№2 Оптимальный шаг градиентного спуска

$$Q(\vec{w}) = (X\vec{w} - \vec{y})^T (X\vec{w} - \vec{y}) \to \min_w$$

$$\vec{w}^{(k)} = \vec{w}^{(k-1)} - \eta \nabla_w Q(\vec{w}^{(k-1)})$$

$$Q(\vec{w}^{(k-1)} - \eta \nabla_w Q(\vec{w}^{(k-1)})) \to \min_w$$

$$Q(\vec{w}^{(k-1)}) = (X\vec{w}^{(k-1)} - \vec{y})^T (X\vec{w}^{(k-1)} - \vec{y})$$

$$\nabla_w Q(\vec{w}^{(k-1)}) = X^T (X\vec{w}^{(k-1)} - \vec{y}) + (X\vec{w}^{(k-1)} - \vec{y})X \overset{def}{\equiv} \vec{q}$$

$$\vec{w}^{(k)} = \vec{w}^{(k-1)} - \eta\vec{q}$$

$$Q(\vec{w}^{(k-1)} - \eta\vec{q}) \to \min_w$$

$$Q(\vec{w}^{(k-1)} - \eta\vec{q}) = \left(X(\vec{w}^{(k-1)} - \eta\vec{q}) - \vec{y}\right)^T \left(X(\vec{w}^{(k-1)} - \eta\vec{q}) - \vec{y}\right) =$$

$$= \left(\vec{w}^{(k-1)T} X^T - \eta\vec{q}^T X^T - \vec{y}^T\right)\left(X\vec{w}^{(k-1)} - \eta X\vec{q} - \vec{y}\right)$$

$$\frac{\partial Q}{\partial \eta} = -\vec{w}^{(k-1)T} X^T \cdot X\vec{q} + 2\eta\, \vec{q}^T X^T X \vec{q} - \vec{q}^T X^T X \vec{w}^{(k-1)} + \vec{q}^T X^T \vec{y} + \vec{y}^T X \vec{q} = 0$$

$$\eta = \frac{\vec{q}^T X^T X \vec{w}^{(k-1)} + \vec{w}^{(k-1)T} X^T X \vec{q} - \vec{q}^T X^T \vec{y} - \vec{y}^T X \vec{q}}{2\vec{q}^T X^T X \vec{q}} =$$

$$= \frac{\vec{q}^T X^T (X\vec{w}^{(k-1)} - \vec{y}) + (\vec{w}^{(k-1)T} X^T - \vec{y}^T) X\vec{q}}{2\vec{q}^T X^T X \vec{q}}$$