8^2. Проверим выполнение условий окончания:

$$\left| \frac{15{,}142 - 15{,}123}{15{,}123} \right| = 0{,}0013 < \varepsilon_1 = 0{,}003; \quad \left| \frac{1{,}65 - 1{,}6125}{1{,}6125} \right| = 0{,}023 < \varepsilon_1 = 0{,}03.$$

Поиск закончен. Полученное приближенное решение $x^* \cong \bar{x} = 1{,}6125$.

Найдем аналитически координату точки минимума с помощью необходимых условий безусловного экстремума: $\frac{d f(x)}{dx} = 4x - \frac{16}{x^2} = 0$. Отсюда $x^* = \sqrt[3]{4} = 1{,}5874$. В этой точке $f''(x^*) = 4 + \frac{32}{(x^*)^3} = 12 > 0$, т.е. достаточные условия безусловного минимума выполняются. ■

5.2. МЕТОД КОНФИГУРАЦИЙ

Постановка задачи

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

Стратегия поиска

Метод конфигураций, или метод Хука–Дживса (Hooke–Jeeves), представляет собой комбинацию *исследующего поиска* с циклическим изменением переменных и ускоряющего *поиска по образцу*. Исследующий поиск ориентирован на выявление локального поведения целевой функции и определение направления ее убывания вдоль «оврагов» [28]. Полученная информация используется при поиске по образцу при движении вдоль «оврагов» [5].

Исследующий поиск начинается в некоторой начальной точке x^0, называемой *старым базисом*. В качестве множества направлений поиска выбирается множество координатных направлений. Задается величина шага, которая может быть различной для разных координатных направлений и переменной в процессе поиска. Фиксируется первое координатное направление и делается шаг в сторону увеличения соответствующей переменной. Если значение функции в пробной точке меньше значения функции в исходной точке, шаг считается удачным. В противном случае необходимо вернуться в предыдущую точку и сделать шаг в противоположном направлении с последующей проверкой поведения функции. После перебора всех координат исследующий поиск завершается. Полученная точка называется *новым базисом* (на рис. 5.12 в точке x^0 произведен исследующий поиск и получена точка x^1 – новый базис). Если исследующий поиск с данной величиной шага неудачен, то она уменьшается и процедура продолжается. Поиск заканчивается, когда текущая величина шага станет меньше некоторой величины.

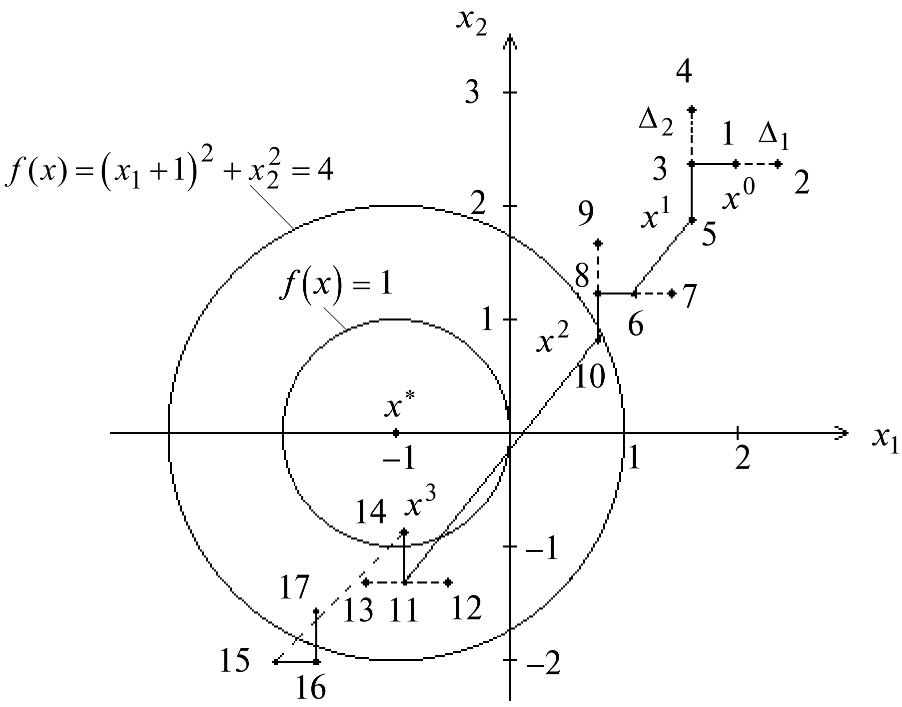

Рис. 5.12

Поиск по образцу заключается в движении по направлению от старого базиса к новому (от точки x^0 через точку x^1, из точки x^1 через точку x^2, из x^2 через x^3 на рис. 5.12). Величина ускоряющего шага задается ускоряющим множителем λ. Успех поиска по образцу определяется с помощью исследующего поиска из полученной точки (например из точек 6, 11, 15 на рис. 5.12). Если при этом значение в наилучшей точке меньше, чем в точке предыдущего базиса, то поиск по образцу удачен (точки 6, 11 – результат удачного поиска по образцу, а точка 15 – неудачного). Если поиск по образцу неудачен, происходит возврат в новый базис, где продолжается исследующий поиск с уменьшенным шагом. На рис. 5.12 удачный поиск отображается сплошными линиями, а неудачный – штриховыми, числа соответствуют порождаемым алгоритмом точкам.

Обозначим через $d_1,\dots,d_n$ – координатные направления:

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

При поиске по направлению d_i меняется только переменная x_i, а остальные переменные остаются зафиксированными.

Алгоритм

Шаг 1. Задать начальную точку x^0, число $\varepsilon > 0$ для остановки алгоритма, начальные величины шагов по координатным направлениям $\Delta_1,\dots,\Delta_n \ge \varepsilon$, ускоряющий множитель $\lambda > 0$, коэффициент уменьшения шага $\alpha > 1$. Положить $y^1 = x^0$, $i = 1$, $k = 0$.

Шаг 2. Осуществить исследующий поиск по выбранному координатному направлению:

а) если $f(y^i+\Delta_i d_i)<f(y^i)$, т.е. $f(y_1^i,\ldots,y_i^i+\Delta_i,\ldots,y_n^i)<f(y_1^i,\ldots,y_i^i,\ldots,y_n^i)$, шаг считается удачным. В этом случае следует положить $y^{i+1}=y^i+\Delta_i d_i$ и перейти к шагу 3;

б) если в п. “а” шаг неудачен, то делается шаг в противоположном направлении. Если $f(y^i-\Delta_i d_i)<f(y^i)$, т.е. $f(y_1^i,\ldots,y_i^i-\Delta_i,\ldots,y_n^i)<f(y_1^i,\ldots,y_i^i,\ldots,y_n^i)$, шаг считается удачным. В этом случае следует положить $y^{i+1}=y^i-\Delta_i d_i$ и перейти к шагу 3;

в) если в пп. “а” и “б” шаги неудачны, положить $y^{i+1}=y^i$.

Шаг 3. Проверить условия:

а) если $i<n$, то положить $i=i+1$ и перейти к шагу 2 (продолжить исследующий поиск по оставшимся направлениям);

б) если $i=n$, проверить успешность исследующего поиска:

- если $f(y^{n+1})<f(x^k)$, перейти к шагу 4;
- если $f(y^{n+1})\ge f(x^k)$, перейти к шагу 5.

Шаг 4. Провести поиск по образцу. Положить

$$x^{k+1}=y^{n+1},\quad y^1=x^{k+1}+\lambda(x^{k+1}-x^k),\quad i=1,\quad k=k+1$$

и перейти к шагу 2.

Шаг 5. Проверить условие окончания:

а) если все $\Delta_i\le\varepsilon$, то поиск закончить: $x^*\cong x^k$;

б) для тех i, для которых $\Delta_i>\varepsilon$, уменьшить величину шага: $\Delta_i=\dfrac{\Delta_i}{\alpha}$. Положить $y^1=x^k$, $x^{k+1}=x^k$, $k=k+1$, $i=1$ и перейти к шагу 2.

З а м е ч а н и я 5.9.

1. В алгоритме можно использовать одинаковую величину шага по координатным направлениям, т.е. вместо $\Delta_1,\Delta_2,\ldots,\Delta_n$ применять Δ.

2. Существует модификация метода, где при исследующем поиске и поиске по образцу используется одномерная минимизация. Тогда, если функция $f(x)$ дифференцируема, метод сходится к стационарной точке [5].

Пример 5.8. Найти минимум функции $f(x)=4(x_1-5)^2+(x_2-6)^2$ методом Хука – Дживса.

☐ 1. На рис. 5.13 изображены линии уровня функции: семейство эллипсов, описываемое уравнением

$$f(x)=4(x_1-5)^2+(x_2-6)^2=\text{const}.$$

Если $\text{const}=4$, имеем $(x_1-5)^2+\dfrac{(x_2-6)^2}{4}=1$, т.е. $a=1, b=2,\ x_{10}=5,\ x_{20}=6$. Зададим $x^0=(8,\,9)^T$ – старый базис; $\varepsilon=0{,}3$; $\Delta_1=1$, $\Delta_2=2$, $\alpha=2,\ \lambda=1$. Положим $k=0, i=1, y^1=(8,\,9)^T$.

2^0. Так как $f\left(y^1+\Delta_1 d_1\right)=f(9,9)=73>f\left(y^1\right)=f\left(x^0\right)=45$, то шаг неудачен.

Так как $f\left(y^1-\Delta_1 d_1\right)=f(7,9)=25<f\left(y^1\right)=f\left(x^0\right)=45$, то этот шаг удачен: $y^2 = y^1 - \Delta_1 d_1 = (7,9)^T$.

3^0. Поскольку $i=1<2=n$, то положим $i=2$ и перейдем к шагу 2.

2^1. Так как $f\left(y^2+\Delta_2 d_2\right)=f(7,11)=41>f\left(y^2\right)=25$, то шаг неудачен.

Так как $f\left(y^2-\Delta_2 d_2\right)=f(7,7)=17<f\left(y^2\right)=25$, то этот шаг удачен: $y^3 = y^2 - \Delta_2 d_2 = (7,7)^T$.

3^1. Поскольку $i=2=n=2$ и $f\left(y^3\right)=17<f\left(x^0\right)=45$, то перейдем к шагу 4.

4^0. Положим $x^1 = y^3 = (7,7)^T$ – новый базис, $i=1$, $k=k+1=1$, найдем $y^1 = x^1 + 1\cdot\left(x^1 - x^0\right) = (7,7)^T + \left[(7,7)^T - (8,9)^T\right] = (6,5)^T$. Выполнен поиск по образцу. Перейдем к шагу 2.

2^2. Так как $f\left(y^1+\Delta_1 d_1\right)=f(7,5)=17>f\left(y^1\right)=5$, то шаг неудачен. Так как $f\left(y^1-\Delta_1 d_1\right)=f(5,5)=1<f\left(y^1\right)=17$, то шаг удачный: $y^2 = y^1 - \Delta_1 d_1 = (5,5)^T$.

3^2. Поскольку $i=1<2=n$, то положим $i=i+1=2$ и перейдем к шагу 2.

2^3. Так как $f\left(y^2+\Delta_2 d_2\right)=f(5,7)=1=f\left(y^2\right)=f(5,5)=1$, то шаг неудачен.

Так как $f\left(y^2-\Delta_2 d_2\right)=f(5,3)=9>f\left(y^2\right)=1$, то шаг неудачен. Поэтому $y^3 = y^2 = (5,5)^T$.

3^3. Поскольку $i=2=n=2$ и $f\left(y^3\right)=1<f\left(x^1\right)=17$, то поиск по образцу на шаге 4^0 успешен. Точка $y^3 = (5,5)^T$ становится новым базисом, а точка $x^1 = (7,7)^T$ – старым базисом. Перейдем к шагу 4.

4^1. Положим $x^2 = y^3 = (5,5)^T$ – новый базис, $i=1$, $k=k+1=2$; $y^1 = x^2 + \left(x^2 - x^1\right) = (5,5)^T + \left[(5,5)^T - (7,7)^T\right] = (3,3)^T$ и перейдем к шагу 2.

2^4. Так как $f\left(y^1+\Delta_1 d_1\right)=f(4,3)=13<f\left(y^1\right)=f(3,3)=25$, то шаг удачен: $y^2 = y^1 + \Delta_1 d_1 = (4,\,3)^T$.

3^4. Поскольку $i=1<2=n$, то положим $i=i+1=2$ и перейдем к шагу 2.

2^5. Так как $f\left(y^2+\Delta_2 d_2\right)=f(4,5)=5<f\left(y^2\right)=13$, то шаг удачен: $y^3 = y^2 + \Delta_2 d_2 = (4,5)^T$.

3^5. Поскольку $i=n$ и $f\left(y^3\right)=5>f\left(x^2\right)=1$, то поиск по образцу на шаге 4^1 неудачен. Перейдем к шагу 5.

5^0. Так как $\Delta_1 = 1 > \varepsilon$, $\Delta_2 = 2 > \varepsilon$, то уменьшим шаг: $\Delta_1 = \dfrac{\Delta_1}{2} = 0{,}5$; $\Delta_2 = \dfrac{\Delta_2}{2} = 1$. Положим $y^1 = x^2 = (5,5)^T$, $x^3 = x^2 = (5,5)^T$, $i = 1$, $k = k + 1 = 3$ и перейдем к шагу 2.

2^6. Так как $f\left(y^1 + \Delta_1 d_1\right) = f(5{,}5;5) = 2 > f\left(y^1\right) = 1$, то шаг неудачен. Поскольку $f\left(y^1 - \Delta_1 d_1\right) = f(4{,}5;5) = 2 > f\left(y^1\right) = 1$, то шаг неудачен. Поэтому $y^2 = y^1 = (5,5)^T$.

3^6. Поскольку $i = 1 < 2 = n$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^7. Так как $f\left(y^2 + \Delta_2 d_2\right) = f(5,6) = 0 < f\left(y^2\right) = f(5,5) = 1$, то шаг удачен: $y^3 = y^2 + \Delta_2 d_2 = (5,6)^T$.

3^7. Поскольку $f\left(y^3\right) = 0 \le f\left(x^3\right) = 1$, то переходим к шагу 4. Точка y^3 становится новым базисом, а точка x^3 – старым.

4^2. Положим $x^4 = y^3 = (5,6)^T$ – новый базис, $i = 1$, $k = k + 1 = 4$; $y^1 = x^4 + \left(x^4 - x^3\right) = (5,6)^T + \left[(5,6)^T - (5,5)^T\right] = (5,7)^T$. Перейдем к шагу 2.

2^8. Так как $f\left(y^1 + \Delta_1 d_1\right) = f(5{,}5;7) = 2 > f\left(y^1\right) = f(5;7) = 1$, то шаг неудачен. Поскольку $f\left(y^1 - \Delta_1 d_1\right) = f(4{,}5;7) = 2 > f\left(y^1\right) = 1$, то шаг тоже неудачен. Тогда $y^2 = y^1 = (5;7)^T$.

3^8. Поскольку $i = 1 < 2 = n$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^9. Так как $f\left(y^2 + \Delta_2 d_2\right) = f(5;8) = 4 > f(5;7) = 1$, то шаг неудачен. Поскольку $f\left(y^2 - \Delta_2 d_2\right) = f(5;6) = 0 < f(5;7) = 1$, то шаг удачен: $y^3 = y^2 - \Delta_2 d_2 = (5{,}6)^T$.

3^9. Поскольку $i = 2 = n = 2$ и $f\left(y^3\right) = f\left(x^4\right) = f(5{,}6) = 0$, то поиск по образцу неудачен. Перейдем к шагу 5.

5^1. Так как $\Delta_1 > \varepsilon$ и $\Delta_2 > \varepsilon$, то уменьшим шаг: $\Delta_1 = \dfrac{0{,}5}{2} = 0{,}25$; $\Delta_2 = 0{,}5$. Положим $y^1 = x^4 = (5;6)^T$, $x^5 = x^4$, $k = k + 1 = 5$, $i = 1$ и перейдем к шагу 2.

2^{10}. Так как $f\left(y^1 + \Delta_1 d_1\right) = f(5{,}25;6) = 0{,}25 > f\left(y^1\right) = 0$, то шаг неудачен. Так как $f\left(y^1 - \Delta_1 d_1\right) = f(4{,}75;6) = 0{,}25 > f\left(y^1\right) = 0$, то шаг неудачен и $y^2 = y^1 = (5;6)^T$.

3^{10}. Поскольку $i = 1 < 2 = n$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

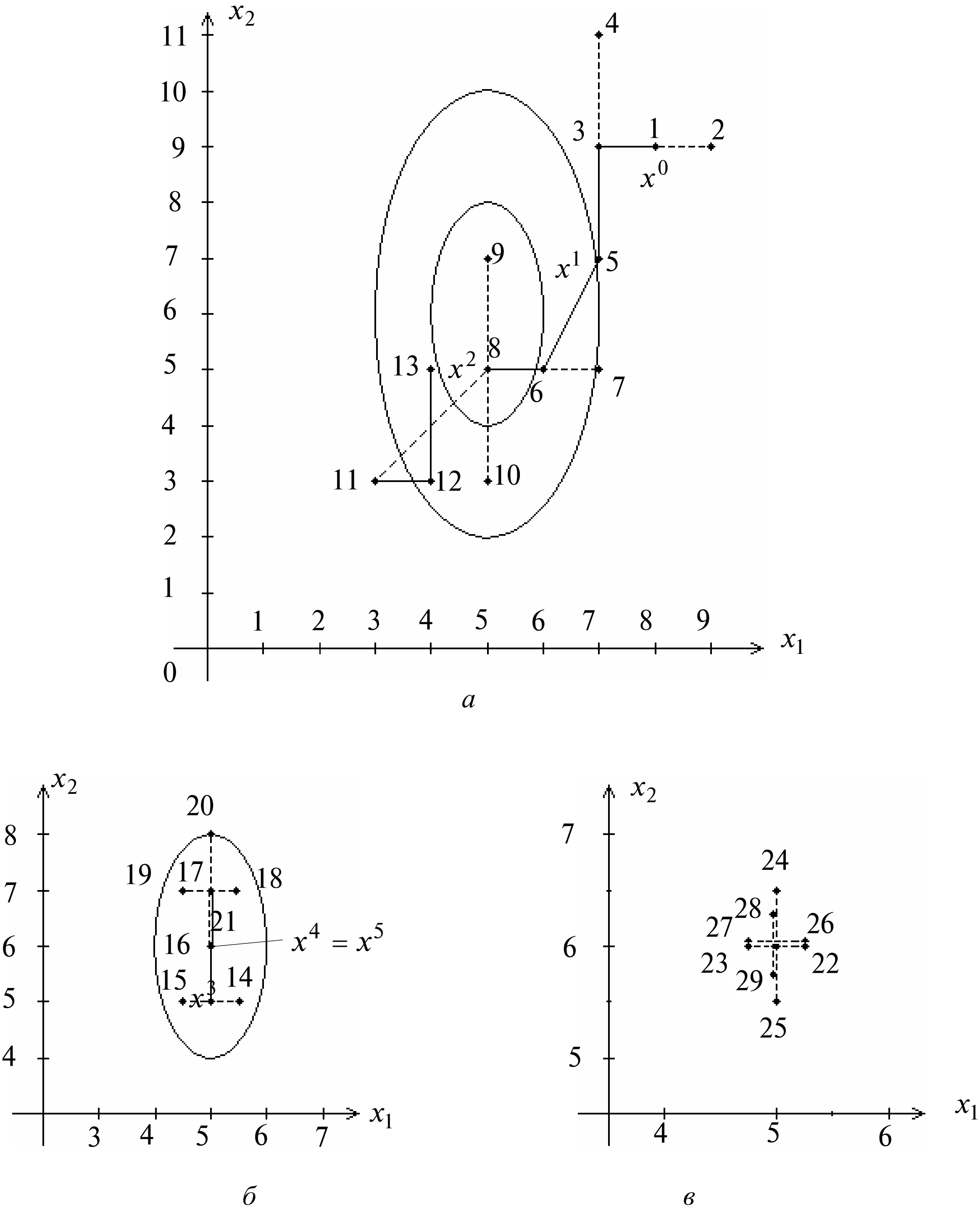

Рис. 5.12

2^{11}. Так как $f\left(y^2 + \Delta_2\, d_2\right) = f(5;6{,}5) = 0{,}25 > f\left(y^2\right) = 0$, то шаг неудачен. Так как $f\left(y^2 - \Delta_2\, d_2\right) = f(5;5{,}5) = 0{,}25 > f\left(y^2\right) = 0$, то шаг неудачен: $y^3 = y^2 = (5;6)^T$.

3^{11}. Поскольку $i = 2 = n = 2$ и $f\left(y^3\right) = 0 = f\left(x^5\right) = 0$, то исследующий поиск неудачен. Перейдем к шагу 5.

5^2. Поскольку $\Delta_1 = 0{,}25 < \varepsilon = 0{,}3$; $\Delta_2 = 0{,}5 > \varepsilon$, то уменьшим шаг Δ_2:
$\Delta_1 = 0{,}25$; $\Delta_2 = \dfrac{0{,}5}{2} = 0{,}25$. Положим $y^1 = x^5 = (5,6)^T$, $x^6 = x^5$, $k = k+1 = 6$, $i = 1$ и перейдем к шагу 2.

2^{12}. Так как $f\left(y^1 + \Delta_1 d_1\right) = f(5{,}25;6) = 0{,}25 > f\left(y^1\right) = 0$ и
$f\left(y^1 - \Delta_1 d_1\right) = f(4{,}75;6) = 0{,}25 > f\left(y^1\right) = 0$, то шаги неудачные и $y^2 = y^1 = (5;6)^T$.

3^{12}. Поскольку $i = 1 < 2 = n$, то положим $i = i+1 = 2$ и перейдем к шагу 2.

2^{13}. Так как $f\left(y^2 + \Delta_2 d_2\right) = f(5;6{,}25) = 0{,}0625 > f\left(y^2\right) = 0$ и
$f\left(y^2 - \Delta_2 d_2\right) = f(5;5{,}875) = 0{,}0625 > f\left(y^2\right) = 0$, шаги неудачны и $y^3 = y^2 = (5;6)^T$.

3^{13}. Поскольку $i = 2 = n = 2$ и $f\left(y^3\right) = 0 = f\left(x^6\right) = 0$, то перейдем к шагу 5.

5^3. Так как $\Delta_1 = 0{,}25 < \varepsilon = 0{,}3$ и $\Delta_2 = 0{,}25 < \varepsilon = 0{,}3$, то поиск завершен: $x^* = (5;6)^T$.

На рис. 5.12,*а–в* последовательно изображены полученные точки; сплошной линией отмечены удачные итерации, а штриховой – неудачные. ■

5.3. МЕТОД ДЕФОРМИРУЕМОГО МНОГОГРАННИКА

Постановка задачи

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

Стратегия поиска

В основу метода деформируемого многогранника, или метода Нелдера–Мида (Nelder–Mead), положено построение последовательности систем $n+1$ точек $x^i(k),\ i = 1,\dots,n+1$, которые являются вершинами выпуклого многогранника. Точки системы $x^i(k+1),\ i = 1,\dots,n+1$, на $(k+1)$-й итерации совпадают с точками системы $x^i(k),\ i = 1,\dots,n+1$, кроме $i = h$, где точка $x^h(k)$ – наихудшая в системе $x^i(k),\ i = 1,\dots,n+1$, т.е. $f(x^h(k)) = \max\limits_{1 \le i \le n+1} f(x^i(k))$. Точка $x^h(k)$ заменяется на другую точку по специальным правилам, описанным ниже. В результате многогранники деформируются в зависимости от структуры линий уровня целевой функции, вытягиваясь вдоль длинных наклонных плоскостей, изменяя направление в изогнутых впадинах и сжимаясь в окрестности минимума. Построение последовательности многогранников заканчивается, когда значения функции в вершинах текущего многогранника отличаются от значения функции в центре тяжести системы $x^i(k),\ i = 1,\dots,n+1;\ i \ne h$, не более чем на величину $\varepsilon > 0$.

Алгоритм

Шаг 1. Задать координаты вершин многогранника $x^1,\ldots,x^{n+1}$; параметры отражения α, сжатия β, растяжения γ; число $\varepsilon > 0$ для остановки алгоритма. Положить $k = 0$ (последующие шаги 2–6 соответствуют текущему номеру k системы точек).

Шаг 2. Среди вершин найти «наилучшую» x^l и «наихудшую» x^h, соответствующие минимальному и максимальному значениям функции:

$$f\left(x^l\right) = \min_{j=1,\ldots,n+1} f\left(x^j\right); \qquad f\left(x^h\right) = \max_{j=1,\ldots,n+1} f\left(x^j\right),$$

а также точку x^s, в которой достигается второе по величине после максимального значение функции $f\left(x^s\right)$.

Шаг 3. Найти «центр тяжести» всех вершин многогранника, за исключением «наихудшей» x^h:

$$x^{n+2} = \frac{1}{n}\left[\sum_{j=1}^{n+1} x^j - x^h\right] = \frac{1}{n}\sum_{\substack{j=1 \\ j\neq h}}^{n+1} x^j .$$

Шаг 4. Проверить условие окончания:

а) если $\sigma = \left\{\frac{1}{n+1}\sum_{j=1}^{n+1}\left[f\left(x^j\right) - f\left(x^{n+2}\right)\right]^2\right\}^{\frac{1}{2}} \le \varepsilon$, процесс поиска можно завершить и в качестве приближенного решения взять наилучшую точку текущего многогранника: $x^* \cong x^l$;

б) если $\sigma > \varepsilon$, продолжать процесс.

Шаг 5. Выполнить операцию *отражения* «наихудшей» вершины через центр тяжести x^{n+2} (рис. 5.13, *а*):

$$x^{n+3} = x^{n+2} + \alpha\left(x^{n+2} - x^h\right).$$

Шаг 6. Проверить выполнение условий:

а) если $f\left(x^{n+3}\right) \le f\left(x^l\right)$, выполнить операцию *растяжения* (рис. 5.13, *б*):

$$x^{n+4} = x^{n+2} + \gamma\left(x^{n+3} - x^{n+2}\right).$$

Найти вершины нового многогранника:

- если $f\left(x^{n+4}\right) < f\left(x^l\right)$, то вершина x^h заменяется на x^{n+4} (при $n = 2$ многогранник будет содержать вершины x^1, x^3, x^6). Затем следует положить $k = k + 1$ и перейти к шагу 2;
- если $f\left(x^{n+4}\right) \ge f\left(x^l\right)$, то вершина x^h заменяется на x^{n+3} (при $n = 2$ многогранник будет содержать вершины x^1, x^3, x^5). Далее следует положить $k = k + 1$ и перейти к шагу 2;

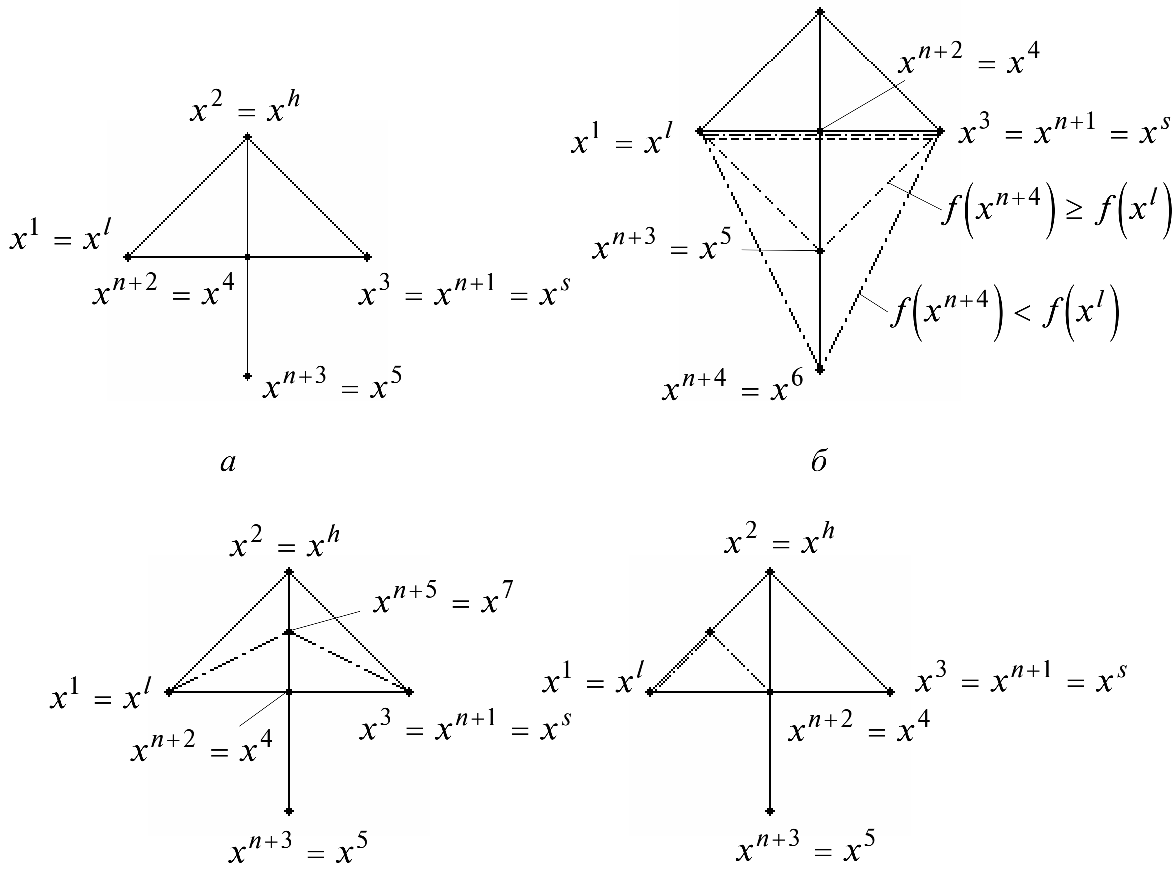

Рис. 5.13

б) если $f\left(x^s\right) < f\left(x^{n+3}\right) \le f\left(x^h\right)$, то выполнить операцию *сжатия* (рис. 5.13, *в*):

$$x^{n+5} = x^{n+2} + \beta\left(x^h - x^{n+2}\right).$$

Заменить вершину x^h на x^{n+5}, положить $k = k + 1$ и перейти к шагу 2 (при $n = 2$ многогранник будет содержать вершины x^1, x^3, x^7);

в) если $f\left(x^l\right) < f\left(x^{n+3}\right) \le f\left(x^s\right)$, то вершину x^h заменить на x^{n+3}. При этом следует положить $k = k + 1$ и перейти к шагу 2;

г) если $f\left(x^{n+3}\right) > f\left(x^h\right)$, выполнить операцию *редукции* (рис. 5.13, *г*). Формируется новый многогранник с уменьшенными вдвое сторонами и вершиной x^l:

$$x^j = x^l + 0{,}5\left(x^j - x^l\right), \quad j = 1, \ldots, n+1.$$

При этом следует положить $k = k + 1$ и перейти к шагу 2.

З а м е ч а н и е 5.10. Нелдер и Мид рекомендуют использовать параметры $\alpha = 1$; $\beta = 0{,}5$; $\gamma = 2$; Павиани (Paviani): $\alpha = 1$; $0{,}4 \le \beta \le 0{,}6$; $2{,}8 \le \gamma \le 3$; Паркинсон и Хатчинсон (Parkinson, Hutchinson): $\alpha = 2$; $\beta = 0{,}25$; $\gamma = 2{,}5$. В последнем случае в рамках операции отражения фактически выполняется растяжение.

Пример 5.9. Найти минимум функции $f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$ методом Нелдера – Мида.

□ 1. Так как $n = 2$, зададим начальный треугольник с вершинами $x^1 = (8,9)^T$; $x^2 = (10,11)^T$; $x^3 = (8,11)^T$. Положим $\alpha = 1$; $\beta = 0{,}5$; $\gamma = 2$; $\varepsilon = 0{,}2$; $k = 0$.

2^0. Так как $f(x^1) = 45, f(x^2) = 125, f(x^3) = 61$, то $x^l = x^1$, $x^h = x^2$, $x^s = x^3$.

3^0. Найдем центр тяжести вершин x^1 и x^3 (середину стороны, противостоящей вершине x^h):

$$x^4 = \frac{1}{2} \cdot \left[x^1 + x^3\right] = \frac{1}{2} \cdot \left[(8,\ 9)^T + (8,\ 11)^T\right] = (8,\ 10)^T; \quad f\left(x^4\right) = 52.$$

4^0. Так как $\sigma = \left\{\frac{1}{3}(49 + 5329 + 81)\right\}^{\frac{1}{2}} = 42{,}657 > \varepsilon$, то процесс продолжается.

5^0. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^2\right) = (8,\ 10)^T + \left[(8,\ 10)^T - (10,\ 11)^T\right] = (6,\ 9)^T; \quad f\left(x^5\right) = f(6,9) = 13.$$

6^0. Поскольку $f\left(x^5\right) = 13 < f\left(x^l\right) = f\left(x^1\right) = 45$, выполним растяжение:

$$x^6 = x^4 + 2 \cdot \left(x^5 - x^4\right) = (8,\ 10)^T + 2 \cdot \left[(6,\ 9)^T - (8,\ 10)^T\right] = (4,\ 8)^T; \quad f\left(x^6\right) = f(4,8) = 8.$$

Так как $f\left(x^6\right) = 8 < f\left(x^l\right) = f\left(x^1\right) = 45$, то вершина x^2 заменяется на x^6. Новый многогранник содержит вершины x^2, x^3, x^6. Положим $k = k + 1 = 1$ и перейдем к шагу 2.

2^1. Имеем вершины $x^1 = (8,9)^T$, $x^2 = (4,8)^T$, $x^3 = (8,11)^T$;

$$f\left(x^1\right) = 45, \quad f\left(x^2\right) = 8; \quad f\left(x^3\right) = 61; \quad x^l = x^2, x^h = x^3, x^s = x^1.$$

3^1. Найдем центр тяжести вершин x^2 и x^1:

$$x^4 = \frac{1}{2} \cdot \left[x^1 + x^2\right] = \frac{1}{2} \cdot \left[(8,\ 9)^T + (4,\ 8)^T\right] = (6;\ 8{,}5)^T; \ f\left(x^4\right) = 10{,}25.$$

4^1. Так как $\sigma = \left\{\frac{1}{3}\left(34{,}75^2 + 50{,}75^2 + 2{,}25^2\right)\right\}^{\frac{1}{2}} = 35{,}53 > \varepsilon$, процесс продолжается.

5^1. Выполним отражение

$$x^5 = x^4 + 1 \cdot \left(x^4 - x^3\right) = (6;8{,}5)^T + \left[(6;8{,}5)^T - (8,11)^T\right] = (4,6)^T; \ f\left(x^5\right) = 4.$$

6^1. Поскольку $f\left(x^5\right) = 4 < f\left(x^l\right) = f\left(x^2\right) = 8$, выполним растяжение.

$$x^6 = x^4 + 2\cdot\left(x^5 - x^4\right) = (6;8{,}5)^T + 2\cdot\left[(4;6)^T - (6;8{,}5)^T\right] = (2;3{,}5)^T;$$

$$f\left(x^6\right) = f(2;3{,}5) = 42{,}25.$$

Так как $f\left(x^6\right) = 42{,}25 > f\left(x^l\right) = f\left(x^2\right) = 8$, то вершина $x^h = x^3$ заменяется на x^5. Новый многогранник содержит вершины x^1, x^2, x^5. Положим $k = 2$ и перейдем к шагу 2.

2^2. Имеем вершины $x^1 = (8,9)^T$, $x^2 = (4,8)^T$, $x^3 = (4,6)^T$;

$$f\left(x^1\right) = 45, \quad f\left(x^2\right) = 8, \quad f\left(x^3\right) = 4; \quad x^l = x^3, x^h = x^1, x^s = x^2.$$

3^2. Найдем центр тяжести вершин x^2 и x^3:

$$x^4 = \frac{1}{2}\cdot\left(x^2 + x^3\right) = \frac{1}{2}\cdot\left[(4,\ 6)^T + (4,\ 8)^T\right] = (4,\ 7)^T; \ f\left(x^4\right) = 5.$$

4^2. Так как $\sigma = \left\{\frac{1}{3}\left(40^2 + 3^2 + 1\right)\right\}^{\frac{1}{2}} = 23{,}2 > \varepsilon$, процесс продолжается.

5^2. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^1\right) = (4,\ 7)^T + \left[(4,\ 7)^T - (8,\ 9)^T\right] = (0,\ 5)^T; \quad f\left(x^5\right) = f(0,5) = 101.$$

6^2. Поскольку $f\left(x^5\right) = 101 > f\left(x^h\right) = 45$, выполним редукцию:

$$x^1 = x^l + 0{,}5\cdot\left(x^1 - x^l\right) = (4,\ 6)^T + \frac{1}{2}\cdot\left[(8,\ 9)^T - (4,\ 6)^T\right] = (6;\ 7{,}5)^T;$$

$$x^2 = x^l + 0{,}5\cdot\left(x^2 - x^l\right) = (4,\ 6)^T + \frac{1}{2}\cdot\left[(4,\ 8)^T - (4,\ 6)^T\right] = (4,\ 7)^T;$$

$$x^3 = x^l + 0{,}5\cdot\left(x^3 - x^l\right) = x^l = (4,6)^T.$$

Положим $k = 3$ и перейдем к шагу 2.

2^3. Вычислим значения функции: $f\left(x^1\right) = 6{,}25,\ f\left(x^2\right) = 5,\ f\left(x^3\right) = 4$.

Поэтому $x^l = x^3, x^h = x^1, x^s = x^2$.

3^3. Найдем центр тяжести вершин x^2 и x^3:

$$x^4 = \frac{1}{2}\cdot\left(x^2 + x^3\right) = \frac{1}{2}\cdot\left[(4,\ 7)^T + (4,\ 6)^T\right] = (4;\ 6{,}5)^T; \quad f\left(x^4\right) = 4{,}25.$$

4^3. Так как $\sigma = \left\{\frac{1}{3}\left(4 + 0{,}75^2 + 0{,}25^2\right)\right\}^{\frac{1}{2}} = 1{,}241 > \varepsilon$, процесс продолжается.

5^3. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^1\right) = (4;6{,}5)^T + \left[(4;6{,}5)^T - (6;7{,}5)^T\right] = (2;5{,}5)^T; \ f\left(x^5\right) = 36{,}25.$$

6^3. Так как $f\left(x^5\right)=36,25>f\left(x^l\right)=f\left(x^1\right)=6,25$, то выполним редукцию:

$$x^1 = x^l + 0,5\cdot\left(x^1 - x^l\right) = (4,\ 6)^T + \frac{1}{2}\cdot\left[(6;\ 7,5)^T - (4,\ 6)^T\right] = (5;\ 6,75)^T\ ;$$

$$x^2 = x^l + 0,5\cdot\left(x^2 - x^l\right) = (4, 6)^T + \frac{1}{2}\cdot\left[(4, 7)^T - (4, 6)^T\right] = (4;\ 6,5)^T\ ;$$

$$x^3 = x^l = x^3 = \left(4, 6\right)^T .$$

Положим $k = k+1 = 4$ и перейдем к шагу 2.

2^4. Вычислим значения функции: $f\left(x^1\right) = 0,5625;\ f\left(x^2\right) = 4,25;\ f\left(x^3\right) = 4$. Поэтому $x^l = x^1, x^h = x^2, x^s = x^3$.

3^4. Найдем центр тяжести вершин x^1 и x^3:

$$x^4 = \frac{1}{2}\cdot\left[x^1 + x^3\right] = \frac{1}{2}\cdot\left[(5;\ 6,75)^T + (4,\ 6)^T\right] = (4,5;\ 6,375)^T\ ;\quad f\left(x^4\right) = 1,14 .$$

4^4. Так как $\sigma = \left\{\frac{1}{3}(0,33 + 9,67 + 8,17)\right\}^{\frac{1}{2}} = 2,46 > \varepsilon$, процесс продолжается.

5^4. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^2\right) = \left(4,5; 6,375\right)^T + \left[\left(4,5; 6,375\right)^T - \left(4; 6,5\right)^T\right] = \left(5; 6,25\right)^T\ ;$$

$$f\left(x^5\right) = 0,0625 .$$

6^4. Так как $f\left(x^5\right) < f\left(x^l\right) = 0,5625$, выполним растяжение:

$$x^6 = x^4 + 2\cdot\left(x^5 - x^4\right) = \left(4,5; 6,375\right)^T + 2\cdot\left[\left(5; 6,25\right)^T - \left(4,5; 6,375\right)^T\right] = \left(5,5; 6,125\right)^T\ ;$$

$$f\left(x^6\right) = 1,015 .$$

Так как $f\left(x^6\right) = 1,015 > f\left(x^l\right) = f\left(x^1\right) = 0,5625$, то вершина $x^h = x^2$ заменяется на x^5. Положим $k = k+1 = 5$ и перейдем к шагу 2.

2^5. Имеем $x^1 = \left(5; 6,75\right)^T, x^2 = \left(5; 6,25\right)^T, x^3 = \left(4, 6\right)^T$;
$f\left(x^1\right) = 0,5625; f\left(x^2\right) = 0,0625; f\left(x^3\right) = 4$. Поэтому $x^l = x^2, x^h = x^3, x^s = x^1$.

3^5. Найдем центр тяжести вершин x^1 и x^2:

$$x^4 = \frac{1}{2}\cdot\left[x^1 + x^2\right] = \frac{1}{2}\cdot\left[(5;\ 6,75)^T + (5;\ 6,25)^T\right] = (5;\ 6,5)^T\ ;\ f\left(x^4\right) = 0,25 .$$

4^5. Так как $\sigma = \left\{\frac{1}{3}(0,097 + 0,035 + 14,062)\right\}^{\frac{1}{2}} = 2,17 > \varepsilon$, процесс продолжается.

5^5. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^3\right) = \left(5; 6,5\right)^T + \left[\left(5; 6,5\right)^T - \left(4, 6\right)^T\right] = \left(6, 7\right)^T;\ f\left(x^5\right) = 5 .$$

6^5. Так как $f\left(x^5\right) > f\left(x^h\right) = f\left(x^3\right) = 4$, выполним редукцию:

$$x^1 = x^2 + 0{,}5\cdot\left(x^1 - x^2\right) = (5; 6{,}25)^T + 0{,}5\cdot\left[(5; 6{,}75)^T - (5; 6{,}25)^T\right] = (5; 6{,}5)^T;$$

$$x^2 = x^l = (5; 6{,}25)^T;$$

$$x^3 = x^2 + 0{,}5\cdot\left(x^3 - x^2\right) = (5; 6{,}25)^T + 0{,}5\cdot\left[(4, 6)^T - (5; 6{,}25)^T\right] = (4{,}5; 6{,}125)^T.$$

Положим $k = k + 1 = 6$ и перейдем к шагу 2.

2^6. Имеем $f\left(x^1\right) = 0{,}25; f\left(x^2\right) = 0{,}0625; f\left(x^3\right) = 1{,}015$. Поэтому $x^l = x^2, x^h = x^3, x^s = x^1$.

3^6. Найдем центр тяжести вершин x^1 и x^2:

$$x^4 = \frac{1}{2}\cdot\left(x^1 + x^2\right) = \frac{1}{2}\cdot\left[(5;\ 6{,}5)^T + (5;\ 6{,}25)^T\right] = (5;\ 6{,}375)^T;\ f\left(x^4\right) = 0{,}14.$$

4^6. Так как $\sigma = \left\{\frac{1}{3}(0{,}012 + 0{,}006 + 0{,}765)\right\}^{\frac{1}{2}} = 0{,}51 > \varepsilon$, процесс продолжается.

5^6. Выполним отражение:

$$x^5 = x^4 + \left(x^4 - x^1\right) = (5; 6{,}375)^T + \left[(5; 6{,}375)^T - (4{,}5; 6{,}125)^T\right] = (5{,}5; 6{,}625)^T;$$

$$f\left(x^5\right) = 1{,}39.$$

6^6. Поскольку $f\left(x^5\right) > f\left(x^h\right) = 1{,}015$, то выполним редукцию:

$$x^1 = x^l + 0{,}5\cdot\left(x^1 - x^l\right) = (5; 6{,}25)^T + 0{,}5\cdot\left[(5; 6{,}5)^T - (5; 6{,}25)^T\right] = (5; 6{,}375)^T;$$

$$x^2 = x^l = (5; 6{,}25)^T;$$

$$x^3 = x^l + 0{,}5\cdot\left(x^3 - x^l\right) = (5; 6{,}25)^T + 0{,}5\cdot\left[(4{,}5; 6{,}125)^T - (5; 6{,}25)^T\right] = (4{,}75; 6{,}18)^T.$$

Положим $k = k + 1 = 7$ и перейдем к шагу 2.

2^7. Так как $f\left(x^1\right) = 0{,}14; f\left(x^2\right) = 0{,}0625; f\left(x^3\right) = 0{,}28$, то получим $x^l = x^2, x^h = x^3, x^s = x^1$.

3^7. Найдем центр тяжести вершин x^1 и x^2:

$$x^4 = \frac{1}{2}\cdot\left(x^1 + x^2\right) = \frac{1}{2}\cdot\left[(5;\ 6{,}375)^T + (5;\ 6{,}25)^T\right] = (5;\ 6{,}31)^T;\ \ f\left(x^4\right) = 0{,}09.$$

4^7. Так как $\sigma = \left\{ \frac{1}{3} \left(0,0025 + 0,0007 + 0,036 \right) \right\}^{\frac{1}{2}} = 0,114 < \varepsilon = 0,2$, процесс закончен: $x^* = x^l = \left(5; 6,25\right)^T$; $f\left(x^*\right) = 0,0625$.

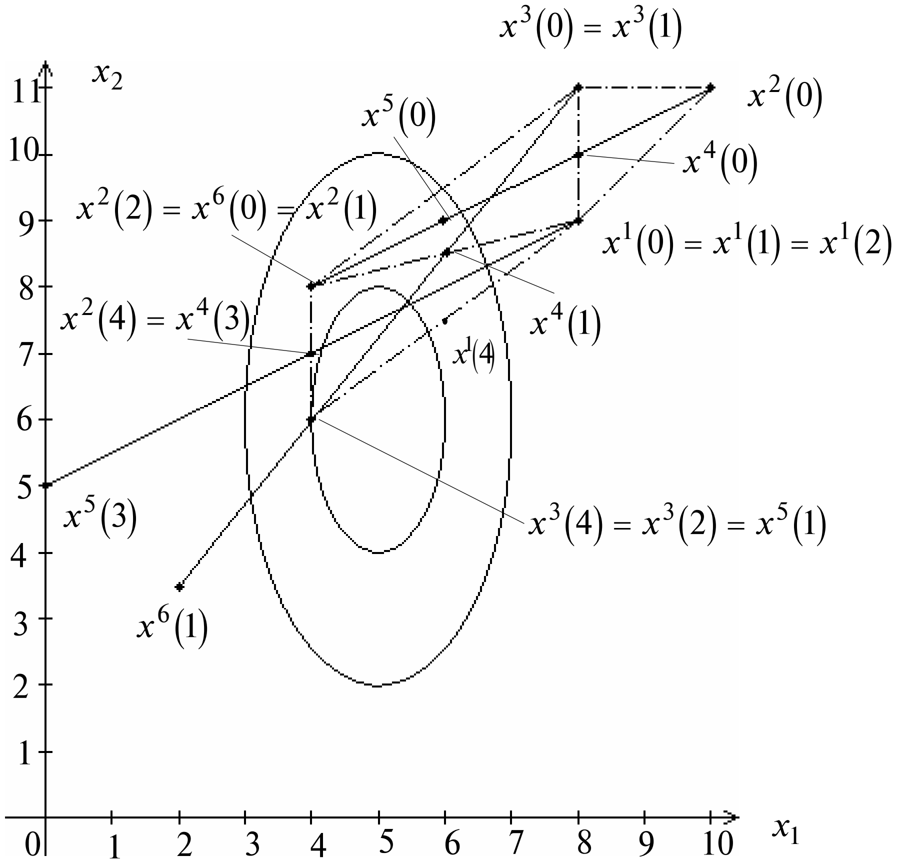

Рис. 5.14

Результаты расчетов до шага 6^2 приведены на рис. 5.14.■

5.4. МЕТОД РОЗЕНБРОКА

Постановка задачи

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

Определение 5.5. Пусть $d_1, d_2, \ldots, d_n$ – линейно независимые векторы, по норме равные единице. Они называются *взаимно ортогональными*, если для всех $i = 1, \ldots, n$ справедливо условие $d_i^T d_j = 0, j \neq i$.

Стратегия поиска

Суть метода Розенброка (Rosenbrock) состоит в следующем. Задается начальная точка. Из нее осуществляется итеративный поиск направления убывания функции с помощью изменяемых дискретных шагов вдоль n линейно независимых и ортогональных направлений. В случае удачного шага в исследуемом направлении его значение на следующей итерации увеличивается с помощью коэффициента растяжения, а в случае неудачи уменьшается за счет умножения на коэффициент сжатия (при этом направление поиска изменяется на противоположное). Поиск в системе текущих направлений проводится до тех пор, пока все возможности уменьшения функции не будут исчерпаны. Если по каждому направлению поиска имеет место неудача, строится новое множество линейно независимых и ортогональных направлений и циклический поиск по отдельным направлениям продолжается. Новые направления поворачиваются по отношению к предыдущим так, что они оказываются вытянутыми вдоль «оврага»[28] (рис. 5.15).

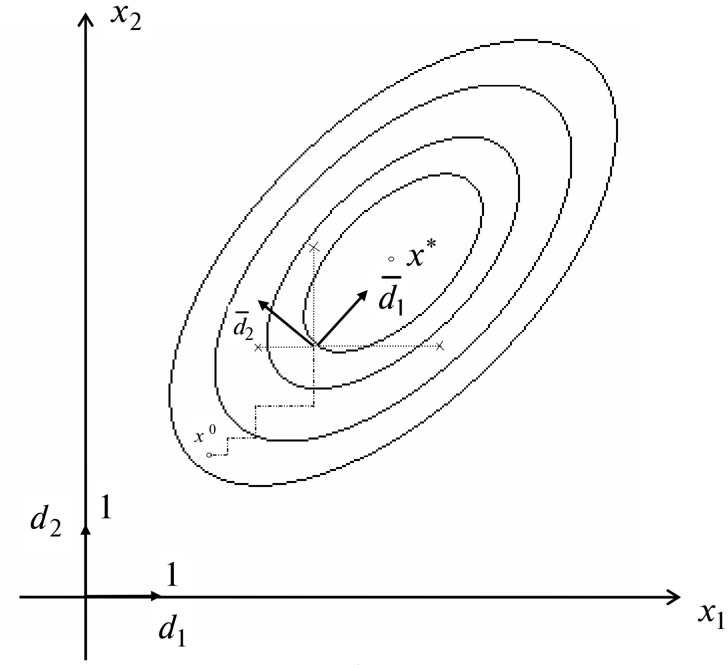

Рис. 5.15

Алгоритм

Шаг 1. Задать начальную точку x^0, число $\varepsilon > 0$ для остановки алгоритма, коэффициенты растяжения $\alpha > 1$ и сжатия $-1 < \beta < 0$, в качестве начальных линейно независимых и ортогональных направлений $d_1, d_2, \ldots, d_n$ выбрать координатные направления

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots,\ d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix};$$

начальную длину шага вдоль каждого из направлений поиска $\Delta_1{}^0, \ldots, \Delta_n{}^0 > 0$; N – максимальное число неудачных серий шагов по всем направлениям на одной итерации. Положить $y^1 = x^0$, $k = 0$, $i = 1$, $\Delta_i = \Delta_i{}^0$ для всех i.

Шаг 2. Сделать шаг по i-му направлению:

а) если $f\left(y^i + \Delta_i\, d_i\right) < f\left(y^i\right)$, шаг считается удачным. В этом случае следует положить $y^{i+1} = y^i + \Delta_i\, d_i$, $\Delta_i = \alpha\, \Delta_i$ и перейти к шагу 3;

б) если $f\left(y^i + \Delta_i\, d_i\right) \geq f\left(y^i\right)$, шаг считается неудачным. Тогда следует положить $y^{i+1} = y^i$, $\Delta_i = \beta\, \Delta_i$ и перейти к шагу 3.

Шаг 3. Проверить выполнение условий:

а) если $i < n$, то положить $i = i + 1$ и перейти к шагу 2 (сделать шаги по оставшимся направлениям);

б) если $i = n$, проверить успешность поиска по текущим ортогональным направлениям:

- если $f\left(y^{n+1}\right) < f\left(y^1\right)$, т.е. хотя бы один спуск по направлению на шаге 2 был успешным, положить: $y^1 = y^{n+1}, i = 1$ и перейти к шагу 2;
- если $f\left(y^{n+1}\right) = f\left(y^1\right)$, т.е. каждый из n последних шагов был неудачным, оценить успешность поиска на текущей итерации:
  - если $f\left(y^{n+1}\right) < f\left(x^k\right)$, т.е. на k-й итерации хотя бы один шаг удачный, то перейти к шагу 4;
  - если $f\left(y^{n+1}\right) = f\left(x^k\right)$, т.е. не было ни одного удачного шага на k-й итерации, процесс поиска приостановить. Если число l последовательно неудачных серий шагов по всем направлениям на текущей итерации не превышает N, проверить условие окончания, а иначе перейти к шагу 4. Проверяются величины Δ_i, использованные во время последней серии шагов. Если $\left|\Delta_i\right| \leq \varepsilon$ для всех i, то найдено приближенное решение задачи: $x^* \cong x^k$. Если $\left|\Delta_i\right| > \varepsilon$ хотя бы для одного i, то положить $y^1 = y^{n+1}$, $i = 1$ и перейти к шагу 2.

Шаг 4. Положить $x^{k+1} = y^{n+1}$ и проверить условие окончания:

а) если $\left\|x^{k+1} - x^k\right\| \leq \varepsilon$, то поиск завершить: $x^* \cong x^{k+1}$;

б) если $\left\|x^{k+1} - x^k\right\| > \varepsilon$, вычислить длины шагов по каждому направлению поиска на k-й итерации $\lambda_1, \ldots, \lambda_n$ из соотношения $x^{k+1} - x^k = \sum_{i=1}^{n} \lambda_i\, d_i$. Далее построить новый набор линейно независимых и взаимно ортогональных направлений поиска $\bar{d}_1, \ldots, \bar{d}_n$ с помощью *процедуры Грама–Шмидта*:

$$a_i = \begin{cases} d_i, & \lambda_i = 0, \\ \sum\limits_{j=i}^{n} \lambda_j\, d_j, & \lambda_i \neq 0, \end{cases} \qquad b_i = \begin{cases} a_i, & i = 1, \\ a_i - \sum\limits_{j=1}^{i-1} \left(a_i^T\, \bar{d}_j\right) \bar{d}_j, & i \geq 2, \end{cases} \qquad \bar{d}_i = \frac{b_i}{\| b_i \|}.$$

Заметим, что если $\lambda_i = 0$, то $\overline{d_i} = d_i$, т.е. новые направления следует вычислять только для тех индексов, для которых $\lambda_i \neq 0$. После нахождения новых направлений следует положить $\overline{d_i} = d_i$, $\Delta_i = \Delta_i^{\,0}$ для всех $i = 1,\ldots,n$, $k = k+1$, $y^1 = x^{k+1}$, $i = 1$ и перейти к шагу 2.

Замечания 5.11.

1. Если шаг 2 удачен, то Δ_i заменяется на $\alpha\,\Delta_i$, т.е. величина шага увеличивается, так как $\alpha > 1$. Неудача приводит к сдвигу в обратном направлении вдоль *i*-го направления при следующей попытке, так как $\beta < 0$.

2. Розенброк рекомендовал следующие коэффициенты растяжения и сжатия: $\alpha = 3,\ \beta = -0{,}5$.

3. Дэвис, Свенн и Кемпи (Davies, Swann, Campey) модифицировали метод Розенброка, применив алгоритмы одномерной минимизации при поиске вдоль каждого направления d_i [5, 41]. Тогда, если функция $f(x)$ дифференцируема, последовательность генерируемых точек сходится к стационарной точке.

Пример 5.10. Найти минимум функции $f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$ методом Розенброка.

□ 1. Зададим начальную точку $x^0 = (8,9)^T$, $\varepsilon = 0{,}6$; $\alpha = 2$, $\beta = -0{,}5$; $d_1 = (1,0)^T, d_2 = (0,1)^T$; $\Delta_1^{\,0} = 1$, $\Delta_2^{\,0} = 2$, $N = 3$. Положим $y^1 = x^0 = (8,9)^T$, $k = 0$, $i = 1$, $\Delta_1 = 1$, $\Delta_2 = 2$.

2^0. Так как $f\left(y^1 + \Delta_1 d_1\right) = f(9,9) = 73 > f\left(y^1\right) = f(8,9) = 45$, шаг неудачен: $y^2 = y^1 = (8,9)^T, \Delta_1 = -0{,}5 \cdot 1 = -0{,}5$.

3^0. Так как $i = 1 < n = 2$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^1. Поскольку $f\left(y^2 + \Delta_2 d_2\right) = f(8,11) = 61 > f\left(y^2\right) = 45$, шаг неудачен:

$$y^3 = y^2 = (8,9)^T,\ \Delta_2 = -0{,}5 \cdot 2 = -1.$$

3^1. Так как $i = 2 = n$, $f\left(y^{n+1}\right) = f\left(y^3\right) = f\left(x^0\right)$, то выполнена одна неудачная серия шагов: $l = 1 < N = 3$. Для выполненной серии шагов $\left|\Delta_1\right| = 1 > \varepsilon = 0{,}6$; $\left|\Delta_2\right| = 2 > \varepsilon = 0{,}6$, поэтому положим $y^1 = y^3 = (8,9)^T, i = 1$ и перейдем к шагу 2.

2^2. Поскольку $f\left(y^1 + \Delta_1 d_1\right) = f(7{,}5;9) = 34 < f\left(y^1\right) = f(8,9) = 45$, шаг удачен:

$$y^2 = y^1 + \Delta_1 d_1 = (7{,}5;9)^T,\ \Delta_1 = \alpha\,\Delta_1 = 2(-0{,}5) = -1.$$

3^2. Так как $i = 1 < n = 2$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^3. Поскольку $f\left(y^2 + \Delta_2 d_2\right) = f(7{,}5;8) = 29 < f\left(y^2\right) = 34$, шаг удачен:

$$y^3 = y^2 + \Delta_2 d_2 = (7{,}5;8)^T,\ \Delta_2 = \alpha\,\Delta_2 = 2(-1) = -2.$$

3^3. Так как $i = 2 = n = 2$, $f\left(y^3\right) = 29 < f\left(y^1\right) = 45$, то положим

$y^1 = y^3 = (7{,}5;8)^T$, $i = 1$ и перейдем к шагу 2.

2^4. Поскольку $f\left(y^1+\Delta_1 d_1\right)=f(6{,}5;8)=13<f\left(y^1\right)=f(7{,}5;8)=29$, шаг удачен: $y^2=y^1+\Delta_1 d_1=(6{,}5;8)^T, \quad \Delta_1=\alpha\,\Delta_1=2(-1)=-2$.

3^4. Так как $i=1<n=2$, то положим $i=i+1=2$ и перейдем к шагу 2.

2^5. Поскольку $f\left(y^2+\Delta_2 d_2\right)=f(6{,}5;6)=9<f\left(y^2\right)=f(6{,}5;8)=13$, шаг удачен: $y^3=y^2+\Delta_2 d_2=(6{,}5;6)^T, \quad \Delta_2=\alpha\,\Delta_2=2(-2)=-4$.

3^5. Так как $i=2=n, f\left(y^3\right)=9<f\left(y^1\right)=29$, положим $y^1=y^3=(6{,}5;6)^T, i=1$ и перейдем к шагу 2.

2^6. Поскольку $f\left(y^1+\Delta_1 d_1\right)=f(4{,}5;6)=1<f\left(y^1\right)=f(6{,}5;6)=9$, шаг удачен:

$$y^2=y^1+\Delta_1 d_1=(4{,}5;6)^T, \quad \Delta_1=\alpha\,\Delta_1=2(-2)=-4.$$

3^6. Так как $i=1<n=2$, то положим $i=i+1=2$ и перейдем к шагу 2.

2^7. Поскольку $f\left(y^2+\Delta_2 d_2\right)=f(4{,}5;2)=17>f\left(y^2\right)=1$, шаг неудачен:

$$y^3=y^2=(4{,}5;6)^T, \quad \Delta_2=\beta\,\Delta_2=-0{,}5(-4)=2.$$

3^7. Так как $i=2=n, \ f\left(y^3\right)=1<f\left(y^1\right)=f(6{,}5;6)=9$, то $y^1=y^3=(4{,}5;6)^T, \ i=1$ и перейдем к шагу 2.

2^8. Поскольку $f\left(y^1+\Delta_1 d_1\right)=f(0{,}5;6)=81>f\left(y^1\right)=f(4{,}5;6)=1$, шаг неудачен: $y^2=y^1=(4{,}5;6)^T, \ \Delta_1=-0{,}5(-4)=2$.

3^8. Так как $i=1<n=2$, то положим $i=i+1=2$ и перейдем к шагу 2.

2^9. Поскольку $f\left(y^2+\Delta_2 d_2\right)=f(4{,}5;8)=5>f\left(y^2\right)=1$, шаг неудачен:

$$y^3=y^2=(4{,}5;6)^T, \ \Delta_2=-0{,}5\cdot 2=-1.$$

3^9. Так как $i=2=n=2$ и $f\left(y^3\right)=f\left(y^1\right)$, но $f\left(y^3\right)=1<f\left(x^0\right)=45$, перейдем к шагу 4.

4^0. Положим $x^1=y^3=(4{,}5;6)^T$. Поскольку $\left\|x^1-x^0\right\|=\left\|(8,\,9)^T-(4{,}5;\,6)^T\right\|=\sqrt{3{,}5^2+3^2}=4{,}61>\varepsilon=0{,}6$, то вычислим λ_1,λ_2 из соотношения $x^1-x^0=\begin{pmatrix}4{,}5\\6\end{pmatrix}-\begin{pmatrix}8\\9\end{pmatrix}=\begin{pmatrix}-3{,}5\\-3\end{pmatrix}=\lambda_1\begin{pmatrix}1\\0\end{pmatrix}+\lambda_2\begin{pmatrix}0\\1\end{pmatrix}$: $\lambda_1=-3{,}5; \ \lambda_2=-3$.

Построим новый набор направлений поиска:

$$a_1=\sum_{j=1}^{n=2}\lambda_j d_j=-3{,}5\begin{pmatrix}1\\0\end{pmatrix}-3\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-3{,}5\\-3\end{pmatrix}; \quad a_2=\sum_{j=2}^{n=2}\lambda_j d_j=-3\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\-3\end{pmatrix};$$

$$b_1=a_1=\begin{pmatrix}-3{,}5\\-3\end{pmatrix}, \text{ так как } i=1; \quad \overline{d}_1=\frac{b_1}{\|b_1\|}=\frac{\begin{pmatrix}-3{,}5\\-3\end{pmatrix}}{4{,}61}=\begin{pmatrix}-0{,}76\\-0{,}65\end{pmatrix};$$

$$b_2 = a_2 - \sum_{j=1}^{2-1}\left(a_2^T \overline{d}_j\right)\overline{d}_j = \begin{pmatrix} 0 \\ -3 \end{pmatrix} - \left(\begin{pmatrix} 0 & -3 \end{pmatrix}\begin{pmatrix} -0{,}76 \\ -0{,}65 \end{pmatrix}\right)\begin{pmatrix} -0{,}76 \\ -0{,}65 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 \\ -3 \end{pmatrix} - 1{,}95\begin{pmatrix} -0{,}76 \\ -0{,}76 \end{pmatrix} = \begin{pmatrix} 1{,}482 \\ -1{,}729 \end{pmatrix}; \qquad \overline{d}_2 = \frac{b_2}{\|b_2\|} = \frac{\begin{pmatrix} 1{,}482 \\ -1{,}729 \end{pmatrix}}{2{,}28} = \begin{pmatrix} 0{,}65 \\ -0{,}76 \end{pmatrix}.$$

Положим $d_1 = \overline{d}_1$, $d_2 = \overline{d}_2$, $\Delta_1 = \Delta_1^0 = 1$, $\Delta_2 = \Delta_2^0 = 2$, $k = k + 1 = 1$, $y^1 = x^1 = (4{,}5; 6)^T$, $i = 1$ и перейдем к шагу 2.

2^{10}. Поскольку $f\left(y^1 + \Delta_1 d_1\right) = f(3{,}74; 5{,}35) = 6{,}766 > f\left(y^1\right) = f(4{,}5; 6) = 1$, шаг неудачен: $y^2 = y^1 = (4{,}5; 6)^T$, $\Delta_1 = -0{,}5 \cdot 1 = -0{,}5$.

3^{10}. Так как $i = 1 < n = 2$, то положим $i = i + 1 = 2$.

2^{11}. Поскольку $f\left(y^2 + \Delta_2 d_2\right) = f(5{,}801; 4{,}481) = 4{,}876 > f\left(y^2\right) = 1$, шаг неудачен: $y^3 = y^2 = (4{,}5; 6)^T$, $\Delta_2 = -0{,}5 \cdot 2 = -1$.

3^{11}. Так как $i = 2 = n$, $f\left(y^3\right) = f\left(y^1\right) = 1$, то оценим успешность поиска на текущей итерации. Так как $f\left(y^3\right) = f\left(x^1\right) = 1$, то на текущей итерации не было ни одного удачного шага. Поскольку $l = 1 < N = 3$, то проверим условие окончания. Имеем $|\Delta_2| = 1 > \varepsilon = 0{,}6$, поэтому положим $y^1 = y^3 = (4{,}5; 6)^T$, $i = 1$ и перейдем к шагу 2.

2^{12}. Поскольку $f\left(y^1 + \Delta_1 d_1\right) = f(4{,}880; 6{,}325) = 0{,}164 < f\left(y^1\right) = f(4{,}5; 6) = 1$, шаг удачен: $y^2 = y^1 + \Delta_1 d_1 = (4{,}880; 6{,}325)^T$, $\Delta_1 = \alpha\,\Delta_1 = -0{,}5 \cdot 2 = -1$.

3^{12}. Так как $i = 1 < n = 2$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^{13}. Поскольку $f\left(y^2 + \Delta_2 d_2\right) = f(4{,}229; 7{,}085) = 3{,}555 > f\left(y^2\right) = 0{,}164$, шаг неудачен: $y^3 = y^2 = (4{,}880; 6{,}325)^T$, $\Delta_2 = \beta\,\Delta_2 = -0{,}5(-1) = 0{,}5$.

3^{13}. Так как $i = 2 = n$, то проверим успешность поиска по текущим ортогональным направлениям: $f\left(y^3\right) = 0{,}164 < f\left(y^1\right) = 1$ – поиск успешный. Положим $y^1 = y^3 = (4{,}880; 6{,}325)^T$, $i = 1$ и перейдем к шагу 2.

2^{14}. Поскольку справедливо неравенство

$$f\left(y^1 + \Delta_1 d_1\right) = f(5{,}639; 6{,}976) = 2{,}586 > f\left(y^1\right) = f(4{,}880; 6{,}325) = 0{,}164,$$

шаг неудачен: $y^2 = y^1 = (4{,}880; 6{,}325)^T$, $\Delta_1 = -0{,}5 \cdot (-1) = 0{,}5$.

3^{14}. Так как $i = 1 < n = 2$, то положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^{15}. Поскольку $f\left(y^2 + \Delta_2 d_2\right) = f(5{,}205; 5{,}946) = 0{,}171 > f\left(y^2\right) = 0{,}164$, шаг неудачен: $y^3 = y^2 = (4{,}880; 6{,}325)^T$, $\Delta_2 = -0{,}5 \cdot 0{,}5 = -0{,}25$.

3^{15}. Так как $i = 2 = n$ и $f\left(y^3\right) = 0{,}164 = f\left(y^1\right)$, то оценим успешность поиска на текущей итерации: $f\left(y^3\right) = 0{,}164 < f\left(x^1\right) = 1$ – на текущей итерации был удачный шаг. Перейдем к шагу 4.

4^1. Положим $x^2 = y^3 = \left(4{,}880; 6{,}325\right)^T$. Так как выполняется условие окончания $\left\| x^2 - x^1 \right\| = \sqrt{\left(4{,}880 - 4{,}5\right)^2 + \left(6{,}325 - 6\right)^2} = 0{,}5 < \varepsilon = 0{,}6$, то процесс поиска завершается: $x^* = x^2 = \left(4{,}880; 6{,}325\right)^T$; $f\left(x^*\right) = 0{,}164$. ■

5.5. МЕТОД СОПРЯЖЕННЫХ НАПРАВЛЕНИЙ

Постановка задачи

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

Определение 5.6. Пусть H - симметрическая матрица размеров ($n \times n$). Векторы $d_1, d_2, \ldots, d_n$ называются *H-сопряженными* или просто *сопряженными*, если $d_i^T H d_j = 0$ при всех $i \neq j$.

Стратегия поиска

В методе сопряженных направлений, или методе Пауэлла (Powell), используется тот факт, что минимум квадратичной функции может быть найден не более чем за *n* шагов при условии, что поиск ведется вдоль сопряженных относительно матрицы Гессе направлений. Так как достаточно большой класс целевых функций может быть представлен в окрестности точки минимума своей квадратичной аппроксимацией, описанная идея применяется и для неквадратичных функций. Задаются начальная точка и направления $d_1, d_2, \ldots, d_n$, совпадающие с координатными. Находится минимум $f(x)$ при последовательном движении по $(n+1)$ направлениям с помощью одного из методов одномерной минимизации (см. разд. 5.1). При этом полученная ранее точка минимума берется в качестве исходной для поиска по следующему направлению, а направление d_n используется как при первом ($d_0 = d_n$), так и при последнем поиске. Находится новое направление поиска, сопряженное с d_n. Оно проходит через точки, полученные при первом и последнем поиске. Заменяется d_1 на d_2, d_2 на d_3 и т.д. Направление d_n заменяется сопряженным направлением, после чего повторяется поиск по $(n+1)$ направлениям, уже не содержащим старого направления d_1. Для квадратичных функций последовательность n^2 одномерных поисков приводит к точке минимума (если все операции выполнены точно).

Построение сопряженного направления для квадратичной функции при $n = 2$ изображено на рис. 5.16. Оно проходит через точки 1 и 3.

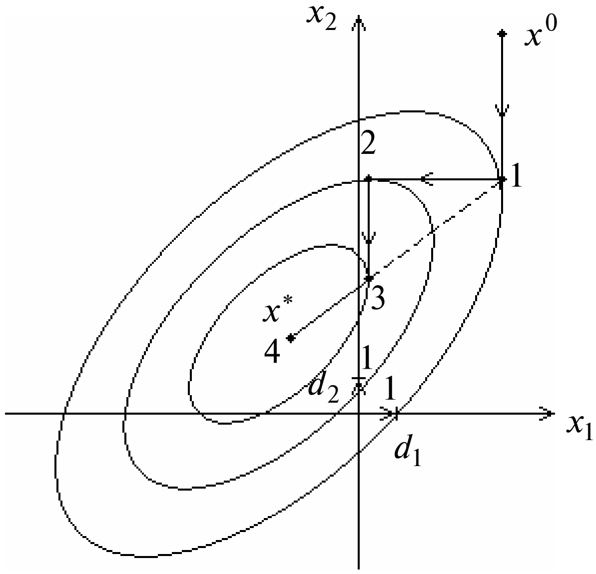

Рис. 5.16

Алгоритм

Шаг 1. Задать начальную точку x^0, число $\varepsilon > 0$ для окончания алгоритма, начальные направления поиска

$$d_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ d_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots,\ d_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Положим $d_0 = d_n, i = 0,\ y^0 = x^0,\ k = 0$.

Шаг 2. Найти $y^{i+1} = y^i + t_i\, d_i$, где шаг t_i находится в результате поиска минимума функции $f\left(y^i + t_i\, d_i\right)$ по t_i одним из методов одномерной минимизации.

Шаг 3. Проверить выполнение условий:

а) если $i < n - 1$, положить $i = i + 1$ и перейти к шагу 2;

б) если $i = n - 1$, проверить успешность поиска по *n* первым направлениям. Если $y^n = y^0$, то поиск завершить: $x^* \cong y^n$, иначе положить $i = i + 1$ и перейти к шагу 2;

в) если $i = n$, проверить успешность поиска по *n* последним направлениям. Если $y^{n+1} = y^1$, поиск завершить: $x^* \cong y^{n+1}$, иначе перейти к шагу 4 для построения сопряженного направления.

Шаг 4. Положить $x^{k+1} = y^{n+1}$ и проверить условие окончания:

а) если $\left\| x^{k+1} - x^k \right\| < \varepsilon$, то поиск завершить: $x^* \cong x^{k+1}$;

б) иначе положить $\overline{d}_0 = \overline{d}_n = y^{n+1} - y^1$ (новое направление);

$\overline{d}_i = d_{i+1}$, $i = 1,\dots,n-1$ (исключается старое направление).

Проверить выполнение условия:

- если $\operatorname{rang}\left(\overline{d}_1,\dots,\overline{d}_n\right) = n$, то новая система направлений линейно независима. В этом случае положить $\overline{d}_i = d_i$, $i = 0,1,\dots,n;$ $k = k+1,$ $i = 0$, $y^0 = x^{k+1}$ и перейти к шагу 2;
- если $\operatorname{rang}\left(\overline{d}_1,\dots,\overline{d}_n\right) < n$, то новая система направлений линейно зависима. Тогда следует продолжать поиск в старых направлениях. Для этого положить $d_i = d_i$, $i = 0,1,\dots,n$; $y^0 = x^{k+1}$, $k = k+1$, $i = 0$ и перейти к шагу 2.

З а м е ч а н и е 5.12. Изложенный алгоритм соответствует описанному в [36]. Существует алгоритм Пауэлла, в котором не гарантируется линейная независимость направлений поиска, а в [5] приведена модификация алгоритма Пауэлла, предложенная Зангвиллом (Zangwill). Последняя модификация гарантирует линейную независимость направлений поиска и сходимость за конечное число шагов.

Пример 5.11. Найти минимум функции $f(x) = 4\left(x_1 - 5\right)^2 + \left(x_2 - 6\right)^2$ методом Пауэлла.

□ 1^0. Зададим начальную точку $x^0 = (8,9)^T$, $d_1 = \begin{pmatrix}1\\0\end{pmatrix}, d_2 = \begin{pmatrix}0\\1\end{pmatrix}$, $\varepsilon = 0{,}1$. Положим $d_0 = d_n = d_2$; $y^0 = x^0$, $i = 0,\ k = 0$.

2^0. Получим $y^1 = y^0 + t_0\, d_0 = (8,9)^T + t_0(0,1)^T = (8, 9 + t_0)^T$. Найдем минимум функции $f(8, 9 + t_0) = 36 + (3 + t_0)^2$ по t_0. Очевидно, $t_0 = -3$, а $y^1 = (8,6)^T$.

3^0. Имеем $i = 0 < 2 = n$, поэтому положим $i = i + 1 = 1$ и перейдем к шагу 2.

2^1. Получим $y^2 = y^1 + t_1\, d_1 = (8,6)^T + t_1(1,0)^T = (8 + t_1, 6)^T$. Найдем минимум функции $f(8 + t_1, 6) = 4(3 + t_1)^2$ по t_1. Он достигается при $t_1 = -3$, тогда $y^2 = (5,6)^T$.

3^1. Имеем $i = 1 = n - 1$, $y^n = y^2 \ne y^0$, поэтому положим $i = i + 1 = 2$ и перейдем к шагу 2.

2^2. Получим $y^3 = y^2 + t_2\, d_2 = (5,6)^T + t_2(0,1)^T = (5, 6 + t_2)^T$. Найдем минимум функции $f(5, 6 + t_2) = t_2^2$ по t_2. Очевидно, $t_2 = 0$, а $y^3 = y^2 = (5,6)^T$.

3^2. Имеем $i = 2 = n,$ $y^3 \ne y^1$. Перейдем к шагу 4.

4^0. Находим $x^1 = y^3 = (5,6)^T$, $\left\|x^1 - x^0\right\| = \sqrt{(8-5)^2 + (9-6)^2} = 4{,}24 > \varepsilon$.

Положим $\overline{d}_0 = \overline{d}_n = \overline{d}_2 = y^3 - y^1 = (5,6)^T - (8,6)^T = (-3,0)^T$; $\overline{d}_1 = d_2 = \begin{pmatrix}0\\1\end{pmatrix}$.

Так как $\operatorname{rang}\begin{pmatrix}-3 & 0\\0 & 1\end{pmatrix} = 2 = n$, то система векторов линейно независима. Положим

$d_2 = \overline{d}_2 = \begin{pmatrix} -3 \\ 0 \end{pmatrix}$, $d_1 = \overline{d}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $d_0 = \overline{d}_0 = \begin{pmatrix} -3 \\ 0 \end{pmatrix}$, $k = k+1$, $i = 0$, $y^0 = x^1 = (5,6)^T$ и перейдем к шагу 2.

2^3. Получим $y^1 = y^0 + t_0\, d_0 = (5,6)^T + t_0\,(-3,0)^T = (5 - 3t_0, 6)^T$. Найдем минимум функции $f(5 - 3t_0, 6) = 36t_0^2$ по t_0. Так как $t_0 = 0$, то $y^1 = (5,6)^T = y^0$.

3^3. Имеем $i = 0 < n - 1 = 1$, поэтому положим $i = i + 1 = 1$ и перейдем к шагу 2.

2^4. Получим $y^2 = y^1 + t_1\, d_1 = (5,6)^T + t_1(0,1)^T = (5, 6 + t_1)^T$. Минимум функции $f(5, 6 + t_1) = t_1^2$ по t_1 достигается при $t_1 = 0$. Тогда $y^2 = (5,6)^T = y^1 = y^0$.

3^4. Имеем $i = 1 = n - 1$, $y^2 = y^0$, поэтому поиск завершается: $x^* \cong y^2 = (5,\, 6)^T$; $f\left(x^*\right) = 0$. ■

5.6. МЕТОДЫ СЛУЧАЙНОГО ПОИСКА

5.6.1. Адаптивный метод случайного поиска

Постановка задачи

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

Стратегия поиска

Задается начальная точка x^0. Каждая последующая точка находится по формуле

$$x^{k+1} = x^k + t_k\, \xi^k,$$

где $t_k > 0$ – величина шага; ξ^k – случайный вектор единичной длины, определяющий направление поиска; k – номер итерации. На текущей итерации при помощи генерирования случайных векторов ξ^k получаются точки, лежащие на гиперсфере радиуса t_k с центром в точке x^k (рис. 5.17). Если значение функции в полученной точке не меньше, чем в центре, шаг считается неудачным (точки y^1, y^2 при поиске из x^0; y^1, y^3 при поиске из x^1). Если число неудачных шагов из текущей точки достигает некоторого числа M, дальнейший поиск продолжается из той же точки, но с меньшим шагом до тех пор, пока он не станет меньше заранее заданной величины R. Если же значение функции в полученной точке меньше, чем в центре, шаг считается удачным и в найденном направлении делается увеличенный шаг, играющий роль ускоряющего шага (как при поиске по образцу в методе конфигураций). Если при этом значение функции снова меньше, чем в центре, направление считается удачным и дальнейший поиск продолжается из этой точки (точки $z^3 = x^1$ при поиске из x^0, $z^4 = x^2$ при поиске из x^1). Если же значение функции не стало меньше, чем в центре, направление считается неудачным и поиск продолжается из старого центра (в точке y^2 при поиске из x^1 функция меньше, чем в x^1, а в точке z^2 уже не меньше, поэтому направление $\left(z^2 - x^1\right)$ – неудачное).

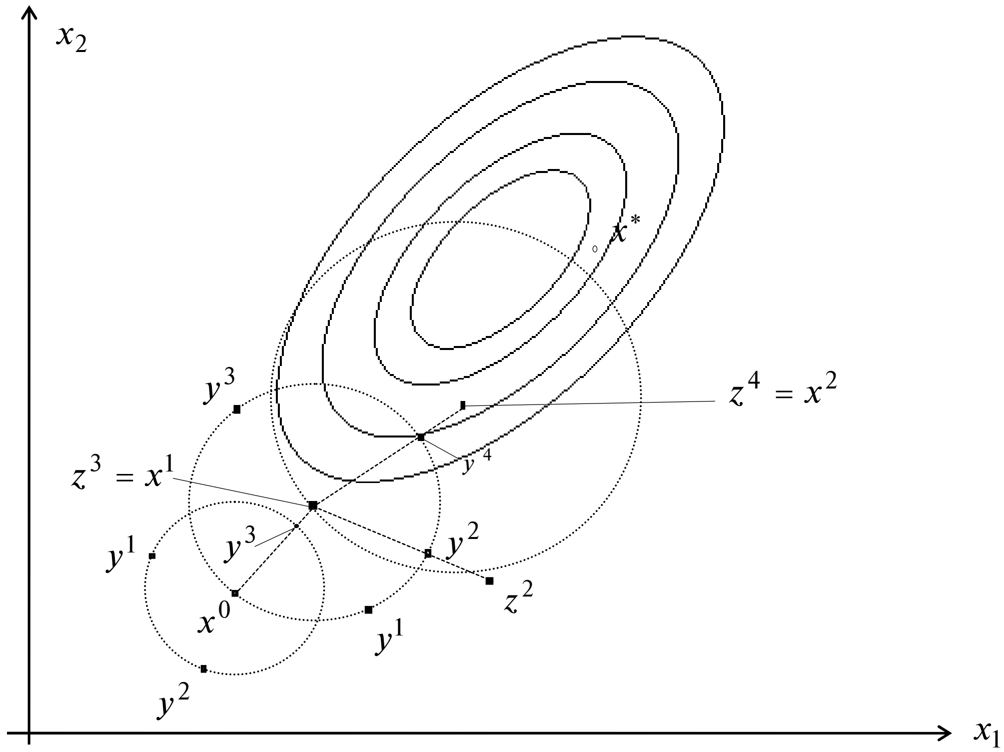

Рис. 5.17

Алгоритм

Шаг 1. Задать начальную точку x^0, коэффициенты расширения $\alpha \ge 1$ и сжатия $0 < \beta < 1$, M – максимальное число неудачно выполненных испытаний на текущей итерации, $t_0 = 1$ – начальную величину шага, R - минимальную величину шага, N – максимальное число итераций. Положить $k = 0, j = 1$.

Шаг 2. Получить случайный вектор $\xi^j = \left(\xi_1{}^j, \dots, \xi_n{}^j\right)^T$, где $\xi_i{}^j$ – случайная величина, равномерно распределенная на интервале $\left[-1,1\right]$.

Шаг 3. Вычислить $y^j = x^k + t_k \dfrac{\xi^j}{\left\|\xi^j\right\|}$.

Шаг 4. Проверить выполнение условий:

а) если $f\left(y^j\right) < f\left(x^k\right)$, шаг удачный. Положить $z^j = x^k + \alpha\left(y^j - x^k\right)$ и определить, является ли текущее направление $y^j - x^k$ удачным:

- если $f(z^j) < f(x^k)$, то направление поиска удачное. Положить $x^{k+1} = z^j$, $t_{k+1} = \alpha t_k$, $k = k+1$ и проверить условие окончания. Если $k < N$, положить $j = 1$ и перейти к шагу 2. Если $k = N$, поиск завершить: $x^* \cong x^k$;
- если $f(z^j) \ge f(x^k)$, направление поиска неудачное, перейти к шагу 5;

б) если $f(y^j) \ge f(x^k)$, шаг неудачный и перейти к шагу 5.

Шаг 5. Оценить число неудачных шагов из текущей точки:

а) если $j < M$, следует положить $j = j + 1$ и перейти к шагу 2;
б) если $j = M$, проверить условие окончания:

- если $t_k \le R$, процесс закончить: $x^* \cong x^k, f(x^*) \cong f(x^k)$;
- если $t_k > R$, положить $t_k = \beta t_k$, $j = 1$ и перейти к шагу 2.

Замечания 5.13.

1. Величина ξ_i^j, равномерно распределенная на интервале $[-1,1]$, генерируется обычно с помошью датчиков псевдослучайных чисел на ЭВМ. Вырабатывается случайная величина η_i^j, равномерно распределенная на $[0,1]$, а затем используется линейное преобразование: $\xi_i^j = 2\eta_i^j - 1$.

2. Шумер и Стейглиц (Schumer, Steiglitz) рекомендуют следующие параметры алгоритма: $\alpha = 1{,}618$; $\beta = 0{,}618$; $M = 3n$. При $\alpha = 1$ точка z^j на шаге 4 совпадает с y^j, т.е. аналог поиска по образцу не производится. Начальный шаг $t_0 \ge R$ можно задать произвольно [36] .

3. Если выполнено условие окончания $t_k \le R$, то в качестве ответа можно использовать любую точку внутри шара с радиусом t_k и центром в точке x^k .

4. Многочисленные варианты случайного поиска изложены в [12] и могут включать элементы обучения, при котором направления убывания функции становятся более вероятными, а другие направления – менее вероятными.

Пример 5.12. Найти минимум функции $f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$ методом адаптивного случайного поиска.

☐ 1^0. Зададим начальную точку $x^0 = (8,9)^T$, $\alpha = 1{,}618$; $\beta = 0{,}618$; $N = 10$; $R = 0{,}8$; $t_0 = 1$; $M = 3$. Положим $k = 0, j = 1$.

2^0. Получим $\xi^1 = (0{,}843; 0{,}374)^T$.

3^0. Вычислим $y^1 = x^0 + t_0 \dfrac{\xi^1}{\|\xi^1\|} = (8,\ 9)^T + \dfrac{(0{,}843;\ 0{,}374)^T}{0{,}922} = (8{,}914;\ 9{,}4)^T$.

4^0. Так как $f(y^1) = 72{,}83 > f(x^0) = 45$, шаг неудачен.

5^0. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^1. Получим $\xi^2 = (0{,}239; 0{,}954)^T$.

3^1. Вычислим $y^2 = x^0 + t_0 \dfrac{\xi^2}{\left\|\xi^2\right\|} = (8,\ 9)^T + \dfrac{(0{,}239;\ 0{,}954)^T}{0{,}983} = (8{,}24;\ 9{,}97)^T$.

4^1. Так как $f\left(y^2\right) = 57{,}75 > f\left(x^0\right) = 45$, шаг неудачен.

5^1. Имеем $j = 2 < M = 3$. Положим $j = j + 1 = 3$ и перейдем к шагу 2.

2^2. Получаем $\xi^2 = (-0{,}159; -0{,}402)^T$.

3^2. Вычислим $y^3 = x^0 + t_0 \dfrac{\xi^3}{\left\|\xi^3\right\|} = (8,\ 9)^T + \dfrac{(-0{,}159;\ -0{,}402)^T}{0{,}432} = (7{,}63;\ 8{,}07)^T$.

4^2. Так как $f\left(y^3\right) = 31{,}95 < f\left(x^0\right) = 45$, шаг удачный. Положим

$$z^3 = x^0 + \alpha\left(y^3 - x^0\right) = (8,9)^T + 1{,}618\left[(7{,}63; 8{,}07)^T - (8,9)^T\right] = (7{,}4; 7{,}49)^T,$$

$f\left(z^3\right) = 25{,}26 < f\left(x^0\right) = 45$, направление удачное. Положим $x^1 = z^3 = (7{,}4; 7{,}49)^T$, $t_1 = \alpha t_0 = 1{,}618 \cdot 1 = 1{,}618$, $k = k + 1 = 1$. Так как $k = 1 < N = 10$, положим $j = 1$ и перейдем к шагу 2.

2^3. Получим $\xi^2 = (0{,}168; -0{,}727)^T$.

3^3. Вычислим

$$y^1 = x^1 + t_1 \frac{\xi^1}{\left\|\xi^1\right\|} = (7{,}4;\ 7{,}49)^T + 1{,}618 \frac{(0{,}168;\ -0{,}727)^T}{0{,}747} = (7{,}67;\ 5{,}82)^T.$$

4^3. Так как $f\left(y^1\right) = 29{,}19 > f\left(x^1\right) = 25{,}26$, шаг неудачный.

5^2. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^4. Получим $\xi^2 = (-0{,}478; -0{,}214)^T$.

3^4. Вычислим

$$y^2 = x^1 + t_1 \frac{\xi^2}{\left\|\xi^2\right\|} = (7{,}4;\ 7{,}49)^T + 1{,}618 \frac{(-0{,}478; -0{,}214)^T}{0{,}524} = (5{,}92;\ 6{,}83)^T.$$

4^4. Так как $f\left(y^2\right) = 4{,}07 < f\left(x^1\right) = 25{,}26$, шаг удачный. Положим

$$z^2 = x^1 + \alpha\left(y^2 - x^1\right) = (7{,}4; 7{,}49)^T + 1{,}618\left[(5{,}92; 6{,}83)^T - (7{,}4; 7{,}49)^T\right] = (5{,}005; 6{,}42)^T,$$

$f\left(z^2\right) = 0{,}176 < f\left(x^1\right) = 25{,}26$, направление удачное.

Положим $x^2 = z^2 = (5{,}005; 6{,}42)^T$, $t_2 = \alpha t_1 = 2{,}618$, $k = k + 1 = 2$. Так как $k = 2 < N = 10$, положим $j = 1$ и перейдем к шагу 2.

2^5. Получим $\xi^1 = (-0{,}361; 0{,}112)^T$.

3^5. Вычислим

$$y^1 = x^2 + t_2 \frac{\xi^1}{\left\|\xi^1\right\|} = (5{,}005;\ 6{,}42)^T + 2{,}618 \frac{(-0{,}361;\ 0{,}112)^T}{0{,}378} = (2{,}93;\ 7{,}19)^T.$$

4^5. Так как $f\left(y^1\right) = 18{,}55 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^3. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^6. Получим $\xi^2 = \left(0{,}674; 0{,}551\right)^T$.

3^6. Вычислим

$$y^2 = x^2 + t_2 \frac{\xi^2}{\left\|\xi^2\right\|} = \left(5{,}005;\ 6{,}42\right)^T + 2{,}618 \frac{\left(0{,}674;\ 0{,}551\right)^T}{0{,}87} = \left(7{,}03;\ 8{,}08\right)^T.$$

4^6. Так как $f\left(y^1\right) = 20{,}81 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^4. Имеем $j = 2 < M = 3$. Положим $j = j + 1 = 3$ и перейдем к шагу 2.

2^7. Получим $\xi^3 = \left(0{,}789; -0{,}742\right)^T$.

3^7. Вычислим

$$y^3 = x^2 + t_2 \frac{\xi^3}{\left\|\xi^3\right\|} = \left(5{,}005;\ 6{,}42\right)^T + 2{,}618 \frac{\left(0{,}789;\ -0{,}742\right)^T}{1{,}083} = \left(6{,}91;\ 4{,}63\right)^T.$$

4^7. Так как $f\left(y^3\right) = 14{,}73 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^5. Имеем $j = 3 = M$. Так как $t_2 = 2{,}618 > R = 0{,}8$, положим $t_2 = 0{,}618\, t_2 = 0{,}618 \cdot 2{,}618 = 1{,}618$, $j = 1$ и перейдем к шагу 2.

2^8. Получим $\xi^1 = \left(-0{,}824; -0{,}193\right)^T$.

3^8. Вычислим

$$y^1 = x^2 + t_2 \frac{\xi^1}{\left\|\xi^1\right\|} = \left(5{,}005;\ 6{,}42\right)^T + 1{,}618 \frac{\left(-0{,}824;\ -0{,}193\right)^T}{0{,}846} = \left(3{,}43;\ 6{,}05\right)^T.$$

4^8. Так как $f\left(y^1\right) = 9{,}86 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^6. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^9. Получаем $\xi^2 = \left(-0{,}08; 0{,}917\right)^T$.

3^9. Вычислим

$$y^2 = x^2 + t_2 \frac{\xi^2}{\left\|\xi^2\right\|} = \left(5{,}005;\ 6{,}42\right)^T + 1{,}618 \frac{\left(-0{,}08;\ 0{,}917\right)^T}{0{,}92} = \left(4{,}86;\ 8{,}03\right)^T.$$

4^9. Так как $f\left(y^2\right) = 4{,}19 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^7. Имеем $j = 2 < M = 3$. Положим $j = j + 1 = 3$ и перейдем к шагу 2.

2^{10}. Получим $\xi^3 = \left(0{,}05; 0{,}171\right)^T$.

3^{10}. Вычислим

$$y^3 = x^2 + t_2 \frac{\xi^3}{\left\|\xi^3\right\|} = \left(5{,}005;\ 6{,}42\right)^T + 1{,}618 \frac{\left(0{,}05;\ 0{,}171\right)^T}{0{,}178} = \left(5{,}46;\ 7{,}97\right)^T.$$

4^{10}. Так как $f\left(y^3\right) = 4{,}73 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^8. Имеем $j = 3 = M$. Так как $t_2 = 1{,}618 > R = 0{,}8$, положим $t_2 = 0{,}618\, t_2 = 0{,}618 \cdot 1{,}618 = 1$, $j = 1$ и перейдем к шагу 2.

2^{11}. Получим $\xi^1 = (0{,}251; -0{,}447)^T$.

3^{11}. Вычислим

$$y^1 = x^2 + t_2 \frac{\xi^1}{\left\|\xi^1\right\|} = (5{,}005;\ 6{,}42)^T + 1 \frac{(0{,}251;\ -0{,}447)^T}{0{,}51} = (5{,}5;\ 5{,}54)^T.$$

4^{11}. Так как $f\left(y^1\right) = 1{,}21 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^9. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^{12}. Получим $\xi^2 = (-0{,}812; 0{,}102)^T$.

3^{12}. Вычислим

$$y^2 = x^2 + t_2 \frac{\xi^2}{\left\|\xi^2\right\|} = (5{,}005;\ 6{,}42)^T + 1 \cdot \frac{(-0{,}812;\ 0{,}102)^T}{0{,}818} = (4{,}01;\ 6{,}54)^T.$$

4^{12}. Так как $f\left(y^2\right) = 4{,}12 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^{10}. Имеем $j = 2 < M = 3$. Положим $j = j + 1 = 3$ и перейдем к шагу 2.

2^{13}. Получим $\xi^3 = (0{,}507; 0{,}537)^T$.

3^{13}. Вычислим

$$y^3 = x^2 + t_2 \frac{\xi^3}{\left\|\xi^3\right\|} = (5{,}005;\ 6{,}42)^T + 1 \frac{(0{,}507;\ 0{,}537)^T}{0{,}738} = (5{,}69;\ 7{,}15)^T.$$

4^{13}. Так как $f\left(y^3\right) = 3{,}23 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^{11}. Имеем $j = 3 = M$. Так как $t_2 = 1 > R = 0{,}8$, то положим $t_2 = 0{,}618\, t_2 = 0{,}618$, $j = 1$ и перейдем к шагу 2.

2^{14}. Получим $\xi^1 = (-0{,}587; 0{,}461)^T$.

3^{14}. Вычислим

$$y^1 = x^2 + t_2 \frac{\xi^1}{\left\|\xi^1\right\|} = (5{,}005;\ 6{,}42)^T + 0{,}618 \frac{(-0{,}587;\ 0{,}461)^T}{0{,}746} = (4{,}52;\ 6{,}8)^T.$$

4^{14}. Так как $f\left(y^1\right) = 1{,}56 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^{12}. Имеем $j = 1 < M = 3$. Положим $j = j + 1 = 2$ и перейдем к шагу 2.

2^{15}. Получим $\xi^2 = (0{,}911; 0{,}018)^T$.

3^{15}. Вычислим

$$y^2 = x^2 + t_2 \frac{\xi^2}{\left\|\xi^2\right\|} = (5{,}005; 6{,}42)^T + 0{,}618 \frac{(0{,}911; 0{,}018)^T}{0{,}9112} = (5{,}62; 6{,}43)^T.$$

4^{15}. Так как $f\left(y^2\right) = 1{,}72 > f\left(x^2\right) = 0{,}176$, шаг неудачен.

5^{13}. Имеем $j = 2 < M = 3$. Положим $j = j + 1 = 3$ и перейдем к шагу 2.

2^{16}. Получим $\xi^3 = (-0{,}07; -0{,}971)^T$.

3^{16}. Вычислим

$$y^3 = x^2 + t_2 \frac{\xi^3}{\left\|\xi^3\right\|} = (5{,}005;\ 6{,}42)^T + 0{,}618 \frac{(-0{,}07;\ -0{,}971)^T}{0{,}973} = (4{,}96;\ 5{,}803)^T.$$

4^{16}. Так как $f\left(y^3\right) = 0{,}046 < f\left(x^2\right) = 0{,}176$, шаг удачен. Положим

$$z^3 = x^2 + \alpha\left(y^3 - x^2\right) = (5{,}005; 6{,}42)^T + 1{,}618\left[(4{,}96; 5{,}803)^T - (5{,}005; 6{,}42)^T\right] =$$

$= (4{,}93; 5{,}42)^T$, $f\left(z^3\right) = 0{,}356 > f\left(x^2\right) = 0{,}176$, направление поиска неудачное. Перейдем к шагу 5.

5^{14}. Имеем $j = 3 = M$, $t_2 = 0{,}618 < R = 0{,}8$. Поэтому $x^* \cong x^2 = (5{,}005; 6{,}42)^T$, $f\left(x^*\right) \cong 0{,}176$ или, более точно, результат содержится в круге с радиусом $t_2 = 0{,}618$ и центром в точке x^2. ■

5.6.2. Метод случайного поиска с возвратом при неудачном шаге

Постановка задачи

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

Стратегия поиска

Задается начальная точка x^0. Каждая последующая точка находится по формуле

$$x^{k+1} = x^k + t_k\, \xi^k,$$

где $t_k > 0$ – величина шага; ξ^k – случайный вектор единичной длины, определяющий направление поиска; k – номер итерации. На текущей итерации при помощи генерирования случайных векторов ξ^k получаются точки, лежащие на гиперсфере радиуса t_k с центром в точке x^k (рис. 5.18). Если значение функции в полученной точке не меньше, чем в центре, шаг считается неудачным (точки y^1, y^2 при поиске из x^0; y^1, y^2, y^3 при поиске из x^1), происходит возврат в текущий центр и поиск продолжается. Если число неудачных шагов из текущей точки достигает некоторого числа M, дальнейший поиск продолжается из той же точки, но с меньшим шагом до тех пор, пока он не станет меньше заранее заданной величины R. Если же значение функции в полученной точке меньше, чем в центре, шаг считается удачным и дальнейший поиск продолжается из этой точки.

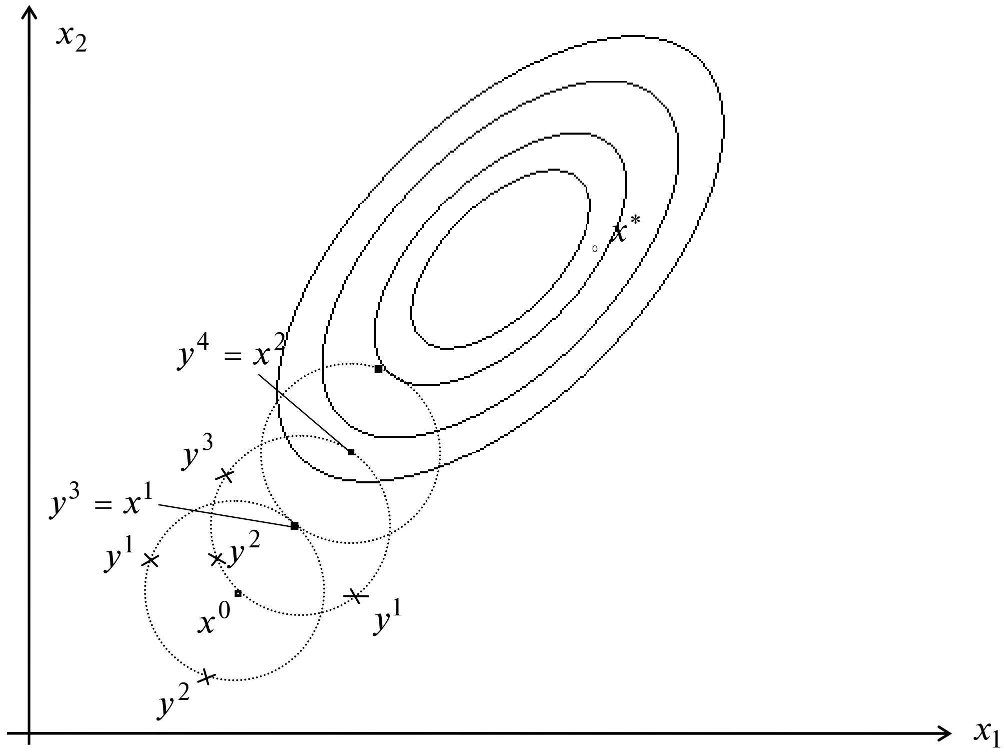

Рис. 5.18

Алгоритм

Шаг 1. Задать начальную точку x^0, коэффициент сжатия $0 < \beta < 1$, M – максимальное число неудачно выполненных испытаний на текущей итерации, t_0 – начальную величину шага, R - минимальную величину шага, N – максимальное число итераций. Положить $k = 0, j = 1$.

Шаг 2. Получить случайный вектор $\xi^j = \left(\xi_1{}^j, \ldots, \xi_n{}^j\right)^T$, где $\xi_i{}^j$ – случайная величина, равномерно распределенная на интервале $[-1,1]$.

Шаг 3. Вычислить $y^j = x^k + t_k \dfrac{\xi^j}{\left\|\xi^j\right\|}$.

Шаг 4. Проверить выполнение условий:

а) если $f\left(y^j\right) < f\left(x^k\right)$, шаг удачный. Положить $x^{k+1} = y^j$, $t_{k+1} = t_k$, $k = k+1$ и проверить условие окончания. Если $k < N$, положить $j = 1$ и перейти к шагу 2. Если $k = N$, поиск завершить: $x^* \cong x^k$;

б) если $f\left(y^j\right) \geq f\left(x^k\right)$, шаг неудачный и перейти к шагу 5.

Шаг 5. Оценить число неудачных шагов из текущей точки:

а) если $j < M$, следует положить $j = j+1$ и перейти к шагу 2;

б) если $j = M$, проверить условие окончания:

- если $t_k \leq R$, процесс закончить: $x^* \cong x^k, f\left(x^*\right) \cong f\left(x^k\right)$;
- если $t_k > R$, положить $t_k = \beta\, t_k$, $j = 1$ и перейти к шагу 2.

5.6.3. Метод наилучшей пробы

Постановка задачи

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

Стратегия поиска

Задается начальная точка x^0. Каждая последующая точка находится по формуле

$$x^{k+1} = x^k + t_k \, \xi^k ,$$

где $t_k > 0$ – величина шага; ξ^k – случайный вектор единичной длины, определяющий направление поиска; k – номер итерации. На текущей итерации при помощи генерирования случайных векторов ξ^k получается M точек $y^1, \ldots, y^M$, лежащих на гиперсфере радиуса t_k с центром в точке x^k (рис. 5.19). Среди полученных точек выбирается точка y^m, в которой значение функции наименьшее. Если в выбранной точке значение функции меньше, чем в центре, то дальнейший поиск продолжается из этой точки. Иначе поиск продолжается из старого центра, но с меньшим шагом до тех пор, пока он не станет меньше заранее заданной величины R.

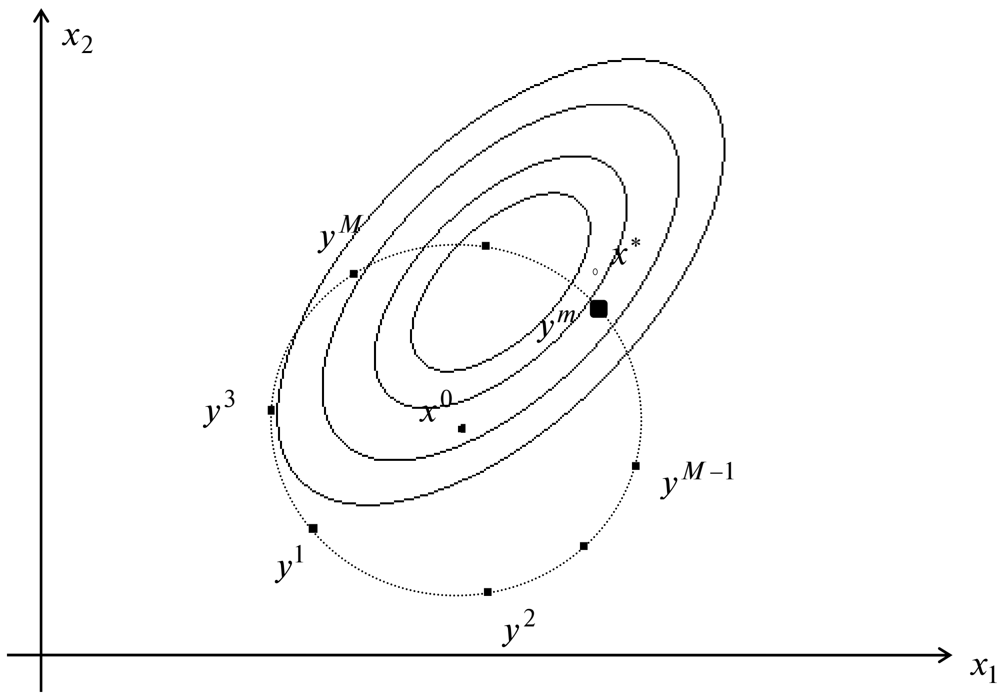

Рис. 5.19

Алгоритм

Шаг 1. Задать начальную точку x^0, коэффициент сжатия $0 < \beta < 1$, M – число испытаний на текущей итерации, $t_0 = 1$ – начальную величину шага, R – минимальную величину шага, N – максимальное число итераций. Положить $k = 0$, $j = 1$.

Шаг 2. Получить M реализаций случайного вектора $\xi^j = \left(\xi_1{}^j, \ldots, \xi_n{}^j\right)^T$, где $j = 1,\ldots, M$, $\xi_i{}^j$ – случайная величина, равномерно распределенная на интервале $\left[-1,1\right]$.

Шаг 3. Вычислить

$$y^j = x^k + t_k \frac{\xi^j}{\left\|\xi^j\right\|}, \quad j = 1,\ldots, M .$$

Шаг 4. Найти y^m из условия $f(y^m) = \min\limits_{1 \le j \le M} f(y^j)$.

Проверить выполнение условий:

а) если $f\left(y^m\right) < f\left(x^k\right)$, шаг удачный. Положить $x^{k+1} = y^m$, $t_{k+1} = t_k$, $k = k+1$ и проверить условие окончания. Если $k < N$, положить $j = 1$ и перейти к шагу 2. Если $k = N$, поиск завершить: $x^* \cong x^k$;

б) если $f\left(y^m\right) \ge f\left(x^k\right)$, шаг неудачный и перейти к шагу 5.

Шаг 5. Проверить условие окончания:

а) если $t_k \le R$, процесс закончить: $x^* \cong x^k$, $f\left(x^*\right) \cong f\left(x^k\right)$;

б) если $t_k > R$, положить $t_k = \beta t_k$, $j = 1$ и перейти к шагу 2.

Замечания 5.14.

1. Существуют варианты данного метода, в которых на шаге 4 полагают $x^{k+1} = y^m$. В этом случае становятся возможными шаги в направлении возрастания функции. Они могут позволить преодолевать локальные минимумы при поиске глобального экстремума.

2. Недостатком метода является учет только наилучшей пробной точки. В отбрасываемых точках содержится полезная информация о поведении целевой функции.

3. Одним из методов учета информации, содержащейся во всех сгенерированных точках, является *алгоритм статистического градиента*. Для каждой из M реализаций $\xi^1,\ldots,\xi^M$ случайного вектора ξ, полученных в точке x^k, вычисляются разности $\Delta f^{\,j} = f(x^k + t_{\text{пр}}\, \xi^k) - f(x^k)$, где $t_{\text{пр}}$ – пробное значение шага. В качестве направления поиска используется вектор статистического антиградиента $d^k = -\dfrac{1}{t_{\text{пр}}} \sum\limits_{j=1}^{M} \xi^j \Delta f^{\,j}$ или вектор $\dfrac{d^k}{\left\|d^k\right\|}$. Далее алгоритм решения совпадает с описанным выше.

Задачи для самостоятельного решения

1. Решить задачу

$$f(x) = x_1{}^3 + x_2{}^2 - 3x_1 - 2x_2 + 2 \to \min$$

методами конфигураций, деформированного многогранника, сопряженных направлений.

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

2. Решить задачу

$$f(x) = (x_1 - 2)^2 + (x_2 - 5)^2 + (x_3 + 2)^2 \to \min$$

методами конфигураций, деформированного многогранника, сопряженных направлений, Розенброка.

Ответ: точное решение $x^* = (2,\ 5,\ -2)^T$.

3. Решить задачу

$$f(x) = x_1^4 + x_2^4 + 2x_1^2 x_2^2 - 4x_1 + 3 \to \min$$

методами конфигураций, деформированного многогранника, сопряженных направлений.

Ответ: точное решение $x^* = (1,\ 0)^T$.

4. Решить задачу

$$f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 \to \min$$

методами конфигураций, деформированного многогранника, сопряженных направлений. В качестве начальной точки рекомендуется взять $(1;\ 1)^T$.

Ответ: точное решение $x^* = (3,\ 2)^T$.

5. Решить задачу

$$f(x) = 1 - 2x_1 - 2x_2 - 4x_1 x_2 + 10x_1^2 + 2x_2^2 \to \min$$

методами конфигураций, деформированного многогранника, сопряженных направлений, Розенброка.

Ответ: точное решение $x^* = (0{,}25;\ 0{,}75)^T$.

6. Методом Свенна найти начальный интервал неопределенности для решения задачи

$$f(x) = x^2 - 6x + 14 \to \min$$

при $x^0 = 0$; $t = 1$; $t = 0{,}1$; $t = 0{,}01$.

Ответ: $L_0 = [1,\ 7]$ при $t = 1$; $L_0 = [1{,}5;\ 6{,}3]$ при $t = 0{,}1$; $L_0 = [1{,}27;\ 5{,}11]$ при $t = 0{,}01$.

7. Методом Свенна найти начальный интервал неопределенности для решения задачи

$$f(x) = x^2 + 6x + 12 \to \min.$$

Ответ: $L_0 = [-8;\ 4]$ при $x^0 = -10,\ t = 2$; $L_0 = [-5,\ 1]$ при $x^0 = 1,\ t = 2$; $L_0 = [-6;\ 0]$ при $x^0 = 1, t = 1$; $L_0 = [-7;\ -1]$ при $x^0 = 0,\ t = 1$.

8. Методами равномерного поиска, деления интервала пополам, дихотомии, золотого сечения, Фибоначчи решить задачу

$$f(x) = x^2 - 6nx + 14 \to \min,\quad L_0 = [-m,\ 4n].$$