Этап №1
Методы безусловной минимизации функции многих переменных

<u>Дано</u>: $f(X) = x^2 + x \cdot y + 2y^2 + 20x + 10y + 2 \to extr$

а) Аналитически отыскать экстремум функции двух переменных
(с использованием аппарата необходимых и достаточных условий экстремума)

<u>Алгоритм решения задачи с использованием необходимых и достаточных условий</u>

необходимые условия экстремума	1. Записать градиент функции $f(X)$: $\nabla f(X) = \left(\frac{\partial f(X)}{\partial x_1}, .., \frac{\partial f(X)}{\partial x_n} \right)^T$. 2. Записать необходимые условия безусловного экстремума – составить систему алгебраических уравнений вида: $\begin{cases} \frac{\partial f(X)}{\partial x_i} = 0, \quad i = 1..n \end{cases}$
достаточные условия экстремума и необходимые условия 2-го порядка	3. Найти стационарные точки функции X^*, решив полученную систему. 4. Составить матрицу Гессе $H(X)$. 5. Вычислить матрицу Гессе в точках X^*. 6. Проверить знакоопределенность матрицы $H(X^*)$ для каждой точки: • $H(X^*) > 0$ - X^*- локальный минимум функции • $H(X^*) < 0$ - X^*- локальный максимум функции • $H(X^*) \geq 0$ - требуются дополнительная проверка на локальный минимум • $H(X^*) \leq 0$ - требуются дополнительная проверка на локальный максимум $H(X^*) <> 0$- в точке X^* нет экстремума.

<u>Исследование знакоопределенности матрицы</u> $H(X^*)$ **:**

- $H(X^*) > 0$ - критерий Сильвестра или на основании определения (все $\lambda_j > 0$)
- $H(X^*) < 0$ - критерий Сильвестра или на основании определения (все $\lambda_j < 0$)
- $H(X^*) \geq 0$ - на основании определения (все $\lambda_j \geq 0$), $H(X^*) \leq 0$ - на основании определения (все $\lambda_j \leq 0$)
- $H(X^*) <> 0$ - на основании определения (λ_j разных знаков).

Критерий Сильвестра (критерий знакоопределенности матрицы)

Матрица является положительно определенной , если все ее диагональные миноры положительны; матрица является отрицательно определенной, если ее диагональные миноры чередуют знак, начиная с **«-»**.

Решение:

Запишем градиент функции:

$$\nabla f(X) = \begin{pmatrix} 2x + y + 20 \\ x + 4y + 10 \end{pmatrix}$$

Запишем необходимые условия экстремума и вычислим координаты стационарных точек:

$$\begin{cases} 2x + y + 20 = 0 \\ x + 4y + 10 = 0 \end{cases} \underset{\substack{(1) \\ (1)-2\cdot(2)}}{\Rightarrow} \begin{cases} 2x + y + 20 = 0 \\ -7y = 0 \end{cases} \Rightarrow \begin{cases} 2x + y + 20 = 0 \\ y = 0 \end{cases} \Rightarrow \begin{cases} x = -10 \\ y = 0 \end{cases}$$

Получена стационарная точка функции $X^* = (-10, \quad 0)^T$

Составим матрицу Гессе:

$$\begin{matrix} \dfrac{\partial^2 f}{\partial x^2} = 2; & \dfrac{\partial^2 f}{\partial x \partial y} = 1; \\ \dfrac{\partial^2 f}{\partial y \partial x} = 1; & \dfrac{\partial^2 f}{\partial y^2} = 4; \end{matrix} \quad \Rightarrow \quad H(X) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$$

Вычислим матрицу Гессе в полученной стационарной точке:

$$H(X^*) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$$

Определим характер полученной стационарной точки, используя критерий Сильвестра:

$$\Delta_1 = 2 > 0$$
$$\Delta_2 = 2 \cdot 4 - 1 \cdot 1 = 7 > 0$$

Так как все диагональные миноры матрицы положительны, матрица Гессе является положительно-определенной $H(X^*) > 0$, и, следовательно, точка $X^* = (-10, \quad 0)^T$ является точкой ***локального минимума*** функции.

Ответ: функции $f(X)$ имеет локальный безусловный минимум в точке с координатами $X^* = (-10, \quad 0)$.

б) Сделать три итерации методом градиентного спуска из начальной точки $X^0 = (0, 0)^T$ ***в направлении экстремума***

<u>Внимание !</u>
Для пунктов б)-е): если при аналитическом решении задачи найден локальный максимум функций, то для численного решения задачи необходимо умножить исходную функцию на (-1) и перейти к задаче поиска минимума, при этом нужно пересчитать градиент и матрицу Гессе для новой функции. В результирующих таблицах значение функции нужно умножить на (-1).

<u>Алгоритм метода градиентного спуска</u>

$\boxed{X^{k+1} = X^k - t_k \nabla f(X^k)}$, здесь:

- $d^k = -\nabla f(X^k)$ - направление антиградиента функции;
- $t_k > 0$ - шаг <u>выбирается</u> из условия убывания функции в точках последовательности: $f(X^{k+1}) < f(X^k)$

<u>Итерация 0</u>

$$X^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$f(X^0) = 0^2 + 0 \cdot 0 + 2 \cdot 0^2 + 20 \cdot 0 + 10 \cdot 0 + 2 = 2$$

$$\nabla f(X^0) = \begin{pmatrix} 2 \cdot 0 + 0 + 20 \\ 0 + 4 \cdot 0 + 10 \end{pmatrix} = \begin{pmatrix} 20 \\ 10 \end{pmatrix}$$

$$\left\| \nabla f(X^0) \right\| = \sqrt{20^2 + 10^2} = 22.3607$$

<u>Итерация 1</u>

Вычислим точку X^1 по формуле: $X^1 = X^0 - t_0 \nabla f(X^0)$. Зададим шаг $t_0 = 0.1$

$$X^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - 0.1 \cdot \begin{pmatrix} 20 \\ 10 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$$

$$f(X^1) = (-2)^2 + (-2) \cdot (-1) + 2 \cdot (-1)^2 + 20 \cdot (-2) + 10 \cdot (-1) + 2 = -40$$

$f(X^1) < f(X^0)$, следовательно, шаг выбран удачно

$$\nabla f(X^1) = \begin{pmatrix} 2 \cdot (-2) + (-1) + 20 \\ (-2) + 4 \cdot (-1) + 10 \end{pmatrix} = \begin{pmatrix} 15 \\ 4 \end{pmatrix}$$

$$\left\| \nabla f(X^1) \right\| = \sqrt{15^2 + 4^2} = 15.52417$$

Итерация 2

Вычислим точку X^2 по формуле: $X^2 = X^1 - t_1 \nabla f(X^1)$. Зададим шаг $t_1 = 0.1$

$$X^2 = \begin{pmatrix} -2 \\ -1 \end{pmatrix} - 0.1 \cdot \begin{pmatrix} 15 \\ 4 \end{pmatrix} = \begin{pmatrix} -3.5 \\ -1.4 \end{pmatrix}$$

$$f(X^2) = (-3.5)^2 + (-3.5)\cdot(-1.4) + 2\cdot(-1.4)^2 + 20\cdot(-3.5) + 10\cdot(-1.4) + 2 = -60.93$$

$f(X^2) < f(X^1)$, следовательно, шаг выбран удачно

$$\nabla f(X^2) = \begin{pmatrix} 2\cdot(-3.5) + (-1.4) + 20 \\ (-3.5) + 4\cdot(-1.4) + 10 \end{pmatrix} = \begin{pmatrix} 11.6 \\ 0.9 \end{pmatrix}$$

$$\left\|\nabla f(X^2)\right\| = \sqrt{11.6^2 + 0.9^2} = 11.63486$$

Итерация 3

Вычислим точку X^3 по формуле: $X^3 = X^2 - t_2 \nabla f(X^2)$. Зададим шаг $t_2 = 0.1$

$$X^3 = \begin{pmatrix} -3.5 \\ -1.4 \end{pmatrix} - 0.1 \cdot \begin{pmatrix} 11.6 \\ 0.9 \end{pmatrix} = \begin{pmatrix} -4.66 \\ -1.49 \end{pmatrix}$$

$$f(X^3) = (-4.66)^2 + (-4.66)\cdot(-1.49) + 2\cdot(-1.49)^2 + 20\cdot(-4.66) + 10\cdot(-1.49) + 2 = -73.0008$$

$f(X^3) < f(X^2)$, следовательно, шаг выбран удачно

$$\nabla f(X^3) = \begin{pmatrix} 2\cdot(-4.66) + (-1.49) + 20 \\ (-4.66) + 4\cdot(-1.49) + 10 \end{pmatrix} = \begin{pmatrix} 9.19 \\ -0.62 \end{pmatrix}$$

$$\left\|\nabla f(X^3)\right\| = \sqrt{9.19^2 + (-0.62)^2} = 9.21089$$

Приведенные вычисления представим в виде таблицы

№	x	y	t	∇_x	∇_y	$\|\nabla f(X)\|$	f
0	0	0	-	20	10	22.3607	2
1	-2	-1	0.1	15	4	15.52417	-40
2	-3.5	-1.4	0.1	11.6	0.9	11.63486	-60.93
3	-4.66	-1.49	0.1	9.19	-0.62	9.21089	-73.0008

в) Сделать одну итерацию методом наискорейшего спуска из начальной точки $X^0 = (0,0)^T$ ***в направлении экстремума***

Алгоритм метода наискорейшего градиентного спуска

$\boxed{X^{k+1} = X^k - t_k \nabla f(X^k)}$, здесь:

- $d^k = -\nabla f(X^k)$ - направление антиградиента функции;
- $t_k > 0$ - вычисляется из условия наибольшего убывания функции в точках последовательности: $t_k = \arg\min[f(X^{k+1})]$

Итерация 0. Итерация 0 совпадает с 0-й итерацией метода градиентного спуска.

Итерация 1

Вычислим точку X^1 по формуле: $X^1 = X^0 - t_0 \nabla f(X^0)$.

$$X^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - t_0 \cdot \begin{pmatrix} 20 \\ 10 \end{pmatrix} = \begin{pmatrix} -20 \cdot t_0 \\ -10 \cdot t_0 \end{pmatrix}$$

Вычислим шаг t_0:

$$f(X^1) = (-20 \cdot t_0)^2 + (-20 \cdot t_0) \cdot (-10 \cdot t_0) + 2 \cdot (-10 \cdot t_0)^2 + 20 \cdot (-20 \cdot t_0) + 10 \cdot (-10 \cdot t_0) + 2 =$$

$$= 400 \cdot t_0^2 + 200 \cdot t_0^2 + 200 \cdot t_0^2 - 400 \cdot t_0 - 100 \cdot t_0 + 2 = \underline{\underline{800 \cdot t_0^2 - 500 \cdot t_0 + 2}}$$

$$\frac{df(X^1)}{dt_0} = 1600 \cdot t_0 - 500 = 0 \;\Rightarrow\; t_0 = \frac{500}{1600} = 0.3125$$

$$X^1 = \begin{pmatrix} -20 \cdot 0.3125 \\ -10 \cdot 0.3125 \end{pmatrix} = \begin{pmatrix} -6.25 \\ -3.125 \end{pmatrix}$$

$$f(X^1) = 800 \cdot 0.3125^2 - 500 \cdot 0.13125 + 2 = -76.125$$

$$\nabla f(X^1) = \begin{pmatrix} 2 \cdot (-6.25) + (-3.125) + 20 \\ (-6.25) + 4 \cdot (-3.125) + 10 \end{pmatrix} = \begin{pmatrix} 4.375 \\ -8.75 \end{pmatrix}$$

$$\left\| \nabla f(X^1) \right\| = \sqrt{4.375^2 + (-8.75)^2} = 9.7828$$

Приведенные вычисления представим в виде таблицы

№	x	y	t	∇_x	∇_y	$\|\nabla f(X)\|$	f
0	0	0	-	20	10	22.3607	10
1	-6.25	-3.125	0.3125	4.375	-8.75	9.7828	-76.125

г) Сделать две итерации методом сопряженных градиентов из начальной точки $X^0 = (0, 0)^T$ в направлении экстремума

Алгоритм метода сопряженных градиентов

$\boxed{X^{k+1} = X^k + t_k d^k}$, здесь:

- $d^0 = -\nabla f(X^0)$

  $d^k = -\nabla f(X^k) + \beta_{k-1} d^{k-1}$

  $$\beta_{k-1} = \frac{\left\|\nabla f(X^k)\right\|^2}{\left\|\nabla f(X^{k-1})\right\|^2}$$

- $t_k > 0$ - вычисляется из условия наибольшего убывания функции в точках последовательности: $t_k = \arg\min[f(X^{k+1})]$

Итерация 0. Итерация 0 совпадает с 0-й итерацией метода градиентного спуска.

Итерация 1. Итерация 1 совпадает с 1-й итерацией метода наискорейшего спуска.

Итерация 2

Вычислим точку X^2 по формулам:

$$X^2 = X^1 + t_1 d^1$$

$$d^1 = -\nabla f(X^1) + \beta_0 d^0, \quad d^0 = -\nabla f(X^0)$$

$$\beta_0 = \frac{\left\|\nabla f(X^1)\right\|^2}{\left\|\nabla f(X^0)\right\|^2}$$

$$\beta_0 = \frac{9.7828^2}{22.3607^2} = 0.19141$$

$$d^1 = -\begin{pmatrix} 4.375 \\ -8.75 \end{pmatrix} + 0.19141 \cdot \begin{pmatrix} -20 \\ -10 \end{pmatrix} = \begin{pmatrix} -8.20313 \\ 6.83594 \end{pmatrix}$$

$$X^2 = \begin{pmatrix} -6.25 \\ -3.125 \end{pmatrix} + t_1 \cdot \begin{pmatrix} -8.20313 \\ 6.83594 \end{pmatrix} = \begin{pmatrix} -6.25 - 8.20313 \cdot t_1 \\ -3.125 + 6.83594 \cdot t_1 \end{pmatrix}$$

Вычислим шаг t_1:

$$f(X^2) = (-6.25 - 8.20313 \cdot t_1)^2 + (-6.25 - 8.20313 \cdot t_1) \cdot (-3.125 + 6.83594 \cdot t_1) +$$
$$+ 2 \cdot (-3.125 + 6.83594 \cdot t_1)^2 +$$
$$+ 20 \cdot (-6.25 - 8.20313 \cdot t_1) + 10 \cdot (-3.125 + 6.83594 \cdot t_1) + 2 =$$
$$\underline{\underline{= 104.67523 \cdot t_1^2 - 95.70313 \cdot t_1 - 76.125}}$$

$$\frac{df(X^2)}{dt_1} = 209.35059 \cdot t_1 - 95.70313 = 0 \;\Rightarrow\; t_1 = \frac{95.70313}{209.35059} = 0.45714$$

$$X^2 == \begin{pmatrix} -6.25 - 8.20313 \cdot 0.45714 \\ -3.125 + 6.83594 \cdot 0.45714 \end{pmatrix} = \begin{pmatrix} -9.99997 \\ -0.00001 \end{pmatrix} \approx \begin{pmatrix} -10 \\ 0 \end{pmatrix}$$

$$f(X^2) = 104.67523 \cdot 0.45714^2 - 95.70313 \cdot 0.45714 - 76.125 = -98.00001 \approx -98$$

$$\nabla f(X^2) = \begin{pmatrix} 2 \cdot (-9.99997) + (-0.00001) + 20 \\ (-9.99997) + 4 \cdot (-0.00001) + 10 \end{pmatrix} = \begin{pmatrix} 0.00006 \\ 0.00026 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\| \nabla f(X^2) \right\| = \sqrt{0.00006^2 + 0.00026^2} = 0.00027 \approx 0$$

Т.к. $\left\| \nabla f(X^2) \right\| = 0$***, то*** X^2 ***-стационарная точка функции ! Вычисления закончены !***

Приведенные вычисления представим в виде таблицы

№							
0	x	y	t	∇_x	∇_y	$\lVert\nabla f(X)\rVert$	f
	0	0	-	20	10	22.3607	10
			β	d_x	d_y		
			-	-20	-10		
1	x	y	t	∇_x	∇_y	$\lVert\nabla f(X)\rVert$	f
	-6.25	-3.125	0.3125	4.375	-8.75	9.7828	-76.125
			β	d_x	d_y		
			0.19141	-8.20313	6.83594		
2	x	y	t	∇_x	∇_y	$\lVert\nabla f(X)\rVert$	f
	-5	0	0.45714	0	0	0	-98

д) Сделать одну итерацию методом Ньютона из начальной точки $X^0 = (0,0)^T$ в направлении экстремума

<u>Алгоритм метода Ньютона</u>

$\boxed{X^{k+1} = X^k - H^{-1}(X^k)\nabla f(X^k)}$, здесь:

- $d^k = -H^{-1}(X^k)\nabla f(X^k)$ - направление спуска.

<u>Итерация 0</u>. Итерация 0 совпадает с 0-й итерацией метода градиентного спуска.

<u>Итерация 1</u>

Вычислим точку X^1 по формуле: $X^1 = X^0 - H^{-1}(X^0)\nabla f(X^0)$

Вычислим матрицу обратную к матрице Гессе, вычисленной в точке X^0:

$$H(X^0) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \Rightarrow H^{-1}(X^0) = \frac{1}{7}\cdot\begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix} \Rightarrow H^{-1}(X^0) = \begin{pmatrix} 4/7 & -1/7 \\ -1/7 & 2/7 \end{pmatrix}$$

Тогда

$$X^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 4/7 & -1/7 \\ -1/7 & 2/7 \end{pmatrix}\cdot\begin{pmatrix} 20 \\ 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 80/7 - 10/7 \\ -20/7 + 20/7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 10 \\ 0 \end{pmatrix} = \begin{pmatrix} -10 \\ 0 \end{pmatrix}$$

$$f(X^1) = (-10)^2 + 2\cdot(-10)\cdot 0 + 2\cdot 0^2 + 20\cdot(-10) + 10\cdot 0 + 2 = -98$$

$$\nabla f(X^1) = \begin{pmatrix} 2\cdot(-10) + 0 + 20 \\ -10 + 2\cdot 0 + 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\|\nabla f(X^1)\right\| = \sqrt{0^2 + 0^2} = 0$$

Т.к. $\left\|\nabla f(X^1)\right\| = 0$, то X^1-стационарная точка функции ! Вычисления закончены !

Приведенные вычисления представим в виде таблицы

№	x	y	∇_x	∇_y	f	$\|\nabla f(X)\|$
0	0	0	20	10	10	22.3607
1	-10	0	0	0	-98	0