

ЗАДАЧА 3.

Запишем функцию ^{N1} в зависимости от θ и приравняем ее градиент к нулю.

$$F(\theta) = \|Y - X\theta\|^2 + \lambda \|\theta\|^2 =$$

$$= (Y - X\theta)^T (Y - X\theta) + \lambda \theta^T \theta$$

$$\nabla_{\theta} F(\theta) = -X^T Y + X^T X \theta + \lambda \theta = 0 \Rightarrow$$

$$\Rightarrow \boxed{\theta = (X^T X + \lambda E)^{-1} X^T Y}$$

~~или~~

^{N2}
Градиентный спуск:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} = \theta_t - \eta (-X^T Y + X^T X \theta_t + \lambda \theta_t)$$

стохастический градиентный спуск:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} = \theta_t - \eta (-X_I^T Y_I + X_I^T X \theta_t + \lambda \theta_t),$$

где X_I , Y_I — это матрицы из k случайных элементов с повторениями, индексы которых собраны в \mathcal{I} и $\{1, \dots, n\}$.