

Круглов А. И. 605-202 SP 2

5  $X \sim N(\mu, \sigma^2)$ ,  $f(X) = (EX - 1)^2 + (DX - 1)^2 \rightarrow \min_{\mu, \sigma^2}$

Ясно, что  $EX = \mu$ ,  $DX = \sigma^2$ , подставим:  $f(X) = (\mu - 1)^2 + (\sigma^2 - 1)^2$

$$\frac{\partial f}{\partial \mu} = 2(\mu - 1), \quad \frac{\partial f}{\partial (\sigma^2)} = 2(\sigma^2 - 1) \Rightarrow \text{WAS GD: } \mu_{t+1} = \mu_t - \eta \cdot \underbrace{2(\mu_t - 1)}_{\partial f / \partial \mu}, \quad \sigma_{t+1}^2 = \sigma_t^2 - \eta \cdot \underbrace{2(\sigma_t^2 - 1)}_{\partial f / \partial (\sigma^2)},$$

где  $\eta$  — <sup>размер</sup> шаг градиентного спуска (learning rate)

Приведя обе производные к нулю, получим, что минимум  $f$  достигается при  $\mu = \sigma^2 = 1$