Глава 6. МЕТОДЫ ПЕРВОГО ПОРЯДКА

6.1. МЕТОД ГРАДИЕНТНОГО СПУСКА С ПОСТОЯННЫМ ШАГОМ

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0,1,\ldots$, таких, что $f\left(x^{k+1}\right) < f\left(x^k\right)$, $k = 0,1,\ldots$. Точки последовательности $\{x^k\}$ вычисляются по правилу

$$x^{k+1} = x^k - t_k \nabla f\left(x^k\right),\ k = 0,1,\ldots\ , \qquad (6.1)$$

где точка x^0 задается пользователем; $\nabla f\left(x^k\right)$ – градиент функции $f(x)$, вычисленный в точке x^k; величина шага t_k задается пользователем и остается постоянной до тех пор, пока функция убывает в точках последовательности, что контролируется путем проверки выполнения условия $f\left(x^{k+1}\right) - f\left(x^k\right) < 0$ или $f\left(x^{k+1}\right) - f\left(x^k\right) < -\varepsilon \left\| \nabla f\left(x^k\right) \right\|^2, 0 < \varepsilon < 1$ [28]. Построение последовательности $\{x^k\}$ заканчивается в точке x^k, для которой $\left\| \nabla f\left(x^k\right) \right\| < \varepsilon_1$, где ε_1 – заданное малое положительное число, или $k \ge M$, где M – предельное число итераций, или при двукратном одновременном выполнении двух неравенств $\left\| x^{k+1} - x^k \right\| < \varepsilon_2$, $\left| f\left(x^{k+1}\right) - f\left(x^k\right) \right| < \varepsilon_2$, где ε_2 – малое положительное число.

Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается с помощью дополнительного исследования, которое описано ниже.

Алгоритм

Шаг 1. Задать x^0, $0 < \varepsilon < 1$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M – предельное число итераций. Найти градиент функции в произвольной точке $\nabla f(x) = \left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n} \right)^T$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f\left(x^k\right)$.

Шаг 4. Проверить выполнение критерия окончания $\left\| \nabla f\left(x^k\right) \right\| < \varepsilon_1$:

а) если критерий выполнен, расчет закончен, $x^* = x^k$;
б) если критерий не выполнен, то перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \geq M$:

а) если неравенство выполнено, то расчет окончен: $x^* = x^k$;
б) если нет, то перейти к шагу 6.

Шаг 6. Задать величину шага t_k.

Шаг 7. Вычислить $x^{k+1} = x^k - t_k \nabla f\left(x^k\right)$.

Шаг 8. Проверить выполнение условия

$$f\left(x^{k+1}\right) - f\left(x^k\right) < 0 \quad \left(\text{или } f\left(x^{k+1}\right) - f\left(x^k\right) < -\varepsilon \left\| \nabla f\left(x^k\right) \right\|^2\right):$$

а) если условие выполнено, то перейти к шагу 9;
б) если условие не выполнено, положить $t_k = \dfrac{t_k}{2}$ и перейти к шагу 7.

Шаг 9. Проверить выполнение условий

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

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;
б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 3.

Геометрическая интерпретация метода для $n = 2$ приведена на рис. 6.1.

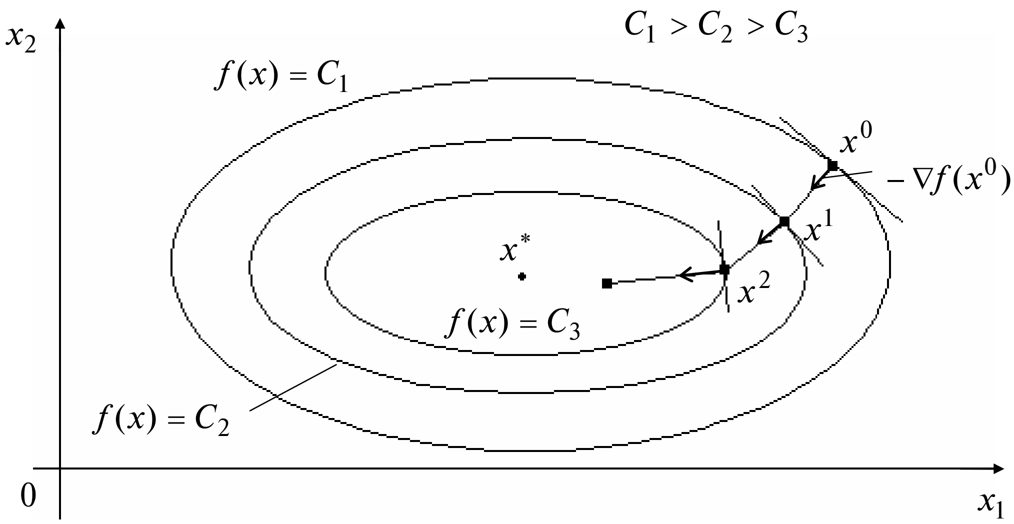

Рис. 6.1

Сходимость

Утверждение 6.1 [28]**.** *Пусть функция* $f(x)$ *дифференцируема и ограничена снизу на* R^n*, а ее градиент удовлетворяет условию Липшица* $\| \nabla f(x) - \nabla f(y) \| \le L \| x - y \|$*,* $\forall x, y \in R^n$*, где* $L > 0$*. Тогда при произвольной начальной точке* $x^0 \in R^n$ *для метода градиентного спуска с постоянным шагом имеем*

$$\lim_{k \to \infty} \left\| \nabla f\left(x^k\right) \right\| = 0. \tag{6.2}$$

З а м е ч а н и я 6.1.

1. Утверждение 6.1 гарантирует сходимость последовательности $\left\{x^k\right\}$ к стационарной точке x^*, где $\nabla f\left(x^*\right) = 0$. Следовательно, найденная в результате применения метода точка x^* нуждается в дополнительном исследовании с целью ее классификации.

2. Метод градиентного спуска гарантирует сходимость последовательности $\left\{x^k\right\}$ к точке минимума для сильно выпуклых функций [28].

3. При решении примеров итерационный процесс подбора удачной величины t_k отражается в индексации шагов 7 и 8. Первый индекс совпадает с номером k, а второй – с числом делений текущей величины t_k пополам.

Скорость сходимости

Оценки скорости сходимости получены только для сильно выпуклых функций, когда последовательность $\left\{x^k\right\}$ сходится к точке минимума $f(x)$ со скоростью геометрической прогрессии:

$$f\left(x^k\right) - f\left(x^*\right) \le q^k\left(f\left(x^0\right) - f\left(x^*\right)\right), \qquad \left\|x^k - x^*\right\| \le C\left(\sqrt{q}\right)^k,$$

где $q \in (0,1)$, $C > 0$ – константы [39].

Процедура решения задачи

1. Используя алгоритм градиентного спуска с постоянным шагом, найти точку x^k, в которой выполнен по крайней мере один из критериев окончания расчетов.

2. Провести анализ точки x^k с целью установить, является ли точка x^k найденным приближением решения задачи. Процедура анализа определяется наличием у функции $f(x)$ непрерывных вторых производных. Если $f(x) \in C^2$, то следует провести проверку выполнения достаточных условий минимума: $H(x^*) > 0$. Если $H\left(x^k\right) > 0$, то точка x^k есть найденное приближение искомой точки x^*. Если $f(x) \in C^1$, то следует проверить функцию $f(x)$ на выпуклость в Q-окрестности точки x^k, используя критерий выпуклости для функций $f(x) \in C^1$: *функция* $f(x)$ *выпукла* (*строго выпукла*) *в том и только в*

том случае, если $f(x+y) \ge f(x) + (\nabla f(x), y)$, $\forall x, y \in Q$; $(f(x+y) > f(x) + (\nabla f(x), y))$; (эквивалентное определение см. в гл. 1). Если функция $f(x)$ выпукла (строго выпукла), то x^k есть найденное приближение точки x^*.

З а м е ч а н и е 6.2. Если требуется найти глобальный минимум функции $f(x)$, то для строго выпуклой $f(x)$ решение этой задачи аналогично поиску локального минимума функции. В случае, когда $f(x)$ имеет несколько локальных минимумов, поиск глобального минимума осуществляется в результате перебора всех локальных минимумов.

Пример 6.1. Найти локальный минимум функции

$$f(x) = 2x_1^2 + x_1x_2 + x_2^2.$$

□ I. Определим точку x^k, в которой выполнен по крайней мере один из критериев окончания расчетов.

1. Зададим x^0, ε_1, ε_2, M: $x^0 = (0{,}5;1)^T$, $\varepsilon_1 = 0{,}1$; $\varepsilon_2 = 0{,}15$; $M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Положим $k = 0$.

3^0. Вычислим $\nabla f(x^0)$: $\nabla f(x^0) = (3; 2{,}5)^T$.

4^0. Вычислим $\|\nabla f(x^0)\|$: $\|\nabla f(x^0)\| = 3{,}9 > 0{,}1$. Перейдем к шагу 5.

5^0. Проверим условие $k \ge M$: $k = 0 < 10 = M$. Перейдем к шагу 6.

6^0. Зададим $t_0 = 0{,}5$.

7^0. Вычислим x^1: $x^1 = (0{,}5;1)^T - 0{,}5(3; 2{,}5)^T = (-1;-0{,}25)^T$; $f(x^1) = 2{,}31$.

8^0. Сравним $f(x^1)$ с $f(x^0) = 2$. Имеем $f(x^1) > f(x^0)$. Вывод: условие $f(x^{k+1}) < f(x^k)$ для $k = 0$ не выполняется. Зададим $t^0 = 0{,}25$, перейдем к повторению шагов 7, 8.

7^{01}. Вычислим x^1: $x^1 = (0{,}5;1)^T - 0{,}25(3; 2{,}5)^T = (-0{,}25; 0{,}375)^T$; $f(x^1) = 0{,}171$.

8^{01}. Сравним $f(x^1)$ и $f(x^0)$. Вывод: $f(x^1) < f(x^0)$. Перейдем к шагу 9.

9^0. Вычислим $\|x^1 - x^0\|$ и $|f(x^1) - f(x^0)|$:

$$\|x^1 - x^0\| = 0{,}976 > 0{,}15; \quad |f(x^1) - f(x^0)| = 1{,}829 > 0{,}15.$$

Вывод: положим $k = 1$ и перейдем к шагу 3.

3^1. Вычислим $\nabla f(x^1)$: $\nabla f(x^1) = (-0{,}625; 0{,}51)^T$.

4^1. Вычислим $\|\nabla f(x^1)\|$: $\|\nabla f(x^1)\| = 0{,}81 > 0{,}1$. Перейдем к шагу 5.

5^1. Проверим условие $k \ge M$: $k = 1 < 10 = M$. Перейдем к шагу 6.

6^1. Зададим $t_1 = 0{,}25$.

7^1. Вычислим x^2: $x^2 = \left(-0{,}25; 0{,}375\right)^T - 0{,}25\left(-0{,}625; 0{,}5\right)^T = \left(-0{,}094; 0{,}25\right)^T$; $f\left(x^2\right) = 0{,}056$.

8^1. Сравним $f\left(x^2\right)$ с $f\left(x^1\right)$. Вывод: $f\left(x^2\right) < f\left(x^1\right)$. Перейдем к шагу 9.

9^1. Вычислим $\left\|x^2 - x^1\right\|$ и $\left|f\left(x^2\right) - f\left(x^1\right)\right|$:

$$\left\|x^2 - x^1\right\| = 0{,}2 > 0{,}15; \qquad \left|f\left(x^2\right) - f\left(x^1\right)\right| = 0{,}115 < 0{,}15.$$

Вывод: положим $k = 2$ и перейдем к шагу 3.

3^2. Вычислим $\nabla f\left(x^2\right)$: $\nabla f\left(x^2\right) = \left(-0{,}126; 0{,}406\right)^T$.

4^2. Вычислим $\left\|\nabla f\left(x^2\right)\right\|$: $\left\|\nabla f\left(x^2\right)\right\| = 0{,}425 > 0{,}1$. Перейдем к шагу 5.

5^2. Проверим условие $k \ge M$: $k = 2 < 10 = M$, перейдем к шагу 6.

6^2. Зададим $t_2 = 0{,}25$.

7^2. Вычислим x^3: $x^3 = \left(-0{,}094; 0{,}25\right)^T - 0{,}25\left(-0{,}126; 0{,}406\right)^T = \left(-0{,}063; 0{,}15\right)^T$; $f\left(x^3\right) = 0{,}021$.

8^2. Сравним $f\left(x^3\right)$ и $f\left(x^2\right)$. Вывод: $f\left(x^3\right) < f\left(x^2\right)$. Перейдем к шагу 9.

9^2. Вычислим $\left\|x^3 - x^2\right\|$ и $\left|f\left(x^3\right) - f\left(x^2\right)\right|$:

$$\left\|x^3 - x^2\right\| = 0{,}105 < 0{,}15; \qquad \left|f\left(x^3\right) - f\left(x^2\right)\right| = 0{,}035 < 0{,}15.$$

Вывод: положим $k = 3$ и перейдем к шагу 3.

3^3. Вычислим $\nabla f\left(x^3\right)$: $\nabla f\left(x^3\right) = \left(-0{,}102; 0{,}237\right)^T$.

4^3. Вычислим $\left\|\nabla f\left(x^3\right)\right\|$: $\left\|\nabla f\left(x^3\right)\right\| = 0{,}257 > 0{,}1$. Перейдем к шагу 5.

5^3. Проверим условие $k \ge M$: $k = 3 < 10 = M$, перейдем к шагу 6.

6^3. Зададим $t_3 = 0{,}25$.

7^3. Вычислим x^4: $x^4 = \left(-0{,}063; 0{,}15\right)^T - 0{,}25\left(-0{,}102; 0{,}237\right)^T = \left(-0{,}038; 0{,}091\right)^T$; $f\left(x^4\right) = 0{,}0076$.

8^3. Сравним $f\left(x^4\right)$ и $f\left(x^3\right)$: $f\left(x^4\right) < f\left(x^3\right)$.

9^3. Вычислим $\left\|x^4 - x^3\right\|$, $\left|f\left(x^4\right) - f\left(x^3\right)\right|$:

$$\left\|x^4 - x^3\right\| = 0{,}064 < 0{,}15; \qquad \left|f\left(x^4\right) - f\left(x^3\right)\right| = 0{,}015 < 0{,}15.$$

Условия $\left\|x^{k+1} - x^k\right\| < \varepsilon_2$, $\left|f\left(x^{k+1}\right) - f\left(x^k\right)\right| < \varepsilon_2$ выполнены при $k = 2{,}3$. Расчет окончен. Найдена точка $x^4 = \left(-0{,}038; 0{,}091\right)^T$; $f\left(x^4\right) = 0{,}0076$. На рис. 6.2 полученные точки соединены штриховой линией.

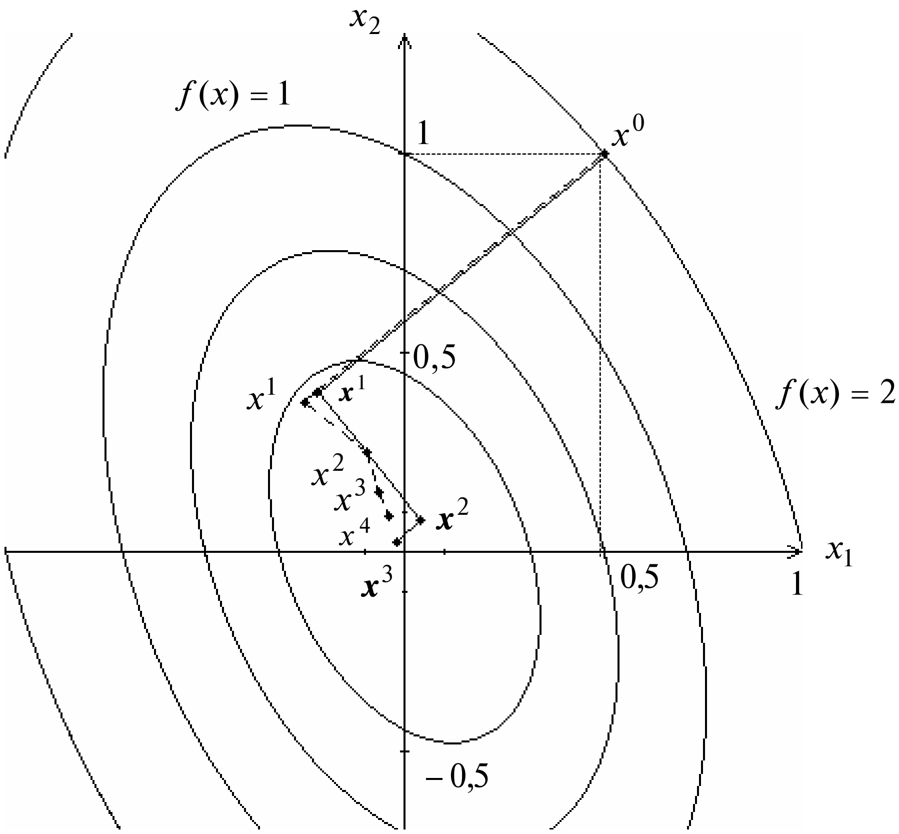

Рис. 6.2

II. Проведем анализ точки x^4.

Функция $f(x) = 2x_1^2 + x_1x_2 + x_2^2$ является дважды дифференцируемой, поэтому проведем проверку достаточных условий минимума в точке x^4. Для этого проанализируем матрицу Гессе $H = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$. Матрица является постоянной и положительно определенной (т.е. $H > 0$), так как оба ее угловых минора $\Delta_1 = 4$ и $\Delta_2 = 7$ положительны. Следовательно, точка $x^4 = (-0{,}038; 0{,}091)^T$ есть найденное приближение точки локального минимума $x^* = (0{,}0)^T$, а значение $f(x^4) = 0{,}0076$ есть найденное приближение значения $f(x^*) = 0$. Заметим, что условие $H > 0$ является одновременно условием строгой выпуклости функции $f(x) = 2x_1^2 + x_1x_2 + x_2^2$ на R^2 (см. гл. 1). Следовательно, $x^4 = (-0{,}038; 0{,}091)^T$, $f(x^4) = 0{,}0076$ есть найденные приближения точки глобального минимума $f(x)$ и ее наименьшего значения на R^2. ■

6.2. МЕТОД НАИСКОРЕЙШЕГО ГРАДИЕНТНОГО СПУСКА

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0,1,\ldots$, таких, что $f\left(x^{k+1}\right) < f\left(x^k\right)$, $k = 0,1,\ldots$. Точки последовательности $\{x^k\}$ вычисляются по правилу

$$x^{k+1} = x^k - t_k \nabla f\left(x^k\right), \tag{6.3}$$

где точка x^0 задается пользователем; величина шага t_k определяется для каждого значения k из условия

$$\varphi(t_k) = f\left(x^k - t_k \nabla f\left(x^k\right)\right) \to \min_{t_k}. \tag{6.4}$$

Задача (6.4) может решаться с использованием необходимого условия минимума $\frac{d\varphi}{dt_k} = 0$ и последующей проверкой достаточного условия минимума $\frac{d^2\varphi}{dt_k{}^2} > 0$. Такой путь может быть выбран либо при достаточно простой минимизируемой функции $\varphi(t_k)$, либо при предварительной аппроксимации достаточно сложной функции $\varphi(t_k) = f\left(x^k - t_k \nabla f\left(x^k\right)\right)$ полиномом $P(t_k)$ (как правило, второй или третьей степени), и тогда условие $\frac{d\varphi}{dt_k} = 0$ замещается условием $\frac{dP}{dt_k} = 0$, а условие $\frac{d^2\varphi}{dt_k{}^2} > 0$ – условием $\frac{d^2 P}{dt_k{}^2} > 0$.

Другой путь решения задачи (6.4) связан с использованием численных методов, когда ищется $\min\limits_{t_k \in [a,b]} \varphi(t_k) = \min\limits_{t_k \in [a,b]} f\left(x^k - t_k \nabla f\left(x^k\right)\right)$ (см. разд. 5.1). Границы интервала $[a, b]$ задаются пользователем. При этом степень близости найденного значения t_k к оптимальному значению t_k^*, удовлетворяющему условиям $\frac{d\varphi}{dt_k} = 0$, $\frac{d^2\varphi}{dt_k^2} > 0$, зависит от задания интервала $[a, b]$ и точности методов одномерной минимизации [27].

Построение последовательности $\{x^k\}$, $k = 0,1,\dots$, заканчивается в точке x^k, для которой $\|\nabla f(x^k)\| < \varepsilon_1$, где ε_1 – заданное число, или, если $k \ge M$, M – предельное число итераций, или при двукратном одновременном выполнении неравенств $\|x^{k+1} - x^k\| < \varepsilon_2$, $|f(x^{k+1}) - f(x^k)| < \varepsilon_2$, где ε_2 – малое положительное число.

Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки локального минимума x^* решается путем дополнительного исследования.

Алгоритм

Шаг 1. Задать x^0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, предельное число итераций M. Найти градиент функции в произвольной точке $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n}\right)^T$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить выполнение критерия окончания $\|\nabla f(x^k)\| < \varepsilon_1$:

а) если критерий выполнен, то $x^* = x^k$;
б) если критерий не выполнен, то перейти к шагу 5.

Шаг 5. Проверить выполнение неравенства $k \ge M$:

а) если неравенство выполнено, то $x^* = x^k$;
б) если нет, то перейти к шагу 6.

Шаг 6. Вычислить величину шага t_k^* из условия

$$\varphi(t_k) = f\left(x^k - t_k \nabla f(x^k)\right) \to \min_{t_k}.$$

Шаг 7. Вычислить $x^{k+1} = x^k - t_k^* \nabla f(x^k)$.

Шаг 8. Проверить выполнение условий

$$\|x^{k+1} - x^k\| < \varepsilon_2, \quad |f(x^{k+1}) - f(x^k)| < \varepsilon_2:$$

а) если оба условия выполнены при текущем значении k и $k = k - 1$, то расчет окончен, $x^* = x^{k+1}$;
б) если хотя бы одно из условий не выполнено, то положить $k = k + 1$ и перейти к шагу 3.

Геометрическая интерпретация метода для $n = 2$ приведена на рис. 6.3.

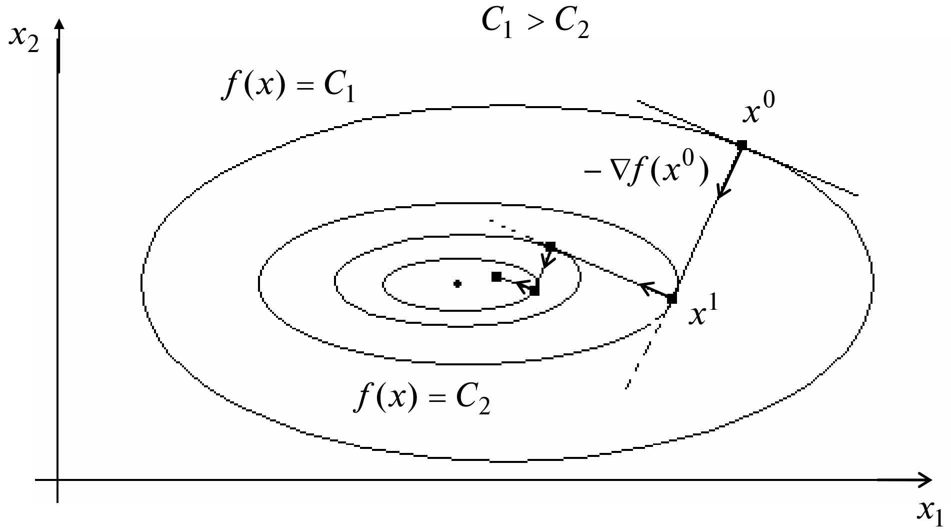

Рис. 6.3

Сходимость

Утверждение 6.2. *Пусть функция* $f(x)$ *удовлетворяет условиям утверждения 6.1. Тогда при произвольной начальной точке* $x^0 \in R^n$ *для метода наискорейшего градиентного спуска имеем* $\left\| \nabla f\left(x^k\right) \right\| \to 0$ при $k \to \infty$ [28].

З а м е ч а н и я 6.3.

1. Утверждение гарантирует сходимость последовательности $\left\{x^k\right\}$ к стационарной точке x^*, где $\nabla f\left(x^*\right) = 0$. Следовательно, найденная в результате применения метода точка x^* нуждается в дополнительном исследовании с целью ее классификации.

2. Метод наискорейшего спуска гарантирует сходимость последовательности $\left\{x^k\right\}$ к точке минимума для сильно выпуклых функций [28].

Скорость сходимости

Оценки скорости сходимости получены только для сильно выпуклых функций, когда последовательность $\left\{x^k\right\}$ сходится к точке минимума функции $f(x)$ со скоростью геометрической прогрессии (линейная сходимость): $\left\|x^{k+1} - x^k\right\| \le \dfrac{M - m}{M + m} \left\|x^k - x^*\right\|$, где M и m – оценки наибольшего и наименьшего собственных значений матрицы $H(x)$ функции $f(x)$ [28].

З а м е ч а н и я 6.4.

1. Процедура решения задачи совпадает с описанной в разд. 6.1.

2. Относительно процедуры поиска глобального минимума функции $f(x)$ остается справедливым замечание 6.2.

Пример 6.2. Найти локальный минимум функции

$$f\left(x\right)=2x_1{}^2+x_1x_2+x_2{}^2 .$$

☐ I. Определим точку x^k, в которой выполнен по крайней мере один из критериев окончания расчетов.

1. Зададим x^0, ε_1, ε_2, M: $x^0=\left(0{,}5;1\right)^T$; $\varepsilon_1=0{,}1$; $\varepsilon_2=0{,}15$; $M=10$. Найдем градиент функции в произвольной точке $\nabla f(x)=(4x_1+x_2;x_1+2x_2)^T$.

2. Положим $k=0$.

3^0. Вычислим $\nabla f\left(x^0\right)$: $\nabla f\left(x^0\right)=\left(3;2{,}5\right)^T$.

4^0. Вычислим $\left\|\nabla f\left(x^0\right)\right\|$: $\left\|\nabla f\left(x^0\right)\right\|=3{,}9>0{,}1$. Перейдем к шагу 5.

5^0. Проверим условие $k\ge M$: $k=0<10=M$, перейдем к шагу 6.

6^0. Следующую точку найдем по формуле

$$x^1=x^0-t_0\nabla f(x^0)=(0{,}5;\,1)^T-t_0(3;\,2{,}5)^T=(0{,}5-3t_0;1-2{,}5\cdot t_0)^T .$$

Подставим полученные выражения $x_1^1=0{,}5-3t_0$, $x_2^1=1-2{,}5\cdot t_0$ для координат в $f(x)$: $\varphi(t_0)=2\cdot(0{,}5-3t_0)^2+(0{,}5-3t_0)\cdot(1-2{,}5\cdot t_0)+(1-2{,}5\cdot t_0)^2$. Найдем минимум функции $\varphi(t_0)$ по t_0 с помощью необходимых условий безусловного экстремума:

$$\frac{d\varphi(t_0)}{dt_0}=4\cdot(0{,}5-3t_0)\cdot(-3)+(-3)\cdot(1-2{,}5t_0)+(-2{,}5)\cdot(0{,}5-3t_0)+2\cdot(1-2{,}5\cdot t_0)\cdot(-2{,}5)=$$

$=-15{,}25+63{,}25\cdot t_0=0$. Отсюда $t_0{}^*\cong 0{,}24$. Так как $\dfrac{d^2\varphi(t_0)}{dt_0^2}=63{,}25>0$, найденное значение шага обеспечивает минимум функции $\varphi(t_0)$ по t_0.

Заметим, что можно получить формулу для вычисления наилучшей величины шага t_k^* на любой итерации из условия

$$\varphi\left(t_k\right)=f\left(x^k-t_k\,\nabla f\left(x^k\right)\right)\to\min_{t_k} .$$

Имеем

$$\nabla f\left(x^k\right)=\left(4x_1^k+x_2^k;x_1^k+x_2^k\right)^T;x^k-t_k\nabla f\left(x^k\right)=\left[x_1^k-t_k\left(4x_1^k+x_2^k\right);x_2^k-t_k\left(x_1^k+x_2^k\right)\right]^T,$$

$$\varphi(t_k)=2\left(x_1^k-t_k(4x_1^k+x_2^k)\right)^2+\left(x_1^k-t_k(4x_1^k+x_2^k)\right)\left(x_2^k-t_k(x_1^k+x_2^k)\right)+$$
$$+\left(x_2^k-t_k(x_1^k+x_2^k)\right)^2 .$$

Из условия $\dfrac{d\varphi}{dt_k}=0$ получаем

$$t_k^* = \frac{\left(4x_1{}^k + x_2{}^k\right)^2 + \left(x_1{}^k + 2x_2{}^k\right)^2}{4\left(4x_1{}^k + x_2{}^k\right)^2 + 2\left(4x_1{}^k + x_2{}^k\right)\left(x_1{}^k + 2x_2{}^k\right) + 2\left(x_1{}^k + 2x_2{}^k\right)^2}.$$

Определим t_0^*: $t_0^* = 0{,}24$.

7^0. Найдем $x^1 = x^0 - t_0^* \nabla f\left(x^0\right)$: $x^1 = (0{,}5;1)^T - 0{,}24(3;2{,}5)^T = (-0{,}22;0{,}4)^T$.

8^0. Вычислим $\left\|x^1 - x^0\right\|$: $\left\|x^1 - x^0\right\| = 0{,}937 > 0{,}15$. Вычислим $\left|f\left(x^1\right) - f\left(x^0\right)\right|$: $\left|f\left(x^1\right) - f\left(x^0\right)\right| = 1{,}83 > 0{,}15$. Вывод: положим $k = 1$ и перейдем к шагу 3.

3^1. Вычислим $\nabla f\left(x^1\right)$: $\nabla f\left(x^1\right) = (-0{,}48;0{,}58)^T$.

4^1. Вычислим $\left\|\nabla f\left(x^1\right)\right\| = 0{,}752 > 0{,}1$.

5^1. Проверим условие $k \geq M$: $k = 1 < 10 = M$.

6^1. Определим t_1^*: $t_1^* = 0{,}546$ (см. п. 6^0).

7^1. Найдем $x^2 = x^1 - t_1^* \nabla f\left(x^1\right)$:

$$x^2 = (-0{,}22;0{,}4)^T - 0{,}546\,(-0{,}48;0{,}58)^T = (0{,}04;0{,}08)^T.$$

8^1. Вычислим $\left\|x^2 - x^1\right\|$, $\left|f\left(x^2\right) - f\left(x^1\right)\right|$:

$$\left\|x^2 - x^1\right\| = 0{,}41 > 0{,}15; \quad \left|f\left(x^2\right) - f\left(x^1\right)\right| = 0{,}156 > 0{,}15.$$

Положим $k = 2$ и перейдем к шагу 3.

3^2. Вычислим $\nabla f\left(x^2\right)$: $\nabla f\left(x^2\right) = (0{,}24;0{,}2)^T$.

4^2. Вычислим $\left\|\nabla f\left(x^2\right)\right\|$: $\left\|\nabla f\left(x^2\right)\right\| = 0{,}312 > 0{,}1$.

5^2. Проверим условие $k \geq M$: $k = 2 < 10 = M$.

6^2. Определим t_2^*: $t_2^* = 0{,}24$ (см. п. 6^0).

7^2. Найдем $x^3 = x^2 - t_2^* \nabla f\left(x^2\right)$:

$$x^3 = (0{,}04;0{,}08)^T - 0{,}24\,(0{,}24;0{,}2)^T = (-0{,}0176;0{,}032)^T.$$

8^2. Вычислим $\left\|x^3 - x^2\right\|$, $\left|f\left(x^3\right) - f\left(x^2\right)\right|$:

$$\left\|x^3 - x^2\right\| = 0{,}0749 < 0{,}15; \quad \left|f\left(x^3\right) - f\left(x^2\right)\right| = 0{,}0116 < 0{,}15.$$

Положим $k = 3$ и перейдем к шагу 3.

3^3. Вычислим $\nabla f\left(x^3\right)$: $\nabla f\left(x^3\right) = (-0{,}012;-0{,}0816)^T$.

4^3. Вычислим $\left\|\nabla f\left(x^3\right)\right\|$: $\left\|\nabla f\left(x^3\right)\right\| = 0{,}082 < 0{,}1$. Расчет окончен. Найдена точка $x^3 = (-0{,}0176;0{,}032)^T$, $f\left(x^3\right) = 0{,}00127$. На рис. 6.2 полученные точки выделены и соединены сплошной линией.

II. Проведем анализ точки x^3.

В примере 6.1 было показано, что функция $f(x)$ строго выпуклая, следовательно, точка x^3 является найденным приближением точки глобального минимума x^*. ■

6.3. МЕТОД ПОКООРДИНАТНОГО СПУСКА

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия решения задачи состоит в построении последовательности точек $\{x^k\}$, $k = 0,1,\ldots$, таких, что $f(x^{k+1}) < f(x^k)$, $k = 0,1,\ldots$. Точки последовательности $\{x^k\}$ вычисляются по циклам в соответствии с правилом

$$x^{jk+1} = x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1}, \tag{6.5}$$

где j – номер цикла вычислений; $j = 0,1,2,\ldots$; k – номер итерации внутри цикла, $k = 0,1,\ldots,n-1$; e_{k+1}, $k = 0,1,\ldots,n-1$ – единичный вектор, $(k+1)$-я проекция которого равна 1; точка x^{00} задается пользователем; величина шага t_k выбирается из условия

$$f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1} \right) - f\left(x^{jk}\right) < 0 \text{ или } f\left(x^{jk+1}\right) - f\left(x^{jk}\right) < -\varepsilon \left\| \nabla f\left(x^{jk}\right) \right\|^2 .$$

Если выбранное условие при текущем t_k не выполняется, шаг уменьшается вдвое и точка $x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1}$ вычисляется заново. Легко видеть, что при фиксированном j за одну итерацию с номером k изменяется только одна проекция точки x^{jk}, имеющая номер $k+1$, а в течение всего цикла с номером j, т.е. начиная с $k = 0$ и кончая $k = n-1$, изменяются все n проекций точки x^{j0}. После этого точке x^{jn} присваивается номер $x^{j+1,0}$ и она берется за начальную точку для вычислений в $(j+1)$-м цикле. Расчет заканчивается в точке x^{jk} при выполнении по крайней мере одного из трех критериев окончания счета: $\left\| \nabla f\left(x^{jk}\right) \right\| < \varepsilon_1$ или $j \ge M$, или двукратного выполнения неравенств $\left\| x^{jk+1} - x^{jk} \right\| < \varepsilon_2$, $\left| f\left(x^{jk+1}\right) - f\left(x^{jk}\right) \right| < \varepsilon_2$.

Полученные в результате вычислений точки могут быть записаны как элементы последовательности $\left\{x^l\right\}$, где $l = n \cdot j + k$ – порядковый номер точки, т.е. $\left\{x^l\right\} = \left\{x^0 = x^{00}, x^1 = x^{01}, \ldots, x^n = x^{0n} = x^{10}, x^{n+1} = x^{11}, x^{n+2} = x^{12}, \ldots\right\}$.

Алгоритм

Шаг 1. Задать x^{00}, $\varepsilon > 0, \varepsilon_1 > 0, \varepsilon_2 > 0$, предельное число *M* циклов счета, кратное *n*, где *n* – размерность вектора x. Найти градиент $\nabla f(x)$.

Шаг 2. Задать номер цикла $j = 0$.

Шаг 3. Проверить условие $j \ge M$:

а) если $j \ge M$, то расчет окончен и $x^* = x^{jk}$;
б) если нет, то перейти к шагу 4.

Шаг 4. Задать $k = 0$.

Шаг 5. Проверить условие $k \le n - 1$:
а) если $k \le n - 1$, то перейти к шагу 6;
б) если $k = n$, то положить $j = j + 1$, $x^{j+1,k} = x^{jn}$ и перейти к шагу 3.

Шаг 6. Вычислить $\nabla f\left(x^{jk}\right)$.

Шаг 7. Проверить выполнение критерия окончания $\left\| \nabla f\left(x^{jk}\right) \right\| < \varepsilon_1$:

а) если критерий выполнен, то расчет окончен и $x^* = x^{jk}$;
б) если нет, то перейти к шагу 8.

Шаг 8. Задать t_k.

Шаг 9. Вычислить точку x^{jk+1}: $x^{jk+1} = x^{jk} - t_k \left(\dfrac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1}$.

Шаг 10. Проверить выполнение условия

$$f\left(x^{jk+1}\right) - f\left(x^{jk}\right) < 0 \quad \left(\text{или } f\left(x^{jk+1}\right) - f\left(x^{jk}\right) < -\varepsilon \left\| \nabla f\left(x^{jk}\right) \right\|^2\right).$$

а) если условие выполнено, то перейти к шагу 11;

б) если нет, то положить $t_k = \dfrac{t_k}{2}$ и перейти к шагу 9.

Шаг 11. Проверить выполнение условий

$$\left\| x^{jk+1} - x^{jk} \right\| < \varepsilon_2, \qquad \left| f\left(x^{jk+1}\right) - f\left(x^{jk}\right) \right| < \varepsilon_2:$$

а) если в двух последовательных циклах с номерами j и $j - 1$ оба условия выполняются, то расчет в точке x^{jk+1} окончен и $x^* = x^{jk+1}$;
б) если хотя бы одно из условий не выполнено, положить $k = k + 1$ и перейти к шагу 5.

Геометрическая интерпретация метода для $n = 2$ приведена на рис. 6.4.

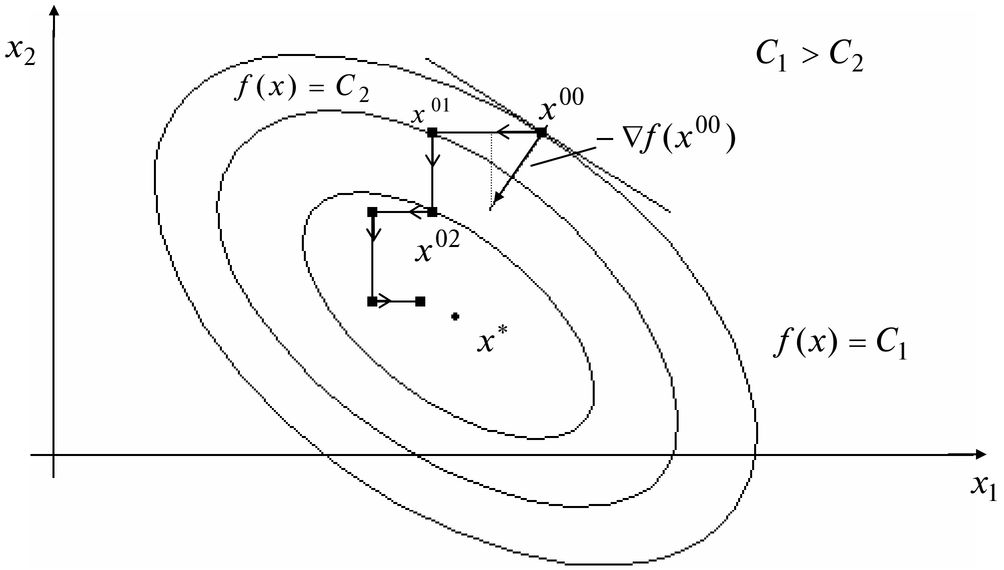

Рис. 6.4

З а м е ч а н и я 6.5.

1. Если функция $f(x)$ удовлетворяет условиям утверждения 6.1, то построение последовательности $\{x^k\}$ по методу покоординатного спуска обеспечивает выполнение условия $\left\| \nabla f(x^k) \right\| \to 0$ при $k \to \infty$ [28].

2. Найденная в результате применения метода точка x^* нуждается в дополнительном исследовании с целью ее классификации.

3. Скорость сходимости метода оценивается как линейная (см. гл. 4).

4. Относительно процедуры решения задачи и поиска глобального минимума справедливо замечание 6.4.

Пример 6.3. Найти локальный минимум функции

$$f(x) = 2x_1^2 + x_1 x_2 + x_2^2 .$$

□ I. Определим точку x^{jk}, в которой выполнен по крайней мере один из критериев окончания расчетов.

1. Зададим $x^{00}, \varepsilon, \varepsilon_1, \varepsilon_2, M$: $x^{00} = (0{,}5;1)^T, \varepsilon = 0, \varepsilon_1 = 0{,}1; \varepsilon_2 = 0{,}15; M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Зададим $j = 0$.

3^0. Проверим выполнение условия $j \ge M$: $j = 0 < 10 = M$.

4^0. Зададим $k = 0$.

5^0. Проверим выполнение условия $k \le n-1$: $k = 0 < 1 = n-1$.

6^0. Вычислим $\nabla f(x^{00})$: $\nabla f(x^{00}) = (3; 2{,}5)^T$.

7^0. Проверим условие $\left\| \nabla f(x^{00}) \right\| < \varepsilon_1$: $\left\| \nabla f(x^{00}) \right\| = 3{,}8 > 0{,}1$.

8^0. Зададим $t_0 = 0{,}5$.

9^0. Вычислим $x^{01} = x^{00} - t_0 \left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{00}} \cdot e_1$, где

$\frac{\partial f\left(x^{00}\right)}{\partial x_1} = \left(4x_1 + x_2\right)\Big|_{x^{00}} = 2 + 1 = 3$, $e_1 = (1,0)^T$. Отсюда $x^{01} = (-1;1)^T$.

10^0. Проверим условие $f\left(x^{01}\right) - f\left(x^{00}\right) < 0$: $f\left(x^{01}\right) - f\left(x^{00}\right) = 2 - 2 = 0$.

Вывод: положим $t_0 = 0{,}25$ и перейдем к шагу 9.

9^{01}. Вычислим x^{01} с шагом $t_0 = 0{,}25$: $x^{01} = (-0{,}25;1)^T$.

10^{01}. Проверим условие $f\left(x^{01}\right) - f\left(x^{00}\right) < 0$:

$$f\left(x^{01}\right) - f\left(x^{00}\right) = 0{,}875 - 2 = -1{,}125 < 0.$$

11^0. Проверим условия $\left\| x^{01} - x^{00} \right\| < \varepsilon_2$, $\left| f\left(x^{01}\right) - f\left(x^{00}\right) \right| < \varepsilon_2$:

$$\left\| x^{01} - x^{00} \right\| = 0{,}75 > 0{,}15, \quad \left| f\left(x^{01}\right) - f\left(x^{00}\right) \right| = 1{,}125 > 0{,}15.$$

Положим $k = 1$ и перейдем к шагу 5.

5^1. Проверим условие $k \le n - 1$: $k = 1 = n - 1$.

6^1. Вычислим $\nabla f\left(x^{01}\right)$: $\nabla f\left(x^{01}\right) = (0;1{,}75)^T$.

7^1. Проверим условие $\left\| \nabla f\left(x^{01}\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{01}\right) \right\| = 1{,}75 > 0{,}1$.

8^1. Зададим $t_1 = 0{,}5$.

9^1. Вычислим $x^{02} = x^{01} - t_1 \left(\frac{\partial f(x)}{\partial x_2} \right)_{x = x^{01}} \cdot e_2$, где $e_2 = (0,1)^T$;

$\frac{\partial f\left(x^{01}\right)}{\partial x_2} = \left(x_1 + 2x_2\right)\Big|_{x^{01}} = -0{,}25 + 2 = 1{,}75$. Отсюда $x^{02} = (-0{,}25;0{,}125)^T$.

10^1. Проверим условие $f\left(x^{02}\right) - f\left(x^{01}\right) < 0$:

$$f\left(x^{02}\right) - f\left(x^{01}\right) = 0{,}109 - 0{,}875 = -0{,}766 < 0.$$

11^1. Проверим условия $\left\| x^{02} - x^{01} \right\| < \varepsilon_2$, $\left| f\left(x^{02}\right) - f\left(x^{01}\right) \right| < \varepsilon_2$:

$$\left\| x^{02} - x^{01} \right\| = 0{,}875 > 0{,}15, \quad \left| f\left(x^{02}\right) - f\left(x^{01}\right) \right| = 0{,}766 > 0{,}15.$$

Положим $k = 2$ и перейдем к шагу 5.

5^2. Проверим условие $k \le n - 1$: $k = 2 > n - 1$. Зададим $j = 1, x^{10} = x^{02}$, перейдем к шагу 3.

3^1. Проверим условие $j \ge M$: $j = 1 < 10 = M$.

4^1. Зададим $k = 0$.

5^2. Проверим условие $k \le n - 1$: $k = 0 < 1 = n - 1$.

6^2. Вычислим $\nabla f\left(x^{10}\right)$: $\nabla f\left(x^{10}\right) = \nabla f\left(x^{02}\right) = (-0{,}875; 0{,}00)^T$.

7^2. Проверим условие $\left\| \nabla f\left(x^{10}\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{10}\right) \right\| = 0{,}875 > 0{,}1$.

8^2. Зададим $t_0 = 0{,}25$.

9^2. Вычислим $x^{11} = x^{10} - t_0 \left(\dfrac{\partial f(x)}{\partial x_1} \right)_{x = x^{10}} \cdot e_1$: $x^{11} = (-0{,}03; 0{,}125)^T$.

10^2. Проверим условие $f\left(x^{11}\right) - f\left(x^{10}\right) < 0$:

$$f\left(x^{11}\right) - f\left(x^{10}\right) = 0{,}01 - 0{,}109 = -0{,}099 < 0.$$

11^2. Проверим условия $\left\| x^{11} - x^{10} \right\| < \varepsilon_2$, $\left| f\left(x^{11}\right) - f\left(x^{10}\right) \right| < \varepsilon_2$:

$$\left\| x^{11} - x^{10} \right\| = 0{,}22 > 0{,}15, \quad \left| f\left(x^{11}\right) - f\left(x^{10}\right) \right| = 0{,}099 < 0{,}15.$$

Положим $k = 1$ и перейдем к шагу 5.

5^3. Проверим условие $k \le n - 1$: $k = 1 = n - 1$.

6^3. Вычислим $\nabla f\left(x^{11}\right)$: $\nabla f\left(x^{11}\right) = (0{,}005; 0{,}22)^T$.

7^3. Проверим условия $\left\| \nabla f\left(x^{11}\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{11}\right) \right\| = 0{,}22 > 0{,}1$.

8^3. Зададим $t_1 = 0{,}25$.

9^3. Вычислим $x^{12} = x^{11} - t_0 \left(\dfrac{\partial f(x)}{\partial x_2} \right)_{x = x^{11}} \cdot e_2$: $x^{12} = (-0{,}03; 0{,}07)^T$.

10^3. Проверим условие $f\left(x^{12}\right) - f\left(x^{11}\right) < 0$:

$$f\left(x^{12}\right) - f\left(x^{11}\right) = 0{,}0046 - 0{,}01 = -0{,}0054 < 0.$$

11^3. Проверим условия $\left\| x^{12} - x^{11} \right\| < \varepsilon_2$, $\left| f\left(x^{12}\right) - f\left(x^{11}\right) \right| < \varepsilon_2$:

$$\left\| x^{12} - x^{11} \right\| = 0{,}055 < 0{,}15, \quad \left| f\left(x^{12}\right) - f\left(x^{11}\right) \right| = 0{,}0054 < 0{,}15.$$

Зададим $k = 2$ и перейдем к шагу 5.

5^4. Проверим условие $k \le n - 1$: $k = 2 > n - 1$. Положим $j = 2, x^{20} = x^{12}$ и перейдем к шагу 3.

3^2. Проверим условие $j \ge M$: $j = 2 < 10 = M$.

4^2. Зададим $k = 0$.

5^4. Проверим условие $k \le n - 1$: $k = 0 < 1 = n - 1$.

6^4. Вычислим $\nabla f\left(x^{20}\right)$: $\nabla f\left(x^{20}\right) = \nabla f\left(x^{12}\right) = (-0{,}05; 0{,}11)^T$.

7^4. Проверим условие $\left\| \nabla f\left(x^{20}\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{20}\right) \right\| = 0{,}12 > 0{,}1$.

8^4. Зададим $t_0 = 0{,}25$.

9^4. Вычислим $x^{21} = x^{20} - t_0 \left(\frac{\partial f(x)}{\partial x_1} \right)_{x=x^{20}} \cdot e_1$: $x^{21} = (-0{,}02; 0{,}07)^T$.

10^4. Проверим условие $f\left(x^{21}\right) - f\left(x^{20}\right) < 0$: $0{,}0043 - 0{,}046 = -0{,}0003 < 0$, перейдем к шагу 11.

11^4. Проверим условия $\left\|x^{21} - x^{20}\right\| < \varepsilon_2$, $\left| f\left(x^{21}\right) - f\left(x^{20}\right)\right| < \varepsilon_2$:

$$\left\|x^{21} - x^{20}\right\| = 0{,}01 < 0{,}15, \qquad \left| f\left(x^{21}\right) - f\left(x^{20}\right)\right| = 0{,}0003 < 0{,}15.$$

Условия $\left\|x^{jk+1} - x^{jk}\right\| < \varepsilon_2$, $\left| f\left(x^{jk+1}\right) - f\left(x^{jk}\right)\right| < \varepsilon_2$ выполнены в двух последовательных циклах с номерами $j = 2$ и $j - 1 = 1$. Расчет окончен, найдена точка $x^{21} = (-0{,}02; 0{,}07)^T$; $f\left(x^{21}\right) = 0{,}0043$.

На рис. 6.5 полученные точки соединены штриховой линией.

II. Проведем анализ точки x^{21}.

В примере 6.1 было показано, что функция $f(x)$ строго выпукла, имеет единственный минимум и, следовательно, точка $x^{21} = (-0{,}02; 0{,}07)^T$ является найденным приближением точки глобального минимума. ■

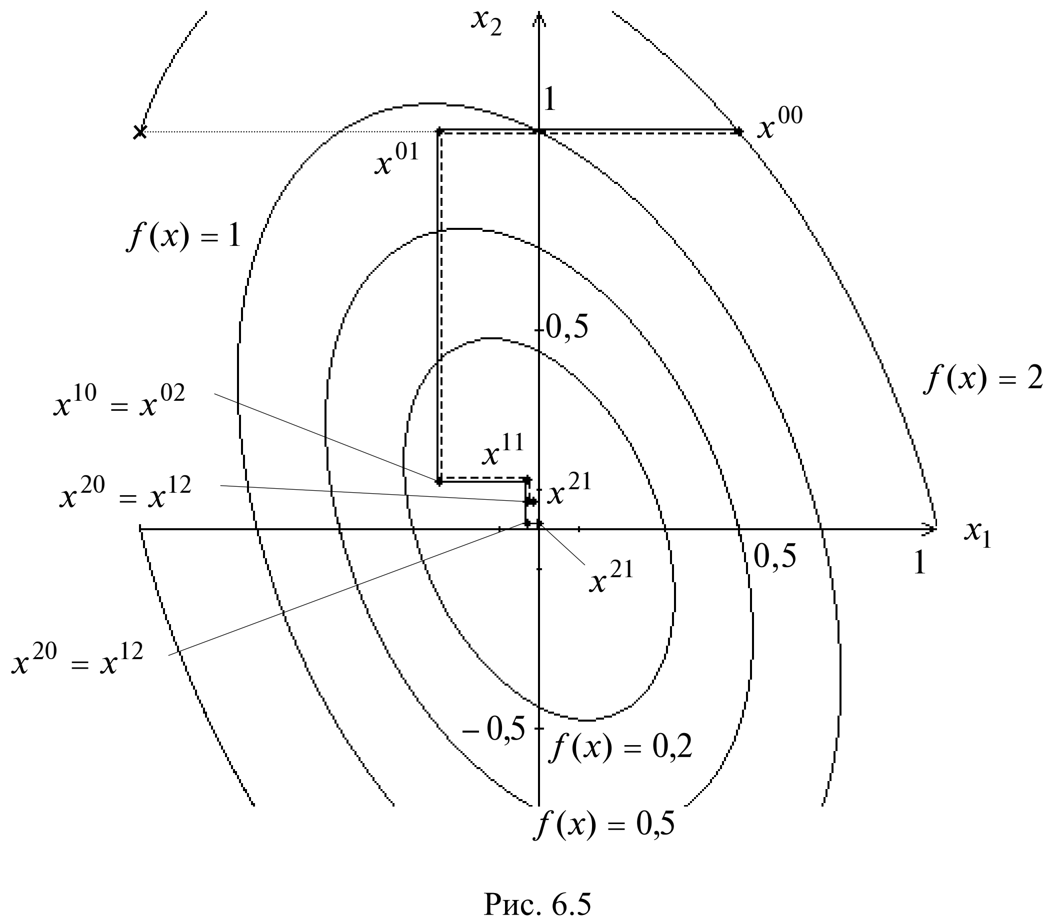

Рис. 6.5

6.4. МЕТОД ГАУССА–ЗЕЙДЕЛЯ

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия метода Гаусса–Зейделя (Gauss–Seidel) состоит в построении последовательности точек $\left\{x^k\right\}$, $k = 0,1,\ldots$, таких, что $f\left(x^{k+1}\right) < f\left(x^k\right)$, $k = 0,1,\ldots$. Точки последовательности $\left\{x^k\right\}$ вычисляются по правилу

$$x^{jk+1} = x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}}\right)_{x = x^{jk}} \cdot e_{k+1}, \tag{6.6}$$

где j – номер цикла вычислений, $j = 0,1,2,\ldots$; k – номер итерации внутри цикла, $k = 0,1,\ldots,n-1$; e_{k+1} – единичный вектор, $(k+1)$-я проекция которого равна 1; точка x^{00} задается пользователем, величина шага t_k выбирается из условия

$$\varphi(t_k) = f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}}\right)_{x = x^{jk}} \cdot e_{k+1}\right) \to \min_{t_k}.$$

Данная задача является задачей одномерной минимизации функции $\varphi(t_k) = f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}}\right)_{x = x^{jk}} \cdot e_{k+1}\right)$ и может быть решена либо с использованием условий $\frac{d\varphi}{dt_k} = 0$, $\frac{d^2\varphi}{dt_k{}^2} > 0$, либо численно с использованием методов одномерной минимизации, как задача $\varphi(t_k) \to \min\limits_{t_k \in [a,b]}$ (см. разд. 5.1).

Если уравнение $\frac{d\varphi}{dt_k} = 0$ имеет высокую степень и корни его трудно определить, можно аппроксимировать функцию $\varphi(t_k)$ полиномом $P(t_k)$ второй или третьей степени и определить t_k^* из условий $\frac{dP}{dt_k} = 0$, $\frac{d^2 P}{dt_k{}^2} > 0$.

При численном решении задачи определения величины шага степень близости найденного значения t_k к оптимальному значению t_k^*, удовлетворяющему условиям

$\frac{d\varphi}{dt_k} = 0$, $\frac{d^2\varphi}{dt_k^2} > 0$, зависит от задания интервала $[a, b]$ и точности методов одномерной минимизации [27].

Легко видеть, что при фиксированном j за одну итерацию с номером k изменяется только одна проекция точки x^{jk}, имеющая номер $k+1$, а в течение всего цикла с номером j, т.е. начиная с $k = 0$ и кончая $k = n-1$, изменяются все n проекций точки x^{j0}. После этого точке x^{jn} присваивается номер $x^{j+1,0}$ и она берется за начальную точку для вычислений в $(j+1)$-м цикле.

Расчет заканчивается в точке x^{jk} при выполнении по крайней мере одного из трех критериев окончания счета: $\left\| \nabla f\left(x^{jk}\right) \right\| < \varepsilon_1$ или $k \geq M$, или двукратного выполнения неравенств $\left\| x^{jk+1} - x^{jk} \right\| < \varepsilon_2$, $\left| f\left(x^{jk+1}\right) - f\left(x^{jk}\right) \right| < \varepsilon_2$. Здесь $\varepsilon_1, \varepsilon_2$ – малые положительные числа, M - предельное число циклов итераций.

Полученные в результате вычислений точки могут быть записаны как элементы последовательности $\left\{x^l\right\}$, где $l = n \cdot j + k$ – порядковый номер точки, т.е. $\left\{x^l\right\} = \left\{x^0 = x^{00}, x^1 = x^{01}, \ldots, x^n = x^{0n} = x^{10}, x^{n+1} = x^{11}, x^{n+2} = x^{12}, \ldots\right\}$.

Алгоритм

Шаг 1. Задать x^{00}, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$; предельное число M циклов счета, кратное n, где n – размерность вектора x. Найти градиент $\nabla f(x)$.

Шаг 2. Задать номер цикла $j = 0$.

Шаг 3. Проверить условие $j \geq M$:

а) если $j \geq M$, то расчет окончен и $x^* = x^{jk}$;

б) если $j < M$, то перейти к шагу 4.

Шаг 4. Задать $k = 0$.

Шаг 5. Проверить условие $k \leq n-1$:

а) если $k \leq n-1$, то перейти к шагу 6;

б) если $k = n$, то положить $j = j+1$ и перейти к шагу 3.

Шаг 6. Вычислить $\nabla f\left(x^{jk}\right)$.

Шаг 7. Проверить выполнение условия $\left\| \nabla f\left(x^{jk}\right) \right\| < \varepsilon_1$:

а) если условие выполнено, то расчет окончен и $x^* = x^{jk}$;

б) если нет, то перейти к шагу 8.

Шаг 8. Вычислить t_k^* из условия

$$\varphi(t_k) = f\left(x^{jk} - t_k \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1} \right) \to \min_{t_k}.$$

Шаг 9. Вычислить $x^{jk+1} = x^{jk} - t_k^* \left(\frac{\partial f(x)}{\partial x_{k+1}} \right)_{x = x^{jk}} \cdot e_{k+1}$.

Шаг 10. Проверить выполнение условий

$$\left\| x^{jk+1} - x^{jk} \right\| < \varepsilon_2, \quad \left| f\left(x^{jk+1}\right) - f\left(x^{jk}\right) \right| < \varepsilon_2:$$

а) если оба условия выполнены в двух последовательных циклах с номерами j и $j-1$, то расчет окончен, найдена точка $x^* = x^{jk+1}$;

б) если не выполняется хотя бы одно условие, положить $k = k+1$ и перейти к шагу 5.

Геометрическая интерпретация метода для $n = 2$ приведена на рис. 6.6.

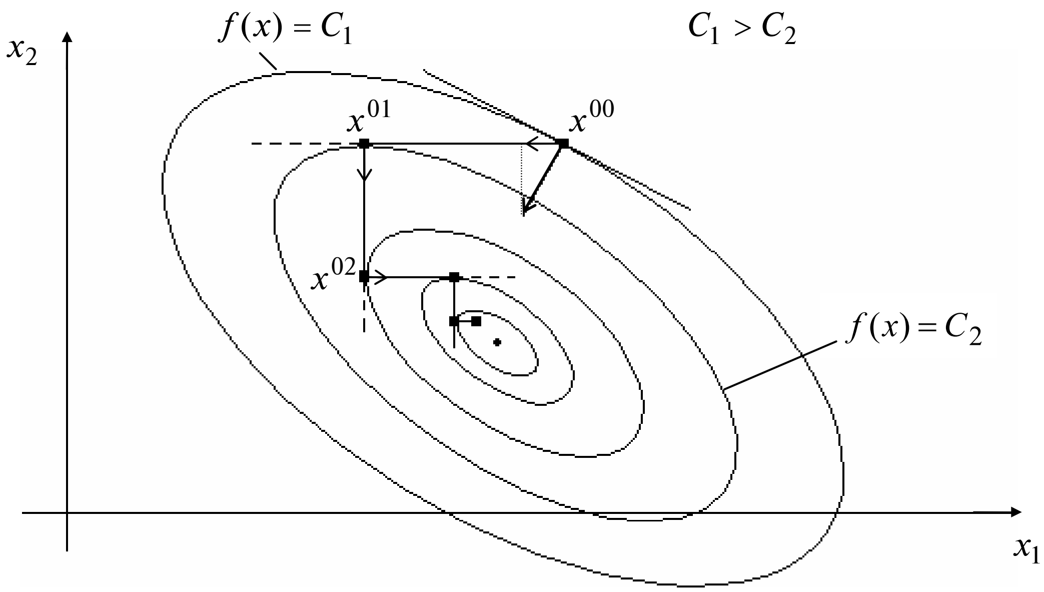

Рис. 6.6

З а м е ч а н и я 6.6.

1. Относительно свойств последовательности $\left\{x^k\right\}$, $k = 0,1,\ldots$, полученной по методу Гаусса–Зейделя, справедлив п.1 замечаний 6.5.

2. Процедура решения задачи совпадает с описанной в разд. 6.1.

Пример 6.4. Найти локальный минимум функции

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

☐ I. Определим точку x^{jk}, в которой выполнен хотя бы один из критериев окончания расчетов.

1. Зададим $x^{00}, \varepsilon_1, \varepsilon_2, M$: $x^{00} = (0{,}5;1)^T$, $\varepsilon_1 = 0{,}1$; $\varepsilon_2 = 0{,}15$; $M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Зададим $j = 0$.

3^0. Проверим выполнение условия $j \ge M$: $j = 0 < 10 = M$.

4^0. Зададим $k = 0$.

5^0. Проверим выполнение условия $k \le n-1$: $k = 0 < 1 = n-1$.

6^0. Вычислим $\nabla f\left(x^{00}\right)$: $\nabla f\left(x^{00}\right) = (3; 2{,}5)^T$.

7^0. Проверим условие $\left\| \nabla f\left(x^{00}\right)\right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{00}\right)\right\| = 3{,}9 > 0{,}1$.

8^0. Определим величину шага t_0^* из условия

$$\varphi(t_0) = f\left(x^{j0} - t_0 \left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{j0}} \cdot e_1 \right) \to \min_{t_0}.$$

Воспользуемся формулой (6.6) при $k = 0, j = 0$: $x^{01} = x^{00} - t_0 \left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{00}} \cdot e_1$.

Поскольку $\left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{00}} = 4x_1 + x_2 \Big|_{x = x^{00}} = 2 + 1 = 3$, $e_1 = (1;\, 0)^T$, то $x^{01} = (0{,}5;\, 1)^T - t_0 \cdot 3 \cdot (1;\, 0)^T = (0{,}5 - 3t_0;\, 1)^T$ или $x_1^{01} = 0{,}5 - 3t_0, x_2^{01} = 1$. Подставляя полученные выражения в $f(x)$, имеем $\varphi(t_0) = 2(0{,}5 - 3t_0)^2 + (0{,}5 - 3t_0) \cdot 1 + 1$.

Из необходимого условия экстремума $\frac{d\varphi(t_0)}{dt_0} = 4 \cdot (0{,}5 - 3t_0) \cdot (-3) - 3 = 0$ или $36t_0 - 9 = 0$ находим $t_0^* = \frac{1}{4}$. Так как $\frac{d^2\varphi(t_0)}{dt_0^{\,2}} = 36 > 0$, то найденное значение шага обеспечивает минимум функции $\varphi(t_0)$ по t_0.

Можно показать, что в силу структуры заданной функции $f(x)$ величина шага в направлении $-\left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{jk}} \cdot e_1$ не зависит от x^k, является постоянной и равной $\frac{1}{4}$.

9^0. Определим $x^{01} = x^{00} - t_0^* \left(\frac{\partial f(x)}{\partial x_1} \right)_{x = x^{00}} \cdot e_1$: $x^{01} = \left(-0{,}25;1\right)^T$.

10^0. Проверим условия $\left\| x^{01} - x^{00} \right\| < \varepsilon_2$, $\left| f\left(x^{01}\right) - f\left(x^{00}\right) \right| < \varepsilon_2$:

$$\left\| x^{01} - x^{00} \right\| = 0{,}25 > 0{,}15, \quad \left| f\left(x^{01}\right) - f\left(x^{00}\right) \right| = |\, 0{,}875 - 2 \,| = 1{,}125 > 0{,}15.$$

Положим $k = 1$ и перейдем к шагу 5.

5^1. Проверим условие $k \le n - 1$: $k = 1 = n - 1$.

6^1. Вычислим $\nabla f\left(x^{01}\right)$: $\nabla f\left(x^{01}\right) = \left(0;1{,}75\right)^T$.

7^1. Проверим условие $\left\| \nabla f\left(x^{01}\right)\right\| < \varepsilon_1$: $\left\| \nabla f\left(x^{01}\right)\right\| = 1{,}75 > 0{,}1$.

8^1. Определим величину шага t_1^* из условия

$$\varphi(t_1) = f\left(x^{j1} - t_1 \left(\frac{\partial f(x)}{\partial x_2} \right)_{x = x^{j1}} \cdot e_2 \right) \to \min_{t_1}.$$

Воспользуемся формулой (6.6) при $k = 1, j = 0$: $x^{02} = x^{01} - t_1 \left(\frac{\partial f(x)}{\partial x_2} \right)_{x = x^{01}} \cdot e_2$.

Поскольку $\left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{01}} = x_1 + 2x_2\Big|_{x=x^{01}} = -0{,}25 + 2 = 1{,}75$, $e_2 = (0;\, 1)^T$, то

$x^{02} = (-0{,}25;\, 1)^T - t_1 \cdot 1{,}75 \cdot (0;\, 1)^T = (-0{,}25;\, 1 - 1{,}75 \cdot t_1)^T$ или $x_1^{02} = -0{,}25; x_2^{02} = 1 - 1{,}75 \cdot t_1$.
Подставляя полученные выражения в $f(x)$, имеем

$$\varphi(t_1) = 2(-0{,}25)^2 + (-0{,}25) \cdot (1 - 1{,}75 \cdot t_1) + (1 - 1{,}75 \cdot t_1)^2 .$$

Из необходимого условия экстремума

$$\frac{d\varphi(t_1)}{dt_1} = 0{,}25 \cdot 1{,}75 + 2 \cdot (1 - 1{,}75 \cdot t_1) \cdot (-1{,}75) = 0 \quad \text{или} \quad 2 \cdot 1{,}75^2 t_1 - 1{,}75^2 = 0$$

находим $t_1^* = \frac{1}{2}$. Так как $\frac{d^2\varphi(t_1)}{dt_1{}^2} = 2 \cdot 1{,}75^2 > 0$, то найденное значение шага обеспечивает минимум функции $\varphi(t_1)$ по t_1.

Можно показать, что в силу структуры функции $f(x)$ величина шага в направлении $-\left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{jk}} \cdot e_2$ остается постоянной и равной $\frac{1}{2}$.

9^1. Вычислим $x^{02} = x^{01} - t_1^* \left(\frac{\partial f(x)}{\partial x_2}\right)_{x=x^{01}} \cdot e_2$: $x^{02} = (-0{,}25; 0{,}125)^T$.

10^1. Проверим условия $\left\| x^{02} - x^{01} \right\| < \varepsilon_2$, $\left| f(x^{02}) - f(x^{01}) \right| < \varepsilon_2$:

$\left\| x^{02} - x^{01} \right\| = 0{,}875 > 0{,}15$, $\left| f(x^{02}) - f(x^{01}) \right| = |\, 0{,}12 - 0{,}875 \,| = 0{,}755 > 0{,}15$.

Положим $k = 2$ и перейдем к шагу 5.

5^2. Проверим условие $k \le n - 1$: $k = 2 = n$. Положим $j = 1, x^{10} = x^{02}$ и перейдем к шагу 3.

3^1. Проверим условие $j \ge M$: $j = 1 < 10 = M$.

4^1. Зададим $k = 0$.

5^3. Проверим условие $k \le n - 1$: $k = 0 < 1 = n - 1$.

6^3. Вычислим $\nabla f(x^{10})$: $\nabla f(x^{10}) = \nabla f(x^{02}) = (-0{,}875; 0{,}00)^T$.

7^3. Проверим условие $\left\| \nabla f(x^{10}) \right\| < \varepsilon_1$: $\left\| \nabla f(x^{10}) \right\| = 0{,}875 > 0{,}1$.

8^3. Полагаем $t_0^* = 0{,}25$ (см. п. 8^0).

9^3. Вычислим $x^{11} = x^{10} - t_0^* \left(\frac{\partial f(x)}{\partial x_1}\right)_{x=x^{10}} \cdot e_1$: $x^{11} = (-0{,}03; 0{,}125)^T$.

10^3. Проверим условия $\left\| x^{11} - x^{10} \right\| < \varepsilon_2$, $\left| f(x^{11}) - f(x^{10}) \right| < \varepsilon_2$:

$\left\| x^{11} - x^{10} \right\| = 0{,}22 > 0{,}15$, $\left| f(x^{11}) - f(x^{10}) \right| = |\, 0{,}013 - 0{,}1375 \,| = 0{,}124 < 0{,}15$.

Положим $k = 1$ и перейдем к шагу 5.

5^4. Проверим условие $k \le n-1$: $k = 1 = n-1$.

6^4. Вычислим $\nabla f\left(x^{11}\right)$: $\nabla f\left(x^{11}\right) = (0{,}005; 0{,}22)^T$.

7^4. Проверим условие $\left\|\nabla f\left(x^{11}\right)\right\| < \varepsilon_1$: $\left\|\nabla f\left(x^{11}\right)\right\| = 0{,}22 > 0{,}1$.

8^4. Зададим $t_1^* = 0{,}5$ (см. п. 8^1).

9^4. Вычислим $x^{12} = x^{11} - t_1^* \left(\dfrac{\partial f(x)}{\partial x_2}\right)_{x=x^{11}} \cdot e_2$: $x^{12} = (-0{,}03; 0{,}015)^T$.

10^4. Проверим условия $\left\|x^{12} - x^{11}\right\| < \varepsilon_2$, $\left|f\left(x^{12}\right) - f\left(x^{11}\right)\right| < \varepsilon_2$:

$$\left\|x^{12} - x^{11}\right\| = 0{,}11 < 0{,}15, \quad \left|f\left(x^{12}\right) - f\left(x^{11}\right)\right| = \left|0{,}0015 - 0{,}013\right| = 0{,}0115 < 0{,}15.$$

Положим $k = 2$ и перейдем к шагу 5.

5^5. Проверим условие $k \le n-1$: $k = 2 = n$. Положим $j = 2$, $x^{20} = x^{12}$ и перейдем к шагу 3.

3^2. Проверим условие $j \ge M$: $j = 2 < 10 = M$.

4^2. Зададим $k = 0$.

5^6. Проверим условие $k \le n-1$: $k = 0 < 1 = n-1$.

6^5. Вычислим $\nabla f\left(x^{20}\right)$: $\nabla f\left(x^{20}\right) = (-0{,}105; 0)^T$.

7^5. Проверим условие $\left\|\nabla f\left(x^{20}\right)\right\| < \varepsilon_1$: $\left\|\nabla f\left(x^{20}\right)\right\| = 0{,}105 > \varepsilon_1$.

8^5. Зададим $t_0^* = 0{,}25$ (см. п. 8^0).

9^5. Вычислим $x^{21} = x^{20} - t_0^* \left(\dfrac{\partial f(x)}{\partial x_1}\right)_{x=x^{20}} \cdot e_1$: $x^{21} = (-0{,}004; 0{,}015)^T$.

10^5. Проверим условия $\left\|x^{21} - x^{20}\right\| < \varepsilon_2$, $\left|f\left(x^{21}\right) - f\left(x^{20}\right)\right| < \varepsilon_2$:

$$\left\|x^{21} - x^{20}\right\| = 0{,}026 < 0{,}15, \quad \left|f\left(x^{21}\right) - f\left(x^{20}\right)\right| = \left|0{,}000197 - 0{,}0015\right| = 0{,}0013 < 0{,}15.$$

Условия $\left\|x^{jk+1} - x^{jk}\right\| < \varepsilon_2$, $\left|f\left(x^{jk+1}\right) - f\left(x^{jk}\right)\right| < \varepsilon_2$ выполнены в двух последовательных циклах с номерами $j = 2$ и $j - 1 = 1$. Расчет окончен, найдена точка $x^{21} = (-0{,}004; 0{,}015)^T$; $f\left(x^{21}\right) = 0{,}000197$. На рис. 6.5 полученные точки последовательности $x^{00} \to x^{01} \to x^{02} = x^{10} \to x^{11} \to x^{12} = x^{20} \to x^{21}$ соединены сплошной линией. Очевидно, метод Гаусса–Зейделя сходится быстрее, чем метод покоординатного спуска.

II. Проведем анализ точки x^{21}.

Точка x^{21} является найденным приближением точки глобального минимума $f(x)$, так как функция $f(x)$ строго выпуклая (см. пример 6.1). ■

6.5. МЕТОД ФЛЕТЧЕРА–РИВСА

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия метода Флетчера–Ривса (Fletcher–Reeves) состоит в построении последовательности точек $\left\{x^k\right\}$, $k = 0,1,\ldots$, таких, что $f\left(x^{k+1}\right) < f\left(x^k\right)$, $k = 0,1,\ldots$. Точки последовательности $\left\{x^k\right\}$ вычисляются по правилу:

$$x^{k+1} = x^k + t_k\, d^k,\ k = 0,1,\ldots; \tag{6.7}$$

$$d^k = -\nabla f\left(x^k\right) + \beta_{k-1} d^{k-1}; \tag{6.8}$$

$$d^0 = -\nabla f\left(x^0\right); \tag{6.9}$$

$$\beta_{k-1} = \frac{\left\| \nabla f\left(x^k\right)\right\|^2}{\left\| \nabla f\left(x^{k-1}\right)\right\|^2}. \tag{6.10}$$

Точка x^0 задается пользователем, величина шага t_k определяется для каждого значения k из условия

$$\varphi(t_k) = f\left(x^k + t_k d^k\right) \to \min_{t_k}. \tag{6.11}$$

Решение задачи одномерной минимизации (6.11) может осуществляться либо из условия $\dfrac{d\varphi}{dt_k} = 0$, $\dfrac{d^2\varphi}{dt_k{}^2} > 0$, либо численно с использованием методов одномерной минимизации, когда решается задача

$$\varphi\left(t_k\right) \to \min_{t_k \in [a,b]}. \tag{6.12}$$

При численном решении задачи определения величины шага степень близости найденного значения t_k к оптимальному значению t_k^*, удовлетворяющему условиям $\dfrac{d\varphi}{dt_k} = 0$, $\dfrac{d^2\varphi}{dt_k^2} > 0$, зависит от задания интервала $[a, b]$ и точности методов одномерной минимизации [27].

Вычисление величины β_{k-1} по формуле (6.10) обеспечивает для квадратичной формы $f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ построение последовательности *H*-сопряженных направлений $d^0, d^1, \ldots, d^k, \ldots$, для которых $\left(d^j, Hd^i\right) = 0$, $\forall i, j = 0,1,\ldots,k; i \neq j$. При этом в точках последовательности $\left\{x^k\right\}$ градиенты функции $f(x)$ взаимно перпендикулярны, т.е. $\left(\nabla f\left(x^{k+1}\right), \nabla f\left(x^k\right)\right) = 0$, $k = 0,1,\ldots$.

Для квадратичных функций $f(x)$ с матрицей $H > 0$ метод Флетчера–Ривса является конечным и сходится к точке минимума за число шагов, не превышающее n – размерность вектора x .

При минимизации неквадратичных функций метод не является конечным, при этом следует отметить, что погрешности в решении задачи (6.11) приводят к нарушению не только перпендикулярности градиентов, но и *H*-сопряженности направлений. Для неквадратичных функций, как правило, используется *алгоритм Полака–Рибьера* (Polak–Ribiere), когда в формулах (6.7) – (6.9) величина β_{k-1} вычисляется следующим образом:

$$\beta_{k-1} = \begin{cases} \dfrac{\left(\nabla f\left(x^k\right), \left[\nabla f\left(x^k\right) - \nabla f\left(x^{k-1}\right)\right]\right)}{\left\|\nabla f\left(x^{k-1}\right)\right\|^2}, & k \notin J, \\ 0, \quad k \in J, & \end{cases}$$

где $J = \{0, n, 2n, \ldots\}$. В отличие от алгоритма Флетчера–Ривса алгоритм Полака–Рибьера предусматривает использование итерации наискорейшего градиентного спуска через каждые n шагов. Построение последовательности $\left\{x^k\right\}$ заканчивается в точке, для которой $\left\|\nabla f\left(x^k\right)\right\| < \varepsilon_1$, где ε_1 – заданное число, или при $k \geq M$, где M – предельное число итераций, или при двукратном одновременном выполнении двух неравенств $\left\|x^{k+1} - x^k\right\| < \varepsilon_2$, $\left|f\left(x^{k+1}\right) - f\left(x^k\right)\right| < \varepsilon_2$, где ε_2 – малое положительное число.

Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое было описано в разд. 6.1.

Алгоритм

Шаг 1. Задать x^0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M - предельное число итераций. Найти градиент $\nabla f(x)$.

Шаг 2. Положить $k = 0$.

Шаг 3. Вычислить $\nabla f\left(x^k\right)$.

Шаг 4. Проверить выполнение критерия окончания $\left\|\nabla f\left(x^k\right)\right\| < \varepsilon_1$:

а) если критерий выполняется, то расчет окончен и $x^* = x^k$;

б) если нет, то перейти к шагу 5.

Шаг 5. Проверить условие $k \ge M$:

а) если неравенство выполняется, то расчет окончен и $x^* = x^k$;

б) если нет, то при $k = 0$ перейти к шагу 6, а при $k \ge 1$ перейти к шагу 7.

Шаг 6. Определить $d^0 = -\nabla f\left(x^0\right)$.

Шаг 7. Определить

$$\beta_{k-1} = \frac{\left\| \nabla f\left(x^k\right) \right\|^2}{\left\| \nabla f\left(x^{k-1}\right) \right\|^2}, \qquad \left[\beta_{k-1} = \begin{cases} \dfrac{\left(\nabla f\left(x^k\right), \left[\nabla f\left(x^k\right) - \nabla f\left(x^{k-1}\right) \right] \right)}{\left\| \nabla f\left(x^{k-1}\right) \right\|^2}, & k \notin J \\ 0, \quad k \in J \end{cases} \right].$$

Шаг 8. Определить $d^k = -\nabla f\left(x^k\right) + \beta_{k-1} d^{k-1}$.

Шаг 9. Найти t_k^* из условия $\varphi(t_k) = f\left(x^k + t_k\, d^k\right) \to \min\limits_{t_k}$.

Шаг 10. Вычислить $x^{k+1} = x^k + t_k^*\, d^k$.

Шаг 11. Проверить выполнение условий

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

а) в случае выполнения обоих условий в двух последовательных итерациях с номерами k и $k-1$ расчет окончен, найдена точка $x^* = x^{k+1}$;

б) если не выполняется хотя бы одно из условий, положить $k = k+1$ и перейти к шагу 3.

Геометрическая интерпретация метода для $n = 2$ изображена на рис. 6.7.

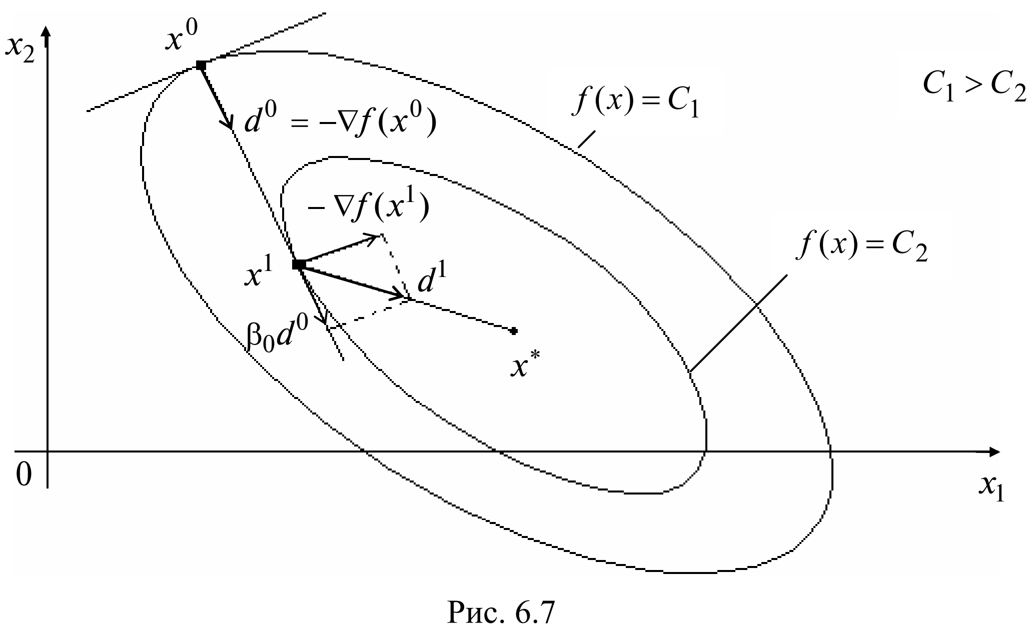

Рис. 6.7

Сходимость

Утверждение 6.3. *Если квадратичная функция* $f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}\, x_i x_j$ *с неотрицательно определенной матрицей H достигает своего минимального значения на* R^n*, то метод Флетчера–Ривса обеспечивает отыскание точки минимума не более чем за n шагов* [39].

Утверждение 6.4. *Если функция* $f(x)$ *ограничена снизу, а ее градиент удовлетворяет условию Липшица* $\| \nabla f(x) - \nabla f(y) \| \le L \| x - y \| \quad \forall x, y \in R^n$*, то в методе Полака–Рибьера* $\lim\limits_{k \to \infty} \left\| \nabla f\left(x^k\right) \right\| = 0$ [28].

З а м е ч а н и я 6.7.

1. Утверждение 6.4 гарантирует сходимость последовательности $\left\{x^k\right\}$ к стационарной точке x^*, где $\nabla f\left(x^*\right) = 0$. Следовательно, найденная в результате применения метода точка x^* нуждается в дополнительном исследовании с целью классификации этой точки.

2. Метод Полака–Рибьера гарантирует сходимость последовательности $\left\{x^k\right\}$ к точке минимума для сильно выпуклых функций.

3. Поиск глобального минимума $f(x)$ может быть осуществлен в соответствии с замечанием 6.2.

4. Процедура решения задачи совпадает с описанной в разд. 6.1.

Скорость сходимости

Оценки скорости сходимости получены только для сильно выпуклых функций, когда последовательность $\left\{x^k\right\}$ сходится к точке минимума функции $f(x)$ со скоростью $\left\| x^{k+n} - x^* \right\| \le C \left\| x^k - x^* \right\|^2$, $k \in \{0, n, 2n, \ldots\}$ [39].

Пример 6.5. Найти локальный минимум функции

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

□ I. Определим точку x^k, в которой выполнен по крайней мере один из критериев окончания расчетов.

1. Зададим $x^0, \varepsilon_1, \varepsilon_2, M$: $x^0 = (0{,}5;1)^T$, $\varepsilon_1 = 0{,}1; \varepsilon_2 = 0{,}15; M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Положим $k = 0$.

3^0. Вычислим $\nabla f\left(x^0\right)$: $\nabla f\left(x^0\right) = (3; 2{,}5)^T$.

4^0. Проверим условие $\left\| \nabla f\left(x^0\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^0\right) \right\| = 3{,}9 > 0{,}1$.

5^0. Проверим условие $k \ge M$: $k = 0 < 10 = M$.

6^0. Определим $d^0 = -\nabla f\left(x^0\right)$: $d^0 = -\left(3; 2{,}5\right)^T$.

9^0. Определим t_0^* из условия $f\left(x^0 + t_0\, d^0\right) \to \min\limits_{t_0}$: $t_0^* = 0{,}24$ (см. пример 6.2, так как первая итерация выполняется по методу наискорейшего спуска).

10^0. Вычислим $x^1 = x^0 + t_0^*\, d^0$: $x^1 = \left(-0{,}22; 0{,}4\right)^T$.

11^0. Проверим условия $\left\| x^1 - x^0 \right\| < \varepsilon_2$, $\left| f\left(x^1\right) - f\left(x^0\right) \right| < \varepsilon_2$:

$$\left\| x^1 - x^0 \right\| = 0{,}937 > 0{,}15; \qquad \left| f\left(x^1\right) - f\left(x^0\right) \right| = \left| 0{,}17 - 2 \right| = 1{,}83 > 0{,}15.$$

Положим $k = 1$ и перейдем к шагу 3.

3^1. Вычислим $\nabla f\left(x^1\right)$: $\nabla f\left(x^1\right) = \left(-0{,}48; 0{,}58\right)^T$.

4^1. Проверим условие $\left\| \nabla f\left(x^1\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^1\right) \right\| = 0{,}752 > 0{,}1$.

5^1. Проверим условие $k \ge M$: $k = 1 < 10 = M$.

7^1. Определим $\beta_0 = \dfrac{\left\| \nabla f\left(x^1\right) \right\|^2}{\left\| \nabla f\left(x^0\right) \right\|^2}$: $\beta_0 = 0{,}0373$.

8^1. Определим $d^1 = -\nabla f\left(x^1\right) + \beta_0 d^0$:

$$d^1 = -\left(-0{,}48; 0{,}58\right)^T - 0{,}0373\left(3; 2{,}5\right)^T = \left(0{,}368; -0{,}673\right)^T.$$

9^1. Определим t_1^* из условия $f\left(x^1 + t_1\, d^1\right) \to \min\limits_{t_1}$. Воспользуемся формулой

$$x^2 = x^1 + t_1\, d^1 = (-0{,}22;\ 0{,}4)^T + t_1 (0{,}368;\ -0{,}673)^T = (-0{,}22 + 0{,}368\, t_1;\ 0{,}4 - 0{,}673\, t_1)^T.$$

Подставляя полученное выражение в $f(x)$, имеем

$$\varphi(t_1) = 2 \cdot (-0{,}22 + 0{,}368\, t_1)^2 + (-0{,}22 + 0{,}368\, t_1) \cdot (0{,}4 - 0{,}673\, t_1) + (0{,}4 - 0{,}673\, t_1)^2.$$

Применяя необходимое условие безусловного экстремума

$$\frac{d\varphi(t_1)}{dt_1} = 4 \cdot (-0{,}22 + 0{,}368\, t_1) \cdot 0{,}368 + 0{,}368 \cdot (0{,}4 - 0{,}673\, t_1) +$$
$$+ (-0{,}22 + 0{,}368\, t_1) \cdot (-0{,}673) + 2 \cdot (0{,}4 - 0{,}673\, t_1) \cdot (-0{,}673) = 0,$$

находим $t_1^* \cong 0{,}595$. Поскольку $\dfrac{d^2\varphi(t_1)}{d\, t_1^{\,2}} = 0{,}952226 > 0$, найденное значение шага обеспечивает минимум функции $\varphi(t_1)$ по t_1.

10^1. Вычислим $x^2 = x^1 + t_1^* d^1$: $x^2 = (0{,}0010; 0{,}000)^T$.

11^1. Проверим условия $\left\| x^2 - x^1 \right\| < \varepsilon_2$, $\left| f(x^2) - f(x^1) \right| < \varepsilon_2$:

$$\left\| x^2 - x^1 \right\| = 0{,}456 > 0{,}15; \quad \left| f(x^2) - f(x^1) \right| = 0{,}17 > 0{,}15.$$

Положим $k = 2$ и перейдем к шагу 3.

3^2. Вычислим $\nabla f(x^2)$: $\nabla f(x^2) = (0{,}003; 0{,}006)^T$.

4^2. Проверим условие $\left\| \nabla f(x^2) \right\| < \varepsilon_1$: $\left\| \nabla f(x^2) \right\| = 0{,}0067 < 0{,}1$. Расчет окончен. Найдена точка $x^2 = (0{,}001; 0)^T$; $f(x^2) = 2 \cdot 10^{-6}$. На рис. 6.8 полученные точки соединены штриховой линией.

II. Проведем анализ точки x^2.

Используем утверждение 6.1. Функция $f(x) = 2x_1^2 + x_1 x_2 + x_2^2$ есть квадратичная функция двух переменных, имеющая положительно определенную матрицу вторых производных $H = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$. Это позволяет сделать вывод, что функция $f(x)$ строго выпукла, следовательно, имеет единственный минимум, приближение которого $x^2 = (0{,}001; 0)^T$ найдено за две итерации. ■

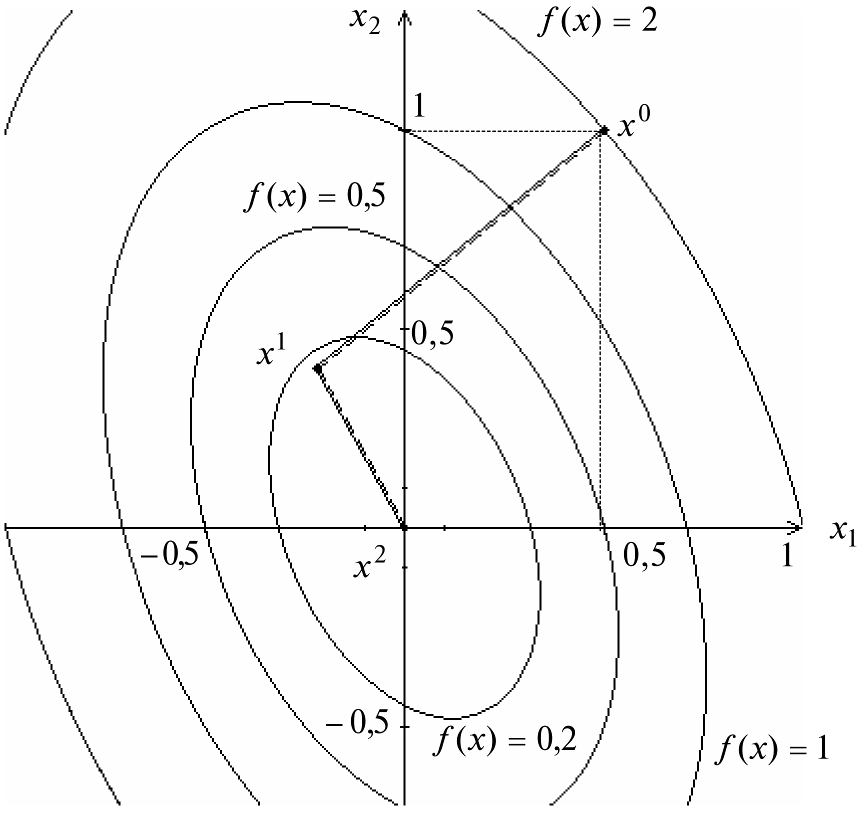

Рис. 6.8

6.6. МЕТОД ДЭВИДОНА–ФЛЕТЧЕРА–ПАУЭЛЛА

Постановка задачи

Пусть дана функция $f(x)$, ограниченная снизу на множестве R^n и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции $f(x)$ на множестве допустимых решений $X = R^n$, т.е. найти такую точку $x^* \in R^n$, что

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

Стратегия поиска

Стратегия метода Дэвидона–Флетчера–Пауэлла (Davidon–Fletcher–Powell) состоит в построении последовательности точек $\left\{x^k\right\}$, $k = 0,1,\ldots$, таких, что $f\left(x^{k+1}\right) < f\left(x^k\right)$, $k = 0,1,\ldots$. Точки последовательности $\left\{x^k\right\}$ вычисляются по правилу

$$x^{k+1} = x^k - t_k A^k \nabla f\left(x^k\right),\; k = 0,1,\ldots\,, \tag{6.13}$$

где A^k есть матрица размера $n \times n$, которая вычисляется по правилу

$$A^{k+1} = A^k + A^k{}_c, \quad A^0 = E\,, \tag{6.14}$$

$$A^k{}_c = \frac{\Delta x^k \left(\Delta x^k\right)^T}{\left(\Delta x^k\right)^T \Delta g^k} - \frac{A^k \Delta g^k \left(\Delta g^k\right)^T A^k}{\left(\Delta g^k\right)^T A^k \Delta g^k}, \tag{6.15}$$

где $\Delta x^k = x^{k+1} - x^k$, $\Delta g^k = \nabla f\left(x^{k+1}\right) - \nabla f\left(x^k\right)$.

Точка x^0 задается пользователем, величина шага t_k определяется из условия

$$\varphi(t_k) = f\left(x^k - t_k A^k \nabla f\left(x^k\right)\right) \to \min_{t_k}. \tag{6.16}$$

Решение задачи (6.16) может осуществляться как из условий $\dfrac{d\varphi}{dt_k} = 0$, $\dfrac{d^2\varphi}{dt_k{}^2} > 0$ или $\dfrac{dP}{dt_k} = 0$, $\dfrac{d^2 P}{dt_k{}^2} > 0$, где $P(t_k)$ – полином, аппроксимирующий функцию $\varphi(t_k)$, так и численно, т.е. путем поиска решения задачи $\varphi(t_k) \to \min\limits_{t_k \in [a,b]}$ методами одномерной минимизации.

Формулы (6.14), (6.15) при аналитическом решении задачи (6.16) обеспечивают построение последовательности $\left\{A^k\right\}$ положительно определенных матриц, таких, что $A^k \to H^{-1}\left(x^*\right)$ при $k \to \infty$. Следствием этого для квадратичной функции $f(x) = \dfrac{1}{2}(Hx, x) + (b, x)$, где $H > 0$, является тот факт, что направления d^k, $k = 0,1,\ldots$, будут H-сопряженными и, следовательно, алгоритм Д–Ф–П сойдется не более чем за n шагов.

Для неквадратичных функций $f(x)$ алгоритм перестает быть конечным и его сходимость зависит от точности решения задачи (6.16). Глобальную сходимость алгоритма можно гарантировать лишь при его обновлении через каждые n шагов, т.е. когда в формуле (6.13)

$$A^k = \begin{cases} E, \ k \in J; \ J = \{0, n, 2n, \ldots\}, \\ A^{k-1} + A_c^{k-1}, \ k \notin J. \end{cases}$$

Построение последовательности $\{x^k\}$ заканчивается в точке x^k, для которой $\left\|\nabla f(x^k)\right\| < \varepsilon_1$, где ε_1 – заданное число, или при $k \ge M$, где M – предельное число итераций), или при двукратном одновременном выполнении двух неравенств $\left\|x^{k+1} - x^k\right\| < \varepsilon_2$, $\left| f(x^{k+1}) - f(x^k)\right| < \varepsilon_2$, где ε_2 – малое положительное число.

Вопрос о том, может ли точка x^k рассматриваться как найденное приближение искомой точки минимума, решается путем проведения дополнительного исследования, которое было описано в разд. 6.1.

Алгоритм

Шаг 1. Задать x^0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M – предельное число итераций. Найти градиент $\nabla f(x)$.

Шаг 2. Положить $k = 0, A^0 = E$.

Шаг 3. Вычислить $\nabla f(x^k)$.

Шаг 4. Проверить критерий окончания $\left\|\nabla f(x^k)\right\| < \varepsilon_1$:

а) если критерий выполнен, то расчет закончен и $x^* = x^k$;
б) если нет, то перейти к шагу 5.

Шаг 5. Проверить условие $k \ge M$:

а) если неравенство выполнено, то расчет закончен и $x^* = x^k$;
б) если нет, перейти при $k = 0$ к шагу 10, а при $k \ge 1$ к шагу 6.

Шаг 6. Вычислить $\Delta g^k = \nabla f(x^{k+1}) - \nabla f(x^k)$.

Шаг 7. Вычислить $\Delta x^k = x^{k+1} - x^k$.

Шаг 8. Вычислить

$$A^k{}_c = \frac{\Delta x^k (\Delta x^k)^T}{(\Delta x^k)^T \Delta g^k} - \frac{A^k \Delta g^k (\Delta g^k)^T A^k}{(\Delta g^k)^T A^k \Delta g^k}.$$

Шаг 9. Вычислить $A^{k+1} = A^k + A^k{}_c$.

Шаг 10. Определить $d^k = -A^k \nabla f\left(x^k\right)$.

Шаг 11. Вычислить t_k^* из условия $\varphi(t_k) = f\left(x^k - t_k A^k \nabla f\left(x^k\right)\right) \to \min\limits_{t_k}$.

Шаг 12. Вычислить $x^{k+1} = x^k - t_k^* A^k \nabla f\left(x^k\right)$.

Шаг 13. Проверить условия $\left\|x^{k+1} - x^k\right\| < \varepsilon_2$, $\left| f\left(x^{k+1}\right) - f\left(x^k\right)\right| < \varepsilon_2$:

а) если оба неравенства выполняются в двух последовательных итерациях с номерами k и $k-1$, то расчет окончен и $x^* = x^{k+1}$;
б) в противном случае положить $k = k+1$ и перейти к шагу 3.

Сходимость

Утверждение 6.5. *Алгоритм метода Д–Ф–П в применении к квадратичной функции* $f(x) = \frac{1}{2}(Hx, x) + (b, x)$ *с положительно определенной матрицей Гессе H обеспечивает отыскание минимума* $x^* = H^{-1}b$ *не более чем за* n *шагов* [39].

З а м е ч а н и я 6.8.

1. Если минимизируемая функция $f(x)$ не является квадратичной и удовлетворяет условиям утверждения 6.1, то последовательность $\left\{x^k\right\}$, построенная по методу Д–Ф–П с обновлением, такова, что $\left\| \nabla f\left(x^k\right)\right\| \to 0$ при $k \to \infty$. Следовательно, найденная в результате применения метода точка x^* нуждается в дополнительном исследовании с целью ее классификации.

2. Если $f(x)$ дважды непрерывно дифференцируема и $H\left(x^*\right) > 0$, то метод Д–Ф–П с обновлением сходится к точке локального минимума x^* со сверхлинейной скоростью [39].

3. Если в дополнение к условиям п. 2 справедливо $\| H(x)y \| \le k \| y \|$ $\forall y \in R^n$ в окрестности точки x^*, то последовательность $\left\{x^k\right\}$ сходится к точке x^* с квадратичной скоростью [39].

4. Поиск глобального минимума $f(x)$ осуществляется в соответствии с замечанием 6.2. Процедура решения задачи совпадает с изложенной в разд. 6.1.

Пример 6.6. Найти локальный минимум

$$f(x) = 2x_1^2 + x_1x_2 + x_2^2.$$

□ I. Определим точку x^k, в которой выполнен по крайней мере один из критериев окончания расчетов.

1. Зададим $x^0, \varepsilon_1, \varepsilon_2, M$: $x^0 = (0{,}5;1)^T$, $\varepsilon_1 = 0{,}1; \varepsilon_2 = 0{,}15; M = 10$. Найдем градиент функции в произвольной точке $\nabla f(x) = (4x_1 + x_2; x_1 + 2x_2)^T$.

2. Положим $k = 0$, $A^0 = E$.

3^0. Вычислим $\nabla f\left(x^0\right)$: $\nabla f\left(x^0\right) = (3; 2{,}5)^T$.

4^0. Проверим выполнение условия $\left\| \nabla f\left(x^0\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^0\right) \right\| = 3{,}9 > 0{,}1$.

5^0. Проверим условие $k \ge M$: $k = 1 < 10 = M$; так как $k = 0$, перейдем к шагу 10.

10^0. Определим $d^0 = -A^0 \cdot \nabla f\left(x^0\right)$:

$$d^0 = -A^0 \cdot \nabla f\left(x^0\right) = -E\, \nabla f\left(x^0\right) = -\nabla f\left(x^0\right) = -(3; 2{,}5)^T.$$

11^0. Вычислим t_0^* из условия $f\left(x^0 - t_0\, \nabla f\left(x^0\right)\right) \to \min\limits_{t_0}$:

$$f\left(x^0 - t_0\, \nabla f\left(x^0\right)\right) = 2(0{,}5 - 3t_0)^2 + (0{,}5 - 3t_0)(1 - 2{,}5t_0) + (1 - 2{,}5t_0)^2 = \varphi(t_0).$$

Из условия $\dfrac{d\varphi}{dt_0} = 0$ находим $t_0^* = 0{,}24$ (см. пример 6.2, так как первая итерация выполняется по методу наискорейшего спуска).

12^0. Вычислим $x^1 = x^0 - t_0^*\, \nabla f\left(x^0\right)$: $x^1 = (-0{,}22; 0{,}4)^T$.

13^0. Проверим условия $\left\| x^1 - x^0 \right\| < \varepsilon_2$, $\left| f\left(x^1\right) - f\left(x^0\right) \right| < \varepsilon_2$:

$$\left\| x^1 - x^0 \right\| = 0{,}94 > 0{,}15; \qquad \left| f\left(x^1\right) - f\left(x^0\right) \right| = |\, 0{,}17 - 2 \,| = 1{,}83 > 0{,}15.$$

Положим $k = 1$ и перейдем к шагу 3.

3^1. Вычислим $\nabla f\left(x^1\right)$: $\nabla f\left(x^1\right) = (-0{,}48; 0{,}58)^T$.

4^1. Проверим условие $\left\| \nabla f\left(x^1\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^1\right) \right\| = 0{,}75 > 0{,}1$.

5^1. Проверим условие $k \ge M$: $k = 1 < 10 = M$; перейдем к шагу 6.

6^0. Вычислим $\Delta g^0 = \nabla f\left(x^1\right) - \nabla f\left(x^0\right)$: $\Delta g^0 = (-3{,}48; -1{,}92)^T$.

7^0. Вычислим $\Delta x^0 = x^1 - x^0$: $\Delta x^0 = (-0{,}72; -0{,}6)^T$.

8^0. Вычислим $A_c^0 = \dfrac{\Delta x^0 \left(\Delta x^0\right)^T}{\left(\Delta x^0\right)^T \Delta g^0} - \dfrac{A^0 \Delta g^0 \left(\Delta g^0\right)^T A^0}{\left(\Delta g^0\right)^T A^0 \Delta g^0}$:

$$A_c^0 = \frac{(-0{,}72; -0{,}6)^T (-0{,}72; -0{,}6)}{(-0{,}72; -0{,}6)(-3{,}48; -1{,}92)^T} - \frac{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (-3{,}48; -1{,}92)^T (-3{,}48; -1{,}92) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}{(-3{,}48; -1{,}92) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} (-3{,}48; -1{,}92)^T} =$$

$$= \begin{pmatrix} -0{,}625 & -0{,}305 \\ -0{,}305 & -0{,}134 \end{pmatrix}.$$

9^0. Вычислим $A^1 = A^0 + A_c^0$: $A^1 = \begin{bmatrix} 0{,}375 & -0{,}305 \\ -0{,}305 & 0{,}866 \end{bmatrix}$.

10^1. Определим $d^1 = -A^1 \cdot \nabla f\left(x^1\right)$: $d^1 = \left(0{,}356; -0{,}648\right)^T$.

11^1. Вычислим t_1^* из условия $f\left(x^1 - t_1 A^1 \nabla f\left(x^1\right)\right) \to \min\limits_{t_1}$. Воспользуемся формулой $x^2 = x^1 - t_1\, A^1 \nabla f\left(x^1\right) = x^1 + t_1 d^1 = (-0{,}22;\, 0{,}4)^T + t_1 (0{,}356;\, -0{,}648)^T =$ $= (-0{,}22 + 0{,}356 \cdot t_1;\, 0{,}4 - 0{,}648 \cdot t_1)^T$. Подставив полученные выражения для координат $x_1^2 = -0{,}22 + 0{,}356 \cdot t_1$, $x_2^2 = 0{,}4 - 0{,}648 \cdot t_1$ в функцию $f(x)$, получим

$$\varphi(t_1) = 2 \cdot (-0{,}22 + 0{,}356 \cdot t_1)^2 + (-0{,}22 + 0{,}356 \cdot t_1) \cdot (0{,}4 - 0{,}648 \cdot t_1) + (0{,}4 - 0{,}648 \cdot t_1)^2 .$$

Применив необходимое условие безусловного экстремума

$$\begin{aligned} \frac{d\varphi(t_1)}{dt_1} &= 4 \cdot (-0{,}22 + 0{,}356 \cdot t_1) \cdot 0{,}356 + 0{,}356 \cdot (0{,}4 - 0{,}648 \cdot t_1) + \\ &+ (-0{,}22 + 0{,}356 \cdot t_1) \cdot (-0{,}648) - 2 \cdot 0{,}648 \cdot (0{,}4 - 0{,}648 \cdot t_1) = 0 \end{aligned}$$

или $0{,}885376 \cdot t_1 - 0{,}54672 = 0$, найдем $t_1^* \cong 0{,}618$. Поскольку $\dfrac{d^2\varphi(t_1)}{dt_1{}^2} = 0{,}885376 > 0$, найденное значение шага обеспечивает минимум функции $\varphi(t_1)$ по t_1.

12^1. Вычислим $x^2 = x^1 - t_1^*\, A^1 \nabla f\left(x^1\right)$: $x^2 = \left(0{,}000; 0{,}000\right)^T$.

13^1. Проверим условия $\left\| x^2 - x^1 \right\| < \varepsilon_2$, $\left| f\left(x^2\right) - f\left(x^1\right) \right| < \varepsilon_2$:

$$\left\| x^2 - x^1 \right\| = 0{,}45 > 0{,}1; \qquad \left| f\left(x^2\right) - f\left(x^1\right) \right| = 0{,}17 > 0{,}15 .$$

Полагаем $k = 2$, переходим к шагу 3.

3^2. Вычислим $\nabla f\left(x^2\right)$: $\nabla f\left(x^2\right) = \left(0{,}000; 0{,}000\right)^T$.

4^2.Проверим условие $\left\| \nabla f\left(x^2\right) \right\| < \varepsilon_1$: $\left\| \nabla f\left(x^2\right) \right\| = 0 < 0{,}1$. Расчет окончен в точке $x^2 = \left(0; 0\right)^T$. На рис. 6.8 полученные точки соединены сплошной линией.

II. Проведем анализ точки x^2.

Используем утверждение 6.5. Функция $f\left(x\right) = 2x_1{}^2 + x_1 x_2 + x_2{}^2$ является квадратичной формой с положительно определенной матрицей Гессе $H = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$. Следовательно, точка x^2 есть найденное приближение точки минимума $x^* = \left(0, 0\right)^T$.■

6.7. МЕТОД КУБИЧЕСКОЙ ИНТЕРПОЛЯЦИИ

Постановка задачи

Пусть дана функция $f(x)$ одной переменной, ограниченная снизу на множестве R и имеющая непрерывную производную.

Требуется найти безусловный минимум функции $f(x)$, т.е. такую точку $x^* \in R$, что $f(x^*) = \min\limits_{x \in R} f(x)$.

Стратегия поиска

Задается начальная точка и с помощью серии пробных шагов находятся две точки, первые производные в которых имеют противоположные знаки. По величине функции и ее первых производных в полученных точках строится интерполяционный полином третьей степени. В качестве приближения точки минимума берется точка минимума полинома. Процесс поиска заканчивается, если производная в точке минимума полинома достаточно мала или процедура становится неэффективной.

Алгоритм

Шаг 1. Задать начальную точку x^0, величину шага $\Delta > 0$ и малые положительные числа ε_1 и ε_2.

Шаг 2. Вычислить производную $f'\left(x^0\right)$.

Шаг 3. Проверить знак производной в точке x^0:

а) если $f'\left(x^0\right) < 0$, вычислить $x^{k+1} = x^k + 2^k \Delta$, $k = 0,1,\ldots$, вплоть до точки x^M, в которой $f'\left(x^{M-1}\right) f'\left(x^M\right) \le 0$;

б) если $f'\left(x^0\right) > 0$, вычислить $x^{k+1} = x^k - 2^k \Delta$, $k = 0,1,\ldots$, вплоть до точки x^M, в которой $f'\left(x^{M-1}\right) f'\left(x^M\right) \le 0$.

Шаг 4. Положить $x_1 = x^{M-1}$, $x_2 = x^M$ и вычислить $f(x_1) = f_1$, $f'(x_1) = f_1'$, $f(x_2) = f_2$, $f'(x_2) = f_2'$.

Шаг 5. Найти точку минимума кубического интерполяционного полинома по формуле

$$\bar{x} = \begin{cases} x_2, & \mu < 0, \\ x_2 - \mu(x_2 - x_1), & 0 \le \mu \le 1, \\ x_1, & \mu > 1, \end{cases}$$

где $\mu = \dfrac{f_2' + w - z}{f_2' - f_1' + 2w}$; $\quad z = \dfrac{3(f_1 - f_2)}{x_2 - x_1} + f_1' + f_2'$; $\quad w = \begin{cases} \left(z^2 - f_1' f_2'\right)^{\frac{1}{2}}, & x_1 < x_2, \\ -\left(z^2 - f_1' f_2'\right)^{\frac{1}{2}}, & x_1 > x_2, \end{cases}$

и значение $f(\bar{x})$.

Шаг 6. Проверить условие убывания:

а) если $f(\bar{x}) < f(x_1)$, перейти к шагу 7;

б) если $f(\bar{x}) \geq f(x_1)$, вычислять $\bar{x}$ по формуле $\bar{x} = \bar{x} - \frac{1}{2}(\bar{x} - x_1)$ до тех пор, пока не будет выполнено неравенство $f(\bar{x}) \leq f(x_1)$.

Шаг 7. Проверить выполнение условий окончания:

$$\left| f'(\bar{x}) \right| \leq \varepsilon_1, \qquad \left| \frac{\bar{x} - x_1}{\bar{x}} \right| \leq \varepsilon_2:$$

а) если оба условия выполнены, процедура закончена и $x^* \cong \bar{x}$;

б) если хотя бы одно из условий не выполнено, положить либо $x_1 = \bar{x}$, $x_2 = x_1$, если $f'(\bar{x}) f'(x_1) < 0$, либо $x_1 = \bar{x}$, $x_2 = x_2$, если $f'(\bar{x}) f'(x_2) < 0$. Перейти к шагу 5.

Замечания 6.9.

1. На шагах 2 и 3 реализуется эвристическая процедура поиска границ интервала неопределенности, где изменение знака производной свидетельствует о переходе через точку минимума.

2. Формула, используемая на шаге 5, гарантирует, что точка $\bar{x}$ не выйдет за границы интервала $[x_1, x_2]$.

3. На шаге 6 проверяется, действительно ли точка $\bar{x}$ является приближением к минимуму.

4. На шаге 7 из трех точек x_1, x_2, $\bar{x}$ выбираются две, в которых знаки первых производных различны, после чего процедура кубической интерполяции повторяется.

5. Интерполяционный полином третьей степени строится по двум точкам вместо обычных четырех, так как в каждой точке используется информация о производной.

Пример 6.7. Найти минимум функции $f(x) = 2x^2 + \frac{16}{x}$ методом кубической интерполяции.

□ 1. Зададим $x^0 = 1$; $\Delta = 1$; $\varepsilon_1 = 0{,}01$; $\varepsilon_2 = 0{,}03$.

2. Вычислим $f'(x) = 4x - \frac{16}{x^2}$; $f'\left(x^0\right) = f'(1) = -12 < 0$.

3. Так как $f'(1) < 0$, то $x^1 = x^0 + \Delta = 1 + 1 = 2$. Вычислим $f'\left(x^1\right) = 4$. Поэтому $f'\left(x^0\right) f'\left(x^1\right) < 0,\ M = 1$.

4^0. Положим $x_1 = x^{M-1} = x^0 = 1$, $x_2 = x^M = x^1 = 2$ и вычислим

$$f'(x_1) = f'(1) = f_1' = -12; \qquad f(x_1) = f_1 = 18;$$

$$f'(x_2) = f'(2) = f_2' = 4; \qquad f(x_2) = f_2 = 16.$$

5^0. Вычислим

$$z=\frac{3(18-16)}{1}+(-12)+4=-2\,; \qquad w=\left[\,4-(-12)\cdot 4\,\right]^{\frac{1}{2}}=52^{\frac{1}{2}}=7{,}211\,;$$

$$\mu=\frac{4+7{,}211-(-2)}{4-(-12)+2\cdot 7{,}211}=0{,}4343\,; \quad \overline{x}=2-0{,}4343\,(2-1)=1{,}5657\,; \quad f(\overline{x})=15{,}1219\,.$$

6^0. Проверим условие убывания. Так как $f(\overline{x})=15{,}219<f(x_1)=18$, то перейдем к шагу 7.

7^0. Проверим условие окончания: $\left|\,f'(\overline{x})\right|=\left|f'(1{,}5657)\right|=\left|-0{,}264\,\right|>\varepsilon_1=0{,}01$. Условие не выполняется. Так как справедливо $f'(x_1)\,f'(x_2)=-0{,}264\cdot 4<0$, то $x_1=\overline{x}=1{,}5657;\ x_2=x_2=2$. Перейдем к шагу 5.

5^1. Вычислим $\quad z=\dfrac{3(15{,}1219-16)}{0{,}4343}+(-0{,}264)+4=-2{,}3296\,;$

$$w=\left[(2{,}3296)^2-(-0{,}264)\cdot 4\,\right]^{\frac{1}{2}}=2{,}5462\,; \quad \mu=\frac{4+2{,}5462-(-2{,}3296)}{4-(-0{,}264)+2\cdot 2{,}5462}=0{,}9486\,;$$

$$\overline{x}=2-0{,}9486\,(2-1{,}5657)=1{,}588\,; \qquad f(\overline{x})=f(1{,}588)=15{,}119\,.$$

6^1. Проверим условие убывания. Так как $f(\overline{x})=15{,}119<f(x_1)=15{,}1219$, то перейдем к шагу 7.

7^1. Проверим условия окончания:

$$\left|\,f'(\overline{x})\right|=\left|\,f'(1{,}588)\right|=0{,}0072<\varepsilon_1=0{,}01 \text{ (выполняется);}$$

$$\left|\,\frac{1{,}588-1{,}5657}{1{,}588}\right|=0{,}014<0{,}03 \text{ (выполняется).}$$

Поэтому расчет окончен и $x^*\cong\overline{x}=1{,}588$. Точная координата точки минимума $x^*=1{,}588$ (см. пример 5.8). Отсюда следует, что применение кубической интерполяции дало лучший результат, чем применение квадратичной интерполяции). ■

Задачи для самостоятельного решения

1. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя решить задачу:

$$f(x)=x_1^{\,3}-x_1x_2+x_2^{\,2}-2x_1+3x_2-4\to\min\,,\ x^0=(0,0)^T\,.$$

Ответ: точное решение $x^*=\left(\frac{1}{2},-\frac{5}{4}\right)^T$.

2. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя решить задачу:

$$f(x) = (x_2 - x_1^2)^2 + (1 - x_1)^2 \to \min, \; x^0 = (0,0)^T .$$

Ответ: точное решение $x^* = (1, 1)^T$.

3. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя из начальных точек $x^0 = (0{,}5; 0)^T$ и $x^0 = (-0{,}1; -0{,}5)^T$ решить задачу:

$$f(x) = [(x_2 + 1)^2 + x_1^2] \cdot [x_1^2 + (x_2 - 1)^2] \to \min .$$

Ответ: точное решение $x^* = (0{,}1)^T$ из точки $x^0 = (0{,}5; 0)^T$; $x^* = (0; -1)^T$ из точки $x^0 = (-0{,}1; -0{,}5)^T$.

4. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя из начальных точек $x^0 = (0{,}3)^T$ и $x^0 = (3{,}0)^T$ решить задачу:

$$f(x) = (x_2^2 + x_1^2 - 1)^2 + (x_1 + x_2 - 1)^2 \to \min .$$

Ответ: точное решение $x^* = (0{,}1)^T$ из точки $x^0 = (0{,}3)^T$; $x^* = (1{,}0)^T$ из точки $x^0 = (3{,}0)^T$.

5. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя из начальных точки $x^0 = (0{,}1; 0{,}5)^T$ решить задачу:

$$f(x) = -x_1^2 \exp[1 - x_1^2 - 20{,}25(x_1 - x_2)^2] \to \min .$$

Ответ: точное решение $x^* = (1, 1)^T$.

6. Методами градиентного и наискорейшего градиентного спуска, методами Флетчера–Ривса и Д–Ф–П, методами покоординатного спуска и Гаусса–Зейделя из начальной точки $x^0 = (0{,}1)^T$ решить задачу:

$$f(x) = -x_1 x_2 \exp[\,-(x_1 + x_2)] \to \min .$$

Ответ: точное решение $x^* = (1, 1)^T$.