

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

- ▶ Повторение алгоритма градиентного спуска на каждом примере (3-5 раз)

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

- ▶ Повторение алгоритма градиентного спуска на каждом примере (3-5 раз)
- ▶ Правило момента

$$\Delta i = W_i - W_{i-1}$$

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i) + \alpha \Delta i, \quad \alpha \approx 0.1$$

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

- ▶ Повторение алгоритма градиентного спуска на каждом примере (3-5 раз)
- ▶ Правило момента

$$\Delta i = W_i - W_{i-1}$$

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \alpha \approx 0.1$$

- ▶ Стимуляция нейронов

$$W_{i+1} = (1 - \gamma)W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \gamma \approx 10^{-4}$$

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

- ▶ Повторение алгоритма градиентного спуска на каждом примере (3-5 раз)
- ▶ Правило момента

$$\Delta i = W_i - W_{i-1}$$

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \alpha \approx 0.1$$

- ▶ Стимуляция нейронов

$$W_{i+1} = (1 - \gamma)W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \gamma \approx 10^{-4}$$

- ▶ Адаптивный выбор ε ?

Ускорение обучения

- ▶ Правильный выбор коэффициента

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i)$$

(начать с $\varepsilon = 1$)

- ▶ Повторение алгоритма градиентного спуска на каждом примере (3-5 раз)
- ▶ Правило момента

$$\Delta i = W_i - W_{i-1}$$

$$W_{i+1} = W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \alpha \approx 0.1$$

- ▶ Стимуляция нейронов

$$W_{i+1} = (1 - \gamma)W_i - \varepsilon \nabla f(W_i) + \alpha \Delta_i, \quad \gamma \approx 10^{-4}$$

- ▶ Адаптивный выбор ε ?
- ▶ Сопряженные градиенты, обучение на гессиане??