Глава 4

Нелинейное программирование

Задача нелинейного программирования — это задача выбора таких неотрицательных значений некоторых переменных, подчиненных системе ограничений в форме неравенств, при которых достигается максимум или минимум данной функции [1]. В обозначениях, принятых в разделе 2.2, задача нелинейного программирования на максимум состоит в следующем: требуется найти

$$\max_{\mathbf{x}} F(\mathbf{x}) \tag{4.0.1}$$

при условии, что $\mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}$, $\mathbf{x} \geqslant \mathbf{0}$,

или в развернутом виде найти

$$\max_{x_1, x_2, \ldots, x_n} F(x_1, x_2, \ldots, x_n)$$

при условии, что

$$\begin{gathered} g_1(x_1, x_2, \ldots, x_n) \leqslant b_1 \\ g_2(x_1, x_2, \ldots, x_n) \leqslant b_2 \\ \vdots \qquad\qquad \vdots \\ g_m(x_1, x_2, \ldots, x_n) \leqslant b_m \\ x_1 \geqslant 0, x_2 \geqslant 0, \ldots, x_n \geqslant 0. \end{gathered} \tag{4.0.2}$$

[1] Основные сведения по нелинейному программированию изложены в [1, 2, 3, 4, 5]. В большинстве книг нелинейное программирование рассматривается вслед за линейным программированием. В нашей книге принят обратный порядок изложения: линейному программированию посвящена гл. 5, следующая за главой о нелинейном программировании, поскольку нелинейное программирование тесно связано с классическими задачами математического программирования, рассмотренными в гл. 3, тогда как линейное программирование примыкает к теории игр (гл. 6). Задачи линейного программирования, характеризующиеся тем, что в них как целевая функция, так и функции ограничений линейны, рассматриваются как частный вид задач нелинейного программирования.

Величины $x_1, x_2, \ldots, x_n$ — составляющие n-мерного вектора-столбца $\mathbf{x}$ — представляют собой *инструментальные* переменные. Функция $F(\cdot)$ — это *целевая функция*, а *функции ограничений* $g_1(\cdot), g_2(\cdot), \ldots, g_m(\cdot)$ суть составляющие m-мерного вектора-столбца $\mathbf{g}(\cdot)$. Вектор $\mathbf{b}$ содержит *константы ограничений* $b_1, b_2, \ldots, b_m$. Предположим, что m и n — конечные числа; заданы $m+1$ непрерывно дифференцируемых и не содержащих случайных элементов функций $F(\cdot)$, $g_1(\cdot)$, $g_2(\cdot), \ldots, g_m(\cdot)$; вектор $\mathbf{b}$ состоит из заданных вещественных чисел; $\mathbf{x}$ — это любой вектор с вещественными компонентами, удовлетворяющий $m+n$ ограничениям из (4.0.1)[1].

Сделаем несколько замечаний относительно задачи нелинейного программирования. Во-первых, отметим, что в отличие от задачи классического программирования величины m и n не связаны никакими ограничениями. Во-вторых, выбор знаков неравенств $\leqslant$ совершенно условен. Например, умножая неравенство $x_1 - 2x_2 \geqslant 7$ на -1, можно превратить его в $-x_1 + 2x_2 \leqslant -7$, изменив тем самым знак неравенства на противоположный. В-третьих, отметим, что ограничение в форме равенства, например $x_3 + 8x_7 = 12$, можно заменить парой ограничений неравенств $x_3 + 8x_7 \leqslant 12$ и $-x_3 - 8x_7 \leqslant -12$. В-четвертых, условие неотрицательности всех инструментальных переменных $x_1, x_2, \ldots, x_n$ не является обязательным. Если на значения некоторой переменной, например x_9, ограничения не налагаются (то есть она может принимать отрицательные, положительные и нулевые значения), то эту переменную можно заменить двумя переменными: $x_9' \geqslant 0$ и $x_9'' \geqslant 0$, полагая $x_9 = x_9' - x_9''$, так что новая формулировка задачи не будет включать переменную x_9. Таким образом, классическую задачу математического программирования (3.0.1) можно рассматривать как частную задачу нелинейного программирования, в которой условие неотрицательности переменных отсутствует, а ограничения неравенства заменены ограничениями в форме равенств.

[1] Задачи, в которых число переменных или число ограничений бесконечно, называются задачами бесконечномерного нелинейного программирования [6]. Задачи, в которых $\mathbf{b}$, $F(\cdot)$ или $\mathbf{g}(\cdot)$ включают случайные элементы, относятся к классу задач стохастического нелинейного программирования [7, 8].

Геометрически каждое из n условий неотрицательности

$$x_j \geqslant 0, \quad j = 1, 2, \ldots, n \tag{4.0.3}$$

определяет полупространство неотрицательных значений независимых переменных, а пересечение всех таких полупространств представляет собой подмножество n-мерного евклидова пространства, называемое *неотрицательным ортантом*. Например, в E^2 неотрицательный ортант — это неотрицательный квадрант, т. е. первый квадрант плюс соответствующие полуоси. Каждое из m ограничений-неравенств

$$g_i(x_1, x_2, \ldots, x_n) \leqslant b_i, \quad i = 1, 2, \ldots, m \tag{4.0.4}$$

также определяет множество точек в n-мерном евклидовом пространстве, а пересечение этих m множеств с неотрицательным ортантом составляет *допустимое множество*

$$X = \{\mathbf{x} \in E^n \mid \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}, \ \mathbf{x} \geqslant \mathbf{0}\}. \tag{4.0.5}$$

Поверхности (линии) постоянного уровня и направления скорейшего роста, дающие геометрическое представление о целевой функции, рассмотрены в гл. 2. Геометрически задача нелинейного программирования состоит в отыскании точки или множества точек из допустимого множества, где достигается поверхность наибольшего уровня. Так как по предположению целевая функция непрерывна и допустимое множество замкнуто, то, согласно теореме Вейерштрасса (раздел 2.2), решение (глобальный максимум) задачи существует, если допустимое множество не пустое и ограниченное. Решение может принадлежать либо границе, либо внутренней части допустимого множества, как это показано на рис. 2.4.

Важную роль в задачах нелинейного программирования играют условия выпуклости. Задачей выпуклого программирования часто называют задачу, в которой функции ограничений выпуклые, а целевая функция вогнутая. В этом случае, согласно теореме из раздела 2.2, локальный максимум целевой функции, находящийся внутри допустимого множества или на его границе, является глобальным максимумом, а множество точек, на которых достигается глобальный максимум, выпукло. Если дополнительно предполагается, что целевая функция строго вогнута, то задача имеет единственное решение [1].

[1] В более общем случае квазивыпуклого программирования, когда функции ограничений квазивыпуклы, а целевая функция квазивогнута, локальный максимум является глобальным максимумом, и множество всех точек максимума (как локальных, так и глобальных максимумов) выпуклое. Если предположить, кроме того, что целевая функция строго квазивогнутая, то решение задачи является единственным. См. [9].

4.1. ЗАДАЧА НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ ПРИ ОГРАНИЧЕНИЯХ НЕОТРИЦАТЕЛЬНОСТИ

Если на переменные налагаются только условия неотрицательности ($m = 0$), то основная задача (4.0.1) сводится к выбору неотрицательных значений инструментальных переменных, максимизирующих функцию, т. е. к задаче вида: найти

$$\max_{\mathbf{x}} F(\mathbf{x}) \quad \text{при условии, что } \mathbf{x} \geqslant \mathbf{0}. \tag{4.1.1}$$

Одним из подходов к решению этой задачи является разложение в ряд Тейлора, использованное в параграфе 3.1 для решения классической задачи математического программирования при отсутствии ограничений. Предположим, что в $\mathbf{x}^*$ существует локальный максимум задачи (4.1.1). Тогда в точках, близких к $\mathbf{x}^*$

$$F(\mathbf{x}^*) \geqslant F(\mathbf{x}^* + h\,\Delta\mathbf{x}), \tag{4.1.2}$$

где $\Delta\mathbf{x}$ определяет направление в E^n, а h — любое произвольное малое положительное число. Пусть $F(\mathbf{x})$ — дважды непрерывно дифференцируемая функция, тогда функцию в правой части (4.1.2) можно разложить по формуле Тейлора в окрестности $\mathbf{x}^*$

$$F(\mathbf{x}^* + h\Delta\mathbf{x}) = F(\mathbf{x}^*) + h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\,\Delta\mathbf{x} + \\ + \frac{1}{2!}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})\,\Delta\mathbf{x}, \tag{4.1.3}$$

где $0 < \theta < 1$.

Сопоставляя последние два неравенства, приходим к *основному неравенству*

$$h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\,\Delta\mathbf{x} + \frac{1}{2}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x}) \leqslant 0, \tag{4.1.4}$$

являющемуся необходимым условием существования локального максимума в $\mathbf{x}^*$. Если $\mathbf{x}^* > \mathbf{0}$, т. е. $\mathbf{x}^*$ — внутренняя точка, то основное неравенство выполняется для всех направлений $\Delta\mathbf{x}$. Поэтому, как и в классических задачах математического программирования, условием первого порядка является обращение в нуль первых производных. Допустим теперь, что одна из инструментальных переменных принимает граничное значение $x_j^* = 0$. Пусть приращения всех остальных переменных равны 0. Тогда, поскольку при $x_j^* = 0$, единственным допустимым направлением является такое, при котором $\Delta x_j \geqslant 0$, то из основного неравенства следует, что

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*)\,\Delta x_j \leqslant 0. \qquad (4.1.5)$$

Следовательно, условие первого порядка состоит в том, что

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*) \leqslant 0, \text{ если } x_j^* = 0. \qquad (4.1.6)$$

Итак, в точке внутреннего максимума ($x_j^* > 0$) производная по x_j обращается в 0, а если максимум достигается на границе ($x_j^* = 0$), то первая производная меньше или равна 0. Но так как либо производная равна 0 (во внутренней точке), либо соответствующая переменная принимает нулевое значение (в граничной точке), то их произведение в точке максимума всегда равно 0

$$\frac{\partial F}{\partial x_j}(\mathbf{x}^*)\, x_j^* = 0. \qquad (4.1.7)$$

Сумма этих произведений также равна 0

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\,\mathbf{x}^* = \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(\mathbf{x}^*)\, x_j^* = 0. \qquad (4.1.8)$$

Условие обращения в нуль суммы указанных произведений на самом деле требует, чтобы каждое слагаемое равнялось нулю (т. е. соотношение (4.1.7) должно выполняться для всех j). Это следует из того, что все переменные неотрицательны, а частные производные первого порядка не превышают нуля. Таким образом, локальный максимум в $\mathbf{x}^*$ определяется с помощью $2n + 1$ условий первого порядка

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) \leqslant 0$$
$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\,\mathbf{x}^* = 0 \qquad (4.1.9)$$
$$\mathbf{x}^* \geqslant 0.$$

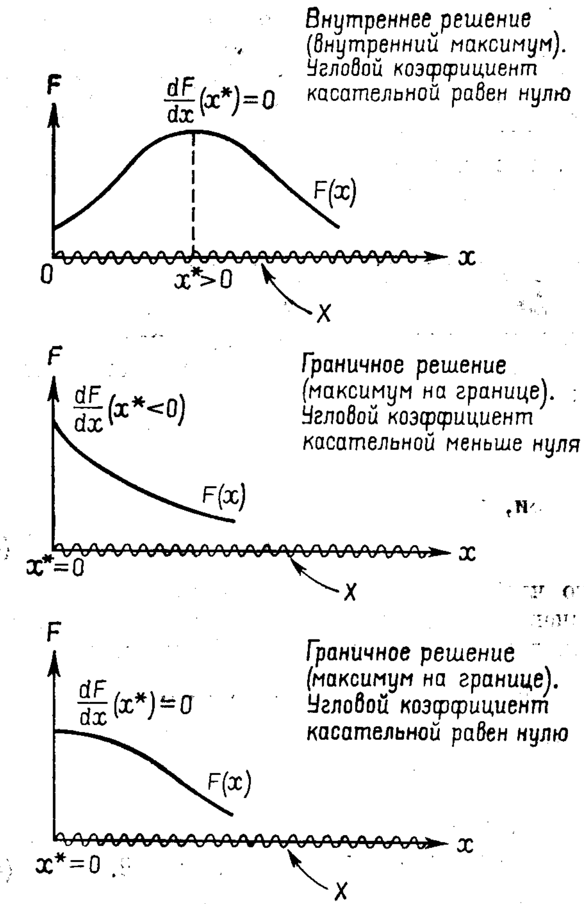

Рис. 4.1. Три возможных решения задачи максимизации функции одной неотрицательной переменной.

Из этих условий следует приведенный выше результат: в точке максимума частная производная либо равна нулю, если соответствующая переменная положительна, либо меньше или равна нулю, если соответствующая переменная обращается в нуль

$$\left.\begin{aligned} &\frac{\partial F}{\partial x_j}(\mathbf{x}^*) = 0, \text{ если } x_j^* > 0 \\ &\frac{\partial F}{\partial x_j}(\mathbf{x}^*) \leqslant 0, \text{ если } x_j^* = 0 \end{aligned}\right\} j = 1, 2, \ldots, n. \qquad (4.1.10)$$

На рис. 4.1 проиллюстрированы все возможные решения задачи в одномерном случае: внутреннее решение-максимум достигается во внутренней точке, где угловой коэффициент касательной равен нулю; граничное решение, при котором угловой коэффициент касательной меньше нуля, и граничное решение, при котором угловой коэффициент касательной равен нулю.

4.2. УСЛОВИЯ КУНА — ТАККЕРА

Используя результаты предыдущего параграфа, можно проанализировать общую задачу нелинейного программирования:

найти

$$\max_{\mathbf{x}} F(\mathbf{x})$$

при условии, что

$$\mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}, \quad \mathbf{x} \geqslant \mathbf{0}. \tag{4.2.1}$$

Для того чтобы преобразовать ограничения-неравенства в ограничения в форме равенств, введем вектор, состоящий из m вспомогательных («свободных») переменных

$$\mathbf{s} \equiv \mathbf{b} - \mathbf{g}(\mathbf{x}) = (s_1, s_2, \ldots, s_m)'. \tag{4.2.2}$$

Теперь задача состоит в отыскании

$$\max_{\mathbf{x},\, \mathbf{s}} F(\mathbf{x})$$

при условии, что

$$\mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{b}, \quad \mathbf{x} \geqslant \mathbf{0}, \quad \mathbf{s} \geqslant \mathbf{0}, \tag{4.2.3}$$

причем неотрицательность вспомогательных переменных обеспечивает выполнение ограничений-неравенств. Если бы выражения (4.2.3) не включали $m + n$ ограничений-неравенств, то наша задача являлась бы классической задачей математического программирования. Функция Лагранжа такой задачи имела бы вид

$$L' = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s}), \tag{4.2.4}$$

где $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ — вектор множителей Лагранжа (как и в предыдущей главе).

Необходимые условия первого порядка состояли бы в том, что все первые частные производные функции L' по $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{s}$ должны обращаться в нуль. Но так как $\mathbf{x}$ и $\mathbf{s}$ неотрицательны, то соответствующие условия на значения первых производных по этим $m + n$ переменным заменяются на условия, которые можно получить по формулам (4.1.9) предыдущего параграфа. В соответствии с этим условия первого порядка для существования локального максимума задачи (4.2.3) состоят в следующем:

$$\begin{aligned}
&\frac{\partial L'}{\partial \mathbf{x}} = \frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} \leqslant \mathbf{0}\\
&\frac{\partial L'}{\partial \mathbf{x}}\mathbf{x} = \left(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)\mathbf{x} = 0\\
&\mathbf{x} \geqslant \mathbf{0}\\
&\frac{\partial L'}{\partial \mathbf{y}} = \mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s} = \mathbf{0}\\
&\frac{\partial L'}{\partial \mathbf{s}} = -\mathbf{y} \leqslant \mathbf{0}\\
&\frac{\partial L'}{\partial \mathbf{s}}\mathbf{s} = -\mathbf{y}\mathbf{s} = 0\\
&\mathbf{s} \geqslant \mathbf{0},
\end{aligned} \tag{4.2.5}$$

где все переменные, функции и производные вычисляются при $\mathbf{x}^*$, $\mathbf{y}^*$ и $\mathbf{s}^*$. Заменяя вектор вспомогательных переменных $\mathbf{s}$ на $\mathbf{b} - \mathbf{g}(\mathbf{x})$, приходим к условиям Куна — Таккера

$$\begin{aligned}
&\left(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right) \leqslant \mathbf{0}\\
&\left(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)\mathbf{x} = 0\\
&\mathbf{x} \geqslant \mathbf{0}\\
&(\mathbf{b} - \mathbf{g}(\mathbf{x})) \geqslant \mathbf{0}\\
&\mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) = 0\\
&\mathbf{y} \geqslant \mathbf{0}.
\end{aligned} \tag{4.2.6}$$

Те же условия можно получить, если определить функцию Лагранжа для исходной задачи (4.2.1) как сумму целевой функции и скалярного произведения векторов множителей Лагранжа и разности между константами ограничений

$$L = L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})), \tag{4.2.7}$$

или в развернутом виде

$$L(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) = F(x_1, x_2, \ldots, x_n) + \\ + \sum_{i=1}^{m} y_i (b_i - g_i(x_1, x_2, \ldots, x_n)). \qquad (4.2.8)$$

Условия Куна — Таккера записываются теперь следующим образом:

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) = \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}\frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*) \leqslant 0$$

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*)\,\mathbf{x}^* = \left(\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}^*\frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*)\right)\mathbf{x}^* = 0$$

$$\mathbf{x}^* \geqslant 0$$

$$\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = \mathbf{b} - \mathbf{g}(\mathbf{x}^*) \geqslant 0 \qquad (4.2.9)$$

$$\mathbf{y}^*\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = 0$$

$$\mathbf{y}^* \geqslant 0.$$

Эти условия являются необходимыми и достаточными для существования (строгого) локального максимума, если целевая функция (строго) вогнутая, а функции ограничений выпуклые, и если, кроме того, выполнено некоторое условие регулярности, налагаемое на ограничения, которое будет введено в следующем параграфе. Условия Куна — Таккера можно записать в развернутом виде как $2m + 2n + 2$ соотношений:

$$\frac{\partial L}{\partial x_j} = \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_n} \leqslant 0, \quad j = 1, 2, \ldots, n \qquad (4.2.10)$$

$$\sum_{j=1}^{n} \frac{\partial L}{\partial x_j} x_j = \sum_{j=1}^{n}\left(\frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j}\right) x_j = 0 \qquad (4.2.11)$$

$$x_j \geqslant 0, \quad j = 1, 2, \ldots, n \qquad (4.2.12)$$

$$\frac{\partial L}{\partial y_i} = b_i - g_i(\cdot) \geqslant 0, \quad i = 1, 2, \ldots, m \qquad (4.2.13)$$

$$\sum_{i=1}^{m} y_i \frac{\partial L}{\partial y_i} = \sum_{i=1}^{m} y_i (b_i - g_i(\cdot)) = 0 \qquad (4.2.14)$$

$$y_i \geqslant 0, \quad i = 1, 2, \ldots, m, \qquad (4.2.15)$$

где все переменные, функции и производные вычисляются в $(\mathbf{x}^*, \mathbf{y}^*)$.

Для того чтобы понять важный смысл этих условий, отметим прежде всего, что соотношения (4.2.12) и (4.2.13) выражают соответственно ограничения неотрицательности и ограничения-неравенства исходной задачи нелинейного программирования. Отметим, что вследствие неравенств (4.2.10) и (4.2.12) каждое из слагаемых суммы (4.2.11) должно равняться нулю, так что

$$\text{либо } \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i\left(\frac{\partial g_i}{\partial x_j}\right) = 0, \text{ либо } x_j = 0 \qquad (4.2.16)$$

(либо эти равенства выполняются одновременно)

$$j = 1, 2, \ldots, n.$$

Иначе говоря, либо каждое из соотношений между производными выполняется как равенство, либо переменная равна нулю, либо обе эти возможности осуществляются одновременно. Таким образом,

$$\left.\begin{aligned} &\frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j} \leqslant 0, \text{ но} \\ &\frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j} = 0, \text{ если } x_j^* > 0, \\ &x_j^* \geqslant 0, \text{ но } x_j^* = 0, \text{ если } \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j} < 0 \end{aligned}\right\} j = 1, 2, \ldots, n. \qquad (4.2.17)$$

Каждое из слагаемых суммы (4.2.14) также равно нулю благодаря неравенствам (4.2.13) и (4.2.15), так что

либо $y_i = 0$, либо $g_i(\mathbf{x}^*) = b_i$ (или же эти неравенства выполняются одновременно), $i = 1, 2, \ldots, m$. (4.2.18)

Иначе говоря, либо множитель Лагранжа равен нулю, либо соответствующее ограничение-неравенство выполняется как равенство, либо и то и другое выполняется одновременно. Итак,

$$\left.\begin{aligned} g_i(\mathbf{x}^*) \leqslant b_i, \text{ но } g_i(\mathbf{x}^*) = b_i, \text{ если } y_i^* > 0 \\ y_i^* \geqslant 0, \text{ но } y_i^* = 0, \text{ если } g_i(\mathbf{x}^*) < b_i \end{aligned}\right\} i = 1, 2, \ldots, m. \tag{4.2.19}$$

Условия (4.2.17) и (4.2.19), представляющие собой иной способ формулирования условий Куна — Таккера, называют *условиями дополняющей нежесткости*. Совершенно ясно, что

$$L(\mathbf{x}^*, \mathbf{y}^*) = F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = F(\mathbf{x}^*), \tag{4.2.20}$$

т. е. значение функции Лагранжа в точке решения совпадает с оптимальным значением целевой функции, так как, согласно (4.2.14), $\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = 0$.

Дадим геометрическую интерпретацию условий Куна-Таккера. Будем рассматривать эти условия в форме (4.2.5). Если ввести второй вектор дополнительных переменных

$$\mathbf{r} = \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} - \frac{\partial F}{\partial \mathbf{x}} = (r_1, r_2, \ldots, r_n), \tag{4.2.21}$$

то (4.2.5) преобразуются к виду

$$\begin{gathered} \frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}} + \mathbf{r} = \mathbf{0} \\ \mathbf{rx} = 0 \\ \mathbf{r} = \mathbf{0}, \quad \mathbf{x} \leqslant \mathbf{0} \\ \mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s} = \mathbf{0} \\ \mathbf{ys} = \mathbf{0} \\ \mathbf{s} \geqslant \mathbf{0}, \quad \mathbf{y} \geqslant \mathbf{0}, \end{gathered} \tag{4.2.22}$$

где все переменные, функции и производные вычисляются при $\mathbf{x} = \mathbf{x}^*$, $\mathbf{y} = \mathbf{y}^*$, $\mathbf{r} = \mathbf{r}^*$, $\mathbf{s} = \mathbf{s}^*$. Неотрицательность вспомогательных переменных обеспечивает выполнение соответствующих ограничений-неравенств. Первые n соотношений можно записать в форме

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{y}^*\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*) + \mathbf{r}^*(-\mathbf{I}), \tag{4.2.23}$$

где $\mathbf{I}$ — единичная матрица. Геометрический смысл (4.2.23) состоит в том, что в искомой точке градиент целевой функции $\frac{\partial F}{\partial \mathbf{x}}$ должен быть линейной комбинацией градиентов ограничивающих гиперповерхностей. Градиентами ограничений неравенств являются строки матрицы Якоби $\frac{\partial \mathbf{g}}{\partial \mathbf{x}}$, градиентами ограничений неотрицательности — строки матрицы $(-\mathbf{I})$, а коэффициентами линейной комбинации — неотрицательные множители Лагранжа $\mathbf{y}^*$ и вспомогательные переменные $\mathbf{r}^*$. Таким образом, если решение лежит на границе, то направление скорейшего роста представляет собой линейную комбинацию векторов-нормалей к поверхности, взятых с неотрицательными коэффициентами.

Рассмотрим теперь характерный пример задачи нелинейного программирования: найти

$$\max_{x_1, x_2} F(x_1, x_2) = -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2$$

при условии, что

$$\begin{gathered} x_1 + x_2 \leqslant 1 \\ 8x_1^2 + x_2^2 \leqslant 2 \\ x_1 \geqslant 0, \ x_2 \geqslant 0. \end{gathered} \tag{4.2.24}$$

Поскольку здесь целевая функция строго вогнутая, а функции ограничений выпуклые, то система неравенств, входящих в условия Куна — Таккера, имеет единственное решение в точке глобального максимума. Функция Лагранжа для этой задачи имеет вид

$$\begin{aligned} L(x_1, x_2, y_1, y_2) = -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + \\ + 80x_2 + y_1(1 - x_1 - x_2) + y_2(2 - 8x_1^2 - x_2^2), \end{aligned} \tag{4.2.25}$$

а условия Куна — Таккера —

$$\begin{gathered} \frac{\partial L}{\partial x_1} = -16x_1 + 12x_2 - 50 - y_1 - 16y_2x_1 \leqslant 0 \\ \frac{\partial L}{\partial x_2} = -20x_2 + 12x_1 + 80 - y_1 - 2y_2x_2 \leqslant 0 \\ \frac{\partial L}{\partial x_1}x_1 + \frac{\partial L}{\partial x_2}x_2 = (-16x_1 + 12x_2 - 50 - y_1 - 16y_2x_1)x_1 + \\ + (-20x_2 + 12x_1 + 80 - y_1 - 2y_2x_2)x_2 = 0 \\ x_1 \geqslant 0 \\ x_2 \geqslant 0 \\ \frac{\partial L}{\partial y_1} = 1 - x_1 - x_2 \geqslant 0 \\ \frac{\partial L}{\partial y_2} = 2 - 8x_1^2 - x_2^2 \geqslant 0 \\ y_1\frac{\partial L}{\partial y_1} + y_2\frac{\partial L}{\partial y_2} = y_1(1 - x_1 - x_2) + y_2(2 - 8x_1^2 - x_2^2) = 0 \\ y_1 \geqslant 0 \\ y_2 \geqslant 0. \end{gathered} \tag{4.2.26}$$

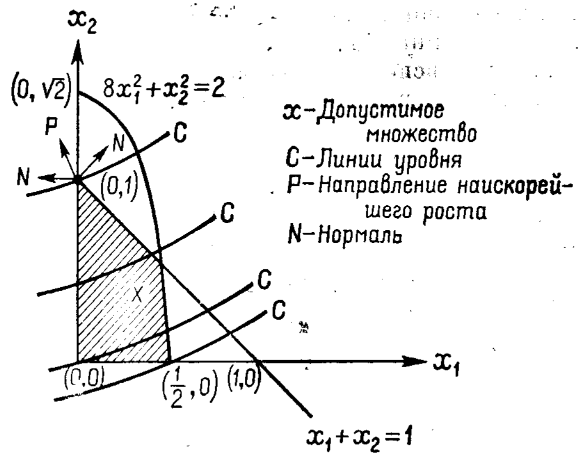

Рис. 4.2. **Геометрическое решение задачи нелинейного программирования.**

Хотя эти условия полностью характеризуют решение, тем не менее они малопригодны для практического отыскания решений задачи. Рассмотрим несколько различных допустимых точек. Точка начала координат (0, 0) не удовлетворяет условиям Куна — Таккера, так как в этой точке $y_1 = 0$, $y_2 = 0$, а $\frac{\partial L}{\partial x_2} = 80$. Точка $\left(\frac{1}{2}, 0\right)$ также не удовлетворяет условиям, так как здесь $y_1 = 0$, а $\frac{\partial L}{\partial x_2} = 86$. Однако точка (0, 1) является решением задачи. В этой точке

$$
\begin{aligned}
\mathbf{x}^* &= (0,\ 1)' \\
\mathbf{y}^* &= (60,\ 0) \\
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) &= (-98,\ 0) \\
\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) &= (0,\ 1)' \\
F(\mathbf{x}^*) &= 70 \\
\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) &= (-38,\ 60).
\end{aligned}
\tag{4.2.27}
$$

На рис. 4.2. решение представлено в геометрической форме. Отметим, что в точке максимума направление скорейшего роста лежит между нормалями (векторы нормалей направлены прочь от допустимого множества). Отметим, кроме того, что второе ограничение в данном случае оказалось излишним.

4.3. ТЕОРЕМА КУНА — ТАККЕРА

Подход Куна — Таккера к общей задаче нелинейного программирования: найти

$$\max_{\mathbf{x}} F(\mathbf{x})$$

$$\text{при условии, что } \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b},\ \mathbf{x} \geqslant 0, \tag{4.3.1}$$

изложенный в предыдущем параграфе, состоит в том, что вводится вектор-строка множителей Лагранжа $\mathbf{y} = (y_1, y_2, \ldots, y_m)$, число которых равно числу ограничений-неравенств и определяется функциями Лагранжа

$$L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})). \tag{4.3.2}$$

В этом случае условия Куна — Таккера записываются как

$$
\begin{gathered}
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) \leqslant \mathbf{0} \qquad \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) \geqslant \mathbf{0} \\
\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*)\,\mathbf{x}^* = 0 \qquad \mathbf{y}^* \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = 0 \\
\mathbf{x}^* \geqslant \mathbf{0} \qquad \mathbf{y}^* \geqslant \mathbf{0}.
\end{gathered}
\tag{4.3.3}
$$

Из сопоставления этих неравенств с условиями существования максимума (4.1.9) видно, что $(\mathbf{x}^*, \mathbf{y}^*)$ является седловой точкой функции Лагранжа, т. е. точкой, максимизирующей функцию по совокупности всех неотрицательных переменных $\mathbf{x}$ и минимизирующей ее по совокупности всех неотрицательных множителей Лагранжа $\mathbf{y}$

$$
\begin{gathered}
L(\mathbf{x}, \mathbf{y}^*) \leqslant L(\mathbf{x}^*, \mathbf{y}^*) \leqslant L(\mathbf{x}^*, \mathbf{y}) \\
\text{при всех } \mathbf{x} \geqslant \mathbf{0},\ \mathbf{y} \geqslant \mathbf{0}.
\end{gathered}
\tag{4.3.4}
$$

Задача отыскания неотрицательных векторов $(\mathbf{x}^*, \mathbf{y}^*)$, удовлетворяющих (4.3.4) называется *задачей о седловой точке* [1].

[1] Тот факт, что $(\mathbf{x}^*, \mathbf{y}^*)$ представляет собой седловую точку функции $L(\mathbf{x}, \mathbf{y})$ и что два набора условий (4.3.3) симметричны, есте-

По теореме Куна — Таккера **x** является решением задачи нелинейного программирования, если ($\mathbf{x}^*$, $\mathbf{y}^*$) является решением задачи о седловой точке, и при некоторых условиях $\mathbf{x}^*$ является решением задачи нелинейного программирования лишь в том случае, когда существует такой вектор **y**, что ($\mathbf{x}^*$, $\mathbf{y}^*$) является решением задачи о седловой точке.

В первой части теоремы утверждается, что если ($\mathbf{x}^*$, $\mathbf{y}^*$) представляет собой седловую точку, как в (4.3.4), то $\mathbf{x}^*$ суть решение задачи нелинейного программирования. Пусть ($\mathbf{x}^*$, $\mathbf{y}^*$) — седловая точка, тогда $\mathbf{x}^*$ максимизирует функцию Лагранжа (относительно всех $\mathbf{x} \geqslant 0$), а $\mathbf{y}^*$ минимизирует ее

$$F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \quad (4.3.5)$$

$$F(\mathbf{x}^*) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \leqslant F(\mathbf{x}^*) + \mathbf{y}(\mathbf{b} - \mathbf{g}(x^*)). \quad (4.3.6)$$

Запишем последнее неравенство в следующем виде:

$$(\mathbf{y} - \mathbf{y}^*)(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) \geqslant 0, \quad \mathbf{y} \geqslant 0. \quad (4.3.7)$$

Поскольку компоненты могут быть сколько угодно большими, то $\mathbf{x}^*$ должно удовлетворять ограничениям-неравенствам

$$\mathbf{g}(\mathbf{x}^*) \leqslant \mathbf{b}. \quad (4.3.8)$$

С другой стороны, если положить в (4.3.7) $\mathbf{y} = \mathbf{0}$, то

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x}^*)) = 0, \quad (4.3.9)$$

так как $\mathbf{y}^* \geqslant 0$, и $\mathbf{b} - \mathbf{g}(\mathbf{x}^*) \geqslant 0$. Используя (4.3.9), неравенство (4.3.5) можно записать как

$$F(\mathbf{x}^*) \geqslant F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})), \quad \mathbf{x} \geqslant 0. \quad (4.3.10)$$

Если **x** — допустимый вектор, то

$$F(\mathbf{x}^*) \geqslant F(\mathbf{x}), \quad (4.3.11)$$

так как $\mathbf{y}^*$ неотрицателен. Следовательно, вектор $\mathbf{x}^*$ максимизирует $F(\cdot)$ в классе допустимых значений **x**, являясь

ственным образом приводит к двойственным задачам: $\max\limits_{\mathbf{x}} L(\mathbf{x}, \mathbf{y})$ при условии, что $\dfrac{\partial L}{\partial \mathbf{y}} \geqslant 0$, $\mathbf{x} \geqslant 0$; $\min\limits_{\mathbf{y}} L(\mathbf{x}, \mathbf{y})$ при условии, что $\dfrac{\partial L}{\partial \mathbf{x}} \leqslant 0$, $\mathbf{y} \geqslant 0$. Этот подход оказался исключительно плодотворным в теории линейного программирования (см. гл. 5), однако в теории нелинейного программирования он не столь удачен. О двойственности в нелинейном программировании см. [10].

тем самым решением задачи нелинейного программирования. Следует заметить, что при доказательстве первой части (достаточности) теоремы Куна — Таккера не потребовалось никаких специальных предложений относительно функций $F(\cdot)$ и $\mathbf{g}(\cdot)$.

Доказательство второй части (необходимости) теоремы Куна — Таккера существенно опирается на некоторые предположения о функциях $F(\cdot)$ и $\mathbf{g}(\cdot)$. Эта часть теоремы справедлива, если $F(\cdot)$ — вогнутая функция, функции $\mathbf{g}(\cdot)$ — выпуклые и если выполнено специальное условие регулярности ограничений, состоящее в том, что в некоторой точке допустимого множества все ограничения-неравенства выполняются как строгие неравенства, т. е. если существует такой вектор $\mathbf{x}^0$, что $\mathbf{x}^0 \geqslant \mathbf{0}$ и $\mathbf{g}(\mathbf{x}^0) < \mathbf{b}$ [1]. Пусть при этих предположениях $\mathbf{x}^*$ является решением задачи нелинейного программирования, т. е.

$$\mathbf{x}^* \geqslant 0, \ \mathbf{g}(\mathbf{x}^*) \leqslant \mathbf{b} \quad \text{и} \quad F(\mathbf{x}^*) \geqslant F(\mathbf{x})$$

при всех **x** таких, что

$$\mathbf{x} \geqslant 0, \quad \mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}. \quad (4.3.12)$$

Определим два множества в $(m+1)$-мерном пространстве

$$A = \left\{ \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \middle| \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \leqslant \begin{pmatrix} F(\mathbf{x}) \\ \mathbf{b} - \mathbf{g}(\mathbf{x}) \end{pmatrix} \right\} \text{ при некоторых } \mathbf{x} \geqslant 0$$

$$B = \left\{ \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} \middle| \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} > \begin{pmatrix} F(\mathbf{x}^*) \\ 0 \end{pmatrix} \right\}, \quad (4.3.13)$$

где a_0 и b_0 — скаляры, а **a** и **b** m-мерные векторы-строки. На рис. 4.3 представлен соответствующий пример для $m = n = 1$. Здесь заштрихованная часть оси x — это допустимое множество; x^* — точка решения. Множество A — это часть плоскости, которая ограничена точками с координатами по оси $0x$, равными $b - g(x)$, а по оси $0y$, равными $F(x)$. Множество B — это внутренняя часть квадранта, вершина которого имеет координату по оси $0x$, равную 0, а по оси $0y$, равную $F(x^*)$. Так как $F(\cdot)$ — вогнутая функция, а $g(\cdot)$ — выпуклая, то множество A

[1] Обсуждение указанного здесь условия регулярности ограничений и других возможных условий регулярности представлено в [11]. Доказательство необходимости, приведенное в настоящей книге, в котором не требуется, чтобы целевая функция или функции ограничений были дифференцируемыми, принадлежит Удзаве [12].

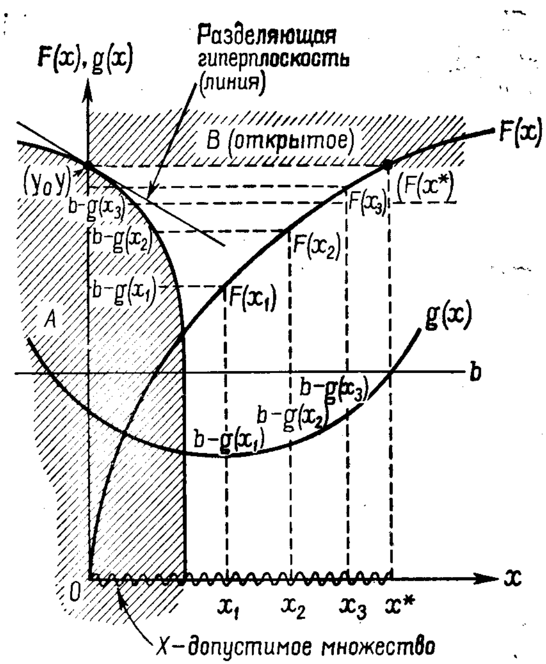

Рис. 4.3. Множества A и B для задачи нелинейного программирования при $m = n = 1$.

здесь, как и в более общем случае, выпукло. Множество B, как внутренняя часть ортанта, также выпукло.

Поскольку $\mathbf{x}^*$ является решением задачи нелинейного программирования, то эти два множества не имеют общих точек. Следовательно, по теореме о существовании разделяющей гиперплоскости для непересекающихся выпуклых множеств существует ненулевой вектор-строка $(y_0, \mathbf{y})$, где y_0 — скаляр, а $\mathbf{y}$ m-мерный вектор, такой, что для всех

$$\begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \text{ из } A \text{ и } \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix} \text{ из } B$$

$$(y_0, \mathbf{y}) \begin{pmatrix} a_0 \\ \mathbf{a} \end{pmatrix} \leqslant (y_0, \mathbf{y}) \begin{pmatrix} b_0 \\ \mathbf{b} \end{pmatrix}. \qquad (4.3.14)$$

Так как вектор $(y_0, \mathbf{y})$ неотрицателен по самому определению множества B и так как точка $(F(\mathbf{x}^*), \mathbf{0})'$ лежит на границе B, то для всех $\mathbf{x} \geqslant \mathbf{0}$

$$y_0 F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant y_0 F(\mathbf{x}^*). \qquad (4.3.15)$$

Покажем, что $y_0 > 0$. Так как, согласно условию регулярности ограничений в некоторой точке допустимого множества, все ограничения-неравенства выполняются как строгие неравенства, то из (4.3.15) следует, что $\mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant 0$ при всех $\mathbf{x} \geqslant \mathbf{0}$. Если $y_0 = 0$, то из (4.3.15) следует, что $\mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant 0$ при любом $\mathbf{x} \geqslant \mathbf{0}$. Но так как $\mathbf{y}$ — неотрицательный вектор, то последнее противоречит предположению о том, что существует $\mathbf{x}^0 \geqslant 0$, такой, что $\mathbf{g}(\mathbf{x}^0) < \mathbf{b}$.

Поскольку $y_0 > 0$, то обе части неравенства (4.3.15) можно разделить на y_0, так что при любом $\mathbf{x} \geqslant 0$

$$\cancel{y_0}F(\mathbf{x}) + \mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant \cancel{y_0}F(\mathbf{x}^*), \qquad (4.3.16)$$

где $\mathbf{y}^* = \left(\frac{1}{y_0}\right)\mathbf{y} \geqslant \mathbf{0}$. В частности, если $\mathbf{x} = \mathbf{x}^*$, то

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) \leqslant 0, \qquad (4.3.17)$$

а так как $\mathbf{g}(\mathbf{x}^*) \leqslant \mathbf{b}$ и $\mathbf{y}^* \geqslant \mathbf{0}$, то

$$\mathbf{y}^*(\mathbf{b} - \mathbf{g}(\mathbf{x})) = 0. \qquad (4.3.18)$$

Следовательно, если функция Лагранжа определяется как

$$L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})), \qquad (4.3.19)$$

то из (4.3.16), (4.3.18) и неотрицательности $\mathbf{y}$ следует, что $(\mathbf{x}^*, \mathbf{y}^*)$ представляет собой седловую точку для $L(\mathbf{x}, \mathbf{y})$ при $\mathbf{x} \geqslant \mathbf{0}$, $\mathbf{y} \geqslant \mathbf{0}$. Таким образом доказана вторая часть теоремы [1]. Итак, при указанных условиях $\mathbf{x}^*$ будет решением задачи нелинейного программирования (4.3.1)

---

[1] Отметим, что теорема остается верной и в том случае, когда условие регулярности ограничений не выполнено, если при этом вводится множитель Лагранжа y_0, соответствующий целевой функции, так что функция Лагранжа имеет вид:

$$\bar{L} = y_0 F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})).$$

Такой подход предложен в работе Джона [13]. Если условие регулярности ограничений выполнено, то в точке решения $y_0 > 0$, так что можно разделить $\bar{L}$ на y_0. В результате получим функцию L, то есть y_0 можно принять равным единице.

Условие регулярности ограничений можно рассмотреть, используя те же геометрические представления, что и при интерпретации выражения (4.2.23). Если указанное условие регулярности ограничений не выполнено, то решением задачи может быть точка, являющаяся вершиной острого выступа, нормали в которой направлены

g(x) ≤ b   b-g(x)

тогда, и только тогда, когда существует вектор $\mathbf{y}^*$, такой, что $(\mathbf{x}^*, \mathbf{y}^*)$ является решением задачи о седловой точке (4.3.4).

Рассмотрим теперь задачу о седловой точке при дополнительном предположении о дифференцируемости функций $F(\mathbf{x})$ и $\mathbf{g}(\mathbf{x})$, которое не использовалось при доказательстве. Первая часть задачи о седловой точке заключается в отыскании максимума по неотрицательным переменным $\mathbf{x}$. Применяя (4.1.9), получаем следующие условия:

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) \leqslant 0$$

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*)\, \mathbf{x}^* = 0, \tag{4.3.20}$$

$$\mathbf{x}^* \geqslant 0.$$

Для решения второй части задачи о седловой точке—минимизации $L(\mathbf{x}^*, \mathbf{y})$ по неотрицательным множителям Лагранжа $\mathbf{y}$ требуется выполнение условий

$$\frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) \geqslant 0$$

$$\mathbf{y}^* \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = 0, \quad \mathbf{y}^* \geqslant 0. \tag{4.3.21}$$

Полученные соотношения представляют собой условия Куна — Таккера (4.3.3).

4.4. ИНТЕРПРЕТАЦИЯ МНОЖИТЕЛЕЙ ЛАГРАНЖА

Как и в предыдущей главе, множители Лагранжа можно истолковывать как характеристики изменений оптимального значения целевой функции при изменениях констант ограничений

в противоположные стороны. Условия первого порядка $\frac{\partial \bar{L}}{\partial \mathbf{x}} = y_0 \frac{\partial F}{\partial \mathbf{x}} - \mathbf{y} \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \leqslant 0$ будут иметь нетривиальное решение относительно множителей Лагранжа. При этом те множители Лагранжа, которые соответствуют ограничениям-неравенствам, определяющим выступ, принимают положительные значения; все другие, включая y_0, равны нулю, а условия первого порядка выполняются как равенства.

$$\mathbf{y}^* = \frac{\partial F^*}{\partial \mathbf{b}}. \tag{4.4.1}$$

Для доказательства следует сначала показать, что $\mathbf{x}^*$ и $\mathbf{y}^*$ можно представить в виде функций от констант-ограничений, а затем продифференцировать функцию Лагранжа по этим константам аналогично тому, как это было сделано в разделе 3.3.

Условия Куна — Таккера можно было бы записать в виде равенств, если бы было известно, какие именно ограничения выполняются как равенства, а какие как неравенства, а также и то, какие именно переменные равны нулю и какие положительны. Допустим, что эти соотношения и переменные перенумерованы так, что в точке оптимума первые m_1 из них выполняются как равенства, а остальные $m - m_1$ как неравенства $(0 \leqslant m_1 \leqslant m)$ и что первые n_1 переменных положительны, а остальные $n - n_1$ равны нулю $(0 \leqslant n_1 \leqslant n)$. Векторы можно расчленить следующим образом:

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} \mathbf{g}^1(\mathbf{x}) \\ \mathbf{g}^2(\mathbf{x}) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}^1 \\ \mathbf{b}^2 \end{pmatrix}, \quad \mathbf{y} = (\mathbf{y}^1 \mathbf{y}^2), \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \end{pmatrix}, \tag{4.4.2}$$

где $\mathbf{g}^1(\mathbf{x})$, $\mathbf{b}^1$ и $\mathbf{y}^1$ состоят из первых элементов векторов $\mathbf{g}(\mathbf{x})$, $\mathbf{b}$ и $\mathbf{y}$ соответственно, а $\mathbf{x}^1$ — из первых n_1 элементов $\mathbf{x}$. Тогда условия Куна — Таккера можно записать как

$$\frac{\partial L}{\partial \mathbf{x}^1} = \frac{\partial F}{\partial \mathbf{x}^1}(\mathbf{x}) - \mathbf{y}^1 \frac{\partial \mathbf{g}^1}{\partial \mathbf{x}^1}(\mathbf{x}) = 0$$

$$\mathbf{x}^2 = 0 \tag{4.4.3}$$

$$\frac{\partial L}{\partial \mathbf{y}^1} = \mathbf{b}^1 - \mathbf{g}^1(\mathbf{x}) = 0$$

$$\mathbf{y}^2 = 0.$$

Очевидно, что (4.4.1) выполняется для последних $m - m_1$ множителей Лагранжа, равных нулю; следовательно,

$$y_i^* = \frac{\partial F^*}{\partial b_i} = 0, \quad i = m_1 + 1,\ m_1 + 2,\ \ldots,\ m. \tag{4.4.4}$$

Последние $m - m_1$ ограничений выполняются как неравенства, так что малые изменения соответствующих констант ограничений не могли бы изменить оптимального значения целевой функции. Что касается первых m_1 множителей Лагранжа, то заметим, что задача приведена к форме классической задачи математического программирования: найти

$$\max_{\mathbf{x}^1} F(\mathbf{x}^1, 0) \quad \text{при условии, что } \mathbf{g}^1(\mathbf{x}^1, 0) = \mathbf{b}^1. \tag{4.4.5}$$

На том же основании, что в разделе 3.3, оказывается возможным представить $\mathbf{x}^1$ и $\mathbf{y}^1$ как функции от $\mathbf{b}^1$ и продифференцировать функцию Лагранжа по $\mathbf{b}^1$. В результате получим, что

$$y_i^* = \frac{\partial F}{\partial b_i} \geqslant 0, \quad i = 1, 2, \ldots, m_1. \tag{4.4.6}$$

В задаче (4.2.24), где $\mathbf{y}^* = (60,0)$, при малом увеличении константы первого ограничения до $1 + \Delta b_1$ оптимальное значение возросло бы до $70 + 60\ \Delta b_1$, тогда как малое увеличение константы второго ограничения никак не повлияло бы на оптимальное значение целевой функции, поскольку это ограничение оказалось нежестким.

4.5. АЛГОРИТМЫ РЕШЕНИЯ

Хотя условия Куна — Таккера дают полную характеристику решения, однако они не содержат конструктивного метода отыскания решения. Например, условия (4.2.26) описывают свойства решения задачи (4.2.24), но не указывают, как найти это решение. Под алгоритмом решения понимают конструктивный метод, с помощью которого можно найти численное решение. В настоящей книге приведены схематические описания лишь нескольких из большого числа существующих алгоритмов решения задач нелинейного программирования [1]. Алгоритмы решения обычно позволяют определить, как изменяются во времени инструментальные переменные, то есть эти переменные отыскиваются в виде функций времени

$$\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))'. \tag{4.5.1}$$

Эти функции часто определяют с помощью дифференциальных уравнений, дающих выражение для скорости изменения переменных $\mathbf{x}(t)$

[1] Более полно алгоритмы решения рассматриваются в [14—17].

$$\dot{\mathbf{x}}(t) = \left(\frac{dx_1(t)}{dt}, \frac{dx_2(t)}{dt}, \ldots, \frac{dx_n(t)}{dt}\right)' = (\dot{x}_1(t), \dot{x}_2(t), \ldots, \dot{x}_n(t))'. \tag{4.5.2}$$

Если значение вектора инструментальных переменных в начальный момент

$$\mathbf{x}(0) = (x_1(0), x_2(0), \ldots, x_n(0))' \tag{4.5.3}$$

является фиксированным, то решение дифференциального уравнения сводится к решению задачи нелинейного программирования

$$\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}^*. \tag{4.5.4}$$

Можно дать классификацию этих алгоритмов, основываясь на свойствах начальной точки $\mathbf{x}(0)$. Если начальная точка не является допустимой, то дифференциальные уравнения смещают ее в пределы допустимого множества, а затем приводят к решению задачи. Этот подход к решению назовем *выбором начальной точки без учета ограничений*. При другом подходе — *выборе начальной точки с учетом ограничений* — начальная точка должна принадлежать допустимому множеству, а дифференциальные уравнения перемещают точку по линиям уровня, все увеличивая значение функции и постепенно приводя к решению.

Многие алгоритмы решения задач нелинейного программирования представляют собой *градиентные методы*. Эти методы основаны на том, что градиент, т. е. вектор

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) = \left(\frac{\partial F}{\partial x_1}(\mathbf{x}), \frac{\partial F}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial F}{\partial x_n}(\mathbf{x})\right), \tag{4.5.5}$$

состоящий из частных производных первого порядка целевой функции, направлен в каждой точке в сторону наибольшего роста целевой функции. Поэтому перемещение в направлении градиента приводит к наибольшему увеличению целевой функции. Решая градиентным методом задачу оптимизации при отсутствии ограничений, мы меняем каждую переменную в любой точке в соответствии с величиной ее частной производной в этой точке

$$\dot{x}_j(t) = \frac{\partial F}{\partial x_j}(\mathbf{x}(t)), \quad j = 1, 2, \ldots, n. \tag{4.5.6}$$

Если предположить, что $F(\cdot)$ выпуклая функция, то с помощью этого метода можно достичь точки максимума.

Для задачи оптимизации при наличии ограничений необходимо модифицировать градиентный метод. Примером такого модифицированного градиентного метода является *градиентный проективный метод*.

При этом методе решения выбор начальной точки осуществляется с учетом ограничений, так что эта точка должна быть допустимой. Перемещение точки происходит в направлении градиента, за исключением тех случаев, когда движение в этом направлении может вывести за пределы допустимого множества. В последнем случае точка смещается в направлении проекции градиента на плоскость, касательную к границе. Перемещение по указанному принципу не выводит точку из допустимого множества и постоянно увеличивает значение целевой функции. Если целевая функция вогнутая, а допустимое множество выпуклое, то указанное перемещение точки в конце концов приводит к решению.

Другим примером градиентного метода в нелинейном программировании является дифференциальный градиентный метод, основанный на условиях Куна — Таккера для функции Лагранжа. Выбор начальной точки в этом методе осуществляется без учета ограничений. Дифференциальные уравнения для переменных $\mathbf{x}(t)$ и для множителей Лагранжа $\mathbf{y}(t)$, рассматриваемых как функции времени, имеют следующий вид

$$\dot{x}_j(t) = \begin{cases} 0, \text{ если } x_j = 0 \text{ и } \dfrac{\partial L}{\partial x_j}(\mathbf{x}(t), \mathbf{y}(t)) < 0 \\ \dfrac{\partial L}{\partial x_j}(\mathbf{x}(t), \mathbf{y}(t)) = \dfrac{\partial F}{\partial x_j}(\mathbf{x}(t)) - \sum\limits_{i=1}^{m} y_i \dfrac{\partial g_i}{\partial x_j}(\mathbf{x}(t)) \\ \qquad\qquad \text{(в противном случае)} \end{cases}$$

$$j = 1, 2, \ldots, n. \tag{4.5.7}$$

$$\dot{y}_i(t) = \begin{cases} 0, \text{ если } y_i = 0 \text{ и } \dfrac{\partial L}{\partial y_i}(\mathbf{x}(t), \mathbf{y}(t)) < 0 \\ -\dfrac{\partial L}{\partial y_i}(\mathbf{x}(t), \mathbf{y}(t)) = -[b_i - g_i(\mathbf{x}(t))] \\ \qquad\qquad \text{(в противном случае)} \end{cases}$$

$$i = 1, 2, \ldots, m.$$

Данный способ решения является градиентным методом, так как скорость изменения каждой из переменных равна

соответствующей частной производной первого порядка целевой функции, измененной на некоторую величину для учета наложенных ограничений. Условия неотрицательности оказываются учтенными за счет того, что процесс перемещения, начинающийся в точке с неотрицательными значениями переменных, не может вывести за пределы неотрицательного ортанта. Условия, накладываемые ограничениями-неравенствами, учтены в результате вычитания из частных производных первого порядка для целевой функции линейных комбинаций элементов соответствующего столбца матрицы Якоби, построенной по функциям ограничений. Взвешивающими коэффициентами в этой линейной комбинации являются множители Лагранжа.

Множители Лагранжа также не могут стать отрицательными числами, поскольку их начальные значения неотрицательны и каждый отдельный множитель возрастает, если соответствующее ограничение не выполнено. Результатом этих свойств множителей Лагранжа является то, что они не позволяют вектору инструментальных переменных выйти из допустимого множества. Данный процесс сводится к решению, начиная с любой начальной точки $(\mathbf{x}(0), \mathbf{y}(0))$ с неотрицательными координатами, если целевая функция строго вогнута, а функции ограничений строго выпуклы.

ЗАДАЧИ

4-А. Решите следующие задачи нелинейного программирования. Дайте геометрическую иллюстрацию решения.

1. Найти

$$\max_{x_1, x_2} 6x_1 - 2x_1^2 + 2x_1x_2 - 2x_2^2$$

при условии, что

$$3x_1 + 4x_2 \leqslant 6$$
$$-x_1 + 4x_2^2 \leqslant 2$$
$$x_1 \geqslant 0, \; x_2 \geqslant 0.$$

2. Найти

$$\max_{x_1, x_2} 3x_1x_2 - x_2^3$$

при условии, что

$$2x_1 + 5x_2 \geqslant 20$$
$$x_1 - 2x_2 = 5$$
$$x_1 \geqslant 0,\ x_2 \geqslant 0.$$

3. Найти

$$\max_{x_1, x_2} x_1 + 2x_2$$

при условии, что

$$3x_1^2 + x_2^2 \leqslant 1$$
$$x_1 - 8x_2 \leqslant -1$$
$$x_1 \geqslant 0,\ x_2 \geqslant 0.$$

4-Б. Задача нелинейного программирования: найти

$$\max_{x_1, x_2} px_1^2 + qx_1x_2$$

при условии, что

$$x_1^2 + rx_2^2 \leqslant 1$$
$$x_1 \geqslant 0,\ x_2 \geqslant 0$$

зависит от трех параметров: p, q, r.

1. Найти решение геометрически, если $p = 0$, $q = r = 1$.
2. Получить условия Куна — Таккера.
3. При каких значениях параметров решение существует?

4-В. Одной из важных задач нелинейного программирования является задача о распределении ограниченного ресурса, суммарная величина которого равна b, между n различными целями:

найти

$$\max_{x_1, x_2, \ldots, x_n} F_1(x_1) + F_2(x_2) + \ldots + F_n(x_n)$$

при условии, что

$$x_1 + x_2 + \ldots + x_n \leqslant b$$
$$x_1 \geqslant 0,\ x_2 \geqslant 0,\ \ldots,\ x_n \geqslant 0.$$

Получить условия дополняющей нежесткости и дать их интерпретацию.

4-Г. Задача квадратического программирования:

найти

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{cx} + \frac{1}{2}\mathbf{x}'\mathbf{Dx}$$

при условии, что

$$\mathbf{Ax} \leqslant \mathbf{b},\quad \mathbf{x} \geqslant 0.$$

Здесь целевая функция является квадратической, а ограничения — линейными. $\mathbf{D}$ — отрицательно определенная матрица [18]. Определить условия Куна — Таккера.

4-Д. Покажите, что если ограничения имеют форму равенств $\mathbf{g}(\mathbf{x}) = \mathbf{b}$, то условия Куна — Таккера сводятся к условиям оптимизации для классической задачи математического программирования.

4-Е. Как изменятся условия Куна — Таккера в следующих случаях:

1. Отсутствуют ограничения неотрицательности.
2. Не накладывается никаких ограничений, кроме ограничений неотрицательности.
3. Переменные неотрицательны и ограничены сверху.

4-Ж. Рассмотрите следующую задачу нелинейного программирования: найти $\max\limits_{x} x$ при условии, что $x^2 \leqslant 0$.

1. Решите эту задачу геометрически.
2. Покажите, что функция Лагранжа не имеет седловых точек. Какое из условий Куна — Таккера не выполняется?

4-З. В задачах квазивыпуклого программирования функция $F(\mathbf{x})$ квазивогнута, а все функции $g_i(\mathbf{x})$ квазивыпуклы [9].

1. Изобразите на схеме, подобной схеме на рис. 4.3, задачу квазивыпуклого программирования, которая не была бы задачей вогнутого программирования.
2. Докажите, что в квазивыпуклом программировании локальный максимум является глобальным максимумом.
3. Докажите, что в квазивыпуклом программировании локальный максимум всегда единственный, если $F(\mathbf{x})$ строго квазивогнута.

4-И. Общую задачу нелинейного программирования (4.0.1) можно представить в виде классической задачи математического программирования, если ввести вспомогательные переменные и вместо условий неотрицательности положить все переменные равными квадратам некоторых новых переменных.

Преобразовав задачу указанным способом, решим ее методами гл. 3. Сопоставьте результаты с условиями Куна — Таккера.

4-К. Предположим, что в общей задаче нелинейного программирования (4.0.1) $\mathbf{x}$ — допустимый вектор, так что $\mathbf{g}(\mathbf{x}) \leqslant \mathbf{b}$, $\mathbf{x} \geqslant 0$. Ясно, что если в окрестности точки существуют другие допустимые точки с большим значением целевой функции, то $\mathbf{x}$ не может быть локальным максимумом. Если $d\mathbf{x} > \mathbf{0}$ и $d\mathbf{g} = (\partial \mathbf{g}/\partial \mathbf{x})(\mathbf{x})\, d\mathbf{x} < 0$, то $\mathbf{x} + d\mathbf{x}$ — это допустимая точка, принадлежащая окрестности $\mathbf{x}$. Следовательно, если существует направление $d\mathbf{x}$, удовлетворяющее этим условиям, такое, что $dF = (\partial F/\partial \mathbf{x})(\mathbf{x})\, d\mathbf{x} > 0$, то $\mathbf{x}$ не может быть локальным максимумом. Таким образом, для существования локального максимума в $\mathbf{x}$ необходимо, чтобы система неравенств

$$\begin{pmatrix} \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) \\ -\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}) \end{pmatrix} d\mathbf{x} > \mathbf{0}, \text{ где } d\mathbf{x} > \mathbf{0},$$

не имела решения. Покажите, что если выполнены условия регулярности ограничений, то из этих неравенств можно вывести условия Куна — Таккера. При доказательстве используются свойства однородных систем линейных неравенств, указанные в разделе Б.6 Приложения Б.

4-Л. Рассмотрим общую задачу нелинейного программирования (4.0.1) с дополнительным ограничением, состоящим в том, что целевая функция не превосходит некоторого значения: $F(\mathbf{x}) \leqslant a$. Определить условия Куна — Таккера в этой задаче; дать интерпретацию множителя Лагранжа, соответствующего дополнительному ограничению.

4-М. Пусть в задаче метода наименьших квадратов (см. задачу 3-О из гл. 3) некоторые коэффициенты должны быть неотрицательными. Сформулируйте задачу нелинейного программирования и найдите условия Куна — Таккера для этого случая.

4.Н. На квадратном участке земли строится N домов. Расположить дома так, чтобы минимальное расстояние между центрами любых двух из них было наибольшим.

4-О. Показать графически, что градиентный проективный метод не позволяет найти глобальный максимум, если целевая функция не вогнута или допустимое множество не выпукло.

4-П. Показать, что если при решении по дифференциальному градиентному методу достигнута точка, удовлетворяющая условиям Куна — Таккера, то из дифференциальных уравнений следует, что переменные больше не изменяются ($\dot{x}_j = 0$ при любом j; $\dot{y}_i = 0$ при любом i).

4-Р. В *задаче векторной максимизации* рассматривается несколько целевых функций $F_1(\mathbf{x})$, $F_2(\mathbf{x})$, ..., $F_q(\mathbf{x})$. Вектор называется эффективным, если он является допустимым и если не существует других допустимых векторов $\mathbf{x}^{**}$, таких, что

$$\left.\begin{aligned} F_k(\mathbf{x}^{**}) &\geqslant F_k(\mathbf{x}^{*}) \text{ при всех } k \\ F_k(\mathbf{x}^{**}) &> F_k(\mathbf{x}^{*}) \text{ при некотором } k \end{aligned}\right\} k = 1, 2, \ldots, q.$$

Иными словами, никакой другой вектор не может увеличивать значений некоторых целевых функций, не уменьшая при этом значений хотя бы одной из оставшихся целевых функций [1].

1. Пусть в задаче имеются две целевые функции и пусть $\mathbf{x}^*$ является решением следующей задачи скалярной оптимизации;

найти

$$\max_{\mathbf{x}} a_1 F_1(\mathbf{x}) + a_2 F_2(\mathbf{x})$$

при условии, что $\mathbf{x} \in \mathbf{X}$,

где a_1 и a_2 — положительные параметры (можно считать, что $a_1 + a_2 = 1$). Доказать, что в этом случае вектор $\mathbf{x}^*$ эффективный. Дайте геометрическую иллюстрацию на плоскости $F_1(\mathbf{x})$, $F_2(\mathbf{x})$.

2. Покажите, что множество всех решений указанной задачи скалярной максимизации при некоторых

[1] [1, 19, 20] Понятие эффективности является центральным понятием в современной теории экономики благосостояния, в которой оно обычно называется «оптимальность по Парето». См. гл. 10 данной книги.

положительных значениях параметров не включает в общем случае всех эффективных точек. Верно ли это, если $F_1(\mathbf{x})$ и $F_2(\mathbf{x})$ вогнутые, а X выпуклое? Рассмотрите случай, когда a_1 (или a_2) может обращаться в нуль.

4-С. В качестве примера задачи векторной максимизации рассмотрим *задачу о выборе портфеля ценных бумаг* [21]. Задача состоит в определении набора $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$, где x_j — доля активов инвестора, вложенных в ценные бумаги, имеющие номер j;

$$j = 1, 2, \ldots, n; \ \mathbf{x} \geqslant 0; \ \sum_{j=1}^{n} x_j = 1.$$

Цели вкладчика капитала можно определить с помощью терминов «доход» и «риск». Доход измеряется как линейная форма от среднего дохода по каждому виду бумаг

$$M(\mathbf{x}) = \mu(\mathbf{x}) = \sum_{j=1}^{n} \mu_j x_j,$$

где μ — вектор средних доходов от каждого вида ценных бумаг. Риск измеряется как квадратичная форма

$$V(\mathbf{x}) = \mathbf{x}' \sum \mathbf{x} = \sum_{j=1}^{n} \sum_{k=1}^{n} \sigma_{jk} x_j x_k,$$

где $\sum$ — заданная квадратная матрица дисперсий и ковариаций доходов размерности $n \times n$. Эта матрица считается положительно определенной. Данный набор ценных бумаг считается *эффективным*, если не существует наборов с большим доходом и с меньшим риском, с большим доходом при том же риске или с меньшим риском при том же уровне дохода.

1. Покажите, что в задаче максимизации дохода: найти

$$\max_{\mathbf{x}} M(\mathbf{x})$$

при условии, что $V(\mathbf{x}) \leqslant \overline{V}$, $\mathbf{x} \geqslant 0$, $\mathbf{1x} = 1$,

где $\overline{V}$ — максимальный риск, можно указать эффективный набор ценных бумаг и что множество решений этой задачи при всех $\overline{V} \geqslant 0$ содержит все эффективные наборы. Каковы в этой задаче условия Куна — Таккера?

2. Покажите, что задача найти

$$\min_{\mathbf{x}} V(\mathbf{x})$$

при условии, что $M(\mathbf{x}) \geqslant \overline{M}$, $\mathbf{x} \geqslant 0$, $\mathbf{1x} = 1$,

где $\overline{M}$ — минимальный доход, имеет решение, являющееся эффективным набором бумаг. Каковы здесь условия Куна — Таккера?

3. Имеются два вида ценных бумаг: один со средним доходом в 20% и дисперсией, равной 5, другой — со средним доходом в 50% и дисперсией, равной 15, ковариация их составляет 5. На $(M(\mathbf{x}), V(\mathbf{x}))$ плоскости найти геометрическое место точек, соответствующих эффективным наборам. Покажите, как влияет на решение изменение ковариации. Для этого найдите геометрические места точек решения, полагая ковариацию равной -8, -2, 0, 3 и оставляя неизменными все остальные параметры.