

Круглов А.И. БОС-202

1) Loss в Ridge-регрессии: $L(\theta) = \|y - X\theta\|^2 + \lambda \|\theta\|^2$ (λ - сила регуляриз.)

$$\theta_{k+1} = \text{prox}_{\lambda} (\theta_k - \eta \cdot \nabla L(\theta_k)) = \text{prox}_{\lambda}(p) \quad (\text{с ленин})$$

Проксимальный оператор для Ridge-регрессии: $] g(x, p) = \frac{1}{2}(x-p)^2 + \lambda x^2,$

$$\text{тогда } \partial g(x, p) = x - p + 2\lambda x < 0, \quad x = \frac{p}{1+2\lambda} \quad \Rightarrow$$

$$\theta_{k+1} = \frac{1}{1+2\lambda} (\theta_k - \eta \cdot \nabla L(\theta_k)), \quad \text{где } \nabla L(\theta_k) = -X^T(y - X\theta_k) + 2\lambda\theta_k$$

По формуле видно, что при увеличении λ коэффициенты θ уменьшаются, т.е.

λ - действительно "сила регуляризации"