① Рассмотрим одномерную линейную регрессию без свободного параметра $y(x) = \theta \cdot x$, где $x \in \mathbb{R}$, $\theta \in \mathbb{R}$. Предполагаем, что данные получены по правилу $y_i = \theta x_i + \varepsilon_i$, где $i \in \{1, ..., n\}$, числа $x_i \in \mathbb{R}$ - неслучайны; ε_i - случайные ошибки.

Посчитаем оценку $\hat{\theta}$ методом МНК по явной формуле:

$$\hat{\theta} = (X^T \cdot X)^{-1} X^T \cdot Y = (X_1^2 + ... + X_n^2)^{-1} \cdot (X_1 y_1 + ... + X_n y_n) = \frac{\sum\limits_{k=1}^{n} X_k y_k}{\sum\limits_{k=1}^{n} X_k^2}$$

Отметим, что: $X = (x_1, ..., x_n)^T$; $Y = (y_1, ..., y_n)^T$

② Выпишем формулы итераций GD и SGD для поиска $\hat{\theta}$.

Пусть $F(\theta) = \|Y - \theta X\|^2$ - ф-я потерь. Посчитаем $\nabla F(\theta)$:

$$F(\theta) = (Y - X\theta)^T \cdot (Y - X\theta) = Y^T Y - 2Y^T X \theta + \theta^T X X \theta.$$

$$\nabla F(\theta) = -2Y^T X^T + 2X^T X \theta = 2X^T \cdot (X\theta - Y) = 2 \cdot \sum\limits_{k=1}^{n} X_k \cdot (X_k \cdot \theta - y_k). \text{ Тогда.}$$

GD: $\theta_{t+1} = \theta_t - \eta \cdot \nabla F(\theta) = \theta_t - \eta \sum\limits_{k=1}^{n} X_k (X_k \theta - y_k)$ (Коэф-т η "съел" коэф 2).

SGD: $\theta_{t+1} = \theta_t - \eta \cdot \sum\limits_{k \in I} X_k (X_k \cdot \theta - y_k)$, где $I = \{k_1, ..., k_m\}$, где $k_1, ..., k_m \sim \mathcal{U}\{1, ..., n\}$.