① $f(x) = x \cdot \ln|x|$

$f'(x) = \ln|x| + \frac{1}{x} \cdot x = \ln|x| + 1 = 0$

$\ln|x| = -1$

$x_0 = e^{-1}$

$f''(x) = \frac{1}{x} \Rightarrow f''(x_0) = e > 0 \Rightarrow x_0 = e^{-1}$ - точка минимума

② $f(x, y) = y^2 + 2xy - 4x - 2y - 3$

$\begin{cases} f(x,y)'_x = 2y - 4 \\ f(x,y)'_y = 2y + 2x - 2 \end{cases} \Rightarrow \begin{cases} 2y - 4 = 0 \\ 2y + 2x - 2 = 0 \end{cases} \Rightarrow \begin{cases} y = 2 \\ x = -1 \end{cases}$

$f(x,y)''_{xy} = 2$; $f(x,y)''_{xx} = 0$; $f(x,y)''_{yx} = f(x,y)''_{yy} = 2$;

т. е $H = \begin{pmatrix} 0 & 2 \\ 2 & 2 \end{pmatrix}$ $\Delta_1 = 0$ $\Delta_2 = -4 \Rightarrow$ экстремума в т. $(-1, 2)$ нет.

③ $f(x, y) = (1 + 2x + 3y)^{2023}$ в т. $(0, 0)$

$\nabla f = \begin{cases} 2023 \cdot 2 \cdot (1 + 2x + 3y)^{2022} \\ 2023 \cdot 3 \cdot (1 + 2x + 3y)^{2022} \end{cases}$

$\nabla f \Big|_{\substack{x=0 \\ y=0}} = \begin{pmatrix} 4046 \\ 6069 \end{pmatrix}$

Т.к до второго члена требуется разложить в ряд Тейлора, то:

$f(x, y) \approx f(0, 0) + (4046 \quad 6069) \begin{pmatrix} x - 0 \\ y - 0 \end{pmatrix} + \bar{o}(\sqrt{x^2 + y^2}) \approx$

$\approx 1 + 4046x + 6069y + \bar{o}(\sqrt{x^2 + y^2})$

④ $f(x) = ax^2 + bx + c$, где $a > 0$

$\nabla f(x) = 2ax + b$

$2ax + b = 0$

$x = -\frac{b}{2a}$ - т. минимума, т.к. $a > 0$ (парабола открыта вверх)

$x^t = x^{t-1} - \nabla f(x^{t-1}) \cdot \gamma$

$\gamma = \frac{x^{t-1} - x^t}{\nabla f(x^{t-1})}$, т. к. $x^{t-1} = x_0$, а $x^t = -\frac{b}{2a}$, то.

$$\gamma = \frac{x_0 + \frac{b}{2a}}{2a x_0 + b}$$

⑤ $w_t = w_{t-1} + (\nabla Q(w_t))^2$

1) Ошибка знака градиента:

Если он хочет минимизировать функцию $Q(w)$, то нужно идти против направления градиента. Иначе (если функция не ограничена сверху), алгоритм будет увеличивать значения $Q(w)$, что приведет к расхождению весов. Т.е исправление:

$$w_t = w_{t-1} - (\nabla Q(w_t))^2$$

2) Ошибка квадрата градиента:

Квадрат градиента игнорирует знак $\nabla Q(w_t)$. А т.к. градиентный спуск движется против направления градиента, а $(\nabla Q(w_t))^2$ всегда имеет положительное направление, то веса будут меняться произвольно, т.е исправление:

$$w_t = w_{t-1} - \nabla Q(w_t)$$

⑥ Мира Мурати, Сергей Брин, Аркадий Волож, Андрей Карпаты, Ли Кайфу.