

Лабораторная работа «НЕЛИНЕЙНАЯ ОПТИМИЗАЦИЯ»

Тема 1. ПРИМЕНЕНИЕ ПОИСКОВЫХ МЕТОДОВ ОПТИМИЗАЦИИ

Цель. Изучить методы поиска минимума одномерных унимодальных функций

Контрольные вопросы

1. К какому классу методов относится алгоритм деления пополам?
2. Описать схему алгоритма деления пополам.
3. К какому классу методов относится алгоритм Фибоначчи?
4. Описать схему алгоритма Фибоначчи.
5. Каковы свойства золотого сечения?
6. Описать схему алгоритма золотого сечения.
7. Как сравнить эффективность алгоритмов одномерной условной оптимизации?

Краткие теоретические сведения и примеры выполнения заданий:

ПАССИВНЫЙ МЕТОД ПОИСКА МИНИМУМА

Метод оптимизации называется пассивным, когда все точки x_i , $i = 1, N$, вычислений характеристик задачи (в данном случае значений целевой функции) выбираются одновременно до начала вычислений.

Если N четное, т.е. $N = 2l$, $l = 1, 2, \dots$, то наилучшее (в смысле максимального уменьшения длины отрезка локализации) размещение точек x_i , $i = 1, N$, получается разбиением их на равноотстоящие ε -пары, т.е.

$$x_{2j-1} = a + \frac{b-a}{N/2+1} j - \frac{\varepsilon}{2}, \quad x_{2j} = a + \frac{b-a}{N/2+1} j + \frac{\varepsilon}{2}, \quad j = \overline{1, N/2},$$

где ε — некоторое малое положительное число. При этом

$$L_N = \frac{b-a}{N/2+1} + \frac{\varepsilon}{2} = \frac{L_0}{l+1} + \frac{\varepsilon}{2}$$

Если N нечетное, т.е. $N = 2l + 1$, $l = 1, 2, \dots$, то наилучшим является равномерное распределение точек, т.е.

$$x_i = a + \frac{b-a}{N+1} i, \quad i = \overline{1, N}, \quad L_N = 2 \frac{b-a}{N+1} = \frac{L_0}{l+1}.$$

Нетрудно заметить, что использование нечетного числа точек при пассивном методе поиска неэффективно.

После определения точек x_i , $i = 1, N$, вычисляются значения функции $f(x_i)$. Пусть $f(x_k) = \min f(x_i)$. Тогда, полагая $x_0 = a$, $x_{N+1} = b$, определяется итоговый отрезок локализации $[x_{k-1}, x_{k+1}]$. Точка x_k принимается за аппроксимацию (оценку) точки минимума x^* , значение функции $f(x_k)$ - за

оценку $f^* = f(x^*)$, т е-

$$x^* \equiv x_k, f^* \equiv f(x_k).$$

Пример. Определить с помощью пассивного поиска минимум функции $f(x) = x + 1/x$, заданной на отрезке $[0,2]$:

а) при $N=6$, $e=0,1$; б) при $N=7$.

Решение.

а) $N=6$, $e=0,1$.

Определяем пары точек x_{2j-1}, x_{2j}

$$x_{2j-1} = 0 + \frac{2-0}{3+1}j - \frac{0,1}{2} = 0,5j - 0,05, \quad j = \overline{1,3};$$

$$x_{2j} = 0 + \frac{2-0}{3+1}j + \frac{0,1}{2} = 0,5j + 0,05, \quad j = \overline{1,3}.$$

Результаты вычислений x и $f(x)$ заносим в табл.

Номер отсчета	1	2	3	4	5	6
x	0,45	0,55	0,95	1,05	1,45	1,55
$f(x)$	2,67	2,37	2,0026	2,0024	2,14	2,20

Поскольку $f(x_4) = \min_{i=1,6} f(x_i)$, то полагаем $\Delta_6 = [x_3, x_5] = [0,95; 1,45]$, $x^* \equiv x_4 = 1,05$, $f^* \equiv f(x_4) = 2,0024$.

Ответ: $\Delta_6 = [0,95; 1,45]$, $x^* \equiv 1,05$, $f^* \equiv 2,0024$.

б) $N=7$.

Определяем x_i - помощью соотношения

$$x_i = 0 + \frac{2-0}{7+1}i = 0,25i, \quad i = \overline{1,7}.$$

Результаты вычислений x и $f(x)$ заносим в табл.

Номер отсчета	1	2	3	4	5	6	7
x	0,25	0,5	0,75	1	1,25	1,5	1,75
$f(x)$	4,25	2,50	2,08	2,00	2,05	2,17	2,32

Поскольку $f(x_4) = \min_{i=1,7} f(x_i)$, то полагаем $\Delta_7 = [x_3, x_5] = [0,75; 1,25]$, $x^* \equiv x_4 = 1$, $f^* \equiv f(x_4) = 2$.

Ответ: $\Delta_7 = [0,75; 1,25]$, $x^* \equiv 1$, $f^* \equiv 2$.

АКТИВНЫЕ МЕТОДЫ ПОИСКА МИНИМУМА

Метод дихотомии (половинного деления)

В данном случае общее количество вычислений $f(x)$ четное, т.е. $N = 2l$, $l = 1, 2, \dots$, на j -м шаге (j -й итерации) производится пара вычислений $x_1^{(j)}$ и $x_2^{(j)}$, отстоящих на расстоянии $\varepsilon / 2$ по обе стороны от середины текущего отрезка локализации $[a^{(j-1)}, b^{(j-1)}]$. Если $f_1^{(j)} \leq f_2^{(j)}$, то отбрасывается часть отрезка, расположенная справа от $x_2^{(j)}$; если $f_1^{(j)} > f_2^{(j)}$, то отбрасывается часть отрезка, расположенная слева от $x_1^{(j)}$.

Используются два условия окончания вычислений:

- а) выполнение заданного количества вычислений N ;
- б) достижение заданной величины δ уменьшения отрезка локализации.

Итак, алгоритм поиска минимума унимодальной функции методом дихотомии заключается в следующем.

1. Задаются N (либо δ) и ε , полагается $j=1$.

2. На j -й итерации вычисляются

$$x_1^{(j)} = \frac{1}{2}(a^{(j-1)} + b^{(j-1)}) - \frac{\varepsilon}{2}, \quad x_2^{(j)} = \frac{1}{2}(a^{(j-1)} + b^{(j-1)}) + \frac{\varepsilon}{2},$$

$$f_1^{(j)} = f(x_1^{(j)}), \quad f_2^{(j)} = f(x_2^{(j)}).$$

Если $f_1^{(j)} \leq f_2^{(j)}$, то $a^{(j)} = a^{(j-1)}$, $b^{(j)} = x_2^{(j)}$.

Если $f_1^{(j)} > f_2^{(j)}$, то $a^{(j)} = x_1^{(j)}$, $b^{(j)} = b^{(j-1)}$.

3. Проверяется условие окончания вычислений:

$$\text{а) } j=N/2 \text{ либо б) } \frac{L_{2j}}{L_0} \leq \delta.$$

Если оно выполняется, то определяются итоговый отрезок локализации, оценки точки минимума x^* и величины минимума $f^* = f(x^*)$, и вычисления завершаются.

Если условие не выполняется, то полагается $j = j + 1$ и осуществляется переход к п.2.

Отметим, что для определения оценки точки минимума надо рассмотреть все исследованные точки итогового отрезка локализации и выбрать ту из них, для которой значение функции минимально.

Пример. Определить методом дихотомии минимум функции $f(x) = x^4 - 6x^2 + 10$, заданной на отрезке $[1, 3]$, при $N=8$, $\varepsilon=0,1$.

Решение.

В данном случае будут выполнены $N/2=4$ итерации.

Результаты вычислений заносим в табл. 4.3.

Таблица 4.3

Номер итерации	$x_1^{(j)}$	$x_2^{(j)}$	$f_1^{(j)}$	\leq $>$	$f_2^{(j)}$	$\alpha^{(j)}$	$b^{(j)}$
0	—	—	—		—	1	3
1	1,95	2,05	1,644	<	2,446	1	2,05
2	1,475	1,575	1,680	>	1,270	1,475	2,05
3	1,713	1,813	1,004	<	1,082	1,475	1,813
4	1,594	1,694	1,211	>	1,017	1,594	1,813

Поскольку $j=N/2=4$, то вычисления завершаются.

Точка минимума локализована на отрезке $[1,594; 1,813]$. На данном отрезке исследованы 4 точки:

$$\left. \begin{aligned} \alpha^{(4)} = 1,594 &\rightarrow f(\alpha^{(4)}) = 1,211; \\ b^{(4)} = 1,813 &\rightarrow f(b^{(4)}) = 1,082; \\ x_2^{(4)} = 1,694 &\rightarrow f(x_2^{(4)}) = 1,017; \\ x_1^{(3)} = 1,713 &\rightarrow f(x_1^{(3)}) = 1,004; \end{aligned} \right\} x^* \cong x_1^{(3)} = 1,713, f^* \cong f(x_1^{(3)}) = 1,004.$$

Ответ:

$$\Delta_x = [1,594; 1,813], x^* \cong 1,713, f^* \cong 1,004.$$

Метод Фибоначчи

Метод Фибоначчи является наилучшим (в смысле максимального уменьшения длины отрезка локализации) среди активных методов поиска. Согласно методу Фибоначчи, на первом шаге (первой итерации) проводятся два вычисления значений $f(x)$ в точках $x_1^{(1)}$ и $x_2^{(1)}$ (причем $x_1^{(1)} < x_2^{(1)}$, расположенных симметрично относительно середины отрезка $[a, b]$). По результатам вычислений одна из частей отрезка ($[a, x_1^{(1)}]$ либо $[x_2^{(1)}, b]$) отбрасывается, при этом одна из точек (соответственно $x_2^{(1)}$ либо $x_1^{(1)}$) уже проведенных вычислений остается внутри отрезка $\Delta_2 = \Delta^{(1)}$. На каждом последующем шаге (последующей итерации) точка очередного вычисления выбирается симметрично оставшейся точки. Таким образом, на первой итерации проводятся два вычисления значений $f(x)$, на каждой последующей - одно вычисление. Поэтому при заданном количестве вычислений N будет выполнено $N - 1$ шагов (итераций).

При вычислении $x_1^{(j)}$ и $x_2^{(j)}$, $j = \overline{1, N-1}$, используются числа Фибоначчи, определяемые следующим образом:

$$F_0 = F_1 = 1, F_k = F_{k-1} + F_{k-2}, k = 2, 3, \dots$$

Условием окончания вычислений является выполнение заданного

количества вычислений N .

Итак, алгоритм поиска минимума унимодальной функции методом Фибоначчи заключается в следующем.

1. Задается N , определяются числа Фибоначчи F_k , $k = \overline{0, N+1}$, выбирается ε из условия

$$\varepsilon < \frac{b-a}{F_{N+1}}.$$

Полагается $j=1$.

2. На j -й итерации вычисляются

$$x_1^{(j)} = a^{(j-1)} + \frac{F_{N-j+1}}{F_{N-j+1}}(b^{(j-1)} - a^{(j-1)}) - \frac{(-1)^{N-j+1}}{F_{N-j+1}}\varepsilon,$$

$$x_2^{(j)} = a^{(j-1)} + \frac{F_{N-j}}{F_{N-j+1}}(b^{(j-1)} - a^{(j-1)}) + \frac{(-1)^{N-j+1}}{F_{N-j+1}}\varepsilon,$$

$$f_1^{(j)} = f(x_1^{(j)}), f_2^{(j)} = f(x_2^{(j)}).$$

Если $f_1^{(j)} \leq f_2^{(j)}$, то $a^{(j)} = a^{(j-1)}$, $b^{(j)} = x_2^{(j)}$, $x_2^{(j+1)} = x_1^{(j)}$.

Если $f_1^{(j)} > f_2^{(j)}$, то $a^{(j)} = x_1^{(j)}$, $b^{(j)} = b^{(j-1)}$, $x_1^{(j+1)} = x_2^{(j)}$.

3. Проверяется условие окончания вычислений

$$j = N - 1.$$

Если оно выполняется, то определяются итоговый отрезок локализации, оценки точки минимума x^* и величины минимума $f^* = f(x^*)$ и вычисления завершаются.

Если условие не выполняется, то полагается $j=j+1$ и осуществляется переход к п.2.

Примечание. На j -й, $j>1$, итерации вычисляется только та точка $x_i^{(j)}$, $i = 1, 2$, которая не была определена на предыдущей итерации.

Отметим, что оценкой точки минимума x^* является та из точек $x_i^{(N-1)}$, $i = 1, 2$, которая осталась внутри итогового отрезка локализации.

Пример. Определить методом Фибоначчи минимум функции $f(x) = x^4 - 6x^2 + 10$, заданной на отрезке $[1, 3]$, при $N=4$.

Решение. В данном случае будут выполнены $N - 1 = 3$ итерации.

Определяем числа Фибоначчи F_k , $k = \overline{1, 5}$:

$$F_0 = F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, F_5 = 8.$$

$$\varepsilon < \frac{b-a}{F_5} = \frac{3-1}{8} = 0,25.$$

Выбираем $\varepsilon=0,1$.

Первая итерация

$$x_1^{(1)} = a^{(0)} + \frac{F_{N-2}}{F_N} (b^{(0)} - a^{(0)}) - \frac{(-1)^N}{F_N} \varepsilon = 1 + \frac{F_2}{F_4} (3 - 1) - \frac{(-1)^4}{F_4} \cdot 0,1 = 1 + \frac{2 \cdot 2 - 0,1}{5} = 1 + 0,78 = 1,78,$$

$$x_2^{(1)} = a^{(0)} + \frac{F_{N-1}}{F_N} (b^{(0)} - a^{(0)}) + \frac{(-1)^N}{F_N} \varepsilon = 1 + \frac{F_3}{F_4} (3 - 1) + \frac{(-1)^4}{F_4} \cdot 0,1 = 1 + \frac{3 \cdot 2 + 0,1}{5} = 1 + 1,22 = 2,22.$$

Вторая итерация

$$x_1^{(2)} = a^{(1)} + \frac{F_{N-3}}{F_{N-1}} (b^{(1)} - a^{(1)}) - \frac{(-1)^{N-1}}{F_{N-1}} \varepsilon = 1 + \frac{F_1}{F_3} (2,22 - 1) - \frac{(-1)^3}{F_3} \cdot 0,1 = 1 + \frac{1 \cdot 1,22 + 0,1}{3} = 1 + 0,44 = 1,44.$$

Третья итерация

$$x_2^{(2)} = a^{(2)} + \frac{F_{N-2}}{F_{N-2}} (b^{(2)} - a^{(2)}) + \frac{(-1)^{N-2}}{F_{N-2}} \varepsilon = 1,44 + \frac{F_1}{F_2} (2,22 - 1,44) + \frac{(-1)^2}{F_2} \cdot 0,1 = 1,44 + \frac{1 \cdot 0,78 + 0,1}{2} = 1,44 + 0,44 = 1,88.$$

Результаты вычислений заносим в табл. 4.4.

Номер итерации	$x_1^{(j)}$	$x_2^{(j)}$	$f_1^{(j)}$	\leq $>$	$f_2^{(j)}$	$a^{(j)}$	$b^{(j)}$
0	—	—	—		—	1	3
1	1,78*	2,22*	1,028	<	4,719	1	2,22
2	1,44*	1,78	1,858	>	1,028	1,44	2,22
3	1,78	1,88*	1,028	<	1,286	1,44	1,88

Примечание. Знаком * помечаем точки $x_i^{(j)}$, $i=1,2$, вычисляемые на j -й итерации.

Поскольку $j=N-1=3$, то вычисления завершаются.

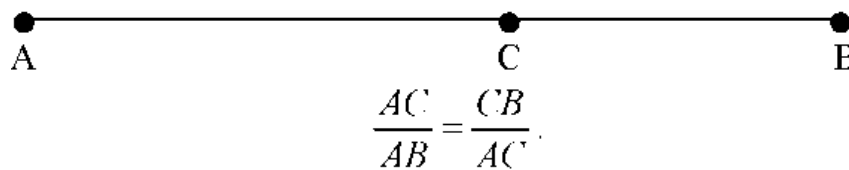
Точка минимума локализована на отрезке $[1,44; 1,88]$, $x^* = x^{(3)} = 1,78$, $f^* = f(x^{(3)}) = 1,028$.

Ответ:

$$\Delta_4 = [1,44; 1,88], \quad x^* \equiv 1,78, \quad f^* \equiv 1,028$$

Метод золотого сечения

Недостатком наиболее эффективного метода Фибоначчи является то, что должно быть задано количество вычислений N . Метод золотого сечения почти столь же эффективен, как и метод Фибоначчи, но при этом не зависит от N . Алгоритм поиска по методу золотого сечения определяется тем же правилом симметрии, что и алгоритм по методу Фибоначчи: на первой итерации выбираются две точки, расположенные симметрично относительно середины исходного отрезка; на каждой последующей итерации выбирается одна точка, расположенная симметрично оставшейся точки. Разница заключается в выборе точек. Метод золотого сечения основан на делении отрезка локализации «золотым сечением», т.е. таком делении, когда отношение большей части отрезка ко всему отрезку равно отношению меньшей части к большей



При таком делении используются две дроби Фибоначчи

$$\phi_1 = \frac{3 - \sqrt{5}}{2} \cong 0,382, \quad \phi_2 = \frac{\sqrt{5} - 1}{2} \cong 0,618,$$

удовлетворяющие условиям

$$\phi_1 + \phi_2 = 1, \quad \phi_1 = (\phi_2)^2.$$

В случае метода золотого сечения используются два условия окончания вычислений:

- а) выполнение заданного количества вычислений N ,
- б) достижение заданной величины δ уменьшения отрезка локализации.

Итак, алгоритм поиска минимума унимодальной функции методом золотого сечения заключается в следующем.

1. Задается N (либо δ), полагается $j=1$.
2. На j -й итерации вычисляются

$$x_1^{(j)} = a^{(j-1)} + \phi_1 (b^{(j-1)} - a^{(j-1)}),$$

$$x_2^{(j)} = a^{(j-1)} + \phi_2 (b^{(j-1)} - a^{(j-1)}),$$

$$f_1^{(j)} = f(x_1^{(j)}), \quad f_2^{(j)} = f(x_2^{(j)}).$$

Если $f_1^{(j)} \leq f_2^{(j)}$, то $a^{(j)} = a^{(j-1)}$, $b^{(j)} = x_2^{(j)}$, $x_2^{(j+1)} = x_1^{(j)}$.

Если $f_1^{(j)} > f_2^{(j)}$, то $a^{(j)} = x_1^{(j)}$, $b^{(j)} = b^{(j-1)}$, $x_1^{(j+1)} = x_2^{(j)}$.

3. Проверяется условие окончания вычислений:

$$\text{а) } j = N - 1 \quad \text{либо} \quad \text{б) } \frac{L_{j+1}}{L_0} \leq \delta.$$

Если оно выполняется, то определяются итоговый отрезок локализации, оценки точки минимума x^* и величины минимума f^* и вычисления завершаются.

Если условие не выполняется, то полагается $j=j+1$ и осуществляется переход к п.2.

Пример. Определить методом золотого сечения минимум функции $f(x) = x^4 - 6x^2 + 10$, заданной на отрезке $[1,3]$, при $N=4$.

Решение.

В данном случае будут выполнены $N - 1 = 3$ итерации. Результаты вычислений заносим в табл. 4.5.

Таблица 4.5

Номер итерации	$x_1^{(j)}$	$x_2^{(j)}$	$f_1^{(j)}$	\leq $>$	$f_2^{(j)}$	$a^{(j)}$	$b^{(j)}$
0	—	—	—		—	1	3
1	1,764*	2,236*	1,012	<	4,999	1	2,236
2	1,472*	1,764	1,694	>	1,012	1,472	2,236
3	1,764	1,944*	1,012	<	1,607	1,472	1,944

Поскольку $j = N - 1 = 3$, то вычисления завершаются.

Точка минимума локализована на отрезке

$$\Delta_4 = [1,472; 1,944], \quad x^* \equiv x_1^{(3)} = 1,764, \quad f^* \equiv f(x_1^{(3)}) = 1,012.$$

Индивидуальные задания 1

1. Определить с помощью пассивного поиска минимум функции $f(x)$, заданной на отрезке $[0, 8]$: а) при $N=16$, $\varepsilon = 0,1$; б) при $N=17$.

2. Определить методом дихотомии минимум функции $f(x)$, заданной на отрезке $[0, 8]$, при $N=16$, $\varepsilon = 0,1$.

3. Определить методом Фибоначчи минимум функции $f(x)$, заданной на отрезке $[0, 8]$, при $N=16$, $\varepsilon = 0,2$.

4. Определить методом золотого сечения минимум функции $f(x)$, заданной на отрезке $[0, 8]$, при $N=16$.

В	$f(x)$	В	$f(x)$	В	$f(x)$	В	$f(x)$
1	$f(x) = x^2 - 3x + 2$	9	$f(x) = x^2 - 3x + 7$	17	$f(x) = x^2 - 3x + 2$	25	$f(x) = x^2 - 3x + 7$
2	$f(x) = x^2 - 5x + 4$	10	$f(x) = x^2 - 5x + 9$	18	$f(x) = x^2 - 5x + 4$	26	$f(x) = x^2 - 5x + 9$
3	$f(x) = x^2 - 7x + 6$	11	$f(x) = x^2 - 7x + 11$	19	$f(x) = x^2 - 7x + 6$	27	$f(x) = x^2 - 7x + 11$
4	$f(x) = x^2 - 9x + 8$	12	$f(x) = x^2 - 9x + 3$	20	$f(x) = x^2 - 9x + 8$	28	$f(x) = x^2 - 9x + 3$
5	$f(x) = x^2 - 11x + 10$	13	$f(x) = x^2 - 11x + 5$	21	$f(x) = x^2 - 11x + 10$	29	$f(x) = x^2 - 11x + 5$
6	$f(x) = x^2 - 13x + 12$	14	$f(x) = x^2 - 13x + 7$	22	$f(x) = x^2 - 13x + 12$	30	$f(x) = x^2 - 13x + 7$
7	$f(x) = x^2 - 15x + 14$	15	$f(x) = x^2 - 3x + 9$	23	$f(x) = x^2 - 15x + 14$	31	$f(x) = x^2 - 3x + 9$
8	$f(x) = x^2 - 7x + 1$	16	$f(x) = x^2 - 5x - 1$	24	$f(x) = x^2 - 7x + 1$	32	$f(x) = x^2 - 5x - 1$

Тема 2.

“НЕЛИНЕЙНАЯ МНОГОМЕРНАЯ ОПТИМИЗАЦИЯ”

Тема

Методы поиска минимума многомерных функций

Цель. Используя математический аппарат теории нелинейного программирования рассчитать оптимальный режим поставок товара для минимизации издержек

Контрольные вопросы

1. Как аналитически найти минимум многомерной функции с ограничениями на область переменных?
2. Сформулировать теорему Куна-Таккера.
3. Какой алгоритм нахождения минимума аналитически без использования теоремы Куна-Таккера?
4. Какой алгоритм нахождения минимума аналитически с использованием теоремы Куна-Таккера?
5. Как определить оптимальный размер партии поставки?
6. Как рассчитать характеристики работы склада в оптимальном режиме?
7. Опишите метод Гаусса-Зейделя.
8. Опишите метод Хука-Дживса.
9. Опишите метод Розенброка.
10. Опишите метод сопряженных направлений.
11. Опишите симплекс-метод.
12. Опишите метод деформируемого многогранника (Нелдера-Мида)
13. Опишите метод наискорейшего спуска.
14. Опишите метод дробления шага.
15. Опишите метод оптимизации Ньютона.
16. Опишите методы случайного поиска

Ход работы.

Постановка задачи.

Склад оптовой торговли отпускает 5 видов товаров. Известны потребности V_i , издержки заказывания K_i , издержки содержания s_i , расход складской площади на единицу товара f_i , а также величина складской площади торгового зала F . Хотя бы одна единица товара каждого вида должна храниться на складе.

Требуется определить оптимальные партии поставок при ограничении на максимальный уровень запаса при условии, что все пять видов продукции поступают на склад от разных поставщиков (раздельная оптимизация)

- 1) **Решить указанным в задании методом.** Выводить промежуточные результаты вычислений (координаты точки и значения функции в точке, полученные на каждой итерации). Выписать полученный ответ.
- 2) **Найти решение на компьютере (например, в Excel).**

Индивидуальные задания 2

Вариант	F	i	1	2	3	4	5	Метод решения
1	1200	V_i	900	700	300	1000	200	метод Гаусса-Зейделя
		K_i	10	5	20	30	6	
		S_i	5	15	10	2	3	
		f_i	16	4	15	22	10	
2	500	V_i	400	600	800	700	200	метод Хука-Дживса
		K_i	10	12	11	9	8	
		S_i	16	8	8	7	4	
		f_i	4	3	5	4	4	
3	500	V_i	700	200	500	150	800	метод Розенброка
		K_i	5	5	20	3	4	
		S_i	15	4	10	2	20	
		f_i	20	5	2	8	4	
4	1500	V_i	3000	5000	6400	1500	80	метод сопряженных направлений
		K_i	4	6	7	6	4	
		S_i	40	6	14	6	16	
		f_i	4	3	5	40	20	
5	900	V_i	900	400	800	200	150	метод наискорейшего спуска
		K_i	5	10	11	7	2	
		S_i	4	7	6	4	2	
		f_i	8	5	6	3	3	
6	800	V_i	4000	2000	8000	600	1500	метод дробления шага
		K_i	10	7	15	110	6	
		S_i	8	70	6	8	20	
		f_i	3	2	2	5	30	
7	1350	V_i	5000	7000	2000	200	800	метод с возвратом при неудачном шаге
		K_i	6	110	7	5	4	
		S_i	15	8	20	4	8	
		f_i	10	5	2	3	4	
8	1000	V_i	48000	22400	6400	8600	2460	метод наилучшей пробы
		K_i	120	160	130	140	110	
		S_i	200	280	260	200	250	
		f_i	1.8	1.6	1.2	1.5	1.4	
9	1250	V_i	3200	2100	5400	7900	2420	метод Гаусса-Зейделя
		K_i	110	150	120	130	100	
		S_i	150	260	240	200	230	
		f_i	14	5	3	4	6	
10	6000	V_i	1350	1210	1150	1300	890	метод Хука-Дживса
		K_i	70	65	80	77	93	
		S_i	11	9	3	7	6	
		f_i	8	9	4	6	7	

11	1000	V_i K_i S_i f_i	500 20 5 10	100 10 10 20	200 5 4 5	150 3 2 2	400 7 20 8	метод Розенброка
12	500	V_i K_i S_i f_i	400 10 16 4	600 12 8 3	800 11 8 5	700 9 7 4	200 8 4 4	метод сопряженных направлений
13	500	V_i K_i S_i f_i	700 5 15 20	200 5 4 5	500 20 10 2	150 3 2 8	800 4 20 4	метод наискорейшего спуска
14	1500	V_i K_i S_i f_i	3000 4 40 4	5000 6 6 3	6400 7 14 5	1500 6 6 40	80 4 16 20	метод дробления шага
15	900	V_i K_i S_i f_i	900 5 4 8	400 10 7 5	800 11 6 6	200 7 4 3	150 2 2 3	метод с возвратом при неудачном шаге
16	1200	V_i K_i S_i f_i	900 10 5 16	700 5 15 4	300 20 10 15	1000 30 2 22	200 6 3 10	метод наилучшей пробы
17	500	V_i K_i S_i f_i	400 10 16 4	600 12 8 3	800 11 8 5	700 9 7 4	200 8 4 4	метод Гаусса-Зейделя
18	500	V_i K_i S_i f_i	700 5 15 20	200 5 4 5	500 20 10 2	150 3 2 8	800 4 20 4	метод Хука-Дживса
19	1500	V_i K_i S_i f_i	3000 4 40 4	5000 6 6 3	6400 7 14 5	1500 6 6 40	80 4 16 20	метод Розенброка
20	900	V_i K_i S_i f_i	900 5 4 8	400 10 7 5	800 11 6 6	200 7 4 3	150 2 2 3	метод сопряженных направлений
21	800	V_i K_i	4000 10	2000 7	8000 15	600 110	1500 6	метод наискорейшего

		Si fi	8 3	70 2	6 2	8 5	20 30	спуска
22	1350	Vi Ki Si fi	5000 6 15 10	7000 110 8 5	2000 7 20 2	200 5 4 3	800 4 8 4	метод дробления шага
23	1000	Vi Ki Si fi	48000 120 200 1.8	22400 160 280 1.6	6400 130 260 1.2	8600 140 200 1.5	2460 110 250 1.4	метод с возвратом при неудачном шаге
24	1250	Vi Ki Si fi	3200 110 150 14	2100 150 260 5	5400 120 240 3	7900 130 200 4	2420 100 230 6	метод наилучшей пробы
25	6000	Vi Ki Si fi	1350 70 11 8	1210 65 9 9	1150 80 3 4	1300 77 7 6	890 93 6 7	метод Гаусса- Зейделя
26	1000	Vi Ki Si fi	500 20 5 10	100 10 10 20	200 5 4 5	150 3 2 2	400 7 20 8	метод Хука-Дживса
27	500	Vi Ki Si fi	400 10 16 4	600 12 8 3	800 11 8 5	700 9 7 4	200 8 4 4	метод Розенброка
28	500	Vi Ki Si fi	700 5 15 20	200 5 4 5	500 20 10 2	150 3 2 8	800 4 20 4	метод сопряженных направлений
28	1500	Vi Ki Si fi	3000 4 40 4	5000 6 6 3	6400 7 14 5	1500 6 6 40	80 4 16 20	метод наискорейшего спуска
30	900	Vi Ki Si fi	900 5 4 8	400 10 7 5	800 11 6 6	200 7 4 3	150 2 2 3	метод дробления шага

Сначала решить задачу указанным методом, используя знания по нелинейной оптимизации.

Порядок выполнения работы на компьютере (в Excel)

1. Оптимизация без ограничений на складские площади.)

Строим таблицу 1.

Таблица 1.

I	Vi	Ki	Si	f	qi0	Ki*Vi/ qi0	Si*qi	fi*qi
1	8000	40	16	20	200,00	1600,00	3200,00	4000,00
2	160	5	4	3	20,00	40,00	80,00	60,00
3	1800	6	6	4	60,00	180,00	360,00	240,00
4	150	6	2	3	30,00	30,00	60,00	90,00
5	200	30	30	15	20,00	300,00	600,00	300,00
						2150,00	4300,00	4690,00

F 1340

L 4300

Найдем оптимальные размеры поставок при отсутствии ограничений по формуле Уилсона.

$$q_i^0 = \sqrt{\frac{2K_i v_i}{s_i}}$$

Заносим вычисления в таблицу.

Рассчитаем суммарные расходы при данном плане поставок.

$$L = \sum_{i=1}^n \left(\frac{K_i v_i}{q_i^0} + \frac{1}{2} s_i q_i^0 \right)$$

Для этого введем дополнительные столбцы $\frac{K_i v_i}{q_i^0}$, $s_i q_i^0$. Далее в отдельной ячейке записываем формулу для расчета.

2. Оптимизация с ограничениями на складские площади.

Так как ограничение накладывается на максимальный уровень запаса, то $h=1$. Проверим существенность ограничения на складские площади ($f=1340 \text{ м}^2$). Для этого сравним необходимое количество складских площадей с имеющимся.

$$h \sum_{i=1}^5 f_i q_i^0 = 4690 (\text{м}^2)$$

Так как полученное значение больше исходного, то ограничение является существенным. Для нахождения скорректированных значений составим оптимизационную модель. Цель – минимизировать суммарные расходы.

$$L = \sum_{i=1}^n \left(\frac{K_i v_i}{q_i} + \frac{1}{2} s_i q_i \right) \rightarrow \min$$

Ограничение вводится на величину складских площадей.

$$h \sum_{i=1}^n f_i q_i \leq f$$

Получили задачу нелинейной оптимизации, которую можно решить средствами EXCEL.

Для расчетов строим таблицу 2. (Копируем таблицу 1 ниже и ставим начальные значения в столбце q, например, равные 1 для того, чтобы начальные значения удовлетворяли области ограничений. **На ноль делить нельзя!**).

Столбцом значений будет столбец q*. Значение целевой функции находится в ячейке L. Правая часть ограничения записывается в отдельную ячейку. В программе «поиск решения» задаем параметры – «нелинейная модель», «неотрицательные значения».

Таблица 2.

I	Vi	Ki	Si	f	qi0	Ki*Vi/ qi0	Si*qi	fi*qi
1	8000	40	16	20	54,39	5883,09	870,29	1087,86
2	160	5	4	3	6,86	116,57	27,45	20,59
3	1800	6	6	4	21,66	498,589	129,97	86,64
4	150	6	2	3	7,50	119,98	15,00	22,50
5	200	30	30	15	8,16	735,30	244,80	122,40
						7353,53	1287,51	1340,00

F 1340

L 7997,281

3. Сведем полученные результаты в таблицу:

результат системы	необходимые складские площади	издержки ра- боты в д.е./год
управление поставками без ограничений	4690	4300
управление поставками с ограничениями на складские площади	1340	7997,28

4. Делаем анализ полученных результатов (указать объём поставки товара каждого вида).

Краткие сведения по предлагаемым методам решения.

Метод Гаусса-Зейделя

Рассматривается следующая многомерная задача безусловной оптимизации (точнее говоря, задача многомерной локальной безусловной оптимизации): найти минимум критерия оптимальности $\Phi(\mathbf{X})$, определенного в n -мерном евклидовом пространстве R^n ,

$$\min_{\mathbf{X} \in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

При решении задачи (1) методом Гаусса-Зейделя (методом покоординатного спуска, методом циклического покоординатного спуска) используются следующие итерационные формулы

$$\begin{aligned} \mathbf{X}_1^{r+1} &= \mathbf{X}^r + \lambda_1^r \mathbf{L}_1, \\ \mathbf{X}_2^{r+1} &= \mathbf{X}_1^{r+1} + \lambda_2^r \mathbf{L}_2, \\ &\dots, \\ \mathbf{X}_n^{r+1} &= \mathbf{X}_{n-1}^{r+1} + \lambda_n^r \mathbf{L}_n = \mathbf{X}^{r+1}, \end{aligned} \quad (2)$$

где вектор \mathbf{L}_i определяет направление вдоль i -й координатной оси и представляет собой n -мерный вектор с компонентами

$$l_{ij} = \begin{cases} 1, & \text{если } j=i, \\ 0, & \text{если } j \neq i, j=1, 2, \dots, n, \end{cases}$$

а величины $\lambda_1^r, \lambda_2^r, \dots, \lambda_n^r$ – определяются из условий

$$\begin{aligned} \min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}^r + \lambda \mathbf{L}_1) &= \Phi(\mathbf{X}^r + \lambda_1^r \mathbf{L}_1) = \Phi(\mathbf{X}_1^{r+1}), \\ \min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_1^{r+1} + \lambda \mathbf{L}_2) &= \Phi(\mathbf{X}_1^{r+1} + \lambda_2^r \mathbf{L}_2) = \Phi(\mathbf{X}_2^{r+1}), \\ &\dots \\ \min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_{n-1}^{r+1} + \lambda \mathbf{L}_n) &= \Phi(\mathbf{X}_{n-1}^{r+1} + \lambda_n^r \mathbf{L}_n) = \Phi(\mathbf{X}_n^{r+1}). \end{aligned} \quad (3)$$

Другими словами, величина $\lambda_i^r, i=1, 2, \dots, n$ представляет собой длину шага, минимизирующего функцию $\Phi(\mathbf{X})$ в направлении \mathbf{L}_i на итерации номер r , исходя из точки, полученной на предыдущем шаге.

Если положить $\mathbf{X}_0^{r+1} = \mathbf{X}^r, \mathbf{X}_n^{r+1} = \mathbf{X}^{r+1}$, то формулы (2), (3) можно записать в виде

$$\mathbf{X}_i^{r+1} = \mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{L}_i, i=1, 2, \dots, n, \quad (4)$$

$$\min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda \mathbf{L}_i) = \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{L}_i) = \Phi(\mathbf{X}_i^{r+1}), i=1, 2, \dots, n. \quad (5)$$

Таким образом, каждая итерация по методу Гаусса-Зейделя включает в себя n шагов. Каждая последующая итерация начинается из точки, полученной на последнем шаге предыдущей итерации. Поиск заканчивается при выполнении одного из стандартных условий окончания итераций:

$$\|X^{r+1} - X^r\| \leq \varepsilon_X, \quad (6)$$

$$|\Phi(X^{r+1}) - \Phi(X^r)| \leq \varepsilon_\Phi \quad (6')$$

Заметим, что задачи (5) даже в случае одноэкстремальной функции $\Phi(X)$ могут быть задачами многоэкстремальной оптимизации и могут быть решены рассмотренными в главе 4 методами решения задач одномерной оптимизации.

Схема метода Гаусса-Зейделя:

Задаем начальную точку X^0 и полагаем $r=0$, $i=1$.

Последовательно для $i=1, 2, \dots, n$ решаем задачи (5), т.е. исходя из предыдущей точки, отыскиваем минимум функции $\Phi(X)$ вдоль i -го координатного направления;

Если условие окончания поиска (6) или (6') выполнено, то полагаем $X^* \approx X^{r+1}$ и заканчиваем вычисления. Иначе - полагаем $r=r+1$ и переходим к п. 2 •

Метод Гаусса-Зейделя иллюстрирует рис. 1, на котором показан фрагмент линий уровня функции $\Phi(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$.

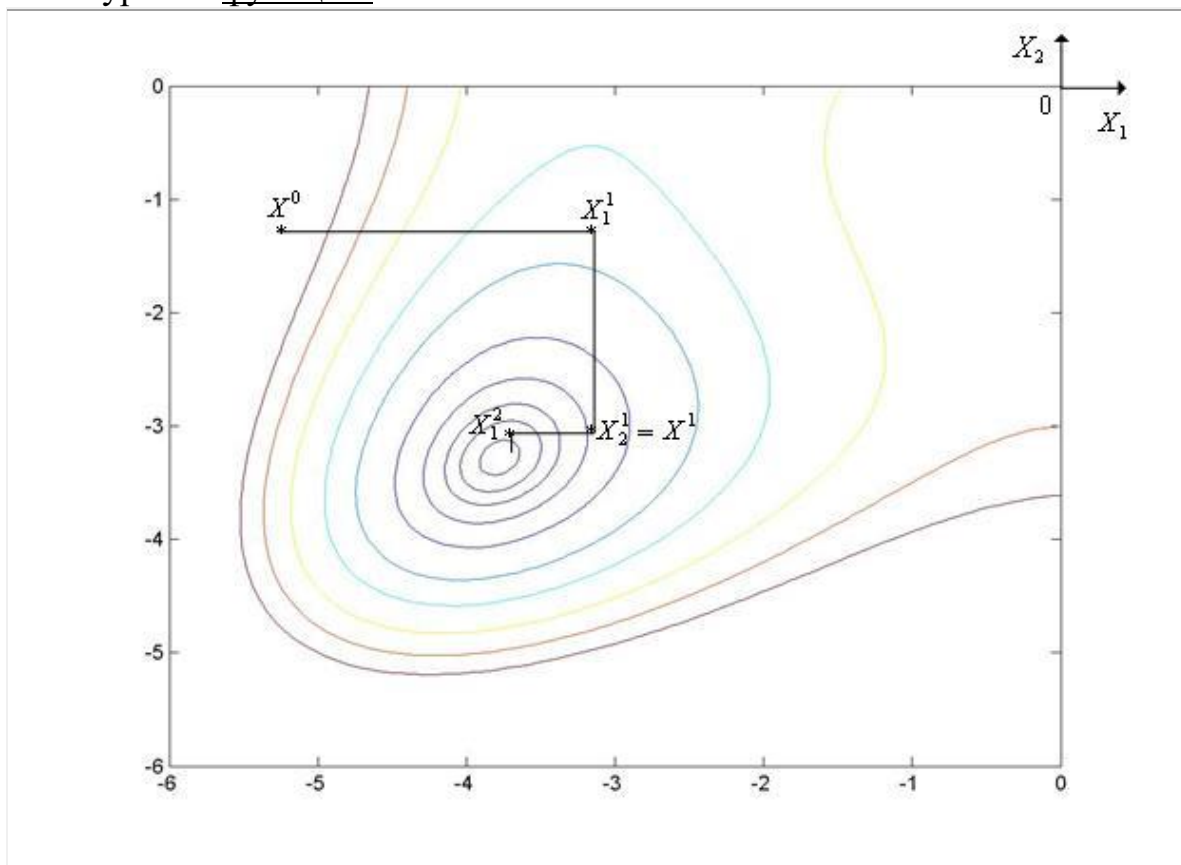


Рис. 1. Траектория поиска минимума методом Гаусса-Зейделя.

На рис. 1 точка X_1^1 представляет собой локальный минимум функции $\Phi(X)$ вдоль оси X_1 при исходной точке X^0 . Точка X_2^1 представляет собой

локальный минимум функции $\Phi(\mathbf{X})$ вдоль оси X_2 при исходной точке \mathbf{X}_1^1 . Отыскание точки \mathbf{X}_2^1 завершает первую итерацию. Следующая итерация начинается из точки $\mathbf{X}^1 = \mathbf{X}_2^1$. И т.д.

Метод Хука-Дживса

Рассмотрим следующую многомерную задачу локальной безусловной оптимизации: найти минимум критерия оптимальности $\Phi(\mathbf{X})$, определенного в n -мерном евклидовом пространстве R^n ,

$$\min_{\mathbf{X} \in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

При решении задачи (1) методом Хука-Дживса (методом конфигураций, методом пробных шагов) используются итерационные формулы, аналогичные формулам, используемым в методе Гаусса-Зейделя

$$\begin{aligned} \tilde{\mathbf{X}}_i^{r+1} &= \tilde{\mathbf{X}}_{i-1}^{r+1} + \lambda_i^r \mathbf{L}_i, \quad i=1, 2, \dots, n, \\ \tilde{\mathbf{X}}_0^{r+1} &= \mathbf{X}^r, \quad \tilde{\mathbf{X}}_n^{r+1} = \tilde{\mathbf{X}}^{r+1} \end{aligned} \quad (2)$$

где принято λ_i^r , $\tilde{\mathbf{X}}_n^{r+1} = \tilde{\mathbf{X}}^{r+1}$, вектор \mathbf{L}_i определяет направление вдоль i -й координатной оси и представляет собой n -мерный вектор с компонентами

$$l_{ij} = \begin{cases} 1, & \text{если } j=i, \\ 0, & \text{если } j \neq i, \quad j=1, 2, \dots, n, \end{cases}$$

величины $\lambda_1^r, \lambda_2^r, \dots, \lambda_n^r$ — определяются из условий

$$\lambda_i^r = \begin{cases} \Delta_i^r, & \text{если } \Phi(\mathbf{X}_{i-1}^{r+1} + \Delta_i^r \mathbf{L}_i) < \Phi(\mathbf{X}_{i-1}^{r+1}), \\ -\Delta_i^r, & \text{если } \Phi(\mathbf{X}_{i-1}^{r+1} - \Delta_i^r \mathbf{L}_i) < \Phi(\mathbf{X}_{i-1}^{r+1}) < \Phi(\mathbf{X}_{i-1}^{r+1} + \Delta_i^r \mathbf{L}_i), \\ 0, & \text{если } \min[\Phi(\mathbf{X}_{i-1}^{r+1} + \Delta_i^r \mathbf{L}_i), \Phi(\mathbf{X}_{i-1}^{r+1} - \Delta_i^r \mathbf{L}_i)] > \Phi(\mathbf{X}_{i-1}^{r+1}). \end{cases} \quad (3)$$

После завершения n шагов выполняется спуск в направлении вектора

$$\begin{pmatrix} \tilde{\mathbf{X}}^{r+1} \\ \mathbf{X}^r \end{pmatrix} \text{ по формуле} \quad \mathbf{X}^{r+1} = \mathbf{X}^r + \alpha^r (\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r), \quad (4)$$

где α^r — ускоряющий множитель. В различных модификациях метода Хука-Дживса множитель α^r может

- приниматься постоянным (обычно, равным 2),
- выбираться из условия $\Phi(\mathbf{X}^{r+1}) < \Phi(\mathbf{X}^r)$,
- находиться из условия локального минимума функции $\Phi(\mathbf{X})$

при движении из точки \mathbf{X}^r в направлении вектора $(\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r)$:

$$\min_{\alpha \in (-\infty, \infty)} \Phi(\mathbf{X}^r + \alpha(\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r)) = \Phi(\mathbf{X}^r + \alpha^r(\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r)) = \Phi(\mathbf{X}^{r+1}). \quad (5)$$

Заметим, что задачи (5) даже в случае одноэкстремальной функции $\Phi(\mathbf{X})$ могут быть многоэкстремальными задачами оптимизации и могут быть решены рассмотренными в главе 4 методами решения задач одномерной оптимизации.

Итерации заканчиваются при выполнении одного из стандартных условий окончания итераций:

$$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| \leq \varepsilon_X, \quad (6)$$

$$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \leq \varepsilon_\Phi \quad (6')$$

Вектор $\Delta^0 = (\Delta_1^0, \Delta_2^0, \dots, \Delta_n^0)$ является вектором свободных параметров метода - вектором «пробных шагов» по всем n координатным осям.

Известна модификация метода Хука-Дживса, в которой точка определяется не процедурами (2), (3), а методом Гаусса-Зейделя.

Схема метода Хука-Дживса:

Задаем начальную точку \mathbf{X}^0 , вектор «пробных» шагов $\Delta^0 = (\Delta_1^0, \Delta_2^0, \dots, \Delta_n^0)$ и полагаем $r=0$.

Последовательно для $i=1, 2, \dots, n$ по формулам (2), (3) находим точки $\tilde{\mathbf{X}}_1^{r+1}, \tilde{\mathbf{X}}_2^{r+1}, \dots, \tilde{\mathbf{X}}_n^{r+1} = \tilde{\mathbf{X}}^{r+1}$.

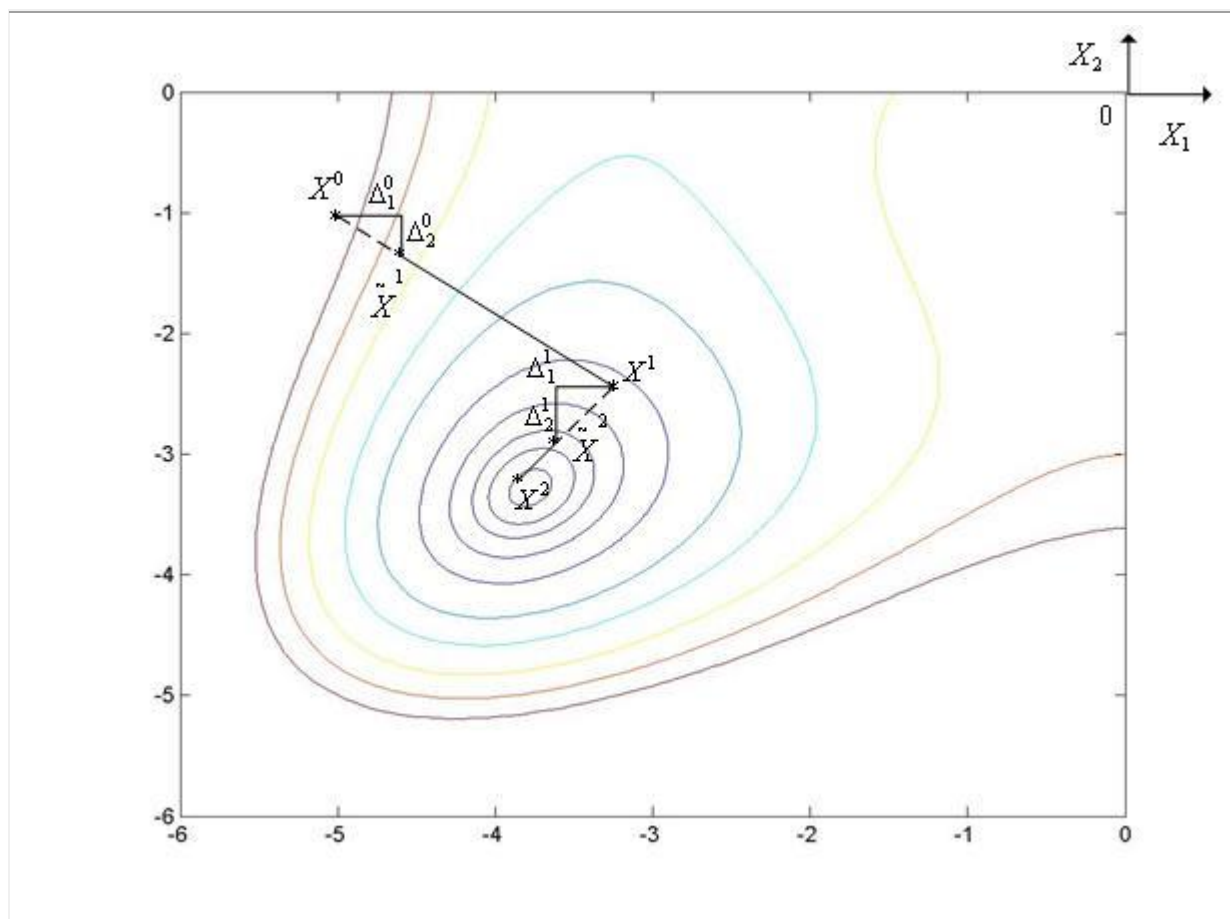
Если $\tilde{\mathbf{X}}^{r+1} \neq \tilde{\mathbf{X}}^r$, то переходим к п. 4). Иначе уменьшаем длины «пробных» шагов $\Delta_1^0, \Delta_2^0, \dots, \Delta_n^0$, например, вдвое и переходим к п.2).

Если условие окончания поиска (6) или (6') выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$ и заканчиваем вычисления. Иначе выполняем спуск в

направлении вектора $(\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r)$ по формуле (4), в которой ускоряющий множитель находится, например, из условия (5). Полагаем $r=r+1$ и переходим к п. 2 •

Метод Хука-Дживса иллюстрирует рис. 1, на котором показаны линии уровня функции. Ускоряющий множитель α^r находится из условия локального минимума функции $\Phi(\mathbf{X})$ при движении из точки \mathbf{X}^r в

направлении вектора $(\tilde{\mathbf{X}}^{r+1} - \mathbf{X}^r)$.



Метод Розенброка

$$\min_{\mathbf{X} \in \mathbb{R}^N} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

При решении задачи (1) методом Розенброка (методом вращающихся координат) используется преобразование на каждой итерации системы координат таким образом, чтобы в новой системе координат одна из осей совпадала с направлением предыдущего шага. Остальные оси новой системы координат обычно находят с помощью процедуры ортогонализации Грамма-Шмидта.

Рассмотрим произвольный набор векторов $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ пространства R^n . Поставим задачу построить на основе этих векторов ортонормированный набор векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ того же пространства R^n .

Напомним, что набор векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ называется ортонормированным, если для любых двух векторов из этого набора

выполняется условие

$$(\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \quad (2)$$

Или, другими словами, набор векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ортонормирован, если эти векторы линейно независимы и скалярное произведение любых двух из них равно единице.

Для построения векторов $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ применим индуктивный подход. Положим, что

$$\mathbf{e}_1 = \frac{\mathbf{p}_1}{\|\mathbf{p}_1\|}, \quad \rho_{11} = \|\mathbf{p}_1\|, \quad (3)$$

где $\|\cdot\|$ – символ евклидовой нормы (длины вектора). Полагая векторы $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{k-1}$ уже построенными будем искать вектор \mathbf{e}_k в виде

$$\rho_{kk} \mathbf{e}_k = \mathbf{p}_k - \sum_{i=1}^{k-1} \rho_{ik} \mathbf{e}_i. \quad (4)$$

Для отыскания неизвестных множителей ρ_{ik} умножим (4) скалярно на вектор $\mathbf{e}_j, j=1, 2, \dots, k-1$:

$$\begin{aligned} \rho_{kk}(\mathbf{e}_k, \mathbf{e}_1) &= 0 = (\mathbf{p}_k, \mathbf{e}_1) - \sum_{i=1}^{k-1} \rho_{ik}(\mathbf{e}_i, \mathbf{e}_1) = (\mathbf{p}_k, \mathbf{e}_1) - \rho_{1k}(\mathbf{e}_1, \mathbf{e}_1), \\ \rho_{kk}(\mathbf{e}_k, \mathbf{e}_2) &= 0 = (\mathbf{p}_k, \mathbf{e}_2) - \sum_{i=1}^{k-1} \rho_{ik}(\mathbf{e}_i, \mathbf{e}_2) = (\mathbf{p}_k, \mathbf{e}_2) - \rho_{2k}(\mathbf{e}_2, \mathbf{e}_2), \\ &\dots \\ \rho_{kk}(\mathbf{e}_k, \mathbf{e}_{k-1}) &= 0 = (\mathbf{p}_k, \mathbf{e}_{k-1}) - \sum_{i=1}^{k-1} \rho_{ik}(\mathbf{e}_i, \mathbf{e}_{k-1}) = (\mathbf{p}_k, \mathbf{e}_{k-1}) - \rho_{k-1,k}(\mathbf{e}_{k-1}, \mathbf{e}_{k-1}). \end{aligned}$$

Поскольку $(\mathbf{e}_j, \mathbf{e}_j) = 1$, имеем

$$\rho_{jk} = (\mathbf{p}_k, \mathbf{e}_j), \quad j \in [1, k-1]. \quad (5)$$

Множитель ρ_{kk} найдем из условия $\|\mathbf{e}_k\| = 1$:

$$\rho_{kk} = \left\| \mathbf{p}_k - \sum_{i=1}^{k-1} \rho_{ik} \mathbf{e}_i \right\|. \quad (6)$$

Процесс перехода от векторов $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ к векторам $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ согласно формулам (3) – (6) называется ортогонализацией Грамма-Шмидта •

Каждая итерация метода Розенброка состоит из двух этапов. В зависимости от модификации метода первый этап может выполняться с использованием различных методов. Рассмотрим применение на первом этапе итерационной формулы метода Гаусса-Зейделя. Приведем формулировку этой формулы, несколько отличную от уже рассмотренной формулировки.

Положим $\mathbf{X}_0^{r+1} = \mathbf{X}^r$, $\mathbf{X}_n^{r+1} = \mathbf{X}^{r+1}$ и пусть $\mathbf{e}_1^r, \mathbf{e}_2^r, \dots, \mathbf{e}_n^r$ – орты системы координат, используемой на r -ой итерации. Тогда итерационную формулу метода Гаусса-Зейделя можно записать в виде

$$\mathbf{X}_i^{r+1} = \mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{e}_i^r = \mathbf{X}_{i-1}^{r+1} + \mathbf{q}_i^r, \quad i=1, 2, \dots, n, \quad (7)$$

где коэффициенты λ_i^r находятся из условий

$$\min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda \mathbf{e}_i^r) = \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{e}_i^r) = \Phi(\mathbf{X}_i^{r+1}), \quad i=1, 2, \dots, n. \quad (8)$$

На втором этапе каждой из итераций система векторов $\mathbf{e}_1^r, \mathbf{e}_2^r, \dots, \mathbf{e}_n^r$ с использованием ортогонализации Грамма-Шмидта заменяется новой системой линейно независимых векторов $\mathbf{e}_1^{r+1}, \mathbf{e}_2^{r+1}, \dots, \mathbf{e}_n^{r+1}$.

Схема метода Розенброка:

1. Задаем начальную точку \mathbf{X}^0 , полагаем $r=0$, $i=1$, и орты исходной системы координат обозначаем $\mathbf{e}_1^0, \mathbf{e}_2^0, \dots, \mathbf{e}_n^0$.
2. Исходя из точки \mathbf{X}^r по формулам (7), (8) выполняем одну итерацию по методу Гаусса-Зейделя – получаем точку \mathbf{X}^{r+1} и совокупность векторов $\mathbf{q}_1^r, \mathbf{q}_2^r, \dots, \mathbf{q}_n^r$.

3. Если одно из стандартных условий окончания итераций

$$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| \leq \varepsilon_X, \quad (9)$$

$$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \leq \varepsilon_\Phi$$

выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$, и заканчиваем вычисления. Иначе переходим к п.4.

4. На основе векторов $\mathbf{q}_1^r, \mathbf{q}_2^r, \dots, \mathbf{q}_n^r$ находим векторы $\mathbf{p}_1^r, \mathbf{p}_2^r, \dots, \mathbf{p}_n^r$:

$$\mathbf{p}_i^r = \sum_{j=i}^n \mathbf{q}_j^r, \quad i=1, 2, \dots, n. \quad (10)$$

5. С помощью процедуры ортогонализации Грамма-Шмидта (3)–(6) выполняем переход от системы векторов $\mathbf{p}_1^r, \mathbf{p}_2^r, \dots, \mathbf{p}_n^r$ к системе векторов $\mathbf{e}_1^{r+1}, \mathbf{e}_2^{r+1}, \dots, \mathbf{e}_n^{r+1}$, полагаем $r=r+1$ и переходим к п. 2 •

Заметим, что из формулы (10) следует равенство $\mathbf{p}_n^r = \mathbf{q}_n^r$.

Метод Розенброка иллюстрирует рис. 1, на котором показаны линии уровня функции,

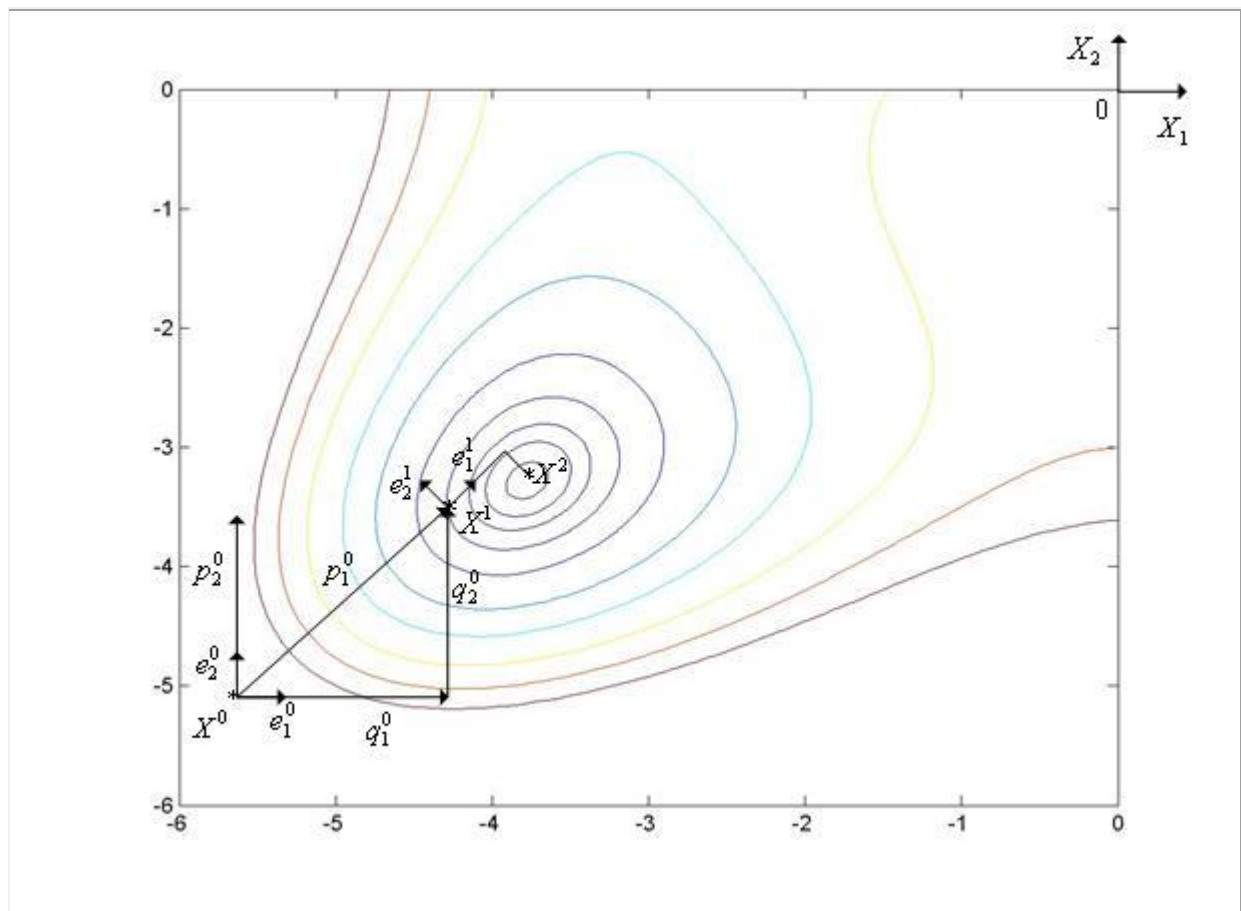


Рис. 1. Траектория поиска минимума методом Розенброка.

Метод сопряженных направлений

Рассмотрим следующую задачу локальной безусловной оптимизации: найти минимум критерия оптимальности $\Phi(\mathbf{X})$, определенного в n -мерном евклидовом пространстве R^n ,

$$\min_{\mathbf{X} \in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

Введем прежде следующие понятия: векторы $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_k$, принадлежащие пространству R^n , называются векторами сопряженными относительно матрицы A ($n \times n$) (A – ортогональными векторами), если $(A\mathbf{p}_i, \mathbf{p}_j) = 0$ для всех $i \neq j, i, j \in [1, k]$.

В методе сопряженных направлений применяется итерационная формула метода Гаусса-Зейделя в виде, близком к использованному в методе Розенброка.

Положим $\mathbf{X}_0^{r+1} = \mathbf{X}^r$ и пусть $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ – орты используемой системы координат. Тогда итерационную формулу метода Гаусса-Зейделя можно записать в виде

$$\mathbf{X}_i^{r+1} = \mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{e}_i = \mathbf{X}_{i-1}^{r+1} + \mathbf{p}_i^r, \quad i=1,2,\dots,n, \quad (2)$$

где коэффициенты λ_i^r находятся из условий

$$\min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda \mathbf{e}_i) = \Phi(\mathbf{X}_{i-1}^{r+1} + \lambda_i^r \mathbf{e}_i) = \Phi(\mathbf{X}_i^{r+1}), \quad i=1,2,\dots,n. \quad (3)$$

Схема метода сопряженных направлений:

Задаем начальную точку \mathbf{X}^0 и полагаем $r=0, i=1$.

Последовательно для $i=1,2,\dots,n$ по формулам (2), (3) находим точки $\mathbf{X}_1^{r+1}, \mathbf{X}_2^{r+1}, \dots, \mathbf{X}_n^{r+1}$.

Исходя из точки \mathbf{X}_n^{r+1} , еще раз находим минимум функции $\Phi(\mathbf{X})$ вдоль первого координатного направления - вычисляем координаты точки

$$\mathbf{X}_{n+1}^{r+1} = \mathbf{X}_n^{r+1} + \tilde{\lambda}_1^r \mathbf{e}_1, \quad (4)$$

где коэффициент $\tilde{\lambda}_1^r$ находится из условия

$$\min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_{n+1}^{r+1} + \lambda \mathbf{e}_1) = \Phi(\mathbf{X}_n^{r+1} + \tilde{\lambda}_1^r \mathbf{e}_1) = \Phi(\mathbf{X}_{n+1}^{r+1}). \quad (5)$$

Исходя из точки \mathbf{X}_1^{r+1} , находим минимум функции $\Phi(\mathbf{X})$ вдоль вектора $\mathbf{p}_{n+1}^r = \mathbf{X}_{n+1}^{r+1} - \mathbf{X}_1^{r+1}$ - вычисляем

$$\mathbf{X}^{r+1} = \mathbf{X}_1^r + \lambda_{n+1}^{r+1} \mathbf{p}_{n+1}^r, \quad (6)$$

где коэффициент λ_{n+1}^{r+1} находится из условия

$$\min_{\lambda \in (-\infty, \infty)} \Phi(\mathbf{X}_1^{r+1} + \lambda \mathbf{p}_{n+1}^r) = \Phi(\mathbf{X}^r + \lambda_{n+1}^{r+1} \mathbf{p}_{n+1}^r) = \Phi(\mathbf{X}^{r+1}). \quad (7)$$

Если одно из стандартных условий окончания итераций

$$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| \leq \varepsilon_X, \quad (8)$$

$$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \leq \varepsilon_\Phi$$

выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$, и заканчиваем вычисления. Иначе - полагаем $r=r+1$ и переходим к п.2) •

Метод сопряженных направлений иллюстрирует рис. 1, на котором показаны линии уровня функции.

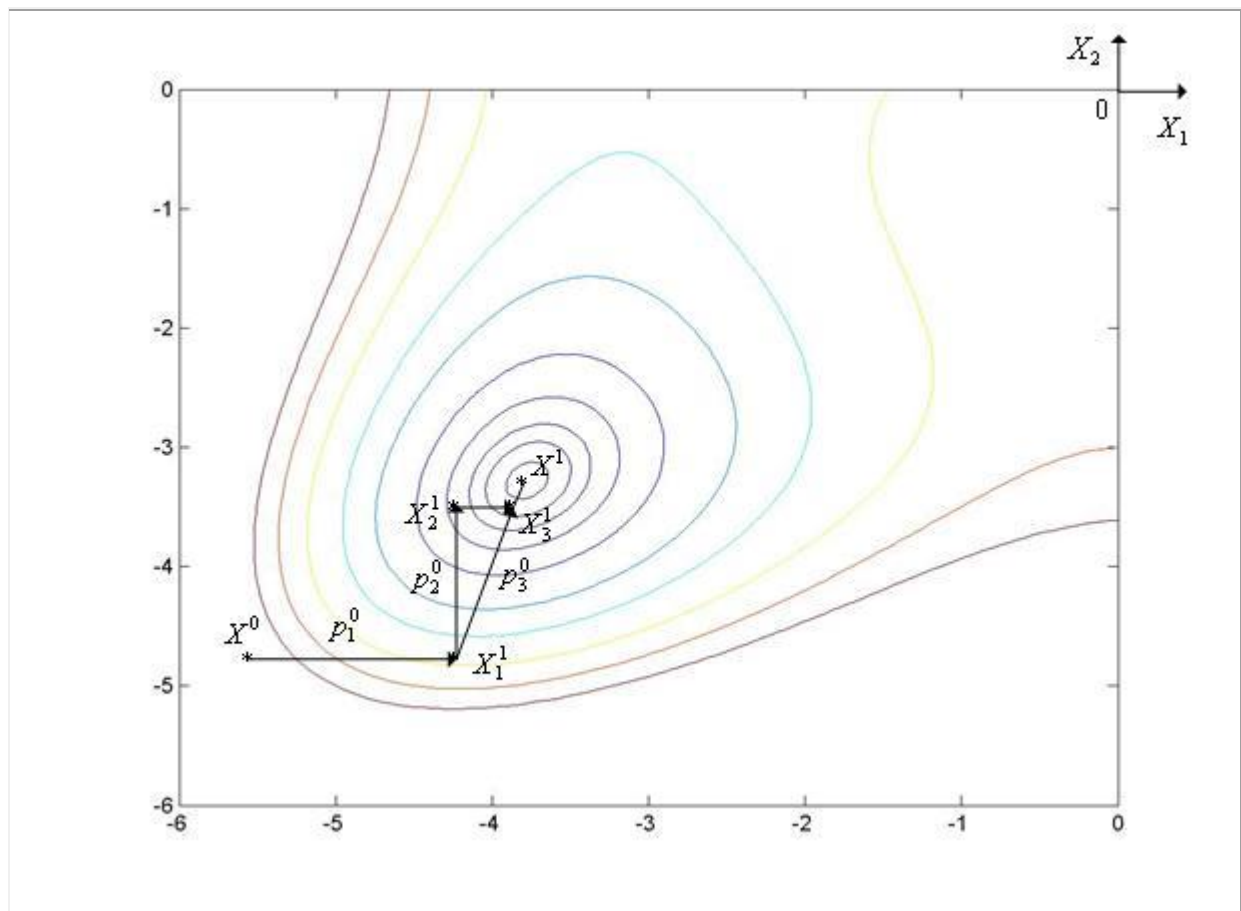


Рис. 1. Траектория поиска минимума функции методом сопряженных направлений.

Метод наискорейшего спуска. Метод дробления шага

Рассматривается следующая многомерная задача локальной безусловной оптимизации: найти минимум критерия оптимальности $\Phi(\mathbf{X})$, определенного в n -мерном евклидовом пространстве R^n ,

$$\min_{\mathbf{X} \in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

Положим, что функция $\Phi(\mathbf{X})$ всюду дифференцируема в n -мерном евклидовом пространстве R^n .

Направление спуска в градиентных методах оптимизации совпадает с направлением антиградиента минимизируемой функции $\Phi(\mathbf{X})$.

Итерационная формула градиентных методов оптимизации имеет вид

$$\mathbf{X}^{r+1} = \mathbf{X}^r + \lambda^r \mathbf{S}^r. \quad (2)$$

Здесь λ^r - длина шага на r -ой итерации в направлении \mathbf{S}^r , где

$$\mathbf{S}^r = -\frac{\nabla \Phi^r}{\|\nabla \Phi^r\|} \quad (3)$$

единичный вектор направления антиградиента функции $\Phi(\mathbf{X})$ в точке \mathbf{X}^r , $\|\cdot\|$ - некоторая векторная норма, например, евклидова. Напомним, что

градиент функции $\Phi(\mathbf{X})$ в точке \mathbf{X}^r есть значение вектора частных производных этой функции в точке \mathbf{X}^r :

$$\nabla \Phi^r = \nabla \Phi(\mathbf{X}^r) = \left(\begin{array}{c} \frac{\partial \Phi(\mathbf{X})}{\partial x_1} \\ \frac{\partial \Phi(\mathbf{X})}{\partial x_2} \\ \vdots \\ \frac{\partial \Phi(\mathbf{X})}{\partial x_n} \end{array} \right) \bigg|_{\mathbf{X}=\mathbf{X}^r}$$

Различные градиентные методы оптимизации отличаются между собой правилами выбора длины шага λ^r .

Градиентный метод наискорейшего спуска.

Градиентный метод наискорейшего спуска в качестве длины шага λ^r использует величину, при которой достигается минимум функции $\Phi(\mathbf{X})$ в направлении \mathbf{S}^r :

$$\Phi(\mathbf{X}^{r+1}) = \Phi(\mathbf{X}^r + \lambda^r \mathbf{S}^r) = \min_{\lambda \in \mathbb{R}^1} \Phi(\mathbf{X}^r + \lambda \mathbf{S}^r). \quad (4)$$

Задача (2) есть одномерная задача оптимизации, которая может быть решена уже рассмотренными методами.

Схема метода:

Задаем начальную точку \mathbf{X}^0 и полагаем счетчик числа итераций $r=0$.

По формуле (3) вычисляем компоненты вектора \mathbf{S}^r .

Каким-либо методом решаем одномерную задачу оптимизации (4) – определяем точку \mathbf{X}^{r+1} .

Вычисляем величину $\Phi(\mathbf{X}^{r+1})$ – значение функции $\Phi(\mathbf{X})$ в точке \mathbf{X}^{r+1} .

Проверяем условие окончания поиска (см. ниже). Если условие окончания поиска выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$ и завершаем итерации. Иначе – полагаем $r=r+1$, переходим к п. 2 •

В качестве критерия окончания поиска можно использоваться одно из стандартных условий окончания итераций

$$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| = \lambda^r \leq \varepsilon_X, \quad (5)$$

$$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \leq \varepsilon_\Phi. \quad (6)$$

В качестве критерия окончания поиска можно использоваться также условие

$$\|\nabla \Phi^r\| \leq \varepsilon_\nabla, \quad (7)$$

где ε_∇ – константа, определяющая требуемую точность решения по градиенту функции $\Phi(\mathbf{X})$.

Градиентный метод наискорейшего спуска иллюстрирует рис. 1.

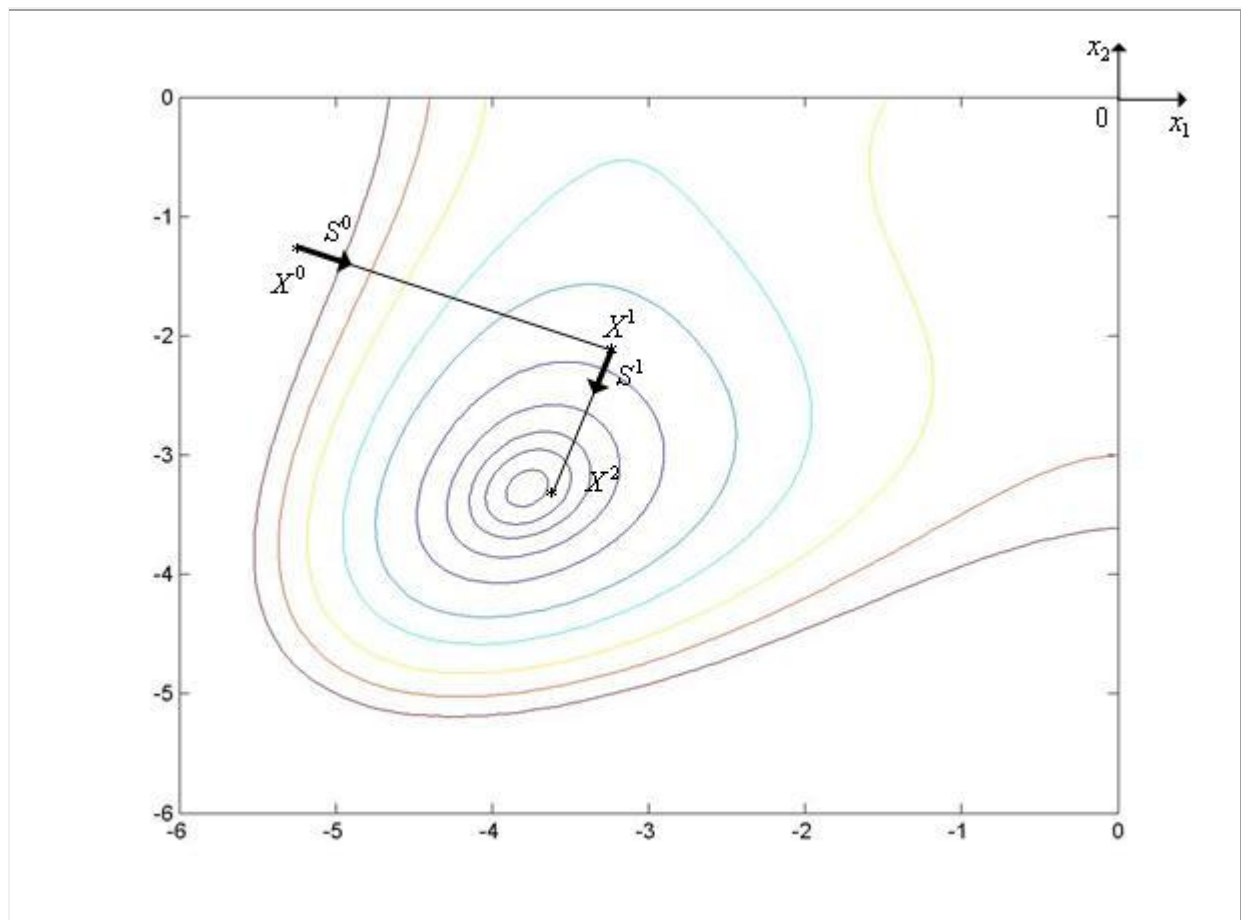


Рис. 1. Траектория поиска минимума функции градиентным методом наискорейшего спуска.

Градиентный метод с дроблением шага

В градиентном методе с дроблением шага точка \mathbf{X}^{r+1} определяется по формуле

$$\mathbf{X}^{r+1} = \mathbf{X}^r + \lambda^r \mathbf{S}^r, \quad (8)$$

где величина шага λ^r находится из условия

$$\Phi(\mathbf{X}^r) - \Phi(\mathbf{X}^{r+1}) \geq 0.5 \lambda^r \|\nabla \Phi^r\|. \quad (9)$$

Схема метода:

Задаем начальную точку \mathbf{X}^0 , начальную величину шага λ^0 и коэффициент дробления шага $\nu \in (0, 1]$. Полагаем счетчик числа итераций $r=0$.

По формуле (8) вычисляем компоненты вектора \mathbf{X}^{r+1} .

Вычисляем величину $\Phi(\mathbf{X}^{r+1})$ - значение функции $\Phi(\mathbf{X})$ в точке \mathbf{X}^{r+1} .

Если условие (9) выполнено, то переходим к следующему пункту. Иначе – переходим к пункту 6.

Полагаем $\lambda^r = \nu \lambda^r$ и переходим к пункту 2.

Проверяем условие окончания поиска (см. ниже). Если условие окончания поиска выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$, и завершаем итерации. Иначе –

полагаем $r=r+1$ переходим к п. 2.

В качестве критерия окончания поиска можно использовать условия (5) – (7).

Градиентный метод дробления шага иллюстрирует рис. 2.

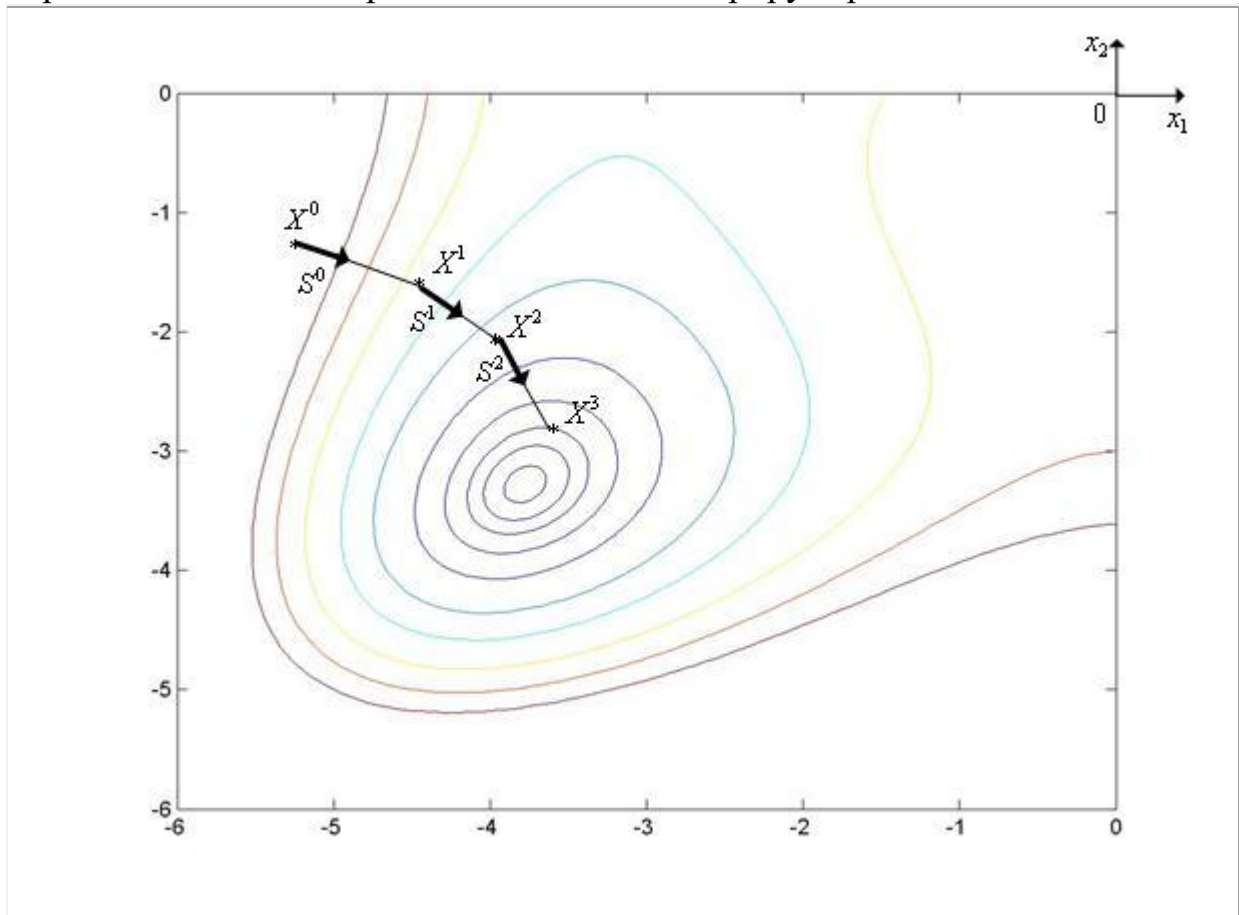


Рис. 2. Траектория поиска минимума функции градиентным методом дробления шага.

Пример

Выполнить несколько итераций (не менее двух) решения двумерной задачи локальной безусловной оптимизации

$$\min_{\mathbf{X} \in \mathbb{R}^2} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*, \quad (10)$$

$$\text{где } \Phi(\mathbf{X}) = \Phi(x, y) = x^2 + y^2 + 3(x+y)^2 = 4x^2 + 4y^2 + 6xy, \quad (11)$$

градиентным методом с дроблением шага, исходя из точки $\mathbf{X}^0 = (x^0, y^0) = (-2.0, 1.0)$.

Принять $\lambda^0 = 1.0$, $v = 0.7$, в качестве нормы вектора градиента использовать евклидову норму.

Траекторию поиска изобразить на рис. 3, на котором приведены линии уровня квадратичной функции (11).

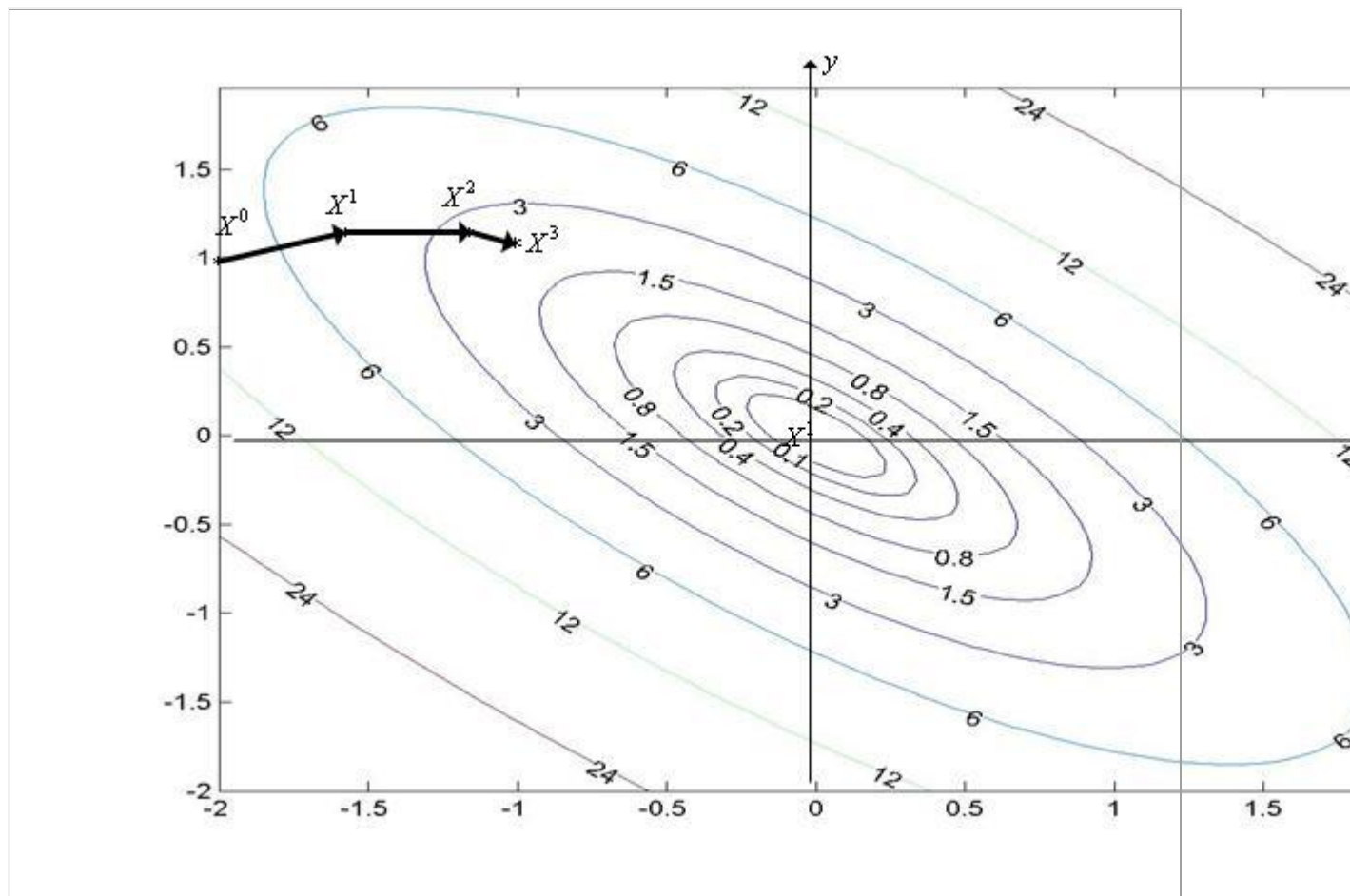


Рис. 3. К примеру. Фрагмент (три итерации) траектории поиска минимума функции (11) градиентным методом с дроблением шага, исходя из точки $X_0=(x_0,y_0)=(-2.0,1.0)$.

Итерация градиентного метода с дроблением шага для задачи (10), (11) имеет вид

$$\mathbf{X}^{r+1} = \begin{pmatrix} x^{r+1} \\ y^{r+1} \end{pmatrix} = \begin{pmatrix} x^r \\ y^r \end{pmatrix} + \lambda^r \begin{pmatrix} s_x^r \\ s_y^r \end{pmatrix} \quad (12)$$

$$\text{где } \begin{pmatrix} s_x^r \\ s_y^r \end{pmatrix} = \mathbf{S}^r = -\frac{\nabla \Phi^r}{\|\nabla \Phi^r\|}, \quad \nabla \Phi^r = \begin{pmatrix} \frac{\partial \Phi(x^r, y^r)}{\partial x} \\ \frac{\partial \Phi(x^r, y^r)}{\partial y} \end{pmatrix} \quad (13)$$

а величина шага λ^r находится из условия

$$\Phi(\mathbf{X}^r) - \Phi(\mathbf{X}^{r+1}) \geq 0.5 \lambda^r \|\nabla \Phi^r\|. \quad (14)$$

Найдем явные выражения для частных производных функции (11):

$$\frac{\partial \Phi(x, y)}{\partial x} = 8x + 6y; \quad \frac{\partial \Phi(x, y)}{\partial y} = 8y + 6x. \quad (15)$$

Таким образом, из (12), (13), (15) имеем искомую итерационную формулу градиентного метода с дроблением шага для задачи (10), (11).

$$\mathbf{X}^{r+1} = \begin{pmatrix} x^{r+1} \\ y^{r+1} \end{pmatrix} = \begin{pmatrix} x^r \\ y^r \end{pmatrix} + \lambda^r \begin{pmatrix} s_x^r \\ s_y^r \end{pmatrix}, \quad s_x^r = -\frac{8x^r + 6y^r}{\|\nabla\Phi^r\|}, \quad s_y^r = -\frac{8y^r + 6x^r}{\|\nabla\Phi^r\|},$$

$$\|\nabla\Phi(\mathbf{X})\| = \sqrt{(8x+6y)^2 + (8y+6x)^2}. \quad (16)$$

Первая итерация ($r=0$).

Из формул (15), (16) последовательно имеем

$$\mathbf{X}^1 = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} x^0 \\ y^0 \end{pmatrix} + \lambda^0 \begin{pmatrix} s_x^0 \\ s_y^0 \end{pmatrix}$$

$$\frac{\partial\Phi(x^0, y^0)}{\partial x} = 8x^0 + 6y^0 = 8(-2) + 6 = -10, \quad \frac{\partial\Phi(x^0, y^0)}{\partial y} = 8y^0 + 6x^0 = 8 + 6(-2) = -4,$$

$$\|\nabla\Phi^0\| = \sqrt{(-10)^2 + (-4)^2} \approx 10.77,$$

$$s_x^0 \approx \frac{10}{10.77} \approx 0.93, \quad s_y^0 \approx \frac{4}{10.77} \approx 0.37,$$

$$\mathbf{X}^1 = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} x^0 \\ y^0 \end{pmatrix} + \lambda^0 \begin{pmatrix} s_x^0 \\ s_y^0 \end{pmatrix} \approx \begin{pmatrix} -2 \\ 1 \end{pmatrix} + 0.5 \begin{pmatrix} 0.93 \\ 0.37 \end{pmatrix} \approx \begin{pmatrix} -1.54 \\ 1.18 \end{pmatrix}.$$

Таким образом, $\mathbf{X}^1 = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} -1.54 \\ 1.18 \end{pmatrix}$ (см. рис. 3).

Условие (14) на первой итерации имеет вид

$$\Phi(\mathbf{X}^0) - \Phi(\mathbf{X}^1) \geq 0.5 \lambda^0 \|\nabla\Phi^0\|.$$

Поскольку

$$\Phi(\mathbf{X}^0) = \Phi(x^0, y^0) = 4(x^0)^2 + 4(y^0)^2 + 6x^0y^0 = 4(-2)^2 + 4 + 6(-2) = 8,$$

$$\Phi(\mathbf{X}^1) = \Phi(x^1, y^1) = 4(x^1)^2 + 4(y^1)^2 + 6x^1y^1 \approx 4(-1.54)^2 + 4(1.18)^2 + 6(-1.54)(1.18) \approx 4.16,$$

левая часть этого неравенства равна $8 - 4.16 = 3.84$. Его правая часть, легко видеть, равна $0.5^2 \cdot 10.77 = 2.69$.

Таким образом, на первой итерации условие (14) выполняется и величина шага λ должна быть изменена: $\lambda^1 = 0.7 \cdot 0.5 = 0.35$.

Вторая итерация ($r=1$).

Аналогично первой итерации последовательно имеем

$$\mathbf{X}^2 = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} + \lambda^1 \begin{pmatrix} s_x^1 \\ s_y^1 \end{pmatrix}$$

$$\frac{\partial\Phi(x^1, y^1)}{\partial x} = 8x^1 + 6y^1 = 8(-1.54) + 6 \cdot 1.18 = -5.24, \quad \frac{\partial\Phi(x^1, y^1)}{\partial y} = 8y^1 + 6x^1 = 8 \cdot 1.18 + 6(-1.54) = 0.20,$$

$$\|\nabla\Phi^1\| = \sqrt{(-5.24)^2 + (0.20)^2} \approx 5.24,$$

$$s_x^1 \approx \frac{5.24}{5.24} \approx 1.0, \quad s_y^1 \approx -\frac{0.2}{5.24} \approx -0.04,$$

$$\mathbf{X}^2 = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} + \lambda^1 \begin{pmatrix} s_x^1 \\ s_y^1 \end{pmatrix} \approx \begin{pmatrix} -1.54 \\ 1.18 \end{pmatrix} + 0.35 \begin{pmatrix} 1.0 \\ -0.04 \end{pmatrix} \approx \begin{pmatrix} -1.19 \\ 1.22 \end{pmatrix}.$$

Таким образом, $\mathbf{X}^2 = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} = \begin{pmatrix} -1.19 \\ 1.22 \end{pmatrix}$ (см. рис. 3).

Условие (14) на второй итерации имеет вид

$$\Phi(\mathbf{X}^1) - \Phi(\mathbf{X}^2) \geq 0.5 \lambda^1 \|\nabla \Phi^1\|.$$

Поскольку

$$\Phi(\mathbf{X}^1) \approx 4.16$$

$$\Phi(\mathbf{X}^2) = \Phi(x^2, y^2) = 4(x^2)^2 + 4(y^2)^2 + 6x^2y^2 \approx 4(-1.19)^2 + 4(1.22)^2 + 6(-1.19)(1.22) \approx 2.91,$$

левая часть этого неравенства равна $4.16 - 2.91 = 1.25$. Его правая часть, легко видеть, равна $0.5 \cdot 0.35 \cdot 5.24 = 0.92$.

Таким образом, на второй итерации условие (14) выполняется и величина шага λ должна быть изменена: $\lambda^2 = 0.7 \cdot 0.35 \approx 0.24$.

Третья итерация ($r=2$).

Аналогично первой итерации последовательно имеем

$$\mathbf{X}^3 = \begin{pmatrix} x^3 \\ y^3 \end{pmatrix} = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} + \lambda^2 \begin{pmatrix} s_x^2 \\ s_y^2 \end{pmatrix}$$

$$\frac{\partial \Phi(x^2, y^2)}{\partial x} = 8x^2 + 6y^2 = 8(-1.19) + 6 \cdot 1.22 = -2.20, \quad \frac{\partial \Phi(x^2, y^2)}{\partial y} = 8y^2 + 6x^2 = 8 \cdot 1.22 + 6(-1.19) = 2.62,$$

$$\|\nabla \Phi^2\| = \sqrt{(-2.20)^2 + (2.62)^2} \approx 3.42,$$

$$s_x^2 \approx \frac{2.20}{3.42} \approx 0.64, \quad s_y^2 \approx -\frac{2.62}{3.42} \approx -0.77,$$

$$\mathbf{X}^3 = \begin{pmatrix} x^3 \\ y^3 \end{pmatrix} = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix} + \lambda^2 \begin{pmatrix} s_x^2 \\ s_y^2 \end{pmatrix} = \begin{pmatrix} -1.19 \\ 1.22 \end{pmatrix} + 0.24 \begin{pmatrix} 0.64 \\ -0.77 \end{pmatrix} \approx \begin{pmatrix} -1.04 \\ 1.03 \end{pmatrix}.$$

$$\mathbf{X}^3 = \begin{pmatrix} x^3 \\ y^3 \end{pmatrix} = \begin{pmatrix} -1.04 \\ 1.03 \end{pmatrix}$$

Таким образом, (см. рис. 3).

Условие (14) на третьей итерации имеет вид

$$\Phi(\mathbf{X}^2) - \Phi(\mathbf{X}^3) \geq 0.5 \lambda^2 \|\nabla \Phi^2\|.$$

Поскольку

$$\Phi(\mathbf{X}^2) \approx 2.91,$$

$$\Phi(\mathbf{X}^3) = \Phi(x^3, y^3) = 4(x^3)^2 + 4(y^3)^2 + 6x^3y^3 \approx 4(-1.04)^2 + 4(1.03)^2 + 6(-1.04)(1.03) \approx 2.14,$$

левая часть этого неравенства равна $2.91 - 2.14 = 0.77$. Его правая часть, легко видеть, равна $0.5 \cdot 0.24 \cdot 3.42 = 0.41$.

Таким образом, на третьей итерации условие (14) выполняется и величина шага λ должна быть изменена: $\lambda^3 = 0.7 \cdot 0.24 \approx 0.17$.

Метод с возвратом при неудачном шаге. Метод наилучшей пробы

Рассматривается следующая многомерная задача локальной безусловной оптимизации: найти минимум критерия оптимальности $\Phi(\mathbf{X})$, определенного в n -мерном евклидовом пространстве R^n ,

$$\min_{\mathbf{X} \in R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*. \quad (1)$$

При решении задачи (1) методом с возвратом при неудачном шаге

(одношаговый метод оптимизации) используется итерационная формула

$$\mathbf{X}^{r+1} = \mathbf{X}^r + \lambda^r \frac{\mathbf{\Psi}^r}{\|\mathbf{\Psi}^r\|}, \quad (2)$$

где λ^r - величина шага на r -ой итерации, $\mathbf{\Psi}^r = (\psi_1^r, \psi_2^r, \dots, \psi_n^r)$ - реализация n -мерного случайного вектора, $\|\cdot\|$ - некоторая векторная норма. Обычно в качестве координат вектора $\mathbf{\Psi}^r$ используют независимые случайные величины, равномерно распределенные в интервале $[-1, 1]$.

Схема метода с возвратом при неудачном шаге.

Задаем начальную точку \mathbf{X}^0 , начальную длину шага λ^0 и полагаем счетчик числа итераций $r=0$.

Задаем начальное значение счетчика числа неудачных попыток $k=1$.

Получаем реализацию случайных чисел $\psi_1^r, \psi_2^r, \dots, \psi_n^r$ - компонент вектора $\mathbf{\Psi}^r$ и по формуле (2) находим пробную точку \mathbf{X}^{r+1} .

Вычисляем значение $\Phi(\mathbf{X}^{r+1})$ функции $\Phi(\mathbf{X})$ в точке \mathbf{X}^{r+1} .

Если $\Phi(\mathbf{X}^{r+1}) < \Phi(\mathbf{X}^r)$, то полагаем $r=r+1$ и переходим к п.3. Иначе – переходим к п.6.

Полагаем $k=k+1$. Если $k < K$, то переходим к п.3. Иначе – переходим к п.7. Здесь K – предельное количество неудачных попыток (свободный параметр метода). Рекомендуется $K=3n$.

Проверяем условие окончания поиска (см. ниже). Если условие окончания поиска выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$ и завершаем итерации. Иначе –

полагаем $r=r+1$, $\lambda^r = \alpha \lambda^r$ и переходим к п. 2. Здесь $\alpha \in (0, 1)$ - коэффициент уменьшения шага (свободный параметр метода) •

В качестве условия окончания поиска можно использоваться одно из стандартных условий окончания итераций:

$$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| = \lambda^r \leq \varepsilon_X, \quad (3)$$

где ε_X - константа, определяющая требуемую точность решения по \mathbf{X} ;

$$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \leq \varepsilon_\Phi, \quad (4)$$

где ε_Φ - константа, определяющая требуемую точность решения по Φ .

Метод с возвратом при неудачном шаге иллюстрирует рис. 1.

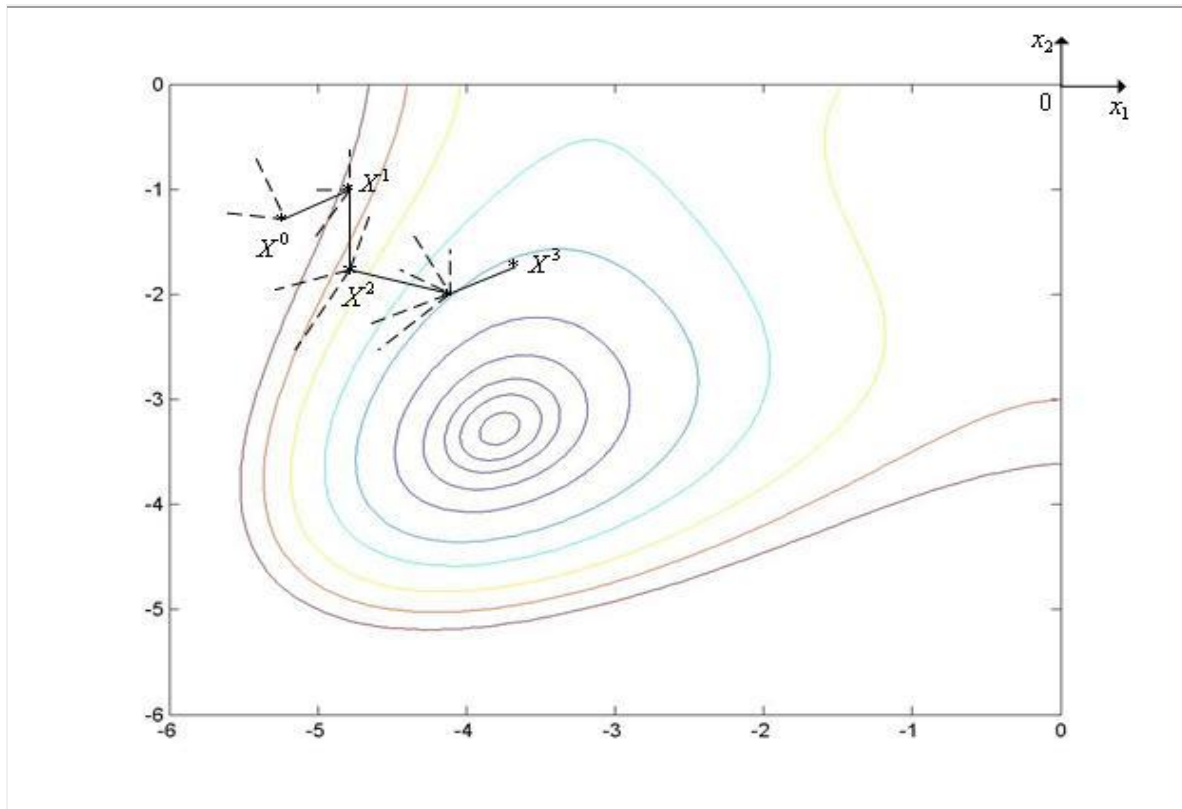


Рис. 1. Траектория поиска минимума функции методом с возвратом при неудачном шаге. Пунктиром на рисунке показаны неудачные шаги.

Модификацией метода с возвратом при неудачном шаге является метод наилучшей пробы (также одношаговый метод оптимизации).

Схема метода наилучшей пробы.

Задаем начальную точку \mathbf{X}^0 , начальную длину шага λ^0 и полагаем счетчик числа итераций $r=0$.

Генерируем M случайных векторов $\Psi_i^r, i \in [1, M]$ и по формуле (2) находим пробные точки $\mathbf{X}_i^{r+1}, i \in [1, M]$

Вычисляем значения $\Phi(\mathbf{X}_i^{r+1})$ функции $\Phi(\mathbf{X})$ в пробных точках $\mathbf{X}_i^{r+1}, i \in [1, M]$ и находим минимальное из этих значений

$$\Phi(\mathbf{X}^{r+1}) = \Phi(\mathbf{X}_k^{r+1}) = \min_{i \in [1, M]} \Phi(\mathbf{X}_i^{r+1}).$$

Если $\Phi(\mathbf{X}^{r+1}) < \Phi(\mathbf{X}^r)$, то полагаем $r=r+1$ и переходим к п.2. Иначе – переходим к п.5.

Проверяем условие окончания поиска (см. (3), (4)). Если условие

окончания поиска выполнено, то полагаем $\mathbf{X}^* \approx \mathbf{X}^{r+1}$ и завершаем итерации. Иначе – полагаем $r=r+1$, $\lambda^r = \alpha \lambda^r$ и переходим к п.2. Здесь $\alpha \in (0, 1)$ – коэффициент уменьшения шага (свободный параметр метода) •

Метод наилучшей пробы иллюстрирует рис. 2.

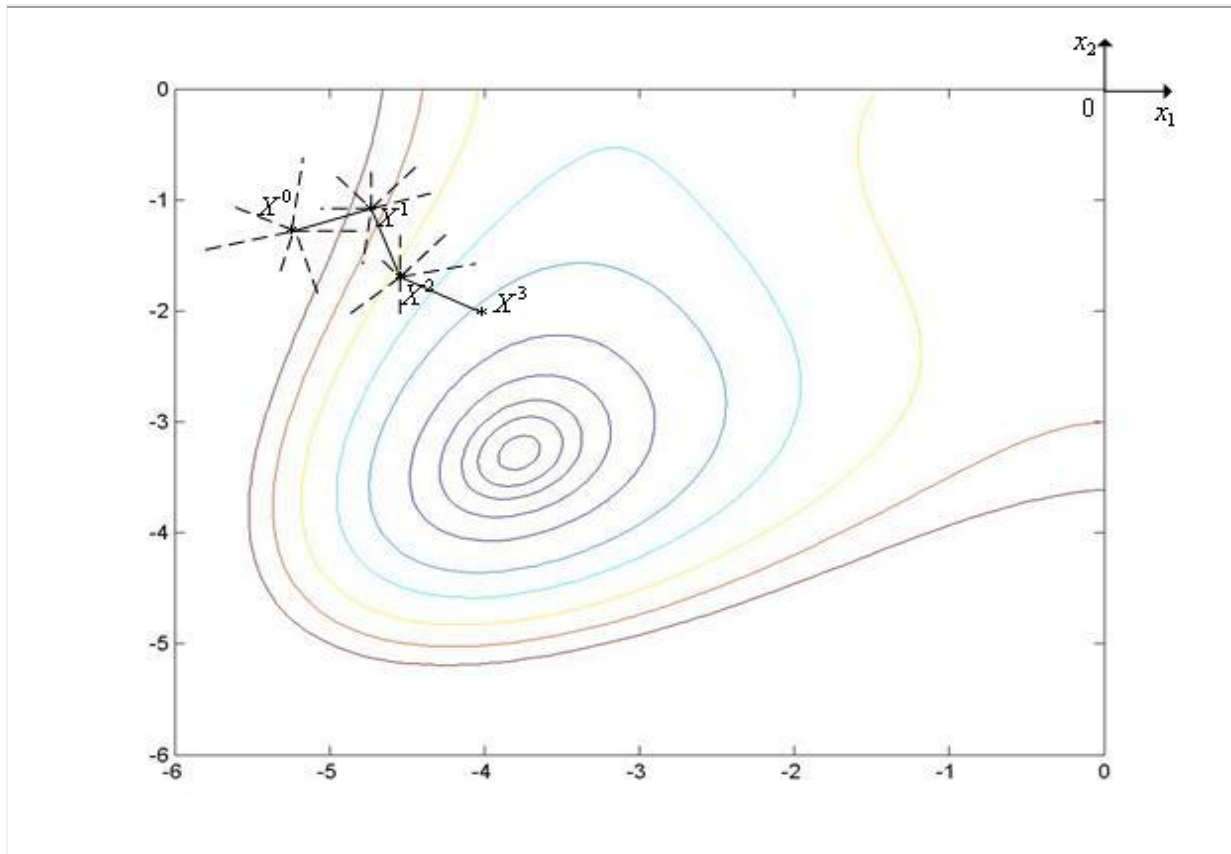


Рис. 2. Траектория поиска минимума функции Химмельблау методом наилучшей пробы. Пунктиром на рисунке показаны неудачные пробы.