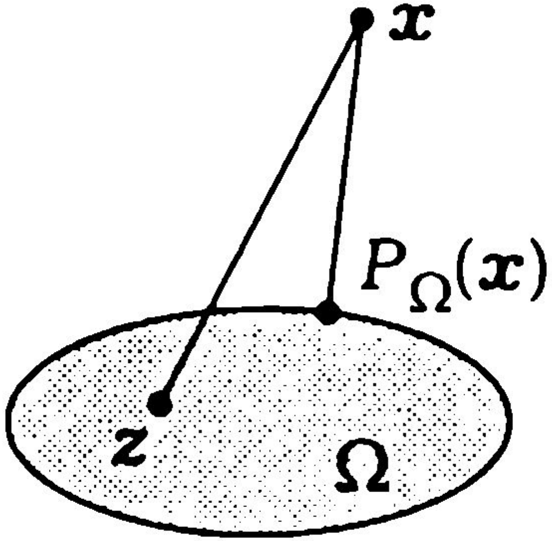
Проектирование точки на множество

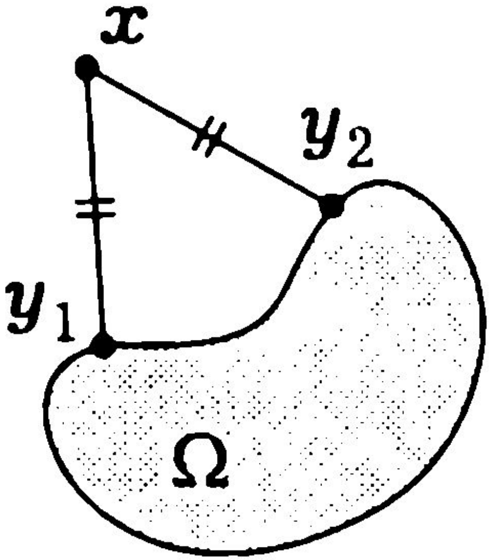
Известно , что непосредственное применение к численному решению общей задачи нелинейного программирования одного из методов спуска может привести к тому, что очередная точка итерационной последовательности выйдет за пределы допустимого множества. Чтобы этого избежать, необходимо корректировать точку, полученную в результате выполнения очередной итерации. Одним из способов такой корректировки является операция проектирования точки на допустимое множество.

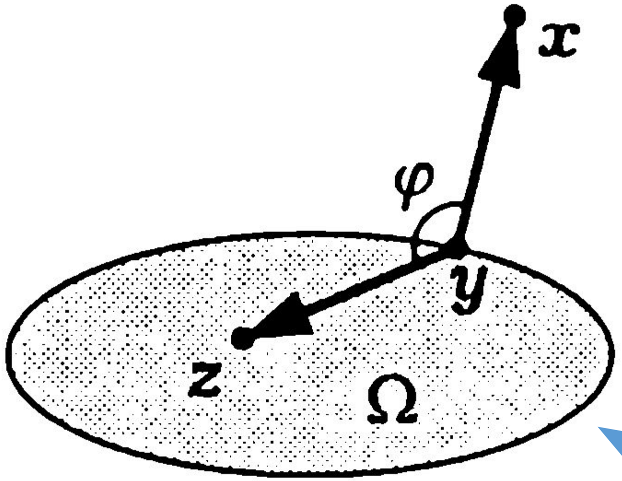
Проекцией точки $\boldsymbol{x} \in \mathbb{R}^n$ ***на множество*** $\Omega \subset \mathbb{R}^n$ называют ближайшую к $\boldsymbol{x}$ точку $\boldsymbol{y}$ этого множества и обозначают ее $P_\Omega(\boldsymbol{x})$. Для точки $\boldsymbol{y} = P_\Omega(\boldsymbol{x})$ выполняются равенства

$$|\boldsymbol{y} - \boldsymbol{x}| = \inf_{\boldsymbol{z} \in \Omega} |\boldsymbol{z} - \boldsymbol{x}| = \rho(\boldsymbol{x}, \Omega). \qquad (5.17)$$

Величину $\rho(\boldsymbol{x}, \Omega)$ называют ***расстоянием от точки*** $\boldsymbol{x}$ ***до множества*** Ω.







Свойства операции проектирования точки на множество.

1. Если $x \in \Omega$, то $P_\Omega(x) = x$.

2. Если Ω — замкнутое множество, то любая точка $x \in \mathbb{R}^n$ имеет проекцию на это множество. Если при этом Ω — выпуклое множество, то проекция любой точки $x \in \mathbb{R}^n$ на Ω единственна.

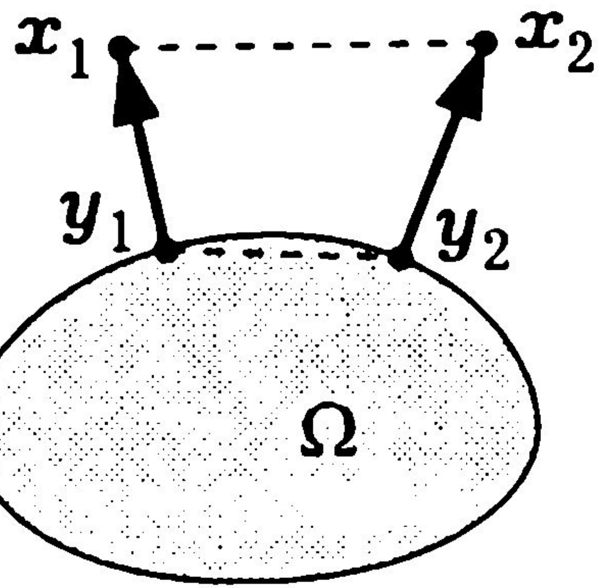
3. Точка y является проекцией точки $x \in \mathbb{R}^n$ на замкнутое выпуклое множество Ω тогда и только тогда, когда для любой точки $z \in \Omega$ выполнено неравенство

$$(z - y, x - y) \leqslant 0. \tag{5.18}$$

4. Если Ω — выпуклое замкнутое множество, $y_1 = P_\Omega(x_1)$ и $y_2 = P_\Omega(x_2)$ — проекции точек x_1 и x_2 на Ω, то

$$|y_1 - y_2| \leqslant |x_1 - x_2|,$$

т.е. длина проекции отрезка на выпуклое множество не превосходит длины самого отрезка



Непосредственно из определения следует, что проекцию заданной точки $x \in \mathbb{R}^n$ на замкнутое множество $\Omega \subset \mathbb{R}^n$ можно найти путем решения задачи минимизации

$$\Phi(z) = |z - x|^2 \to \min, \quad z \in \Omega. \tag{5.19}$$

Согласно свойству 2 эта задача имеет решение, а при дополнительном условии выпуклости Ω решение задачи единственно. Трудность решения задачи связана главным образом с видом множества Ω и способом его описания.

Пример 5.3. Если $\Omega = \{z \in \mathbb{R}^n\colon |z - x_0| \leqslant R\}$ — замкнутый шар с центром в точке $x_0 \in \mathbb{R}^n$ и радиусом R, то из геометрических соображений нетрудно понять, что проекция точки $x \notin \Omega$ на Ω расположена на отрезке, соединяющем точку x с центром шара x_0 (рис. 5.9). Отсюда несложно получить

$$P_\Omega(x) = x_0 + \frac{x - x_0}{|x - x_0|} R. \tag{5.20}$$

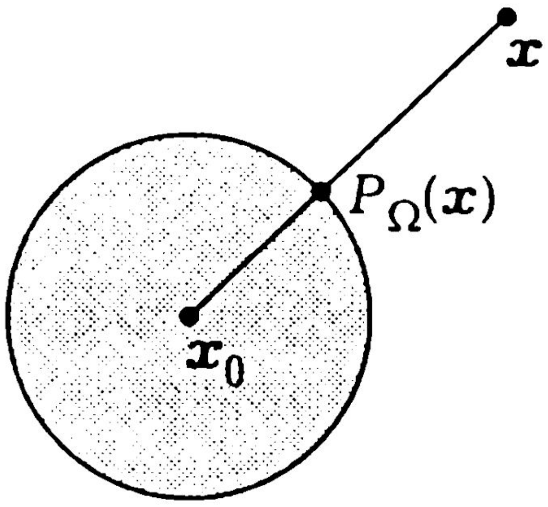


Рис. 5.9

Множество $\widetilde{\Omega} = \{z \in \mathbb{R}^n\colon |z - x_0| = R\}$ точек, равноудаленных от точки x_0 (называемое $(n+1)$-мерной сферой), является замкнутым, но не выпуклым. Для точки x_0 любая точка множества является проекцией. Для точек $x \notin \widetilde{\Omega}$ проекция на сферу единственна и может быть найдена по формуле (5.20).

П р и м е р 5.4. Гиперплоскость

$$\Omega = \{\boldsymbol{z} \in \mathbb{R}^n \colon (\boldsymbol{n}, \boldsymbol{z}) = b\}, \tag{5.21}$$

где $\boldsymbol{n} \in \mathbb{R}^n$ — нормальный вектор этой гиперплоскости, а b — постоянное число, является замкнутым выпуклым множеством. Проекцию точки $\boldsymbol{x} \notin \Omega$ на множество Ω, как и в предыдущем примере, можно найти из геометрических соображений. Из этих соображений нетрудно понять, что вектор $\boldsymbol{y} - \boldsymbol{x}$, где $\boldsymbol{y} = P_\Omega(\boldsymbol{x})$, коллинеарен вектору $\boldsymbol{n}$, а поэтому $\boldsymbol{y} = \boldsymbol{x} + \lambda \boldsymbol{n}$, где λ — некоторое пока не известное число. Подставляя это представление в уравнение гиперплоскости, получаем $(\boldsymbol{n}, \boldsymbol{y}) = (\boldsymbol{n}, \boldsymbol{x}) + \lambda \boldsymbol{n}^2 = b$. Отсюда легко определить значение λ, и в результате мы получаем

$$\boldsymbol{y} = P_\Omega(\boldsymbol{x}) = \boldsymbol{x} + \frac{b - (\boldsymbol{n}, \boldsymbol{x})}{\boldsymbol{n}^2} \boldsymbol{n}. \tag{5.22}$$

Найденное решение подтверждается с помощью теоремы Пифагора. Для любой точки $\boldsymbol{z} \in \Omega$ имеем $(\boldsymbol{n}, \boldsymbol{z}) = (\boldsymbol{n}, \boldsymbol{y})$, откуда $(\boldsymbol{n}, \boldsymbol{z} - \boldsymbol{y}) = 0$, т.е. векторы $\boldsymbol{n}$ и $\boldsymbol{z} - \boldsymbol{y}$, а значит, и векторы $\boldsymbol{y} - \boldsymbol{x}$ и $\boldsymbol{z} - \boldsymbol{y}$, ортогональны. По теореме Пифагора

$$|\boldsymbol{z} - \boldsymbol{x}|^2 = |(\boldsymbol{z} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{x})|^2 = |\boldsymbol{z} - \boldsymbol{y}|^2 + |\boldsymbol{y} - \boldsymbol{x}|^2 \geqslant |\boldsymbol{y} - \boldsymbol{x}|^2.$$

Следовательно, точка $\boldsymbol{y} = P_\Omega(\boldsymbol{x})$ оказывается ближайшей к точке $\boldsymbol{x}$.

Несложно проверить, что проекция точки x на полупространство

$$\Omega^* = \{\boldsymbol{z} \in \mathbb{R}^n \colon (\boldsymbol{n}, \boldsymbol{z}) \leqslant b\} \tag{5.23}$$

(при условии $\boldsymbol{x} \in \Omega^*$) определяется той же формулой (5.22).

Пример 5.6. Рассмотрим n-мерный параллелепипед

$$\Omega = \left\{\boldsymbol{x} \in \mathbb{R}^n \colon a_j \leqslant x_j \leqslant b_j,\ j = \overline{1, n}\right\},$$

где a_j и b_j — заданные числа. Координаты y_j проекции $\boldsymbol{y} = P_\Omega(\boldsymbol{x})$ точки $\boldsymbol{x} = (x_1, \ldots, x_n)$ на множество Ω можно найти по формулам $y_j = \min\left\{\max\left\{x_j, a_j\right\}, b_j\right\}$, $j = \overline{1, m}$, или в развернутой форме

$$y_j = \begin{cases} a_j, & x_j \leqslant a_j; \\ x_j, & a_j < x_j < b_j; \\ b_j, & x_j \geqslant b_j, \end{cases} \qquad j = \overline{1, n}.$$

Метод проекции точки на множество

Сочетание метода градиентного спуска, используемого при безусловной минимизации, и операции проектирования точки на множество составляет существо ***метода проекции точки на множество***

В методе проекции точки на множество релаксационную последовательность $\{x^k\}$ строят с помощью рекуррентного соотношения

$$x^k = P_\Omega(x^{k-1} + \varkappa_k w^k), \tag{5.25}$$

сперва определяя точку $\widetilde{x}^k = x^{k-1} + \varkappa_k w^k$ каким-либо методом градиентного спуска, а затем проектируя ее на допустимое множество. Здесь $w^k = -\operatorname{grad} f_0(x^{k-1})$ — антиградиент целевой функции, вычисленный в точке $x^{k-1} \in \Omega$. Процесс построения релаксационной последовательности начинается с выбора начальной точки $x^0 \in \Omega$.

Если Ω — замкнутое выпуклое множество, процедуры градиентного спуска и последующего проектирования на допустимое множество можно объединить в одной задаче минимизации. Действительно, рассмотрим минимизацию на множестве Ω квадратичной функции

$$\Phi_k(\boldsymbol{z}) = \frac{|\boldsymbol{z} - \boldsymbol{x}^{k-1}|^2}{2} - \varkappa_k(\boldsymbol{z} - \boldsymbol{x}^{k-1}, \boldsymbol{w}^k), \quad \boldsymbol{z} \in \Omega.$$

Эта функция выпуклая как сумма двух выпуклых функций. Она достигает минимума в точке $\boldsymbol{x}^k \in \Omega$, в которой неравенство

$$(\operatorname{grad} \Phi_k(\boldsymbol{x}^k), \boldsymbol{z} - \boldsymbol{x}^k) \geqslant 0$$

выполнено для любой точки $\boldsymbol{z} \in \Omega$. В данном случае $\operatorname{grad} \Phi_k(\boldsymbol{z}) = \boldsymbol{z} - \boldsymbol{x}^{k-1} - \varkappa_k \boldsymbol{w}^k$, поэтому

$$0 \leqslant (\operatorname{grad} \Phi_k(\boldsymbol{x}^k), \boldsymbol{z} - \boldsymbol{x}^k) = (\boldsymbol{x}^k - \boldsymbol{x}^{k-1} - \varkappa_k \boldsymbol{w}^k, \boldsymbol{z} - \boldsymbol{x}^k) = \\ = -(\widetilde{\boldsymbol{x}}^k - \boldsymbol{x}^k, \boldsymbol{z} - \boldsymbol{x}^k).$$

Сравнивая неравенство $(\widetilde{\boldsymbol{x}}^k - \boldsymbol{x}^k, \boldsymbol{z} - \boldsymbol{x}^k) \leqslant 0$ с условием (5.18), убеждаемся, что точка $\boldsymbol{x}^k$ минимума функции $\Phi_k(\boldsymbol{z})$ на множестве Ω является проекцией точки $\widetilde{\boldsymbol{x}}^k = \boldsymbol{x}^{k-1} + \varkappa_k \boldsymbol{w}^k$ на это множество.

Опишем подробнее последовательность действий при выполнении k-й итерации варианта этого метода, в котором используется исчерпывающий спуск в направлении антиградиента. Предположим, что допустимое множество Ω задано ограничением $g(\boldsymbol{x}) = 0$ типа равенства, где $g(\boldsymbol{x})$ — функция, дифференцируемая в $\mathbb{R}^n$. Графически выполнение k-й итерации в двумерном случае показано на рис. 5.10. Здесь Ω есть множество точек плоской гладкой кривой Γ, заданной уравнением $g(\boldsymbol{x}) = 0$.

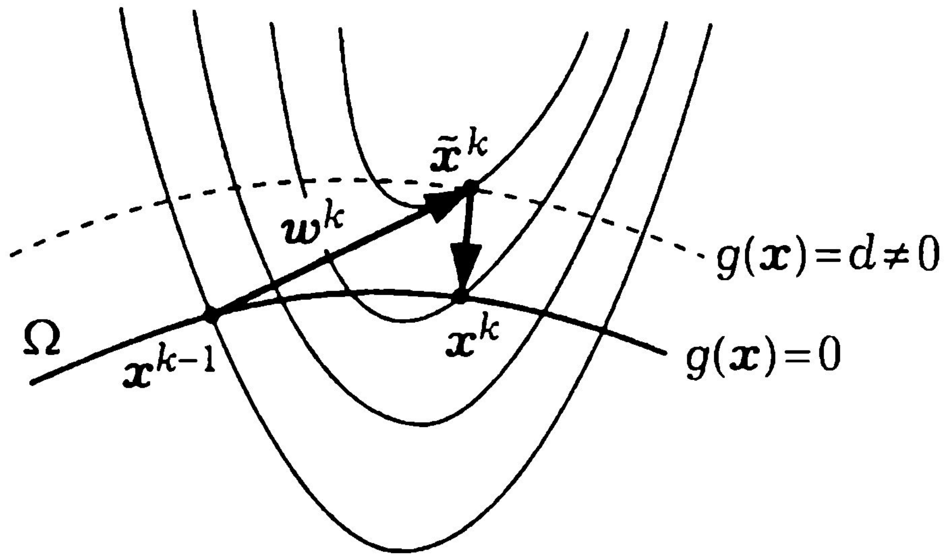


Рис. 5.10

Пример 5.7. Рассмотрим задачу нелинейного программирования

$$f_0(x_1, x_2) = 10x_1^2 - 4x_1x_2 + 7x_2^2 - 4\sqrt{5}(5x_1 - x_2) - 16 \to \min,$$

$$x_1 - x_2 - \sqrt{5} = 0.$$

В качестве начальной точки выберем $\boldsymbol{x}^0 = \left(0,\ -\sqrt{5}\right)$. Проектирование точки $\widetilde{\boldsymbol{x}}^k$ на множество Ω выполняем по формуле (5.22), где $b = \sqrt{5}$ и $\boldsymbol{n} = (1,\ -1)$ — нормальный вектор прямой $x_1 - x_2 - \sqrt{5} = 0$. Условием останова выберем выполнение неравенства $|\boldsymbol{x}^k - \boldsymbol{x}^{k-1}| < \varepsilon_1$, где $\varepsilon_1 = 0{,}01$, а в качестве приближенного решения задачи — последнюю точку релаксационной последовательности: $\boldsymbol{x}^* \approx \boldsymbol{x}^k$ и $f(\boldsymbol{x}^*) \approx f(\boldsymbol{x}^k)$.

Точное решение рассматриваемой задачи: $\boldsymbol{x}^* = \left(\sqrt{5},\ 0\right)$ и $f(\boldsymbol{x}^*) = -66$, где $\sqrt{5} \approx 2{,}2361$. При численном решении задачи используем три варианта выбора значения $\varkappa_k$ в рекуррентном соотношении

$$\boldsymbol{y} = P_\Omega(\boldsymbol{x}) = \boldsymbol{x} + \frac{b - (\boldsymbol{n}, \boldsymbol{x})}{\boldsymbol{n}^2}\boldsymbol{n}$$

$$x^k = P_\Omega(x^{k-1} + \varkappa_k w^k)$$

Таблица 5.2

$\varkappa_k$	N	x^N
$\varkappa = 0{,}10$	6	$(2{,}234,\ -0{,}002)$
$\varkappa = 0{,}05$	7	$(2{,}235,\ -0{,}001)$
$\widetilde{\varkappa}_k$	4	$(2{,}236,\ 0{,}000)$

(5.25): первый и второй варианты — постоянные значения $\varkappa = 0{,}1$ и $\varkappa = 0{,}05$; третий — выбор $\varkappa_k$ с помощью исчерпывающего спуска. Результаты численного решения представлены табл. 5.2. Траектории спуска в трех вариантах метода показаны на рис. 5.11: *а* — $\varkappa = 0{,}1$; *б* — $\varkappa = 0{,}05$; *в* — $\varkappa = \widetilde{\varkappa}_k$ (параметр определяется исчерпывающим спуском). Расчеты показывают, что по числу N итераций и достигнутому приближению к точке минимума в данном примере предпочтение следует отдать варианту метода с использованием исчерпывающего спуска.

Обратим внимание на различие в траекториях поиска точки x^* при постоянных значениях $\varkappa$ (см. рис. 5.11, *а* и *б*). Видно, что при большем из двух значений (при $\varkappa = 0{,}1$) точка минимума целевой функции достигнута за меньшее число итераций.

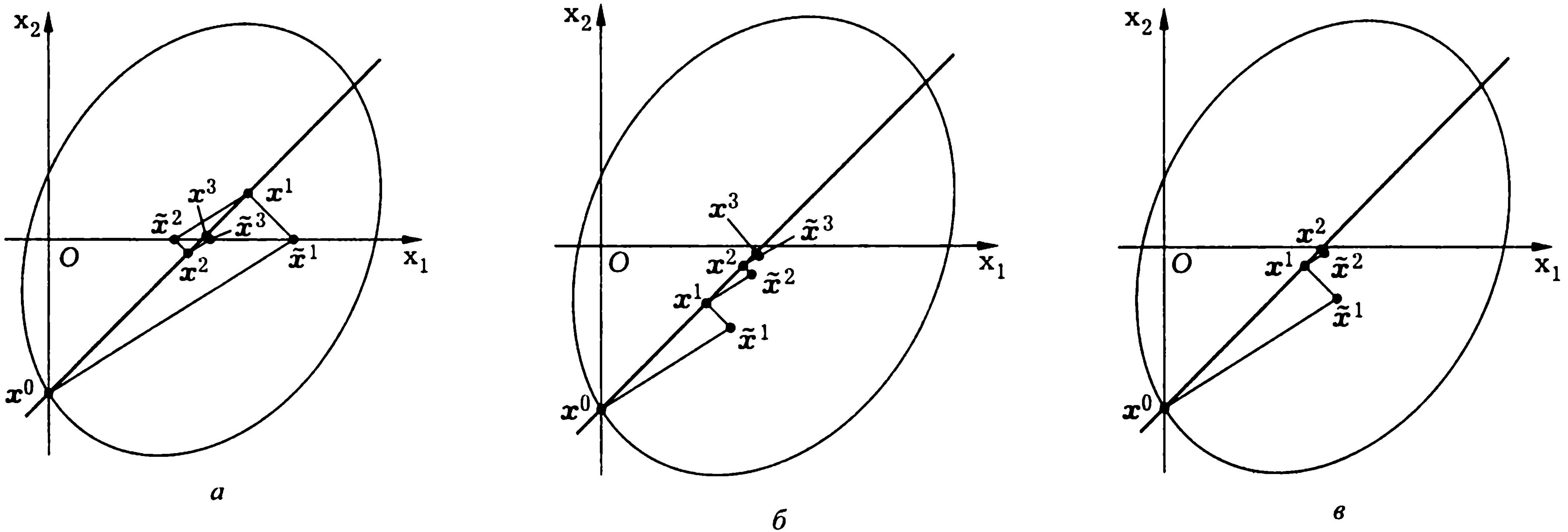


Рис. 5.11

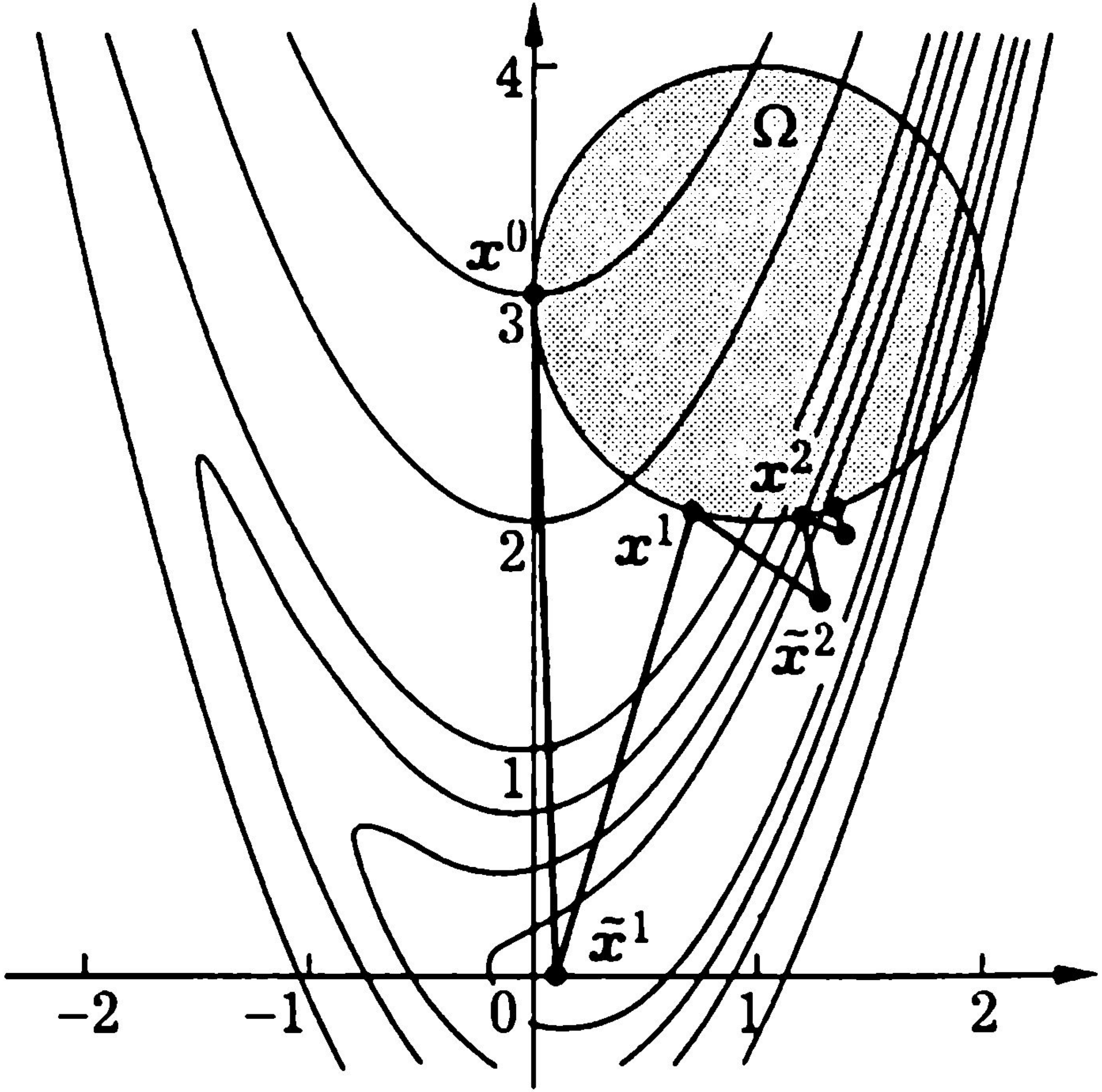
П р и м е р 5.8. Методом проекции точки на множество решим задачу нелинейного программирования

$$f_0(x_1, x_2) = 10(x_1^2 - x_2)^2 + (x_1 - 1)^2 \to \min;$$
$$(x_1 - 1)^2 + (x_2 - 3)^2 \leqslant 1.$$

В качестве начальной выберем точку $\boldsymbol{x}^0 = (0,\ 3) \in \Omega$, а параметр $\varkappa_k$ на каждой итерации будем находить путем исчерпывающего спуска. Проектирование точки $\widetilde{\boldsymbol{x}}^k$ на замкнутый круг Ω осуществляем в соответствии с формулой (5.20). Таким образом,

$$\boldsymbol{x}^k = P_\Omega(\widetilde{\boldsymbol{x}}^k) = \boldsymbol{x}_0 + \frac{\widetilde{\boldsymbol{x}}^k - \boldsymbol{x}_0}{|\widetilde{\boldsymbol{x}}^k - \boldsymbol{x}_0|},$$

где $\boldsymbol{x}_0 = (1,\ 3)$ — центр круга Ω. Прекращение поиска определяется условием $|\boldsymbol{x}^k - \boldsymbol{x}^{k-1}| < \varepsilon_1$, где $\varepsilon_1 = 10^{-3}$ — параметр точности поиска.



4
3
2
1
0
−2
−1
1
2
Ω
x^0
x^1
x^2
$\tilde{x}^1$
$\tilde{x}^2$

Пример 5.9. Методом проекции точки на множество найдем минимум функции $f_0(x_1, x_2)$ из примера 5.8 в прямоугольнике $\Omega \subset \mathbb{R}^2$, заданном соотношениями $x_1 \in [-2{,}5,\ 0]$, $x_2 \in [-1,\ 2]$. В качестве начальной возьмем точку $\boldsymbol{x}^0 = (-2,\ 2)$, а способ выбора параметра $\varkappa_k$ оставим тем же, что и в примере 5.8. Для прямоугольника построение проекции рассмотрено в примере 5.6. На рис. 5.13 показаны первые итерации процесса поиска точки минимума.

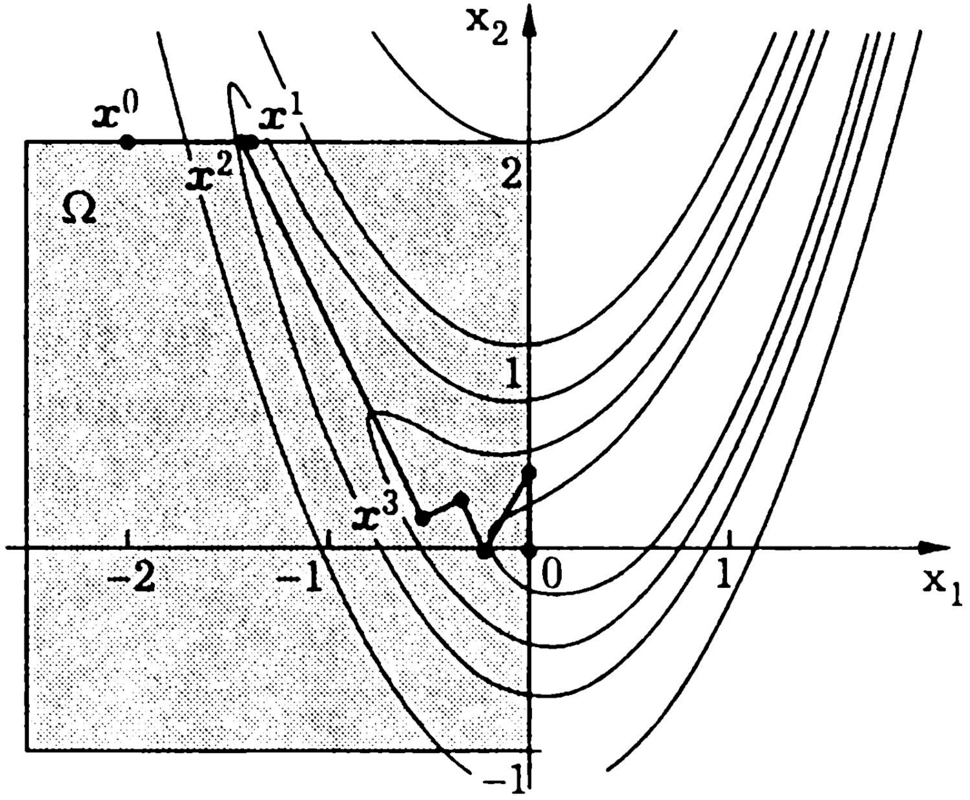


Рис. 5.13

Метод проекции антиградиента

Результатом проектирования точки $\boldsymbol{x} \in \mathbb{R}^n$ на замкнутое множество $\Omega \subset \mathbb{R}^n$ является точка $\boldsymbol{y} \in \Omega$, либо совпадающая с $\boldsymbol{x}$, если $\boldsymbol{x} \in \Omega$, либо лежащая на его границе $\partial\Omega$. Это свойство операции проектирования позволяет при численном решении задачи оптимизации с ограничениями на каждой k-й итерации использовать обычный метод градиентного спуска в направлении антиградиента из точки $\boldsymbol{x}^{k-1}$ в точку $\widetilde{\boldsymbol{x}}^k$, а затем находить проекцию $\boldsymbol{x}^k \in \Omega$ точки $\widetilde{\boldsymbol{x}}^k$ на множество Ω. Если при этом $\widetilde{\boldsymbol{x}}^k \notin \Omega$, то операция проектирования как бы возвращает ее обратно в допустимое множество Ω.

Естествен вопрос: нельзя ли предварительно найти такое *направление спуска*, движение точки по которому при определенных ограничениях вообще не выводило бы ее за пределы допустимого множества Ω?

Сначала рассмотрим случай, когда множество Ω описывается системой линейных ограничений типа равенства:

$$\Omega = \left\{ \boldsymbol{x} \in \mathbb{R}^n \colon (\boldsymbol{a}_i, \boldsymbol{x}) = b_i,\ i = \overline{1, m} \right\}. \tag{5.29}$$

Отметим, что в силу линейности ограничений множество Ω является выпуклым.

Составим вектор $\boldsymbol{b} = (b_1 \ \ldots \ b_m)^{\mathrm{T}}$ и матрицу A размера $m \times n$, строками которой являются матрицы-строки $\boldsymbol{a}_i^{\mathrm{T}}$, $i = \overline{1, m}$. Без ограничения общности можно принять, что $\operatorname{Rg} A = m < n$ В этом случае векторы $\boldsymbol{a}_i$, $i = \overline{1, m}$, линейно независимы.

Отметим, что для любой матрицы A матрица AA^{T} является симметрической, поскольку $(AA^{\mathrm{T}})^{\mathrm{T}} = (A^{\mathrm{T}})^{\mathrm{T}} A^{\mathrm{T}} = AA^{\mathrm{T}}$. Если векторы $\boldsymbol{a}_i \in \mathbb{R}^n$, $i = \overline{1, m}$, линейно независимы, то матрица AA^{T} невырождена и имеет обратную матрицу $(AA^{\mathrm{T}})^{-1}$, причем симметрическую, так как матрица AA^{T} симметрическая.

Квадратная матрица

$$P = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A \tag{5.30}$$

порядка n является симметрической, поскольку

$$P^{\mathrm{T}} = \left(A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A\right)^{\mathrm{T}} = A^{\mathrm{T}}\left(A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}\right)^{\mathrm{T}} =$$

$$= A^{\mathrm{T}}\left((AA^{\mathrm{T}})^{-1}\right)^{\mathrm{T}}(A^{\mathrm{T}})^{\mathrm{T}} = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A = P.$$

Кроме того,

$$PP = \left(A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A\right)A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A =$$

$$= A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(AA^{\mathrm{T}})(AA^{\mathrm{T}})^{-1}A = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A = P,$$

т.е. $P^2 = P$.

Квадратную матрицу M, удовлетворяющую соотношению $M^2 = M$, называют ***идемпотентной матрицей***, а симметрическую идемпотентную матрицу — ***проекционной***. Отметим, что проекционная матрица является неотрицательно определенной, так как

$$(M\boldsymbol{w}, \boldsymbol{w}) = \left(M^2\boldsymbol{w}, \boldsymbol{w}\right) = (M\boldsymbol{w}, M\boldsymbol{w}) = |M\boldsymbol{w}|^2 \geqslant 0.$$

Собственными значениями идемпотентной матрицы могут быть только числа 0 и 1. В самом деле, если $\boldsymbol{x} \neq \boldsymbol{0}$ — собственный вектор идемпотентной матрицы M с собственным значением λ, то

$$\lambda\boldsymbol{x} = M\boldsymbol{x} = M^2\boldsymbol{x} = M(\lambda\boldsymbol{x}) = \lambda^2\boldsymbol{x}.$$

Следовательно, $(\lambda^2 - \lambda)\boldsymbol{x} = \boldsymbol{0}$ и $\lambda^2 - \lambda = 0$, так как $\boldsymbol{x} \neq \boldsymbol{0}$. Уравнению $\lambda^2 - \lambda = 0$ удовлетворяют только числа 0 и 1. Очевидно, что собственными значениями проекционной матрицы также могут быть только числа 0 и 1.

Нетрудно показать, что если P — проекционная матрица, то и $P^* = I - P$ — также проекционная матрица (здесь I — единичная матрица). Действительно,

$$(P^*)^2 = (I - P)^2 = I - 2P + P^2 = I - P = P^* \tag{5.31}$$

и

$$(P^*)^{\mathrm{T}} = (I - P)^{\mathrm{T}} = I - P = P^*. \tag{5.32}$$

С помощью матрицы P любой вектор $w \in \mathbb{R}^n$ можно представить в виде суммы $w = Pw + (I - P)w$, в которой слагаемые Pw и $(I - P)w$ ортогональны. Ортогональность слагаемых вытекает из равенств $P(I - P) = P - P^2 = \Theta$, в соответствии с которыми

$$\big(Pw, (I - P)w\big) = \big(w, P(I - P)w\big) = 0. \tag{5.33}$$

Вектор Pw является проекцией вектора w на линейное подпространство $\mathcal{L}_P = \{x \in \mathbb{R}^n \colon x = Pw\}$. Если матрица P имеет вид (5.30), то базисом в пространстве $\mathcal{L}_P$ являются векторы a_i, $i = \overline{1, m}$, так как в этом случае для любого вектора $y \in \mathcal{L}_P$ имеем

$$P = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A$$

$$y = Pw = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}Aw = A^{\mathrm{T}}d = \sum_{i=1}^{m} d_i a_i, \qquad (5.34)$$

где $d = (AA^{\mathrm{T}})^{-1}Aw$, а d_i, $i = \overline{1, m}$, — координаты вектора d.

Обозначим $P^* = I - P$. Линейное подпространство $\mathcal{L}_{P^*}$ всех векторов вида P^*w, $w \in \mathbb{R}^n$, является ортогональным дополнением линейного подпространства $\mathcal{L}_P$, поскольку для любых $w_1, w_2 \in \mathbb{R}^n$

$$(Pw_1, P^*w_2) = (w_1, P(I - P)w_2) = 0.$$

Поэтому вектор $\boldsymbol{w}^{\perp} = P^{*}\boldsymbol{w}$ ортогонален каждому вектору $\boldsymbol{a}_i$, $i = \overline{1, m}$. Таким образом, вектор $\boldsymbol{w}^{\perp}$ ортогонален нормальному вектору $\boldsymbol{a}_i \in \mathbb{R}^n$ каждой гиперплоскости $(\boldsymbol{a}_i, \boldsymbol{x}) = b_i$, $i = \overline{1, m}$, ограничивающей множество Ω (5.29). Следовательно, направление, задаваемое вектором $\boldsymbol{w}^{\perp}$, параллельно этим гиперплоскостям и движение в этом направлении из любой точки $\boldsymbol{x}^0 \in \Omega$ не выводит за пределы множества Ω.

$$\Omega = \left\{ \boldsymbol{x} \in \mathbb{R}^n \colon (\boldsymbol{a}_i, \boldsymbol{x}) = b_i, \ i = \overline{1, m} \right\}$$

Отметим, что в силу линейности ограничений множество Ω является выпуклым.

Линейные ограничения типа равенства. Пусть в задаче оптимизации допустимое множество Ω определено ограничениями типа равенства, т.е. имеет вид (5.29), а целевая функция $f_0(\boldsymbol{x})$ дифференцируема в $\mathbb{R}^n$. Предположим, что задана точка $\boldsymbol{x}^0 \in \Omega$. Рассмотрим функцию

$$\varphi(\boldsymbol{y}) = f_0(\boldsymbol{x}^0 + P^*\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{R}^n, \tag{5.35}$$

где матрица P определена соотношением (5.30). Для этой функции в силу правила дифференцирования сложной функции и симметричности матрицы P справедливо равенство*

$$\operatorname{grad} \varphi(\boldsymbol{y}) = P^* \operatorname{grad} f_0(\boldsymbol{x}), \tag{5.36}$$

где $\boldsymbol{x}$ и $\boldsymbol{y}$ связаны соотношением

$$P^* = I - P$$

$$\boldsymbol{x} = \boldsymbol{x}^0 + P^*\boldsymbol{y}, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n. \tag{5.37}$$

$$P = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A$$

Теорема 5.2. Если $\boldsymbol{y}^*$ — стационарная точка функции $\varphi(\boldsymbol{y})$, то соответствующая ей точка

$$\boldsymbol{x}^* = \boldsymbol{x}^0 + P^* \boldsymbol{y}^* \tag{5.38}$$

удовлетворяет необходимому условию минимума целевой функции $f_0(\boldsymbol{x})$ на допустимом множестве Ω вида (5.29).

$$\big(\operatorname{grad} f_0(\boldsymbol{x}^*), \boldsymbol{e}\big) \geqslant 0, \quad \boldsymbol{e} \in C(\boldsymbol{x}^*)$$

П р и м е р 5.10. Применим теорему 5.2 для поиска минимума квадратичной функции $f_0(\boldsymbol{x}) = \frac{1}{2}(Q\boldsymbol{x}, \boldsymbol{x})$ с положительно определенной матрицей Q порядка n на множестве Ω вида (5.29).

Поскольку в данном случае $\operatorname{grad} f_0(\boldsymbol{x}) = Q\boldsymbol{x}$, то в соответствии с соотношениями (5.36) и (5.37) можно записать

$$\operatorname{grad} \varphi(\boldsymbol{y}) = P^* Q\big(\boldsymbol{x}^0 + P^* \boldsymbol{y}\big).$$

Из условия $\operatorname{grad} \varphi(\boldsymbol{y}^*) = \boldsymbol{0}$ получим СЛАУ

$$P^* Q P^* \boldsymbol{y} = -P^* Q \boldsymbol{x}^0, \qquad (5.41)$$

которой должна удовлетворять стационарная точка $\boldsymbol{y}^*$ функции $\varphi(\boldsymbol{y})$ (5.35).

Остановимся на частном случае функции двух переменных. Пусть $\boxed{f_0(\boldsymbol{x}) = f_0(x_1, x_2) = x_1^2 + x_2^2}$, а допустимое множество задано одним ограничением $\boxed{x_2 = 2}$. Тогда

$$Q = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

а в представлении (5.29) $m = 1$, $\boldsymbol{a}_1 = (0 \;\; 1)^{\mathrm{T}}$ и $b_1 = 2$.

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \colon (\boldsymbol{a}_i, \boldsymbol{x}) = b_i,\; i = \overline{1, m}\}$$

Ограничению типа равенства удовлетворяет, например, точка $\boldsymbol{x}^0 = (2,\ 2)$, лежащая на прямой, задаваемой уравнением $(\boldsymbol{a}_1, \boldsymbol{x}) = 2$, или $x_2 = 2$. В данном случае имеем $A = \boldsymbol{a}_1^{\mathrm{T}} = (0\ \ 1)$, $AA^{\mathrm{T}} = \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1 = |\boldsymbol{a}_1|^2 = 1$ и

$$P = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot 1 \cdot (0\ \ 1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix};$$

$$P^* = I - P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix};$$

$$P^*QP^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix};$$

$$P^*Q\boldsymbol{x}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}.$$

Система линейных уравнений (5.41) относительно координат точки $y^* = (y_1^*,\ y_2^*)$ принимает вид

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \end{pmatrix},$$

или $2y_1^* + 0 \cdot y_2^* = -4$ и $0 \cdot y_1^* + 0 \cdot y_2^* = 0$. Из первого уравнения находим $y_1^* = -2$, а второму уравнению удовлетворяет произвольное значение y_2^*. Используя равенство (5.38), получаем

$$\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} + P^* \begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ y_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

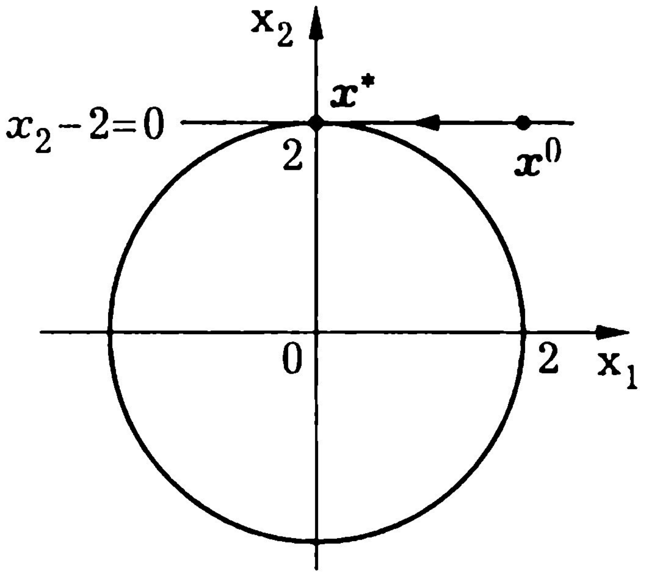


Рис. 5.14

Таким образом, $\boldsymbol{x}^* = (0,\ 2)$ и $f_0(\boldsymbol{x}^*) = 4$. Это точка минимума функции $f_0(\boldsymbol{x})$ на прямой $x_2 = 2$, так как функция строго выпуклая. Полученный результат очевиден с геометрической точки зрения: $\boldsymbol{x}^* = (0,\ 2)$ является точкой касания прямой $x_2 = 2$ и линии уровня $f_0(x_1, x_2) = x_1^2 + x_2^2 = 4$ целевой функции (рис. 5.14).

Рассмотрим процедуру поиска минимума дифференцируемой целевой функции $f_0(\boldsymbol{x})$ на множестве Ω (5.29). Выберем начальную точку $\boldsymbol{x}^0 \in \Omega$, вычислим в этой точке антиградиент $\boldsymbol{w}^1 = -\operatorname{grad} f_0(\boldsymbol{x}^0)$ и антиградиент $\boldsymbol{p}^1 = P^*\boldsymbol{w}^1$ функции $\varphi(\boldsymbol{y})$ в точке $\boldsymbol{y} = \boldsymbol{w}^1$. Если $\boldsymbol{p}^1 = \mathbf{0}$, то градиент функции $\varphi(\boldsymbol{y})$ в точке $\boldsymbol{w}^1$ равен нулевому вектору, т.е. $\boldsymbol{w}^1 = \boldsymbol{y}^*$ — стационарная точка этой функции, а в силу теоремы 5.2 $\boldsymbol{x}^0$ — точка, удовлетворяющая необходимому условию минимума целевой функции на множестве Ω (для выпуклой целевой функции $\boldsymbol{x}^0$ — точка ее наименьшего значения на Ω).

При численном решении целесообразно считать $\boldsymbol{y} = \boldsymbol{w}^1$ стационарной точкой функции $\varphi(\boldsymbol{y})$ при выполнении условия, аналогичного неравенству (3.3): $|\boldsymbol{p}^1| < \varepsilon_3$, где $\varepsilon_3 > 0$ — параметр точности поиска. Если $|\boldsymbol{p}^1| \geqslant \varepsilon_3$, то вектор $\boldsymbol{p}^1$ определяет направление спуска для функции $\varphi(\boldsymbol{y})$ при ее безусловной минимизации. Но этот же вектор задает *возможное направление спуска* для целевой функции, если в соответствии с (5.37) для первой итерации записать

$$\boldsymbol{x}^1 = \boldsymbol{x}^0 + \varkappa_1 P^* \boldsymbol{w}^1 = \boldsymbol{x}^0 + \varkappa_1 \boldsymbol{p}^1, \qquad \varkappa_1 > 0. \tag{5.42}$$

Действительно, точка $\boldsymbol{x}^1$ принадлежит множеству Ω при любом значении $\varkappa_1$, так как, учитывая равенства $(\boldsymbol{a}_i, P^*\boldsymbol{w}) = 0$, $i = \overline{1, m}$, для любого вектора $\boldsymbol{w}^1$ имеем

$$(\boldsymbol{a}_i, \boldsymbol{x}^1) = (\boldsymbol{a}_i, \boldsymbol{x}^0) + (\boldsymbol{a}_i, P^*\boldsymbol{w}^1) = (\boldsymbol{a}_i, \boldsymbol{x}^0) = b_i, \quad i = \overline{1, m}.$$

Поскольку с учетом равенств (5.31) и (5.32)

$$(\boldsymbol{p}^1, \operatorname{grad} f_0(\boldsymbol{x}^0)) = -(P^*\boldsymbol{w}^1, \boldsymbol{w}^1) = -((P^*)^2\boldsymbol{w}^1, \boldsymbol{w}^1) =$$
$$= -(P^*\boldsymbol{w}^1, P^*\boldsymbol{w}^1) = -|\boldsymbol{p}^1|^2 < 0,$$

вектор $\boldsymbol{p}^1$ задает направление спуска и для целевой функции, в котором ее значение уменьшается.

Значение $\varkappa_1$ в соотношении (5.42) можно найти с помощью исчерпывающего спуска в направлении $\boldsymbol{p}^1$ или методом дробления шага, задаваясь некоторым первоначальным значением $\varkappa_0$ и при необходимости уменьшая его так, чтобы добиться выполнения неравенства $f_0(\boldsymbol{x}^1) < f_0(\boldsymbol{x}^0)$. Это позволит начать формировать релаксационную последовательность $\{\boldsymbol{x}^k\}$.

Найдя точку $\boldsymbol{x}^1$ по соотношению (5.42), переходим ко второй итерации и т.д. На k-й итерации, вычисляя в точке $\boldsymbol{x}^{k-1} \in \Omega$ антиградиент $\boldsymbol{w}^k = -\operatorname{grad} f_0(\boldsymbol{x}^{k-1})$, находим возможное направление спуска $\boldsymbol{p}^k = P^* \boldsymbol{w}^k$; в этом направлении (если $|\boldsymbol{p}^k| \geqslant \varepsilon_3$) ищем точку

$$\boldsymbol{x}^k = \boldsymbol{x}^{k-1} + \varkappa_k P^* \boldsymbol{w}^k = \boldsymbol{x}^{k-1} + \varkappa_k \boldsymbol{p}^k, \quad \varkappa_k > 0, \tag{5.43}$$

подбирая соответствующим образом значение $\varkappa_k$. Можно показать, что $\boldsymbol{x}^k \in \Omega$ при любом значении $\varkappa_k$. Следовательно, можно переходить к следующей, $(k+1)$-й итерации. Если $|\boldsymbol{p}^k| < \varepsilon_3$, то итерации прекращаем, полагая $\boldsymbol{w}^* \approx \boldsymbol{w}^k$ и $\boldsymbol{x}^* \approx \boldsymbol{x}^{k-1}$.

Поскольку на каждой k-й итерации вектор $\boldsymbol{p}^k$ является проекцией антиградиента $\boldsymbol{w}^k$ целевой функции на ортогональное дополнение подпространства $\mathcal{L}_P$, то описанную процедуру поиска минимума этой функции на множестве Ω назовем ***методом проекции антиградиента***.