Свиридова О.В., Рыбанов А.А.

Методы многомерной оптимизации

Волжский

2021

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ВОЛЖСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ (ФИЛИАЛ)
ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО
УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

О.В. Свиридова, А.А. Рыбанов

Методы многомерной оптимизации

Электронное учебное пособие

Волжский

2021

УДК 004(07)
ББК 32.97
С 247

Рецензенты:
заведующий кафедрой математики, информатики и естественных наук Волжского филиала Волгоградского государственного университета, канд. физ.-мат. наук, доцент
Полковников А.А.
канд. педагогических наук, доцент
кафедры методики преподавания математики и физики, ИКТ
ФГБОУ ВО «Волгоградский государственный социально-педагогический университет (ВГСПУ)»
Филиппова Е.М.

Издается по решению редакционно-издательского совета
Волгоградского государственного технического университета

Свиридова, О.В.

Методы многомерной оптимизации [Электронный ресурс] : учебное пособие / О. В. Свиридова, А. А. Рыбанов ; Министерство науки и высшего образования Российской Федерации, ВПИ (филиал) ФГБОУ ВО ВолгГТУ. – Электронные текстовые данные (1 файл: 2,55 МБ). – Волжский, 2021. – Режим доступа: http://lib.volpi.ru. – Загл. с титул. экрана.

ISBN 978-5-9948-4052-8

В учебном пособии представлены два раздела многомерной оптимизации – аналитические и численные методы. Каждый раздел пособия содержит теоретическую часть, примеры решения типовых задач, систематизированную подборку контрольных заданий.

Предназначено для студентов, обучающихся по направлению 09.03.01 «Информатика и вычислительная техника» в рамках курса «Методы оптимизации».

Ил. 10, табл. 5, библиограф.: 6 назв.

ISBN 978-5-9948-4052-8

Содержание

Введение……………………………………………………………………4

1. Аналитические методы оптимизации функции нескольких переменных..5

1.1. Безусловный экстремум функции многих переменных 5

1.2. Условный экстремум при ограничениях типа равенств..………………13

1.3. Условный экстремум при ограничениях типа неравенств. 24

1.4. Условный экстремум в замкнутой области……………………………...28

Контрольные задания...……...…………….…..……………………………35

2. Численные методы оптимизации функции нескольких переменных……36

2.1. Методы безусловной минимизации……….…...…………………………36

2.2. Метод покоординатного спуска.…...……...……………………………38

2.3. Метод градиентного спуска………...……...……………………………41

2.4. Метод наискорейшего спуска...45

2.5. Метод сопряженных градиентов ……………...48

2.6. Метод Ньютона 52

Контрольные задания...……...…………….…..……………………………57

Библиографический список……………………………………………59

ВВЕДЕНИЕ

Электронный ресурс «Методы многомерной оптимизации» является учебным пособием, которое содержит теоретический материал по оптимизации функций нескольких переменных и является дополнением к лекциям по дисциплине «Методы оптимизации».

Учебное пособие представляет два раздела многомерной оптимизации – аналитические и численные методы оптимизации функции нескольких переменных.

Электронное учебное пособие разработано в соответствии с требованиями Федерального государственного стандарта высшего профессионального образования и предназначены для студентов бакалавриата, обучающихся по направлению 09.03.01 «Информатика и вычислительная техника». Областью применения данного электронного ресурса является учебный процесс в Волжском политехническом институте (филиал) Волгоградского технического университета, объектом обучения могут выступать студенты направления «Программная инженерия», изучающие дисциплину «Исследование операций», и студенты направления «Информатика и вычислительная техника», изучающие дисциплину «Методы оптимизации».

1. Аналитические методы оптимизации функции нескольких переменных

1.1. Безусловный экстремум функции многих переменных

Постановка задачи

Дана функция многих переменных F (x_1, x_2 ,..., x_n). Если из переменных составить n-мерный вектор-столбец x, то можно записать F(x_1, x_2 ,..., x_n)= F(x). Этот вектор можно рассматривать как точку в n-мерном пространстве R^n ($x \in R^n$). Введем также следующие обозначения: градиент функции - n-мерный вектор-столбец:

$$\frac{dF(x)}{dx} = \begin{bmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \\ --- \\ \frac{\partial F}{\partial x_n} \end{bmatrix}$$

и гессиан функции – n×n-мерная матрица

$$\frac{d^2F(x)}{dx^2} = \begin{bmatrix} \frac{\partial^2 F}{\partial {x_1}^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F}{\partial x_2 x_1} & \frac{\partial^2 F}{\partial {x_2}^2} & \cdots & \frac{\partial^2 F}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1} & \frac{\partial^2 F}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 F}{\partial x_{n^2}} \end{bmatrix}$$

Пусть C – симметрическая n×n-мерная матрица с элементами C_{ij} и Δx – n-мерный вектор с элементами Δx_1, Δx_2, ..., Δx_n, T – знак транспонирования. Тогда выражение

$$q(C, \Delta x) = \Delta x^T C \Delta x = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} \Delta x_i \Delta x_j$$

называется квадратической формой. Квадратическая форма q(C, Δx) (соответственно матрица C) называется

- положительно определенной, если q(C, Δx) >0 при любых Δx≠0;
- неотрицательно определенной, если q(C, Δx)≥0 при любых Δx≠0;
- отрицательно определенной, если q(C, Δx) <0 при любых Δx≠0;
- неположительно определенной, если q(C, Δx) ≤ 0 при любых Δx≠0;
- знаконеопределенной, если q(C, Δx) может иметь разные знаки при разных Δx.

Проверка квадратической формы q(C, Δx) или матрицы C на знакоопределенность может быть выполнена с помощью критерия Сильвестра. Пусть $\Delta_1, \Delta_2, ..., \Delta_n$ - главные миноры матрицы C.

Критерий Сильвестра. Квадратическая форма q(C, Δx) (соответственно матрица C)

- положительно определена, если все $\Delta_i >$ 0 (i=1;…,n);.
- неотрицательно определена, если все $\Delta_i \geq 0$ (i=1;…,n);
- отрицательно определена, если все $(-1)^i \Delta_i > 0$ (i=1;…,n);
- неположительно определена, если все $(-1)^i \Delta_i \geq 0$ (i=1;…,n);
- знаконеопределена, если не выполняются предыдущие условия.

Задача. Для заданной функции F(x) нужно найти точки экстремума.

Решение

1. Составляется система уравнений (первое необходимое условие экстремума)

$$\frac{dF(x)}{dx} = 0$$

или

$$\frac{\partial F(x)}{\partial x_1} = 0,$$
$$\frac{\partial F(x)}{\partial x_2} = 0,$$
$$\ldots$$
$$\frac{\partial F(x)}{\partial x_n} = 0.$$

Эта система уравнений может иметь несколько решений $\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_k$, которые являются точками в R^n.

2. Чтобы решить вопрос, какой экстремум достигается в этих точках, нужно вычислить вторые производные функции F(x) и составить матрицу:

$$C = \frac{d^2 F(\hat{x}_m)}{dx^2} = \left.\frac{d^2 F(x)}{dx^2}\right|_{x=\hat{x}_m}$$

или соответствующую квадратическую форму q(C, Δx). Здесь $\widehat{x}_m$ – исследуемое решение. В результате получается (по второму необходимому условию экстремума):

- Если квадратическая форма q(C, Δx) (или матрица C) положительно определена, то функция F(x) в точке $\widehat{x}_m$ имеет минимум.
- Если квадратическая форма q(C, Δx) (или матрица C) отрицательно определена, то функция F(x) в точке $\widehat{x}_m$ имеет максимум.
- В остальных случаях точка $\widehat{x}_m$ является седловой.

Пример 1.1. Найти точки экстремума функции

$$F(x_1, x_2) = x_1^2 + x_2^2.$$

Решение. Составляем уравнения

$$\frac{\partial F}{\partial x_1} = 2x_1 = 0,$$
$$\frac{\partial F}{\partial x_2} = 2x_2 = 0.$$

Они имеют единственное решение $x_1 = x_2 = 0$. Вторые производные равны

$$\frac{\partial^2 F}{\partial x_1^2} = \frac{\partial^2 F}{\partial x_2^2} = 2, \frac{\partial^2 F}{\partial x_1 \partial x_2} = 0.$$

Поэтому квадратическая форма $q(x_1,x_2)=2x_1^2 + 2x_2^2 > 0$, и, следовательно, данная функция в точке $x_1 = x_2 = 0$ имеет минимум.

Пример 1.2. Найти точки экстремума функции

$$F(x_1,x_2)=x_1^2-x_2^2.$$

Решение. Составляем уравнения

$$\frac{\partial F}{\partial x_1} = 2x_1 = 0,$$
$$\frac{\partial F}{\partial x_1} = -2x_2 = 0.$$

Они имеют единственное решение $x_1 = x_2 = 0$. Вторые производные равны

$$\frac{\partial^2 F}{\partial x_1^2} = 2, \frac{\partial^2 F}{\partial x_2^2} = -2, \frac{\partial^2 F}{\partial x_1 \partial x_2} \cancel{= 0.}$$

Отсюда матрица

$$C = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

и ее главные миноры равны $\Delta_1 = 2$, $\Delta_2 = -4$. Поэтому эта матрица не является знакоопределенной, и, следовательно, точка $x_1 = x_2 = 0$ является седловой.

Пример 1.3. Исследовать на экстремум функцию

$$f(x)=f(x_1,\,x_2)=x_1x_2-\frac{1}{2(x_1+x_2)}.$$

Решение.

1. Определим градиент ∇f(x) функции f (x):

$$\nabla f(x)=\nabla f(x_1,\,x_2)=\left(x_2+\frac{1}{2(x_1+x_2)^2},\;x_1+\frac{1}{2(x_1+x_2)^2}\right),$$

а также матрицу Гессе H(x):

$$H(x)=H(x_1,\,x_2)=\begin{pmatrix}-\dfrac{1}{(x_1+x_2)^3} & 1-\dfrac{1}{(x_1+x_2)^3}\\ 1-\dfrac{1}{(x_1+x_2)^3} & -\dfrac{1}{(x_1+x_2)^3}\end{pmatrix}$$

Найдем стационарные точки функции $f(x)=f(x_1,\ x_2)$, в которых может быть экстремум, используя необходимое условие экстремума:

$$\nabla f(x^*)=0\Rightarrow\begin{cases}x_2+\dfrac{1}{2(x_1+x_2)^2}=0,\\ x_1+\dfrac{1}{2(x_1+x_2)^2}=0,\end{cases}\Rightarrow x_1=x_2\Rightarrow$$

$$\Rightarrow x_1+\frac{1}{8x_1^2}=0\Rightarrow x_1^3=-\frac{1}{8}\Rightarrow x_1=-\frac{1}{2}\Rightarrow x_2=-\frac{1}{2}.$$

Нет точек, в которых частные производные функции f(x) не существуют.

Таким образом, функция имеет одну критическую точку

$$x^*=\left(-\frac{1}{2},\;-\frac{1}{2}\right).$$

2. Матрица Гессе $H(x_1,\ x_2)$ в точке

$$x^*=\left(-\frac{1}{2},\;-\frac{1}{2}\right)$$

имеет вид

$$H(x^*) = H\left(-\frac{1}{2},\ -\frac{1}{2}\right) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Поскольку главный минор второго порядка матрицы H(x)

$$\det H_2(x^*) = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3 < 0,$$

то матрица H(x*) не является ни положительно определенной, ни отрицательно определенной, следовательно, экстремума нет в точке

$$x^* = \left(-\frac{1}{2},\ -\frac{1}{2}\right).$$

Пример 1.4. Исследовать на экстремум функцию

$$f(x) = f(x_1,\ x_2) = e^{2x_1}(x_1 + x_2^{\,2} - 2x_2).$$

Решение

1. Определим градиент ∇f(x) функции f (x):

$$\nabla f(x) = \nabla f(x_1,\ x_2) = \left(e^{2x_1}(2x_1 + 2x_2^2 - 4x_2 + 1),\ e^{2x_1}(2x_2 - 2)\right),$$

а также матрицу Гессе H(x):

$$H(x) = H(x_1,\ x_2) = \begin{pmatrix} e^{2x_1}(4x_1 + 4x_2^2 - 8x_2 + 4) & e^{2x_1}(4x_2 - 4) \\ e^{2x_1}(4x_2 - 4) & e^{2x_1}2 \end{pmatrix}$$

Найдем стационарные точки функции f(x)=f(x_1, x_2), в которых может быть экстремум, используя необходимое условие экстремума:

$$\nabla f(x^*) = 0 \Rightarrow \begin{cases} e^{2x_1}(2x_1 + 2x_2^2 - 4x_2 + 1) = 0, \\ e^{2x_1}(2x_2 - 2) = 0. \end{cases}$$

Из второго уравнения

$$x_2 = 1 \Rightarrow 2x_1 - 1 = 0 \Rightarrow x_1 = \frac{1}{2}.$$

Нет точек, в которых частные производные функции f (x) не существуют.

Таким образом, функция имеет одну критическую точку

$$x^* = \left(\frac{1}{2},\ 1\right).$$

2. Матрица Гессе $H(x_1, x_2)$ в точке

$$x^* = \left(\frac{1}{2},\ 1\right)$$

имеет вид

$$H(x^*) = H\left(\frac{1}{2},\ 1\right) = \begin{pmatrix} 2e & 0 \\ 0 & 2e \end{pmatrix}$$

Поскольку главный минор второго порядка матрицы $H(x)$

$$\det H_2(x^*) = \begin{vmatrix} 2e & 0 \\ 0 & 2e \end{vmatrix} = 4e^2 > 0,$$

главный минор первого порядка матрицы $H(x^*)$

$$\det H_1(x^*) = 2e > 0,$$

то матрица $H(x^*)$ является положительно определенной. Следовательно, в точке

$$x^* = \left(\frac{1}{2},\ 1\right)$$

функция $f(x_1, x_2)$, имеет локальный минимум и принимает значение

$$f_{\min}\left(\frac{1}{2}, 1\right) = -\frac{e}{2}.$$

Пример 1.5. Исследовать на экстремум функцию

$$f(x) = f(x_1,\ x_2) = x_1^2 x_2^2 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + x_1 x_2 + 1.$$

Решение

1. Определим градиент $\nabla f(x)$ функции $f(x)$:

$$\nabla f(x) = \nabla f(x_1, x_2) = \left(2x_1 x_2^2 + x_1 + x_2, 2x_2 x_1^2 + x_2 + x_1\right),$$

а также матрицу Гессе H(x):

$$H(x)=H(x_1,\ x_2)=\begin{pmatrix} 2x_2^2+1 & 4x_1x_2+1 \\ 4x_1x_2+1 & 2x_1^2+1 \end{pmatrix}$$

Найдем стационарные точки функции $f(x)=f(x_1,\ x_2)$, в которых может быть экстремум, используя необходимое условие экстремума:

$$\nabla f(x^*)=0\Rightarrow\begin{cases} 2x_1x_2^2+x_1+x_2=0, \\ 2x_2x_1^2+x_2+x_1=0, \end{cases}\Rightarrow x_1=x_2\Rightarrow$$

$$\Rightarrow 2x_2^3+2x_2=0\Rightarrow x_2=0\Rightarrow x_1=0.$$

Нет точек, в которых частные производные функции f (x) не существуют.

Таким образом, функция имеет одну критическую точку x* = (0; 0).

2. Матрица Гессе $H(x_1,\ x_2)$ в точке x* = (0; 0) имеет вид

$$H(x^*)=H(0,\ 0)=\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Поскольку главный минор второго порядка матрицы H(x)

$$\det H_2(x^*)=\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}=0,$$

а главный минор первого порядка матрицы H(x)

$$\det H_1(x^*)=1>0,$$

то матрица H(x*) является неотрицательно определенной, достаточные условия экстремума не выполняются. Требуются дополнительные исследования.

3. Исследуем значения функции $f(x)=f(x_1,\ x_2)$.

$$f(x_1,\ x_2)=x_1^2x_2^2+\frac{x_1^2}{2}+\frac{x_2^2}{2}+x_1x_2+1=(x_1x_2)^2+\frac{1}{2}(x_1+x_2)^2+1,$$

значит, $f(x_1,\ x_2)>1$, если $x_1\neq 0$, $x_2\neq 0$. Получили, что $f(x_1,\ x_2)>f(0,\ 0)$, следовательно, в точке x* = (0; 0) функция $f(x_1,\ x_2)$ имеет локальный минимум и принимает значение

$$f_{\min}(0,\ 0) = 1.$$

1.2. Условный экстремум при ограничениях типа равенств

Постановка задачи. Для заданной функции многих переменных F(x) найти точки экстремума в некоторой области S пространства R^n, которая задается системой уравнений

$$\begin{aligned} g_1(x) &= 0, \\ g_2(x) &= 0, \\ &\dots \\ g_m(x) &= 0. \end{aligned} \qquad (1.1)$$

При этом m <n и предполагается, что система (1) совместная.

Метод исключения переменных

Решая уравнения (1.1), можно m переменных x выразить через остальные n-m переменных x, т.е. найти зависимость x= F(x). Подставляя последнее в функцию F(x), получаем новую функцию F_0(x), зависящую только от переменных x, на которые не накладывается никаких ограничений. Поэтому задача поиска точек экстремума функции F_0(x) может быть решена методом, описанным выше.

Метод множителей Лагранжа

1. Составляется функция Лагранжа

$$L(x, \mu) = F(x) + \sum_{s=1}^{m} \mu_s g_s(x),$$

где μi – некоторые числа, называемые множителями Лагранжа. Затем находится экстремум этой функции.

2. Первое необходимое условие экстремума приводит к системе равенств:

$$
\begin{aligned}
&\frac{\partial L}{\partial x_1} = \frac{\partial F}{\partial x_1} + \sum_{s=1}^{m} \mu_s \frac{\partial g_s(x)}{\partial x_1} = 0,\\
&\frac{\partial L}{\partial x_2} = \frac{\partial F}{\partial x_2} + \sum_{s=1}^{m} \mu_s \frac{\partial g_s(x)}{\partial x_2} = 0,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\frac{\partial L}{\partial x_n} = \frac{\partial F}{\partial x_n} + \sum_{s=1}^{m} \mu_s \frac{\partial g_s(x)}{\partial x_n} = 0
\end{aligned}
\qquad (1.2)
$$

Объединяя (1.1) и (1.2), получаем n+m уравнений для n+m неизвестных x и μ. Эти уравнения могут иметь несколько решений $(x_1, \mu_1),...,(x_k, \mu_k)$ (k=0,1,…). Чтобы решить вопрос, какой экстремум достигается при каждом из этих решений, нужно вычислить вторые производные функции $L(x, \mu)$ и составить квадратическую форму

$$
q(C, \Delta x) = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} \Delta x_i \Delta x_j \ , \qquad (1.3)
$$

где теперь

$$
C_{ij} = \left[\frac{\partial^2 L}{\partial x_i \partial x_j} \right]_{x=\hat{x}_m, \mu=\hat{\mu}_m} = \left[\frac{\partial^2 F}{\partial x_i \partial x_j} + \sum_{s=1}^{m} \mu_s \frac{\partial^2 g_s(x)}{\partial x_i \partial x_j} \right]_{x=\hat{x}_m, \mu=\hat{\mu}_m} \qquad (1.4)
$$

(x_m, μ_m) – исследуемое решение. Последние обозначения означают, что для вычисления элементов Cij нужно вычислить вторые производные функции $L(x, \mu)$ и затем в полученные выражения вместо x и μ подставить x_m и μ_m.В отличие от (1.3) переменные Δxi – здесь уже не произвольные числа, а они должны удовлетворять уравнениям:

$$
\begin{aligned}
&\frac{\partial g_1}{\partial x_1} \Delta x_1 + \frac{\partial g_1}{\partial x_2} \Delta x_2 + \ldots + \frac{\partial g_1}{\partial x_n} \Delta x_n = 0,\\
&\frac{\partial g_2}{\partial x_1} \Delta x_1 + \frac{\partial g_2}{\partial x_2} \Delta x_2 + \ldots + \frac{\partial g_2}{\partial x_n} \Delta x_n = 0,\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\frac{\partial g_m}{\partial x_1} \Delta x_1 + \frac{\partial g_m}{\partial x_2} \Delta x_2 + \ldots + \frac{\partial g_m}{\partial x_n} \Delta x_n = 0.
\end{aligned}
\qquad (1.5)
$$

которые следуют из (1.1).

В этих уравнениях производные $\partial g_i / \partial x_j$ нужно вычислять при $x = x_m$, $\mu = \mu_m$. Так как m<n, то система (1.5) допускает ненулевые решения. Выражая m переменных Δxi через другие n-m переменных Δxi и подставляя их в (1.3), получаем новую квадратическую форму, которую нужно проверять на знакоопределеность.

Пример 1.6. Пусть имеется прямоугольник со сторонами x_1 и x_2 и пусть

$$x_1+x_2=1$$

При каких x_1 и x_2 площадь прямоугольника $S=x_1 \cdot x_2$ максимальна?

Метод исключения переменных

Из условия $x_1+x_2=1$ находим $x_1=1-x_2$. Подставляя в S, получаем

$$S(x_2) = (1 - x_2)x_2 .$$

Вычисляя первую производную функции $S(x_2)$,

$$\frac{\partial S}{\partial x_2} = 1 - 2x_2 = 0 .$$

Отсюда следует, что экстремум площади S достигается при

$$x_1 = x_2 = \frac{1}{2} .$$

Вторая производная функции S равна

$$\frac{\partial^2 S}{\partial x_2^2} = -2 < 0 .$$

Отсюда следует, что площадь S имеет максимум. Таким образом, из всех прямоугольников с заданным периметром наибольшей площадью обладает квадрат.

Метод множителей Лагранжа

Составляется функция Лагранжа

$$L(x_1, x_2, \mu) = x_1 x_2 + \mu(x_1 + x_2 - 1),$$

где μ – множитель Лагранжа. Первое необходимое условие экстремума приводит к системе равенств

$$\frac{\partial L}{\partial x_1} = x_2 + \mu = 0,$$
$$\frac{\partial L}{\partial x_2} = x_1 + \mu = 0.$$

Решая эти уравнения вместе с условием $x_1+x_2=1$, получаем

$$x_1 = x_2 = \frac{1}{2} \qquad \mu = -\frac{1}{2}$$

Вторые производные функции Лагранжа равны

$$\frac{\partial^2 L}{\partial x_1^2} = \frac{\partial^2 L}{\partial x_2^2} = 0, \frac{\partial^2 L}{\partial x_1 \partial x_2} = 1.$$

Поэтому квадратическая форма равна

$$q(\Delta x_1, \Delta x_2) = \Delta x_1 \Delta x_2 .$$

Но согласно (1.5) и условиею $x_1+x_2=1$ величины Δx_1 и Δx_2 должны удовлетворять условию $\Delta x_1+\Delta x_2 = 0$. Отсюда следует, что $\Delta\ x_1=-\Delta x_2$, и, следовательно, $q\ (\Delta x_1, \Delta\ x_2) =- \Delta\ x_2^2\ <0$, т.е. достигается максимум.

Пример 1.7. Пусть имеется цилиндрическая емкость высотой h и радиусом r (см. рис. 1.1). Объем цилиндра равен $V=\pi r^2 h$, а общая поверхность

$$S = 2\pi r^2 + 2\pi rh = 2\pi r(r+h).$$

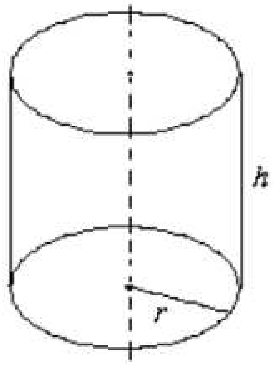

Рис. 1.1. Цилиндрическая емкость

Задача: Пусть поверхность

$$S = 2\pi r(r + h) = S_0 \tag{1.6}$$

задана. Найти такие r и h, при которых объем $V=\pi r^2 h$ максимален.

Решение

Метод исключения переменных

Из (1.6) находим

$$h = \frac{S_0}{2\pi r} - r.$$

Подставляя это в выражение для объема, получаем

$$V = \pi r^2\left(\frac{S_0}{2\pi r} - r\right) = \frac{rS_0}{2} - \pi r^3.$$

Вычисляя первую производную функции V(r), получаем

$$\frac{\partial V}{\partial r} = \frac{S_0}{2} - 3r^2 = 0.$$

Отсюда следует, что экстремум объема V достигается при

$$r = \sqrt{\frac{S_0}{6\pi}}, \quad h = 2r, \tag{1.7}$$

т.е. диаметр цилиндра равен его высоте. Здесь предполагается, что $r>0$. Вторая производная функции V(r) равна

$$\frac{\partial^2 V}{\partial r^2} = -6\pi r < 0,$$

т.е. достигается максимум.

Метод множителей Лагранжа

Составляется функция Лагранжа

$$L(x, \mu) = \pi r^2 h + \mu\left(2\pi r(r + h) - S_0\right),$$

где μ – множитель Лагранжа. Первое необходимое условие экстремума приводит к системе равенств

$$\begin{gathered}\frac{\partial L}{\partial r} = 2\pi rh + 2\pi\mu(2r + h) = 0,\\ \frac{\partial L}{\partial h} = \pi r^2 + 2\pi\mu r = 0.\end{gathered} \tag{1.8}$$

Решая совместно уравнения (1.6) и (1.8), получаем (1.7) и μ=– r /2.

Вычисляя вторые производные функции L(x, μ), получаем с учетом (1.7))

$$\frac{\partial^2 L}{\partial r^2} = 2\pi h + 4\pi\mu = 2\pi r,\ \frac{\partial^2 L}{\partial r \partial h} = 2\pi r + 2\pi\mu = \pi r,\ \frac{\partial^2 L}{\partial r^2} = 0\,.$$

Отсюда квадратическая форма (1.3) равна

$$q(\Delta r, \Delta h) = 2\pi r\left(\Delta r^2 + \Delta r \Delta h\right). \qquad (1.9)$$

Чтобы найти соотношение между Δr и Δh, воспользуемся уравнениями (1.5) и (1.6). Поскольку в данном случае

$$g(r, h) = 2\pi r(r + h) - S_0\,,$$

то из (1.5) имеем с учетом (1.7)

$$\frac{\partial g}{\partial r}\Delta r + \frac{\partial g}{\partial h}\Delta h = 2\pi(2r + h)\Delta r + 2\pi r \Delta h = 8\pi r \Delta r + 2\pi r \Delta h = 0\,.$$

Отсюда получаем, что Δh =–4Δr. Подставляя это в (1.9), получаем окончательно, что q (Δr, Δh) = $-3r\pi\Delta r^2 < 0$, т.е. достигается максимум.

Пример 1.8. Пусть имеется кусок проволоки длиной *l*, который разрезается на два куска длиной x_1 и x_2 соответственно, т.е.

$$x_1 + x_2 = l\,. \qquad (1.10)$$

Из первого куска выгибается квадрат, из второго равносторонний треугольник со сторонами $x_1/4$ и $x_2/3$ соответственно (см. рис. 1.2).

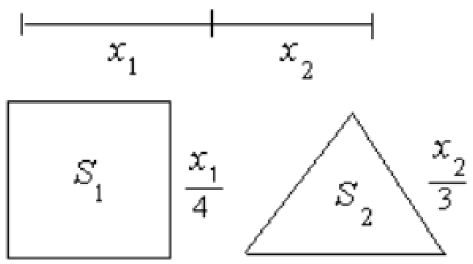

Рис. 1.2. Квадрат и равносторонний треугольник

Задача. Найти x_1 и x_2, при которых суммарная площадь обеих фигур минимальна и максимальна. Используя известные формулы из геометрии, можно подсчитать, что $S_1= c_1x_1^2$ и $S_2 =c_2\, x_2^2$, где

$$c_1=1/16 \text{ и } c_2=\sqrt{3}/36.$$

Главное для дальнейшего то, что c1>c2. В результате общая площадь равна

$$F(x_1, x_2) = c_1x_1^2 + c_2x_2^2. \tag{1.11}$$

Метод исключения переменных

Из (1.10) находим $x_1= l - x_2$. Подставляя это в (1.11), получаем

$$F(x_1) = c_1x_1^2 + c_2(l - x_2)^2,$$

т.е. получаем функцию одной переменной. Первое необходимое условие экстремума приводит к уравнению

$$\frac{\partial F}{\partial x_1} = 2c_1x_1 - 2c_2(l - x_1) = 0.$$

Отсюда получается

$$x_1 = \frac{c_2 l}{c_1 + c_2},\ x_2 = \frac{c_1 l}{c_1 + c_2}. \tag{1.12}$$

Вычисляя вторую производную, получаем

$$\frac{\partial^2 F}{\partial x_1^2} = 2c_1 + 2c_2 > 0,$$

и, следовательно, в точке (1.12) функция $F(x_1, x_2)$ имеет минимум. Проведенное решение не позволяет найти максимум функции $F(x_1, x_2)$. Причина заключается в том, что при решении не учтены требования, чтобы $x_1\geq 0$ и $x_2\geq 0$. Полное решение этой задачи будет получено в следующем разделе. Метод множителей Лагранжа приводить не будем, так как он дает тот же результат. Геометрическое решение задачи легко получить из рис. 1.3, где приведен график функции $F(x_1)$. Видно, что на интервале [0, l] эта функция имеет максимум в угловых точках $x_1=0$ и $x_1=l$. Причем при $x_1=l$ получается глобальный максимум. На рисунке 1.3

$$x_0 = \frac{c_2 l}{c_1 + c_2},$$

т.е. это точка, где достигается минимум.

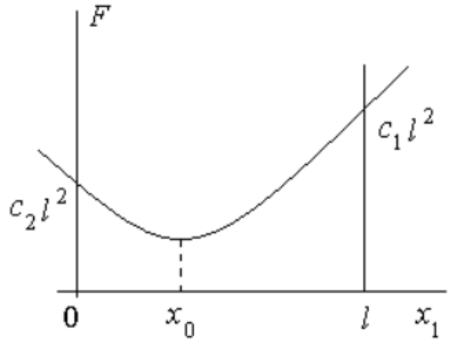

Рис. 1.3. График функции F(x_1)

Пример 1.9. Исследовать на экстремум функцию f(x)=f(x_1, x_2)= x_1^2 + +x_2^2 при условии x_1 + x_2 =1.

Решение. Рассмотрим графический способ решения.

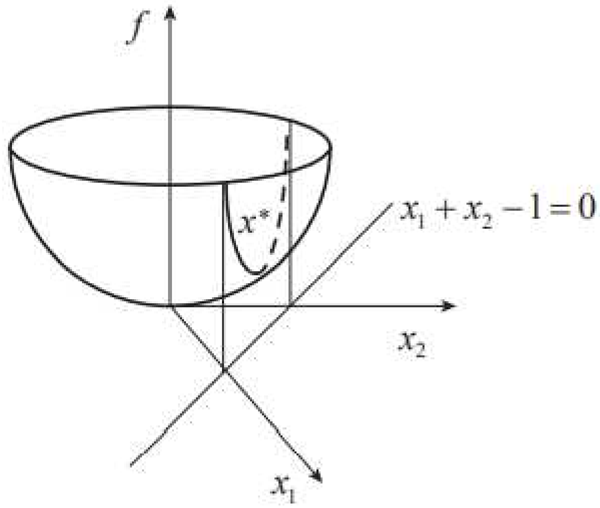

Рис. 1.4. Графическое решение задачи

Точка

$$x^* = \left(\frac{1}{2}, \frac{1}{2}\right)$$

— точка условного минимума функции f(x_1, x_2) на прямой x_1 + x_2 =1.

Рассмотрим аналитический способ решения. Благодаря уравнению связи исключим из функции f(x_1, x_2) переменную x_2 =1 - x_1, что сведет задачу к исследованию функции одной переменной:

$$f(x_1, x_2)\big|_{x_2 = 1 - x_1} = f(x_1) = x_1^2 + (1 - x_1)^2 = 2x_1^2 - 2x_1 + 1.$$

1. Найдем производную функции одной переменной f (x_1):

$$\frac{df(x_1)}{dx_1} = 4x_1 - 2,$$

а также критические точки функции f (x_1):

$$\frac{df(x_1)}{dx_1} = 4x_1 - 2 = 0 \Rightarrow x_1^* = \frac{1}{2}.$$

Таким образом,

$$x^* = \left(\frac{1}{2}, \frac{1}{2}\right)$$

— точка возможного экстремума функции одной переменной.

2. Проверим достаточное условие экстремума функции одной переменной:

$$\frac{d^2 f(x_1^*)}{dx_1^2} = 4 > 0.$$

Следовательно,

$$x_1^* = \frac{1}{2}$$

— точка безусловного минимума для функции f (x_1)=2 x_1^2 -2 x_1 +1.

Соответствующая ей точка

$$x^* = \left(\frac{1}{2}, 1 - \frac{1}{2}\right)$$

— точка условного минимума функции f(x)=f(x_1, x_2)= x_1^2 +x_2^2 при условии x_1 + x_2 =1.

Таким образом, задача нахождения условного экстремума сведена к задаче об отыскании обычного экстремума. Общий метод такого сведения — метод Лагранжа.

Пример 1.10. Фигура X ограничена линиями x_1= 0, x_2 = 0, x_2+x_1^2- 6= 0. Вписать в X прямоугольник наибольшей площади.

Решение. Площадь вписанного в фигуру X прямоугольника S(x):

$$S(x) = S(x_1, x_2) = x_1 x_2,$$

$x_2+x_1^2-6=0$ — условие связи для переменных x_1, x_2.

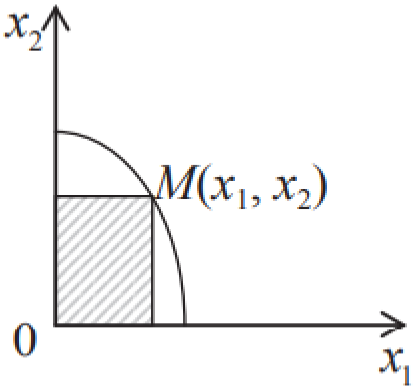

Рис. 1.5. Графическая интерпретация решения задачи

Требуется определить условный экстремум (максимум) функции S(x) при условии $x_2+x_1^2-6=0$. Для этого построим функцию Лагранжа

$$L(x_1,\ x_2,\ \lambda)=x_1x_2+\lambda(x_2+x_1^2-6)$$

и найдем стационарные точки (x_1*, x_2*, λ*) функции L(x_1, x_2, λ):

$$\begin{cases} \dfrac{\partial L(x_1,x_2,\lambda)}{\partial x_1}=x_2+2\lambda x_1, \\ \dfrac{\partial L(x_1,x_2,\lambda)}{\partial x_2}=x_1+\lambda, \\ \dfrac{\partial L(x_1,x_2,\lambda)}{\partial \lambda}=x_2+x_1^2-6, \end{cases} \Rightarrow \begin{cases} x_2+2\lambda x_1=0, \\ x_1+\lambda=0, \\ x_2+x_1^2-6=0. \end{cases}$$

Эта система имеет единственное решение:

$$\begin{cases} x_1^*=\sqrt{2}, \\ x_2^*=4, \\ \lambda^*=-\sqrt{2}. \end{cases}$$

Итак,

$$x^*=(\sqrt{2},\ 4,\ -\sqrt{2})$$

— стационарная точка для функции L(x_1, x_2, λ).

Найдем второй дифференциал d^2L(x_1, x_2, λ) в произвольной точке:

$$d^2L(x_1, x_2, \lambda) = \frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1^{\,2}}(dx_1)^2 +$$

$$+\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_2^{\,2}}(dx_2)^2 + \frac{\partial^2 L(x_1, x_2, \lambda)}{\partial \lambda^2}(d\lambda)^2 +$$

$$+2\left(\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1 \partial x_2}dx_1 dx_2 + \frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1 \partial \lambda}dx_1 d\lambda + \frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_2 \partial \lambda}dx_2 d\lambda\right).$$

Вычислим вторые производные функции Лагранжа $L(x_1, x_2, \lambda)$ в любой точке:

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1^{\,2}} = 2\lambda,$$

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1 \partial x_2} = 1,$$

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_2^{\,2}} = 0,$$

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_1 \partial \lambda} = 2x_1,$$

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial \lambda^2} = 0,$$

$$\frac{\partial^2 L(x_1, x_2, \lambda)}{\partial x_2 \partial \lambda} = 1.$$

С учетом полученных результатов найдем второй дифференциал в стационарной точке

$$x^* = (\sqrt{2},\ 4,\ -\sqrt{2}).$$

$$d^2L(\sqrt{2},\ 4,\ -\sqrt{2}) = -2\sqrt{2}(dx_1)^2 + 2(dx_1 dx_2 + 2\sqrt{2}dx_1 d\lambda + dx_2 d\lambda).$$

Учтем связь x_1 и x_2:

$$x_2 + x_1^{\,2} - 6 = 0,\ dx_2 + 2x_1 dx_1 = 0$$

— в произвольной точке;

в точке

$$(\sqrt{2},\ 4,\ -\sqrt{2}):$$

$$dx_2 = -2\sqrt{2}dx_1 \Rightarrow$$

$$\Rightarrow d^2L(\sqrt{2},\ 4,\ -\sqrt{2}) = -2\sqrt{2}(dx_1)^2 - 4\sqrt{2}(dx_1)^2 +$$
$$+4\sqrt{2}dx_1d\lambda - 4\sqrt{2}dx_1d\lambda = -6\sqrt{2}(dx_1)^2.$$

$$d^2L(\sqrt{2},\ 4,\ -\sqrt{2}) = -6\sqrt{2}(dx_1)^2 < 0 \Rightarrow x^* = (\sqrt{2},\ 4,\ -\sqrt{2})$$

— точка максимума функции Лагранжа.

Таким образом, точка ($\sqrt{2}$; 4) — точка условного максимума функции $S(x_1, x_2)$. Итак, прямоугольник имеет наибольшую площадь при $x_1 = \sqrt{2}$, $x_2 = 4$.

1.3. Условный экстремум при ограничениях типа неравенств

Постановка задачи. Для заданной функции многих переменных F(x) найти точки экстремума в некоторой области S пространства R^n, которая задается системой неравенств

$$\begin{gathered} g_1(x) \le 0, \\ g_2(x) \le 0, \\ \dots \\ g_m(x) \le 0. \end{gathered} \qquad (1.13)$$

При этом предполагается, что система (1.13) совместная.

Решение. С помощью введения новых переменных система (1.13) может быть приведена к системе равенств вида

$$\begin{gathered} g_1(x) + x_{n+1}^2 = 0, \\ g_2(x) + x_{n+2}^2 = 0, \\ \dots\dots \\ g_m(x) + x_{n+m}^2 = 0. \end{gathered} \qquad (1.13)$$

В результате приходим к задаче, решение которой обсуждалось в предыдущей теме.

Пример 1.11. Найти экстремумы функции F(x)=x^2 на интервале [a,b]. Графическое решение приведено на рисунке 1.6.

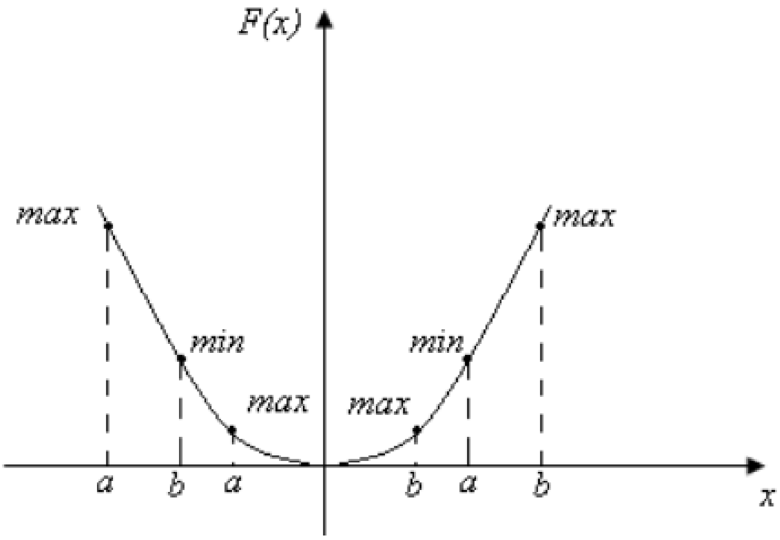

Рис. 1.6. График функции F(x) = x^2

Последние неравенства запишем в виде двух равенств

$$g_1 = x - a - x_1^2 = 0,$$
$$g_2 = b - x - x_2^2 = 0, \quad (1.14)$$

где x_1 и x_2 – новые переменные.

Составим функцию Лагранжа

$$L = x^2 + \mu_1\left(x - a - x_1^2\right) + \mu_2\left(b - x - x_2^2\right)$$

где μ, μ_1 – множители Лагранжа. Условия первого порядка приводят к уравнениям

$$\frac{\partial L}{\partial x} = 2x + \mu_1 - \mu_2 = 0,$$
$$\frac{\partial L}{\partial x_1} = -2\mu_1 x_1 = 0, \quad (1.15)$$
$$\frac{\partial L}{\partial x_2} = -2\mu_2 x_2 = 0.$$

Чтобы решить вопрос об экстремумах функции, вычислим вторые производные функции L(x, μ). Эти производные равны

$$\frac{\partial^2 L}{\partial x^2} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2\mu_1 & 0 \\ 0 & 0 & -2\mu_2 \end{bmatrix}.$$

Поэтому квадратическая форма (1.3) равна

$$q(\Delta x, \Delta x_1, \Delta x_2) = 2\Delta x^2 - 2\mu_1 \Delta x_1^2 - 2\mu_2 \Delta x_2^2. \tag{1.16}$$

При этом величины Δx, Δx_1, Δx_2 согласно (3.5) и (4.3) должны удовлетворять уравнениям

$$\begin{aligned} &\Delta x - 2x_1 \Delta x_1 = 0, \\ &-\Delta x - 2x_2 \Delta x_2 = 0. \end{aligned} \tag{1.17}$$

Решая совместно уравнения (1.14) и (1.15), получаем три варианта решений.

Вариант I. Пусть $x_1=0$. Тогда

$$x = a, \mu_1 \neq 0, x_1 = 0, x_2 = \pm\sqrt{b-a} \neq 0, \mu_2 = 0, \quad \mu_1 = -2a.$$

Далее из (1.17) получаем: $\Delta x=\Delta x_1=0$, $\Delta x_1\neq 0$. Поэтому $q=2a\Delta x_1^2$. В результате получается, что в точках $x=a$ функция имеет максимум при $a<0$ и минимум при $a>0$.

Вариант II. Пусть $x_2=0$. Тогда

$$x = b, \mu_2 \neq 0, x_2 = 0, x_1 = \pm\sqrt{b-a} \neq 0, \mu_1 = 0, \quad \mu_2 = 2b.$$

Далее из (1.17) получаем: $\Delta x=\Delta x_1=0$, $\Delta x_2\neq 0$. Поэтому $q=-2b\Delta x_2^2$. В результате получается, что в точках $x=b$ функция имеет минимум при $b<0$ и максимум при $b>0$.

Вариант III. Пусть $x_1\neq 0$, $x_2\neq 0$. Тогда $x=\mu_1=\mu_2=0$. Поэтому $q=2\Delta x^2>0$ и в точке $x=0$ функция имеет минимум.

Пример 1.12. Рассмотрим пример 1.5. Добавим условия $x_1\geq 0$, $x_2\geq 0$. Эти условия можно преобразовать в равенства, если ввести новые переменные x_3 и x_4:

$$g_1 = x_1 - x_3^2 = 0,$$
$$g_2 = x_2 - x_4^2 = 0. \tag{1.18}$$

Учитывая (1.10), (1.11) и (1.18), составим функцию Лагранжа

$$L = c_1x_1^2 + c_2x_2^2 + \mu(x_1 + x_2 - l) + \mu_1(x_1 - x_3^2) + \mu_2(x_2 - x_4^2),$$

где μ, $\mu 1$, $\mu 2$ – множители Лагранжа. Условия первого порядка приводят к уравнениям

$$\begin{aligned}
&\frac{\partial L}{\partial x_1} = 2c_1x_1 + \mu + \mu_1 = 0,\\
&\frac{\partial L}{\partial x_2} = 2c_2x_2 + \mu + \mu_2 = 0,\\
&\frac{\partial L}{\partial x_3} = -2\mu_1x_3 = 0,\\
&\frac{\partial L}{\partial x_4} = -2\mu_2x_4 = 0.
\end{aligned} \tag{1.19}$$

Вторые производные функции $L(x, \mu)$ равны

$$\frac{\partial^2 L}{\partial x^2} = \begin{vmatrix} 2c_1 & 0 & 0 & 0 \\ 0 & 2c_2 & 0 & 0 \\ 0 & 0 & -2\mu_1 & 0 \\ 0 & 0 & 0 & -2\mu_2 \end{vmatrix}.$$

Поэтому квадратическая форма (1.3) равна

$$q(\Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4) = 2c_1\Delta x_1^2 + 2c_2\Delta x_2^2 - 2\mu_1\Delta x_3^2 - 2\mu_2\Delta x_4^2.$$

При этом величины Δx_1, Δx_2, Δx_3, Δx_4 согласно (1.5) и (1.18) должны удовлетворять уравнениям

$$\begin{aligned}
&\Delta x_1 + \Delta x_2 = 0,\\
&\Delta x_1 - 2x_3\Delta x_3 = 0,\\
&\Delta x_2 - 2x_4\Delta x_4 = 0.
\end{aligned} \tag{1.20}$$

Решая совместно уравнения (1.10), (1.18) и (1.19), получаем три варианта решений.

Вариант I. Пусть x_3=0. Тогда

$$x_1 = 0, x_2 = l,\ x_4 \neq 0, \mu_2 = 0, \mu = -2c_2 l, \mu_1 = 2c_2 l.$$

Далее из (1.20) получаем $\Delta x_1=\Delta x_2=\Delta x_4=0$, $\Delta x_3\neq 0$. Поэтому $q=-2c\Delta x^2<0$. В результате получается, что при $x_1=0$, $x_2=l$ функция имеет максимум.

Вариант II. Пусть x_4=0. Тогда

$$x_1 = l, x_2 = 0,\ x_3 \neq 0, \mu_1 = 0, \mu = -2c_1 l, \mu_2 = 2c_1 l.$$

Далее из (1.20) получаем $\Delta x_1=\Delta x_2=\Delta x_3=0$, $\Delta x_4\neq 0$. Поэтому $q=-2c\Delta x^2<0$. В результате получается, что при $x_1=1$, $x_2=0$ функция имеет максимум.

Вариант III. Пусть $x_3\neq 0$, $x_4\neq 0$. Тогда

$$\mu_1 = \mu_2 = 0, x_1 = \frac{c_2 l}{c_1 + c_2}, x_2 = \frac{c_1 l}{c_1 + c_2}.$$

Поэтому $q=2c\Delta x^2+2c\Delta x^2>0$ и, следовательно, в указанной точке функция имеет минимум. Таким образом, найдено полное решение задачи, которое соответствует рисунку 1.3.

1.4. Условный экстремум в замкнутой области

Рассматривается задача оптимизации целевой функции f(x) на допустимом множестве $X \subset R^n$, где X — непусто и компактно (ограничено и замкнуто):

$$f(x) \to \underset{x \in X}{\text{extr}}.$$

Теорема Вейерштрасса, ее следствия и обобщения для непрерывных (или полунепрерывных) целевых функций определяют условия разрешимости задачи оптимизации. При выполнении этих условий требуется найти решение $x_0 \in X$ задачи (1.4).

Теорема. Если функция f(x): $X \subset R^n \to R$ дифференцируема в замкнутой ограниченной области $X \subset R^n$, то наибольшее

$$\sup_{x \in X} f(x)$$

и наименьшее

$$\inf_{x \in X} f(x)$$

значения функции f (x) на области X достигаются или в критических точках, или на границе области X.

Для нахождения наибольшего и наименьшего значения функции f(x) в области $x \in X$ следует:

— найти критические точки внутри области X, вычислить в них значения функции f (x) ;

— найти наибольшее и наименьшее значения функции f(x) на границе области X;

— сравнить найденные значения и выбрать среди них наибольшее и наименьшее.

Пример 1.13. Найти наибольшее и наименьшее значения функции

$$f(x) = f(x_1, x_2) = e^{x_1^2 + x_2^2}$$

в области X:

$$x_1^2 + x_2^2 \le a^2.$$

Решение

1. Найдем критические точки функции $f(x)=f(x_1, x_2)$ внутри области X.

$$\nabla f(x_1, x_2) = 0 \Rightarrow$$

$$\begin{cases} \dfrac{\partial f(x_1, x_2)}{\partial x_1} = 2x_1 e^{x_1^2 + x_2^2} = 0, \\ \dfrac{\partial f(x_1, x_2)}{\partial x_2} = 2x_2 e^{x_1^2 + x_2^2} = 0, \end{cases} \Rightarrow x_1 = x_2 = 0.$$

Таким образом, функция

$$f(x_1, x_2) = e^{x_1^2 + x_2^2}$$

имеет одну критическую точку $M_1(0; 0)$ внутри области X .

Вычислим значение функции f (x) в точке M_1:

$$f(M_1) = e^{0+0} = 1.$$

2. Найдем наибольшее $f_{\text{наиб}}\big|_{\Gamma}$ и наименьшее $f_{\text{наим}}\big|_{\Gamma}$ значения функции f (x) на границе Г области X .

$$\Gamma : x_1^2 + x_2^2 = a^2.$$

$$f(x)\big|_{x\in\Gamma} = e^{a^2} \Rightarrow f_{\text{наиб}}\big|_{\Gamma} = f_{\text{наим}}\big|_{\Gamma} = e^{a^2}.$$

3. Сравнивая результаты пунктов 1 и 2, получим наибольшее значение функции

$$f_{\text{наиб}} = f_{\text{наиб}}\big|_{\Gamma} = e^{a^2},$$

наименьшее значение функции

$$f_{\text{наим}} = f(0,0) = 1.$$

Пример 1.14. Найти наибольшее и наименьшее значения функции

$$f(x) = f(x_1, x_2) = 2x_1^3 - 6x_1x_2 + 3x_2^2$$

в области X , ограниченной осью Ox_2, прямой $x_2 = 2$ и параболой

$$x_2 = \frac{x_1^2}{2}.$$

Решение

1. Найдем критические точки функции $f(x)=f(x_1, x_2)$ внутри области X:

$$\nabla f(x_1, x_2) = 0 \Rightarrow$$

$$\Rightarrow \begin{cases} \dfrac{\partial f(x_1,x_2)}{\partial x_1} = 6x_1^2 - 6x_2 = 0, \\ \dfrac{\partial f(x_1,x_2)}{\partial x_2} = -6x_1 + 6x_2 = 0, \end{cases} \Rightarrow M_1(0,\ 0),\ M_2(1,\ 1).$$

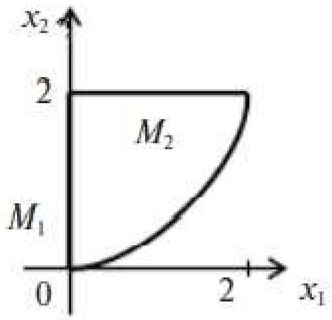

Рис. 1.7. Графическая интерпретация решения задачи

Таким образом, функция

$$f(x_1,\ x_2) = 2x_1^3 - 6x_1x_2 + 3x_2^2$$

имеет одну критическую точку $M_2(1;\ 1)$ внутри области X .

Вычислим значение функции f (x) в точке M_2:

$$f(M_2) = f_1(1,1) = -1.$$

2. Рассмотрим поведение функции на границе области:

$$\text{а) } \begin{cases} x_2 \in [0,\ 2], \\ x_1 = 0, \\ f(0,\ x_2) = 3x_2^2 \text{ — возрастающая функция,} \end{cases}$$

$$f_2(0,\ 0) = 0;\ f_3(0,\ 2) = 12;$$

$$\text{б) } \begin{cases} x_1 \in [0,\ 2], \\ x_2 = 2, \\ f(x_1,\ 2) = 2x_1^3 - 12x_1 + 12, \end{cases}$$

$$\frac{df(x_1,\ 2)}{dx_1} = 6x_1^2 - 12 = 0;\ x_1 = \sqrt{2} \in [0,\ 2].$$

Найдем значения функции $f(x_1,\ 2)$ в точке ($\sqrt{2}$; 2) и в точке (2; 2):

$$f_4(\sqrt{2},\ 2) = 12 - 8\sqrt{2};\ f_5(2,\ 2) = 4;$$

$$
\text{в)}\ \begin{cases} x_2 = \dfrac{x_1^2}{2}, \\ f\left(x_1,\ \dfrac{x_1^2}{2}\right) = \dfrac{3}{4}x_1^4 - x_1^3, \\ x \in [0,\ 2], \end{cases}
$$

$$
\frac{df}{dx} = 3x_1^3 - 3x_1^2 = 0 \Rightarrow x_1^{(1)} = 0,\ x_1^{(2)} = 1.
$$

$$
f_6\left(1,\ \frac{1}{2}\right) = -\frac{1}{4}.
$$

3. Сравнивая полученные значения f_i, i =1, 2, 3, 4, 5, 6, находим наибольшее значение функции

$$
f_{\text{наиб}} = f_3(0,\ 2) = 12,
$$

наименьшее значение функции

$$
f_{\text{наим}} = f_1(1,\ 1) = -1.
$$

Пример 1.15. Найти наибольшее и наименьшее значения функции

$$
f(x) = f(x_1,\ x_2) = x_1x_2 - x_1^2x_2 - \frac{x_1x_2^2}{2} - 3
$$

в области X : $0 \le x_1 \le 1$, $0 \le x_2 \le 2$.

Решение

1. Найдем критические точки функции f(x)=f = (x_1, x_2) внутри области X:

$$
\nabla f(x_1,\ x_2) = 0 \Rightarrow
$$

$$
\Rightarrow \begin{cases} \dfrac{\partial f(x_1,\ x_2)}{\partial x_1} = x_2 - 2x_1x_2 - \dfrac{x_2^2}{2} = 0, \\ \dfrac{\partial f(x_1,\ x_2)}{\partial x_2} = x_1 - x_1^2 - x_1x_2 = 0. \end{cases}
$$

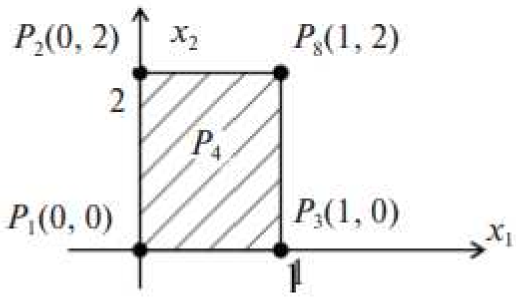

Рис. 1.8. Графическая интерпретация решения задачи

Из первого уравнения:

$$\left[\begin{array}{l} x_2 = 0, \\ x_2 \neq 0 \Rightarrow 1 - 2x_1 - \dfrac{x_2}{2} = 0 \Rightarrow x_2 = 2 - 4x_1. \end{array}\right.$$

Из второго уравнения:

$$\left[\begin{array}{l} x_1 = 0, \\ x_1 \neq 0 \Rightarrow 1 - x_1 - x_2 = 0. \end{array}\right.$$

Следовательно, возможны следующие варианты:

$$P_1(0,\ 0),\ P_2(0,\ 2),\ P_3(1,\ 0),\ P_4\left(\frac{1}{3},\ \frac{2}{3}\right).$$

Таким образом, функция

$$f(x_1,\ x_2) = x_1x_2 - x_1^2x_2 - \frac{x_1x_2^2}{2} - 3$$

имеет одну критическую точку

$$P_4\left(\frac{1}{3},\ \frac{2}{3}\right)$$

внутри области X .

Вычислим значение функции f (x) в точке P_4:

$$f(P_4) = f_1\left(\frac{1}{3},\ \frac{2}{3}\right) = \frac{1}{3}\frac{2}{3} - \frac{1}{9}\frac{2}{3} - \frac{1}{2}\frac{1}{3}\frac{4}{9} - 3 = -\frac{79}{27}.$$

2. Найдем критические точки функции на границе области X :

$$\text{а) }\begin{cases} x_1 = 0, \\ x_2 \in [0,\ 2], \\ f(0,\ x_2) = -3, \end{cases}$$

$$\frac{df(0,\ x_2)}{dx_2} = 0 \Rightarrow \text{критические точки } P_5 = (0,\ 0),\ P_6(0,\ 2);$$

$$\text{б) }\begin{cases} x_1 = 1, \\ x_2 \in [0,\ 2], \\ f(1,\ x_2) = -x_2^2/2 - 3, \end{cases}$$

$$\frac{df(1,\ x_2)}{dx_2} = -x_2 = 0 \Rightarrow x_2 = 0 \Rightarrow P_7(1,\ 0)$$

и, кроме того, граничная точка $P_8(1;\ 2)$.

Найдем значения функции f $(1;\ x_2)$ в точке (1; 0) и в точке (1; 2):

$$f_2(1,\ 0) = -3;\ f_3(1,\ 2) = -5;$$

$$\text{в) }\begin{cases} x_2 = 0, \\ x_1 \in [0,\ 1], \\ f(x_1,\ 0) = -3, \end{cases}$$

$$\frac{df(x_1,\ 0)}{dx_1} = 0 \Rightarrow \text{критические точки } P_9 = (0,\ 0),\ P_{10}(1,\ 0);$$

$$\text{г) }\begin{cases} x_2 = 2, \\ x_1 \in [0,\ 1], \\ f(x_1,\ 2) = -2x_1^2 - 3, \end{cases}$$

$$\frac{df(x_1,\ 2)}{dx_1} = -4x_1 = 0 \Rightarrow x_1 = 0 \Rightarrow \text{критическая точка } P_{11} = (0,\ 2).$$

Найдем значение функции f $(x_1;\ 2)$ 1 2 в точке (0; 2):

$$f_4(0, 2) = -3.$$

3. Сравнивая полученные значения f_i, i =1, 2, 3, 4, находим наибольшее значение функции

$$f_{\text{наиб}} = f_1\left(\frac{1}{3},\ \frac{2}{3}\right) = -\frac{79}{27},$$

наименьшее значение функции

$$f_{\text{наим}} = f_3(1, 2) = -5.$$

Контрольные задания

1. Исследовать на максимум и минимум следующие функции $f(x_1,x_2)$:

а) $f(x) = f(x_1,\ x_2) = -x_1^2 - x_1x_2 - x_2^2 + x_1 + x_2;$

б) $f(x) = f(x_1,\ x_2) = x_1^3 + x_2^3 - 3ax_1x_2;$

в) $f(x) = f(x_1,\ x_2) = e^{2x_1}(x_1 + x_2^2 + 2x_2)$.

2. Найти условные экстремумы функций:

а) $f(x) = f(x_1,\ x_2) = \dfrac{1}{x_1} + \dfrac{1}{x_2}$ при условии $x_1 + x_2 = 2;$

б) $f(x) = f(x_1,\ x_2) = \dfrac{1}{x_1} - \dfrac{3}{x_2}$ при условии $3x_1 - x_2 = 6;$

в) $f(x) = f(x_1,\ x_2) = x_1x_2^2$ при условии $x_1 + 2x_2 = 1$.

3. Найти наибольшее и наименьшее значения функции $f(x_1,x_2)$ в области X:

а) $f(x) = f(x_1,\ x_2) = x_1 - 2x_2 - 3,$
$X: 0 \le x_1 \le 1,\quad 0 \le x_2 \le 1,\quad x_1 + x_2 \le 1;$

б) $f(x) = f(x_1,\ x_2) = x_1^2 + x_2^2 - 12x_1 + 16x_2,\ X: x_1^2 + x_2^2 \le 25;$

в) $f(x) = f(x_1,\ x_2) = x_1x_2,\ X: x_1^2 + x_2^2 \le 1$.

2. Численные методы оптимизации функции

2.1. Методы безусловной минимизации

Рассматривается задача безусловной минимизации целевой функции f(x): $R^n \to R$ (без ограничений):

$$f(x) \to \min_{x \in R^n}.$$

Идея методов приближенного решения поставленной задачи состоит в построении последовательности точек

$$x^{(0)},\ x^{(1)},\ \ldots,\ x^{(k)},\ \ldots,$$

удовлетворяющих условию

$$f(x^{(0)}) \ge f(x^{(1)}) \ge \ldots \ge f(x^{(k)}) \ge \ldots.$$

Такие последовательности { $x^{(k)}$ } называются релаксационными, а методы — методами спуска.

Общая схема метода спуска

Пусть $x^{(0)}$ — начальная точка. Последовательные приближения

$$x^{(1)},\ x^{(2)},\ \ldots,\ x^{(k)},\ \ldots$$

определяют следующим образом:

— в точке $x^{(k)}$ выбирают направление спуска $y^{(k)} \in R^n$;

— находят (k +1)-е приближение по формуле

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} y^{(k)},\ k = 0,\ 1,\ 2,\ \ldots,$$

где $\alpha^{(k)}$ — величина шага спуска.

Все методы спуска различаются:

— либо выбором направления спуска;

— либо способом движения вдоль направления спуска.

Основная задача при выборе параметров $\alpha^{(k)}$, $y^{(k)}$ каждого метода спуска — это обеспечение последовательного убывания значений целевой функции f (x) в точках $x^{(k)}$, k = 0, 1, 2, ...

Теорема 2.1. Пусть функция f(x) дифференцируема в точке $x^* \in R^n$. Тогда любой вектор $y \in R^n$, удовлетворяющий условию

$$(\nabla f(x^*), y) < 0, \tag{2.1}$$

определяет направление убывания функции f (x) в точке x*, то есть существует такое число α > 0, что f (x*+ α y) < f (x*).

Доказательство

Пусть $y \in R^n$ — произвольный вектор, удовлетворяющий условию (2.1). Тогда существует такое число α > 0, что в α-окрестности точки x* для дифференцируемой функции f(x) справедливо разложение (в точке x*)

$$f(x^* + \alpha y) - f(x^*) = (\nabla f(x^*), \alpha y) + r(\alpha y), \tag{2.2}$$

причем

$$\lim_{\alpha \to 0} \frac{r(\alpha y)}{\alpha \|y\|} = 0.$$

Из условия (2.1) следует, что знак

$$f(x^* + \alpha y) - f(x^*)$$

определяется первым слагаемым (в α-окрестности точки x*) в выражении (2.2):

$$f(x^* + \alpha y) - f(x^*) < 0 \Rightarrow f(x^* + \alpha y) < f(x^*).$$

Замечание. Наибыстрейшее убывание функции f(x) в точке x* происходит в направлении вектора

$$y^* = -\nabla f(x^*)$$

— антиградиента функции f (x) в точке x* .

Это свойство антиградиента функции $-\nabla f(x)$ легло в основу градиентных методов.

Определение. Метод, использующий для своей реализации значения f(x), а также ее производных до m-го порядка включительно, называется *методом m-го порядка*. Будем рассматривать наиболее применяющиеся методы:

0–го порядка, использующие только значения f(x);

1–го порядка, использующие значения f(x), а также значения ее первых производных;

2–го порядка, использующие значения f(x), а также значения ее первых и вторых производных.

2.2. Метод покоординатного спуска

В практических задачах нередки случаи, когда минимизируемая функция либо не обладает нужной гладкостью, либо является гладкой, но вычисление производных с нужной точностью является слишком трудоемким. В таких случаях применяют методы минимизации, которые требуют лишь вычисления значений функций, то есть методы 0-го порядка. Из методов этого типа рассмотрим метод покоординатного спуска.

Суть метода состоит в том, что, задав начальное приближение, выбирается направление движения по одной из покоординатных осей, причем, на последующих шагах идет циклический перебор направлений по координатным осям.

Наиболее распространенным является метод покоординатного спуска с дроблением шага.

Обозначим через e_i= (0,...,1,...) единичный координатный (базисный) вектор, у которого i-я координата равна 1, а остальные равны 0.

Пологаем

$$d_k = e_{i_k}\ ;\ i_k = k - n[k/n] + 1,\ [k/n]$$

– целая часть числа k / n . Будем иметь

$$d_0 = e_1,\ \mathrm{d}_1 = \mathrm{e}_2, \ldots, \mathrm{d}_{\mathrm{n}\text{-}1} = \mathrm{e}_{\mathrm{n}}, \ldots$$

Опишем подробно одну итерацию.

Пусть получено x^k. Будем искать x^{k+1}.

Вычислим значение функции f (x) в точке ($x^k + \alpha_k * d_k$) и проверим неравенство

$$f(X_k - \alpha_k * d_k) < f(X^k).$$

В случае выполнения последнего неравенства полагаем

$$X^{k+1} = (X_k - \alpha_k * d_k), \quad \alpha_{k+1} = \alpha_k.$$

В случае невыполнения обоих неравенств полагаем

$$X^{k+1} = X_k, \quad \alpha_{k+1} = \begin{cases} \lambda * \alpha_k & \text{при } i_k = n, 2n, \\ \alpha_k & \text{при } i_k \neq n, 2n, \end{cases}$$

λ – параметр метода; $0 < \lambda < 1$.

Последние условия означают, что если за один цикл из n итераций при переборе направлений всех координатных осей e_1 , e_2 ,..., e_n с шагом α_k реализовалась хотя бы одна удачная итерация, то длина шага α_k не дробится и сохраняется по крайне мере на протяжении следующего цикла из n итераций.

Если же среди последних n итераций не оказалось ни одной удачной, то шаг α_k дробится.

Сходимость метода обеспечена для гладких функций, несмотря на то, что это метод 0-го порядка. Оказывается, что если f (X) не является гладкой, то метод покоординатного спуска может не сходиться к решению.

Другой вариант метода покоординатного спуска может состоять в получении α_k как решения задачи одномерной минимизации:

$$f(X_k + \alpha_k * d_k) \rightarrow \min.$$

Этот вариант имеет смысл применять в том случае, когда α_k можно найти явно.

Хотя скорость сходимости метода покоординатного спуска невысокая, благодаря его простоте и скромным требованиям к гладкости этот метод довольно широко применяется на практике.

На рисунке 2.1 приведена блок-схема метода покоординатного спуска для функции дух переменных с оптимизацией длины шага.

Пример 2.1. Найти минимум функции

$$f(x_1, x_2) = 9 * x_1^2 + x_2^2 - 18 * x_1 + 6 * x_2 + 18$$

Пологаем

$$X^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad X^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha_0 * \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix}.$$

Находим α_0 как решение задачи одномерной минимизации, а именно, ищем α_0, обеспечивающее минимум

$$9 * (\alpha_0 - 1)^2 + 9, \text{ это будет } \alpha_0 = 1.$$

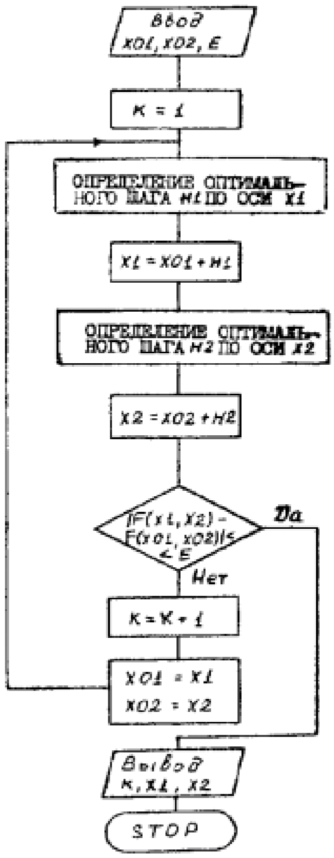

Рис. 2.1. Блок-схема метода покоординатного спуска

Отсюда

$$X^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Далее

$$X^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_1 * \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha_1 \end{pmatrix}; \; f(x^2) = (\alpha_1 + 3)^2 \quad \Rightarrow \alpha_1 = \text{-}3$$

$$X^2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

В результате выполнения 2-й итерации получили точное решение.

2.3. Метод градиентного спуска

Метод градиентного спукска относится к методам 1-го порядка.

Процедура построения итерационной последовательности $\{x^{(k)}\}$ организована следующим образом:

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} y^{(k)},\ k = 0,\ 1,\ 2,\ \ldots,$$

$$y^{(k)} = -\nabla f(x^{(k)}),$$

где величины шага спуска $\alpha^{(k)} > 0$ выбираются достаточно малыми для того, чтобы выполнялось условие последовательного убывания значений целевой функции f (x) в точках $x^{(k)}$, k = 0, 1, 2, ...

Теорема 2.2. Пусть функция f (x) ограничена снизу, дифференцируема на R^n и ее градиент $\nabla f(x)$ удовлетворяет условию Липшица с постоянной L> 0:

$$\|\nabla f(x') - \nabla f(x'')\| \le L\|x' - x''\| \text{ для } \forall x', x'' \in R^n.$$

Тогда для любой начальной точки $x^{(0)}$ в итерационной процедуре:

$$x^{(k+1)} = x^{(k)} + \alpha y^{(k)}, \qquad (2.3)$$

$$y^{(k)} = -\nabla f(x^{(k)}),\ k = 0,\ 1,\ 2,\ \ldots$$

можно выбрать такое число $\alpha > 0$, постоянное для всех k, что

$$\lim_{k\to\infty} \|\nabla f(x^{(k)})\| = 0,$$

причем последовательность $\{x^{(k)}\}$ будет релаксационной.

Замечание. Теорема 2.2 гарантирует сходимость последовательности $\{x^{(k)}\}$:

— либо к значению функции f(x) в некоторой стационарной точке x*;

— либо к точной нижней грани $\inf\{f(x): x \in R^n\}$, при этом не исключается случай, когда ничего определенного сказать о сходимости последовательности $\{x^{(k)}\}$ нельзя.

Теорема 2.3. Пусть функция f(x) удовлетворяет условию теоремы 2.2 и существует постоянная C такая, что множество уровня

$$L(C) = \{x \in R^n : f(x) \le C\}$$

не пусто и компактно. Тогда для любой начальной точки $x^{(0)} \in L(C)$ последовательность $\{x^{(k)}\}$, определенная по формуле (2.3), будет релаксационной, причем каждая ее предельная точка x* удовлетворяет условию $\nabla f(x^*)=0$.

Следствие. Если дополнительно функция f(x) выпукла, то всякая предельная точка последовательности $\{x^{(k)}\}$ будет точкой минимума функции f (x) на R^n.

Приведем анализ метода градиентного спуска:

1. Итерационная процедура спуска (2.3) с постоянным шагом α = const проста, но при реализации требует знания величины шага α, которая на практике определяется подбором. При этом выбор большого шага α может привести к нарушению условия убывания целевой функции

$$f(x^{(k+1)}) = f(x^{(k)} + \alpha y^{(k)}) < f(x^{(k)}). \tag{2.4}$$

С другой стороны, достаточно малый шаг обеспечит выполнение неравенства (2.4), но для достижения требуемой точности потребуется большое количество итераций. Поэтому сначала фиксируют некоторое значение шага спуска α > 0 и производят вычисления по формуле (2.3). При этом если неравенство (2.4) не выполняется, то для текущей итерации величину шага спуска α уменьшают (дробят) до тех пор, пока условие убывания функции (2.4) не будет выполнено, и продолжают вычисления.

2. *Условие окончания процедуры спуска*. Выполняется требование на точность решения задачи минимизации

$$\left\| x^{(k+1)} - x^{(k)} \right\| \le \varepsilon_1, \text{ или}$$

$$\left| f(x^{(k+1)}) - f(x^{(k)}) \right| \le \varepsilon_2,$$

где $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ — заданные заранее числа.

Учитывая условия (2.3), получим

$$\left\| x^{(k+1)} - x^{(k)} \right\| \le \alpha \left\| y^{(k)} \right\| = \alpha \left\| \nabla f(x^{(k)}) \right\|,$$

отсюда следует, что можно использовать в качестве условия окончания итерационной процедуры градиентного спуска неравенства:

$$\left| \frac{\partial f(x^{(k)})}{\partial x_i} \right| \le \varepsilon_3,\ i = 1,\ 2,\ \ldots,\ n, \text{ или}$$

$$\left\| \nabla f(x^{(k)}) \right\| \le \varepsilon_3,\ \varepsilon_3 > 0.$$

3. *Выбор решения*. В качестве приближения для x^* принимается последняя вычисленная точка $x^{(k)}$. Тогда значение $f(x^{(k)})$ приближенно определяет величину $f^* = f(x^*)$.

Пример 2.2

Минимизировать целевую функцию

$$f(x) = f(x_1, x_2) = x_1^{\,2} + 2x_2^{\,2} + e^{x_1 + x_2},\ x \in R^2,$$

методом градиентного спуска, завершив расчет при

$$\left| \frac{\partial f(x^{(k)})}{\partial x_i} \right| \le 0.05,\ i = 1, 2.$$

Решение

Выберем начальное приближение $x^{(0)}=(0;\ 0)$ и величину шага спуска $\alpha=1$, построим последовательность (2.3) (с дроблением шага спуска α), записывая результаты вычислений в таблицу.

Таблица 1

k	$x_1^{(k)}$	$x_2^{(k)}$	$f(x^{(k)})$	$\frac{\partial f(x^{(k)})}{\partial x_1}$	$\frac{\partial f(x^{(k)})}{\partial x_2}$	α	Примечание
0	0	0	1	1	1	1	
	–1	–1	3.145	–	–		Условие (2.4) нару□шено. Уменьшим α в 2 раза
	0	0	1	1	1	0.5	

	–0.5	–0.5	1.118	–	–		Условие (2.4) нару□шено. Уменьшим α в 2 раза
	0	0	1	1	1	0.25	
1	–0.25	–0.25	0.794	0.106	–0.393	0.25	Условие (2.4) выполнено
2	–0.277	–0.152	0.774	0.098	0.045	0.25	Условие (2.4) выполнено
3	–0.301	–0.163	0.772	0.026	0.023	–	Точность достигнута

Следовательно, получим

$$x^* \approx x^{(3)} = (-0.301,\ -0.163),$$

$$f^* \approx f(-0.301,\ -0.163) = 0.772.$$

Пример 2.3

Найти минимум функции

$$f(x_1, x_2) = x_1^2 + 2x_2^2 + e^{x_1 + x_2},$$

завершив вычисления при погрешности $\varepsilon = 0{,}5$, выбрав начальное приближение $X_1(0) = 0$ и $X_2(0) = 0$, коэффициент шага $H = 0{,}1$.

Решение. Необходимые начальные данные приведены в условии задачи. Для вычислений выберем работу с постоянным шагом без коррекции *(H = Const)*.

Результаты вычислений занесем в таблицу, где *Df/Dx1* и *Df/Dx2* - значения частных производных в соответствующих точках.

Таблица 2

П	X_1	X_2	$F(X_1, X_2)$	*Df/D x1*	*Df/D x2*	*|Grad f|*
1	*0,0000*	*0,0000*	*1,0000*	*1,0000*	*1,0000*	*1,4142*
2	*-0,1000*	*-0,1000*	*0,8487*	*0,6187*	*0,4187*	*0,7471*
3	*-0,1619*	*-0,1419*	*0,8045*	*0,4143*	*0,1706*	*0,4480*
4	*-0,2033*	*-0,1589*	*0,7880*	*0,2895*	*0,0604*	*0,2957*
5	*-0,2323*	*-0,1650*	*0,7806*	*0,2077*	*0,0123*	*0,2080*
6	*-0,2530*	*-0,1662*	*0,7768*	*0,1515*	*-0,0072*	*0,1517*
7	*-0,2682*	*-0,1655*	*0,7748*	*0,1118*	*-0,0138*	*0,1126*
8	*-0,2794*	*-0,1641*	*0,7737*	*0,0831*	*-0,0146*	*0,0844*
9	*-0,2877*	*-0,1626*	*0,7731*	*0,0621*	*-0,0131*	*0,0635*
10	*-0,2939*	*-0,1613*	*0,7727*	*0,0466*	*-0,0110*	*0,0479*

В последней точке модуль градиента меньше заданной погрешности, поэтому поиск прекращается.

Итак,

$(x^*; y^*) \approx (-0,2539; -0.1613)$

и

$f^* \approx 0,7727.$

2.4. Метод наискорейшего спуска

Процедура построения последовательности $\{x^{(k)}\}$ организована следующим образом:

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} y^{(k)},\ k = 0,\ 1,\ 2,\ \ldots,$$

$$y^{(k)} = -\nabla f(x^{(k)}).$$

В качестве величины шага спуска $\alpha^{(k)} > 0$ выбирают решение задачи одномерной минимизации:

$$g^{(k)}(\alpha^{(k)}) = \min_{\alpha > 0} g^{(k)}(\alpha),\ g^{(k)}(\alpha) = f(x^{(k)} + \alpha y^{(k)}). \qquad (2.5)$$

Таким образом, в данном методе спуск производится в направлении наибыстрейшего убывания целевой функции и при этом с максимально возможным шагом.

Теорема 2.4. Пусть функция f (x) удовлетворяет условию теоремы 2.2. Тогда для любой начальной точки $x^{(0)}$ метод наискорейшего спуска приводит к построению последовательности $\{x^{(k)}\}$, удовлетворяющей условию

$$\lim_{k \to \infty} \left\| \nabla f(x^{(k)}) \right\| = 0.$$

Если $y^{(k)} \neq 0$, то задача минимизации (2.5) имеет решение лишь при $\alpha^{(k)} > 0$, так как

$$\left. \frac{dg^{(k)}(\alpha)}{d\alpha} \right|_{\alpha=0} = \frac{\partial f(x^{(k)})}{\partial y^{(k)}} = (\nabla f(x^{(k)}), y^{(k)}) = -\left\| y^{(k)} \right\|^2 < 0,$$

то есть при $\alpha^{(k)} > 0$ имеем уменьшение значения функции f (x) . Пусть на k-м шаге метода наискорейшего спуска $\alpha^{(k)} > 0$ — решение задачи (2.5), тогда

$$\left.\frac{dg^{(k)}(\alpha)}{d\alpha}\right|_{\alpha=\alpha^{(k)}} = (\nabla f(x^{(k+1)}), y^{(k)}) = 0 \Rightarrow (y^{(k+1)},\ y^{(k)}) = 0.$$

Откуда следует, что направление спуска на (k +1)-й итерации ортогонально направлению спуска на предыдущей k-й итерации. Таким образом, кривая движения по методу наискорейшего спуска представляет собой ломаную, соседние звенья которой взаимно ортогональны.

Эффективность градиентных методов зависит от вида минимизируемой функции. Метод наискорейшего спуска сойдется за одну итерацию при любом начальном приближении для функции $f(x) = f(x_1, x_2) = x_1^2 + x_2^2$ (см. рисунок ниже). Но сходимость будет очень медленной, например, в случае функции вида $f(x) = f(x_1, x_2) = x_1^2 + 100x_2^2$. В тех ситуациях, когда линии уровня минимизируемой функции представляют собой прямолинейный или, хуже того, криволинейный «овраг», эффективность метода оказывается очень низкой.

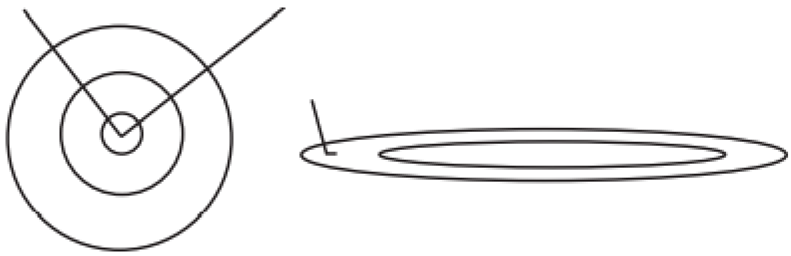

Рис. 2.2. Траектории спуска в зависимости от вида функций

Пример 2.4

Минимизировать целевую функцию

$$f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + e^{x_1+x_2},\ x \in R^2,$$

методом наискорейшего спуска, завершив вычисления при

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 0.05,\ i = 1,\ 2.$$

Решение

Шаг 0. Полагаем начальное приближение $x^{(0)} = (0, 0)$, тогда $\nabla f(x^{(0)}) = (1; 1)$,

$$g^{(0)}(\alpha) = f(0 - 1\alpha,\ 0 - 1\alpha) = 3\alpha^2 + e^{-2\alpha}.$$

Для нахождения точки минимума $a^{(0)}$ функции $g^{(0)}(\alpha)$ используем метод перебора:

Таблица 3

α	0.18	0.20	0.22	0.24	0.26
$g^{(0)}(\alpha)$	0.795	0.790	0.789	0.791	0.797

Таким образом, $a^{(0)} = 0{,}22$, откуда получим

$$x^{(1)} = (0, 0) - 0.22 \cdot (1, 1) = (-0.22,\ -0.22).$$

Шаг 1.

$$\nabla f(x^{(1)}) = (0.204,\ -0.236),$$

$$g^{(1)}(\alpha) = (-0.22 - 0.204\alpha)^2 + (-0.22 + 0.236\alpha)^2 + e^{-0.44+0.032\alpha}.$$

Минимизируем $g^{(1)}(\alpha)$:

Таблица 4

α	0.28	0.30	0.32	0.34	0.36
$g^{(1)}(\alpha)$	0.77401	0.77384	0.77380	0.77387	0.77401

Таким образом, $\alpha^{(1)} = 0{,}32$, откуда получим

$$x^{(2)} = (-0.22,\ -0.22) - 0.32 \cdot (0.204,\ -0.236) = (-0.2853,\ -0.1445).$$

Шаг 2 .

$$\nabla f(x^{(2)}) = (0.08007,\ 0.07268),$$

$$g^{(2)}(\alpha) = (-0.2853 - 0.08007\alpha)^2 + (-0.1445 - 0.07268\alpha)^2 + e^{-0.429-0.15275\alpha}.$$

Минимизируем $g^{(2)}(\alpha)$:

Таблица 5

α	0.20	0.22	0.24	0.26	0.28
g(0)(α)	0.77273	0.77241	0.77240	0.77241	0.77244

Таким образом, $\alpha^{(2} = 0{,}24$, откуда получим

$$x^{(3)} = (-0.3045,\ -0.1619),$$

$$\nabla f(x^{(3)}) = (0.01821,\ -0.02051),$$

то есть требуемая точность достигнута.

Следовательно, получим:

$$x^* \approx x^{(3)} = (-0.305,\ -0.162),\ f^* \approx f(-0.305,\ -0.162) = 0.772.$$

2.5. Метод сопряженных градиентов

Процедура построения последовательности $\{x^{(k)}\}$ к точке минимума функции f (x) организована следующим образом:

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} p^{(k)},\ k = 0,\ 1,\ 2,\ \ldots, \tag{2.6}$$

$$p^{(0)} = \nabla f(x^{(0)}),\ p^{(k)} = \nabla f(x^{(k)}) + \beta^{(k)} p^{(k-1)},\ k = 1,\ 2,\ \ldots. \tag{2.7}$$

В качестве величины шага спуска $\alpha^{(k)} > 0$ выбирают решение задачи одномерной минимизации:

$$g^{(k)}(\alpha^{(k)}) = \min_{\alpha > 0} g^{(k)}(\alpha),\ g^{(k)}(\alpha) = f(x^{(k)} - \alpha p^{(k)}). \tag{2.8}$$

Параметр $\beta^{(k)}$ определяется по формуле

$$\beta^{(k)} = \frac{\left\|\nabla f(x^{(k)})\right\|^2}{\left\|\nabla f(x^{(k-1)})\right\|^2} = \frac{\sum\limits_{i=1}^{n} \dfrac{\partial f(x^{(k)})}{\partial x_i}^2}{\sum\limits_{i=1}^{n} \dfrac{\partial f(x^{(k-1)})}{\partial x_i}^2},\ k = 1,\ 2,\ \ldots$$

Таким образом, метод сопряженных градиентов отличается от метода наискорейшего спуска только выбором направления спуска ($-p^{(k)}$ вместо $y^{(k)} = -\nabla f(x^{(k)})$. При этом $p^{(k)}$ из выражения (2.7) определяется не только антиградиентом $-\nabla f(x^{(k)})$, но и направлением спуска $-p^{(k-1)}$ на предыдущем шаге.

Это позволяет более полно, чем в градиентных методах, рассмотренных выше, учитывать особенности функции f(x) при построении последовательных приближений (2.6) к ее точке минимума.

Теорема 2.5. Пусть функция f(x) ограничена снизу и ее градиент $\nabla f(x)$ удовлетворяет условию Липшица с постоянной $L > 0$:

$$\|\nabla f(x') - \nabla f(x'')\| \le L\|x' - x''\| \text{ для } \forall x', x'' \in R^n.$$

Тогда для последовательности $\{x^{(k)}\}$, построенной по формулам (2.6)–(2.8) справедливо равенство

$$\lim_{k\to\infty}\|\nabla f(x^{(k)})\| = 0.$$

Условием окончания итерационной процедуры метода сопряженных градиентов является выполнение требования на точность решения задачи минимизации:

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le \varepsilon_3,\ i = 1,\ 2,\ \ldots,\ n, \text{ или}$$

$$\|\nabla f(x^{(k)})\| \le \varepsilon_3,\ \varepsilon_3 > 0.$$

Часто для уменьшения влияния накапливающихся погрешностей вычислений через каждые N итераций (2.6) полагают $\beta_{mN} = 0$, $m = 0, 1, \ldots$, то есть производят обновление метода (N — параметр алгоритма).

Пример 2.5

Минимизировать целевую функцию

$$f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2},\ x \in R^2,$$

методом сопряженных градиентов, завершив вычисления при

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 0.03,\ i = 1, 2.$$

Решение

Шаг 0. Полагаем начальное приближение $x^{(0)}= (1; 1)$, тогда

$$\nabla f(x^{(0)}) = (2,4),\ p^{(0)} = \nabla f(x^{(0)}) = (2,4),$$

$$g^{(0)}(\alpha) = f(1-2\alpha,\ 1-4\alpha) = 36\alpha^2 - 20\alpha + 3.$$

Для нахождения точки минимума $\alpha^{(0)}$ функции $g^{(0)}(\alpha)$ используем условие

$$\left.\frac{dg^{(0)}(\alpha)}{d\alpha}\right|_{\alpha=\alpha^{(0)}} = 0, \text{ то есть } \alpha^{(0)} = \frac{5}{18}.$$

Откуда

$$x^{(1)} = (1,1) - \frac{5}{18}\cdot(2,4) = \left(\frac{4}{9},\ -\frac{1}{9}\right).$$

Шаг 1.

$$\nabla f(x^{(1)}) = \left(\frac{8}{9},\ -\frac{4}{9}\right), \beta^{(1)} = \frac{4}{81},$$

$$p^{(1)} = \left(\frac{8}{9},\ -\frac{4}{9}\right) + \frac{4}{81}(2,4) = \left(\frac{80}{81},\ -\frac{20}{81}\right),$$

$$g^{(1)}(\alpha) = \frac{800}{729}\alpha^2 - \frac{80}{81}\alpha + \frac{2}{9}.$$

Минимизируем $g^{(1)}(\alpha)$ и получаем, что

$$\alpha^{(1)} = \frac{9}{20},$$

откуда получаем

$$x^{(2)} = \left(\frac{4}{9},\ -\frac{1}{9}\right) - \frac{9}{20}\left(\frac{80}{81},\ -\frac{20}{81}\right) = (0,0).$$

$$\nabla f(x^{(2)}) = (0,0),$$

требуемая точность достигнута.

Следовательно, получим

$$x^* \approx x^{(2)} = (0,0),\ f^* \approx f(0,0) = 0.$$

Пример 2.6

Минимизировать целевую функцию

$$f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + x_1x_2 - 7x_1 - 7x_2,\ x \in R^2,$$

методом сопряженных градиентов, завершив вычисления при

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 0.05,\ i = 1, 2.$$

Решение

Шаг 0. Полагаем начальное приближение $x^{(0)}$= (0; 0), тогда

$$\nabla f(x^{(0)}) = (-7,\ -7),$$
$$p^{(0)} = \nabla f(x^{(0)}) = (-7,\ -7),$$
$$g^{(0)}(\alpha) = f(0 + 7\alpha,\ 0 + 7\alpha) = 98(2\alpha^2 - \alpha).$$

Для нахождения точки минимума $\alpha^{(0)}$ функции $g^{(0)}(\alpha)$ используем условие

$$\left.\frac{dg^{(0)}(\alpha)}{d\alpha}\right|_{\alpha=\alpha^{(0)}} = 0,$$

то есть

$$\alpha^{(0)} = \frac{1}{4}.$$

Откуда получим

$$x^{(1)} = (0, 0) - \frac{1}{4}\cdot(-7,\ -7) = \left(\frac{7}{4},\ \frac{7}{4}\right).$$

Шаг 1.

$$\nabla f(x^{(1)}) = \left(-\frac{7}{4},\ \frac{7}{4}\right), \beta^{(1)} = \frac{1}{16},$$

$$p^{(1)} = \left(-\frac{7}{4},\ \frac{7}{4}\right) + \frac{1}{16}(-7,\ -7) = \left(-\frac{35}{16},\ \frac{21}{16}\right),$$

$$g^{(1)}(\alpha) = \frac{49}{32}\left(\frac{7}{2}\alpha^2 - 4\alpha - 392\right).$$

Минимизируем $g^{(1)}(\alpha)$ и получаем, что

$$\alpha^{(1)} = \frac{4}{7},$$

откуда получим

$$x^{(2)} = \left(\frac{7}{4},\ \frac{7}{4}\right) - \frac{4}{7}\left(-\frac{35}{16},\ \frac{21}{16}\right) = (3, 1).$$

$$\nabla f(x^{(2)}) = (0, 0),$$

то есть требуемая точность достигнута.

Следовательно, получим

$$x^* \approx x^{(2)} = (3, 1),\ f^* \approx f(3, 1) = -14.$$

2.6. Метод Ньютона

Основными преимуществами этого метода по сравнению с градиентными методами являются более высокая скорость сходимости, сохранение работоспособности при овражном характере минимизируемой функции.

Пусть известно приближение $x^{(k)} \in R^n$. Если целевая функция f (x) дважды непрерывно дифференцируема на R^n, то в окрестности точки $x^{(k)}$ справедливо разложение

$$f(x) = f^k(x) + r(x^{(k)},\ h), \text{ причем } \lim_{\alpha \to 0} \frac{r(x^{(k)},\ h)}{\|h\|^2} = 0,$$

где h= x- $x^{(k)}$, через $f^{(k)}(x)$ обозначена следующая квадратичная функция:

$$f^k(x) = f(x^{(k)}) + (\nabla f(x^{(k)}),\ h) + \frac{1}{2}\left(h,\ H(x^{(k)})h\right).$$

Предположим, что матрица Гессе $H(x^{(k)})$ положительно определенная. В качестве очередной точки $x^{(k+1)}$ выберем точку глобального минимума квадратичной функции $f^{(k)}(x)$:

$$x^{(k+1)} = x^{(k)} - H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)}), \qquad (2.9)$$

причем

$$f^k(x^{(k+1)}) = f(x^{(k)}) - \frac{1}{2}\left(H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)}),\ \nabla f(x^{(k)})\right).$$

Поскольку для любого $y \in R^n$, $y \neq 0$ справедливо $(y, H^{-1} y) > 0$, то

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

Причем

$$f^k(x^{(k)}) = f(x^{(k)}),$$

то есть вектор

$$h = x^{(k+1)} - x^{(k)} = -H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)})$$

определяет направление убывания $f^{(k)}(x)$ в точке $x^{(k)}$, а также направление убывания и целевой функции f(x):

$$f(x^{(k+1)}) - f(x^{(k)}) < 0. \qquad (2.10)$$

Действительно,

$$\nabla f^k(x) = \nabla f(x^{(k)}) + H(x^{(k)})h \Rightarrow (\nabla f^k(x^{(k)}), h^{(k)}) = (\nabla f(x^{(k)}), h^{(k)}),$$

при этом

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

то есть $h^{(k)}$ — направление убывания $f^{(k)}(x)$ в точке $x^{(k)}$:

$$(\nabla f^k(x^{(k)}),\ h^{(k)}) < 0 \Rightarrow (\nabla f(x^{(k)}),\ h^{(k)}) < 0.$$

Неравенство (2.10) будет справедливо лишь для точки $x^{(k+1)}$, близкой к $x^{(k)}$.

Таким образом, формула (2.9) может быть использована для организации итерационного процесса (при некоторых ограничениях, обеспечивающих ее сходимость).

Теорема 2.6. Пусть функция f(x) дважды непрерывно дифференцируема на R^n, x^* — стационарная точка функции f (x), матрица Гессе $H(x^*)$ не вырожденная. Тогда существует такая окрестность точки x*, что для любого начального приближения $x^{(0)}$ из этой окрестности последовательность $\{x^{(k)}\}$, построенная по формуле (2.9):

$$x^{(k+1)} = x^{(k)} - H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)}),\ k = 0, 1, 2,\ \ldots,$$

сходится к x* .

Таким образом, для любого $x^{(0)} \in U(x^*)$ получим последовательность $\{x^{(k)}\}$ (2.9), каждая предельная точка x* которой удовлетворяет условию

$$\nabla f(x^*) = 0.$$

Итак, если начальное приближение $x^{(0)}$ выбрано достаточно близким к точке минимума функции f(x), то последовательность, построенная по методу Ньютона (2.9), будет сходиться к этой точке. Однако, не зная положения точки минимума, такой выбор $x^{(0)}$ практически осуществить невозможно.

Анализ метода Ньютона

Приведем схему анализа метода Ньютона:

1. В общем случае процесс вычисления на каждой итерации матрицы вторых производных $H^{-1}(x^{(k)})$ и ее обращения требует значительных вычислительных затрат, что резко снижает эффективность применения метода Ньютона.

2. Сходимость процесса гарантируется, только если начальная точка $x^{(0)}$ достаточно близка к точке экстремума.

3. Условие окончания итерационной процедуры метода Ньютона — выполнение требования на точность решения задачи минимизации:

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le \varepsilon_3,\ i = 1, 2,\ \ldots,\ n, \text{ или}$$

$$\left\|\nabla f(x^{(k)})\right\| \le \varepsilon_3,\ \varepsilon_3 > 0.$$

4. Выбор решения. В качестве приближения для x* принимается по-

следняя вычисленная точка $x^{(k)}$. Тогда значение $f(x^{(k)})$ приближенно определяет величину $f^* = f(x^*)$.

Пример 2.7

Найти точку минимума целевой функции

$$f(x) = \sqrt{(1+x^2)},$$

$$x \in R.$$

Решение

Шаг 0. Выберем начальное приближение $x^{(0)} = c > 0$.

Вычислим первую и вторую производные функции f (x):

$$f'(x) = \frac{x}{\sqrt{(1+x^2)}},$$

$$f''(x) = \frac{1}{(1+x^2)^{3/2}} \Rightarrow (f''(x))^{-1} = (1+x^2)^{3/2}.$$

Шаг 1. По формуле (2.9) вычислим $x^{(1)}$:

$$x^{(1)} = x^{(0)} - (f''(x^{(0)}))^{-1} f'(x^{(0)}) =$$

$$= c - (1+c^2)^{3/2} \frac{c}{\sqrt{(1+c^2)}} = c - c(1+c^2) = -c^3.$$

Шаг 2. Вычислим $x^{(2)}$:

$$x^{(2)} = x^{(1)} - (f''(x^{(1)}))^{-1} f'(x^{(1)}) =$$

$$= -c^3 - (1+(-c^3)^2)^{3/2} \frac{-c^3}{\sqrt{(1+(-c^3)^2)}} =$$

$$= -c^3 + c^3(1+c^6) = c^9.$$

В итоге при соответствующем выборе начального приближения $x^{(0)}$:

— при $\left|x^{(0)}\right| < 1 \Rightarrow x^{(k)} \to x_{\min} = 0;$

— при $\left|x^{(0)}\right| = 1 \Rightarrow x^{(k)} = (-1)^k;$

— при $\left|x^{(0)}\right| > 1 \Rightarrow \left|x^{(k)}\right| \to \infty.$

Таким образом, в данном примере использование метода Ньютона при выборе начального приближения $x^{(0)} = c > 0$ приводит к построению последовательности $\{x^{(k)}\}$, удаляющейся от точки минимума.

Пример 2.8

Найти точку минимума целевой функции

$$f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + e^{x_1 + x_2},\ x \in R^2,$$

методом Ньютона, выбрав в качестве начального приближения $x^{(0)}$=(0,3012259; - 0,1629096), с точностью

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 10^{-5}, i = 1, 2.$$

Решение

Шаг 0. Начальное приближение $x^{(0)}$=(0,3012259; - 0,1629096), тогда

$$\nabla f(x^{(0)}) = (0.02622655,\ -0.02296005),$$

$$H(x^{(0)}) = \begin{pmatrix} 0.39319151 & 0.62867835 \\ 0.62867835 & 0.22329787 \end{pmatrix}$$

Найдем обратную матрицу $H^{-1}(x^{(0)})$:

$$H^{-1}(x^{(0)}) = \begin{pmatrix} 0.39319151 & -0.053404226 \\ -0.053404226 & 0.22329787 \end{pmatrix}$$

Шаг 1. По формуле (2.9) вычислим $x^{(1)}$:

$$x^{(1)} = x^{(0)} - H^{-1}(x^{(0)}) \cdot \nabla f(x^{(0)}) =$$

$$= (-0.3012259,\ -0.1629096) - \begin{pmatrix} 0.39319151 & -0.053404226 \\ -0.053404226 & 0.22329787 \end{pmatrix} \times$$

$$\times (0.02622655,\ -0.02296005) = (-0.3127641,\ -0.1563821).$$

$$\nabla f(x^{(1)}) = (7.9 \cdot 10^{-6},\ 7.9 \cdot 10^{-6}),$$

то есть требуемая точность достигнута.

Следовательно, получим

$$x^* \approx x^{(1)} = (-0.3127641,\ -0.1563821).$$

Метод Ньютона — Рафсона — это модифицированный метод Ньютона, позволяющий устранить его недостаток, как возможность несходимости последовательности { $x^{(k)}$} (2.9) к точке x* .

В методе Ньютона — Рафсона последовательность { $x^{(k)}$} строится по формуле

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} H^{-1}(x^{(k)}) \nabla f(x^{(k)}),\ k = 0,\ 1,\ 2,\ \ldots \qquad (2.11)$$

В качестве величины шага спуска $\alpha^{(k)} > 0$ выбирают решение задачи одномерной минимизации

$$g^{(k)}(\alpha^{(k)}) = \min_{\alpha>0} g^{(k)}(\alpha),$$
$$g^{(k)}(\alpha) = f\left(x^{(k)} - \alpha H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)})\right). \qquad (2.12)$$

Теорема 2.7. Пусть функция f(x) дважды непрерывно дифференцируема на R^n, существуют такие числа $M \geq m > 0$, что выполняется неравенство:

$$m\|s\|^2 \leq (s, H(x)s) \leq M\|s\|^2 \text{ для } \forall x, s \in R^n,$$

где H(x) — матрица Гессе. Тогда последовательность $\{x^{(k)}\}$, построенная по формулам (2.11), (2.12),

$$x^{(k+1)} = x^{(k)} - H^{-1}(x^{(k)}) \cdot \nabla f(x^{(k)}),\ k = 0,\ 1,\ 2,\ \ldots,$$

сходится к точке минимума x* независимо от выбора начального приближения $x^{(0)}$.

В качестве другой модификации метода Ньютона можно применять следующую итерационную схему:

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} H^{-1}(x^{(0)}) \cdot \nabla f(x^{(k)}),\ k = 0,\ 1,\ 2,\ \ldots, \qquad (2.13)$$

где $\alpha^{(k)} > 0$ — решение задачи одномерной минимизации:

$$g^{(k)}(\alpha^{(k)}) = \min_{\alpha>0} g^{(k)}(\alpha),\ g^{(k)}(\alpha) = f\left(x^{(k)} - \alpha H^{-1}(x^{(0)}) \cdot \nabla f(x^{(k)})\right).$$

В этом случае для построения направления спуска используется один раз вычисленная и обращенная матрица вторых производных $H(x^{(0)})$. Если матрица $H(x^{(0)})$ положительно определенная, то итерационный процесс (2.13) является модификацией градиентного спуска и сходится независимо от выбора начального приближения $x^{(0)}$.

Контрольные задания

1. Совершить один шаг градиентного спуска из точки x(0) с шагом α и сравнить значения f(x(0)) и f (x(1)):

а) $f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + e^{x_1+x_2},\ x^{(0)} = (1,1),\ \alpha = 0.1;$

б) $f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + e^{x_1+x_2},\ x^{(0)} = (1,1),\ \alpha = 0.265;$

в) $f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 2x_2^{\ 2} + e^{x_1+x_2},\ x^{(0)} = (1,1),\ \alpha = 0.5;$

2. Минимизировать функции f (x)= f(x_1, x_2) методом наискорейшего спуска, заканчивая вычисления при

$$\left|\frac{\partial f(x^{(k)})}{\partial x_i}\right| \le 0.01,\ i = 1, 2:$$

а) $f(x) = f(x_1,\ x_2) = 7x_1^{\ 2} + 2x_1x_2 + 5x_2^{\ 2} + x_1 - 10x_2;$

б) $f(x) = f(x_1,\ x_2) = 3x_1^{\ 2} - 3x_1x_2 + 4x_2^{\ 2} - 2x_1 + x_2;$

в) $f(x) = f(x_1,\ x_2) = x_1^{\ 2} + 4x_1x_2 + 17x_2^{\ 2} + 5x_2;$

Библиографический список.

1. Измайлов А.Ф. Численные методы оптимизации. М.: Физико-математическая литература, 2005. - 300 с.

2. Черноруцкий И.Г. Методы оптимизации в теории управления. СПб.: Питер, 2004. - 255 с.

3. Пантелеев А.В. Методы оптимизации в примерах и задачах. Высш. шк. М.: 2005. - 544 с.

4. Аттетков А.В. Методы оптимизации: учеб. для вузов. М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. - 440 с.

5. Струченков В.И. Методы оптимизации в прикладных задачах. М.: СОЛОН□ПРЕСС, 2009. – 320 с.

6. Параев Ю.И., Панасенко Е.А. Методы оптимизации – Томск: Изд-во ТУСУР, 2012, – с. 20.

Электронное учебное издание

Ольга Викторовна **Свиридова**
Александр Александрович **Рыбанов**

Методы многомерной оптимизации

Учебное пособие

Электронное издание сетевого распространения

Редактор Матвеева Н.И.

Темплан 2021 г. Поз. № 1.
Подписано к использованию 26.04.2021. Формат 60x84 1/16.
Гарнитура Times. Усл. печ. л. 3,75.

Волгоградский государственный технический университет.
400005, г. Волгоград, пр. Ленина, 28, корп. 1.

ВПИ (филиал) ВолгГТУ.
404121, г. Волжский, ул. Энгельса, 42а.