Глава 3

Классическая задача математического программирования

Классическая задача математического программирования — это задача выбора таких значений некоторых переменных, подчиненных системе ограничений в форме равенств, при которых достигается максимум или минимум данной функции [1]. В обозначениях раздела 2.2 классическая задача на максимум состоит в следующем: требуется найти

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

$$\text{при условии, что } \mathbf{g}(\mathbf{x}) = \mathbf{b}, \tag{3.0.1}$$

или в развернутом виде найти

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

при условии, что

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= b_1 \\ g_2(x_1, x_2, \ldots, x_n) &= b_2 \\ &\cdots \\ g_m(x_1, x_2, \ldots, x_n) &= b_m. \end{aligned} \tag{3.0.2}$$

Переменные $x_1, x_2, \ldots, x_n$ — составляющие n-мерного вектора-столбца $\mathbf{x}$ — представляют собой средства, *инструментальные величины*. 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-столбец $\mathbf{b}$ содержит *константы ограничений* $b_1, b_2, \ldots, b_m$.

[1] Основные сведения о задачах классического математического программирования изложены в [1, 2, 3]. См. также библиографию к гл. 12 этой книги. Указанные там источники представляют интерес, поскольку большинство современных математиков рассматривают классическое математическое программирование лишь как введение к вариационному исчислению.

Предположим, что число переменных n и число ограничений m суть конечные числа и что $n > m$. Разность $n - m$ называется числом *степеней свободы* задачи. Предположим также, что 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 фиксированных действительных чисел. Вектор $\mathbf{x}$ — любой вектор с вещественными компонентами, удовлетворяющий m ограничениям из (3.0.1).

Геометрически каждое из равенств, входящих в систему ограничений

$$g_i(x_1, x_2, \ldots, x_n) = b_i, \; i = 1, 2, \ldots, m, \tag{3.0.3}$$

определяет множество точек в n-мерном евклидовом пространстве m, а пересечение всех m множеств представляет собой *допустимое множество*

$$X = \{\mathbf{x} \in E^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{b}\}. \tag{3.0.4}$$

Так же как в гл. 2, определим поверхности (линии) уровня целевой функции и направление скорейшего роста (уравнения (2.3.1) и (2.3.2)). Геометрическое истолкование задачи состоит в том, что в допустимом множестве отыскивается точка (или множество точек), где достигается поверхность наибольшего уровня целевой функции (т. е. поверхность, наиболее удаленная в направлении скорейшего роста). Так как целевая функция непрерывна, а допустимое множество замкнуто, то, согласно теореме Вейерштрасса (раздел 2.2), решение существует в том случае, если допустимое множество будет непустым и ограниченным.

3.1. ЗАДАЧИ ОПТИМИЗАЦИИ ПРИ ОТСУТСТВИИ ОГРАНИЧЕНИЙ

Особый класс задач составляют задачи оптимизации при отсутствии ограничений ($m = 0$). В случае когда $m = 0$, а функция зависит от скалярного аргумента ($n = 1$), задача состоит в отыскании вещественного числа x, максимизирующего $F(x)$. Если в такой задаче внутренний локальный максимум достигается в x^*, то это значит, что для всех точек $x^* + \Delta x$, близких к x^*,

$$F(x^*) \geqslant F(x^* + \Delta x), \qquad (3.1.1)$$

где Δx — произвольное малое приращение аргумента.

Пусть функция $F(x)$ имеет непрерывные и конечные производные первого и второго порядков. Тогда в окрестности точки x^* можно разложить функцию, стоящую в правой части неравенства (3.1.1), по формуле Тейлора (с остаточным членом)

$$F(x^* + \Delta x) = F(x^*) + \frac{dF}{dx}(x^*)\,\Delta x + \frac{1}{2!}\,\frac{d^2F}{dx^2}(x^* + \theta\Delta x)\,(\Delta x)^2, \qquad (3.1.2)$$

где

$$0 < \theta < 1.$$

Подставив это разложение функции в (3.1.1), придем к *основному неравенству*

$$\frac{dF}{dx}(x^*)\,\Delta x + \frac{1}{2}\,\frac{d^2F}{dx^2}(x^* + \theta\Delta x)\,(\Delta x)^2 \leqslant 0, \qquad (3.1.3)$$

которое должно выполняться для любого произвольно малого приращения Δx. Если Δx больше нуля, то, разделив обе части основного неравенства на Δx и переходя к пределу при Δx, стремящемся к 0, получим, что

$$\frac{dF}{dx}(x^*) \leqslant 0. \qquad (3.1.4)$$

Но аналогичное рассуждение при отрицательных Δx приводит к неравенству

$$\frac{dF}{dx}(x^*) \geqslant 0. \qquad (3.1.5)$$

Таким образом, из основного неравенства получаем в качестве необходимого условия первого порядка, что первая производная в точке локального максимума обращается в нуль

$$\frac{dF}{dx}(x^*) = 0. \qquad (3.1.6)$$

Если условия первого порядка выполнены, то из основного неравенства следует, что

$$\frac{d^2F}{dx^2}(x^* + \theta\Delta x) \leqslant 0, \qquad (3.1.7)$$

так как $(\Delta x)^2$ всегда больше 0. Поскольку (3.1.7) выполняется для всех Δx, а вторая производная по предположению непрерывна, приходим к необходимому условию второго порядка

$$\frac{d^2F}{dx^2}(x^*) \leqslant 0, \qquad (3.1.8)$$

т. е. в точке локального максимума производная второго порядка отрицательна или равна нулю.

Итак, условия (3.1.6) и (3.1.8) представляют собой соответственно условия первого и второго порядков, необходимые для существования локального максимума в x^*.

Достаточные условия наличия в x^* строгого локального максимума состоят в том, что первая производная в этой точке обращается в нуль, а вторая производная строго меньше нуля. Иначе говоря, если выполнены условия

$$\begin{aligned} &\frac{dF}{dx}(x^*) = 0 \\ &\frac{d^2F}{dx^2}(x^*) < 0, \end{aligned} \qquad (3.1.9)$$

то x^* есть точка строго локального максимума

$$F(x^*) > F(x^* + \Delta x). \qquad (3.1.10)$$

Достаточность этих условий можно доказать с помощью основного неравенства либо более простым способом, используя теорему о среднем значении. По теореме о среднем значении

$$F(x^* + \Delta x) = F(x^*) + \frac{dF}{dx}(x^* + \theta\Delta x)\,\Delta x, \qquad (3.1.11)$$

где $0 < \theta < 1$. Функция $F(x)$ непрерывно дифференцируема, а ее первая производная равна нулю и строго убывает при x^*. Следовательно, если $\Delta x > 0$, то

$$\frac{dF}{dx}(x^* + \theta\Delta x) < 0, \qquad (3.1.12)$$

тогда как если $\Delta x < 0$, то

$$\frac{dF}{dx}(x^* + \theta\Delta x) > 0. \qquad (3.1.13)$$

В любом случае

$$\frac{dF}{dx}(x^* + \theta\Delta x)\,\Delta x < 0, \qquad (3.1.14)$$

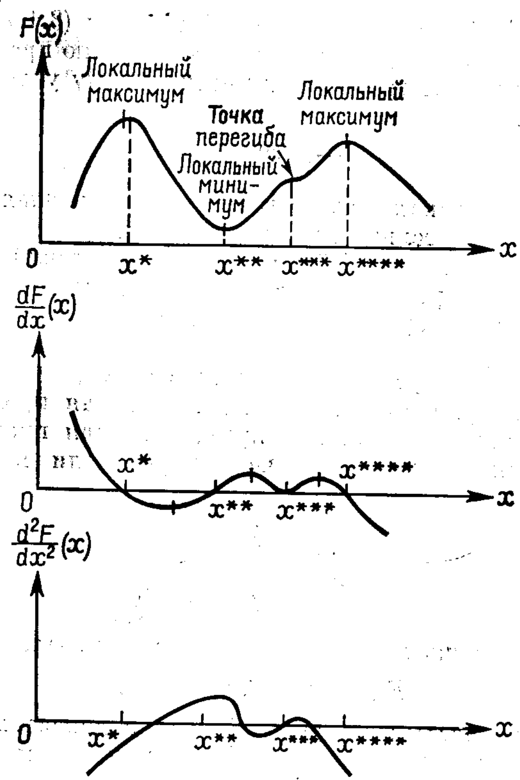

Рис. 3.1. Максимизация при отсутствии ограничений для случая одной независимой переменной.

так что из (3.1.11) следует

$$F(x^*) > F(x^* + \Delta x). \qquad (3.1.15)$$

На рис. 3.1 ход решения представлен в геометрической форме. Угловой коэффициент касательной к кривой $F(x)$ в точке x^* равен нулю, причем величина углового коэффициента убывает, так что точка x^* удовлетворяет условиям (3.1.9) и, следовательно, является точкой строгого локального максимума. Эти условия выполнены и в x^{****}, которая также является точкой строгого локального максимума. В точках x^{**} и x^{***} выполняется условие первого порядка — угловой коэффициент касательной равен нулю, но условие второго порядка не

выполнено, так как угловой коэффициент касательной в x^{**} возрастает, а в x^{***} остается постоянным. Точка x^{**} есть точка строгого локального минимума, а точка x^{***}, в которой первая и вторая производные обращаются в нуль, представляет собой один из случаев точки перегиба. Из примера с точкой x^{***} ясно, что хотя условие первого порядка (3.1.6) и условие второго порядка (3.1.8) и являются необходимыми, однако выполнения одних только этих условий еще недостаточно для существования максимума. Другим примером этого может служить $F(x) = x^3$ при $x = 0$.

Задачу оптимизации при отсутствии ограничений для функции векторного аргумента $m = 0$, $n > 1$ можно исследовать способом, подобным изложенному выше. Формулировка задачи:

найти

$$\max_{\mathbf{x}} F(\mathbf{x}) = F(x_1, x_2, \ldots, x_n). \qquad (3.1.16)$$

Допустим, что локальный максимум существует в $\mathbf{x}^*$:

$$F(\mathbf{x}^*) \geqslant F(\mathbf{x}^* + h\,\Delta\mathbf{x}). \qquad (3.1.17)$$

Это означает, что

$$F(x_1^*, x_2^*, \ldots, x_n^*) \geqslant$$
$$\geqslant F(x_1^* + h\,\Delta x_1, x_2^* + h\,\Delta x_2, \ldots, x_n^* + h\,\Delta x_n), \qquad (3.1.18)$$

где h — произвольное малое положительное число, Δx_j — произвольное изменение x_j $(j = 1, 2, \ldots, n)$, а $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)'$ определяет некоторое направление в E^n. Функцию в правой части (3.1.17) можно рассматривать как функцию h. Разложение по формуле Тейлора в окрестности точки $h = 0$ дает

$$F(\mathbf{x}^* + h\Delta\mathbf{x}) = F(\mathbf{x}^*) + h\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*)\,\Delta\mathbf{x} +$$
$$+ \frac{1}{2!}h^2(\Delta\mathbf{x})'\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x}), \qquad (3.1.19)$$

где $0 < \theta < 1$. Вектор $\frac{\partial F}{\partial \mathbf{x}}$ в (3.1.19) есть градиент функции, а матрица $\frac{\partial^2 F}{\partial \mathbf{x}^2}$ — матрица Гессе:

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

$$\frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_2^2}(\mathbf{x}) & \dots & \frac{\partial^2 F}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 F}{\partial x_n \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 F}{\partial x_n^2}(\mathbf{x}) \end{bmatrix}. \quad (3.1.20)$$

Выпишем (3.1.19) в развернутом виде

$$F(x_1^* + h\Delta x_1, x_2^* + h\Delta x_2, \dots, x_n^* + h\Delta x_n) =$$

$$= F(x_1^*, x_2^*, \dots, x_n^*) + h \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(x_1^*, x_2^*, \dots, x_n^*) \Delta x_j +$$

$$+ \frac{1}{2} h^2 \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_k}(x_1^* + \theta h \Delta x_1, x_2^* + \theta h \Delta x_2, \dots, x_n^* +$$

$$+ \theta h \Delta x_n) \Delta x_j \Delta x_k. \quad (3.1.21)$$

Сопоставляя выражения (3.1.17) и (3.1.19), придем к *основному неравенству*

$$h \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) \Delta \mathbf{x} + \frac{1}{2} h^2 (\Delta \mathbf{x})' \frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^* + \theta h \Delta \mathbf{x}) (\Delta \mathbf{x}) \leqslant 0, \quad (3.1.22)$$

которое должно выполняться для всех направлений $\Delta \mathbf{x}$ и для всех малых положительных чисел h. Чтобы вывести из основного неравенства необходимое условие первого порядка, разделим обе части (3.1.22) на h и перейдем к пределу при h, стремящемся к 0. В результате получаем, что в точке локального максимума градиент функции равен нулевому вектору

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = 0. \quad (3.1.23)$$

Точка, в которой все частные производные первого порядка равны нулю, называется *стационарной*. Следовательно, по доказанному точка локального максимума всегда является стационарной. В качестве еще одного следствия из основного неравенства получаем необходимое условие второго порядка, а именно, что в точке локального максимума матрица Гессе отрицательно определена или отрицательно полуопределена:

$$(\Delta \mathbf{x})' \frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^*) (\Delta \mathbf{x}) \leqslant 0 \quad (3.1.24)$$

для всех $\Delta \mathbf{x}$.

Для того чтобы строгий локальный максимум достигался в $\mathbf{x}^*$, достаточно, чтобы точка $\mathbf{x}^*$ была стационарной точкой, в которой матрица Гессе отрицательно определена. Иначе говоря, если выполнены условия

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = 0 \quad (3.1.25)$$

$$(\Delta \mathbf{x})' \frac{\partial^2 F}{\partial \mathbf{x}^2}(\mathbf{x}^*) (\Delta \mathbf{x}) < 0,$$

то $\mathbf{x}^*$ является точкой строгого локального максимума

$$F(\mathbf{x}^*) > F(\mathbf{x}^* + \Delta \mathbf{x}). \quad (3.1.26)$$

Условия первого порядка для существования локального максимума функции двух переменных при отсутствии ограничений, т. е. когда требуется найти

$$\max_{x_1, x_2} F(x_1, x_2), \quad (3.1.27)$$

состоят в том, чтобы точка $\mathbf{x}^* = (x_1^*, x_2^*)'$ была стационарной точкой

$$\frac{\partial F}{\partial x_1}(\mathbf{x}^*) = 0, \quad \frac{\partial F}{\partial x_2}(\mathbf{x}^*) = 0. \quad (3.1.28)$$

Необходимые условия второго порядка в данном случае заключаются в том, что матрица Гессе отрицательно определена или отрицательно полуопределена, а это эквивалентно выполнению следующих условий, налагаемых на главные миноры матрицы:

$$\frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) \leqslant 0 \quad (3.1.29)$$

$$\begin{vmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\mathbf{x}^*) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\mathbf{x}^*) & \frac{\partial^2 F}{\partial x_2^2}(\mathbf{x}^*) \end{vmatrix} \geqslant 0. \quad (3.1.30)$$

В завершение анализа данного случая укажем, что если в стационарной точке

$$\frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) \geqslant 0 \quad (3.1.31)$$

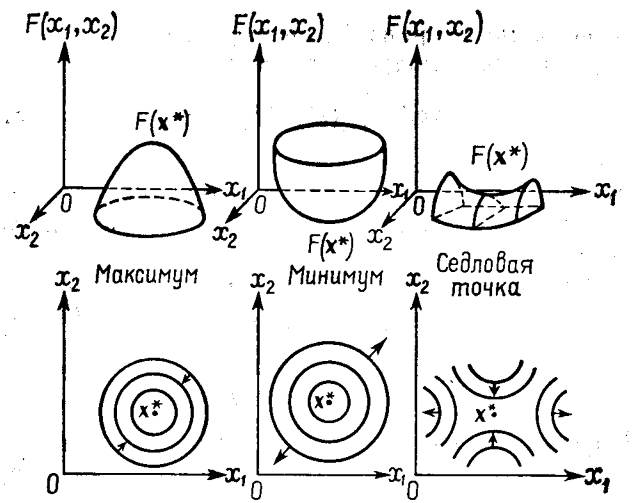

Рис. 3.2. Максимум, минимум и седловая точка в задаче оптимизации функции двух переменных при отсутствии ограничений. В нижней части рисунка изображены линии уровня и направления наискорейшего роста.

и если выполнено (3.1.30), то $\mathbf{x}^*$ будет точкой локального минимума, если же (3.1.30) не выполнено, то $\mathbf{x}^*$ — седловая точка. Эти три варианта проиллюстрированы на рис. 3.2. В верхней части диаграммы непосредственно изображены все три типа стационарных точек, а в нижней части показаны линии уровня и направления скорейшего роста. Отметим, что в изображенной на рис. 3.2 седловой точке достигается минимум по переменной x_1 и максимум по переменной x_2.

3.2. МЕТОД МНОЖИТЕЛЕЙ ЛАГРАНЖА

Одним из наиболее эффективных методов решения классических задач программирования является *метод множителей Лагранжа* [4, 5, 6]. Этому методу в настоящей книге придается особое значение, поскольку он будет неоднократно использоваться в качестве основного подхода к решению почти всех видов задач оптимизации; кроме того, с его помощью можно получить ценную информацию о том, в какой степени оптимальное значение целе-

вой функции чувствительно к изменениям констант ограничений. Последнее обстоятельство позволяет придавать множителям Лагранжа важный экономический смысл в задачах рациональной экономической деятельности.

Для первоначального ознакомления с указанным методом рассмотрим задачу на максимум с одной степенью свободы, в которой $n = 2$, $m = 1$:

найти

$$\max_{x_1, x_2} F(x_1, x_2) \tag{3.2.1}$$

при условии, что $g(x_1, x_2) = b$.

Допустим, что в $\mathbf{x}^* = (x_1^*, x_2^*)'$ существует локальное решение задачи и что в этой точке одна из частных производных функции ограничений не равна нулю. Пусть переменные занумерованы так, что

$$\frac{\partial g}{\partial x_2}(\mathbf{x}^*) \neq 0. \tag{3.2.2}$$

Тогда выражение для полного дифференциала

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0 \tag{3.2.3}$$

можно в окрестности $\mathbf{x}^*$ записать в виде

$$\frac{dx_2}{dx_1} = -\frac{\partial g/\partial x_1}{\partial g/\partial x_2}. \tag{3.2.4}$$

Решая это уравнение, представим x_2 как функцию x_1:

$$x_2 = h(x_1), \quad \text{где} \quad \frac{dh}{dx_1} = -\frac{\partial g/\partial x_1}{\partial g/\partial x_2}. \tag{3.2.5}$$

Теперь исходную задачу можно представить как задачу оптимизации функции одной переменной при отсутствии ограничений; найти

$$\max_{x_1} H(x_1) = F(x_1, h(x_1)). \tag{3.2.6}$$

Согласно результатам предыдущего раздела, условие первого порядка для существования локального максимума функции состоит в том, что

$$\frac{dH}{dx_1} = \frac{\partial F}{\partial x_1} + \frac{\partial F}{\partial x_2}\frac{dh}{dx_1} = 0. \tag{3.2.7}$$

Используя выражение (3.2.5), получим

$$\frac{dH}{dx_1}=\frac{\partial F}{\partial x_1}-\left(\frac{\partial F/\partial x_2}{\partial g/\partial x_2}\right)\frac{\partial g}{\partial x_1}=0. \qquad (3.2.8)$$

Очевидно, что также справедливо

$$\frac{\partial F}{\partial x_2}-\left(\frac{\partial F/\partial x_2}{\partial g/\partial x_2}\right)\frac{\partial g}{\partial x_2}=0. \qquad (3.2.9)$$

Введем новую переменную y

$$y=\frac{\partial F/\partial x_2}{\partial g/\partial x_2}. \qquad (3.2.10)$$

Из условия существования локального максимума необходимо следует, что

$$\frac{\partial F}{\partial x_j}-y\,\frac{\partial g}{\partial x_j}=0, \quad j=1,\,2 \qquad (3.2.11)$$

или, исключая из этих выражений переменную y,

$$\frac{\partial F/\partial x_1}{\partial F/\partial x_2}=\frac{\partial g/\partial x_1}{\partial g/\partial x_2}. \qquad (3.2.12)$$

Геометрическая интерпретация решения дана на рис. 3.3. Полный дифференциал функции при перемещении вдоль линии постоянного уровня $F(x_1,\,x_2)=\text{const}$, т. е.

$$dF=\frac{\partial F}{\partial x_1}\,dx_1+\frac{\partial F}{\partial x_2}\,dx_2=0. \qquad (3.2.13)$$

Из последнего выражения вытекает следующее свойство углового коэффициента касательной к кривой ограничений:

$$\left.\frac{dx_2}{dx_1}\right|_{\substack{\text{линия}\\ \text{уровня}}}=-\frac{\partial F/\partial x_1}{\partial F/\partial x_2}. \qquad (3.2.14)$$

Однако, согласно (3.2.4), оказывается, что

$$\left.\frac{dx_2}{dx_1}\right|_{\substack{\text{кривая}\\ \text{ограничений}}}=-\frac{\partial g/\partial x_1}{\partial g/\partial x_2}. \qquad (3.2.15)$$

Следовательно, из условия первого порядка для максимума (3.2.12) вытекает, что решение достигается в точке касания линии уровня и кривой ограничений, иначе говоря, там, где угловые коэффициенты касательных к этим кривым равны

$$\left.\frac{dx_2}{dx_1}\right|_{\substack{\text{линия}\\ \text{уровня}}}=\left.\frac{dx_2}{dx_1}\right|_{\substack{\text{кривая}\\ \text{ограничений}}}. \qquad (3.2.16)$$

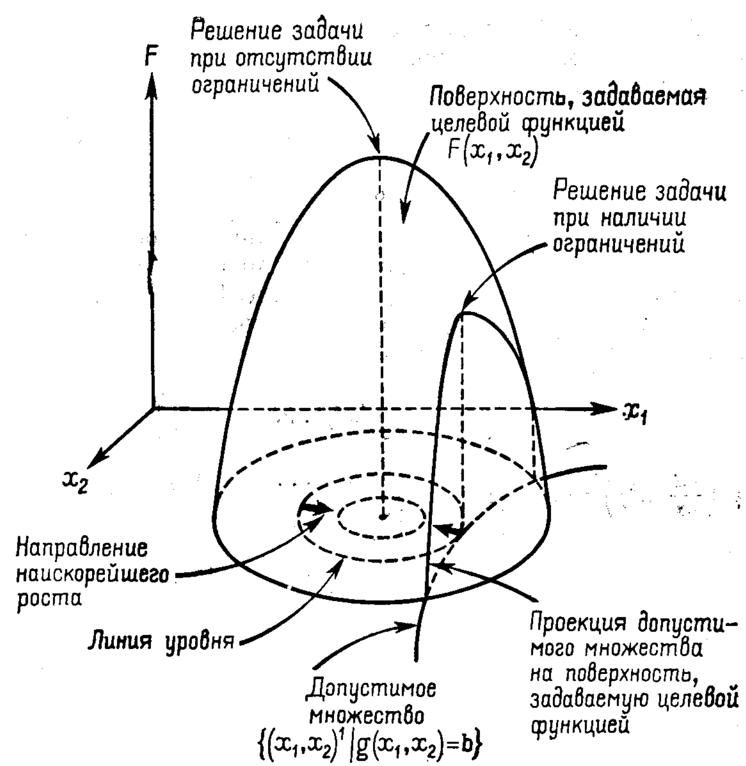

Рис. 3.3. Максимизация функции двух переменных при одном ограничении.

Отметим теперь одно весьма существенное обстоятельство, а именно: необходимые условия (3.2.11) и исходное ограничение можно было бы получить как условия стационарности некоторой точки функции

$$L(x_1,\,x_2,\,y)=F(x_1,\,x_2)+y\,(b-g(x_1,\,x_2)). \qquad (3.2.17)$$

Эти условия состоят в том, что

$$\frac{\partial L}{\partial x_j}=\frac{\partial F}{\partial x_j}-y\,\frac{\partial g}{\partial x_j}=0, \quad j=1,\,2 \qquad (3.2.18)$$

$$\frac{\partial L}{\partial y}=b-g(x_1,\,x_2)=0. \qquad (3.2.19)$$

Переменную y называют *множителем Лагранжа*, а функцию $L(\cdots)$ — *функцией Лагранжа* (лагранжианом).

Аналогичный подход можно применить к классической задаче математического программирования в общей форме: найти

$$\max_{\mathbf{x}} F(\mathbf{x}) \tag{3.2.20}$$

при условии, что $\mathbf{g}(\mathbf{x})=\mathbf{b}$.

Предположим, что задача имеет локальное решение в $\mathbf{x}^*$ и что функции ограничений удовлетворяют *условию Якоби*, т. е. в точке $\mathbf{x}^*$ ранг ρ матрицы Якоби — матрицы частных производных функций ограничений — совпадает с числом строк матрицы

$$\rho\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}}(\mathbf{x}^*)\right)=\rho\begin{pmatrix}\frac{\partial g_1}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}^*) & \ldots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}^*)\\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}^*) & \ldots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}^*)\\ \vdots & \vdots & & \vdots\\ \frac{\partial g_m}{\partial x_1}(\mathbf{x}^*) & \frac{\partial g_m}{\partial x_2}(\mathbf{x}^*) & \ldots & \frac{\partial g_m}{\partial x_n}(\mathbf{x}^*)\end{pmatrix}=m. \tag{3.2.21}$$

Отметим, что в приведенной выше задаче с одной степенью свободы матрица Якоби есть вектор-строка. Ранг этой матрицы равен единице тогда и только тогда, когда хотя бы одна из частных производных функций ограничений не обращается в нуль. При необходимости можно перенумеровать переменные таким образом, чтобы определитель матрицы, составленной из последних столбцов матрицы Якоби, не был бы равен нулю и чтобы вектор инструментальных переменных можно было представить в виде $\mathbf{x}=(\mathbf{x}^1, \mathbf{x}^2)'$, где вектор $\mathbf{x}^1$ состоит из $n-m$ переменных, а вектор $\mathbf{x}^2$ — из m переменных. Поскольку условие Якоби выполнено, то по теореме о неявной функции в окрестности $\mathbf{x}^*$ можно разрешить систему ограничений, представив $\mathbf{x}^2$ в виде функции $\mathbf{x}^1$

$$\mathbf{x}^2=\mathbf{h}(\mathbf{x}^1), \tag{3.2.22}$$

где $\mathbf{h}$ — вектор-столбец, состоящий из m функций. Теперь задачу (3.2.20) можно свести к задаче оптимизации при отсутствии ограничений

$$\max_{\mathbf{x}^1} H(\mathbf{x}^1)=F(\mathbf{x}^1, \mathbf{h}(\mathbf{x}^1)). \tag{3.2.23}$$

Согласно результатам предыдущего параграфа, необходимое условие для существования локального максимума состоит в том, что

$$\frac{\partial H}{\partial\mathbf{x}^1}=\frac{\partial F}{\partial\mathbf{x}^1}+\frac{\partial F}{\partial\mathbf{x}^2}\frac{\partial\mathbf{h}}{\partial\mathbf{x}^1}=\mathbf{0}, \tag{3.2.24}$$

где $\frac{\partial H}{\partial\mathbf{x}^1}$ представляет собой вектор-строку, состоящую из $n-m$ элементов, а $\frac{\partial\mathbf{h}}{\partial\mathbf{x}^1}$ — матрицу размерности $m\times(n-m)$. Так как ограничения можно записать в виде тождества

$$\mathbf{g}(\mathbf{x}^1, \mathbf{h}(\mathbf{x}^1))\equiv\mathbf{b}, \tag{3.2.25}$$

то после дифференцирования получим

$$\frac{\partial\mathbf{g}}{\partial\mathbf{x}^1}+\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\frac{\partial\mathbf{h}}{\partial\mathbf{x}^1}=\mathbf{0}. \tag{3.2.26}$$

Матрица $\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}$ размерности $m\times m$ невырожденная, так как выполняется условие Якоби. Следовательно,

$$\frac{\partial\mathbf{h}}{\partial\mathbf{x}^1}=-\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\right)^{-1}\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^1}\right) \tag{3.2.27}$$

и условия (3.2.24) можно записать следующим образом:

$$\frac{\partial F}{\partial\mathbf{x}^1}-\left(\frac{\partial F}{\partial\mathbf{x}^2}\right)\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\right)^{-1}\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^1}\right)=\mathbf{0}. \tag{3.2.28}$$

Очевидно также, что

$$\frac{\partial F}{\partial\mathbf{x}^2}-\left(\frac{\partial F}{\partial\mathbf{x}^2}\right)\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\right)^{-1}\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\right)=\mathbf{0}. \tag{3.2.29}$$

Полагая теперь

$$\mathbf{y}=\left(\frac{\partial F}{\partial\mathbf{x}^2}\right)\left(\frac{\partial\mathbf{g}}{\partial\mathbf{x}^2}\right)^{-1}=(y_1, y_2, \ldots, y_m), \tag{3.2.30}$$

можно представить необходимые условия (3.2.28) и (3.2.29) в форме

$$\frac{\partial F}{\partial\mathbf{x}}-\mathbf{y}\frac{\partial\mathbf{g}}{\partial\mathbf{x}}=\mathbf{0}. \tag{3.2.31}$$

Эти необходимые условия наряду с исходными ограничениями можно получить, дифференцируя по $\mathbf{x}$ и по $\mathbf{y}$ функцию

$$F(\mathbf{x})+\mathbf{y}(\mathbf{b}-\mathbf{g}(\mathbf{x})). \tag{3.2.32}$$

Перечислим, из каких основных этапов состоит решение по методу Лагранжа классической задачи математического программирования, то есть следующей задачи: найти $\max_{\mathbf{x}} F(\mathbf{x})$ при условии, что $\mathbf{g}(\mathbf{x}) = \mathbf{b}$. (3.2.33)

На первом этапе вводится вектор-строка из m новых переменных

$$\mathbf{y} = (y_1, y_2, \ldots, y_m), \tag{3.2.34}$$

называемых *множителями Лагранжа*. На втором этапе определяется *функция Лагранжа* как сумма целевой функции и скалярного произведения вектора множителей Лагранжа и вектора разности между постоянными ограничениями и функциями ограничений

$$L(\mathbf{x}, \mathbf{y}) = F(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})), \tag{3.2.35}$$

или в развернутом виде

$$L(x_1, x_2, \ldots, x_n; y_1, y_2, \ldots, y_m) = F(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} y_i (b_i - g_i(x_1, x_2, \ldots, x_n)). \tag{3.2.36}$$

Последним этапом является отыскание точки $(\mathbf{x}^*, \mathbf{y}^*)$, в которой все частные производные первого порядка функции Лагранжа обращаются в нуль:

$$\begin{gathered} \frac{\partial L}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{y}^*) = \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) - \mathbf{y}^* \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*) = 0, \\ \frac{\partial L}{\partial \mathbf{y}}(\mathbf{x}^*, \mathbf{y}^*) = \mathbf{b} - \mathbf{g}(\mathbf{x}^*) = 0. \end{gathered} \tag{3.2.37}$$

Первые n из полученных соотношений, совпадающие с условиями (3.2.31), показывают, что градиент целевой функции должен равняться вектору множителей Лагранжа, умноженному на матрицу Якоби для функций ограничений, т. е.

$$\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{y}^* \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*), \tag{3.2.38}$$

или в развернутом виде

$$\frac{\partial F}{\partial x_j}(x_1^*, x_2^*, \ldots, x_n^*) = \sum_{i=1}^{m} y_i^* \frac{\partial g_i}{\partial x_j}(x_1^*, x_2^*, \ldots, x_n^*), \quad j = 1, 2, \ldots, n. \tag{3.2.39}$$

Остальные m условий представляют собой просто систему ограничений

$$\mathbf{g}(\mathbf{x}^*) = \mathbf{b}. \tag{3.2.40}$$

Решая совместно $m + n$ уравнений (3.2.37), получим значения $m + n$ неизвестных: инструментальных переменных $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)'$ и множителей Лагранжа $\mathbf{y}^* = (y_1^*, y_2^*, \ldots, y_m^*)$. Если выполнены некоторые достаточные условия, указанные ниже, то значения переменных $\mathbf{x}^*$ дают локальное решение классической задачи математического программирования. Этот вывод достаточно очевиден, если учесть тот факт, что $\mathbf{x}^*$ удовлетворяет ограничениям [1] и максимизирует функцию Лагранжа, которая в точке $(\mathbf{x}^*, \mathbf{y}^*)$ просто совпадет со значением целевой функции,

$$L(\mathbf{x}^*, \mathbf{y}^*) = F(\mathbf{x}^*). \tag{3.2.41}$$

Дадим геометрическую интерпретацию $m + n$ условий первого порядка

$$\begin{gathered} \mathbf{g}(\mathbf{x}) = \mathbf{b} \\ \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}^*) = \mathbf{y}^* \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*). \end{gathered} \tag{3.2.42}$$

Для этого заметим, что если определить i-ю кривую ограничений как

$$\{\mathbf{x} \in E^n \mid g_i(\mathbf{x}) = b_i\}, \tag{3.2.43}$$

то вектор градиента i-й функции ограничений

$$\frac{\partial g_i}{\partial \mathbf{x}} = \left(\frac{\partial g_i}{\partial x_1}, \frac{\partial g_i}{\partial x_2}, \ldots, \frac{\partial g_i}{\partial x_n}\right), \tag{3.2.44}$$

являющийся не чем иным, как i-й строкой матрицы Якоби, будет ортогональным (направленным по нормали) к этой кривой, так как

$$dg_i(\mathbf{x}) = \frac{\partial g_i}{\partial x}(\mathbf{x})\, d\mathbf{x} = 0, \quad i = 1, 2, \ldots, m. \tag{3.2.45}$$

Таким образом, условия (3.2.42) показывают, что $\mathbf{x}^*$ лежит в допустимом множестве X и что направление скорейшего роста в $\mathbf{x}^*$ (вектор градиента целевой функции) представляет собой линейную комбинацию (взвешенную сумму) нормалей к кривым ограничений (градиентов к кривым ограничений), взвешивающими коэффициентами в которой являются множители Лагранжа $\mathbf{y}^*$.

---

[1] Строгое доказательство приводится в гл. 4.

6—0270

Необходимые условия второго порядка состоят в том, что матрица Гессе

$$\frac{\partial^2 L}{\partial \mathbf{x}^2}=\begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_1\,\partial x_n} \\ \frac{\partial^2 L}{\partial x_2\,\partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \cdots & \frac{\partial^2 L}{\partial x_2\,\partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 L}{\partial x_n\,\partial x_1} & & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}, \qquad (3.2.46)$$

составленная из вторых частных производных лагранжиана по инструментальным переменным, должна быть отрицательно определенной или отрицательно полуопределенной в точке локального максимума ($\mathbf{x}^*$, $\mathbf{y}^*$) при том условии, что

$$d\mathbf{g}=\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}^*)\,d\mathbf{x}=\mathbf{0}. \qquad (3.2.47)$$

Если при указанных условиях матрица Гессе является отрицательно определенной, то условия первого порядка (3.2.42) являются достаточными для существования локального максимума [7]. Условие, что матрица Гессе (3.2.46) является отрицательно определенной при ограничениях (3.2.47), может быть представлено в форме $n-m$ условий, налагаемых на знаки некоторых миноров матрицы размерности $(m+n)\times(m+n)$

$$\begin{pmatrix} 0 & \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \\ \frac{\partial \mathbf{g}'}{\partial \mathbf{x}} & \frac{\partial^2 L}{\partial \mathbf{x}^2} \end{pmatrix}=\left(\begin{array}{ccc|cccc} 0 & 0 & \cdots\ 0 & \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & & & & & \\ 0 & 0 & \cdots\ 0 & \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \\ \hline \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1\,\partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_1\,\partial x_n} \\ \vdots & & & \vdots & & & \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_m}{\partial x_n} & \frac{\partial^2 L}{\partial x_n\,\partial x_1} & \frac{\partial^2 L}{\partial x_n\,\partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{array}\right), \qquad (3.2.48)$$

полученной в результате окаймления матрицы Гессе матрицами Якоби для функций ограничений. Условия локального максимума в такой форме заключаются в том, что последние $n-m$ главных миноров этой окаймленной матрицы Гессе имеют чередующиеся знаки, причем знак первого из них совпадает со знаком $(-1)^{m+1}$.

3.3. ИНТЕРПРЕТАЦИЯ МНОЖИТЕЛЕЙ ЛАГРАНЖА

Решение системы уравнений (3.2.42), выражающих условия первого порядка, содержит, кроме вектора локального оптимума $\mathbf{x}^*$, еще и вектор множителей Лагранжа $\mathbf{y}^*$. Если выполнено условие Якоби, то существует единственный вектор $\mathbf{y}^*$, соответствующий локальному решению $\mathbf{x}^*$. Знание значений множителей Лагранжа не является излишним; напротив, с их помощью можно получить ценную информацию о задаче. Широкое практическое использование метода множителей Лагранжа во многом объясняется именно последним обстоятельством. Множители Лагранжа, соответствующие решению задачи, измеряют чувствительность оптимального значения целевой функции $F^*=F(\mathbf{x}^*)$ к изменениям констант ограничений $\mathbf{b}$:

$$\mathbf{y}^*=\frac{\partial F^*}{\partial \mathbf{b}}, \qquad (3.3.1)$$

т. е.

$$y_i^*=\frac{\partial F^*}{\partial b_i}, \quad i=1,\,2,\,\ldots,\,m. \qquad (3.3.2)$$

Для доказательства соотношений (3.3.1) необходимо сначала показать, что если величины b_i рассматриваются как переменные, то переменные x_j и y_i ($j=1,\,\ldots,\,n$; $i=1,\,\ldots,\,m$) можно представить в виде функций b_i. Рассмотрим с этой целью условия первого порядка (3.2.42), которые можно записать в виде системы $m+n$ уравнений с $2m+n$ неизвестными ($\mathbf{b}$, $\mathbf{y}$, $\mathbf{x}$), если считать b_i ($i=1,\,2,\,\ldots,\,m$) переменными величинами

$$\begin{gathered}\psi^1(\mathbf{b},\,\mathbf{y},\,\mathbf{x})\equiv\mathbf{b}-\mathbf{g}(\mathbf{x})=\mathbf{0},\\ \psi^2(\mathbf{b},\,\mathbf{y},\,\mathbf{x})\equiv\frac{\partial F}{\partial \mathbf{x}}(\mathbf{x})-\mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x})=\mathbf{0}.\end{gathered} \qquad (3.3.3)$$

Запишем матрицу Якоби этой системы уравнений в следующем виде:

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & -\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right) \\ \mathbf{0} & -\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)' & \frac{\partial^2 L}{\partial \mathbf{x}^2} \end{pmatrix}, \qquad (3.3.4)$$

где $\mathbf{I}$ — единичная матрица порядка m. Ранг матрицы Якоби равен m, если выполнены достаточные условия, налагаемые на окаймленную матрицу Гессе (3.2.48). Следовательно, по теореме о неявной функции, решая систему уравнений, составленную из $m+n$ условий первого порядка, можно представить инструментальные переменные и множители Лагранжа в виде функций от постоянных ограничений $\mathbf{b}$

$$\begin{aligned} \mathbf{y} &= \mathbf{y}(\mathbf{b}) \\ \mathbf{x} &= \mathbf{x}(\mathbf{b}). \end{aligned} \tag{3.3.5}$$

Рассмотрим теперь функцию Лагранжа как функцию, зависящую от постоянных ограничений:

$$L(\mathbf{b}) = F(\mathbf{x}(\mathbf{b})) + \mathbf{y}(\mathbf{b})\,[\mathbf{b} - \mathbf{g}(\mathbf{x}(\mathbf{b}))]. \tag{3.3.6}$$

Дифференцирование по $\mathbf{b}$ дает

$$\frac{\partial L}{\partial \mathbf{b}} = \frac{\partial F}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{b}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{b}} + (\mathbf{b} - \mathbf{g}(\mathbf{x}))'\frac{\partial \mathbf{y}'}{\partial \mathbf{b}} + \mathbf{y} =$$
$$= \left(\frac{\partial F}{\partial \mathbf{x}} - \mathbf{y}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)\left(\frac{\partial \mathbf{x}}{\partial \mathbf{b}}\right) + (\mathbf{b} - \mathbf{g}(\mathbf{x}))'\frac{\partial \mathbf{y}'}{\partial \mathbf{b}} + \mathbf{y}. \tag{3.3.7}$$

Первые два члена этого выражения в точке $(\mathbf{x}^*, \mathbf{y}^*)$ обращаются в нуль из-за условий первого порядка (3.2.37), так что $\frac{\partial L}{\partial \mathbf{b}}$ равно вектору множителей Лагранжа. Однако значение функции Лагранжа в точке $(\mathbf{x}^*, \mathbf{y}^*)$ суть оптимальное значение целевой функции (3.2.41). Следовательно,

$$\frac{\partial L}{\partial \mathbf{b}}(\mathbf{x}^*, \mathbf{y}^*) = \frac{\partial F^*}{\partial \mathbf{b}} = \mathbf{y}^*, \tag{3.3.8}$$

что и утверждалось ранее. Таким образом, помимо того, что метод множителей Лагранжа дает решение классической задачи максимизации, он позволяет также проанализировать, насколько оптимальное значение целевой функции чувствительно к изменениям констант ограничений. Например, если какой-то множитель Лагранжа равен нулю, то малые изменения соответствующей константы ограничений не окажут никакого влияния на оптимальное значение целевой функции. Особенно важна интерпретация множителей Лагранжа в задачах рациональной экономической деятельности. В экономических задачах распределения ресурсов целевая функция имеет размерность стоимости, т. е. цены, умноженной на объем

продукции (таковы, например, прибыль, выручка, издержки), а с помощью ограничений устанавливается определенное значение некоторого количества (например, затрат). Поскольку в таких задачах с помощью множителя Лагранжа измеряют чувствительность величины, имеющей размерность стоимости к изменениям некоторого количества, то он имеет размерность цены. По этой причине множитель Лагранжа часто называют *теневой ценой* (данного вида затрат).

ЗАДАЧИ

3-А. Показать на чертеже допустимое множество для классической задачи максимизации с двумя переменными и одним ограничением, где в качестве ограничения берутся следующие соотношения:

1. $x_1 = 10$
2. $2x_1 + 4x_2 = 8$
3. $x_1^2 + 4x_2^2 = 36$
4. $x_1^2 + 3x_1 = 1$
5. $(x_1 - 1)^2 + (x_2 - 6)^2 = 0$
6. $\sin(x_1^2 + x_2^2) = 0$
7. $\ln x_2 = 0$
8. $e^{x_1} - x_1 = 0.$

3-Б. Докажите, что:

1. Следующие две задачи имеют одинаковое решение:

а) найти $\max\limits_{\mathbf{x}} F(\mathbf{x})$

при условии, что $\mathbf{g}(\mathbf{x}) = \mathbf{b}$;

б) найти $\min\limits_{\mathbf{x}} -F(\mathbf{x})$

при условии, что $\mathbf{g}(\mathbf{x}) = \mathbf{b}$.

2. Докажите: если $F(\mathbf{x})$ — строго вогнутая функция, то стационарная точка представляет собой точку строгого локального максимума.

3-В. Решите следующие задачи графически, используя линии уровня, направления скорейшего роста и допустимые множества:

1. Найти $\max x_1 + x_2$

при условии, что $x_1^2 + x_2^2 = 1$.

2. Найти $\max \exp(-(x_1 + 2x_2))$
при условии, что $2x_1 + 3x_2 = 4$.
3. Найти $\max \sin x_1 \cos x_2$
при условии, что $x_1 - x_2 = 0$.

3-Г. Пусть в задаче оптимизации функции скалярного аргумента при отсутствии ограничений ($n = 1$, $m = 0$) первые k производных функции $F(x)$ тождественно равны нулю в x^*. Используя разложения по формуле Тейлора, получить условия для локального максимума. Распространить полученный результат на векторный случай ($n > 1$).

3-Д. Используя результаты решения предыдущей задачи, рассмотрите следующие примеры:

1. Имеет ли функция $F(x) = 1 - \exp(-x^2)$ максимум или минимум при $x = 0$?
2. Найдите максимум функции $F(x_1, x_2) = (x_1 - a_1^2 x_2^2) \times (x_1 - a_2^2 x_2^2)$, где a_1 и a_2 — постоянные, причем $a_1^2 \neq a_2^2$.

3-Е. Определите достаточные условия второго порядка для задачи на максимум с одной степенью свободы (задача (3.2.1)) ($n = 2$, $m = 1$) с помощью окаймленной матрицы Гессе. Покажите, что те же условия можно получить исходя из того, что

$d^2H/dx_1^2 < 0$, где $H(x_1) = F(x_1, h(x_1))$, как и в (3.2.6).

3-Ж. Рассмотрите классическую задачу программирования с квадратической целевой функцией и линейными ограничениями,

т. е. следующую задачу: найти

$$\max_{\mathbf{x}} \mathbf{cx} + \frac{1}{2}\mathbf{x}'\mathbf{Dx}$$

при условии, что $\mathbf{Ax} = \mathbf{b}$,

где $\mathbf{c}$ — фиксированный вектор-строка, $\mathbf{D}$ — фиксированная матрица размерности $n \times n$, отрицательно определенная (и, следовательно, невырожденная), $\mathbf{A}$ — заданная матрица размерности $m \times n$.

1. Построить функцию Лагранжа и найти условия первого порядка.

2. Найти оптимальный вектор $\mathbf{x}^*$ как функцию $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$ и $\mathbf{D}$. Проверить, что $\mathbf{x}^*$ — допустимое решение и что при любом другом допустимом векторе целевая функция принимает меньшее значение, т. е. что $\mathbf{x}^*$ — точка глобального максимума.

3. Определить параметры чувствительности $\frac{\partial \mathbf{x}^*}{\partial \mathbf{b}}$ и $\frac{\partial \mathbf{x}^*}{\partial \mathbf{c}}$ и проверить, что с изменением этих параметров $\mathbf{x}^*$ меняется линейно.

3-З. Рассмотрите задачу отыскания максимума квадратичной формы при условии, что сумма квадратов переменных равна 1,
т. е. найти

$$\max_{\mathbf{x}} F(\mathbf{x}) = \mathbf{x}'\mathbf{Ax}$$

при условии, что $\mathbf{x}'\mathbf{x} = 1$,

где $\mathbf{A}$ — симметрическая матрица. Покажите, что если $\mathbf{x}^*$ есть решение задачи, то $F(\mathbf{x}^*)$ равно наибольшему характеристическому корню матрицы $\mathbf{A}$. Проиллюстрируйте этот результат геометрически при $n = 2$. При каких условиях $F(\mathbf{x}^*) = 0$?

3-И. Рассмотрите задачу: найти

$$\min_{x_1, x_2} x_1^2 + x_2^2$$

при условии, что $(x_1 - 1)^3 - x_2^2 = 0$.

1. Решить задачу геометрически.
2. Показать, что метод множителей Лагранжа в данном случае неприменим. Почему?

3-К. Если воспользоваться методом *ограниченной вариации*, то можно получить необходимые условия для существования решения классических задач математического программирования, используя следующие $m + 1$ соотношений:

$$dF = \frac{\partial F}{\partial \mathbf{x}} d\mathbf{x} = 0$$

$$d\mathbf{g} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x})\, d\mathbf{x} = \mathbf{0},$$

которые должны одновременно выполняться в точке решения. Запишем эти уравнения в виде матричного уравнения

$$\begin{pmatrix} \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) \\ \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}) \end{pmatrix} d\mathbf{x} = \mathbf{0}.$$

Система из этих $m+1$ уравнений относительно n неизвестных имеет нетривиальное решение лишь в том случае, если

$$\rho \begin{pmatrix} \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}) \\ \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}) \end{pmatrix} = m,$$

где ρ — ранг матрицы.

1. Показать, что с помощью указанного метода можно получить $n-m$ необходимых условий.

2. Для случая $n=3$, $m=1$ показать, что необходимые условия, соответствующие данному методу, эквивалентны необходимым условиям, соответствующим методу множителей Лагранжа.

3-Л. Рассмотрим задачу: найти

$$\max F(x_1, x_2, x_3)$$

при условии, что $g(x_1, x_2, x_3) = b$.

Необходимыми условиями для существования максимума в данном случае являются

$$\frac{\partial F}{\partial x_j}(x_1, x_2, x_3) = y\frac{\partial g}{\partial x_j}(x_1, x_2, x_3), \quad j = 1, 2, 3.$$

$$g(x_1, x_2, x_3) = b.$$

Предположим, что в задачу добавлено еще одно ограничение,

имеющее вид $\frac{\partial F/\partial x_1}{\partial F/\partial x_2} = \mu\frac{\partial g/\partial x_1}{\partial g/\partial x_2}$, $\mu \neq 1$.

Каковы будут новые необходимые условия?

3-М. Рассмотрим следующую задачу:

$$\max F(x_1, x_2) = f(x_1, x_2) - w_1x_1 - w_2x_2.$$

Предположим, что матрица Гессе, составленная из частных производных второго порядка — $\partial^2 f/\partial \mathbf{x}^2$, является отрицательно определенной и что w_1 и w_2 — это фиксированные положительные параметры.

1. Каковы в данном случае необходимые условия первого порядка для максимума?

2. Показать, что x_1 можно представить в виде функции w_1 и w_2 и что $(\partial x_1/\partial w_1) < 0$.

3. Пусть к задаче добавлено линейное ограничение $x_2 = b$, где b — фиксированный ненулевой параметр. Найти новое состояние равновесия и показать, что

$$\left(\frac{\partial x_1}{\partial w_1}\right)\Big|_{\text{без дополнительного ограничения}} \leqslant \left(\frac{\partial x_1}{\partial w_1}\right)\Big|_{\text{с дополнительным ограничением}} < 0.$$

Этот результат является иллюстрацией *принципа Ле Шателье* [4].

3-Н. Примером классической задачи математического программирования является задача об определении *оптимального размера партии* товара в теории управления запасами. Имеющийся у фирмы запас некоторого однородного товара $I(t)$ уменьшается с постоянной скоростью dI/dt, т. е. в каждую единицу времени используется одно и то же количество товара. Фирма производит заказы на некоторое количество этого товара x, которое поставляется ей немедленно всякий раз, когда уровень запасов становится равным нулю. Ежегодная потребность в товаре равна A, и фирма производит заказы на поставку ей новых партий товара n раз в течение года, так что

$$A = nx.$$

При хранении товара фирма несет убытки, связанные с расходами на *хранение товара* и *оформление заказов*. Средняя величина запаса товара равна $\frac{x}{2}$, а расходы на хранение одной единицы товара равны C_h, так что общие расходы на хранение составляют $C_h\frac{x}{2}$. Как сказано выше, фирма производит заказы на поставку новых партий товара n раз в течение года. Расходы на оформление одного отдельного заказа равны C_0, поэтому общие расходы на оформление заказов равны C_0n. Таким образом, общие расходы составляют

$$C = C_h\frac{x}{2} + C_0n.$$

1. Показать графически, как меняется во времени уровень запасов. Доказать, что средний уровень запасов равен $x/2$.

2. Минимизировать общие расходы на хранение товара C, выбирая x и n при условии, что $A = nx$. (Указание: воспользоваться методом множителей Лагранжа.) Найдите оптимальную величину поставки (оптимальное значение x) как функцию параметров C_0, C_h и A. Объясните, какой смысл имеет множитель Лагранжа в данной задаче.

3. Третий вид расходов фирмы составляют *штрафные платежи* за невыполненные ею заказы. Этот вид расходов ранее не упоминался, поскольку предполагалось, что у фирмы всегда имеется нужный запас товаров. Предположим теперь, что фирма производит заказы даже не при нулевом уровне запасов, а только тогда, когда уровень не выполненных ею заказов достигает определенного значения U, и что в этот момент она производит все необходимые поставки своим заказчикам. Стоимость одного невыполненного заказа равна C_p. Найти оптимальные уровни x и U.

3-О. *Метод наименьших квадратов*, применяющийся в теории регрессий, состоит в следующем: по имеющимся данным (x_i, y_i) $(i = 1, \ldots, n)$ строится такая кривая $y = a + bx$, на которой достигается минимум суммы квадратов отклонений

$$S(a, b) = \sum_{i=1}^{n} (y_i - (a + bx_i))^2,$$

т. е. минимизируется функция, зависящая от двух параметров: a («отрезок на оси ординат») и b («наклон»).

1. Определить необходимые условия минимизации функции $S(a, b)$. (Уравнения, дающие эти условия, называются *нормальными уравнениями*.) Показать, что выполнены достаточные условия.

2. Распространите результаты на тот случай, когда подбираемая функция является квадратической, т. е. когда

$$y = a + bx + cx^2.$$

3. Распространите полученные результаты на случай *множественной регрессии*, т. е. когда подбираемая функция имеет вид

$$y = \pi_1 x_1 + \pi_2 x_2 + \ldots + \pi_k x_k = \boldsymbol{\pi}\mathbf{x},$$

где y — зависимая переменная; $\boldsymbol{\pi}$ — вектор-строка коэффициентов наклона (свободный член можно получить, полагая $x_k \equiv 1$), а $\mathbf{x}$ есть k-мерный вектор-столбец независимых переменных. В этом случае сумма квадратов отклонений равна

$$S(\pi_1, \ldots, \pi_k) = S(\boldsymbol{\pi}) = \sum_{i=1}^{n} (y_i - \boldsymbol{\pi}\mathbf{x}_i)^2 = (\mathbf{Y} - \boldsymbol{\pi}\mathbf{X})(\mathbf{Y} - \boldsymbol{\pi}\mathbf{X})',$$

где y_i — результат наблюдения значения зависимой переменной в i-й точке выборки $(i = 1, \ldots, n)$; $\mathbf{Y}$ — n-мерный вектор-строка наблюденных значений зависимой переменной в каждой из точек выборки

$$\mathbf{Y} = (y_1, \ldots, y_n);$$

$\mathbf{x}_i$ — вектор наблюденных значений независимых переменных в i-й точке выборки, а $\mathbf{X}$ — матрица размерности $k \times n$, составленная из результатов наблюдений k независимых переменных в n точках выборки

$$\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) = \begin{pmatrix} x_{11} \ldots x_{1n} \\ \vdots \\ x_{k1} \ldots x_{kn} \end{pmatrix}$$

(предполагается, что $\rho(\mathbf{X}) = k < n$).