



МИНОБРНАУКИ РОССИИ

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»**

Факультет информатики и прикладной математики
Кафедра прикладной математики и экономико-математических
методов

ОТЧЁТ

по дисциплине:

«Методы оптимизации»

на тему:

«Задание 15. Метод наискорейшего спуска (градиента)»

Направление: 01.03.02

Обучающийся: Бронников Егор Игоревич

Группа: ПМ-1901

Санкт-Петербург
2021

Дано:

$$f(x_1, x_2, x_3) = (3x_1 - 3x_2 - 5)^2 + (6x_1 - x_2 - x_3 - 2)^2 + (2x_1 + 5x_2 + x_3 - 1)^2$$

Условие:

Найти стационарную точку методом наискорейшего спуска (градиента).

Решение:

Определим, является ли функция $f(x_1, x_2, x_3)$ выпуклой (вогнутой) или нет.

Найдём первые производные функции:

$$\frac{df}{dx_1} = 98x_1 - 10x_2 - 8x_3 - 58$$

$$\frac{df}{dx_2} = -10x_1 + 70x_2 + 12x_3 + 24$$

$$\frac{df}{dx_3} = -8x_1 + 12x_2 + 4x_3 + 2$$

Составим матрицу Гессе $H(X)$ для функции $f(x_1, x_2, x_3)$ и определим знак её угловых миноров:

$$H(X) = \begin{pmatrix} 98 & -10 & -8 \\ -10 & 70 & 12 \\ -8 & 12 & 4 \end{pmatrix}$$

Вычисляем главные миноры:

$$M_1(\mathbf{H}) = 98 > 0, \quad M_2(\mathbf{H}) = 6760 > 0, \quad M_3(\mathbf{H}) = |\mathbf{H}| = 10368 > 0$$

Матрица \mathbf{H} – положительно определённая матрица и, следовательно, $f(x_1, x_2, x_3)$ – выпуклая функция, которая имеет минимум в некоторой точке X^* .

Рассмотрим луч, выходящий из произвольной точки $X = (x_1, x_2, x_3)$, $X' = X + tS$, и направленный в сторону убывания функции $f(x_1, x_2, x_3)$.

Координаты точек на этом луче имеют следующий вид:

$$x_1' = x_1 - t \frac{df}{dx_1} = x_1 - t(98x_1 - 10x_2 - 8x_3 - 58)$$

$$x_2' = x_2 - t \frac{df}{dx_2} = x_2 - t(-10x_1 + 70x_2 + 12x_3 + 24)$$

$$x_3' = x_3 - t \frac{df}{dx_3} = x_3 - t(-8x_1 + 12x_2 + 4x_3 + 2)$$

Построим вспомогательную функцию $\phi(t)$, равную функции $f(x_1, x_2, x_3)$ на рассматриваемом луче:

$$\begin{aligned} \phi(t; X) &= f(x_1', x_2', x_3') = f\left(x_1 - t \frac{df}{dx_1}, x_2 - t \frac{df}{dx_2}, x_3 - t \frac{df}{dx_3}\right) = \\ &= f(x_1 - t(98x_1 - 10x_2 - 8x_3 - 58), x_2 - t(-10x_1 + 70x_2 + 12x_3 + 24), x_3 - t(-8x_1 + 12x_2 + 4x_3 + 2)) \end{aligned}$$

Градиент $f(x_1, x_2, x_3)$ и вектор $S = -\text{grad } f(X)$ определяется выражением:

$$\begin{aligned} \text{grad } f(X) &= \left(\frac{df}{dx_1}, \frac{df}{dx_2}, \frac{df}{dx_3} \right) = \\ &= (98x_1 - 10x_2 - 8x_3 - 58, -10x_1 + 70x_2 + 12x_3 + 24, -8x_1 + 12x_2 + 4x_3 + 2) \end{aligned}$$

Тогда:

$$\begin{aligned} t &= \frac{\text{grad } f(X) \text{grad}^T f(X)}{\text{grad } f(X) H(X) \text{grad}^T f(X)} = \\ &= \frac{(98x_1 - 10x_2 - 8x_3 - 58, -10x_1 + 70x_2 + 12x_3 + 24, -8x_1 + 12x_2 + 4x_3 + 2) \begin{pmatrix} 98x_1 - 10x_2 - 8x_3 - 58 \\ -10x_1 + 70x_2 + 12x_3 + 24 \\ -8x_1 + 12x_2 + 4x_3 + 2 \end{pmatrix}}{(98x_1 - 10x_2 - 8x_3 - 58, -10x_1 + 70x_2 + 12x_3 + 24, -8x_1 + 12x_2 + 4x_3 + 2) \begin{pmatrix} 98 & -10 & -8 \\ -10 & 70 & 12 \\ -8 & 12 & 4 \end{pmatrix} \begin{pmatrix} 98x_1 - 10x_2 - 8x_3 - 58 \\ -10x_1 + 70x_2 + 12x_3 + 24 \\ -8x_1 + 12x_2 + 4x_3 + 2 \end{pmatrix}} \end{aligned}$$

Выберем в качестве начальной точки точку $X^0 = (x_1^0, x_2^0, x_3^0) = (0, 0, 0)$, $f(X^0) = 30$.

$$t_1 = \frac{3944}{400856} = 0.0098$$

$$\begin{aligned}x_1^1 &= x_1^0 - t_1(98x_1^0 - 10x_2^0 - 8x_3^0 - 58) = 0.570659 \\x_2^1 &= x_2^0 - t_1(-10x_1^0 + 70x_2^0 + 12x_3^0 + 24) = -0.236135 \\x_3^1 &= x_3^0 - t_1(-8x_1^0 + 12x_2^0 + 4x_3^0 + 2) = -0.0196779\end{aligned}$$

$$f(X^1) = 10.5976$$

На следующей итерации выполняем те же шаги для точки $X^1 = (x_1^1, x_2^1, x_3^1) = (0.570659, -0.236135, -0.0196779)$

$$t_2 = \frac{32.5349}{127.1346} = 0.255909$$

$$\begin{aligned}x_1^2 &= x_1^1 - t_2(98x_1^1 - 10x_2^1 - 8x_3^1 - 58) = 0.4572 \\x_2^2 &= x_2^1 - t_2(-10x_1^1 + 70x_2^1 + 12x_3^1 + 24) = -0.627119 \\x_3^2 &= x_3^1 - t_2(-8x_1^1 + 12x_2^1 + 4x_3^1 + 2) = 1.38209\end{aligned}$$

$$f(X^2) = 6.4346$$

Создадим функцию в Wolfram Mathematica и при помощи неё проделаем оставшиеся итерации:

```
Clear[gradientDescent]
gradientDescent[f_Symbol, x0_List, iter_Integer, e_Real] := Module[
{
  x = x0,
  var = {x1, x2, x3},
  i = 0,
  xn = {1, 1, 1},
  t, δ1, δ2, δ3, grad,
  grad = Grad[f[var], var];
  While[Abs[f[x] - f[xn]] > e && i < iter,
    δ1 = N[Norm[x - xn] / (1 + Norm[xn])];
    δ2 = N[Abs[f[x] - f[xn]] / (1 + Abs[f[xn]])];
    δ3 = N[Norm[grad /. {x1 → x[[1]], x2 → x[[2]], x3 → x[[3]]}] / (1 + Abs[f[x])]];
    xn = x;
    t = grad.Transpose@grad / (grad.({Coefficient[grad, x1], Coefficient[grad, x2], Coefficient[grad, x3]}).Transpose@grad);
    x = N[x - t*grad /. {x1 → x[[1]], x2 → x[[2]], x3 → x[[3]]}];
    i += 1];
  {N[x], N[f[x]]}
]
```

Останавливаем итерационный процесс, когда достигли близкого к точному решению ответа.

k	x₁	x₂	x₃	f	δ₁	δ₂	δ₃
1	0.570659	-0.236135	-0.0196779	10.5976	-	-	-
2	0.457205	-0.627131	1.38211	6.43455	-	-	0.491821
3	0.675209	-0.531513	1.42642	4.01674	0.902227	0.358958	2.68613
4	0.637267	-0.479546	1.50095	3.75176	0.0936679	0.325212	1.07292
5	0.696941	-0.524312	1.56255	3.53533	0.0369414	0.0528189	0.941633
6	0.643929	-0.508272	1.62556	3.33333	0.0358347	0.0455478	1.06181
7	0.7044	-0.540226	1.68457	3.1433	0.0300755	0.0445399	0.970964
8	0.652914	-0.528995	1.74341	2.96411	0.0320224	0.0438534	1.09495
9	0.710271	-0.558237	1.79918	2.79515	0.0272	0.0432465	1.00087
10	0.661596	-0.547978	1.85462	2.63581	0.0290164	0.0426238	1.12724
100	0.795294	-0.838792	3.56331	0.013391	0.00128129	0.000845678	0.300656

Точное значение целевой функции в точке минимума равно:

$$f(X^*) = f\left(\frac{29}{36}, -\frac{31}{36}, \frac{133}{36}\right) = 0$$

$$X^* = \left(\frac{29}{36}, -\frac{31}{36}, \frac{133}{36}\right) \approx (0.805556, -0.861111, 3.69444)$$

k	x₁	x₂	x₃	f	δ₁	δ₂	δ₃
372	0.805552	-0.861103	3.6944	1.56144×10^{-9}	4.24979×10^{-7}	1.00094×10^{-10}	0.000104124