо оно содержится в подпространстве меньшей размерности, сдвинутом на некоторый вектор.

Доказательство. Зафиксируем точку $y \in X$. Рассмотрим вектора вида $x - y$, $x \in X$. Пусть максимальное число линейно независимых векторов среди них равняется r. Ясно, что $r \leq n$. Пусть линейно независимыми векторами являются $x_i - y$, $i = 1, \ldots, r$.

Рассмотрим случай, когда $r = n$. Тогда для каждого вектора $x \in R^n$ система линейных алгебраических уравнений относительно переменных λ_i, $i = 1, \ldots, n$

$$x - y = \lambda_1\,(x_1 - y) + \lambda_2\,(x_2 - y) + \ldots + \lambda_n\,(x_n - y)$$

имеет единственное решение $\lambda_1(x), \lambda_2(x), \ldots, \lambda_n(x)$, которое линейно зависит от вектора x. Последнее следует из формул Крамера для записи решения системы линейных алгебраических уравнений. Следовательно, функции $\lambda_i(x)$ непрерывно зависят от вектора x. Можем записать

$$x = (1 - \sum_{i=1}^{n} \lambda_i\,(x))\, y + \sum_{i=1}^{n} \lambda_i\,(x)\, x_i. \tag{1.8}$$

Зафиксируем числа $\lambda_i^* > 0$, $i = 0, 1, \ldots, n$, $\lambda_0^* + \lambda_1^* + \lambda_2^* + \ldots + \lambda_n^* = 1$. Поскольку множество X является выпуклым, то точка

$$z = \lambda_0^* y + \lambda_1^* x_1 + \lambda_2^* x_2 + \ldots + \lambda_n^* x_n \in X.$$

Далее, $z - y = \lambda_1^*(x_1 - y) + \lambda_2^*(x_2 - y) + \ldots + \lambda_n^*(x_n - y)$. Следовательно, $\lambda_i^* = \lambda_i(z)$, $i = 1, \ldots, n$. Зафиксируем положительное число ε так, чтобы

$$\varepsilon < \lambda_i^* \text{ для всех } i = 1, \ldots, n \quad \text{и} \quad n\,\varepsilon < \lambda_0^*. \tag{1.9}$$

Тогда из непрерывности функций $\lambda_i(x)$ следует, что существует число $\delta > 0$ такое, что $\lambda_i^* - \varepsilon < \lambda_i(x) < \lambda_i^* + \varepsilon$ для всех номеров $i = 1, \ldots, n$ и для всех векторов $x \in R^n$, $\|x - z\| < \delta$. Отсюда и из неравенств (1.9) получим, что $0 < \lambda_i(x)$ для всех $i = 1, \ldots, n$ и

$$\sum_{i=1}^{n} \lambda_i(x) < \sum_{i=1}^{n} \lambda_i^* + n\,\varepsilon = 1 - \lambda_0^* + n\,\varepsilon < 1.$$

Отсюда и из (8) следует, что $x \in X$ для всех векторов $x \in R^n$, $\|x - z\| < \delta$. Стало быть, $z \in int\, X$.

Рассмотрим случай, когда $r < n$. Построим линейное подпространство

$$E = \{z \in R^n \mid z = \sum_{i=1}^{r} b_i (x_i - y) \colon \forall\, b_i \in R\}. \tag{1.10}$$

Для каждой точки $x \in X$ выполнено включение $x - y \in E$. Следовательно,

$$X \subset y + E = \{x \in R^n \mid x = y + z\colon \forall z \in E\}.$$

Замечание 1. Построенное линейное подпространство (1.10) r-мерно. Множество $(-y) + X \subset E$ содержит точку, являющуюся внутренней относительно подпространства E. Это доказывается, как и выше изложенном случае $r = n$.

Теорема 4. Линейное подпространство (1.10) определяется только множеством X и не зависит от точки $y \in X$ и от векторов $x_i \in X$.

Доказательство. Пусть E_* — линейное подпространство, содержащее множество $(-y) + X$. Тогда оно содержит все вектора $x_i - y$, $i = 1, \ldots, r$. Согласно формуле (1.10), $E \subset E_*$. Стало быть, подпространство E является пересечением всех подпространств, содержащих множество $(-y) + X$. Следовательно, оно не зависит от векторов $x_i \in X$. Покажем, что подпространство E не зависит от точки $y \in X$.

Возьмём точку $w \in X$ и рассмотрим линейное подпространство E^*, которое содержит множество $(-w) + X$. Тогда,

$$(-y + x) = -(-w + \overline{y}) + (-w + x) \in E^*,$$

$$(-w + x) = -(-y + w) + (-y + x) \in E_*$$

для любой точки $x \in X$. Стало быть, $(-y) + X \subset E^*$, $(-w) + X \subset E_*$. Таким образом, совокупность подпространств, содержащих множество $(-y) + X$, и совокупность подпространств, содержащих множество $(-w) + X$, совпадают.

Определение 8. *Несущим подпространством* выпуклого множества X называется пересечение всех линейных подпространств, содержащих множество $(-y) + X$, где y — любая точка из множества X.

Несущее подпространство выпуклого множества X обозначается $Lim\ X$. Размерность несущего подпространства называется размерностью выпуклого множества X и обозначается $dim\ X$.

Теоремы отделимости выпуклых множеств

Теорема 5 (*о строгой отделимости точки от выпуклого множества*). Пусть множество $X \subset R^n$ является выпуклым и замкнутым, а точка $y \in R^n$ ему не принадлежит. Тогда существуют единичный вектор $\psi \in R^n$ и число $\varepsilon > 0$ такие, что

$$<\psi, x> \le <\psi, y> - \varepsilon \quad \text{для всех точек } x \in X. \tag{1.11}$$

Доказательство. Рассмотрим задачу

$$\|x - y\|^2 \to \min,\ x \in X. \tag{1.12}$$

Так как целевая функция принимает неотрицательные значения, то нижняя грань её значений на множестве X является конечным числом, которое обозначим b. Из определения нижней грани следует, что существует последовательность точек $x_k \in X$ такая, что $\|x_k - y\|^2 \to b$. Отсюда следует, что последовательность точек x_k является ограниченной. Следовательно, из неё можно выделить сходящуюся подпоследовательность. Не вводя новых обозначений, считаем, что сама последовательность точек $x_k \to z$. Из замкнутости множества X следует, что $z \in X$. Далее, $\|z - y\|^2 = b$. Поскольку точка $y \notin X$, то число $b > 0$.

Возьмём любую точку $x \in X$. Из выпуклости множества X следует, что точки $\lambda x + (1 - \lambda) z \in X$ при всех $0 \le \lambda \le 1$. Далее, функция

$$f(\lambda) = \|\lambda x + (1 - \lambda) z - y\|^2 =$$

$$=\|z-y\|^2+2\lambda<x-z,z-y>+\lambda^2\|x-z\|^2\ \lambda^2$$

при $0 \le \lambda \le 1$достигает минимум в точке $\lambda = 0$. Поэтому

$$\lim_{\lambda\to 0+}\frac{f(\lambda)-f(0)}{\lambda}\ge 0 \Rightarrow <x-z,\ z-y>\ge 0.$$

Отсюда следует, что для ненулевого вектора $\psi_* = y - z$ выполнено неравенство

$$<x,\psi_*>\le<z,\psi_*>=<y,\psi_*>-<y-z,\psi_*>=<\overline{y},y_*>-\|\psi_*\|^2.$$

Следовательно, единичный вектор $\psi = \dfrac{\psi_*}{\|\psi_*\|}$ и число $\varepsilon = \|\psi_*\|$ удовлетворяют неравенству (1.11).

Теорема 6 (*об отделимости точки от выпуклого множества*). Пусть множество $X \subset R^n$ является выпуклым, а точка $y \in R^n$ ему не принадлежит. Тогда существует единичный вектор $y \in R^n$ такой, что

$$<\psi,x>\le<\psi,y>\text{ для всех точек } x\in X. \tag{1.13}$$

Доказательство. Пусть множество X не содержит внутренней точки. Возьмём $w \in X$. Тогда множество $(-w) + X$ лежит в несущем подпространстве $Lim\ X$, размерность которого меньше n. Существует единичный вектор $y \in R^n$ такой, что $<\psi, x> = 0$ для всех точек $x \in Lim\ X$. Можно считать, что $<\psi, y> \ge 0$ (иначе возьмём вектор $(-\psi)$). Тогда условие (1.13) будет выполнено.

Пусть множество X имеет внутреннюю точку. Если $y \notin \overline{X}$, то, используя выпуклость и замкнутость замыкания $\overline{X}$, из теоремы 1 получим существование требуемого вектора ψ.

Пусть $y \in \overline{X}$, но $y \notin X$. Возьмём любую точку $z \in int\ X$. Тогда при любом числе $t > 1$ точка $\overline{x} = (1 - t)\ z + t\ y \notin \overline{X}$. В самом деле, если это не так, то из условий $x \in \overline{X}$, $z \in int\ X$ и из равенства $y = (1 - \lambda)\ z + \lambda x$ при $\lambda = t^{-1} \in (0, 1)$ следует, что точка $y \in int\ X$. Получили противоречие. Устремим число $t \to 0$ и построим последовательность $x_k \notin \overline{X}$ и $x_k \to y$. По теореме 5 существуют единичные вектора ψ_k такие, что для всех точек $x \in X$ выполнены неравенства $<\psi_k,x>\le<\psi_k,x_k>$. Переходя, если нужно к сходящейся подпоследовательности, можем считать, что $\psi_k \to \psi$. Поэтому, переходя к пределу в предыдущем неравенстве, получим требуемое неравенство (1.13).

Теорема 7 (*об отделимости выпуклых множеств*). Пусть заданы два выпуклых непересекающихся множества X и Y в пространстве R^n. Тогда существует единичный вектор $\psi \in R^n$ такой, что

$$<\psi,x>\le<\psi,y>\text{ для всех точек } x\in X \text{ и } y\in Y. \tag{1.14}$$

Доказательство. Рассмотрим множество

$$Z = \{z \in R^n \mid z = x - y: x \in X \text{ и } y \in Y\}. \qquad (1.15)$$

Покажем, что множество Z является выпуклым. В самом деле, пусть точки z_1 и z_2 принадлежат множеству Z. Тогда $z_i = x_i - y_i$, где $x_i \in X$ и $y_i \in Y$. Поэтому для числа $\lambda \in (0,1)$ точка

$$z = (1-\lambda)(x_1 - y_1) + \lambda(x_2 - y_2) =$$
$$(1-\lambda)x_1 + \lambda x_2 - ((1-\lambda)y_1 + \lambda y_2) = x - y.$$

Из выпуклости множеств X и Y следует, что точки

$$x = (1-\lambda)x_1 + \lambda x_2 \in X \text{ и } y = (1-\lambda)y_1 + \lambda y_2 \in Y.$$

Стало быть, точка $z \in Z$.

Из того, что $X \cap Y \neq \varnothing$ следует, что нулевой вектор $0 \notin Z$. Применяя теорему 6, найдём единичный вектор $\psi \in R^n$ такой, что $<\psi, z> \leq 0$ для всех точек $z \in Z$. Отсюда и из определения множества Z (1.15) следует, что неравенство $<\psi, x - y> \leq 0$ выполнено для любых точек $x \in X$ и $y \in Y$. Отсюда получим требуемое неравенство (1.14).

Теорема 8 (*о строгой отделимости выпуклых множеств*)**.** Пусть заданы два выпуклых замкнутых непересекающихся множества X и Y в пространстве R^n, причём одно из них ограничено. Тогда существуют единичный вектор $\psi \in R^n$ и число $\varepsilon > 0$ такие, что

$$<\psi, x> \leq <\psi, y> - \varepsilon \text{ для всех точек } x \in X \text{ и } y \in Y. \qquad (1.16)$$

Доказательство. Рассмотрим множество Z (1.15) и покажем, что оно является замкнутым. Рассмотрим в этом множестве сходящуюся последовательность точек $z_k = x_k - y_k$, $x_k \in X$ и $y_k \in Y$. Пусть, например, множество Y является ограниченным. Тогда из последовательности точек y_k можно выделить сходящуюся подпоследовательность. Не вводя новых обозначений, считаем, что $y_k \to y \in Y$. Следовательно, являясь суммой двух сходящихся последовательностей, последовательность точек x_k сходится к некоторой точке $x \in X$. Следовательно, $z_k \to x - y \in Z$.

По условию теоремы точка $0 \notin Z$. Отсюда и из теоремы 6 следует, существуют единичный вектор $\psi \in R^n$ и число $\varepsilon > 0$ такие, что $<\psi, z> \leq -\varepsilon$ для всех точек $z \in Z$. Отсюда и из определения множества Z (1.15) следует, что неравенство $<\psi, x - y> \leq -\varepsilon$ выполнено для любых точек $x \in X$ и $y \in Y$. Отсюда получим требуемое неравенство (1.16).

2. Выпуклые функции и их свойства

Будем рассматривать линейное вещественное пространство E (не обязательно конечномерное). Как и в случае конечномерного пространства, R^n множество $X \subset E$ называется выпуклым, если из того, что точки $x \in X$ и $y \in X,$ а число $0 < \lambda < 1$ следует, что $(1-\lambda)x + \lambda y \in X$.

Определение 1. Функция $f : X \to R$, определённая на выпуклом множестве $X \subset E$, называется *выпуклой*, если для любых точек $x \in X$ и $y \in X$ и для любого числа $0 < \lambda < 1$ выполнено *неравенство Иенсена*:

$$f((1-\lambda)\,x + \lambda y) \le (1-\lambda)\,f(x) + \lambda f(y). \qquad (2.1)$$

Зафиксируем любые точки $x \in X$ и $p \in E$. Рассмотрим функцию одного переменного

$$\varphi\,(t) = f(x + t\,p),\ t \in I,\ I = \{t \in R\colon x + tp \in X\}. \qquad (2.2)$$

Теорема 1. Если множество X является выпуклым, то множество $I \subset R$, определенное формулой (2.2), является непустым и выпуклым.

Доказательство. Точка $t = 0$ принадлежит множеству I. Возьмём любые точки $t_1 \in I$, $t_2 \in I$. Это значит, что точки : $x + t_i\,p \in X$, $i = 1, 2$. Возьмём любое число $0 < \lambda < 1$. Тогда

$$x + ((1-\lambda)\,t_1 + \lambda t_2)\,p = (1-\lambda)\,(x + t_1 p) + \lambda\,(x + t_2\,p) \in X.$$

Следовательно, точка $(1-\lambda)\,t_1 + \lambda\,t_2 \in I$.

Отметим, что выпуклым множеством на прямой является промежуток (отрезок, конечный или бесконечный интервал, конечный или бесконечный полуинтервал).

Теорема 2. Функция $f(x)$ выпукла тогда и только тогда, когда для любых $x \in X$ и $p \in E$ выпуклой является функция (2.2).

Доказательство. Пусть функция $f(x)$ выпукла. Возьмём число $\lambda \in (0,1)$ и точки $t_1, t_2 \in I$. Тогда

$$\varphi\,(\lambda t_1 + (1-\lambda)\,t_2) = f\,(x + (\lambda t_1 + (1-\lambda)\,t_2)\,p\,) =$$
$$= f\,(\lambda(x + t_1 p) + (1-\lambda)\,(x + t_2\,p\,)) \le$$

$$\leq \lambda f(x + t_1 p) + (1 - \lambda) f(x + t_2 p) = \lambda\, \varphi(t_1) + (1 - \lambda)\, \varphi(t_2).$$

Таким образом,

$$\varphi(\lambda t_1 + (1 - \lambda) t_2) \leq \lambda \varphi(t_1) + (1 - \lambda)\, \varphi(t_2). \qquad (2.3)$$

Пусть для любых $x \in X$ и $p \in E$ функция (2.2) является выпуклой. Покажем, что и функция $f(x)$ является выпуклой. Возьмём любые точки $x \in X$ и $y \in X$, а также число $\lambda \in (0,1)$. Тогда, используя неравенство (2.3), будем иметь

$$f((1 - \lambda)x + \lambda y) = f(x + \lambda(y - x)) = \varphi(\lambda 1 + (1 - \lambda)\, 0) \leq$$
$$\leq \lambda\varphi(1) + (1 - \lambda)\, \varphi(0) = (1 - \lambda) f(x) + \lambda f(y).$$

Изучим вначале выпуклые функции одного переменного, определённые на промежутке. В этом случае геометрический смысл неравенства Иенсена

$$f((1 - \lambda)\, x_1 + \lambda\, x_2) \leq (1 - \lambda) f(x_1) + \lambda f(x_2) \qquad (2.4)$$

заключается в том, что кусок графика, соединяющий две любые точки на этом графике, лежит ниже хорды, соединяющей эти точки (рис. 2.1).

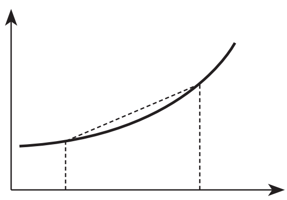


Рис. 2.1

Теорема 2. Пусть функция $f(x)$ является выпуклой на промежутке. Тогда для любых точек $x_0 < x_1 < x_2$ из этого промежутка выполнено неравенство

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \leq \frac{f(x_2) - f(x_0)}{x_2 - x_0} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}. \qquad (2.5)$$

Доказательство. Возьмём число $\lambda = \dfrac{x_2 - x_1}{x_2 - x_0} \in [0,1]$. Тогда

$$x_1 = (1 - \lambda)\, x_2 + \lambda x_0.$$

Отсюда и из неравенства (2.4) получим, что

$$f(x_1) \leq (1 - \lambda) f(x_2) + \lambda f(x_0) \Rightarrow f(x_1) - f(x_0) \leq (1 - \lambda)\, (f(x_2) - f(x_0)).$$

Подставляя в последнее неравенство значение числа λ, получим

$$\frac{f(x_1)-f(x_0)}{x_1-x_0}\le\frac{f(x_2)-f(x_0)}{x_2-x_0}.$$

Далее, $f(x_1)\le(1-\lambda)f(x_2)+\lambda f(x_0)\Rightarrow f(x_2)-f(x_1)\ge\lambda(f(x_2)-f(x_1))$. Подставляя сюда значение числа λ, получим второе неравенство в (2.5)

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}\ge\frac{f(x_2)-f(x_0)}{x_2-x_0}.$$

Замечание 1. На рис. 2.2 даётся геометрическая иллюстрация неравенств (2.5). Для углов α_i, которые образуют соответствующие хорды, выполнены неравенства $\alpha_1\le\alpha_2\le\alpha_3$. Следовательно, $\operatorname{tg}\alpha_1\le\operatorname{tg}\alpha_2\le\operatorname{tg}\alpha_3$. С другой стороны,

$$tg\alpha_1=\frac{f(x_1)-f(x_0)}{x_1-x_0},\quad tg\alpha_2=\frac{f(x_2)-f(x_0)}{x_2-x_0},\quad tg\alpha_3=\frac{f(x_2)-f(x_1)}{x_2-x_1}.$$

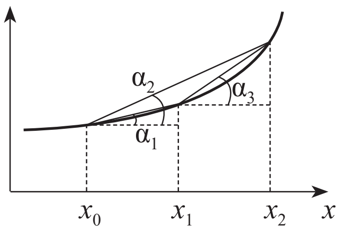


Рис. 2.2

Пример 1. Функция $f(x)=|x|$ является выпуклой на всей числовой оси. В самом деле, из свойства модуля следует, что при $0<\lambda<1$

$$\left|\lambda x_1+(1-\lambda)x_2\right|\le|\lambda x_1|+\left|(1-\lambda)x_2\right|=\lambda|x_1|+(1-\lambda)|x_2|.$$

Теорема 3. Пусть функция $f\colon(a,b)\to R$ является выпуклой. Тогда в каждой точке $x\in(a,b)$ существуют правая и левая производные

$$f'_+(x)=\lim_{\tau\to0+}\frac{f(x+\tau)-f(x)}{\tau},\quad f'_-(x)=\lim_{h\to0-}\frac{f(x+h)-f(x)}{h}\qquad(2.6)$$

и для любых точек $a<x_1<x_2<b$ выполнены неравенства

$$f'_-(x_1)\le f'_+(x_1)\le f'_-(x_2)\le f'_+(x_2).\qquad(2.7)$$

Доказательство. Возьмём число $x\in(a,b)$. Зафиксируем два числа $a_1\in(a,x)$ и $b_1\in(x,b)$. Для достаточно малых чисел $h_2<h_1<0<\tau_1<\tau_2$ выполнены соотношения $a_1<x+h_2<x+h_1<x<x+\tau_1<x+\tau_2<b_1$. Отсюда и из неравенств (2.5) следует, что

$$\frac{f(x)-f(x+h_1)}{-h_1} \le \frac{f(x+\tau_1)-f(x)}{\tau_1} \Rightarrow$$

$$\Rightarrow \frac{f(x+h_1)-f(x)}{h_1} \le \frac{f(x+\tau_1)-f(x)}{\tau_1};$$

$$\frac{f(x)-f(x+h_2)}{-h_2} \le \frac{f(x)-f(x+h_1)}{-h_1} \Rightarrow$$

$$\Rightarrow \frac{f(x+h_2)-f(x)}{h_2} \le \frac{f(x+h_1)-f(x)}{h_1}.$$

Далее

$$\frac{f(x+\tau_1)-f(x)}{\tau_1} \le \frac{f(x+\tau_2)-f(x)}{\tau_2}.$$

Поэтому

$$\frac{f(x+h_2)-f(x)}{h_2} \le \frac{f(x+h_1)-f(x)}{h_1} \le$$

$$\frac{f(x+\tau_1)-f(x)}{\tau_1} \le \frac{f(x+\tau_2)-f(x)}{\tau_2}. \qquad (2.8)$$

Следовательно, функция $\varphi(\tau)=\dfrac{f(x+\tau)-f(x)}{\tau}$ убывает при $\tau\to 0+$ и ограничена снизу. Поэтому существует правая производная $f'_+(x)$ (2.6).

Функция $g(h)=\dfrac{f(x+h)-f(x)}{h}$ возрастает при $h\to 0-$ и ограничена сверху. Поэтому существует левая производная $f'_-(x)$ (2.6).

Из неравенств (2.8) следует $f'_-(x)\le f'_+(x)$. Покажем, что $f'_+(x_1)\le f'_-(x_2)$ при $a<x_1<x_2<b$. Возьмём числа $h<0<\tau$ так, чтобы $x_1<x_1+\tau<x_2+h<x_2$. Тогда из неравенств (2.5) получим, что

$$\frac{f(x_1+\tau)-f(x_1)}{\tau} \le \frac{f(x_2+h)-f(x_1+\tau)}{x_2+h-x_1-\tau} \le \frac{f(x_2)-f(x_2+h)}{-h} =$$

$$\frac{f(x_2+h)-f(x_2)}{h}.$$

Перейдём в этом неравенстве к пределу при $\tau\to 0+$ и при $h\to 0-$. Получим требуемое неравенство $f'_+(x_1)\le f'_-(x_2)$.

Теорема 4 (*критерий выпуклости дифференцируемой функции*). Пусть функция f: $(a, b)\to R$ является дифференцируемой. Тогда она является выпуклой на интервале (a, b) в том и только том случае, если её производная не убывает.

Доказательство. Зафиксируем две точки $a<x_1<x_2<b$. Возьмём достаточно малые положительные числа h_1 и h_2 такие, чтобы $x_1 +$

$h_1 < x_2 - h_2$. Тогда из неравенства (2.5) будем иметь

$$\frac{f(x_1+h_1)-f(x_1)}{h_1} \le \frac{f(x_2)-f(x_2-h_2)}{h_2}.$$

Устремляя в этом неравенстве числа h_1 и h_2 к нулю, получим $f'(x_1) \le f'(x_2)$.

Пусть теперь производная не убывает. Возьмём две точки $a < x_1 < x_2 < b$ рассмотрим функцию

$$\varphi(\lambda) = f((1-\lambda)\,x_1 + \lambda x_2) - (1-\lambda)\,f(x_1) - \lambda f(x_2),\ 0 \le \lambda \le 1. \quad (2.9)$$

Если покажем, что $\varphi(\lambda) \le 0$ при всех $0 \le \lambda \le 1$, то получим требуемое неравенство (2.4). Применяя теорему о дифференцируемости сложной функции, будем иметь $\varphi'(\lambda) = f'((1-\lambda)\,x_1 + \lambda x_2)\,(x_2 - x_1) + f(x_1) - f(x_2) =$

$$= (x_2 - x_1)\left(-\frac{f(x_2)-f(x_1)}{x_2-x_1} + f'\big((1-\lambda)x_1+\lambda x_2\big)\right).$$

Применив теорему Лагранжа, найдём точку $c \in (x_1, x_2)$ такую, что

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} = f'(c).$$

Следовательно, $\varphi'(\lambda) = (x_2 - x_1)\,(f'((1-\lambda)\,x_1 + \lambda\,x_2) - f'(c))$. Отсюда видно, что $\varphi'(\lambda_*) = 0$ при $\lambda_* = \dfrac{c-x_1}{x_2-x_1}$. Так как производная функции $f(x)$ не убывает, то $\varphi'(\lambda) \le 0$ при $(1-\lambda)x_1 + \lambda x_2 \le c$ и $\varphi'(\lambda) \ge 0$ в противном случае. Это значит, что $\varphi'(\lambda) \le 0$ при $0 \le \lambda \le \lambda_*$ и $\varphi'(\lambda) \ge 0$ в противном случае. Стало быть, функция $\varphi(\lambda)$ убывает при $0 \le \lambda \le \lambda_*$ и она возрастает при $\lambda_* \le \lambda \le 1$. Далее, из формулы (2.9) следует, что $\varphi(0) = \varphi(1) = 0$. Следовательно, $\varphi(\lambda) \le 0$ при всех $0 \le \lambda \le 1$ (рис. 2.3).

Теорема 5 (*критерий выпуклости два раза дифференцируемой функции*). Пусть функция $f\colon (a, b) \to R$ два раза дифференцируема. Тогда она является выпуклой на интервале (a, b) в том и только том случае, если её вторая производная $f''(x) \ge 0$ для всех точек $x \in (a, b)$.

Доказательство. Согласно теореме 4 функция $f(x)$ является выпуклой тогда и только тогда, когда её производная $f'(x)$ не убывает. Следовательно, функция $f(x)$ является выпуклой тогда и только тогда, когда её вторая производная $f''(x) \ge 0$ для всех точек $x \in (a, b)$.

Пример 2. Функция $f(x) = x^2$ является выпуклой на всей числовой оси.

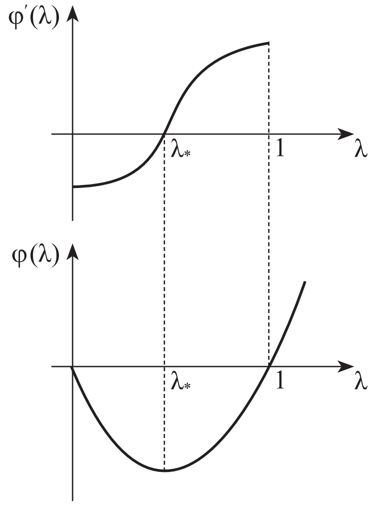


Рис. 2.3

Пример 3. Линейная функция является выпуклой на всей числовой оси.

Теорема 6. Пусть на интервале (a, b) функция $f(x)$ является дифференцируемой и выпуклой (вогнутой) и пусть в точке $c \in (a, b)$ производная $f'(c) = 0$. Тогда

$$f(x) \geq f(a) \qquad (f(x) \leq f(a)) \text{ для всех } x \in (a, b). \qquad (2.10)$$

Доказательство. Рассмотрим случай выпуклой функции. Тогда её производная не убывает. Это значит, что

$$f'(x) \leq f'(c) = 0, \forall x \in (a, c);\ f'(x) \geq f'(c) = 0, \forall x \in (c, b).$$

Стало быть, слева от точки c функция убывает, а справа от этой точки она возрастает. Отсюда получаем неравенство (2.10).

Замечание. Для выпуклых функций точка локального минимума является и точкой *абсолютного минимума*.

Арифметические действия с выпуклыми (вогнутыми) функциями

Пусть функция является выпуклой (вогнутой) на интервале. Тогда при умножении её на неотрицательное число получается выпуклая (вогнутая) функция. В самом деле, при умножении неравенств (2.1) и (2.4) на неотрицательное число их знак сохраняется.

Упражнение. Используя выпуклость множества X, показать, что I является промежутком, то есть если точки $t_1, t_2 \in I$, то при любом числе $\lambda \in [0,1]$ точка $\lambda t_1 + (1 - \lambda) t_2 \in I$.

Замечание. Если функция f определена на всём пространстве R^n, то промежуток I является всей числовой осью.

Упражнение. Показать, что если функция f выпукла, то для любых $\bar{x}_i \in X, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1$ выполнено неравенство $f(\sum_{i=1}^{k} \lambda_i \bar{x}_i) \leq \sum_{i=1}^{k} \lambda_i f(\bar{x}_i)$.

Пусть $F : E \times E \to R$ — билинейная симметричная функция, то есть

$$F(x, y) = F(y, x), \quad F(ax + by, z) = aF(x, z) + bF(y, z)$$

для любых точек x, y, z из пространства E и для любых чисел a и b. Пусть $g : E \to R$ — линейная функция.

Упражнение. Показать, что если функция f вогнута, то для любых $\bar{x}_i \in X, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1$ выполнено неравенство $f(\sum_{i=1}^{k} \lambda_i \bar{x}_i) \geq \sum_{i=1}^{k} \lambda_i f(\bar{x}_i)$.

Упражнение. Показать, что функция $f(\bar{x})$ является вогнутой тогда и только тогда, когда функция $\varphi(t)$ (2.2) является вогнутой по переменной t для любых $\bar{x} \in X, \bar{p} \in R^n$.

Теорема 7 (*критерий выпуклости (вогнутости) дифференцируемой функции*). Пусть функция $f(\bar{x})$ имеет непрерывные частные производные. Тогда она является выпуклой (вогнутой) тогда и только тогда, когда для любых $\bar{x} \in X, \bar{p} \in R^n$ выражение

$$\sum_{i-1}^{n} \frac{f(\bar{x} + t\bar{p})}{\partial x_i} p_i. \tag{2.11}$$

возрастает (убывает) по t.

Доказательство. Функция f выпукла (вогнута) тогда и только тогда, когда функция $\varphi(t)$ (2.2) выпукла (вогнута). По теореме о дифференцируемости сложной функции функция (2.2) дифференцируема и её производная равна

$$\varphi'(t) = \sum_{i-1}^{n} \frac{f(\bar{x} + t\bar{p})}{\partial x_i} p_i .$$

Функция $\varphi(t)$ выпукла (вогнута) тогда и только тогда, когда её производная $\varphi'(t)$ возрастает (убывает) по t.

Теорема 8. Пусть выпуклая (вогнутая) функция $f(\bar{x})$ имеет непрерывные частные производные. Тогда для любых $\bar{x}, \bar{y} \in X$ выполнено равенство

$$f(\bar{y}) - f(\bar{x}) \geq \sum_{i=1}^{n} \frac{f(\bar{x})}{\partial x_i}(y_i - x_i) \left(f(\bar{y}) - f(\bar{x}) \leq \sum_{i=1}^{n} \frac{f(\bar{x})}{\partial x_i}(y_i - x_i) \right). \tag{2.12}$$

Доказательство проведём для выпуклости. Из неравенства (2.1) имеем, что

$$\frac{f(\bar{x} + \lambda(\bar{y} - \bar{x})) - f(\bar{x})}{\lambda} \leq f(\bar{y}) - f(\bar{x}) .$$

Устремим $\lambda \to 0$. Тогда в левой части последнего неравенства будет стоять производная $\varphi'(0)$ функции (2) при $\overline{p} = \overline{y} - \overline{x}$. Эта производная определяется формулой (2.11). Отсюда получим неравенство (2.12).

Теорема 9 (*критерий выпуклости (вогнутости) два раза дифференцируемой функции*). Пусть функция $f(\overline{x})$ имеет непрерывные частные производные второго порядка. Тогда она является выпуклой (вогнутой) в том и только том случае, когда для любых $\overline{x} \in X, \overline{p} \in R^n$ выполнено неравенство

$$\sum_{i,j=1}^{n} \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j} p_i p_j \geq 0 \left(\sum_{i,j=1}^{n} \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j} p_i p_j \leq 0 \right). \tag{2.13}$$

Доказательство. Для функции (2.2) имеем

$$\varphi'(t) = \sum_{i=1}^{n} \frac{f(\overline{x} + t\overline{p})}{\partial x_i} p_i \Rightarrow \varphi''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\overline{x} + t\overline{p})}{\partial x_i \partial x_j} p_i p_j .$$

Функция f выпукла (вогнута) тогда и только тогда, когда функция $\varphi(t)$ выпукла (вогнута) $\Leftrightarrow$ $\varphi''(t) \geq 0$ $(\varphi''(t) \leq 0)$ для любых $\overline{x} \in X, \overline{p} \in R^n$, $\forall t \in I$. Полагая $t = 0$, получим условие (2.13).

Пример. Линейная функция $y = a_0 + \sum_{i=1}^{n} a_i x_i$ является как выпуклой, так и вогнутой. Квадрат от линейной функции $f(x_1, ..., x_2) = ($ ...

$= (a_0 + a_1 x_1 + ... + a_n x_n)^2$ является выпуклой функцией. В самом деле,

$$\frac{\partial f(\overline{x})}{\partial x_j} = 2(a_0 + a_1 x_1 + ... + a_n x_n) a_j \Rightarrow \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j} = 2 a_i a_j .$$

Следовательно, условие (2.13) принимает вид

$$\sum_{i,j=1}^{n} a_i a_j p_i p_j \geq 0.$$

Замечание. Сумма конечного числа выпуклых (вогнутых) функций является выпуклой (вогнутой) функцией.

Критерий положительной (отрицательной) определённости квадратичной симметричной матрицы

Обозначим

$$a_{ij} = \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j}.$$

Тогда из равенства смешанных частных производных получим, что

$$a_{ij} = a_{ji}.$$

Следовательно, матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

является *симметричной*. Неравенства (2.13) примут вид

$$\sum_{i,j=1}^{n} a_{ij} p_i p_j \geq 0 \quad \left(\sum_{i,j=1}^{n} a_{ij} p_i p_j \leq 0 \right).$$

3. Итерационные методы поиска минимума функции многих переменных

Рассматривается задача о нахождении точки абсолютного минимума дифференцируемой функции $f : R^n \to R$

$$f(x) \to \min, x \in R^n. \qquad (з)$$

Для нахождения точек абсолютного минимума задачи (з) используются итерационные процессы

$$x_{k+1} = x_k - \gamma_k P_k\,, \ \gamma_k > 0\,, \ P_k \in R^n. \qquad (3.1)$$

Начальное значение x_0, а также числа γ_k и вектора P_k выбираются по какому-то правилу. Этим правилом и задаётся метод.

Если существует точка $\hat{x}$ = absmin (з), то функция *f* ограничена снизу. Поэтому будем рассматривать ограниченные снизу функции.

Примем следующие обозначения: для $x = (x_1, \ldots, x_n) \in R^n$, $y = (y_1, \ldots, y_n) \in R^n$ положим

$\langle x, y\rangle = x_1y_1 + \ldots + x_ny_n$ — скалярное произведение;

$\|x\| = \langle x, x\rangle^{1/2}$ — норма вектора x;

$f'(x) = \left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right)$ — градиент функции *f*.

Необходимым условием абсолютного минимума $\hat{x}$ = absmin (з) является равенство нулю градиента $f'(\hat{x}) = 0$. Естественно возникает вопрос, когда с помощью процессора (3.1) можно построить последовательность точек, на которых значение градиента функции *f* стремится к нулю. Частичный ответ на этот вопрос содержится в решении следующей задачи.

Задача 1. Пусть функция $f : R^n \to R$ — дифференцируема, ограничена снизу и заданы число $\alpha > 0$ и последовательность точек $x_k \in R^n$ такие, что

$$f(x_k) - f(x_{k+1}) \ge \alpha \|f'(x_k)\|^2, \ k = 1, 2, \ldots \qquad (3.2)$$

Доказать, что

$$\|f'(x_k)\| \to 0 . \qquad (3.3)$$

Указание. Ограниченная снизу числовая последовательность $f(x_k)$ не возрастает. Следовательно, она имеет предел, и поэтому разность, стоящая в левой части (3.2), стремится к нулю.

Пример 1. Рассмотрим функцию $f: R \to R$

$$f(x) = \begin{cases} x^2, x \geq 0 \\ x^3(x+1), x < 0 \end{cases}. \qquad (3.4)$$

График этой функции изображён на рис. 3.1.

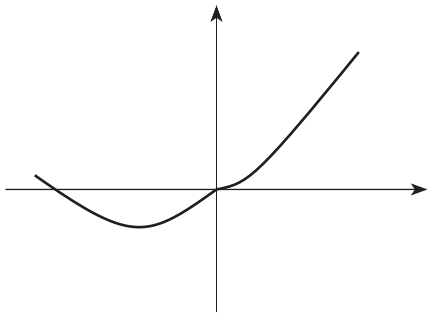


Рис. 3.1

Рассмотрим процесс (3.1), в котором $Pk = x_k$, $1 - \gamma_к = \delta$, $0 < \delta < 1$. В качестве начальной точки возьмём $x_0 = 1$. Тогда $x_k = \delta^k \to 0$ при $k \to \infty$, а неравенство (3.2) примет вид

$$\delta^{2k} - \delta^{2k+2} \geq \alpha \cdot 4\delta^{2k} \Rightarrow \frac{1-\delta^2}{4} \geq \alpha .$$

Таким образом, существует число α, при котором выполнено неравенство (3.2). Следовательно, выполнено условие (3.3). Однако точка $x = 0$ не является точкой минимума функции (3.4).

Рассмотрим теперь выпуклую функцию $f: R^n \to R$. Функция $f: R^n \to R$ называется выпуклой, если выполнено следующее условие: $\forall x \in R^n, \forall y \in R^n, \forall \lambda \in [0,1] \Rightarrow$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) . \qquad (3.5)$$

Если функция f два раза дифференцируема, то обозначим для $x = (x_1, \ldots, x_n)$

$$f''(x)=\begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} \cdots \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \cdots \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} \qquad (3.6)$$

Теорема 1 ([6], с. 21–23). Дважды непрерывно дифференцируемая функция $f: R^n \to R$ является выпуклой в том и только в том случае, если матрица (3.6) положительно определена, то есть

$$\left\langle f''(x)y, y\right\rangle \geq 0, \forall x \in R^n, \forall y \in R^n. \qquad (3.7)$$

Упражнение 1. Доказать эту теорему.

Будем рассматривать дважды непрерывно дифференцируемую функцию $f: R^n \to R$, удовлетворяющую условию сильной выпуклости

$$\exists m > 0 : \left\langle f''(x)y, y\right\rangle \geq m\|y\|^2, \forall x \in R^n, \forall y \in R^n. \qquad (3.8)$$

Упражнение 2. Доказать, что если функция $f(x)$ удовлетворяет условию (3.7), то при любом $m > 0$ условию сильной выпуклости (3.8) удовлетворяет функция $f(x)+m\|x\|^2$.

Задача 2. Доказать, что если дважды непрерывно дифференцируемая функция $f: R^n \to R$ удовлетворяет условию сильной выпуклости (3.8), то

$$f(x) \leq f(y) \Rightarrow \|x-y\| \leq \frac{2}{m}\|f'(y)\|. \qquad (3.9)$$

Указание. Согласно формуле Тейлора, $\exists \lambda \in [0,1]$

$$f(x)-f(y)=\left\langle f'(y), x-y\right\rangle+\frac{1}{2}\left\langle f''(\lambda x+(1-\lambda)y)(x-y),(x-y)\right\rangle.$$

Отсюда, используя неравенство Коши — Буняковского и условие сильной выпуклости (3.8), получим

$$f(x)-f(y) \geq -\|f'(y)\|\,\|x-y\|+\frac{m}{2}\|x-y\|^2.$$

Из этого неравенства следует утверждение (3.9).

Задача 3. Пусть дважды непрерывно дифференцируемая функция f удовлетворяет условию сильной выпуклости (3.8). Доказать, что существует точка absmin(з) и, следовательно, функция f ограничена снизу.

Указание. Зафиксируем точку $y \in R^n$ и рассмотрим множество $Y = \{x \in R^n : f(x) \le f(y)\}$. Из непрерывности функции f следует замкнутость множества Y. Согласно (3.9), множество Y является ограниченным. Следовательно, непрерывная функция f достигает своей нижней грани на множестве Y в некоторой точке $\hat{x} \in Y$. Эта точка является absmin (з).

Теорема 2 ([6], с. 21–23). Непрерывно дифференцируемая функция $f : R^n \to R$ является выпуклой тогда и только тогда, когда

$$f(x) - f(y) \ge \langle f'(y), x - y \rangle, \forall x \in R^n, \forall y \in R^n. \qquad (3.10)$$

Упражнение 3. Доказать, что если непрерывно дифференцируемая функция f является выпуклой, то

$$\hat{x} = abs\min \text{ (з)} \Leftrightarrow f'(\hat{x}) = 0. \qquad (3.11)$$

Задача 4. Пусть дважды непрерывно дифференцируемая функция $f : R^n \to R$ удовлетворяет условию сильной выпуклости (3.8). Тогда для последовательности x_k, удовлетворяющей условию (3.2), выполнено

$$x_k \to \hat{x} = abs\min \text{ (з)}. \qquad (3.12)$$

Указание. Из (3.2) следует, что $f(x_k) \ge f(x_i)$ при $i \ge k$. Поэтому, согласно (3.9),

$$\|x_i - x_k\| \le \frac{2}{m}\|f'(x_k)\|, \forall i \ge k.$$

Отсюда и из (3.3) следует, что последовательность x_k удовлетворяет критерию Коши. Стало быть, $x_k \to \hat{x}$. Из (3.3) и из непрерывности градиента $f'(x)$ следует, что $f'(\hat{x}) = 0$.

Согласно (3.2)

$$\hat{x} = abs\min \text{ (з)}.$$

Градиентные методы

Градиентные методы характеризуются тем, что в формуле (3.1) берётся $P_k = f'(x_k)$ и, стало быть,

$$x_{k+1} = x_k - \gamma_k f'(x_k), \gamma_k > 0. \qquad (3.13)$$

Задача 5. Пусть градиент $f'(x)$ удовлетворяет условию Липшица с константой L.

$$\left\| f'(x) - f'(y) \right\| \le L \left\| x - y \right\|, \forall x \in R^n, \forall y \in R^n . \qquad (3.14)$$

Тогда для процесса (3.13) верна оценка

$$f(x_{k+1}) - f(x_k) \le -\gamma_k (1 - L\gamma_k) \ \left\| f'(x_k) \right\|^2 . \qquad (3.15)$$

Указание. Согласно теореме Лагранжа о конечном приращении функции $f(x_{k+1}) - f(x_k) = f(x_k - \gamma_k f'(x_k)) - f(x_k) =$

$$\left\langle f'(x_k - \lambda\gamma_k f'(x_k)), -\gamma_k f'(x_k) \right\rangle, \lambda \in [0,1]$$

Отсюда и из условия Липшица (3.14) следует, что

$$f(x_{k+1}) - f(x_k) = \left\langle f'(x_k - \lambda\gamma_k f'(x_k)) - f'(x_k), -\gamma_k f'(x_k) \right\rangle$$
$$-\gamma_k \left\langle f'(x_k), f'(x_k) \right\rangle \le L\lambda\gamma_k{}^2 \left\| f'(x_k) \right\|^2 - \gamma_k \left\| f'(x_k) \right\|^2$$

Метод с постоянным шагом

Если численное значение константы L известно, то в процессе (3.13) можно брать $\gamma_k = \gamma, 0 < \gamma < \frac{1}{L}$. Тогда из (3.15) следует (3.2) с $\alpha = \gamma(1 - L\gamma) > 0$.

Метод дробления шага

Применяется тогда, когда численное значение константы Липшица L неизвестно.

Описание метода. Выбираем значения $\gamma > 0, 0 < \varepsilon < 1, 0 < \lambda < 1.$ Пусть значение x_k вычислено.

1. Полагаем $\gamma_k = \gamma$.

2. Вычисляем $x = x_k - \gamma_k f'(x_k)$

3. Проверяем неравенство

$$f(x) - f(x_k) \le -\varepsilon\gamma_k \left\| f'(x_k) \right\|^2 . \qquad (3.16)$$

3.1. Если неравенство (3.16) выполнено, то полагаем $x_{k+1} = x$ и идём к 1.

3.2. Если неравенство (3.16) не выполнено, то полагаем $\gamma_k = \lambda\gamma_k$ и идем к 2.

Обоснование метода. Пусть в процессе счёта приходим к пункту 3.2. Пусть цикл

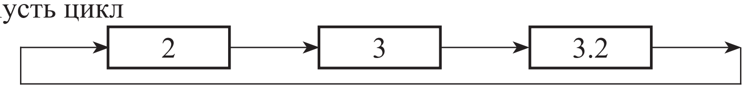


Рис. 3.2

просчитывается $i+1$ раз. Тогда $\gamma_k = \lambda^i \gamma$ и

$$x = x_k - \lambda^i \gamma f'(x_k).$$

Согласно (3.15)

$$f(x) - f(x_k) \le -\lambda^i \gamma (1 - \lambda^i \gamma) \ \|f'(x_k)\|^2.$$

Стало быть, неравенство (3.16) будет выполнено при

$$\varepsilon \lambda^i \gamma \le -\lambda^i \gamma (1 - \lambda^i \gamma) \Leftrightarrow \lambda^i \gamma \le \frac{1-\varepsilon}{L}.$$

Поэтому цикл с рис. 3.2 закончится на конечном шаге, причём

$$\lambda \frac{1-\varepsilon}{L} < \lambda^i \gamma \le \frac{1-\varepsilon}{L} \Rightarrow \lambda \frac{1-\varepsilon}{L} \le \gamma_k \le \frac{1-\varepsilon}{L}.$$

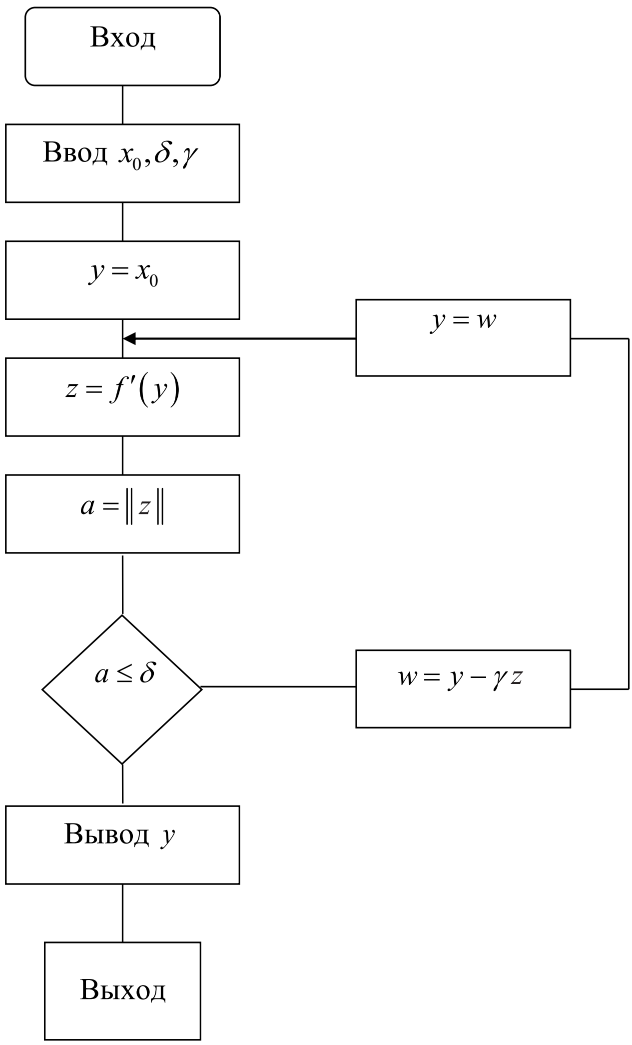


Рис. 3.3. Блок-схема метода с постоянным шагом

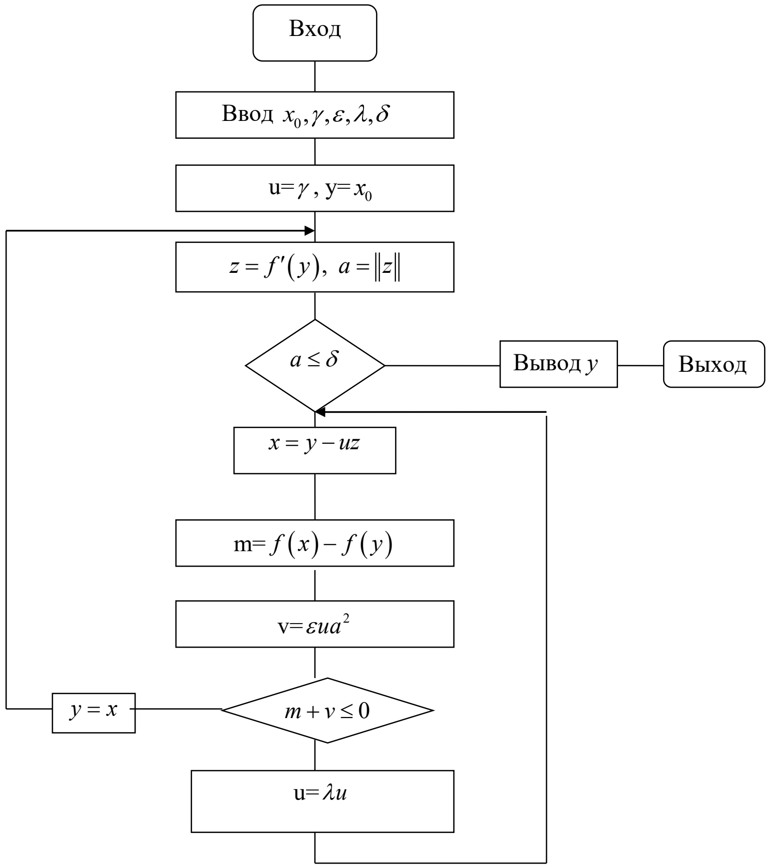


Рис. 3.4. Блок-схема метода с дроблением шага

Отсюда и из неравенства (3.16) следует условие (3.2) с числом

$$\alpha = \varepsilon\lambda \frac{1-\varepsilon}{L} > 0\,.$$

Метод Ньютона с дроблением шага

Будем выбирать вектор $P_k \in R^n$ в формуле (3.1) из условия минимума функции *f* в точке x_{k+1}. Считаем, что функция дважды непрерывно дифференцируема. Применяя формулу Тейлора, будем иметь

$$f(x_{k+1}) - f(x_k) = -\gamma_k \left\langle f'(x_k), P_k \right\rangle + \frac{\gamma_k^{\ 2}}{2} \left\langle f''(x_k) P_k, P_k \right\rangle + O\left(\gamma_k^{\ 2}\right).$$

Вычисляя производную по P_k в правой части и приравнивая её к нулю, будем иметь

$$\gamma_k f''(x_k) P_k = f'(x_k)\,. \qquad (3.17)$$

Допустим, что матрица $f''(x)$ не вырождена. Тогда, вычисляя из (3.17) вектор P_k и подставляя его в (3.1), получим метод Ньютона

$$x_{k+1} = x_k - \left(f''(x_k)\right)^{-1} f'(x_k).$$

Будем рассматривать метод Ньютона с регулировкой шага γ_k, а именно

$$x_{k+1} = x_k - \gamma_k \left(f''(x_k)\right)^{-1} f'(x_k),\ \gamma_k > 0. \tag{3.18}$$

Описание метода. Выбираем значения $\gamma > 0,\ 0 < \varepsilon < 1,\ 0 < \lambda < 1$. Пусть значение x_k вычислено.

1. Полагаем $\gamma_k = \gamma$.

2. Вычисляем $P_k = \left(f''(x_k)\right)^{-1} f'(x_k),\ x = x_k - \gamma_k P_k$.

3. Проверяем неравенство

$$f(x) - f(x_k) \le -\varepsilon\gamma_k \left\langle P_k, f'(x_k)\right\rangle. \tag{3.19}$$

3.1. Если неравенство (3.19) выполнено, то полагаем $x_{k+1} = x$ и идём к 1.

3.2. Если неравенство (3.19) не выполнено, то полагаем $\gamma_k = \lambda\gamma_k$ и идём к 2.

Обоснование метода проведём для дважды непрерывно дифференцируемых функций, удовлетворяющих условию

$$\exists M > m > 0\ :\ m\|y\|^2 \le \left\langle f''(x)y, y\right\rangle \le M\|y\|^2;\ \forall x, y \in R^n. \tag{3.20}$$

Левое неравенство в (3.20) означает условие сильной выпуклости (3.8).

Упражнение 4. Доказать, что правое неравенство в (3.20) выполнено, если все частные производные второго порядка функции f ограничены одним числом для всех $x \in R^n$.

Задача 6. Пусть выполнено условие (3.20). Тогда для обратной матрицы $\left(f''(x)\right)^{-1}$ выполнено

$$\frac{1}{M}\|p\|^2 \le \left\langle \left(f''(x)\right)^{-1} p, p\right\rangle \le \frac{1}{m}\|p\|^2;\ \forall x, p \in R^n. \tag{3.21}$$

Указание. Для положительно определённой матрицы $A = f''(x)$ можно построить ортогональную матрицу T такую, что TT^* = E (E — единичная матрица) и

$$T^*AT=\Lambda=\begin{pmatrix}\lambda_1 & & & & 0\\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & \lambda_n\end{pmatrix}. \tag{3.22}$$

Здесь $\lambda_i>0$ собственные значения матрицы A. Возьмём $y = Tx$, $\forall x\in R^n$. Тогда $\|y\|^2=\langle Tx,Tx\rangle=\langle T^*Tx,x\rangle=\|x\|^2$.

Отсюда и из формулы (3.21) следует

$$m\|x\|^2\le\sum_{i=1}^{n}\lambda_i x_i^{\,2}\le M\|x\|^2\Rightarrow m\le\lambda_i\le M\ .$$

Возьмём $p = Tx$, $\forall x\in R^n$. Тогда

$$\left\langle A^{-1}p,p\right\rangle=\langle T^*A^{-1}Tx,x\rangle=\left\langle \Lambda^{-1}x,x\right\rangle=\sum\nolimits_{i=1}^{n}\lambda_i^{-1}x_i^{\,2}.$$

Следовательно,

$$\frac{1}{M}\|p\|^2=\frac{1}{M}\sum_{i=1}^{n}x_i^2\le\left\langle A^{-1}p,p\right\rangle\le\frac{1}{m}\sum\nolimits_{i=1}^{n}x_i^{\,2}=\frac{1}{m}\|p\|^2.$$

Отметим, что из (3.21) следует невырожденность матрицы $\left(f''(x)\right)^{-1}$.

Обоснование алгоритма. Оценим вначале левую часть неравенства (3.19). Применяя формулу Тейлора, получим

$$f(x)-f(x_k)=f(x_k-\gamma_k P_k)-f(x_k)=$$
$$=\left\langle f'(x_k),-\gamma_k P_k\right\rangle+\frac{1}{2}\left\langle f''(z)(-\gamma_k P_k),(-\gamma_k P_k)\right\rangle.$$

Здесь z — некоторая точка из R^n. Отсюда и из неравенства (3.20) следует, что

$$f(x)-f(x_k)\le-\gamma_k\left\langle f'(x_k),P_k\right\rangle+\frac{M}{2}\gamma_k^{\,2}\left\langle P_k,P_k\right\rangle,$$

$$\left\langle f'(x_k),P_k\right\rangle=\left\langle f''(x_k)P_k,P_k\right\rangle\ge m\left\langle P_k,P_k\right\rangle.$$

Из этих двух неравенств получим оценку

$$f(x)-f(x_k)\le-\gamma_k\left(1-\frac{M}{2m}\gamma_k\right)\left\langle f'(x_k),P_k\right\rangle. \tag{3.23}$$

Пусть в процессе счета приходим к пункту 3.2. Пусть цикл

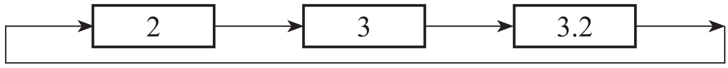


Рис. 3.5

просчитывается $i+1$ раз. Тогда $\gamma_k = \lambda^i \gamma$ и, как следует из (3.23), неравенство (3.19) будет выполнено, если

$$1 - \frac{M}{2m}\lambda^i \gamma \geq \varepsilon \Rightarrow \lambda^i \gamma \leq \frac{1-\varepsilon}{M} 2m = \delta .$$

Стало быть, цикл с рисунка (3.5) закончится на конечном шаге i, причём

$$\lambda\delta \leq \lambda^i \gamma \leq \delta \Rightarrow \lambda\delta \leq \gamma_k \leq \delta .$$

Обозначим

$$\beta = \min\left(\varepsilon\gamma; \varepsilon\lambda\delta\right).$$

Тогда в процессе счёта выполнено неравенство

$$f\left(x_{k+1}\right) - f\left(x_k\right) \leq -\beta\left\langle P_k, f'\left(x_k\right)\right\rangle =$$

$$= -\beta\left\langle \left(f''\left(x_k\right)\right)^{-1} f'\left(x_k\right), f'\left(x_k\right)\right\rangle \leq -\frac{\beta}{M}\left\| f'\left(x_k\right)\right\|^2 .$$

Здесь использовано неравенство (3.19).

Пример. Рассмотрим задачу о нахождении минимума квадратичной функции

$$f\left(x\right) = \frac{1}{2}\left\langle Ax, x\right\rangle + \left\langle c, x\right\rangle; \; A = A^*, \; \left\langle Ax, x\right\rangle > 0 \text{ при } x \neq 0 .$$

Тогда $f'\left(x\right) = Ax + c$, $f''\left(x\right) = A$. Градиентный метод (3.13) примет вид

$$x_{k+1} = x_k - \gamma_k A x_k - \gamma_k c = \left(E - \gamma_k A\right) x_k - y_k c .$$

Метод Ньютона (3.17) записывается в виде

$$x_{k+1} = x_k - \gamma_k A^{-1}\left(A x_k + c\right) = \left(1 - \gamma_k\right) x_k - \gamma_k A^{-1} c .$$

4. Минимизация квадратичной функции

Пусть E является линейным пространством, а $F{:}E{\times}E{\to}R$ — билинейная симметричная функция, то есть

$$F(x, y) = F(y, x),\ F(ax + by, z) = aF(x, z) + bF(y, z)$$

для любых точек x, y, z из пространства E и для любых чисел a и b. Пусть $g{:}E{\to}R$ — линейная функция.

Утверждение 1. Квадратичная функция

$$f(x) = \frac{1}{2} F(x, x) + g(x) \tag{4.1}$$

является выпуклой тогда и только тогда, когда

$$F(p, p) \geq 0 \text{ для любого вектора } p \in E. \tag{4.2}$$

Доказательство. Возьмём любой вектор $p{\in}E$ и рассмотрим функцию

$$\varphi(t) = f(x + tp) = \frac{1}{2}t^2F(p, p) + t\,(F(x, p) + g(p)) + f(x). \tag{4.3}$$

Эта квадратичная функция переменного t является выпуклой тогда и только тогда, когда $\varphi''(t) = F(p, p) \geq 0$.

Утверждение 2. Пусть выполнено условие (4.2). Точка $x_*{\in} E$ является точкой абсолютного минимума квадратичной функции (4.1) тогда и только тогда, когда

$$F(x_*, p) + g(p) = 0 \text{ для любого вектора } p{\in}E. \tag{4.4}$$

Доказательство. Пусть точка $x_*{\in}E$ является точкой абсолютного минимума функции $f(x)$. Возьмём любой вектор $p{\in}E$. Тогда в точке $t = 0$ функция $\varphi(t) = f(x_* + tp)$ достигает абсолютный минимум. Следовательно, $\varphi'(0) = 0$. Отсюда получим условие (4.4).

Пусть в точке $x_*{\in}E$ выполнено условие (4.4). Возьмём любую точку $x{\in}E$ и положим $p = x - x_*$. Для этого вектора p функция $\varphi(t) = f(x_0 + tp)$ является выпуклой. Из условия (4.4) следует, что $\varphi'(0) = 0$. Стало быть,

$$f(x) = \varphi(1) \geq \varphi(0) = f(x_0).$$

Утверждение 3. Пусть квадратичная функция (4.1) ограничена снизу. Тогда, если вектор $p \in E$ удовлетворяет равенству $F(p, p) = 0$, то $F(x, p) + g(p) = 0$ для любой точки $x \in E$.

Доказательство. Существует число N такое, что $f(x) \geq N$ для всех $x \in E$. Отсюда и из (4.3) следует, что $\varphi(t) \geq N$ для всех $t \in R$. В рассматриваемом случае это неравенство принимает вид $\varphi(t) = t(F(x, p) + g(p)) + f(x) \geq N$ для всех $t \in R$. Отсюда следует требуемое равенство.

Утверждение 4. Пусть квадратичная функция (4.1) ограничена снизу числом N. Тогда, если вектор $p \in E$ удовлетворяет неравенству $F(p, p) > 0$, то

$$f(x) \geq N + \frac{\left(F(x,p)+g(p)\right)^2}{2F(p,p)} \text{ для любой точки } x \in E.$$

Доказательство. Минимальное значение функции (4.3) достигается в точке $t = -(F(x, p) + g(p)) / (F(p, p))$ и равно $f(x) - 0{,}5(F(x, p) + g(p))^2 / (F(p, p))$. Отсюда и из неравенства $f(x + tp) \geq N$ для любого числа $t \in R$ следует требуемое неравенство.

Пример 1. Рассмотрим пример, который показывает, что из ограниченности снизу выпуклой квадратичной функции не следует, что она достигает своей нижней грани. Пусть E — пространство непрерывных функций $x{:}[-1, 1] \to R$,

$$F(x(\cdot), y(\cdot)) = \int_{-1}^{1} x(t)y(t)dt, \quad g(x(\cdot)) = \int_{-1}^{1} x(t)e(t)dt, \quad e(t) = sign\ t.$$

Тогда функционал

$$f(x(\cdot)) = 0{,}5F(x(\cdot), x(\cdot)) + g(x(\cdot)) = \frac{1}{2}\int_{-1}^{1} x^2(t)dt + \int_{-1}^{1} x(t)e(t)dt =$$

$$= \frac{1}{2}\int_{-1}^{1}\left(x(t)+e(t)\right)^2 dt - \frac{1}{2}\int_{-1}^{1} e^2(t)dt = \frac{1}{2}\int_{-1}^{1}\left(x(t)+e(t)\right)^2 dt - 1 \geq -1.$$

ограничен снизу. Однако точки абсолютного минимума не существует, поскольку функция $x(t) = e(t)$ является разрывной в точке $t = 0$. Нижняя грань значений этого функционала равна (-1) и достигается на последовательности непрерывных функций

$$x_n(t) = 1 \text{ при } -1 \leq t \leq -n^{-1},\ x_n(t) = -nt$$
$$\text{при } |t| \leq n^{-1},\ x_n(t) = -1 \text{ при } n^{-1} \leq t \leq 1.$$

В самом деле,

$$f(x_n(\cdot)) = \frac{1}{2}\int_{-1}^{1}\left(x_n(t)+e(t)\right)^2 dt - 1 = \frac{1}{2}\int_{-\frac{1}{n}}^{0}\left(-nt-1\right)^2 dt +$$

$$+\ \frac{1}{2}\int_0^{\frac{1}{n}}(-nt+1)^2\,dt\ = \frac{8}{6n}-1\to -1$$

при $n\to\infty$.

Рассмотрим случай, когда пространство $E = R^n$. В этом случае выпуклая квадратичная функция

$$f(\overline{x}) = \frac{1}{2} <F\overline{x},\ \overline{x}> + <\overline{g},\ \overline{x}>. \tag{4.5}$$

Здесь F — симметричная, положительно определенная квадратичная матрица, $\overline{g}$ — фиксированный вектор в R^n.

Градиент этой функции равен

$$f'(\overline{x}) = F\overline{x} + \overline{g}. \tag{4.6}$$

Поскольку функция (4.5) является выпуклой, то точка $\overline{x}_* \in R^n$ является точкой абсолютного минимума функции (4.5) тогда и только тогда, когда её градиент $f'(\overline{x}_*) = \overline{0}$. Таким образом, для нахождения точки минимума квадратичной функции (4.5) нужно решить систему уравнений

$$F\overline{x} + \overline{g} = \overline{0}. \tag{4.7}$$

Теорема 1. Пусть функция (4.5) ограничена снизу. Тогда у неё существует точка абсолютного минимума.

Доказательство. Допустим, что абсолютного минимума не существует. Тогда из (4.7) получим, что $F\overline{x} + \overline{g} \neq \overline{0}$ для любой точки $\overline{x} \in R^n$. Следовательно, вектор $(-\overline{g})$ не принадлежит линейному подпространству $X = \{\overline{z} = F\overline{x}:\ \overline{x} \in R^n\}$. Из теоремы отделимости точки от выпуклого замкнутого множества следует, что существуют вектор $\overline{w} \in R^n$ и положительное число ε такие, что

$$<F\overline{x},\ \overline{w}> \le <-\overline{g},\ \overline{w}> - \varepsilon$$

для любой точки $\overline{x} \in R^n$. Отсюда, учитывая равенство $<F\overline{x},\ \overline{w}> = <\overline{x},\ F\overline{w}>$, получим, что $F\overline{w} = \overline{0}$ и $<\overline{g},\ \overline{w}> \le -\varepsilon$. Тогда из формулы (4.5) будем иметь

$$f(t\overline{w}) = <Ft\overline{w},\ t\overline{w}> + <\overline{g},\ t\overline{w}> = t<\overline{g},\ \overline{w}> \le -\varepsilon t.$$

Следовательно, $f(t\overline{w}) \to -\infty$ при $t \to +\infty$. Получили противоречие с условием ограниченности снизу функции (4.5).

Теорема 2. Если существует точка $\overline{x} \in R^n$ и вектор $\overline{p} \in R^n$ такие, что

$$<F\overline{p},\ \overline{p}> = 0 \text{ и } <f'(\overline{x}),\ \overline{p}> \neq 0, \tag{4.8}$$

то функция (4.5) не ограничена снизу.

Доказательство. Из формул (4.5) и (4.6) следует, что

$$f(\overline{x} + \alpha\overline{p}) = f(\overline{x}) + \alpha <f'(\overline{x}),\ \overline{p}> + \frac{\alpha^2}{2} <F\overline{p},\ \overline{p}>. \tag{4.9}$$

Отсюда и из условий (4.8) следует, что $f(\overline{x} + \alpha\overline{p}) \to -\infty$ либо при $\alpha \to -\infty$, либо при $\alpha \to +\infty$.

Отметим ещё, что верна формула

$$f'(\overline{x} + \alpha\overline{p}) = f'(\overline{x}) + \alpha F\overline{p}. \tag{4.10}$$

Из линейной алгебры известно, что если матрица F является симметричной, то существуют n попарно ортогональных собственных векторов этой матрицы. Это значит, что $F\overline{\psi}_i = \lambda_i\overline{\psi}_i, \lambda_i \in R$ — собственные значения матрицы F. Следовательно, $< F\overline{\psi}_i, \overline{\psi}_j >= \lambda_i < \overline{\psi}_i, \overline{\psi}_j >= 0$ при $i \neq j$.

В связи с этим дадим следующее определение.

Определение 1. Векторы $\overline{p}_i \in R^n$, $i = \overline{1, n}$ называются сопряжённым относительно матрицы F базисом в R^n, если они линейно независимы и

$$< F\overline{p}_i, \overline{p}_j >= 0 \text{ при } i \neq j. \tag{4.11}$$

Если известен сопряжённый относительно матрицы F базис, то задача о нахождении точки абсолютного минимума функции (4.5) легко решается. В самом деле, зафиксируем любую точку $\overline{x} \in R^n$. Тогда для любой точки $\overline{y} \in R^n$ верно разложение

$$\overline{y} = \overline{x} + \sum_{i=1}^{n} \alpha_i \overline{p}_i\,,\ \alpha_i \in R^n.$$

Подставим это разложение в формулу (4.5). Получим

$$f(\overline{y}) = f(\overline{x}) + \sum_{i=1}^{n} \alpha_i < f'(\overline{x}), \overline{p}_i > + \frac{1}{2}\sum_{i,j=1}^{n} < F\overline{p}_i, \overline{p}_j > \alpha_i\alpha_j\,.$$

Отсюда и из условия сопряжённости (4.11) следует, что

$$f(\overline{y})=f(\overline{x})+\sum_{i=1}^{n}\alpha_i < f'(\overline{x}),\overline{p}_i > \; + \; \frac{1}{2}\sum_{i=1}^{n} < F\overline{p}_i,\overline{p}_i > \alpha_i^2 . \quad (4.12)$$

Приравняем к нулю производные по переменным α_i. Будем иметь

$$\alpha_i < F\overline{p}_i,\overline{p}_i > + < f'(\overline{x}),\overline{p}_i > = 0 . \quad (4.13)$$

Случай 1. Пусть при некотором номере $i=\overline{1,n}$ выполнены условия

$$< F\overline{p}_i,\overline{p}_i > = 0, \; < f'(\overline{x}),\overline{p}_i > \neq 0 .$$

Тогда из теоремы 2 следует, что функция (4.5) не ограничена снизу.

Случай 2. Пусть при некотором номере $i=\overline{1,n}$ выполнены условия

$$< F\overline{p}_i,\overline{p}_i > = 0, \; < f'(\overline{x}),\overline{p}_i > = 0 .$$

Тогда любое число α_i удовлетворяет равенству (4.13).

Случай 3. Пусть $< F\overline{p}_i,\overline{p}_i > \neq 0$ при некотором номере $i=\overline{1,n}$. Тогда из (4.13) получим, что

$$\alpha_i = -\frac{< f'(\overline{x}),\overline{p}_i >}{< F\overline{p}_i,\overline{p}_i >} . \quad (4.14)$$

Следствие 1. Чтобы квадратичная функция (4.5) была ограничена снизу, необходимо и достаточно, чтобы $< f'(\overline{x}),\overline{p}_i > = 0$ для всех точек $\overline{x}\in R^n$ и тех векторов $\overline{p}_i, i=\overline{1,n}$, для которых $< F\overline{p}_i,\overline{p}_i > = 0$.

Метод сопряжённых направлений отыскания точки абсолютного минимума выпуклой квадратичной функции

Берём в качестве начальной точки любую точку $\overline{x}_1\in R^n$. Если $f'(\overline{x}_1)=\overline{0}$, то нашли точку абсолютного минимума функции $f(\overline{x})$.

Пусть $f'(\overline{x}_1)\neq\overline{0}$. Полагаем

$$\overline{p}_1 = -f'(\overline{x}_1) . \quad (4.15)$$

Пусть $< F\overline{p}_1,\overline{p}_1 > = 0$. Поскольку

$$< f'(\overline{x}_1),\overline{p}_1 > = -\left\|f'(\overline{x}_1)\right\|^2 \neq 0 ,$$

то по теореме 2 функция $f(\bar{x})$ не ограничена снизу.

Пусть $< F\bar{p}_1, \bar{p}_1 >> 0$ и $f'(\bar{x}_1) \neq \bar{0}$. Полагаем

$$\bar{x}_2 = \bar{x}_1 + \alpha_1 \bar{p}_1, \tag{4.16}$$

где число α_1 выбирается из условия минимума по α функции $f(\bar{x}_1 + \alpha\bar{p}_1)$. Приравнивая к нулю производную по α этой функции, получим уравнение $< f'(\bar{x}_1 + \alpha\bar{p}_1), \bar{p}_1 >= 0$ для определения параметра α. Отсюда, используя формулу (4.10), находим, что

$$\alpha_1 = -\frac{< f'(\bar{x}_1), \bar{p}_1 >}{< F\bar{p}_1, \bar{p}_1 >} = \frac{\|f'(\bar{x}_1)\|^2}{< F\bar{p}_1, \bar{p}_1 >} > 0. \tag{4.17}$$

Отметим, что выполнено равенство

$$< \bar{p}_1, f'(\bar{x}_2) >= 0. \tag{4.18}$$

Полагаем

$$\bar{p}_2 = -f'(\bar{x}_2) + \beta_1 \bar{p}_1. \tag{4.19}$$

Потребуем, чтобы выполнялось равенство

$$< F\bar{p}_1, \bar{p}_2 >= 0. \tag{4.20}$$

Отсюда, используя формулу (4.19), находим значение числа β_1

$$\beta_1 = \frac{< f'(\bar{x}_2), F\bar{p}_1 >}{< F\bar{p}_1, \bar{p}_1 >}. \tag{4.21}$$

Покажем, что

$$< f'(\bar{x}_2), \bar{p}_2 >= -\|f'(\bar{x}_2)\|^2. \tag{4.22}$$

В самом деле,

$$< f'(\bar{x}_2), \bar{p}_2 >= < f'(\bar{x}_2), -f'(\bar{x}_2) + \beta_1 \bar{p}_1 >=$$
$$= -\|f'(\bar{x}_2)\|^2 + \beta_1 < f'(\bar{x}_2), \bar{p}_1 >.$$

Отсюда, используя формулу (4.18), получим требуемое равенство (4.22).

Если $f'(\bar{x}_2) = \bar{0}$, то нашли точку абсолютного минимума функции $f(\bar{x})$.

Пусть $f'(\overline{x}_2) \neq \overline{0}$. Тогда векторы $\overline{p}_1$ и $\overline{p}_2$ являются линейно независимыми. В самом деле, пусть $c_1\overline{p}_1 + c_2\overline{p}_2 = \overline{0}$. Умножим это равенство на вектор $f'(\overline{x}_2)$. Тогда, используя равенства (4.18) и (4.22), получим, что $c_2 = 0$. Следовательно, $c_1 = 0$.

Продолжаем процесс дальше. Рассмотрим построенную точку $\overline{x}_2 \in R^n$, для которой $f'(\overline{x}_2) \neq \overline{0}$. Отсюда и из формулы (4.22) следует, что $< f'(\overline{x}_2), \overline{p}_2 > \neq 0$. Поэтому, если $< F\overline{p}_2, \overline{p}_2 >= 0$, то по теореме 2 функция $f(\overline{x})$ не ограничена снизу.

Пусть $< F\overline{p}_2, \overline{p}_2 >> 0$ и $f'(\overline{x}_2) \neq \overline{0}$. Полагаем

$$\overline{x}_3 = \overline{x}_2 + \alpha_2 \overline{p}_2, \qquad (4.23)$$

где число α_2 выбирается из условия минимума по α функции $f(\overline{x}_2 + \alpha\overline{p}_2)$. Приравнивая к нулю производную по α этой функции, получим уравнение $< f'(\overline{x}_2 + \alpha\overline{p}_2), \overline{p}_2 >= 0$ для определения параметра α. Отсюда, используя формулу (4.10), находим, что

$$\alpha_2 = -\frac{< f'(\overline{x}_2), \overline{p}_2 >}{< F\overline{p}_2, \overline{p}_2 >} = \frac{\|f'(\overline{x}_2)\|^2}{< F\overline{p}_2, \overline{p}_2 >} > 0. \qquad (4.24)$$

Отметим, что выполнено равенство

$$< \overline{p}_2, f'(\overline{x}_3) >= 0. \qquad (4.25)$$

Полагаем

$$\overline{p}_3 = -f'(\overline{x}_3) + \beta_2 \overline{p}_2. \qquad (4.26)$$

Потребуем, чтобы выполнялось равенство

$$< F\overline{p}_2, \overline{p}_3 >= 0. \qquad (4.27)$$

Отсюда, используя формулу (4.26), находим, что

$$\beta_2 = \frac{< f'(\overline{x}_3), F\overline{p}_2 >}{< F\overline{p}_2, \overline{p}_2 >}. \qquad (4.28)$$

Покажем, что

$$< \overline{p}_1, f'(\overline{x}_3) >= 0. \qquad (4.29)$$

В самом деле, из формул (4.10) и (4.23) следует, что

$$< \overline{p}_1, f'(\overline{x}_3) >= < \overline{p}_1, f'(\overline{x}_2) + \alpha_2 F\overline{p}_2 >= < \overline{p}_1, f'(\overline{x}_2) > + \alpha_2 < \overline{p}_1, F\overline{p}_2 >.$$

Отсюда, используя формулы (4.18) и (4.20), получим требуемое равенство (4.29).

Покажем, что

$$< f'(\overline{x}_3), \overline{p}_3 >= -\left\| f'(\overline{x}_3) \right\|^2 . \tag{4.30}$$

В самом деле, используя формулу (4.26), получим, что

$$< f'(\overline{x}_3), \overline{p}_3 >= \ < f'(\overline{x}_3), -f'(\overline{x}_3) + \beta_2 \overline{p}_2 > .$$

Отсюда, используя формулу (4.25), получим требуемое равенство (4.30).

Докажем равенство

$$< F\overline{p}_1, \overline{p}_3 >= 0 . \tag{4.31}$$

Из формулы (4.26) получим, что

$$< F\overline{p}_1, \overline{p}_3 >= \ < F\overline{p}_1, -f'(\overline{x}_3) + \beta_2 \overline{p}_2 > .$$

Отсюда и из равенства (4.20) следует, что

$$< F\overline{p}_1, \overline{p}_3 >= -< F\overline{p}_1, f'(\overline{x}_3) > . \tag{4.32}$$

Далее, из формул (4.16) и (4.19) следует, что

$$\overline{p}_2 = -f'(\overline{x}_1 + \alpha_1 \overline{p}_1) + \beta_1 \overline{p}_1 .$$

Отсюда и из формулы (4.10) получим, что

$$\overline{p}_2 = -f'(\overline{x}_1) - \alpha_1 F\overline{p}_1 + \beta_1 \overline{p}_1 .$$

Из формулы (4.17) следует, что число α_1>0. Поэтому, учитывая (4.15), получим

$$F\overline{p}_1 = \frac{1}{\alpha_1}\left((\beta_1 + 1)\overline{p}_1 - \overline{p}_2\right).$$

Отсюда и из равенства (4.32) будем иметь

$$-< F\overline{p}_1, \overline{p}_3 >= \frac{1}{\alpha_1} < (\beta_1 + 1)\overline{p}_1 - \overline{p}_2, f'(\overline{x}_3) > .$$

Согласно равенствам (4.25) и (4.29) выражение, стоящее в правой части этой формулы, равно нулю. Поскольку $< F\overline{p}_1, \overline{p}_3 >=< \overline{p}_1, F\overline{p}_3 >$, то равенство (4.31) доказано.

Пусть $f'(\overline{x}_3) \neq \overline{0}$. Покажем, что в этом случае векторы $\overline{p}_1, \overline{p}_2, \overline{p}_3$ являются линейно независимыми. Умножим равенство

$c_1\overline{p}_1 + c_2\overline{p}_2 + c_3\overline{p}_3 = \overline{0}$ на вектор $f'(\overline{x}_3)$. Тогда, учитывая формулы (4.25), (4.29) и (4.30), получим, что $c_3\|f'(\overline{x}_3)\|^2 = 0$. Поскольку $f'(\overline{x}_3) \neq \overline{0}$, то $c_3 = 0$. Из линейной независимости векторов $\overline{p}_1$ и $\overline{p}_2$ следует, что оставшиеся числа c_1 и c_2 также равны нулю.

Допустим, что для числа $i = \overline{3, n-1}$ построены векторы $\overline{x}_1 \in R^n$, …, $\overline{x}_i \in R^n$, $\overline{p}_1 \in R^n$, …, $\overline{p}_i \in R^n$, числа α_1, …, α_{i-1}, β_1, …, β_{i-1} такие, что

$$f'(\overline{x}_s) \neq \overline{0},\ <F\overline{p}_s, \overline{p}_s> > 0,\ s = \overline{1, i};\ <F\overline{p}_s, \overline{p}_j> = 0$$
$$\text{при } s \neq j,\ s, j = \overline{1, i}; \tag{4.33}$$

$$<\overline{p}_s, f'(\overline{x}_j)> = 0,\ j<s,\ j = \overline{1, i-2},\ s = \overline{1, i-1};$$
$$<\overline{p}_i, f'(\overline{x}_i)> = -\|f'(\overline{x}_i)\|^2; \tag{4.34}$$

$$\overline{x}_{s+1} = \overline{x}_s + \alpha_s\overline{p}_s,\quad \overline{p}_{s+1} = -f'(\overline{x}_{s+1}) + \beta_s\overline{p}_s,\ s = \overline{1, i-1}; \tag{4.35}$$

$$\alpha_s = -\frac{<f'(\overline{x}_s), \overline{p}_s>}{<F\overline{p}_s, \overline{p}_s>} = \frac{\|f'(\overline{x}_s)\|^2}{<F\overline{p}_s, \overline{p}_s>},$$

$$\beta_s = -\frac{<F\overline{p}_s, f'(\overline{x}_{s+1})>}{<F\overline{p}_s, \overline{p}_s>},\ s = \overline{1, i-1}. \tag{4.36}$$

Векторы $\overline{p}_1, ..., \overline{p}_i$ являются линейно независимыми.

Рассмотрим построенную точку $\overline{x}_i \in R^n$, для которой $f'(\overline{x}_i) \neq \overline{0}$. Отсюда и из второй формулы в (4.34) следует, что $<f'(\overline{x}_i), \overline{p}_i> \neq 0$. Поэтому, если $<F\overline{p}_i, \overline{p}_i> = 0$, то по теореме 2 функция $f(\overline{x})$ не ограничена снизу.

Пусть $<F\overline{p}_i, \overline{p}_i> > 0$ и $f'(\overline{x}_i) \neq \overline{0}$. Полагаем

$$\overline{x}_{i+1} = \overline{x}_i + \alpha_i\overline{p}_i, \tag{4.37}$$

где число α_i выбирается из условия минимума по α функции $f(\overline{x}_i + \alpha\overline{p}_i)$. Приравнивая к нулю производную по α этой функции, получим уравнение $<f'(\overline{x}_i + \alpha\overline{p}_i), \overline{p}_i> = 0$ для определения параметра α. Отсюда, используя формулу (4.10), находим, что

$$\alpha_i = -\frac{<f'(\overline{x}_i), \overline{p}_i>}{<F\overline{p}_i, \overline{p}_i>} = \frac{\|f'(\overline{x}_i)\|^2}{<F\overline{p}_i, \overline{p}_i>} > 0. \tag{4.38}$$

Здесь также учтена вторая формула в (4.34). Отметим, что выполнено равенство

$$<\overline{p}_i, f'(\overline{x}_{i+1})>=0\,. \tag{4.39}$$

Полагаем

$$\overline{p}_{i+1}=-f'(\overline{x}_{i+1})+\beta_i\overline{p}_i\,. \tag{4.40}$$

Потребуем, чтобы выполнялось равенство

$$<F\overline{p}_i, \overline{p}_{i+1}>=0\,. \tag{4.41}$$

Отсюда, используя формулу (4.40), находим, что

$$\beta_i=\frac{<f'(\overline{x}_{i+1}), F\overline{p}_i>}{<F\overline{p}_i, \overline{p}_i>}\,. \tag{4.42}$$

Покажем, что

$$<\overline{p}_s, f'(\overline{x}_{i+1})>=0 \ \text{при всех}\ s=\overline{1,i-1}\,. \tag{4.43}$$

В самом деле, из формул (4.10) и (4.37) следует, что

$$<\overline{p}_s, f'(\overline{x}_{i+1})>=\ <\overline{p}_s, f'(\overline{x}_i)+\alpha_i F\overline{p}_i>=\ <\overline{p}_s, f'(\overline{x}_i)>+\alpha_2<\overline{p}_s, F\overline{p}_i>.$$

Отсюда, используя третью формулу (4.33) и первую формулу (4.34), получим требуемое равенство (4.43).

Покажем, что

$$<f'(\overline{x}_{i+1}), \overline{p}_{i+1}>=-\left\|f'(\overline{x}_{i+1})\right\|^2. \tag{4.44}$$

В самом деле, используя формулу (4.40), получим, что

$$<f'(\overline{x}_{i+1}), \overline{p}_{i+1}>=\ <f'(\overline{x}_{i+1}), -f'(\overline{x}_{i+1})+\beta_i\overline{p}_i>.$$

Отсюда и из равенства (4.39) следует требуемая формула (4.44).

Докажем равенство

$$<F\overline{p}_s, \overline{p}_{i+1}>=0 \ \text{при всех}\ s=\overline{1,i-1}\,. \tag{4.45}$$

Из формулы (4.40) получим, что

$$<F\overline{p}_s, \overline{p}_{i+1}>=\ <F\overline{p}_s, -f'(\overline{x}_{i+1})+\beta_i\overline{p}_i>.$$

Отсюда и из последних равенств в (4.33) следует, что

$$<F\overline{p}_s, \overline{p}_{i+1}>=-<F\overline{p}_s, f'(\overline{x}_{i+1})>. \tag{4.46}$$

Далее, из формул (4.35) и (4.10) следует, что

$$\overline{p}_{s+1} = -f'(\overline{x}_s + \alpha_s \overline{p}_s) + \beta_s \overline{p}_s = -f'(\overline{x}_s) + \alpha_s F\overline{p}_s + \beta_s \overline{p}_s . \quad (4.47)$$

Пусть $1 < s < i - 1$. Тогда из второй формулы в (4.35) следует, что

$$-f'(\overline{x}_s) = \overline{p}_s - \beta_{s-1}\overline{p}_{s-1} .$$

Подставим то равенство в формулу (4.47). Получим

$$\overline{p}_{s+1} = (1 + \beta_s)\overline{p}_s - \beta_{s-1}\overline{p}_{s-1} + \alpha_s F\overline{p}_s .$$

Умножим это равенство на вектор $f'(\overline{x}_{i+1})$. Тогда, учитывая равенства (4.29) и (4.43), получим, что $0 = \alpha_s < F\overline{p}_s, f'(\overline{x}_{i+1}) >$. Согласно (4.33) и (4.36) число $\alpha_s > 0$. Стало быть, $< F\overline{p}_s, f'(\overline{x}_{i+1}) >= 0$. Отсюда и из формулы (4.46) получим требуемое равенство (4.45).

Пусть $s = 1$. Тогда из формулы (4.47) следует, что

$$\overline{p}_2 = -f'(\overline{x}_1) + \alpha_1 F\overline{p}_1 + \beta_1 \overline{p}_1 = \overline{p}_2 = \alpha_1 F\overline{p}_1 + (1 + \beta_1)\overline{p}_1 .$$

Здесь учтена формула (4.15). Умножим это равенство на вектор $f'(\overline{x}_{i+1})$. Получим, что $0 = \alpha_s < F\overline{p}_s, f'(\overline{x}_{i+1}) >$. Число $\alpha_1 > 0$. Поэтому, учитывая формулу (4.46), получим требуемое равенство (4.45).

Если $f'(\overline{x}_{i+1}) = \overline{0}$, то нашли точку абсолютного минимума функции $f'(\overline{x})$. Пусть $f'(\overline{x}_{i+1}) \neq \overline{0}$. Покажем, что в этом случае векторы $\overline{p}_1, ..., \overline{p}_{i+1}$ являются линейно независимыми. Умножим равенство $c_1\overline{p}_1 + ... + c_{i+1}\overline{p}_{i+1} = \overline{0}$ на вектор $f'(\overline{x}_{i+1})$. Тогда, учитывая формулы (4.39), (4.43) и (4.44), получим, что $c_{i+1}\left\|f'(\overline{x}_{i+1})\right\|^2 = 0$. Поскольку $f'(\overline{x}_{i+1}) \neq \overline{0}$, то $c_{n+1} = 0$. Из линейной независимости векторов $\overline{p}_1, ..., \overline{p}_i$ следует, что оставшиеся числа $c_1, \ldots, c_i$ также равны нулю.

5. Применение метода множителей Лагранжа при решении экстремальных задач со связями

Начиная с 1755 года Лагранж рассматривал экстремальные задачи с ограничениями и для их решения применял своё правило, согласно которому с помощью неопределённых коэффициентов задача с ограничениями сводится к нахождению стационарных точек в новой задаче без всяких ограничений.

Задано множество X произвольной природы и функции $f_s::X\to R$, $s=\overline{0,n}$. Рассмотрим задачу

$$f_0(x) \to \min, f_i(x)\le 0,\ i=\overline{1,k};\ f_j(x) = 0,\ j=\overline{k+1,m};\ x\in X. \qquad (5.1)$$

Определение 1. Точка $x_*\in X$, удовлетворяющая связям в задаче (5.1), называется точкой абсолютного минимума в задаче (5.1) ($x_* = \text{absmin}$ (5.1)), если для любой точки $x\in X$, удовлетворяющей связям в задаче (5.1), выполнено неравенство $f_0(x_*) \le f_0(x)$.

Для нахождения точек абсолютного минимума в задаче (5.1) используется правило множителей Лагранжа. Для произвольных параметров $\lambda_s \in R, s=\overline{0,m}$ составляется функция Лагранжа

$$L(x,\ \lambda)=\sum_{s=0}^{m}\lambda_s f_s(x)\,. \qquad (5.2)$$

Эти числа λ_s называются *множителями Лагранжа*.

Теорема 1 (достаточные условия абсолютного минимума в задаче со связями). Пусть существуют точка $x_*\in X$, удовлетворяющая связям в задаче (5.1), и набор множителей Лагранжа $\lambda_s \in R, s=\overline{0,m}$ такие, что

1) $\lambda_0 > 0$, $\lambda_i \ge 0$ при $i=\overline{1,k}$ (условия согласования знаков);

2) $\lambda_i f_i(x_*) = 0$, $i=\overline{1,k}$ (условия дополняющей нежёсткости);

3) $L(x_*, \lambda) \le L(x, \lambda)$ для любого $x\in X$ (условие минимума функции Лагранжа).

Тогда эта точка x_* является точкой абсолютного минимума в задаче (5.1).

Доказательство. Возьмём любую точку $x\in X$, удовлетворяющую связям в задаче (5.1). Тогда, используя условия дополняющей нежёсткости и связи типа равенств в задаче (5.1), получим равенство $\lambda_0 f_0(x_*) = L(x_*, \lambda)$. Отсюда и из условия минимума функции Лагранжа следует, что $\lambda_0 f_0(x_*) \le L(x, \lambda)$. Поскольку точка x удовлетворяет связям в задаче (5.1), а множители Лагранжа удовлетворяют условию согласования знаков, то $L(x,\lambda) \le \lambda_0 f_0(x)$. Так как $\lambda_0>0$, то получаем требуемое неравенство $f_0(x_*) \le f_0(x)$.

Пример 1. Распределить массу в n-ступенчатой ракете так, чтобы при заданной начальной массе m_0 и заданной конечной скорости v ракеты полезная масса ракеты была максимальной.

Выведем вначале *уравнение Мещерского*, которое описывает движение ракеты. Ракета является механической системой с переменной массой. Будем считать её материальной точкой переменной массы.

В момент времени t ракета имеет вектор скорости $\overline{v}\ (t)$ и массу $m(t)$. Следовательно, её вектор количества движения равен $\overline{K}\ (t) = m(t)\overline{v}\ (t)$. В момент времени $t + \Delta t$ вектор скорости ракеты и её масса соответственно равны $\overline{v}\ (t + \Delta t)$ и $m(t + \Delta t)$. За время Δt от ракеты отделилась масса, количество которой равно $m(t) - m(t + \Delta t)$, со скоростью $\overline{w}$ относительно корпуса ракеты. Вектор количества движения этой системы в момент времени $t + \Delta t$ равняется

$$\overline{K}\ (t + \Delta t) = m(t + \Delta t)\overline{v}\,(t + \Delta t) + (\overline{v}\,(t + \Delta t) + \overline{w}\,)(m(t) - m(t + \Delta t)).$$

Изменение количества движения равняется импульсу внешней силы, которой является сила тяжести. Таким образом,

$$\overline{K}\ (t + \Delta t) - \overline{K}\ (t) = m(t)\overline{g}\,\Delta t.$$

Здесь вектор $\overline{g}$ определяется внешними силами, приложенными к ракете. Подставляя в это равенство значения векторов $\overline{K}\ (t + \Delta \mathrm{t})$ и $\overline{K}\,(t)$, получим

$$m(t)(\overline{v}\,(t + \Delta t) - \overline{v}\,(t)) + (m(t) - m(t + \Delta t))\ \overline{w} = m(t)\Delta t\overline{g}\,.$$

Разделим это равенство на Δt и перейдём к пределу при $\Delta t \to 0$. Получим уравнение *Мещерского*

$$\overline{v}'\ (t) = \overline{w}\ \frac{m'(t)}{m(t)} + \overline{g}\,. \tag{5.3}$$

Введём вектор

$$\overline{u}\ \ (t) = \overline{w}\ \frac{m'(t)}{m(t)},$$

который называется *реактивной силой*. Тогда уравнение Мещерского записывается в следующем виде:

$$\overline{v}'(t) = \overline{u}\ \ (t) + \overline{g}\,.$$

Пусть ракета движется по прямой при отсутствии внешних сил и величина скорости отделения реактивной массы относительно корпуса ракеты является постоянной w. Тогда её уравнение движения принимает вид

$$v' = -w\,\frac{m'(t)}{m(t)} \Rightarrow v(t) = v(0) + w\ln\frac{m(0)}{m(t)}.$$

Далее, масса ракеты $m(t) = m_T(t) + m_k$. Здесь посредством $m_T(t)$ обозначено оставшееся количество массы топлива в момент времени t, а m_k — неизменяемая часть массы ракеты. Поэтому максимальная скорость, которую достигнет ракета, равна

$$v_k = v_0 + w\,\ln\frac{m_k + m_T}{m_k}, \tag{5.4}$$

где $v_0 = v(0)$ — начальная скорость ракеты, $m_T = m_T(0)$ — начальный запас топлива. Формула (5.4) является формулой *Циолковского*.

Перейдём к решению задачи. Обозначим через m_i суммарную начальную массу топлива и корпуса i-й ступени ракеты. Тогда её масса корпуса равна $q_i m_i$, а начальная масса топлива равна $(1 - q_i)\, m_i$, где число $0 < q_i < 1$.

Пусть m_* — полезная масса ракеты. Тогда начальная масса всей ракеты равна $m_0 = m_* + m_1 + m_2 + \ldots + m_n$. Вначале работает двигатель первой ступени. После того как топливо первой ступени всё израсходовано, оставшаяся масса ракеты равна $m_* + q_1 m_1 + m_2 + \ldots + m_n$. Из формулы Циолковского получим, что приобретённая скорость ракеты равна

$$v_1 = v_0 + w_1 \ln\frac{m_0}{m_* + q_1 m_1 + m_2 + \ldots + m_n}.$$

Затем отделяется корпус первой ступени и включается двигатель второй ступени. В результате работы двигателя второй ступени достигается скорость

$$v_2 = v_1 + w_2 \ln\frac{m_* + m_2 + m_3 + \ldots + m_n}{m_* + q_2 m_2 + \ldots + m_n}.$$

Затем отделяется корпус второй ступени и включается двигатель третьей ступени и так далее. В результате работы двигателя n-й ступени достигается скорость

$$v_n = v_{n-1} + w_n \ln\frac{m_* + m_n}{m_* + q_n m_n}.$$

Таким образом,

$$v_n = v_0 + w_n \ln\frac{m_* + m_n}{m_* + q_n m_n} + w_{n-1} \ln\frac{m_* + m_{n-1} + \ldots + m_n}{m_* + q_{n-1} m_{n-1} + m_n} +$$

$$+ w_2 \ln \frac{m_* + m_2 + m_3 + ... + m_n}{m_* + q_2 m_2 + ... + m_n} + w_1 \ln \frac{m_0}{m_* + q_1 m_1 + m_2 + ... + m_n}.$$

Обозначим

$$x_1 = \frac{m_0}{m_* + m_2 + ... + m_n}; \; x_2 = \frac{m_* + m_2 + m_3 + ... + m_n}{m_* + m_3 + ... + m_n}; \; \ldots; \; x_n = \frac{m_* + m_n}{m_*}. \quad (5.5)$$

Эти числа удовлетворяют неравенствам $x_i > 1$, $i = 1, \ldots, n$. Отметим, что, задав числа $x_i > 1$ и величину полезной массы m_*, из формул (5.5) можем восстановить массы степеней ракеты

$$m_n = (x_n - 1)m_*, \; m_{n-1} = (x_{n-1} - 1)(m_* + m_n), \ldots,$$
$$m_1 = (x_1 - 1)\,(m_* + m_n + m_{n-1} + \ldots + m_2).$$

Из формул (5.5) следует, что

$$x_1 x_2 \ldots x_n = \frac{m_0}{m_*}. \quad (5.6)$$

Далее,

$$\frac{m_* + m_n}{m_* + q_n m_n} = \frac{x_n}{1 + q_n(x_n - 1)}, \ldots, \frac{m_* + m_2 + m_3 + ... + m_n}{m_* + m_3 + ... + m_n} =$$

$$\frac{x_2}{1 + q_2(x_2 - 1)},$$

$$\frac{m_0}{m_* + m_2 + ... + m_n} = \frac{x_1}{1 + q_1(x_1 - 1)}.$$

Следовательно,

$$v = v_0 + w_n \ln \frac{x_n}{1 + q_n(x_n - 1)} + w_{n-1} \ln \frac{x_{n-1}}{1 + q_{n-1}(x_{n-1} - 1)} + \ldots +$$

$$+ w_2 \ln \frac{x_2}{1 + q_2(x_2 - 1)} + w_1 \ln \frac{x_1}{1 + q_1(x_1 - 1)}. \quad (5.7)$$

Наша цель заключается в максимизации полезной массы m_*. Отсюда, используя формулу (5.6), переформулируем нашу цель в виде минимизации функции $x_1 x_2 \ldots x_n$. Эта цель равносильна минимизации $ln(x_1 x_2 \ldots x_n) = \sum_{i=1}^{n} \ln x_i$.

Таким образом, имеем следующую задачу:

$$\sum_{i=1}^{n} \ln x_i \to min,\ \sum_{i=1}^{n} a_i \ln \frac{x_i}{1+q_i(x_i-1)} - 1 = 0,\ (x_1, x_2, \ldots, x_n) \in X. \quad (5.8)$$

Здесь обозначено

$$a_i = \frac{w_i}{v - v_0} > 0,\ X = \{(x_1, x_2, \ldots, x_n) : x_i > 1,\ i = 1, \ldots, n\}.$$

Найдём условие на параметры ракеты, при выполнении которых существует такое распределение массы среди её ступеней, что заданная конечная скорость v достигается. Это равносильно тому, что на некотором наборе чисел $x_i > 1$, $i = 1, \ldots, n$ удовлетворяется связь в задаче (5.8). Функция

$$g(x) = \ln \frac{x}{1+q(x-1)},\ 0 < q < 1,$$

строго возрастает, поскольку её производная $g'(x) = \dfrac{1-q}{x(1+q(x-1))} > 0.$

Далее, $g(1) = 0$ и $g(x) \to -\ln q$ при $x \to +\infty$. Отсюда получим требуемое условие разрешимости связи в задаче (5.8):

$$-\sum_{i=1}^{n} a_i \ln q_i > 1. \quad (5.9)$$

Запишем для задачи (5.8) функцию Лагранжа с $\lambda_0 = 1$, $\lambda_1 = -\lambda$

$$L(x_1, x_2, \ldots, x_n, \lambda) = \sum_{i=1}^{n} \left(\ln x_i - \lambda a_i \ln \frac{x_i}{1+q_i(x_i-1)}\right) + \lambda. \quad (5.10)$$

Чтобы найти минимальное значение функции (5.10) на множестве X, найдём минимальное значение при $x \geq 1$ у каждой из функций

$$f_i(x) = \ln x - \lambda a_i \ln \frac{x}{1+q_i(x-1)}.$$

Её производная равна

$$f_i'(x) = \frac{1}{x} - \lambda a_i \frac{1-q_i}{x(1+q_i(x-1))} = \frac{(1-\lambda a_i)(1-q_i)+q_i x}{x(1+q_i(x-1))}.$$

При $x \geq 1$ знак этой производной определяется знаком выражения $\varphi_i(x) = (1 - \lambda a_i)(1 - q_i) + q_i x$. Поскольку $\varphi_i(x)$ возрастает с ростом x и $\varphi_i(x) \to +\infty$ при $x \to +\infty$, то, если $\varphi_i(1) \geq 0$, то $\varphi_i(x) > 0$ при $x > 1$. Это выполнено тогда и только тогда, когда $\lambda \le \dfrac{1}{a_i(1-q_i)}$. В этом случае $f_i'(x) > 0$ при $x > 1$ и, следовательно, минимальное значение функция $f_i(x)$ достигает в точке $x(\lambda) = 1$. Пусть $\lambda > \dfrac{1}{a_i(1-q_i)}$. Тогда $\varphi_i(1) < 0$. Из условия $\varphi_i(x) = 0$ находим, что

$$x_i(\lambda) = \frac{(\lambda a_i - 1)(1-q_i)}{q_i} > 1.$$

Таким образом, минимальное значение функции $f_i(x)$ достигается в точке $x_i(\lambda)$, где

$$x_i(\lambda) = \frac{(\lambda a_i - 1)(1-q_i)}{q_i} > 1 \text{ при } \lambda > \frac{1}{a_i(1-q_i)};$$

$$x_i(\lambda) = 1 \text{ при } \lambda \le \frac{1}{a_i(1-q_i)}. \tag{5.11}$$

Подставим функцию $x_i(\lambda)$ в связь в задаче (5.8). Получим уравнение $\varphi(\lambda) = 1$ для определения множителя λ. Здесь обозначено

$$\varphi(\lambda) = \sum_{i=1}^{n} a_i \ln \frac{x_i(\lambda)}{1 + q_i(x_i(\lambda) - 1)}.$$

Эта функция непрерывно зависит от множителя λ. Из формулы (5.10) следует, что

$$\varphi(\lambda) = 0 \text{ при } \lambda \le \min_{1\le i\le n} \frac{1}{a_i(1-q_i)} \text{ и } \varphi(\lambda) = \sum_{i=1}^{n} a_i \ln \frac{\lambda a_i - 1}{\lambda a_i q_i}$$

$$\text{при } \lambda > \max_{1\le i\le n} \frac{1}{a_i(1-q_i)}.$$

Отсюда и из условия (5.9) следует, что $\varphi(\lambda) \to -\sum_{i=1}^{n} a_i \ln q_i > 1$ при $\lambda \to +\infty$. Следовательно, при некотором λ требуемое равенство $\varphi(\lambda) = 1$ будет достигаться.

Рассмотрим случай, когда $a_i = a$ и $q_i = q$ для всех $i = 1, \ldots, n$. Тогда из формул (5.11) следует, что $x_i(\lambda) = x(\lambda)$ для всех $i = \overline{1, n}$, где

$$x(\lambda) = \frac{(\lambda a - 1)(1-q)}{q} > 1 \text{ при } \lambda > \frac{1}{a(1-q)} \text{ и } x(\lambda) = 1 \text{ при } \lambda \le \frac{1}{a(1-q)}.$$

Далее,

$$\varphi(\lambda) = 0 \text{ при } \lambda \le \frac{1}{a(1-q)} \text{ и } \varphi(\lambda) = na\ln\frac{\lambda a-1}{\lambda aq} \text{ при } \lambda > \frac{1}{a(1-q)}.$$

Следовательно, равенство $\varphi(\lambda) = 1$ будет достигаться при

$$\frac{\lambda a-1}{\lambda aq} = e^{\frac{1}{na}} \Rightarrow \lambda = \frac{1}{a(1-qe^{\frac{1}{na}})}.$$

Условие разрешимости (5.9) в рассматриваемом случае принимает вид $(-na\ln q) > 1$ или

$$e^{-\frac{1}{na}} - q > 0. \tag{5.12}$$

Отсюда получим для оценки необходимого числа ступеней ракеты

$$n > -\frac{1}{a\ln q}.$$

Из неравенства (5.12) следует, что $1 > qe^{\frac{1}{na}}$. Поэтому $a(1-q) > a(1-qe^{\frac{1}{na}}) > 0$. Стало быть,

$$\lambda = \frac{1}{a(1-qe^{\frac{1}{na}})} > \frac{1}{a(1-q)}.$$

Следовательно,

$$x(\lambda) = \frac{(\lambda a-1)(1-q)}{q} = (1-q)\frac{e^{\frac{1}{na}}}{1-qe^{\frac{1}{na}}} = \frac{1-q}{e^{-\frac{1}{na}}-q}.$$

Отсюда и из формулы (5.6) находим оптимальное значение полезной массы

$$m_0 = m_*\left(\frac{1-q}{e^{-\frac{1}{na}}-q}\right)^n. \tag{5.13}$$

Из формул (5.6) следует, что оптимальное распределение массы среди ступеней ракеты удовлетворяет следующим равенствам:

$$\frac{m_0}{m_*+m_2+\ldots+m_n} = \frac{m_*+m_2+m_3+\ldots+m_n}{m_*+m_3+\ldots+m_n} = \frac{m_*+m_n}{m_*}.$$

Вычислим значение предела выражения (5.13) при числе ступеней $n\to\infty$. Обозначим

$$T=\frac{1}{n},\ \varphi(T)=\left(\frac{e^{-\frac{T}{a}}-q}{1-q}\right)^{\frac{1}{T}}.$$

Поэтому, $\lim\limits_{n\to\infty}\left(\dfrac{1-q}{e^{-\frac{1}{na}}-q}\right)^{n}=\lim\limits_{T\to0}\dfrac{1}{\varphi(T)}$. Вычислим

$$\lim_{T\to0}\ln\varphi(T)=\lim_{T\to0}\frac{1}{T}\left(\ln(e^{-\frac{T}{a}}-q)-\ln(1-q)\right)=\left|\frac{0}{0}\right|=$$

$$\lim_{T\to0}\frac{e^{-\frac{T}{a}}(-\frac{1}{a})}{e^{-\frac{T}{a}}-q}=-\frac{1}{a(1-q)}.$$

Следовательно, $\lim\limits_{n\to\infty}\left(\dfrac{1-q}{e^{-\frac{1}{na}}-q}\right)^{n}=e^{\frac{1}{a(1-q)}}$.

Рассмотрим задачу о запуске спутника на орбиту Земли. Для этого необходимо, чтобы запускаемый на орбиту спутник приобрёл скорость $v\geq7{,}9$ км/с (первая космическая скорость). Рассмотрим случай, когда требуемая конечная скорость $v=10{,}5$ км/с. Начальная скорость $v_0=0$. Для современных видов топлива характерная величина $w=3$ км/с. Тогда $a=\frac{1}{a}=\frac{v}{w}=3{,}5$. Считаем, что 90 % массы каждой ступени ракеты составляет масса топлива. Это значит, что $q=0{,}1$. Тогда формула (5.12) примет вид $m_*=m_0\left(\dfrac{0{,}9}{e^{-\frac{3{,}5}{n}}-0{,}1}\right)^{n}$. Возьмём m_*, равную одной тонне. Тогда требуемая масса n ступенчатой ракеты равна $m_0=\left(\dfrac{0{,}9}{e^{-\frac{3{,}5}{n}}-0{,}1}\right)^{n}$.

Составим таблицу зависимости требуемой массы от числа ступеней:

Число ступеней, n	2	3	4	5
Масса, m_0	149	77	65	60

Отсюда видно, что целесообразно иметь три ступени вместо двух. Использование четырёх и более ступеней затруднительно в силу сложности и больших затрат на создание более сложных двигателей.

6. Правило множителей Лагранжа в гладкой конечномерной задаче на условный экстремум

В этом параграфе рассматривается применение правила множителей Лагранжа для решения гладких задач на условный экстремум в конечномерных пространствах. Доказывается теорема о необходимом условии локального минимума для этого класса задач.

Рассмотрим конечномерную задачу с ограничениями

$$f_0(\bar{x}) \to \min, f_i(\bar{x}) = 0,\ i = \overline{1,\nu},\ f_j(\bar{x}) \le 0,\ j = \overline{\nu+1,m},\ \bar{x} \in X\ . \quad (6.1)$$

Здесь множество $X \subset R^n$, а функции $f_s : X \to R$, $s = \overline{1,m}$.

Обычно ищутся точки *локального минимума* в задаче (6.1), а затем среди них отыскивается решение задачи (6.1).

Определение 1. Точка $\bar{x}^* \in X$ является точкой *локального минимума* в задаче (6.1), если она удовлетворяет всем связям этой задачи и у неё существует такая окрестность S, что для любой точки $\bar{x} \in X \bigcap S$, также удовлетворяющей всем связям в задаче (6.1), выполнено неравенство $f_0(\bar{x}^*) \le f_0(\bar{x})$.

Запишем функцию *Лагранжа* для задачи (6.1)

$$L(\bar{x},\bar{\lambda}) = \sum_{s=0}^{m} \lambda_s f_s(\bar{x}). \quad (6.2)$$

Теорема 1. Пусть точка $\bar{x}^* \in X$ локального минимума в задаче (6.1) является внутренней точкой множества X, а все функции имеют непрерывные частные производные по всем переменным в окрестности этой точки. Тогда существует *ненулевой* набор множителей Лагранжа такой, что выполнены следующие условия:

1) $\lambda_0 \ge 0, \ldots, \lambda_{\nu+1} \ge 0, \ldots, \lambda_m \ge 0$ (*условие согласования знаков*);

2) $\lambda_j f_j(\bar{x}^*) = 0$, $j = \nu+1, \ldots, m$ (*условия дополняющей нежёсткости*);

3) $\dfrac{\partial L(\bar{x}^*, \bar{\lambda})}{\partial x_q} = 0$, $q = 1, \ldots, n$ (*условия стационарности*).

Замечание 1. Как следует из вида функции Лагранжа (6.2) в векторной форме, условия стационарности можно записать в следующем виде:

$$\sum_{s=0}^{m} \lambda_s f_s'\left(\overline{x}^*\right) = \overline{0}\ .$$

Доказательство теоремы. Пусть векторы $f_i'\left(\overline{x}^*\right)$, $i=\overline{1,\nu}$ — линейно зависимы. Тогда существует ненулевой набор чисел λ_i, $i=\overline{1,\nu}$ такой, что

$$\sum_{i=1}^{\nu} \lambda_i f_i'\left(\overline{x}^*\right) = \overline{0}\ .$$

Возьмём эти числа в качестве множителей Лагранжа при $i=\overline{1,\nu}$, а остальные все множители Лагранжа возьмём равными нулю. Тогда полученный набор будет удовлетворять доказываемым условиям.

Пусть векторы $f_i'\left(\overline{x}^*\right)$, $i=\overline{1,\nu}$ являются линейно независимыми. Тогда можно построить векторы $\overline{a}_p \in R^n$, $p=\overline{1,\nu}$, которые удовлетворяют следующим равенствам:

$$\left\langle f_i'\left(\overline{x}^*\right), \overline{a}_p \right\rangle = \delta_{ip} = \begin{cases} 1 \text{ при } i=p, \\ 0 \text{ при } i \neq p, \end{cases} \quad i,p=\overline{1,\nu}. \tag{6.3}$$

Пусть в некоторых связях типа неравенства в точке $\overline{x}^*$ достигаются равенства. Меняя, если нужно, нумерацию, считаем, что

$$f_{\nu+1}(\overline{x}^*)=0,\ldots,f_l(\overline{x}^*)=0, f_{l+1}(\overline{x}^*)<0,\ldots,f_m(\overline{x}^*)<0. \tag{6.4}$$

Рассмотрим в пространстве R^{l+1} два множества:

$$Z=\{\,\overline{z} = (z_0, z_1, \ldots, z_l) : z_s=\langle f_s'\left(\overline{x}^*\right), \overline{b}\,\rangle, s=0, 1, \ldots, 1, \forall \overline{b} \in R^n\},$$

$$U=\{\,\overline{u} = (u_0, 0, \ldots, 0, u_{\nu+1}, \ldots, u_l) : u_0<0, u_{\nu+1}<0, \ldots, u_l<0\}.$$

Эти множества являются выпуклыми. Покажем, что их пересечение пусто. Допустим противное. Тогда найдётся вектор $\overline{b} \in R^n$ такой, что

$$\left\langle f_0'\left(\overline{x}^*\right), \overline{b}\right\rangle < 0; \left\langle f_i'\left(\overline{x}^*\right), \overline{b}\right\rangle = 0, i=\overline{1,\nu}; \left\langle f_j'\left(\overline{x}^*\right), \overline{b}\right\rangle < 0,\ j=\overline{\nu+1,l}\,. \tag{6.5}$$

Рассмотрим функции

$$\psi_i\left(t, y_1, \ldots, y_\nu\right) = f_i\left(\overline{x}^* + t\overline{b} + \sum_{p=1}^{\nu} y_p \overline{a}_p\right), i=\overline{1,\nu}\,.$$

Выполнены равенства $\psi_i(0, 0, \ldots, 0) = 0,\ i=\overline{1,\nu}$. По условию теоремы все функции $f_i(\overline{x})$ имеют непрерывные частные производные по всем переменным x_q в окрестности точки $\overline{x}^*$. Поэтому, по теореме о дифференцируемости сложной функции, все функции ψ_i $(t, y_1, \ldots, y_\nu)$ имеют непрерывные частные производные по всем переменным в окрестности нулевой точки $t = 0, y_1 = 0, \ldots, y_\nu = 0$. Из равенств (6.3) получим, что матрица, составленная из частных производных $\dfrac{\partial\psi_i(0,0,\ldots,0)}{\partial y_p} = d_{ip}$ является единичной и, следовательно, невырожденной. Согласно теореме о неявной функции на некотором интервале $(-c, c)$ определены дифференцируемые функции $y_p(t),\ p=\overline{1,\nu}$, которые удовлетворяют равенствам

$$\psi_i\left(t, y_1(t),\ldots,y_\nu(t)\right) = f_i\left(\overline{x}^* + t\overline{b} + \sum_{p=1}^{\nu} y_p(t)\overline{a}_p\right) = 0,\ i=\overline{1,\nu}. \quad (6.6)$$

Дифференцируя эти равенства по переменной t в точке $t = 0$, получим

$$< f_i'\left(\overline{x}^*\right), \overline{b} > + \sum_{p=1}^{\nu} y_p'(0) < f_i'\left(\overline{x}^*\right), \overline{a}_p >= 0,\ i=\overline{1,\nu}.$$

Отсюда, используя формулы (6.3) и равенства в соотношениях (6.5), получим, что все $y_i'(0) = 0, i=\overline{1,\nu}$.

Обозначим

$$\overline{x}(t) = \overline{x}^* + t\overline{b} + \sum_{p=1}^{\nu} y_p(t)\overline{a}_p.$$

Тогда $\overline{x}(0) = \overline{x}^*$ и $\overline{x}'(0) = \overline{b}$.

Из равенств (6.6) следует, что $f_i(\overline{x}(t)) = 0, i=\overline{1,\nu}$. Пусть индекс s принимает одно из значений $0, \nu+1, \ldots, l$. Верна формула

$$f_s(\overline{x}(t)) = f_s(\overline{x}(0)) + t < f_s'(\overline{x}(0)), \overline{x}'(0) > + o(t) =$$
$$= f_s\left(\overline{x}^*\right) + t < f_s'\left(\overline{x}^*\right), \overline{b} > + o(t).$$

Здесь $o(t) \to 0$ при $t \to 0$. Отсюда и из неравенств в соотношениях (6.5), получим, что при всех достаточно малых положительных числах t выполнены неравенства $f_0(\overline{x}(t)) < f_s\left(\overline{x}^*\right)$ и $f_s(\overline{x}(t)) < 0, s=\overline{\nu+1,l}$. Если же индекс s принимает одно из значений, равное $l + 1, \ldots, m$, то предыдущее неравенство будет также выполнено при всех достаточно малых t, поскольку для этих индексов выполнены неравенства (6.4).

Таким образом, точки $\overline{x}(t)$ при всех достаточно малых положительных числах t удовлетворяют всем связям и $f_0(\overline{x}(t))<f_0(\overline{x}^*)$. Следовательно, точка $\overline{x}^*$ не является точкой локального минимума в задаче (6.1).

Итак, выпуклые множества Z и U не пересекаются. По теореме отделимости выпуклых множеств существует ненулевой вектор $\overline{c} = (l_0, \ldots, l_l)$ такой, что $\langle \overline{c}, \overline{u} \rangle \le \langle \overline{c}, \overline{z} \rangle$ для всех векторов $\overline{u} \in U$ и $\overline{z} \in Z$.

Отсюда, используя вид этих множеств, получим

$$l_0 u_0 + l_{v+1} u_{v+1} + \ldots + l_l u_l \le \left\langle \sum_{s=0}^{l} \lambda_s f_s'(\overline{x}^*), \overline{b} \right\rangle$$

для любых $\overline{b} \in R^n$ и любых $u_0 > 0$, $u_{v+1} > 0$, …, $u_{li} > 0$. Отсюда получим, что

$$l_0 \ge 0, l_{v+1} \ge 0, \; \ldots, \; l_l \ge 0, \; \sum_{s=0}^{l} \lambda_s f_s'(\overline{x}^*) = \overline{0}.$$

Положим $l_{l+1} = 0, \; \ldots, \; l_m = 0$. Тогда условия 1–3, сформулированные в доказываемой теореме, будут выполнены. Таким образом, теорема доказана.

Рассмотрим задачу *при наличии только линейных связей типа равенств и неравенств*

$$f_0(\overline{x}) \to \min, < \overline{g}_i, \overline{x} > -b_i \le 0,$$
$$i = \overline{1,k}, < \overline{g}_j, \overline{x} > -b_j = 0, \; j = \overline{k+1,m}, \overline{x} \in R^n. \tag{6.7}$$

Как видно из записи задачи (6.7), связи типа равенств/неравенств переставлены местами. Это сделано в целях более удобного изложения дальнейшего материала.

Теорема 2 (о неравенстве нулю множителя λ_0 в гладкой конечномерной задаче с линейными связями). Пусть точка $\overline{x}^*$ является решением задачи (6.8). Тогда существует набор множителей Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ такой, что выполнены следующие условия:

1) $\lambda_0 > 0, \lambda_1 \ge 0, \ldots, \lambda_k \ge 0$;

2) $\lambda_i \left(< \overline{g}_i, \overline{x}^* > -b_i \right) = 0, i = \overline{1,k}$;

3) $f_0'\left(\overline{x}^*\right) + \sum_{s=1}^{m} \lambda_s \overline{g}_s = \overline{0}$.

Доказательство. Допустим, что связи типа равенств являются зависимыми. Осуществляя их перенумеровку, можно считать, что каждая связь

$$< \overline{g}_j, \overline{x}^* > -b_j = 0, \; j = \overline{k+q+1, m}$$

является следствием связей

$$< \overline{g}_j, \overline{x}^* > -b_j = 0, \; j = \overline{k+1, k+q}.$$

Тогда задача (6.7) равносильно задаче, в которой стоят связи типа равенств при $j=\overline{k+1,k+q}$. По теореме 1 для этой задачи существует ненулевой набор множителей Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_{k+q}$, удовлетворяющий условиям 1–3 в теореме 2 для этой задачи. Очевидно, что этот же набор множителей вместе с числами $\lambda_{k+q+1}=0, \ldots, \lambda_m=0$ удовлетворяет условиям 1–3 в теореме 2 для задачи (6.7).

Будем считать, что связи типа равенств являются независимыми. Согласно теореме 1 существует ненулевой набор множителей Лагранжа $\lambda_0 \geq 0$, $\lambda_i \geq 0$, $i=\overline{1,k}$, $\lambda_j \in R$, $j=\overline{k+1,m}$ такой, что выполнены условия 2 и 3.

Вначале рассмотрим случай, когда связи типа неравенств отсутствуют. Допустим, что множитель $\lambda_0=0$. Тогда условие 3 принимает вид $\sum\limits_{j=k+1}^{m}\lambda_j\overline{g}_j=\overline{0}$. Отсюда следует, что функция $\sum\limits_{j=k+1}^{m}\lambda_j\left(\left\langle\overline{g}_j,\overline{x}\right\rangle-b_j\right)=const$ для любой точки $\overline{x}\in R^n$. Точка $\overline{x}^*$ удовлетворяет связям в задаче (6.7). Поэтому $\sum\limits_{j=k+1}^{m}\lambda_j\left(\left\langle\overline{g}_j,\overline{x}\right\rangle-b_j\right)=0$ для любой точки $\overline{x}\in R^n$. Стало быть, связи типа равенств зависимы. Получили противоречие.

Доказательство неравенства $\lambda_0>0$ при наличии линейных связей типа неравенств проведём индукцией по числу этих связей. Когда число таких связей равно нулю, утверждение было доказано выше. Пусть утверждение верно для числа связей, равного $k-1$. Докажем его для числа k.

По теореме 1 для задачи (6.7) существует ненулевой набор множителей Лагранжа λ_s, $s=\overline{0,m}$ такой, что выполнены условия 1–3. Допустим, что $\lambda_0=0$. Тогда из условия 3 получим, что

$$\sum_{s=1}^{m}\lambda_s\overline{g}_s=\overline{0}. \tag{6.8}$$

Из этого равенства следует, что $\sum\limits_{s=1}^{m}\lambda_s\left(\left\langle\overline{g}_s,\overline{x}\right\rangle-b_s\right)=const$. Далее, $\sum\limits_{s=1}^{m}\lambda_s\left(\left\langle\overline{g}_s,\overline{x}^*\right\rangle-b_s\right)=\overline{0}$. Поэтому $\sum\limits_{s=1}^{m}\lambda_s\left(\left\langle\overline{g}_s,\overline{x}\right\rangle-b_s\right)=\overline{0}$ для любой точки $\overline{x}\in R^n$. Стало быть, что среди чисел $\lambda_0\geq 0$, $i=\overline{1,k}$ имеется положительное. В противном случае связи типа равенств являются зависимыми.

Пусть, например, $\lambda_1>0$. Рассмотрим задачу

$$f_0\left(\overline{x}\right)\to\min,\left\langle\overline{g}_1,\overline{x}\right\rangle-b_1=0,\left\langle\overline{g}_i,\overline{x}\right\rangle-b_i\leq 0,$$
$$i=\overline{2,k},\left\langle\overline{g}_j,\overline{x}\right\rangle-b_j=0,\ j=\overline{k+1,m}.$$

Из условий $\lambda_1\left(\left\langle \overline{g}_1, \overline{x}^*\right\rangle - b_1\right) = 0$ и $\lambda_1 > 0$ следует, что точка $\overline{x}^*$ удовлетворяет связям в этой задаче и является её решением. В этой задаче число связей типа неравенств равно $k-1$. По индукционному предположению существуют числа λ_s^*, $s=\overline{1,m}$ такие, что выполнены следующие условия:

1) $\lambda_2^* \geq 0$, …, $\lambda_k^* \geq 0$;

2) $\lambda_i^*\left(\left\langle \overline{g}_i, \overline{x}^*\right\rangle - b_i\right) = 0, i=\overline{2,k}$;

3) $f_0'\left(\overline{x}^*\right) + \sum_{s=1}^{m} \lambda_s^* \overline{g}_s = \overline{0}$.

Отсюда видно, что если $\lambda_1^* \geq 0$, то найденный набор чисел λ_s^*, $s=\overline{1,m}$ является искомым.

Пусть $\lambda_1^* < 0$. Рассмотрим числа $\lambda_s^0 = \lambda_s^* - \dfrac{\lambda_1^*}{\lambda_1}\lambda_s$, $s=\overline{1,m}$. Тогда $\lambda_1^0 = 0$. Поскольку $\lambda_1 > 0$, $\lambda_1^* < 0$ и $\lambda_i \geq 0$, $\lambda_i^* \geq 0$ при $i=\overline{2,k}$, то $\lambda_i^0 \geq 0$ при $i=\overline{2,k}$.

При $i=\overline{2,k}$ имеем

$$\lambda_i^0\left(< \overline{g}_i, \overline{x}^* > -b_i\right) = \lambda_i^*\left(< \overline{g}_i, \overline{x}^* > -b_i\right) - \frac{\lambda_1^*}{\lambda_1}\lambda_i\left(< \overline{g}_i, \overline{x}^* > -b_i\right) = 0.$$

Соотношение $\lambda_1^0\left(\left\langle \overline{g}_1, \overline{x}^*\right\rangle - b_1\right) = 0$ следует из равенства $\lambda_1^0 = 0$.

Далее,

$$\sum_{s=1}^{m} \lambda_s^0 \overline{g}_s = \sum_{s=1}^{m}\left(\lambda_s^* - \frac{\lambda_1^*}{\lambda_1}\lambda_s\right)\overline{g}_s = \sum_{s=1}^{m}\lambda_s^*\overline{g}_s - \frac{\lambda_1^*}{\lambda_1}\sum_{s=1}^{m}\lambda_s\overline{g}_s = \sum_{s=1}^{m}\lambda_s^*\overline{g}_s .$$

Здесь использовано равенство (6.9). Отсюда, используя условие оптимальности 3, получим требуемое равенство

$$f_0'\left(\overline{x}^0\right) + \sum_{s=1}^{m}\lambda_s^0\overline{g}_s = f_0'\left(\overline{x}^*\right) + \sum_{s=1}^{m}\lambda_s^*\overline{g}_s = \overline{0}.$$

Пример 1 (*задача потребительского выбора*). Рассмотрим задачу о поведении потребителя на рынке, на котором имеется набор n видов товаров и услуг. Предположим, что приобретённое количество каждого товара и каждой услуги можно охарактеризовать неотрицательным числом. Тогда набор услуг можно задавать вектором $\overline{x} = (x_1, x_2, \ldots, x_n)$, где неотрицательное число x_i определяет количество i-го вида товара или услуги.

Когда потребитель идёт на рынок приобрести необходимые ему товары, то один набор товаров для него предпочтительнее другого. Будем считать, что задана функция $u(\overline{x})$, которая определяет такое предпочтение, а именно: набор $\overline{x}$ *предпочтительнее* набора

$\overline{y}$ тогда и только тогда, когда $u(\overline{x}) \geq u(\overline{y})$. Эта функция называется *функцией полезности.*

Свойства функции полезности

Экономисты требуют, чтобы функция полезности была «правильной», то есть чтобы она удовлетворяла ряду условий.

Условие ненасыщаемости. Ненасыщаемость соответствует неспособности индивида быть полностью удовлетворённым. Экономисты это формулируют следующим образом: *«не имеет значения, как много товаров и услуг находится в индивидуальных владениях, из новых товаров и услуг человек всегда сумеет извлечь дополнительную пользу»*. Это означает, что $u(x_1,...,x_n) > u(y_1,...,y_n)$ при $x_i \geq y_i$, $i = \overline{1,n}$, причём хотя бы для одного номера $j = \overline{1,n}$ стоит строгое неравенство $x_j > y_j$. Из условия ненасыщаемости следует, что при любом Δx_i выполнено неравенство

$$\frac{u(x_1,...,x_{i-1},x_i+\Delta x_i,x_{i+1},...,x_n) - u(x_1,...,x_{i-1},x_i,x_{i+1},.,...,x_n)}{\Delta x_i} > 0 .$$

Эта величина равняется доле приращённой полезности, приходящейся на каждую новую единицу i-го вида товара. Предельное значение этого отношения при $\Delta x_i \to 0$ называется *предельной полезностью* i-го вида товара. Таким образом, предельная полезность i-го вида товара является частной производной функции полезности по переменной x_i. Из условия ненасыщаемости следует, что эта частная производная неотрицательна. Обычно полагают, что

$$\frac{\partial u(x_1,...,x_{i-1,}x_i,x_{i+1},...,x_n)}{\partial x_i} > 0, \quad i = \overline{1,n} . \tag{6.9}$$

Условие убывающей предельной полезности. Это соответствует тому обстоятельству, что с приобретением всё большего количества какого-либо потребительского товара каждая новая единица этого товара обеспечивает всё меньший дополнительный вклад в общую полезность. Это условие называется *первым законом Госсена* по имени немецкого экономиста XIX в. Госсена, на практике изучившего действие этого закона. Это условие называют и *законом убывающей предельной полезности* — предельная полезность i-го товара убывает с увеличением объёма этого товара при постоянных количествах других видов товаров. Это означает, что

$$\frac{\dfrac{\partial u(x_1,...,x_{i-1},x_i+\Delta x_i,x_{i+1},...,x_n)}{\partial x_i} - \dfrac{\partial u(x_1,...,x_{i-1},x_i,x_{i+1},...,x_n)}{\partial x_i}}{\Delta x_i} < 0$$

при любом Δx_i. Предел этого отношения является второй частной производной и, следовательно,

$$\frac{\partial^2 u(x_1,...,x_{i-1},x_i,x_{i+1},...,x_n)}{\partial x_i^2} \leq 0, \quad i = \overline{1,n} .$$

Замечание 2. В случае одного переменного из условия убывающей предельной полезности следует вогнутость функции полезности.

Замечание 3. Иногда от функции полезности требуют, чтобы она отражала следующее свойство: *лучше иметь комбинацию товаров пусть даже в меньших количествах, чем один из этих товаров наименьшей полезности.* Чтобы формализовать это условие, рассмотрим вектор $\overline{e}_i$, у которого единица стоит на i-м месте, а все остальные координаты равны нулю. Тогда любой набор представим в виде $\overline{x} = \sum_{i=1}^{n} x_i \overline{e}_i$. Обозначим через $Q = \sum_{i=1}^{n} x_i$ общее количество товаров. Тогда сформулированное условие принимает следующий вид: $u(\overline{x}) \ge \min_{1 \le i \le n} u(Q\overline{e}_i)$.

Покажем, что, если функция полезности вогнута, то это неравенство выполнено. В самом деле, положим $\lambda_i = \dfrac{x_i}{Q} \ge 0, \sum_{i=1}^{n} x_i = 1$. Тогда $\overline{x} = \sum_{i=1}^{n} \lambda_i (Q\,\overline{e}_i)$. Поэтому

$$u(\overline{x}) = u\left(\sum_{i=1}^{n} \lambda_i (Q\,\overline{e}_i)\right) \ge \sum_{i=1}^{n} \lambda_i\, u(Q\,\overline{e}_i) \ge \min_{1 \le i \le n} u(Q\,\overline{e}_i).$$

Пусть множество X доступных товаров имеет вид

$$X = \{\overline{x} = (x_1, \dots, x_n) : x_i \ge a_i,\ i = 1, \dots, n\}.$$

Здесь a_i — заданные неотрицательные числа. Например, их набор задаёт потребительскую корзину. На множестве X определена функция полезности $u(\overline{x})$. Считаем, что

$$u(\overline{x}) > 0, \text{ если } x_i > a_i \text{ при всех } i = \overline{1,n}, \tag{6.10}$$

а на границе этого множества функция полезности обращается в нуль, то есть

$$u(x_1, \dots, a_i, \dots, x_n) = 0,\ i = \overline{1,n}. \tag{6.11}$$

Далее, считаем, что функция полезности является непрерывной функцией на множестве X, имеющей в каждой внутренней точке этого множества непрерывные частные производные по всем переменным, удовлетворяющие условиям ненасыщаемости (6.9).

Перейдём к задаче потребительского выбора. На рынке существует цена $p_i > 0$ на i-й вид товара, $i = \overline{1,n}$. Потребитель, имея сумму денег

$$Q > a_1p_1 + a_2p_2 + \dots + a_np_n, \tag{6.12}$$

стремится выбрать набор товаров стоимостью, не превосходящей этой суммы, и наибольшей полезности. Имеем задачу

$$-u(\overline{x}) \to \min, x_1p_1 + x_2p_2 + \ldots + x_np_n - Q \leq 0, -x_i + a_i \leq 0,\ i = \overline{1,n}\ . \qquad (6.13)$$

Множество точек $\overline{x}$, которое задаётся ограничениями в задаче (6.13), является замкнутым и ограниченным. Функция полезности является непрерывной. По теореме Вейерштрасса решение $\overline{x}^*$ в задаче (6.13) существует.

Связи в задаче (6.13) являются линейными. Поэтому составим функцию Лагранжа с $\lambda_0 = 1$

$$L(\overline{x},\overline{\lambda}) = -\overline{u}(\overline{x}) + \lambda_1\left(\sum_{i=1}^{n} x_i p_i - Q\right) + \sum_{i=1}^{n}\lambda_{i+1}(-x_i + a_i) \qquad (6.14)$$

и запишем условия оптимальности:

1) $\lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \ldots,\ \lambda_{n+1} \geq 0$;

2) $\lambda_1\left(\sum_{i=1}^{n} x_i^* p_i - Q\right) = 0$, $\lambda_{i+1}(-x_i^* + a_i) = 0,\quad i = \overline{1,n}$;

3) $-\dfrac{\partial u(\overline{x}^*)}{\partial x_i} + \lambda_1 p_i - \lambda_{i+1} = 0, i = \overline{1,n}\cdot$

Из условий (6.10) и (6.11) следует, что все $x_i^* > a_i$. Поэтому из условий дополняющей нежёсткости 2) получим, что $\lambda_{i+1} = 0$ при всех $i = \overline{1,n}$. С учётом этого замечания перепишем условия 1)–3). Будем иметь

1) $\lambda_1 \geq 0$;

2) $\lambda_1\left(\sum_{i=1}^{n} x_i^* p_i - Q\right) = 0$;

3) $\dfrac{\partial u(\overline{x}^*)}{\partial x_i} = \lambda_1 p_i,\ i = \overline{1,n}$.

Покажем, что $\lambda_1 > 0$. Допустим, что $\lambda_1 = 0$. Тогда из условия 3) противоречат условиям ненасыщаемости (6.9).

Обозначим $\lambda_1 = \lambda$. Таким образом, окончательно имеем следующие условия:

$$\sum_{i=1}^{n} x_i^* p_i = Q;\quad \frac{\partial u(\overline{x}^*)}{\partial x_i} = \lambda p_i,\quad i = \overline{1,n};\ \lambda > 0. \qquad (6.15)$$

Точка $\overline{x}^*$ называется *точкой спроса*. Отсюда можно сделать следующие выводы:

1. В точке спроса потребитель расходует все свои деньги.

2. Вторые равенства (6.15) приблизительно можно записать в следующем виде:

$$\frac{u(x_1^*,\ldots,x_{i-1}^*,x_i^* + \Delta x_i, x_{i+1}^*,\ldots,x_n^*) - u(x_1^*,\ldots,x_{i-1}^*,x_i^*,x_{i+1}^*,\ldots,x_n^*)}{p_i \Delta x_i} \approx \lambda\ .$$

В числителе дроби, стоящей слева в этом равенстве, стоит приращение полезности, а в знаменателе — цена приращённого количества товара. Таким образом, доля приращённой полезности, приходящейся на единицу денежных затрат в точке спроса, одинакова по всем видам товаров.

3. Рассмотрим случай, когда функция полезности $u(\overline{x})$ является вогнутой. Возьмём набор $\overline{x}^0 \in X$, который имеет ту же полезность, что и точка спроса. Оценим его стоимость. Функция

$$\varphi(\overline{x}) = u(\overline{x}) - \lambda(\sum_{i=1}^{n} x_i p_i - Q)$$

как сумма вогнутых функций является вогнутой. Из равенств (6.15) следует, что

$$\frac{\partial\varphi(\overline{x}^*)}{\partial x_i} = \frac{\partial u(\overline{x}^*)}{\partial x_i} - \lambda p_i = 0\cdot$$

Эти условия являются условиями глобального максимума вогнутой функции $\varphi(\overline{x})$ на множестве X. Следовательно, $\varphi(\overline{x}) \le \varphi(\overline{x}^*)$ при всех $\overline{x} \in X$. Значит,

$$u(\overline{x}^0) - \lambda(\sum_{i=1}^{n} x_i^0 p_i - Q) \le u(\overline{x}^*) - \lambda(\sum_{i=1}^{n} x_i^* p_i - Q).$$

Поскольку $u(\overline{x}^0) = u(\overline{x}^*)$, получим $\sum_{i=1}^{n} x_i^0 p_i \ge \sum_{i=1}^{n} x_i^* p_i$.

Таким образом, любой набор товаров той же полезности, что и в точке спроса, имеет бóльшую стоимость. Другими словами, индивидууму невыгодно изменять структуру потребления, поскольку всякое такое изменение ухудшает его благосостояние (*второй закон Госсена*).

Пример 2. Выведем формулу для функции полезности, предполагая, что процентный прирост полезности линейно зависит от процентных приростов количеств товаров. С этой целью обозначим $z_i = x_i - a_i$ при всех $i = \overline{1,n}$. Тогда имеем

$$\frac{\Delta u(\overline{z})}{u(\overline{z})} \approx \sum_{i=1}^{n} k_i \frac{\Delta z_i}{z_i}, k_i > 0\cdot$$

Зафиксируем номер $i = \overline{1,n}$ и положим все $\Delta z_j = 0$ для $j \ne i$. Тогда, переходя к пределу при $\Delta z_i \to 0$, получим

$$\frac{\partial u(\overline{z})}{\partial z_i} = k_i \frac{u(\overline{z})}{z_i} \text{ при всех } i = \overline{1,n}\,.$$

Решая это уравнение при $i = 1$ и учитывая, что $u(0, z_2, \ldots, z_n) = 0$, получим, что

$$u(z_1, \ldots, z_n) = f(z_2, \ldots, z_n)\, z_1^{k_1}\,.$$

Функция $f(z_2, \ldots, z_n)$ удовлетворяет уравнению

$$\frac{\partial f(z_2,\ldots,z_n)}{\partial z_i} = k_i \frac{f(z_2,\ldots,z_n)}{z_i} \text{ при всех } i = \overline{2,n}\,.$$

Отсюда находим, что $f(z_2, \ldots, z_n) = \lambda(z_3, \ldots, z_n)\, z_2^{k_2}$. Продолжая эту процедуру, в итоге получим

$$u(x_1,\ldots,x_i,\ldots,x_n) = c(x_1 - a_1)^{k_1}\ldots(x_i - a_i)^{k_i}\ldots(x_n - a_n)^{k_n}.$$

Здесь число $c > 0$. Эта функция называется *функцией Стоуна*.

Функция Стоуна удовлетворяет сформулированным выше требованиям (6.9)−(6.11). Найдём для неё точку спроса. Для этого запишем равенства (6.15)

$$\sum_{j=1}^{n}(x_j^* - a_j)p_j = Q - \sum_{j=1}^{n} a_j p_j, \tag{6.16}$$

$$k_i(x_1^* - a_1)^{k_1}\ldots(x_i^* - a_i)^{k_i - 1}\ldots(x_n^* - a_n)^{k_n} = \lambda p_i,\ i = \overline{1,n}. \tag{6.17}$$

Умножив каждое i-е равенство (6.17) на $x_i^* - a_i$, получим

$$x_i^* - a_i = \frac{k_i}{p_i}\frac{u(\overline{x}^*)}{\lambda},\ i = \overline{1,n}. \tag{6.18}$$

Умножим каждое i-е равенство (6.17) на $x_i^* - a_i$ и просуммируем. Учитывая равенство (6.16), получим

$$u(\overline{x}^*)\sum_{j=1}^{n} k_j = \lambda\left(Q - \sum_{j=1}^{n} a_j p_j\right).$$

Отсюда и из формул (6.18) получим точку спроса

$$x_i^* = a_i + \frac{\gamma_i\left(Q - \sum_{j=1}^{n} a_j p_j\right)}{p_i}, \text{ где обозначено } \gamma_i = \frac{k_i}{\sum_{j=1}^{n} k_j}. \tag{6.19}$$

Другими словами, потребитель вначале приобретает необходимое количество a_i каждого i-го вида товара. Из оставшейся суммы денег $Q - \sum_{j=1}^{n} a_j p_j$ он выделяет на дополнительное приобретение i-го вида товара сумму, равную $\gamma_i\left(Q - \sum_{j=1}^{n} a_j p_j\right)$. На эту сумму по цене p_i он и приобретает дополнительное количество $\dfrac{\gamma_i\left(Q - \sum_{j=1}^{n} a_j p_j\right)}{p_i}$ i-го вида товара.

Замечание 4. Задача потребительского спроса имеет решение. Необходимым условиям оптимальности в этой задаче удовлетворяет единственная точка, определяемая равенствами (6.19). Поэтому она и будет точкой спроса.

Значение функции Стоуна в точке спроса равно

$$U(Q) = u(\overline{x}^*) = \left(\frac{k_1}{p_1}\right)^{k_1}\ldots\left(\frac{k_n}{p_n}\right)^{k_n}\left(\frac{Q - \sum_{j=1}^{n} p_j a_j}{k}\right)^{k},\ k = \sum_{j=1}^{n} k_j. \tag{6.20}$$

Одним из подходов к построению налоговой политики может служить принцип «*равенства жертв*», который заключается в том, что потерянная за счёт сбора налога полезность одинакова для всех. Если обозначить через $T(Q)$ количество суммы налога с годового дохода Q, то из принципа «*равенства жертв*» получим уравнение для определения функции налога

$$U(Q)-U(Q-T(Q))=C=const\ .$$

Применительно к функции (6.20) это равенство примет следующий вид:

$$\left(Q-\sum_{j=1}^{n}p_j a_j\right)^k-\left(Q-T\left(Q\right)-\sum_{j=1}^{n}p_j a_j\right)^k=C.$$

Отсюда находим функцию налога

$$T(Q)=Q-\sum_{j=1}^{n}p_j a_j-\sqrt[k]{\left(Q-\sum_{j=1}^{n}p_j a_j\right)^k-C}\ .$$

Рассмотрим случай, когда величина

$$\mu=\frac{C}{\left(Q-\sum\limits_{j=1}^{n}p_j a_j\right)^k}$$

является небольшой. Тогда, пользуясь разложением $\sqrt[k]{1-\mu}\approx 1-\dfrac{\mu}{k}$, получим

$$T\left(Q\right)=\frac{C}{k}\left(Q-\sum_{j=1}^{n}p_j a_j\right)^{1-k}.$$

7. Теория двойственности

Задано множество X произвольной природы и функции $f_s : X \to \mathrm{R}$, $s = \overline{0,n}$. Рассмотрим задачу

$$f_0(x) \to \min,\ f_i(x) \le 0,\ i = \overline{1,k};\ f_j(x) = 0,\ j = \overline{k+1,m};\ x \in X. \quad (7.1)$$

В пространстве R^m рассмотрим множество

$R_+^m = \{\overline{y} = (y_1, \dots, y_k, y_{k+1}, \dots, y_m) : y_1 \ge 0, \dots, y_k \ge 0, y_{k+1} \in R, \dots, y_m \in R\}$.

При каждом $\overline{y} \in R_+^m$ запишем функцию Лагранжа для задачи (7.1)

$$L(x, \overline{y}) = f_0(x) + \sum_{s=1}^{m} y_s f_s(x), \quad (7.2)$$

с единичным множителем Лагранжа при целевой функции. Положим

$$\varphi(\overline{y}) = \inf_{x \in X} \left(f_0(x) + \sum_{s=1}^{m} y_s f_s(x) \right). \quad (7.3)$$

Функция $\varphi(\overline{y})$ называется *двойственной функцией Лагранжа*, а задача

$$\varphi(\overline{y}) \to \max,\ y_1 \ge 0, \dots, y_k \ge 0, y_{k+1} \in R, \dots, y_m \in R \quad (7.4)$$

называется *двойственной задачей* к задаче (7.1). Задача (7.1) называется прямой задачей.

Теорема 1. Для любой точки $x_+ \in X$, удовлетворяющей связям в задаче (7.1), и для любого набора чисел

$$\overline{y} \in R_+^m = \{\overline{y} \in R^m : y_1 \ge 0, \dots, y_k \ge 0, y_{k+1} \in R, \dots, y_m \in R\}$$

выполнено неравенство

$$\varphi(\overline{y}) \le f_0(x_+). \quad (7.5)$$

Доказательство. Пусть точка $x_+ \in X$ удовлетворяет связям в задаче (7.1), а набор чисел $\overline{y} \in R_+^m$. Тогда $f_0(x_+) \ge f_0(x_+) + \sum_{s=1}^{m} y_s f_s(x) \ge \varphi(\overline{y})$.

Следствие 1. Обозначим

$$D = \{x \in X : f_i(x) \le 0, i = \overline{1,k}; f_j(x) = 0, j = \overline{k+1,m}\}.$$

Тогда из неравенства (7.5) следует, что

$$\sup_{\bar{y} \in R_+^m} \varphi(\bar{y}) \le \inf_{x \in D} f_0(x). \qquad (7.6)$$

Пример 1. Пусть $E = R$. Рассматривается задача

$$-x \to \min,\ x^2 \le 0,\ x \ge 0.$$

Положим $X = \{x \in R\colon x \ge 0\}$, $f_0(x) = -x$, $f_1(x) = x^2$. Получим задачу вида (7.1) при отсутствии связей типа равенств. Тогда при любом $y_1 = y \ge 0$ функция (7.3) равняется

$$\varphi(y) = \inf_{x \ge 0} [f_0(x) + yf_1(x)] = \inf_{x \ge 0} [-x + yx^2] = \begin{cases} -\infty, y = 0 \\ -\dfrac{1}{4y}, y > 0 \end{cases}.$$

В рассматриваемом примере множество $D = \{0\}$. Поэтому

$$\sup_{y \ge 0} \varphi(y) = 0 = f_0(0) = \inf_{x \in D} f_0(x).$$

Следовательно, в рассмотренном примере в соотношении (7.6) стоит равенство.

Пример 2. Пусть $E = R$. Рассматривается задача

$$-2x_1 + x_2 \to \min,\ x_1 + x_2 - 3 \le 0,\ (x_1, x_2) \in X,$$

где множество $X = \{(0,0); (0,4); (4,4); (4,0); (1,2); (2,1)\}$. При любом числе $y \ge 0$

$$\varphi(y) = \inf_{(x_1, x_2) \in X} [-2x_1 + x_2 + y(x_1 + x_2 - 3)] =$$

$$= \min(-3y; y + 4; 5y - 4; y - 8; 0; -3;) =$$

$$= \min(-3y; y + 4; 5y - 4; y - 8; -3) = \min(-3y; 5y - 4; y - 8; -3).$$

Отсюда, используя неравенство $5y - 4 > y - 8$ при $y \ge 0$, получим, что

$$\varphi(y) = \min(-3y; y - 8; -3) = \begin{cases} y - 8, & \text{при } 0 \le y \le 2, \\ -3y, & \text{при } 2 \le y. \end{cases}$$

Поэтому $\sup\limits_{y \ge 0} \varphi(y) = -6$.

Множество $D = \{(0,0); (1,2); (2,1)\}$. Следовательно,

$$\inf_{x \in D} f_0(x) = \min(0; -3) = -3.$$

В данном примере в соотношении (7.6) стоит строгое неравенство.

Если в соотношении (7.6) стоит строгое неравенство, то говорят, что происходит *разрыв двойственности*.

Теорема 2 (о двойственности). Пусть в задаче (7.1) существует точка $x_0 \in X$, удовлетворяющая связям в этой задаче, и существует набор множителей Лагранжа $\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_m$ такой, что

1) $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$;

2) $\lambda_i f_i(x_0) = 0,\ i = \overline{1,k}$;

3) $f_0(x_0) + \sum\limits_{s=1}^{m} \lambda_s f_s(x_0) \leq f_0(x) + \sum\limits_{s=1}^{m} \lambda_s f_s(x)$ для любой точки $x \in X$.

Тогда числа $y_s = \lambda_s$, $s = \overline{1,m}$ являются решением двойственной задачи (7.4) и в соотношении (7.6) стоит равенство.

Доказательство. Имеем

$$\varphi(\overline{y}) = \inf_{x \in X}\left(f_0(x) + \sum_{s=1}^{m} y_s f_s(x) \right) = f_0(x_0) + \sum_{s=1}^{m} y_s f_s(x_0) = f_0(x_0) \geq \inf_{x \in D} f_0(x).$$

Отсюда и из неравенства (7.6) следует, что точка $\overline{\lambda}$ является решением двойственной задачи.

Отметим ещё, что двойственная задача является конечномерной.

Пример 3. Рассмотрим задачу о подъёме ракеты на высоту не ниже заданной при условии минимального расхода топлива. Уравнение Мещерского, описывающее вертикальный подъём ракеты, на которую действуют только реактивная сила и сила тяжести, имеет вид

$$y' = v,\ v' = u - g,\ u = -w\frac{m'(t)}{m(t)}. \tag{7.7}$$

Здесь y — координата ракеты. Считаем, что скорость w отделения топлива от корпуса ракеты постоянна. Предполагаем также, что величина реактивной тяги ограничена

$$0 \leq u \leq a. \tag{7.8}$$

Выведем формулу для величины топлива, израсходованного при формировании реактивной тяги $u(t)$, $0 \leq t \leq T$. Из формулы (7.7) имеем, что

$$\int_0^T u(t)dt = -\int_0^T w\frac{m'(t)}{m(t)}dt = w\ln\frac{m(0)}{m(T)} = w\ln\frac{m_{\text{топлива}} + m_{\text{корпуса}}}{m_{\text{корпуса}}}.$$

Следовательно, задача о минимизации массы топлива сводится к задаче о минимизации интеграла $\int\limits_0^T u(t)dt$.

Считаем, что начальное положение ракеты задаётся равенствами $y(0) = 0$, $v(0) = 0$. Тогда из формул (7.7) получим, что при заданном управлении $u(t)$ высота $y(T)$ и скорость $v(T)$ определяются следующими формулами:

$$y(T) = \int_0^T (T-t)(u(t)-g)d\,t,\ v(T) = \int_0^T (u(t)-g)dt\,.$$

По условию задачи должно выполняться неравенство $y(T) \geq h$, где h — заданная высота. Поэтому имеется ограничение на выбор реактивной тяги

$$\int_0^T (t-T)\,u(t)d\,t - b \leq 0,\ b = -h - g\frac{T^2}{2}\,.$$

Максимально возможная высота достигается при управлении $u(t) = a$ и равна $(a-g)\dfrac{T^2}{2}$. Поэтому должно выполняться неравенство $(a-g)\dfrac{T^2}{2} \geq h$. Если в этом неравенстве стоит знак равенства, то существует только одно управление $u(t) = a$, поднимающее ракету на высоту не ниже заданной. Это управление и будет оптимальным. Поэтому считаем, что

$$(a-g)\frac{T^2}{2} > h \Leftrightarrow -a\frac{T^2}{2} - b < 0. \qquad (7.9)$$

Ищем управление в классе кусочно-непрерывных функций. В качестве линейного пространства E возьмём множество кусочно-непрерывных функций u:$[0,\ T] \to R$. Введём в рассмотрение функции f_i: $E \to R$, определённые формулами

$$f_0(u(\cdot)) = \int_0^T u(t)dt \ \text{ и } f_1(u(\cdot)) = \int_0^T (t-T)\,u(t)d\,t - b.$$

Учитывая ограничение (7.8), положим

$$X = \{u(\cdot) \in E\text{: } 0 \leq u(t) \leq a \text{ при всех } 0 \leq t \leq T\}.$$

Таким образом, получили задачу

$$\int_0^T u(t)dt \to \min\,,\ \int_0^T (t-T)\,u(t)dt - b \leq 0,\ u(\cdot) \in X. \qquad (7.10)$$

Выпишем двойственную функцию Лагранжа

$$\varphi(y) = \inf_{u(\cdot)\in X}\Big[\int_0^T u(t)dt + y\Big(\int_0^T (t-T)\,u(t)dt - b\Big)\Big] =$$

$$\inf_{u(\cdot)\in X}\Big[\int_0^T (1 + y(t-T))\,u(t)dt - yb\Big]. \qquad (7.11)$$

Если $0 \le y \le \dfrac{1}{T}$, то $1 + y(t-T) > 0$ при всех $0 \le t < T$. Поэтому минимальное значение у интеграла в выражении (7.11) достигается на функции $u(t) \equiv 0$. Следовательно, в рассматриваемом случае $\varphi(y) = -yb$.

Пусть $\dfrac{1}{T} < y$. Тогда $1 + y(t-T) < 0$ при $0 \le t < t_*$ и $1 + y(t-T) > 0$ при $t_* < t \le T$, где $t_* = T - \dfrac{1}{y}$. Следовательно, минимальное значение у интеграла в выражении (7.11) достигается на функции $u(t) = a$ при $0 \le t < t_*$ и $u(t) = 0$ при $t_* < t \le T$. Подставляя эту функцию в формулу (7.11), получим вид значение двойственной функции Лагранжа при этих значениях переменной y. Таким образом,

$$\varphi(y) = -yb \text{ при } 0 \le y \le \frac{1}{T} \text{ и}$$

$$\varphi(y) = a\Big(T - \frac{1}{2y} - \frac{yT^2}{2}\Big) - yb \text{ при } \frac{1}{T} \le y \qquad (7.12)$$

Вычисляя производную функции (7.12), находим, что максимальное значение этой функции достигается в точке $y = \lambda$, где

$$\lambda = \frac{1}{\sqrt{T^2 + \dfrac{2b}{a}}}. \qquad (7.13)$$

Отметим, что согласно условию (7.9), подкоренное выражение в этой формуле больше нуля. Далее,

$$\varphi(\lambda) = a\Big(T - \frac{1}{2\lambda} - \frac{\lambda T^2}{2}\Big) - \lambda b.$$

Используя полученное решение в двойственной задаче, найдём оптимальное решение в исходной задаче. Покажем, что оптимальным решением является функция

$$u_0(t) = a \text{ при } 0 \le t < t_* \text{ и } u_0(t) = 0 \text{ при } t_* < t \le T, \text{ где } t_* = T - \frac{1}{\lambda}. \qquad (7.14)$$

В самом деле, множитель $\lambda > 0$. Далее,

$$\int_0^T \left(t-T\right) u_0(t)d\,t - b = a \int_0^{T-\frac{1}{\lambda}} \left(t-T\right) d\,t - b =$$

$$a[\frac{1}{2}(T-\frac{1}{\lambda})^2 - (T-\frac{1}{\lambda})T] - b = a[\frac{1}{2\lambda^2} - \frac{1}{2}T^2] - b = 0.$$

Здесь была использована формула (7.13). Далее, согласно (7.11)

$$\int_0^T u_0(t)dt + \lambda(\int_0^T \left(t-T\right) u_0(t)dt - b) \le \int_0^T u(t)dt + \lambda(\int_0^T \left(t-T\right) u(t)dt - b)$$

для любой функции $u(\cdot) \in X$. По теореме о достаточных условиях абсолютного минимума в задаче со связями найденная функция $u_0(t)$ является решением в рассматриваемой задаче.

8. Задача выпуклого программирования и теорема Куна — Таккера

Рассмотрим задачу *выпуклого программирования при наличии линейных связей типа равенств*

$$f_0(x) \to min, \ f_1(x) \le 0, \ldots, f_k(x) \le 0, f_{k+1}(x) = 0, \ldots, f_m(x) = 0, x \in X. \quad (8.1)$$

Здесь X — выпуклое множество в линейном пространстве E; $f_i : X \to R$ — выпуклые функции, $i = \overline{0,k}$, а каждая из функций f_j: $E \to R$, $j = \overline{k+1,m}$ является афинной, то есть

$$f_j(x) = g_j(x) - b_j. \quad (8.2)$$

Здесь b_j — заданные числа, а g_j:$E \to R$ — линейные функции. Условие линейности функции g_j:$E \to R$ означает, что для любых точек $x_s \in E$ и для любых чисел $c_s \in R$, $s = 1, 2$ выполнено равенство $g_j(c_1x_1 + c_2x_2) = c_1g_j(x_1) + c_2g_j(x_2)$. Отсюда следует, что каждая функция (8.2) удовлетворяет условию

$$f_j(\lambda x_1 + (1 - \lambda)x_2) = \lambda f_j(x_1) + (1 - \lambda)f_j(x_2)$$
$$\text{для любых } \lambda \in R \text{ и } x_s \in E, s = 1, 2. \quad (8.3)$$

Функция Лагранжа задачи (8.1) имеет вид

$$L(x, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m) = \lambda_0 f_0(x) + \lambda_1 f_1(x) + \lambda_k f_k(x) + \ldots + \lambda_m f_m(x),$$

где числа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ являются *множителями Лагранжа*.

Теорема 1 (теорема Куна – Таккера). Пусть точка x_* является решением задачи (8.1). Тогда существует ненулевой набор множителей Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ такой, что выполнены следующие условия:

1) $\lambda_0 \ge 0, \lambda_1 \ge 0, \ldots, \lambda_k \ge 0$ (условия согласования знаков);

2) $\lambda_1 f_1(x_*) = 0, \ldots, \lambda_k f_k(x_*) = 0$ (условия дополняющей нежёсткости);

3) $L(x_*, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m) = \min\limits_{x \in X} L(x, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m)$ (условие минимума функции Лагранжа).

Отметим, что с учётом условий дополняющей нежёсткости и связей типа равенства условие минимума функции Лагранжа принимает вид

$$\lambda_0 f_0(x) + \lambda_1 f_1(x) + \ldots + \lambda_k f_k(x) + \ldots + \lambda_m f_m(x) \ge \lambda_0 f_0(x_*)$$
$$\text{для любых } x \in X. \qquad (8.4)$$

Доказательство. В пространстве R^{m+1} рассмотрим множество

$$Y = \{(y_0, y_1, \ldots, y_m) \mid \exists z \in X: f_0(z) < f_0(x_*) + y_0, f_i(z) \le y_i,$$
$$i = \overline{1,k}, f_j(z) = y_j, j = \overline{k+1,m}\}. \qquad (8.5)$$

Точка $(1, 0, \ldots, 0) \in Y$ (в формуле (8.5) нужно взять $z = x_*$). Покажем, что множество Y является выпуклым. Возьмём два набора $(y_0, y_1, \ldots, y_m) \in Y$ и $(y_0^*, y_1^*, \ldots, y_m^*) \in Y$. Тогда найдутся точки $z \in X$ и $z^* \in X$ такие, что

$$f_0(z) < f_0(x_*) + y_0, f_i(z) \le y_i, i = \overline{1,k}, f_j(z) = y_j, j = \overline{k+1,m};$$

$$f_0(z^*) < f_0(x_*) + y_0^*, f_i(z^*) \le y_i, i = \overline{1,k}, f_j(z^*) = y_j, j = \overline{k+1,m}.$$

Умножим соотношения первой группы на число $\lambda \in (0, 1)$, а соотношения второй группы — на число $1 - \lambda$ и сложим. Учитывая выпуклость рассматриваемых функций $f_i(x)$ при $i = \overline{1,k}$, а также равенства (8.3) для функций $f_j(x)$, $j = \overline{k+1,m}$, будем иметь

$$f_0(\lambda z + (1-\lambda)z^*) \le \lambda f_0(z) + (1-\lambda) f_0(z^*) < f_0(x_*) + \lambda y_0 + (1-\lambda) y_0^*,$$
$$f_i(\lambda z + (1-\lambda)z^*) \le \lambda f_i(z) + (1-\lambda) f_i(z^*) \le \lambda y_i + (1-\lambda) y_i^*, i = \overline{1,k},$$
$$f_j(\lambda z + (1-\lambda)z^*) = \lambda f_j(z) + (1-\lambda) f_j(z^*) = \lambda y_j + (1-\lambda) y_j^*, j = \overline{k+1,m}.$$

Следовательно, $(\lambda y_0 + (1-\lambda)y_0^*, \lambda y_1 + (1-\lambda)y_1^*, \ldots, \lambda y_m + (1-\lambda)y_m^*) \in Y$.

Нулевой набор $(0, 0, \ldots, 0) \notin Y$. В самом деле, в противном случае найдётся точка $z \in X$ такая, что $f_0(z) < f_0(x_*)$, $f_i(z) \le 0$, $i = \overline{1,k}$, $f_j(z) \le 0, j = \overline{k+1,m}$.

Это противоречит тому, что точка x_* является решением задачи (8.1).

Согласно теореме отделимости точки от выпуклого множества найдётся *ненулевой* набор чисел $(\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m)$ такой, что

$$\lambda_0 y_0 + \lambda_1 y_1 + \ldots + \lambda_m y_m \ge 0 \text{ для любых } (y_0, y_1, \ldots, y_m) \in Y. \quad (8.6)$$

Набор $(1, 0, \ldots, 0)$ принадлежит множеству Y (нужно в формуле (8.5) взять точку $z = x_*$).Подставим его в неравенство (8.6) Получим неравенство $\lambda_0 \ge 0$.

Для любого числа $\varepsilon > 0$ набор $(\varepsilon, 0, \ldots, 0, 1, 0, \ldots, 0)$, у которого i-м месте, $i = \overline{1,k}$, стоит единица, принадлежит множеству Y (нужно в формуле (8.5) взять точку $z = x_*$). Подставим этот набор в неравенство (8.6). Будем иметь $\lambda_0\varepsilon + \lambda_i \geq 0$ для любого числа $\varepsilon > 0$. Устремляя число ε к нулю, получим условие согласования знаков $\lambda_i \geq 0$, $i = \overline{1,k}$.

Для любого числа $\varepsilon > 0$ набор (ε, 0, …, 0, $f_i(x_*)$, 0, …, 0), у которого на i-м месте, $i = \overline{1,k}$, стоит число $f_i(x_*)$, принадлежит множеству Y (нужно в формуле (8.5) взять $z = x_*$). Подставим этот набор в неравенство (8.6). Получим неравенство $\lambda_0\varepsilon + \lambda_i f_i(x_*) \geq 0$ для любых чисел $\varepsilon > 0$. Устремляя число ε к нулю, получим неравенство $\lambda_i f_i(x_*) \geq 0$. Отсюда, учитывая, что $\lambda_i \geq 0$ и $f_i(x_*) \leq 0$, получим условие дополняющей нежёсткости $\lambda_i f_i(x_*) = 0$.

Для любой точки $x \in X$ и для любого числа $\varepsilon > 0$ набор ($f_0(x) - f_0(x_*) + \varepsilon$, $f_1(x)$, …, $f_k(x)$, $f_{k+1}(x)$, …, $f_m(x)$) принадлежит множеству Y (нужно в формуле (8.5) взять $z = x$). Подставим этот набор в неравенство (8.6). Получим

$$\lambda_0 f_0(x) + \lambda_1 f_1(x) + \ldots + \lambda_k f_k(x) + \lambda_{k+1} f_{k+1}(x) + \ldots$$
$$+ \lambda_m f_m(x) - \lambda_0 f_0(x_*) + \lambda_0 \varepsilon \geq 0$$

для любых $\varepsilon > 0$. Устремляя число ε к нулю, придём к требуемому неравенству (8.4).

Замечание 1. Как следует из теоремы о достаточных условиях абсолютного минимума в задаче со связями, если существуют точка $x_* \in X$, удовлетворяющая связям в задаче (8.1), и набор множителей Лагранжа λ_s, $s = \overline{0,m}$ такие, что выполнены условия 1–3 в теореме Куна — Таккера, причём $\lambda_0 > 0$, то эта точка x_* является решением задачи (8.1). Поэтому достаточны условия, при выполнении которых множитель Лагранжа $\lambda_0 > 0$.

Рассмотрим случай, когда в задаче (8.1) отсутствуют связи типа равенства.

Теорема 2. Пусть существует точка $x_0 \in X$ такая, что выполнены *условия Слейтера* $f_1(x_0) < 0$, …, $f_k(x_0) < 0$. Тогда в любом ненулевом наборе чисел $\lambda_0 \geq 0$ и $\lambda_i \geq 0$, $i = \overline{1,k}$, удовлетворяющих неравенству (8.4), множитель $\lambda_0 > 0$.

Доказательство. Допустим, что $\lambda_0 = 0$. Тогда неравенство (8.4) принимает вид $\lambda_1 f_1(x) + \ldots + \lambda_k f_k(x) \geq 0$ для любых $x \in X$. Можно считать, что, например, $\lambda_1 > 0$. Тогда $\lambda_1 f_1(x_0) + \ldots + \lambda_k f_k(x_0) \leq \lambda_1 f_1(x_0) < 0$. Получили противоречие.

Пример 1. Пусть $E = R$, $f_0(x) = -x$, $f_1(x) = x^2$, $X = \{x \in R\colon x \geq 0\}$. Задача

$$-x \to \min,\ x^2 \leq 0,\ x \in X$$

является выпуклой. Очевидно, что точка $x_* = 0$ является решением этой задачи. Согласно теореме Куна — Таккера существует ненулевой набор множителей $\lambda_0 \geq 0$, $\lambda_1 \geq 0$ такой, что $(-\lambda_0 x + \lambda_1 x^2) \geq 0$ при всех $x \geq 0$. Допустим, что $\lambda_0 > 0$. Тогда при достаточно малых числах $x > 0$ функция $(-\lambda_0 x + \lambda_1 x^2) < 0$. Полученное противоречие показывает, что $\lambda_0 = 0$.

Определение 1. Назовём аффинные функции $f_s^*: E \to R, s = \overline{1,q}$, задаваемые формулами (8.2) $f_s^*(x) = g_s^*(x) - b_s^*$, *независимыми*, если из того, что для некоторого набора чисел $c_s \in R$, $s = \overline{1,q}$ выполнены равенства

$$\sum_{s=1}^{q} c_s f_s^*(x) = 0 \text{ для любого } x \in E, \tag{8.7}$$

следует, что все эти числа $c_s = 0$, $s = \overline{1,q}$.

Из формул $f_s^*(x) = g_s^*(x) - b_s^*$ следует, что тождество (8.7) равносильно следующим условиям:

$$\sum_{s=1}^{q} c_s g_s^*(x) = 0 \text{ для любого } x \in E \text{ и } \sum_{s=1}^{q} c_s b_s^* = 0. \tag{8.8}$$

Лемма 1. Пусть аффинные функции $f_s^*: E \to R, s = \overline{1,q}$, таковы, что для некоторого набора чисел $c_s \in R$, $s = \overline{1,q}$ выполнено неравенство $\sum_{s=1}^{q} c_s f_s^*(x) \geq b$ для любого $x \in E$. Здесь b некоторое число. Тогда $\sum_{s=1}^{q} c_s f_s^*(x) = const$ для любого $x \in E$.

Доказательство. Используя формулы $f_s^*(x) = g_s^*(x) - b_s^*$, перепишем заданное неравенство в следующем виде:

$$\sum_{s=1}^{q} c_s g_s^*(x) \geq \sum_{s=1}^{q} c_s b_s^* + b \text{ для любого } x \in E. \tag{8.9}$$

Допустим, что $\sum_{s=1}^{q} c_s g_s^*(x_0) \neq 0$ в некоторой точке $x_0 \in E$. Тогда

$$\sum_{s=1}^{q} c_s g_s^*(tx_0) = t \sum_{s=1}^{q} c_s g_s^*(x_0) \to -\infty \text{ при } t \to -\infty \text{ или при } t \to +\infty.$$

Стало быть, неравенство (8.9) не выполнено. Таким образом, $\sum_{s=1}^{q} c_s g_s^*(x) = 0$ в каждой точке $x \in E$. Следовательно, $\sum_{s=1}^{q} c_s f_s^*(x) = -\sum_{s=1}^{q} c_s b_s^* = const$.

Следствие 1. Пусть в задаче (8.1) максимальное число независимых линейных связей типа равенств равно q. Осуществляя перенумеровку этих связей, считаем, что связи $f_{k+1}(x) = 0$, …,

$f_{k+q}(x) = 0$ являются независимыми, а любые $q + 1$ связи являются зависимыми. Отсюда следует, что существуют числа F_s^j, $s = \overline{1,q}$, $j = \overline{1, m-k-q}$ такие, что для любой точки $x \in E$ выполнены равенства

$$f_{k+q+1}(x) = F_1^1 f_{k+1}(x) + \ldots + F_q^1 f_{k+q}(x), \ldots, f_m(x) =$$
$$= F_1^{m-k-q} f_{k+1}(x) + \ldots + F_q^{m-k-q} f_{k+q}(x).$$

Поэтому задача (8.1) равносильна задаче

$$f_0(x) \to min, f_1(x) \le 0, \ldots, f_k(x) \le 0, f_{k+1}(x) = 0, \ldots, f_{k+q}(x) = 0, x \in X. \quad (8.10)$$

По теореме Куна — Таккера для задачи (8.10) существует ненулевой набор множителей Лагранжа λ_0, λ_1, …, λ_k, …, λ_{k+q}, удовлетворяющий условиям 1–3 в теореме Куна — Таккера для задачи (8.10). Очевидно, что этот же набор множителей вместе с числами $\lambda_{k+q+1} = 0$, …, $\lambda_m = 0$ удовлетворяет условиям 1–3 в теореме Куна — Таккера для задачи (8.1).

Рассмотрим случай, когда в задаче (8.10) множество $X = E$.

Теорема 3. Пусть существует точка $x_0 \in E$ такая, что $f_1(x_0) < 0$, …, $f_k(x_0) < 0, f_{k+1}(x_0) = 0$, …, $f_{k+q}(x_0) = 0$. Тогда в любом ненулевом наборе чисел $\lambda_0 \ge 0$, $\lambda_i \ge 0$, $i = \overline{1,k}$ и $\lambda_j \in R$, $j = \overline{k+1, k+q}$, удовлетворяющих неравенству (8.4) для задачи (8.10) при $X = E$ $\lambda_0 f_0(x) + \lambda_1 f_1(x) + \lambda_k f_k(x) + \ldots + \lambda_{k+q} f_{k+q}(x) \ge \lambda_0 f_0(x_*)$ для любых $x \in E$, множитель $\lambda_0 > 0$.

Доказательство. Допустим, что $\lambda_0 = 0$. Тогда предыдущее неравенство принимает вид $\lambda_1 f_1(x) + \ldots + \lambda_k f_k(x) + \ldots + \lambda_{k+q} f_{k+q}(x) \ge 0$ для любой точки $x \in E$. Поскольку связи типа равенств в задаче (8.10) независимы, то среди чисел $\lambda_i \ge 0$, $i = \overline{1,k}$, имеется ненулевое. Пусть, например, $\lambda_1 > 0$. Тогда получим противоречие

$$0 \le \lambda_1 f_1(x_0) + \ldots + \lambda_k f_k(x_0) + \ldots + \lambda_{k+q} f_{k+q}(x_0) =$$
$$\lambda_1 f_1(x_0) + \ldots + \lambda_k f_k(x_0) \le \lambda_1 f_1(x_0) < 0.$$

Следствие 2. Пусть точка x_* является решением задачи (8.1) при $X = E$ и существует точка $x_0 \in E$ такая, что $f_1(x_0) < 0$, …, $f_k(x_0) < 0$, $f_{k+1}(x_0) = 0$, …, $f_m(x_0) = 0$ Тогда существует набор множителей Лагранжа $\lambda_0 > 0$, $\lambda_1 \ge 0$, …, $\lambda_k \ge 0$, $\lambda_{k+1} \in R$, …, $\lambda_m \in R$ такой, что выполнены условия теоремы Куна — Таккера.

Рассмотрим задачу *выпуклого программирования при наличии только линейных связей типа равенств и неравенств*

$$f_0(x) \to min, f_1(x) \le 0, \ldots, f_k(x) \le 0, f_{k+1}(x) = 0, \ldots, f_m(x) = 0, x \in E. \quad (8.11)$$

Здесь каждая из функций $f_s\colon E \to R$, $s = \overline{1,m}$ является афинной.

Теорема 4 (о неравенстве нулю множителя λ_0 в выпуклой задаче с линейными связями). Пусть точка x_* является решением задачи (8.11). Тогда существует набор множителей Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ такой, что выполнены следующие условия:

1) $\lambda_0 > 0, \lambda_1 \geq 0, \ldots, \lambda_k \geq 0$;

2) $\lambda_1 f_1(x_*) = 0, \ldots, \lambda_k f_k(x_*) = 0$;

3) $L(x_*, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m) = \min\limits_{x \in E} L(x, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m)$.

Доказательство. В соответствии со следствием 1 можем считать, что связи типа равенств являются независимыми. Согласно теореме Куна — Таккера существует ненулевой набор множителей Лагранжа $\lambda_0 \geq 0$, $\lambda_i \geq 0, i = \overline{1,k}$, $\lambda_j \in R, j = \overline{k+1,m}$ такой, что выполнены условия 2 и 3.

Рассмотрим вначале случай, когда связи типа неравенств отсутствуют. Допустим, что множитель $\lambda_0 = 0$. Тогда условие 3 принимает вид $\sum\limits_{j=k+1}^{m} \lambda_j f_j(x) \geq 0$ для любой точки $x \in E$. Применяя лемму 1, получим, что функция $\sum\limits_{j=k+1}^{m} \lambda_j f_j(x) = const$ для любой точки $x \in E$. Точка x_* удовлетворяет связям в задаче (8.1). Поэтому $\sum\limits_{j=k+1}^{m} \lambda_j f_j(x)$ $= 0$ для любой точки $x \in E$. Стало быть, связи типа равенств зависимы. Получили противоречие.

Доказательство неравенства $\lambda_0 > 0$ при наличии линейных связей типа неравенств проведём индукцией по числу этих связей. Когда число таких связей равно нулю, утверждение было доказано. Пусть утверждение верно, когда число таких связей равно $k-1$. Докажем его для числа k.

По теореме Куна — Таккера для задачи (8.11) существует ненулевой набор множителей Лагранжа $\lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ такой, что выполнены следующие условия:

1) $\lambda_0 \geq 0, \lambda_1 \geq 0, \ldots, \lambda_k \geq 0$;

2) $\lambda_1 f_1(x_*) = 0, \ldots, \lambda_k f_k(x_*) = 0$;

3) $L(x_*, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m) = \min\limits_{x \in E} L(x, \lambda_0, \lambda_1, \ldots, \lambda_k, \ldots, \lambda_m)$.

Допустим, что $\lambda_0 = 0$. Тогда из условия 3, применяя лемму 1, получим

$$\sum_{s=1}^{m} \lambda_s f_s(x) = 0 \text{ для любой точки } x \in E. \qquad (8.12)$$

Отсюда следует, что среди чисел $\lambda_i \geq 0$, $i = \overline{1,k}$ имеется положительное. В противном случае связи типа равенств являются зависимыми. Пусть, например, $\lambda_1 > 0$. Рассмотрим задачу

$$f_0(x) \to \min,$$

$$f_1(x) = 0, f_2(x) \le 0, \ldots, f_k(x) \le 0, f_{k+1}(x) = 0, \ldots, f_m(x) = 0,\ x \in E. \quad (8.13)$$

Из условий $\lambda_1 f_1(x_*) = 0$ и $\lambda_1 > 0$ следует, что x_* удовлетворяет связям в этой задаче. Поэтому точка x_* является решением задачи (8.13). В этой задаче число связей типа неравенств равно $k-1$. По индукционному предположению существуют числа $\lambda_s^*, s=\overline{1,m}$ такие, что выполнены следующие условия:

1) $\lambda_2^* \ge 0, \ldots, \lambda_k^* \ge 0$;

2) $\lambda_2^* f_2(x_*) = 0, \ldots, \lambda_k^* f_k(x_*) = 0$;

$$f_0(x_*) + \sum_{s=1}^{m} \lambda_s^* f_s(x_*) \le f_0(x) + \sum_{s=1}^{m} \lambda_s^* f_s(x)$$

$$\text{для любой точки } x \in E. \quad (8.14)$$

Отсюда видно, что если $\lambda_1^* \ge 0$, то найденный набор чисел $\lambda_s^*, s=\overline{1,m}$ является искомым.

Пусть $\lambda_1^* < 0$. Рассмотрим числа $\lambda_s^0 = \lambda_s^* - \dfrac{\lambda_1^*}{\lambda_1}\lambda_s, s=\overline{1,m}$. Тогда $\lambda_1^0 = 0$. Поскольку $\lambda_1 > 0$, $\lambda_1^* < 0$ и $\lambda_i \ge 0$, $\lambda_i^* \ge 0$ при $i=\overline{2,k}$, то $\lambda_i^0 \ge 0$ при $i=\overline{2,k}$.

При $i=\overline{2,k}$ имеем

$$\lambda_i^0 f_i(x_*) = \lambda_i^* f_i(x_*) - \frac{\lambda_1^*}{\lambda_1}\lambda_i f_i(x_*) = 0.$$

Соотношение $\lambda_1^0 f_1(x_*) = 0$ следует из равенства $\lambda_1^0 = 0$. Далее, для любой точки $x \in E$ выполнено равенство

$$f_0(x) + \sum_{s=1}^{m} \lambda_s^0 f_s(x) = f_0(x) + \sum_{s=1}^{m} \lambda_s^* f_s(x) - \frac{\lambda_1^*}{\lambda_1} \sum_{s=1}^{m} \lambda_s f_s(x).$$

Отсюда, используя тождество (8.12), получим, что

$$f_0(x) + \sum_{s=1}^{m} \lambda_s^0 f_s(x) = f_0(x) + \sum_{s=1}^{m} \lambda_s^* f_s(x).$$

для любой точки $x \in E$. Отсюда и из (8.14) получим требуемое неравенство

$$f_0(x_*) + \sum_{s=1}^{m} \lambda_s^0 f_s(x_*) = f_0(x_*) + \sum_{s=1}^{m} \lambda_s^* f_s(x_*) \le f_0(x) + \sum_{s=1}^{m} \lambda_s^* f_s(x)$$

$$= f_0(x) + \sum_{s=1}^{m} \lambda_s^0 f_s(x)$$

для любой точки $x \in E$.

9. Задача квадратичного программирования с линейными связями

Рассмотрим общую задачу квадратичного программирования в пространстве R^n

$$\frac{1}{2}<F\overline{x}, \overline{x}>+<\overline{g}, \overline{x}>\to\min, <\overline{g}_i, \overline{x}>-b_i\leq 0, \; i=\overline{1,k}. \qquad (9.1)$$

Здесь $\overline{g}_i \in R^n$ — заданные векторы, а b_i — заданные числа.

Замечание 1. Каждая линейная связь типа равенств $<\overline{g}_*, \overline{x}>-b_*=0$ может быть записана с помощью двух линейных связей типа неравенств $<\overline{g}_*, \overline{x}>-b_*\leq 0$ и $<-\overline{g}_*, \overline{x}>-(-b_*)\leq 0$. Поэтому к задаче (9.1) сводится задача и при наличии линейных связей типа равенств.

Теорема 1. Пусть в задаче квадратичного программирования (9.1) связи являются совместными. Тогда либо существует точка абсолютного минимума у целевой функции, либо целевая функция не ограничена снизу на множестве, определяемом связями в этой задаче.

Доказательство. Если в задаче (9.1) существует точка абсолютного минимума, то, очевидно, что целевая функция ограничена снизу на множестве, определяемом связями в этой задаче.

Пусть целевая функция в задаче (9.1) ограничена снизу на множестве, определяемом связями в этой задаче. Покажем, что существует точка абсолютного минимума. Доказательство проведём индукцией по числу связей. При $k = 0$ это утверждение было доказано ранее.

Пусть утверждение верно для любой задачи (9.1) при числе связей, равном k. Докажем его для задачи

$$\frac{1}{2}<F\overline{x}, \overline{x}>+<\overline{g}, \overline{x}>\to\min, <\overline{g}_i, \overline{x}>-b_i\leq 0, \; i=\overline{1,k+1}. \qquad (9.2)$$

Если, например, последняя связь является следствием первых k связей, то задача (9.2) эквивалентна задаче (9.2) с k связями и, следовательно, для неё верно утверждение доказываемой теоремы.

Рассмотрим случай, когда существует точка $\overline{x}_* \in R^n$, удовлетворяющая первым k связям, но $<\overline{g}_{k+1}, \overline{x}_*>-b_{k+1}>0$. Поскольку считается, что связи в задаче (9.2) совместны, то существует точка

$\overline{x} \in R^n$, удовлетворяющая всем этим связям. Тогда при любом числе $\lambda \in [0, 1]$ точка $\overline{x}_\lambda = \lambda \overline{x} + (1 - \lambda) \overline{x}_*$ удовлетворяет первым k связям в задаче (9.2). Покажем, что при некотором числе $\lambda \in (0,1)$ эта точка удовлетворяет равенству $< \overline{g}_{k+1}, \overline{x}_\lambda > - b_{k+1} = 0$. С этой целью рассмотрим функцию $l(\lambda) = < \overline{g}_{k+1}, \overline{x}_\lambda > - b_{k+1}$. Тогда

$$l(0) = < \overline{g}_{k+1}, \overline{x}_* > - b_{k+1} > 0, \quad l(1) = < \overline{g}_{k+1}, \overline{x} > - b_{k+1} \leq 0.$$

Отсюда следует, что требуемое число λ существует.

Таким образом, связи в задаче

$$\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} > \to \min,$$
$$< \overline{g}_i, \overline{x} > - b_i \leq 0, \; i = \overline{1,k}, < \overline{g}_{k+1}, \overline{x} > - b_{k+1} = 0 \tag{9.3}$$

являются совместными.

В рассматриваемом случае значения целевой функции ограничены снизу на множестве X, определяемом связями в задаче (9.2). Поскольку связи в задаче (9.3) определяют множество, содержащееся во множестве X, то значения целевой функции в задаче (9.3) также ограничены снизу. Покажем, что в задаче (9.3) существует точка абсолютного минимума. Возьмём любую точку $\overline{x}_+$, удовлетворяющую связям в задаче (9.3). Сделаем замену переменных $\overline{z} = \overline{x} - \overline{x}_+$. Тогда

$$\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} > = < F\overline{z}, \overline{z} > + < F\overline{x}_+ + \overline{g}, \overline{z} >$$
$$+ < F\overline{x}_+, \overline{x}_+ > + < \overline{g}, \overline{x}_+ >.$$

Получим эквивалентную задачу

$$\frac{1}{2} < F\overline{z}, \overline{z} > + < F\overline{x}_+ + \overline{g}, \overline{z} > \to \min,$$

$$< \overline{g}_i, \overline{z} > - (b_i - < \overline{g}_i, \overline{x}_+ >) \leq 0, i = \overline{1,k}; < \overline{g}_{k+1}, \overline{z} > = 0. \tag{9.4}$$

Учитывая последнюю линейную связь в задаче (9.4), мы можем рассматривать эту задачу в линейном подпространстве размерности $n - 1$, ортогональном вектору $\overline{g}_{k+1}$. В этом подпространстве задача (9.3) имеет k связей. По индукционному предположению у неё существует решение. Следовательно, и задача (9.3) имеет точку абсолютного минимума, которую обозначим $\overline{x}^0$. Далее рассмотрим два случая.

Случай 1. В задаче (9.1) значение целевой функции ограничено снизу. Тогда существует точка абсолютного минимума, которую обозначим $\overline{x}_0$. Если $<\overline{g}_{k+1}, \overline{x}_0> -b_{k+1} \le 0$, то эта точка является точкой абсолютного минимума в задаче (2).

Пусть $<\overline{g}_{k+1}, \overline{x}_0> -b_{k+1} > 0$. Покажем, что точкой абсолютного минимума задачи (9.2) является решение задачи (9.3), то есть точка $\overline{x}^0$. Эта точка удовлетворяет всем связям в задаче (9.2).

Возьмём любую точку $\overline{x}$, удовлетворяющую связям в задаче (9.5). При любом числе $\lambda \in [0, 1]$ рассмотрим точку $\overline{x}_\lambda = \lambda \overline{x}_0 + (1 - \lambda)\overline{x}$. Поскольку точки $\overline{x}_0$ и $\overline{x}$ удовлетворяют неравенствам в задаче (9.1), то этим неравенствам удовлетворяет точка $\overline{x}_\lambda$. Покажем, что при некотором числе $\lambda_* \in [0, 1]$ точка $\overline{x}_{\lambda_*}$ удовлетворяет равенству $<\overline{g}_{k+1}, \overline{x}_{\lambda_*}> -b_{k+1} = 0$. С этой целью рассмотрим функцию $l(\lambda) = <\overline{g}_{k+1}, \overline{x}_\lambda> -b_{k+1}$, которая удовлетворяет следующим условиям:

$$l(0) = <\overline{g}_{k+1}, \overline{x}> -b_{k+1} \le 0,\ l(1) = <\overline{g}_{k+1}, \overline{x}_0> -b_{k+1} > 0.$$

Отсюда следует, что требуемое число λ_* существует. Далее, используя выпуклость целевой функции, получим неравенство

$$\frac{1}{2}<F\overline{x}^0, \overline{x}^0> + <\overline{g}, \overline{x}^0> \le \frac{1}{2}<F\overline{x}_{\lambda_*}, \overline{x}_{\lambda_*}> +$$

$$<\overline{g}, \overline{x}_{\lambda_*}> \le \lambda_*(<F\overline{x}_0, \overline{x}_0> + <\overline{g}, \overline{x}_0>) + (1 - \lambda_*)(\frac{1}{2}<F\overline{x}, \overline{x}> +$$

$$+ <\overline{g}, \overline{x}>) \le \lambda_*(\frac{1}{2}<F\overline{x}, \overline{x}> + <\overline{g}, \overline{x}>) +$$

$$+ (1 - \lambda_*)(\frac{1}{2}<F\overline{x}, \overline{x}> + <\overline{g}, \overline{x}>) = \frac{1}{2}<F\overline{x}, \overline{x}> + <\overline{g}, \overline{x}>.$$

Случай 2. В задаче (9.1) значение целевой функции не ограничено снизу. Тогда существует последовательность точек $\overline{x}_m$, удовлетворяющая связям в задаче (1), такая, что $\frac{1}{2}<F\overline{x}_m, \overline{x}_m> + <\overline{g}, \overline{x}_m> \to -\infty$. Поскольку значения целевой функции ограничены снизу на множестве X, определяемом связями в задаче (9.2), то среди точек $\overline{x}_m$ не может быть бесконечного числа точек, удовлетворяющих неравенству $<\overline{g}_{k+1}, \overline{x}_m> -b_{k+1} \le 0$. Итак, считаем, что $<\overline{g}_{k+1}, \overline{x}_m> -b_{k+1} > 0$ для всех достаточно больших номеров m. Возьмём любую точку $\overline{x}$, удовлетворяющую связям в задаче (9.2). Возьмём достаточно большой номер m такой, чтобы $<\overline{g}_{k+1}, \overline{x}_m> - b_{k+1} > 0$ и

$$\frac{1}{2}<F\overline{x}_m, \overline{x}_m> + <\overline{g}, \overline{x}_m> \le \frac{1}{2}<F\overline{x}, \overline{x}> + <\overline{g}, \overline{x}>.$$

Как и выше, показывается, что точка $\overline{x}_\lambda = \lambda \overline{x}_m + (1 - \lambda)\overline{x}$ при некотором числе $\lambda \in [0, 1]$ удовлетворяет связям в задаче (9.3). Следовательно,

$$\frac{1}{2} < F\overline{x}^0, \overline{x}^0 > + < \overline{g}, \overline{x}^0 > \le \frac{1}{2} < F\overline{x}_\lambda, \overline{x}_\lambda > +$$

$$< \overline{g}, \overline{x}_\lambda > \le \lambda(\frac{1}{2} < F\overline{x}_m, \overline{x}_m > +$$

$$+ < \overline{g}, \overline{x}_m >) + (1-\lambda)(\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} >) \le \lambda(\frac{1}{2} < F\overline{x}, \overline{x} > +$$

$$< \overline{g}, \overline{x} >) + (1-\lambda_*)(\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} >) = \frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} >.$$

Стало быть, точка $\overline{x}^0$ является точкой абсолютного минимума в задаче (9.2).

Выпишем задачу квадратичного программирования (9.1) при наличии связей типа равенств

$$\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} > \to \min,$$

$$< \overline{g}_i, \overline{x} > - b_i \le 0,\ i = \overline{1,k};\ < \overline{g}_j, \overline{x} > - b_j \le 0,\ j = \overline{k+1,m}. \quad (9.5)$$

При положительно определённой матрице F эта задача является задачей выпуклого программирования. Для нее верна теорема Куна — Таккера с множителем $\lambda_0 = 1$. Функция Лагранжа

$$L(\overline{x}, \lambda_1, \ldots, \lambda_m) = \frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} > + \sum_{i=1}^{m} \lambda_i \left(< \overline{g}_i, \overline{x} > - b_i \right)$$

является выпуклой по переменной $\overline{x}$. Поэтому теорема Куна — Таккера принимает следующий вид.

Теорема 2. Для того чтобы точка $\overline{x}_*$, удовлетворяющая связям в задаче (5), являлась решением этой задачи, необходимо и достаточно, чтобы существовал набор множителей Лагранжа $\lambda_1, \ldots, \lambda_k, \ldots, \lambda_m$ такой, что выполнены следующие условия:

1) $\lambda_1 \ge 0, \ldots, \lambda_k \ge 0$;

2) $\lambda_1(< \overline{g}_1, \overline{x}_* > - b_1) = 0, \ldots, \lambda_k(< \overline{g}_k, \overline{x}_* > - b_k) = 0$;

3) $F\overline{x}_* + \overline{g} + \lambda_1 \overline{g}_1 + \ldots + \lambda_m \overline{g}_m = \overline{0}$.

Запишем двойственную задачу для задачи (9.1). Двойственная функция Лагранжа равна

$$\varphi(\overline{y}) = \inf_{x \in R^n} (\frac{1}{2} < F\overline{x}, \overline{x} > + < \overline{g}, \overline{x} > + \sum_{i=1}^{m} y_i \left(< \overline{g}_i, \overline{x} > - b_i \right)).$$

Поскольку при любом наборе чисел $\overline{y} = (y_1, \ldots, y_m)$ функция

Лагранжа является выпуклой функцией по переменной $\overline{x}$, то её минимум достигается в точке $\overline{x}_0$, где градиент равен нулю, то есть

$$F\overline{x}_0 + \overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i = \overline{0}.$$

Допустим, что матрица F является невырожденной. Тогда

$$\overline{x}_0 = -F^{-1}(\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i).$$

Следовательно,

$$\varphi(\overline{y}) = \frac{1}{2} < F^{-1}(\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i)), (\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i) > -$$

$$- < \overline{g}, F^{-1}(\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i) > - y_1(< \overline{g}_1, F^{-1}(\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i) > + b_1) - \ldots -$$

$$- y_m(< \overline{g}_m, F^{-1}(\overline{g} + \sum_{i=1}^{m} y_i \overline{g}_i) > + b_m)).$$

Матрица F^{-1} является положительно определённой. В самом деле, для любого вектора $\overline{p} \in R^n$ выполнено $< F^{-1}\overline{p}, \overline{p} > = < F(F^{-1}\overline{p}), F^{-1}\overline{p} > \geq 0$. Стало быть, двойственная функция Лагранжа $\varphi(\overline{y})$ является квадратичной функцией, а двойственная задача

$$\varphi(\overline{y}) \to \max, y_1 \geq 0, \ldots, y_k \geq 0, y_{k+1} \in R, \ldots, y_m \in R$$

является задачей квадратичного программирования с линейными связями.

10. Задача линейного программирования и исследование линейных неравенств

Рассмотрим задачу линейного программирования (ЛП) в пространстве R^n

$$\langle \overline{g}, \overline{x} \rangle \to \min, \ \langle \overline{g}_i, \overline{x} \rangle - b_i \le 0, \ i = \overline{1,k};$$
$$\langle \overline{g}_j, \overline{x} \rangle - b_j = 0, \ j = \overline{k+1,m}. \qquad (10.1)$$

Эта задача является частным случаем рассмотренной ранее задачи квадратичного программирования с нулевой матрицей F, поэтому для неё верны следующие теоремы.

Теорема 1 (критерий существования решения в задаче ЛП). Пусть в задаче линейного программирования (10.1) связи являются совместными. Тогда либо существует точка абсолютного минимума у целевой функции, либо целевая функция не ограничена снизу на множестве, определяемом связями в этой задаче.

Задача (10.1) является задачей выпуклого программирования с линейными связями. Функция Лагранжа с множителем $\lambda_0 = 1$ в этом случае равняется

$$L(\overline{x}, \overline{\lambda}) = \left\langle \overline{g} + \sum_{S=1}^{m} \lambda_s \overline{g}_s, \overline{x} \right\rangle - \sum_{S=1}^{m} \lambda_s b_s .$$

Теорема 2 (условия оптимальности в задаче ЛП). Чтобы точка $\overline{x}_*$, удовлетворяющая связям в задаче (10.1), являлась решением в этой задаче, необходимо и достаточно, чтобы существовали множители Лагранжа λ_s, $s = 1, \ldots, m$ такие, что

1) $\lambda_i \ge 0$, $i = \overline{1,k}$;

2) $\lambda_i (\langle \overline{g}_i, \overline{x}_* \rangle - b_i) = 0$, $i = \overline{1,k}$;

3) $\overline{g} + \sum_{S=1}^{m} \lambda_s \overline{g}_s = \overline{0}$.

Теорема 3 (критерий существования решения в задаче ЛП). Пусть связи в задаче (10.1) являются совместными. Решение в задаче (10.1) существует тогда и только тогда, когда существуют числа $\lambda_i \ge 0$, $i=\overline{1,k}$ и $\lambda_j \in R$, $j=\overline{k+1,m}$ такие, что

$$\overline{g} + \sum_{S=1}^{m} \lambda_s \overline{g}_s = \overline{0}\,. \qquad (10.2)$$

Доказательство. Пусть решение в задаче (10.1) существует. Тогда из теоремы 2 следует существование требуемых чисел λ_s, удовлетворяющих равенству (10.2).

Пусть существуют числа λ_s, первые k из которых неотрицательны, удовлетворяющие равенству (10.2). Тогда для любой точки $\overline{x}$, удовлетворяющей связям в задаче (10.1), выполнено

$$\left\langle \overline{g}, \overline{x} \right\rangle = -\sum_{S=1}^{m} \left\langle \lambda_s \overline{g}_s, \overline{x} \right\rangle \ge -\sum_{S=1}^{m} \lambda_s b_s\,.$$

Таким образом, значения целевой функции в задаче (10.1) ограничены снизу. Следовательно, решение существует.

Теорема 3 (критерий совместности системы линейных неравенств). Чтобы система линейных неравенств

$$\left\langle \overline{g}_i, \overline{x} \right\rangle - b_i \le 0\,,\ i=\overline{1,k} \qquad (10.3)$$

была *несовместной,* необходимо и достаточно, чтобы существовали числа $\lambda_i \ge 0$, $i=\overline{1,k}$ такие, что

$$\sum_{i=1}^{k} \lambda_i \overline{g}_i = \overline{0}\,,\ \sum_{i=1}^{k} \lambda_i b_i = -1\,. \qquad (10.4)$$

Доказательство. Допустим, что выполнены равенства (10.4) для некоторых неотрицательных чисел λ_i, но система неравенств (10.3) имеет решение $\overline{x}$. Умножая неравенства (10.3) на эти неотрицательные числа λ_i и складывая, получим $\sum_{i=1}^{k} \lambda_i b_i \ge 0$. Полученное неравенство противоречит второму равенству в (10.4).

Доказательство необходимости проведём индукцией по числу неравенств. В случае одного неравенства условие несовместности означает, что $\overline{g}_1 = \overline{0}$ и $b_1 < 0$. Тогда для числа $\lambda_1 = -\dfrac{1}{b_1} > 0$ условия (10.4) выполнены.

Пусть утверждение верно для случая $k-1$ неравенств. Рассмотрим k несовместных неравенств (10.3). Если их первые $k-1$ неравенств являются несовместными, то по индукционному предположению существуют числа $\lambda_i \geq 0$, $i=\overline{1,k-1}$ такие, что $\sum_{i=1}^{k-1}\lambda_i \overline{g}_i = \overline{0}$, $\sum_{i=1}^{k-1}\lambda_i b_i = -1$. Возьмём $\lambda_k = 0$. Тогда равенства (10.4) выполнены.

Считаем, что первые $k-1$ неравенств являются совместными. Из условия несовместности системы неравенств (10.3) следует, что для любой точки $\overline{x}$, удовлетворяющей первым $k-1$ неравенствам, выполнено $\langle \overline{g}_k, \overline{x}\rangle > b_k$. Стало быть, задача линейного программирования

$$\langle \overline{g}_k, \overline{x}\rangle \to \min,\ \langle \overline{g}_i, \overline{x}\rangle - b_i \leq 0,\ i=\overline{1,k-1}$$

имеет решение $\overline{x}_*$. По теореме 2 существуют числа $\lambda_i \geq 0$, $i=\overline{1,k-1}$ такие, что

$$\lambda_i(\langle \overline{g}_i, \overline{x}_*\rangle - b_i) = 0,\ i=\overline{1,k-1} \text{ и } \overline{g}_k + \sum_{i=1}^{k-1}\lambda_i \overline{g}_i = \overline{0}.$$

Таким образом, первое равенство (10.4) выполнено. Суммируя условия дополняющей нежёсткости, получим, что $\langle \overline{g}_k, \overline{x}_*\rangle = -\sum_{i=1}^{k-1}\lambda_i b_i$. Отсюда, используя неравенство $\langle \overline{g}_k, \overline{x}_*\rangle > b_k$, получим, что $\sum_{i=1}^{k-1}\lambda_i b_i + b_k < 0$. Умножая все числа λ_i на одно и то же положительное число, добьёмся того, чтобы новые числа λ_i удовлетворяли требуемым условиям (10.4). Теорема доказана.

Рассмотрим неравенства

$$\langle \overline{g}, \overline{x}\rangle - b \leq 0, \tag{10.5}$$

$$\langle \overline{g}_i, \overline{x}\rangle - b_i \leq 0,\ i=\overline{1,k}. \tag{10.6}$$

Определение 1. Неравенство (10.5) называется *неравенством-следствием* системы совместных неравенств (10.6), если каждое решение системы неравенств (10.6) является решением неравенства (10.5).

Теорема 5 (о неравенстве-следствии). Неравенство (10.5) является неравенством-следствием системы совместных неравенств (10.6) тогда и только тогда, когда существуют числа $\lambda_i \geq 0$, $i=\overline{1,k}$ такие, что

$$\overline{g} = \sum_{i=1}^{k}\lambda_i \overline{g}_i,\quad b \geq \sum_{i=1}^{k}\lambda_i b_i. \tag{10.7}$$

Доказательство. Пусть для некоторого набора неотрицательных чисел λ_i выполнены соотношения (10.7). Возьмём точку $\overline{x}$, которая удовлетворяет системе неравенств (10.6). Тогда, умножая неравенства (10.6) на неотрицательные числа λ_i и складывая их, получим $\left\langle \sum_{i=1}^{k} \lambda_i \overline{g}_i, \overline{x} \right\rangle - \sum_{i=1}^{k} \lambda_i b_i \le 0$. Отсюда и из соотношений (10.7) следует, что эта точка $\overline{x}$ удовлетворяет неравенству (10.5).

Допустим теперь, что точка $\overline{x}$, удовлетворяющая системе неравенств (10.6), удовлетворяет и неравенству (10.5). Тогда $\langle -\overline{g}, \overline{x} \rangle \ge -b$ на любом решении системы неравенств (10.6). Следовательно, задача

$$\langle -\overline{g}, \overline{x} \rangle \to \min, \quad \langle -\overline{g}_i, \overline{x} \rangle - b_i \le 0, \quad i = \overline{1,k}$$

имеет решение $\overline{x}_*$. По теореме 2 существуют числа $\lambda_i \ge 0$, $i = \overline{1,k}$ такие, что

$$\lambda_i(\langle \overline{g}_i, \overline{x}_* \rangle - b_i) = 0, \quad i = \overline{1,k} \quad \text{и} \quad (-\overline{g}) + \sum_{i=1}^{k} \lambda_i \overline{g}_i = \overline{0}.$$

Стало быть, первое равенство в (10.7) выполнено. Учитывая это равенство и просуммировав условия дополняющей нежёсткости, получим

$$\langle \overline{g}, \overline{x}_* \rangle = \sum_{i=1}^{k} \lambda_i b_i.$$

Отсюда и из неравенства $\langle \overline{g}, \overline{x}_* \rangle - b \le 0$ следует второе условие в (10.7).

Запишем двойственную задачу к задаче линейного программирования (10.1). Используя вид функции Лагранжа

$$L(\overline{x}, \overline{\lambda}) = \left\langle \overline{g} + \sum_{s=1}^{m} \lambda_s \overline{g}_s, \overline{x} \right\rangle - \sum_{s=1}^{m} \lambda_s b_s,$$

запишем двойственную функцию Лагранжа

$$\varphi(\overline{y}) = \inf_{\overline{x} \in R^n} L(\overline{x}, \overline{y}).$$

Она равна

$$\varphi(\overline{y}) = -\infty \quad \text{при} \quad \overline{g} + \sum_{s=1}^{m} y_s \overline{g}_s \ne \overline{0},$$

$$\varphi(\overline{y}) = -\sum_{s=1}^{m} y_s b_s \quad \text{при} \quad \overline{g} + \sum_{s=1}^{m} y_s \overline{g}_s = \overline{0}. \qquad (10.8)$$

Стало быть, двойственная задача

$$-\sum_{s=1}^{m} y_s b_s \to \max, \ y_1 \geq 0, \ldots, y_k \geq 0, \ y_{k+1} \in R, \ldots, y_m \in R,$$

$$\bar{g} + \sum_{s=1}^{m} y_s \bar{g}_s = \bar{0} \qquad (10.9)$$

также является задачей линейного программирования.

Замечание 1. Если в задаче (10.1) существует решение, то применима теорема 2. По теореме двойственности множители Лагранжа в задаче (10.1) являются решением в задаче (10.9), причём значения целевых функций на оптимальных элементах совпадают.

Пусть в задаче (10.9) существует решение. Тогда по критерию существования решения в задаче линейного программирования решение в задаче (10.1) существует.

Запишем двойственную задачу к задаче линейного программирования (10.9). Для этого запишем её в виде (10.1)

$$\sum_{s=1}^{m} y_s b_s \to \min, \ -y_1 \leq 0, \ldots, -y_k \leq 0, \ y_{k+1} \in R, \ldots, y_m \in R,$$

$$\bar{g} + \sum_{s=1}^{m} y_s \bar{g}_s = \bar{0}. \qquad (10.10)$$

Обозначим через g_i координаты вектора $\bar{g}$, а через g_{si} — координаты вектора $\bar{g}_s$. Тогда функция Лагранжа для задачи (10.10) равна

$$L(\bar{y}, \bar{\lambda}) = \sum_{s=1}^{m} y_s b_s + \sum_{s=1}^{k} (-y_s \lambda_s) + \sum_{i=1}^{n} \lambda_{i+m} (g_i + \sum_{s=1}^{m} y_s g_{si}).$$

Двойственная задача имеет вид

$$\varphi(\bar{\lambda}) = \inf_{\bar{y} \in R^m} L(\bar{y}, \bar{\lambda}).$$

Эта функция равна $\varphi(\bar{\lambda}) = -\infty$, если $b_s - \lambda_s + \sum_{i=1}^{n} \lambda_{i+m} g_{si} \neq 0$ хотя бы при одном индексе $s = \overline{1,k}$ или $b_s + \sum_{i=1}^{n} \lambda_{i+m} g_{si} \neq 0$ хотя бы при одном индексе $s = \overline{k+1,m}$; $\varphi(\bar{\lambda}) = \sum_{i=1}^{n} \lambda_{i+m} g_i$, если $b_s - \lambda_s + \sum_{i=1}^{n} \lambda_{i+m} g_{si} = 0$ при всех $s = \overline{1,k}$ и $b_s + \sum_{i=1}^{n} \lambda_{i+m} g_{si} = 0$ при всех $s = \overline{k+1,m}$.

Следовательно, двойственная задача к задаче (10.10) имеет вид

$$\sum_{i=1}^{n} \lambda_{i+m} g_i \to \max, \ \lambda_1 \geq 0, \ldots, \lambda_k \geq 0, \ \lambda_{k+1} \in R, \ldots, \lambda_{m+n} \in R,$$

$b_s - \lambda_s + \sum_{i=1}^{n} \lambda_{i+m}\, g_{si} = 0$ при всех $s = \overline{1,k}$, $b_s + \sum_{i=1}^{n} \lambda_{i+m}\, g_{si} = 0$ при всех $s = \overline{k+1,m}$.

Обозначим $x_i = -\lambda_{i+m}$. Тогда предыдущую задачу можно записать в следующем виде:

$$\sum_{i=1}^{n} x_i\, g_i \to \min,\ \sum_{i=1}^{n} x_i\, g_{si} - b_s \le 0,\ s = \overline{1,k};\ \sum_{i=1}^{n} x_i\, g_{si} - b_s = 0$$

при всех $s = \overline{k+1,m}$.

Получили задачу (10.1), записанную в координатной форме.

Рассмотрим *классическую задачу* линейного программирования

$$\sum_{i=1}^{n} c_i x_i \to \max,\ \sum_{i=1}^{n} a_{ji} x_i - b_j \le 0,\ j = \overline{1,m};\ x_1 \ge 0, \ldots, x_n \ge 0. \quad (10.11)$$

Введём в рассмотрение вектора $\overline{g} = (-c_1, \ldots, -c_n)$, $\overline{g}_j = (a_{j1}, \ldots, a_{jn})$, $j = \overline{1,m}$; $\overline{g}_{m+s} = (0, \ldots, -1, \ldots, 0)$, $s = \overline{1,n}$. Здесь число (–1) стоит на s-м месте. Тогда задача (10.11) примет вид

$$\langle \overline{g}, \overline{x} \rangle \to \min,\ \langle \overline{g}_j, \overline{x} \rangle - b_j \le 0,\ j = \overline{1,m};\ \langle \overline{g}_{m+s}, \overline{x} \rangle \le 0,\ s = \overline{1,n}.$$

Двойственная задача (10.9) принимает следующий вид:

$$\sum_{j=1}^{m} y_j b_j \to \min,\ y_1 \ge 0, \ldots, y_m \ge 0,\ y_{m+1} \ge 0, \ldots, y_{m+n} \ge 0,$$

$$\overline{g} + \sum_{q=1}^{m+n} y_q \overline{g}_q = \overline{0}.$$

Связи типа равенств запишем в явном виде $-c_i + \sum_{j=1}^{m} y_j a_{ji} - y_{m+i} = 0$, $i = \overline{1,n}$. Учитывая, что числа $y_{m+i} \ge 0$, получим окончательный вид двойственной задачи

$$\sum_{j=1}^{m} y_j b_j \to \min,\ \sum_{j=1}^{m} y_j a_{ji} - c_i \ge 0,\ i = \overline{1,n};\ y_1 \ge 0, \ldots, y_m \ge 0. \quad (10.12)$$

Замечание 2. Из теоремы 3 получим, что решение в задаче (10.11) существует тогда и только тогда, когда существуют числа $y_1 \ge 0, \ldots, y_{m+n} \ge 0$ такие, что $\overline{g} + \sum_{q=1}^{m+n} \lambda_q \overline{g}_q = 0$. Расписывая, как и выше, эти равенства, получим, что решение в задаче (10.11) существует тогда и только тогда, когда существуют числа $y_1 \ge 0, \ldots, y_m \ge 0$ такие, что

$$\sum_{j=1}^{m} \lambda_j a_{ji} \ge c_i, \quad i = \overline{1, n}. \tag{10.13}$$

Замечание 3. Приведём матричные записи прямой и двойственной задачи линейного программирования. Для этого введём матрицу

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

и векторы

$$\overline{x} = \begin{pmatrix} x_1 \\ . \\ x_n \end{pmatrix}, \ \overline{c} = \begin{pmatrix} c_1 \\ . \\ c_n \end{pmatrix}, \ \overline{y} = \begin{pmatrix} y_1 \\ . \\ y_m \end{pmatrix}, \ \overline{b} = \begin{pmatrix} b_1 \\ . \\ b_m \end{pmatrix}.$$

Тогда задачи (10.11) и (10.12) принимают следующий вид:

$$\langle \overline{c}, \overline{x} \rangle \to \max, \ A\overline{x} \le \overline{b}, \ \overline{x} \ge \overline{0}; \tag{10.14}$$

$$\langle \overline{b}, \overline{y} \rangle \to \min, \ A^* \overline{y} \ge \overline{c}, \ \overline{y} \ge \overline{0}. \tag{10.15}$$

Здесь посредством A^* обозначена транспонированная матрица.

11. Метод линеаризации для нахождения минимального значения функции при линейных ограничениях

Рассматривается задача о нахождении минимума непрерывно дифференцируемой функции $f: R^n \to R$ на множестве, которое задаётся системой линейных неравенств и равенств:

$$f(\overline{x}) \to \min, \langle \overline{g}_i, \overline{x} \rangle - b_i \le 0, i = \overline{1,k},$$
$$\langle \overline{g}_j, \overline{x} \rangle - b_j = 0,\ j = \overline{k+1,m};\ \overline{x} \in R^n. \qquad (11.1)$$

Здесь $\overline{g}_S \in R^n, b_S \in R, s = \overline{1,m}$.

Предполагаем, что связи в задаче (11.1) совместны. Необходимые условия минимума в задаче (11.1) запишем с помощью правила множителей Лагранжа с множителем $\lambda_0 = 1$.

Теорема 1. Пусть точка $\overline{x}^*$ является решением задачи (11.1). Тогда существуют числа $\lambda_1, ..., \lambda_m$ такие, что

(*) $\lambda_1 \ge 0, ..., \lambda_k \ge 0$;

(**) $\lambda_i \left(\langle \overline{g}_i, \overline{x}^* \rangle - b_i \right) = 0, i = \overline{1,k}$;

(***) $f'(\overline{x}^*) + \sum_{s=1}^{m} \lambda_s \overline{g}_s = \overline{0}$.

Введём вспомогательную функцию $\overline{p}: R^n \to R^n$ такую, что если $\overline{p}(\overline{x}^*) = \overline{0}$, то в точке $\overline{x}^*$ выполнены необходимые условия (*)–(***) минимума функции $f(\overline{x})$. С этой целью в каждой точке $\overline{x} \in R^n$ рассматривается задача квадратичного программирования

$$\frac{1}{2} \langle \overline{p}, \overline{p} \rangle + \langle f'(\overline{x}), \overline{p} \rangle \to \min,$$

$$\langle \overline{g}_i, \overline{p} + \overline{x} \rangle - b_i \le 0, i = \overline{1,k}, \langle \overline{g}_j, \overline{p} + \overline{x} \rangle - b_j = 0,\ j = \overline{k+1,m}. \qquad (11.2)$$

Лемма 1. В каждой точке $\overline{x} \in R^n$ линейные относительно вектора $\overline{p} \in R^n$ связи (11.2) являются совместными.

Доказательство. Существует точка $\overline{x}_0 \in R^n$, которая удовлетворяет связям в задаче (11.1). Следовательно, точка $\overline{p} = \overline{x}_0 - \overline{x}$ удовлетворяет связям (11.2).

Отметим, что задача (11.2) является задачей квадратичного программирования.

Лемма 2. В каждой точке $\overline{x} \in R^n$ задача (11.2) имеет решение $\overline{p}(\overline{x})$.

Доказательство. Из неравенства

$$\frac{1}{2}\langle \overline{p}, \overline{p} \rangle + \langle f'(\overline{x}), \overline{p} \rangle \geq \frac{1}{2}\|\overline{p}\|^2 - \|f'(\overline{x})\|.\|\overline{p}\| \geq -\frac{1}{2}\|f'(\overline{x})\|^2$$

следует, что целевая функция в задаче (11.2) ограничена снизу. Следовательно, в задаче (11.2) решение $\overline{p}(\overline{x})$ существует.

Применяя к задаче (2) правило множителей Лагранжа, можно утверждать, что в каждой $\overline{x} \in R^n$ существуют числа $u_s(\overline{x})$, $s = \overline{1,m}$ такие, что

$$u_i(x) \geq 0, i = \overline{1,k}; \tag{11.3}$$

$$u_i(x)\left(\langle \overline{g}_i, \overline{p}(\overline{x}) \rangle + \langle \overline{g}_i, \overline{x} \rangle - b_i\right) = 0, i = \overline{1,k}; \tag{11.4}$$

$$\overline{p}(\overline{x}) + f'(\overline{x}) + \sum_{s=1}^{m} u_s(\overline{x})\overline{g}_s = \overline{0}. \tag{11.5}$$

Отсюда следует, что, если в точке $\overline{x}^* \in R^n$ вектор $\overline{p}(\overline{x}^*) = \overline{0}$, то числа $\lambda_s = u_s(\overline{x}^*)$, $s = \overline{1,m}$ удовлетворяют условиям (*)–(***), а сама эта точка $\overline{x}^*$ удовлетворяет связям в задаче (11.1). Стало быть, точка $\overline{x}^*$ удовлетворяет необходимым условиям минимума в этой задаче. Таким образом, нужно построить последовательность точек $\overline{x}_q$, удовлетворяющих связям в задаче (11.1), чтобы $\overline{p}(\overline{x}_q) \to \overline{0}$ при $q \to \infty$. Для этого применяется итерационный процесс

$$\overline{x}_{q+1} = \overline{x}_q + \gamma_q \overline{p}(\overline{x}_q), \ 0 < \gamma_q < 1. \tag{11.6}$$

Лемма 3. Пусть точка $\overline{x}_q$ удовлетворяет связям в задаче (11.1). Тогда точка $\overline{x}_{q+1}$ (11.6) удовлетворяет связям в задаче (11.1).

Доказательство. Умножим каждую связь в задаче (11.1) при $\overline{x} = \overline{x}_q$ на неотрицательное число $1-\gamma_q$, а каждую связь в задаче (11.2) при $\overline{p} = \overline{p}(\overline{x}_q)$, $\overline{x} = \overline{x}_q$ умножим на γ_q и результаты сложим. Получим

$$\left\langle \overline{g}_i, \overline{x}_q + \gamma_q \overline{p}(\overline{x}_q) \right\rangle - b_i \leq 0, i = \overline{1,k} \Leftrightarrow \left\langle \overline{g}_i, \overline{x}_{q+1} \right\rangle - b_i \leq 0, i = \overline{1,k};$$

$$\left\langle \overline{g}_j, \overline{x}_q + \gamma_q \overline{p}(\overline{x}_q) \right\rangle - b_j = 0, j = \overline{k+1,m} \Leftrightarrow \left\langle \overline{g}_j, \overline{x}_{q+1} \right\rangle - b_j = 0, j = \overline{k+1,m}.$$

Таким образом, если начальная точка $\overline{x}_0$ удовлетворяет связям в задаче (11.1), то последовательность точек $\overline{x}_q$ будет удовлетворять связям в задаче (11.1).

Следующая лемма даёт достаточные условия стремления к нулю последовательности $\overline{p}\left(\overline{x}_q\right)$.

Лемма 4. Пусть функция $f\left(\overline{x}\right)$ ограничена снизу на последовательности $\overline{x}_q \in R^n$ и пусть существует число $\alpha > 0$ такое, что

$$f\left(\overline{x}_q\right) - f\left(\overline{x}_{q+1}\right) \geq \alpha\left\|\overline{p}\left(\overline{x}_q\right)\right\|^2, \; q > 1. \qquad (11.7)$$

Тогда выполнено условие $\overline{p}\left(\overline{x}_q\right) \to \overline{0}$ при $q \to \infty$.

Доказательство. Ограниченная снизу числовая последовательность $f\left(\overline{x}_q\right)$ не возрастает. Следовательно, она имеет предел и поэтому разность, стоящая в левой части доказываемого неравенства (11.7), стремится к нулю.

В дальнейшем потребуется значение проекции градиента функции $f\left(\overline{x}\right)$ на направление $\overline{p}\left(\overline{x}\right)$, а именно

$$\left\langle f'\left(\overline{x}\right), \overline{p}\left(\overline{x}\right)\right\rangle = \sum_{s=1}^{m} u_s\left(\overline{x}\right)\left(\left\langle \overline{g}_s, \overline{x}\right\rangle - b_s\right) - \left\|\overline{p}\left(\overline{x}\right)\right\|^2. \qquad (11.8)$$

В самом деле, умножим векторное равенство (11.6) скалярно на $\overline{p}\left(\overline{x}\right)$. Получим

$$\left\langle f'\left(\overline{x}\right), \overline{p}\left(\overline{x}\right)\right\rangle = -\sum_{s=1}^{m} u_s\left(\overline{x}\right)\left\langle \overline{g}_s, \overline{p}\left(\overline{x}\right)\right\rangle - \left\|\overline{p}\left(\overline{x}\right)\right\|^2. \qquad (11.9)$$

Далее, умножим каждую j-ю связь типа равенства в задаче (11.2) на $u_j\left(\overline{x}\right)$ и просуммируем их вместе с равенствами (11.4). Будем иметь

$$-\sum_{s=1}^{m} u_s\left(\overline{x}\right)\left\langle \overline{g}_s, \overline{p}\left(\overline{x}\right)\right\rangle = \sum_{s=1}^{m} u_s\left(\overline{x}\right)\left(\left\langle \overline{g}_s, \overline{x}\right\rangle - b_s\right).$$

Подставим последнее выражение в формулу (11.9). Получим требуемое равенство (11.8).

Обоснование алгоритма

Будем предполагать, что градиент $f'\left(\overline{x}\right)$ удовлетворяет условию Липшица с константой $L > 0$

$$\left\|f'\left(\overline{x}\right) - f'\left(\overline{y}\right)\right\| \leq L\left\|\overline{x} - \overline{y}\right\|, \forall \overline{x} \in R^n, \forall \overline{y} \in R^n. \qquad (11.10)$$

Пример 1. Пусть $f(\overline{x}) = \langle A\overline{x}, \overline{x} \rangle + \langle \overline{g}, \overline{x} \rangle$. Здесь $A = (a_{ij})$ — симметричная матрица, $\overline{g}$ — вектор. Тогда

$$f'(\overline{x}) = 2A\overline{x} + \overline{g} \Rightarrow \left\| f'(\overline{x}) - f'(\overline{y}) \right\| =$$

$$= 2\left\| A(\overline{x} - \overline{y}) \right\| \le L\left\| \overline{x} - \overline{y} \right\|, L = 2\left(\sum_{i,j=1}^{n} a_{ij}{}^2 \right)^{1/2}.$$

Лемма 5. Пусть функция $f(\overline{x})$ удовлетворяет условию (11.10), а точка $\overline{x}_q$ удовлетворяет связям в задаче (11.1). Тогда для любого числа $0 < \gamma_q < 1$ верно неравенство

$$f\left(\overline{x}_q + \gamma_q \overline{p}(\overline{x}_q)\right) \le f(\overline{x}_q) - \gamma_q (1 - L\gamma_q) \left\| \overline{p}(\overline{x}_q) \right\|^2. \qquad (11.11)$$

Доказательство. По теореме о среднем существует число $0 \le \lambda \le 1$ такое, что

$$f\left(\overline{x}_q + \gamma_q \overline{p}(\overline{x}_q)\right) = f(\overline{x}_q) + \gamma_q \left\langle f'(\overline{y}), \overline{p}(\overline{x}_q) \right\rangle, \overline{y} = \overline{x}_q - \lambda\gamma_q \overline{p}(\overline{x}_q). \ (11.12)$$

Используя неравенство Коши — Буняковского и условие Липшица (11.10), получим

$$\left\langle f'(y), p(\overline{x}_q) \right\rangle - \left\langle f'(\overline{x}_q), \overline{p}(\overline{x}_q) \right\rangle \le$$

$$\le \left\| f'(\overline{y}) - f'(\overline{x}_q) \right\| \cdot \left\| \overline{p}(\overline{x}_q) \right\| \le L \left\| \overline{y} - \overline{x}_q \right\| \cdot \left\| \overline{p}(\overline{x}_q) \right\|.$$

Отсюда и из второй формулы в (11.12) следует, что

$$\left\langle f'(y), p(\overline{x}_q) \right\rangle \le \left\langle f'(\overline{x}_q), \overline{p}(\overline{x}_q) \right\rangle + \gamma_q \left\| \overline{p}(\overline{x}_q) \right\|^2.$$

Подставим это неравенство в первую формулу (11.12). Получим

$$f\left(\overline{x}_q + \gamma_q \overline{p}(\overline{x}_q)\right) \le f(\overline{x}_q) + \gamma_q \left\langle f'(\overline{x}_q), \overline{p}(\overline{x}_q) \right\rangle + \gamma_q{}^2 L \left\| \overline{p}(\overline{x}_q) \right\|^2. \ (11.13)$$

Точка $\overline{x}_q$ удовлетворяет связям в задаче (11.1), а числа $u_i(\overline{x}) \ge 0$, $i = \overline{1,k}$. Поэтому из формулы (11.10) получим, что

$$\left\langle f'(\overline{x}_q), \overline{p}(\overline{x}_q) \right\rangle \le -\left\| \overline{p}(\overline{x}_q) \right\|^2.$$

Отсюда и из (11.13) следует требуемое неравенство (11.11).

Метод с постоянным шагом

Пусть численное значение константы Липшица L в (11.10) известно. Возьмём

$$\gamma_q = \gamma, 0 < \gamma < \min\left(1; \frac{1}{L}\right).$$

Тогда из (11.11) следует неравенство (11.7) с $\alpha = \gamma(1 - L\gamma) > 0$. Начальное значение $\overline{x}_0$ выбирается так, чтобы выполнялись ограничения в задаче (11.1).

Метод с дроблением шага

Пусть численное значение константы Липшица L в (11.10) неизвестно.

Начальное значение $\overline{x}_0$ выбирается так, чтобы выполнялись ограничения в задаче (11.1).

Описание метода:

Выбираем числа $0 < \varepsilon < 1, 0 < \gamma < 1, 0 < \lambda < 1$.

Пусть значение $\overline{x}_q$ вычислено.

1. Полагаем $\gamma_q = \gamma$.

2. Вычисляем $\overline{x} = \overline{x}_q + \gamma_q \overline{p}\left(\overline{x}_q\right)$.

3. Проверяем неравенство

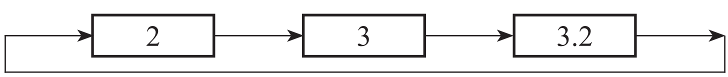
$$\text{(A)} \qquad f\left(\overline{x}\right) - f\left(\overline{x}_q\right) \le -\varepsilon\gamma_q \left\|\overline{p}\left(\overline{x}_q\right)\right\|^2$$

3.1. Если неравенство (А) выполнено, то полагаем $\overline{x}_q = \overline{x}$ и идём к 1.

3.2. Если неравенство (А) не выполнено, то полагаем $\gamma_q = \lambda\gamma_q$ и идём к 2.

Обоснование метода

Пусть в процессе счёта приходим к пункту 3.2. В этом случае цикл

(Б) 

просчитывается некоторое i + 1 раз. Тогда $\gamma_q = \lambda^i\gamma$ и $x = x_k - \lambda^i \gamma p\left(x_k\right)$. Из неравенства (11.11) получим, что

$$f\left(\overline{x}\right) - f\left(\overline{x}_q\right) \le -\lambda^i\gamma\left(1 - L\lambda^i\gamma\right)\left\|\overline{p}\left(\overline{x}_q\right)\right\|^2.$$

Отсюда видно, что неравенство (А) будет выполнено при

$$\varepsilon\lambda^i\gamma \le \lambda^i\gamma\left(1 - L\lambda^i\gamma\right) \Leftrightarrow \lambda^i\gamma \le \frac{1-\varepsilon}{L}.$$

Поэтому цикл (Б) закончится на конечном шаге, причём

$$\lambda\frac{1-\varepsilon}{L} < \lambda\gamma \le \frac{1-\varepsilon}{L} \Rightarrow \lambda\frac{1-\varepsilon}{L} < \gamma_q \le \frac{1-\varepsilon}{L}.$$

Отсюда и из неравенства (А) следует, что в процессе счета выполняется неравенство (11.7) с

$$\alpha = \varepsilon\min\left(\gamma; \lambda\frac{1-\varepsilon}{L}\right).$$

12. Проблема моментов и её приложения в задачах управления движением

Задача о выводе ракеты в заданный момент времени на заданную высоту при достижении заданной скорости и минимизации расхода топлива в формализованном виде имеет

$$\int_0^T u^2(t)dt \to \min, \int_0^T (T-t)u(t)dt = c_1, \int_0^T u(t)dt = c_2 . \qquad (12.1)$$

Здесь c_1 и c_2 — заданные числа.

Обозначим через $L_2[0,T]$ линейное пространство измеримых функций $h:[0,T]\to R$ с нормой

$$\rho\big(h(\cdot)\big) = \sqrt{\int_0^T h^2(t)dt} .$$

Известно, что общий вид линейного функционала $\varphi_u : L_2[0,T]\to R$ следующий:

$$<\varphi_u, h(\cdot)> = \int_0^T h(t)u(t)dt .$$

Его норма равна

$$\rho_*(\varphi_u) = \sqrt{\int_0^T u^2(t)dt} .$$

Задачу (12.1) можно рассматривать в следующей формализации: заданы функции $h_1(t) = T-t, h_2(t) = 1$ как элементы пространства $L_2[0,T]$ и числа c_1 и c_2. Среди всех линейных функционалов $\varphi_u : L_2[0,T]\to R$, удовлетворяющих ограничениям $<\varphi_u, h_1(\cdot)> = c_1$, $<\varphi_u, h_2(\cdot)> = c_2$, требуется найти такой, норма $\rho_*(\varphi_u)$ которого является минимальной.

Сформулируем теперь рассматриваемую задачу в общем виде. Задано линейное нормированное пространство E с нормой $\|h\|$, $h\in E$. Заданы элементы $h_i \in E, i=\overline{1,m}$ и числа $c_i \in R, i=\overline{1,m}$, среди которых хотя бы одно отлично от нуля.

Проблема моментов. Среди всех линейных функционалов $\varphi : E\to R$, удовлетворяющих ограничениям

$$\varphi(h_1) = c_1, \ldots, \varphi(h_m) = c_m, \tag{12.2}$$

требуется найти такой, норма $\|\varphi\|_*$ которого является минимальной.

Замечание 1. Функционал $\varphi : E \to R$ называется линейным, если

1) $\varphi(\alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 \varphi(h_1) + \alpha_2 \varphi(h_2)$ для любых $h_1, h_2 \in E; \alpha_1, \alpha_2 \in R$;

2) существует число $M > 0$ такое, что $|\varphi(h)| \le M \|h\|$ для $\forall h \in E$.

Нормой линейного функционала $\varphi : E \to R$ называется число $\|\varphi\|_*$, которое определяется равенством

$$\|\varphi\|_* = \sup_{\|h\| \le 1} |\varphi(h)| = \sup_{h \ne 0} \frac{|\varphi(h)|}{\|h\|}.$$

Из этого определения следует, что

$$|\varphi(h)| \le \|\varphi\|_* \|h\| \text{ для } \forall h \in E. \tag{12.3}$$

Проведём оценку нормы линейного функционала, удовлетворяющего ограничениям (12.2). С этой целью возьмём любой набор чисел $\lambda_i, i = \overline{1, m},$ удовлетворяющий равенству

$$\sum_{i=1}^{m} \lambda_i c_i = 1. \tag{12.4}$$

Умножим каждое i-е ограничение в (12.2) на число λ_i и сложим. Получим

$$\varphi\left(\sum_{i=1}^{m} \lambda_i h_i\right) = 1.$$

Отсюда и из неравенства (12.3) следует, что

$$1 \le \|\varphi\|_* \left\| \sum_{i=1}^{m} \lambda_i h_i \right\|.$$

Обозначим

$$f_0 = \inf_{\{\lambda_i\}} \left\| \sum_{i=1}^{m} \lambda_i h_i \right\|, \sum_{i=1}^{m} \lambda_i c_i = 1. \tag{12.5}$$

Тогда из предыдущего неравенства получим, что

$$\|\varphi\|_* f_0 \ge 1. \tag{12.6}$$

Следовательно, если $f_0 = 0$, то не существует линейного функционала $\varphi: E \to R$, удовлетворяющего ограничениям (12.2).

Введём в рассмотрение функцию

$$f(\lambda_1, \lambda_2, ..., \lambda_m) = \left\| \sum_{i=1}^{m} \lambda_i h_i \right\|. \qquad (12.7)$$

Теорема 1. Функция (12.7) удовлетворяет условию Липшица

$$| f(\lambda_1, \lambda_2, ..., \lambda_m) - f(\nu_1, \nu_2, ..., \nu_m) | \le L \sum_{i=1}^{m} | \lambda_i - \nu_i |, \quad L = \max_{1 \le i \le m} \| h_i \|.$$

Доказательство. Из неравенства

$$| \|h\| - \|h_*\| | \le | \|h - h_*\| |, \; h, h_* \in E$$

следует, что

$$| f(\lambda_1, \lambda_2, ..., \lambda_m) - f(\nu_1, \nu_2, ..., \nu_m) | \le \left\| \sum_{i=1}^{m} (\lambda_i - \nu_i) h_i \right\|.$$

Отсюда, применяя неравенство треугольника

$$\left\| \sum_{i=1}^{m} (\lambda_i - \nu_i) h_i \right\| \le \sum_{i=1}^{m} | \lambda_i - \nu_i | \|h_i\|,$$

получим требуемое неравенство

$$| f(\lambda_1, \lambda_2, ..., \lambda_m) - f(\nu_1, \nu_2, ..., \nu_m) | \le \sum_{i=1}^{m} | \lambda_i - \nu_i | \|h_i\| \le L \sum_{i=1}^{m} | \lambda_i - \nu_i |.$$

Теорема 2. Пусть вектор $h_i \in E, i = \overline{1, m}$ является линейно независимым. Тогда нижняя грань в задаче (12.5) достигается и её значение $f_0 > 0$.

Доказательство. Из линейной независимости векторов h_i следует, что для любого ненулевого набора чисел $\lambda_i, i = \overline{1, m}$ значение функции (12.7) строго больше нуля. В самом деле,

$$f(\lambda_1, \lambda_2, ..., \lambda_m) = 0 \Leftrightarrow \sum_{i=1}^{m} \lambda_i h_i = 0 \Leftrightarrow \lambda_i = 0 \text{ для всех } i = \overline{1, m}.$$

В пространстве R^m рассмотрим множество

$$S = \{(\lambda_1, \lambda_2, ..., \lambda_m) \in R^m : \max_{1 \le i \le m} | \lambda_i | = 1\}.$$

Это множество является замкнутым и ограниченным. Поскольку функция f удовлетворяет условию Липшица, то она является непрерывной. Отсюда и из неравенства $f(\lambda_1, \lambda_2, ..., \lambda_m) > 0$ следует, что её минимальное значение на множестве S строго больше нуля. Стало быть, существует число $A > 0$ такое, что

$$\left\| \sum_{i=1}^{m} \lambda_i h_i \right\| \ge A > 0 \text{ при } \max_{1 \le i \le m} | \lambda_i | = 1.$$

Пусть $\lambda_1, \lambda_2, ..., \lambda_m$ — произвольный ненулевой набор чисел. Обозначим $\lambda^* = \max\limits_{1 \le i \le m} |\lambda_i|$, $\lambda_i^* = \dfrac{\lambda_i}{\lambda^*}$. Тогда $\max\limits_{1 \le i \le m} \lambda_i^* = 1$. Следовательно,

$$\left\| \sum_{i=1}^{m} \frac{\lambda_i}{\lambda^*} h_i \right\| \ge A \Rightarrow \left\| \sum_{i=1}^{m} \lambda_i h_i \right\| \ge A \max_{1 \le i \le m} |\lambda_i| . \qquad (12.8)$$

Очевидно, что последнее неравенство выполнено и для нулевого набора чисел λ_i.

Возьмём число $\varepsilon > 0$ и рассмотрим множество

$$X = \{(\lambda_1, \lambda_2, ..., \lambda_m) \in R^m : \sum_{i=1}^{m} \lambda_i c_i = 1, \left\| \sum_{i=1}^{m} \lambda_i h_i \right\| \le f_0 + \varepsilon\} .$$

Из неравенства (12.8) следует, что это множество ограничено, а из непрерывности функции *f* следует его замкнутость. Очевидно, что

$$f_0 = \inf_{\overline{\lambda} \in X} f(\lambda_1, \lambda_2, ..., \lambda_m), \overline{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_m) .$$

Множество *X* является замкнутым и ограниченным, а функция *f* является непрерывной. Поэтому функция *f* достигает на множестве *X* своего наименьшего значения на некотором наборе $\lambda_i^0, i = \overline{1, m}$. Это значит, что

$$f_0 = \left\| \sum_{i=1}^{m} \lambda_i^0 h_i \right\|, \sum_{i=1}^{m} \lambda_i^0 c_i = 1 .$$

Поскольку значение функции *f* строго больше нуля, то и $f_0 > 0$.

Рассмотрим задачу

$$\|\varphi\|_* \to \min, \varphi(h_i) = c_i, i = \overline{1, m} \qquad (12.9)$$

Теорема 3. Пусть вектор $h_i \in E, i = \overline{1, m}$ является линейно-независимым. Тогда решение в задаче (12.9) существует и норма оптимального функционала равна f_0^{-1}.

Доказательство. Из равенства (12.6) следует, что для любого линейного функционала $\varphi : E \to R$, удовлетворяющего ограничениям (12.2), выполняется неравенство $\|\varphi\|_* \ge f_0^{-1}$. Построим функционал $\varphi_0 : E \to R$, удовлетворяющий ограничениям (12.2), нормой $\|\varphi\|_* \ge f_0^{-1}$.

С этой целью рассмотрим линейное пространство $E_m \subset E$, натянутое на векторы $h_i, i = \overline{1, m}$. В этом пространстве строим линейный функционал

$$\varphi(\sum_{i=1}^{m}\lambda_i h_i)=\sum_{i=1}^{m}\lambda_i c_i\ .$$

Тогда $\varphi(h_i)=c_i,\ i=\overline{1,m}$. Его норма равна

$$\|\varphi\|_{E_m}=\sup_{\{\lambda_i\}}\frac{\left|\sum_{i=1}^{m}\lambda_i c_i\right|}{\left\|\sum_{i=1}^{m}\lambda_i h_i\right\|}\ .$$

Сделаем в этой формуле замену $\nu_i=\dfrac{\lambda_i}{\sum_{i=1}^{m}\lambda_S c_S}$. Получим

$$\|\varphi\|_{E_m}=\sup_{\{\nu_i\}}\frac{1}{\left\|\sum_{i=1}^{m}\nu_i h_i\right\|}=\frac{1}{\inf\limits_{\{\nu_i\}}\left\|\sum_{i=1}^{m}\nu_i h_i\right\|}=\left|\sum_{i=q}^{m}\nu_i c_i=1\right|=\frac{1}{f_0}\ .$$

Применяя теорему Хана — Банаха, продолжим этот функционал φ на всё пространство E с сохранением нормы. Получим требуемый функционал φ_0.

Рассмотрим случай, когда вектора $h_i,\ i=\overline{1,m}$ являются линейно зависимыми. Считаем, что первые k векторов линейно независимы, а для остальных верно разложение

$$h_j=\sum_{i=1}^{k}\alpha_{ji}h_i,\ j=\overline{k+1,m};a_{ji}\in R\ . \qquad (12.10)$$

Теорема 4. Чтобы задача (12.9) имела решение, необходимо и достаточно, чтобы выполнялись равенства

$$c_j=\sum_{i=1}^{k}\alpha_{ji}c_i,\ j=\overline{k+1,m}\ . \qquad (12.11)$$

Доказательство. *Необходимость*. Пусть задача (12.9) имеет решение $\varphi_0:E\to R$. Следовательно, при $j=\overline{k+1,m}$ выполнены равенства

$$c_j=\varphi_0(h_j)=\left|(12.10)\right|=\varphi_0(\sum_{i=1}^{k}\alpha_{ji}h_i)=\sum_{i=1}^{k}\alpha_{ji}\varphi_0(h_i)=\sum_{i=1}^{k}\alpha_{ji}c_i\ .$$

Достаточность. Пусть выполнены неравенства (12.11). Тогда набор чисел $c_i,\ i=\overline{1,k}$ ненулевой. В противном случае все числа $c_S=0,\ s=\overline{1,m}$.

Пусть функционал $\varphi:E\to R$ удовлетворяет связи $\varphi(h_i)=c_i,\ i=\overline{1,k}$. Тогда при $j=\overline{k+1,m}$

$$\varphi(h_j) = |(12.10)| = \varphi(\sum_{i=1}^{k} \alpha_{ji} h_i) = \sum_{i=1}^{k} \alpha_{ji} \varphi(h_i) = \sum_{i=1}^{k} \alpha_{ji} c_i = c_j.$$

Следовательно, задача (12.9) эквивалентна задаче

$$\|\varphi\|_* \to \min, \varphi(h_i) = c_i, i = \overline{1,k} . \qquad (12.12)$$

Задача (12.12) имеет решение. Поэтому задача (12.9) также имеет решение и оптимальные функционалы в этих задачах совпадают.

Как было показано раньше, норма $\|\varphi_0\|_*$ оптимального функционала $\varphi_0 : E \to R$ равна f_0^{-1}, где

$$f_0 = \min_{\{\lambda_i\}} \left\| \sum_{i=1}^{k} \lambda_i h_i \right\|, \sum_{i=1}^{k} \lambda_i c_i = 1 . \qquad (12.13)$$

Обозначим

$$f_0 = \min_{\{\lambda_i\}} \left\| \sum_{i=1}^{m} \lambda_i h_i \right\|, \sum_{i=1}^{m} \lambda_i c_i = 1 . \qquad (12.14)$$

Теорема 5. Пусть выполнены равенства (12.11). Тогда $f_0 = f_0^*$.

Доказательство. Если брать в (12.14) наборы чисел $\lambda_i, i = \overline{1,m},$ у которых числа $\lambda_j = 0, j = \overline{k+1,m}$, то, согласно формуле (12.13), получим неравенство $f_0 \le f_0^*$.

Пусть числа $\lambda_i, i = \overline{1,m}$ являются решением задачи (12.14). Положим

$$\hat{\lambda}_i = \lambda_i + \sum_{j=k+1}^{m} \alpha_{ji} \lambda_j, i = \overline{1,k} .$$

Тогда

$$\sum_{i=1}^{k} \hat{\lambda}_i c_i = \sum_{i=1}^{k} \lambda_j c_i + \sum_{i=1}^{k} (\sum_{j=k+1}^{m} \alpha_{ji} \lambda_j) c_i = \sum_{i=1}^{k} \lambda_j c_i + \sum_{i=1}^{k} (\sum_{j=k+1}^{m} \alpha_{ji} c_i) \lambda_j .$$

Отсюда, используя равенства (12.11), получим, что

$$\sum_{i=1}^{k} \hat{\lambda}_i c_i = \sum_{i=1}^{k} \lambda_i c_i + \sum_{j=k+1}^{m} \lambda_j c_j = 1.$$

Далее,

$$\sum_{i=1}^{k} \hat{\lambda}_j h_i = \sum_{i=1}^{k} \lambda_i h_i + \sum_{i=1}^{k} (\sum_{j=k+1}^{m} \alpha_{ji} \lambda_j) h_i = \sum_{i=1}^{k} \lambda_i h_i + \sum_{j=k+1}^{m} (\sum_{i=1}^{k} \alpha_{ji} h_i) \lambda_j .$$

Подставляя сюда формулы (12.10), будем иметь

$$\sum_{i=1}^{k}\hat{\lambda}_i h_i = \sum_{i=1}^{k}\lambda_i h_i + \sum_{j=k+1}^{m}\lambda_j h_j = \sum_{s=1}^{k}\lambda_S h_S .$$

Из этих равенств и из формулы (12.14) получим неравенство $f_0 \le f_0^*$.

Теорема 6. Чтобы задача (12.9) имела решение, необходимо и достаточно, чтобы в задаче (12.14) число $f_0 > 0$.

Доказательство. *Необходимость*. Пусть задача (12.9) имеет решение. Тогда, согласно теореме 4, выполнены равенства (12.11). Из теоремы 5 следует, что $f_0 > 0$.

Достаточность. Пусть $f_0 > 0$, но для некоторого номера $j = \overline{k+1,m}$ равенство (12.11) не выполнено. Пусть, например, $c_m \ne \sum_{i=1}^{k}\alpha_{mi}c_i$.

Возьмём пока произвольное число $\lambda \in R$ и положим

$$\lambda_i = -\lambda_{mi} \text{ при } i = \overline{1,k}\text{ ; } \lambda_{k+1} = 0,\dots,\lambda_{m-1} = 0, \lambda_m = \lambda .$$

Тогда

$$\sum_{S=1}^{k}\lambda_S h_S = \sum_{i=1}^{k}(-\alpha_{mi}h_i)\lambda + \lambda h_m = (h_m - \sum_{i-1}^{k}\alpha_{mi}h_i)\lambda = 0,$$

$$\sum_{S=1}^{k}\lambda_S c_S = \sum_{i=1}^{k}(-\alpha_{mi}c_i)\lambda + \lambda c_m = (c_m - \sum_{i-1}^{k}\alpha_{mi}c_i)\lambda .$$

Поэтому, если взять число

$$\lambda = \frac{1}{c_m - \sum_{i-1}^{k}\alpha_{mi}c_i},$$

то получим, что

$$\sum_{s=1}^{m}\lambda_S h_S = 0, \sum_{s=1}^{m}\lambda_S c_S = 1 .$$

Отсюда и из формулы (12.15) следует равенство $f_0 = 0$.

Пусть $\lambda_i^0, i = \overline{1,m}$ — решение в задаче (12.14). Положим

$$h_0 = \sum_{i=1}^{m}\lambda_i^0 h_i . \qquad (12.15)$$

Тогда $f_0 = \|h_0\|$.

Теорема 7 (*принцип максимума*). Оптимальный функционал $\varphi_0 : E \to R$ в задаче (12.9) выделяется среди всех линейных функционалов $\varphi : E \to R$, норма которых $\|\varphi\|_* = \dfrac{1}{f_0}$, следующим свойством:

$$\varphi_0(h_0) = \max_{\|\varphi\|_* = \frac{1}{f_0}} \varphi(h_0) = 1. \qquad (12.16)$$

Доказательство. Имеем

$$\varphi_0(h_0) = \varphi_0\left(\sum_{i=1}^{m} \lambda_i^0 h_i\right) = \sum_{i=1}^{m} \lambda_i^0 \varphi(h_i) = \sum_{i=1}^{m} \lambda_i^0 c_i = 1. \qquad (12.17)$$

Далее, для любого линейного функционала $\varphi : E \to R$ с нормой $\|\varphi\|_* = \dfrac{1}{f_0}$ выполнено неравенство

$$\varphi(h_0) \le \|\varphi\|_* \|h_0\| = \frac{1}{f_0} f_0 = 1. \qquad (12.18)$$

Рассмотрим случай, когда пространство E является *гильбертовым*. Обозначим через $<,>$ скалярное произведение в E. Тогда $E = E^*$ и, следовательно, $\varphi(h) = <\varphi, h>$, $\|h\| = \sqrt{<h,h>}$, $\|h\|_* = \sqrt{<h,h>}$.

Задача (12.16) принимает следующий вид:

$$<\varphi_0, h_0> = \max_{\varphi} <\varphi, h_0> = 1, \sqrt{<\varphi,\varphi>} = \frac{1}{f_0}. \qquad (12.19)$$

Покажем, что задача (12.19) имеет единственное решение и оно равно

$$\varphi_0 = \frac{1}{f_0^2} h_0. \qquad (12.20)$$

В самом деле,

$$<\varphi_0, h_0> = \frac{1}{f_0^2} <h_0, h_0> = \frac{1}{f_0^2} f_0^2 = 1, <\varphi_0, \varphi_0> = \frac{1}{f_0^4} <h_0, h_0> = \frac{1}{f_0^2}.$$

Допустим, что имеется ещё φ такой, что

$$<\varphi_0, h_0> = 1, <\varphi, \varphi> = \frac{1}{f_0^2}.$$

Тогда, обозначая $\psi = \varphi - \varphi_0$, будем иметь, что $\varphi = \varphi_0 + \psi$. Далее,

$$1 = <\varphi_0, h_0> = <\varphi - \psi, h_0> = <\varphi, h_0> - <\psi, h_0> = 1 - <\psi, h_0>.$$

Следовательно, $<\psi, h_0> = 0$. Отсюда и из формулы (12.20) получим, что $<\varphi_0, \psi> = 0$. Далее,

$$\frac{1}{f_0^2} = <\varphi, \varphi> = <\varphi_0 + \psi, \varphi_0 + \psi> =$$

$$= <\varphi_0, \varphi_0> + 2 <\varphi_0, \psi> + <\psi, \psi> = \frac{1}{f_0^2} - <\psi, \psi>.$$

Отсюда получим, что $<\psi,\psi>=0$ и, следовательно, $\psi=0$.

Таким образом, решение задачи (12.19) задаётся формулой (12.20).

Запишем задачу (12.15), с помощью которой находится элемент h_0. Поскольку минимизация нормы вектора эквивалентна задаче минимизации квадрата нормы, то получим следующую задачу:

$$<\sum_{i=1}^{m}\lambda_i h_i,\sum_{i=1}^{m}\lambda_i h_i>\to\min,\ \sum_{i=1}^{m}\lambda_i c_i=1.$$

Обозначим

$$a_{ij}=<h_i,h_j>;\quad i,j=\overline{1,m}. \tag{12.21}$$

Тогда будем иметь следующую задачу:

$$f_0^2=\inf_{\{\lambda_i\}}\sum_{i,j=1}^{m}a_{ij}\lambda_i\lambda_j,\ \ \sum_{i=1}^{m}\lambda_i c_i=1. \tag{12.22}$$

Целевая функция в этой задаче является неотрицательной квадратичной функцией. Если же векторы $h_i, i=\overline{1,m}$ являются линейно независимыми, то эта целевая функция является положительно определённой квадратичной функцией. В этом случае матрица $A=\{a_{ij}\}$ является невырожденной. Поскольку в этой задаче имеется только линейная связь, то запишем функцию Лагранжа с $\lambda_0=1$

$$L=\sum_{i,j=1}^{m}a_{ij}\lambda_i\lambda_j+\lambda\left(\sum_{i=1}^{m}\lambda_i c_i-1\right).$$

Это выпуклая функция. Поэтому условия минимума получаются из условий равенства нулю её частных производных. Запишем эти условия:

$$2\sum_{j=1}^{m}a_{ij}\lambda_j+\lambda c_i=0,\ \ i=\overline{1,m}.$$

В случае если матрица A не является вырожденной, получим

$$\bar{\lambda}=-\frac{\lambda}{2}A^{-1}\bar{c},\ \ \bar{\lambda}=\begin{pmatrix}\lambda_1\\ \dots\\ \lambda_m\end{pmatrix},\ \ \bar{c}=\begin{pmatrix}c_1\\ \dots\\ c_m\end{pmatrix}.$$

Обозначим скалярное произведение векторов $\bar{\varphi}$ и $\bar{\psi}$ из R^m посредством $(\bar{\varphi},\bar{\psi})$. Тогда связь (12.22) можно записать в виде $(\bar{\lambda},\bar{c})=1$. Подставим найденные выражения для $\bar{\lambda}$ в эту связь. Получим

$$1 = (\bar{\lambda}, \bar{c}) = -\frac{\lambda}{2}\left(A^{-1}\bar{c}, \bar{c}\right) \Rightarrow \lambda = -2\left(A^{-1}\bar{c}, \bar{c}\right)^{-1}.$$

Следовательно,

$$\bar{\lambda}^0 = \frac{A^{-1}\bar{c}}{\left(A^{-1}\bar{c}, \bar{c}\right)}. \tag{12.23}$$

Из формул (12.22) и (12.23) получим, что

$$f_0^2 = \left(A\bar{\lambda}^0, \bar{\lambda}^0\right) = \left(\frac{\bar{c}}{\left(A^{-1}\bar{c}, \bar{c}\right)}, \frac{A^{-1}\bar{c}}{\left(A^{-1}\bar{c}, \bar{c}\right)}\right) = \frac{1}{\left(A^{-1}\bar{c}, \bar{c}\right)}.$$

Стало быть, норма оптимального линейного функционала равна

$$\|\varphi_0\|_* = \frac{1}{f_0} = \sqrt{\left(A^{-1}\bar{c}, \bar{c}\right)}. \tag{12.24}$$

Обозначим элементы матрицы A^{-1} через $\hat{a}_{ij}$. Тогда

$$h_0 = \frac{1}{\left(A^{-1}\bar{c}, \bar{c}\right)} \sum_{i=1}^{m} \left(\sum_{j=1}^{m} \hat{a}_{ij} c_j\right) h_i.$$

Отсюда, используя формулы (12.20) и (12.24), получим, что оптимальный линейный функционал равен

$$\varphi_0 = \sum_{i=1}^{m} \left(\sum_{j=1}^{m} \hat{a}_{ij} c_j\right) h_i. \tag{12.25}$$

Управление с минимальной энергией

Сформулированная задача (12.1) является частным случаем задачи управления с минимальной энергией

$$\int_{\tau}^{T} \sum_{s=1}^{n} u_s^2(t)dt \to \min, \quad \int_{\tau}^{T} \sum_{s=1}^{n} h_{is}(t) u_s(t)dt = c_i, \quad i = \overline{1,m}. \tag{12.26}$$

Здесь $\tau < T$ — фиксированные начальный и конечный моменты времени; c_s, $s = \overline{1,m}$ — заданный ненулевой набор чисел; каждая из заданных функций $h_{si} : [\tau, T] \to R$ является измеримой с суммируемым на этом отрезке квадратом.

Обозначим через $L_2^{(n)}[\tau, T]$ линейное пространство измеримых на $[\tau, T]$ вектор — функций $\bar{h}(t) = \left(h_1(t), \ldots, h_n(t)\right)$ с нормой

$$\left\|\bar{h}(\cdot)\right\| = \left(\int_{\tau}^{T} \sum_{s=1}^{n} h_s^2(t)dt\right)^{1/2}.$$

Это пространство является гильбертовым со скалярным произведением

$$\left\langle \overline{h}(\cdot), \overline{v}(\cdot) \right\rangle = \int_{\tau}^{T} \sum_{s=1}^{n} h_s(t) v_s(t) dt.$$

Каждый линейный ограниченный функционал задаётся формулой

$$\left\langle \varphi_u, \overline{h}(\cdot) \right\rangle = \int_{\tau}^{T} \sum_{s=1}^{n} h_s(t) u_s(t) dt,$$

здесь $\overline{u}(t) = \left(u_1(t), \ldots, u_n(t)\right)$ — вектор-функция из пространства $L_2^{(n)}\left[\tau, T\right]$. Норма функционала φ_u задаётся формулой

$$\left\| \varphi_u \right\|_* = \left(\int_{\tau}^{T} \sum_{s=1}^{n} u_s^2(t) dt \right)^{1/2}.$$

Числа a_{ij}, определяемые формулой (12.21), в рассматриваемом случае имеют вид

$$a_{ij} = \int_{\tau}^{T} \sum_{s=1}^{n} h_{is}(t) h_{js}(t) dt. \qquad (12.27)$$

Из формулы (12.25) найдём вектор-функцию $\overline{u}^{(0)}(t) = \left(\overline{u}_1^{(0)}(t), \ldots, \overline{u}_n^{(0)}(t)\right)$, которая определяет оптимальный функционал

$$u_s^{(0)}(t) = \sum_{i=1}^{m} \left(\sum_{j=1}^{m} a_{ij} c_j \right) h_{is}(t), \quad s = \overline{1, n}. \qquad (12.28)$$

Пример 2. Рассмотрим задачу (12.1). Имеем

$$h_1(t) = T - t, \, h_2(t) = 1.$$

Из формулы (12.27) следует, что

$$a_{11} = \int_0^T h_1^2(t) dt = \int_0^T (T - t)^2 dt = \frac{T^3}{3}; \quad a_{22} = \int_0^T h_2^2(t) dt = T;$$

$$a_{12} = \int_0^T h_1(t) h_2(t) dt = \int_0^T (T - t) dt = \frac{T^2}{2} = a_{21}. \qquad (12.29)$$

Определитель матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \dfrac{T^3}{3} & \dfrac{T^2}{2} \\ \dfrac{T^2}{2} & T \end{pmatrix}$$

равен $\dfrac{T^4}{12} \neq 0,$ при $T \neq 0.$ Обратная матрица равна

$$A^{-1} = \begin{pmatrix} \dfrac{12}{T^3} & -\dfrac{6}{T^2} \\ -\dfrac{6}{T^2} & \dfrac{4}{T} \end{pmatrix}.$$

Поэтому формула (12.28) принимает вид

$$u_s^{(0)}(t) = \sum_{i=1}^{2} \left(\sum_{j=1}^{2} a_{ij} c_j \right) h_i(t) = \left(\sum_{j=1}^{2} a_{ij} c_j \right) (T - t) + \sum_{j=1}^{2} a_{2j} c_j.$$

Отсюда и из (12.29) получим, что

$$u_s^{(0)}(t) = \left(\frac{12}{T^3} c_1 - \frac{6}{T^2} c_2 \right) (T - t) + \left(-\frac{6}{T^2} c_1 + \frac{4}{T} c_2 \right) =$$

$$= \frac{2}{T} \left(\frac{3}{T} c_1 - c_2 \right) + \frac{6}{T^2} \left(c_2 - \frac{2}{T} c_1 \right) t.$$

Управление с минимальной силой

Пример 3. Рассмотрим задачу о вертикальном взлёте ракеты в заданный момент времени T на заданную высоту H с достижением нулевой скорости. Уравнение Мещерского имеет вид

$$\dot{x} = v, \dot{v} = u, \text{ где } u = -g - w\frac{\dot{m}(t)}{m(t)}.$$

В начальный момент времени $x(0) = 0$ и $v(0) = 0.$ Величина $|u(t)|$ определяет величину перегрузок, испытываемых приборами. Поэтому целью выбора управления наряду с выше сформулированными является минимизация наибольшей величины $|u(t)|.$

Таким образом, имеем задачу

$$\sup_{0 \le t \le T} |u(t)| \to \min, \quad \int_0^T (T - t) u(t) dt = H, \quad \int_0^T u(t) dt = 0. \qquad (12.30)$$

В качестве пространства E возьмём пространство $L_1[0, T],$ которое является пространством функций $h: [0, T] \to R,$ суммируемых по Лебегу, с нормой

$$\|h(\cdot)\| = \int_0^T |h(t)| dt.$$

Общий вид линейного функционала $\varphi: L_1[0, T] \to R$ задаётся формулой

$$\varphi_u(h(\cdot)) = \int_0^T h(t) u(t) dt,$$

где $u(t)$ — почти всюду ограниченная измеримая функция и

$$\|\varphi_u\|_* = vrai \max_{[0,T]} |u(t)|.$$

Последняя формула означает следующее. Существует множество $I \subset [0,T]$, мера которого равна нулю, и число $M > 0$ такие, что

$$|u(t)| \le M \text{ для всех } t \in [0,T], t \notin I.$$

Норма строится так: берётся множество $\Upsilon \subset [0, T] \subset \mu(\Upsilon) = 0$ (мера Лебега его равна нулю). Вычисляется

$$\inf_{\Upsilon:\mu(\Upsilon)=0} \sup_{t\in[0,T]\setminus\Upsilon} |u(t)| = vrai \max_{[0,T]} |u(t)|.$$

Рассмотрим задачу в общей постановке

$$\sup_{\tau\le t\le T} |u(t)| \to \min, \ \int_\tau^T h_i(t)u(t)dt = c_i, \ i = \overline{1,m}.$$

Решаем задачу (12.15)

$$f_0 = \inf_{\{\lambda_i\}} \int_\tau^T \left| \sum_{i=1}^m \lambda_i h_i(t) \right| dt, \ \sum_{i=1}^m \lambda_i c_i = 1. \qquad (12.31)$$

Пусть $\lambda_i^{(0)}, i = \overline{1,m}$ — решение этой задачи и $f_0 > 0$. Обозначим

$$h^{(0)}(t) = \sum_{i=1}^m \lambda_i^{(0)} h_i(t). \qquad (12.32)$$

Тогда из принципа максимума получим, что

$$\int_\tau^T u^{(0)}(t)h^{(0)}(t)dt = \max \int_\tau^T u(t)h^{(0)}(t)dt, \ \sup_{\tau\le t\le T} |u(t)| = \frac{1}{f_0}.$$

Отсюда получим, что

$$u^{(0)}(t) = \frac{1}{f_0} sign h^{(0)}(t). \qquad (12.33)$$

Пример 4. Решим пример 3. Имеем $h_1(t) = T - t, h_2(t) = 1$. Поэтому задача (12.31) принимает вид

$$f_0 = \inf_{\{\lambda_i\}} \int_\tau^T |\lambda_1(T-t) + \lambda_2| dt, \ \lambda_1 H = 1.$$

Отсюда получим, что

$$\int_0^T |\lambda_1(T-t) + \lambda_2| dt = \int_0^T \left| \frac{1}{H}(T-t) + \lambda_2 \right| dt = \frac{1}{H} \int_0^T |t + a| dt$$

Здесь обозначено

$$a = -T - \lambda_2 H. \qquad (12.34)$$

Итак, рассмотрим задачу

$$J(a)=\int_0^T |t+a|\,dt \to \min, a \in R.$$

Имеем $J(0)=\dfrac{T^2}{2}$. Если $a>0$, то $J(a)=\dfrac{T^2}{2}+aT$.
Пусть $a<0$ и $-a \le T$ (рис. 12.1).

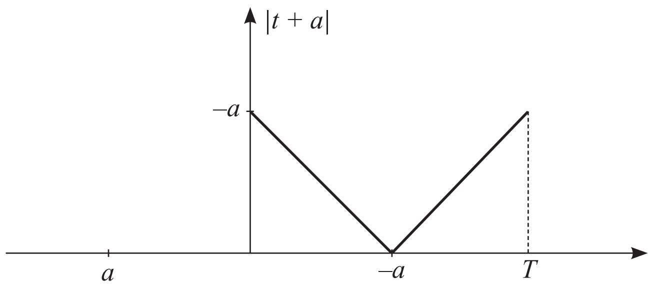


Рис. 12.1

Тогда

$$J(a)=\int_0^T |t+a|\,dt = \frac{a^2}{2}+\frac{(T+a)^2}{2}=\frac{a^2}{2}+\frac{T^2}{2}+Ta+\frac{a^2}{2}=a^2+Ta+\frac{T^2}{2}.$$

Пусть $a<0$ и $-a>T$ (рис. 12.2).

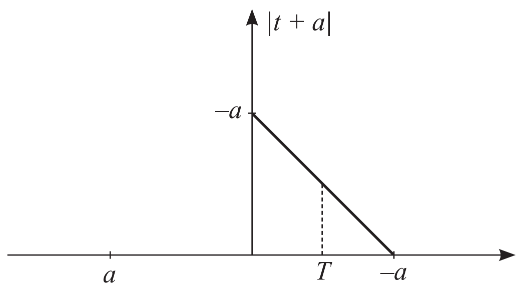


Рис. 12.2

Тогда

$$J(a)=\int_0^T |t+a|\,dt = \int_0^T (-a-t)\,dt = -aT-\frac{T^2}{2}.$$

Таким образом, (рис. 12.3)

$$J(a)=\begin{cases} \dfrac{T^2}{2}+aT, & \text{при } \ a\ge 0 \\ a^2+Ta+\dfrac{T^2}{2}, & \text{при } \ -T\le a<0 \\ -aT-\dfrac{T^2}{2}, & \text{при } \ a<-T \end{cases}$$

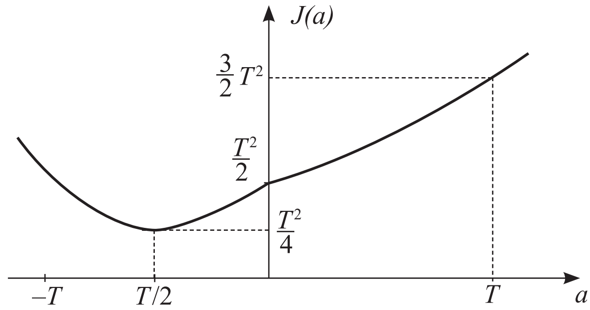


Рис. 12.3

Минимальное значение функции $J(a)$ достигается при $a^{(0)} = -\frac{T}{2}$.

Значит $\lambda_1^{(0)} = \frac{1}{H}$, $\lambda_2^{(0)} = -\frac{a^{(0)} + T}{H} = -\frac{-\frac{T}{2} + T}{H} = -\frac{T}{2H}$.

Далее,

$$f_0 = \int_0^T \left| \lambda_1^{(0)}(T-t) + \lambda_2^{(0)} \right| dt = \frac{1}{H} J\left(a^{(0)}\right) = \frac{T^2}{4H} > 0,$$

$$h^{(0)}(t) = \lambda_1^{(0)}(T-t) + \lambda_2^{(0)} = \frac{1}{H}(T-t) - \frac{T}{2H} = \frac{T}{2H} - \frac{t}{H}.$$

Из формулы (12.33) получим, что

$$u^{(0)}(t) = \frac{4H}{T^2} sign(\frac{T}{2} - t).$$

График управления изображен на рис. 12.4.

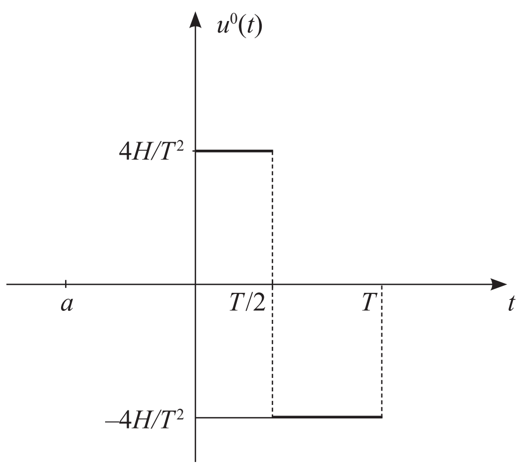


Рис. 12.4

Список рекомендуемой литературы

1. Алексеев, В. М. Сборник задач по оптимизации. Теория. Примеры. Задачи : учеб. пособие / В. М. Алексеев, Э. М. Галеев, В. М. Тихомиров. — 3-е изд., испр. — М. : Физматлит, 2011. — 403 с.

2. Васильев, Ф. П. Методы оптимизации / Ф. П. Васильев. — М. : Факториал Пресс, 2002. — 824 с.

3. Ерёмин, И. И. Двойственность в линейной оптимизации / И. И. Ерёмин. — Екатеринбург : УрО РАН, 2001. — 179 с.

4. Петров, Н. Н. Введение в выпуклый анализ : учеб. пособие / Н. Н. Петров. — Ижевск : Удмурт. гос. ун-т, 2009. — 168 с.

5. Пшеничный, Б. Н. Метод линеаризации / Б. Н. Пшеничный. — М. : Наука, 1983. — 136 с.

6. Пшеничный, Б. Н. Численные методы в экстремальных задачах / Б. Н. Пшеничный, Ю. М. Данилин. — М. : Наука, 1975. — 319 с.

7. Ухоботов, В. И. Правило множителей Лагранжа в задачах вариационного исчисления и оптимального управления : учеб. пособие / В. И. Ухоботов. — Челябинск : Челяб. гос. ун-т, 2006. — 146 с.

Учебное издание

УХОБОТОВ Виктор Иванович
ИЗМЕСТЬЕВ Игорь Вячеславович

ИЗБРАННЫЕ ГЛАВЫ ТЕОРИИ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Учебное пособие

Корректура и вёрстка *М. В. Трифоновой*
Дизайн обложки *О. Харитоновой*

Подписано в печать 16.03.20.
Формат 70×108 $^1/_{16}$. Бумага офсетная.
Усл. печ. л. 7,0. Уч.-изд. л. 10,0.
Тираж 100 экз. Заказ 86.
Цена договорная

Челябинский государственный университет
454001 Челябинск, ул. Братьев Кашириных, 129

Отпечатано в издательстве Челябинского государственного университета
454021 Челябинск, ул. Молодогвардейцев, 57 б