## 随机梯度下降法 Stochastic gradient descent

批梯度下降方法:

$$J_{train}( heta) = rac{1}{2m} \sum_{i+1}^m (h_ heta(x^{(i)}) - y^{(i)})^2 \ Repeat\{ \ // ext{for every j=1, ..., n} \ heta_j := heta_j - lpha rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) x_j^{(i)} \ \}$$

## 随机梯度下降法

- 1. 初始化打乱训练样本
- 2. 代码

$$egin{aligned} Repeat \{ & ext{for i} := 1,..., m \{ \ // ext{for every j} = 0,..., n \ & heta_j := heta_j - lpha(h_{ heta}(x^{(i)} - y^{(i)})x_j^{(i)}) \ & ext{} \} \} \end{aligned}$$

• 批梯度下降: 使用所有m个样本在每次迭代中

• 随机梯度下降:使用1个样本在每次迭代中

• 微型梯度下降:使用b个样本在每次迭代中

## Mini-Batch Gradient Descent 微型梯度下降

假设b=10, m-1000

\$\$

\$\$

$$Repeat\{ \ // ext{for every j=1, 11,21,31,..., 991} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha rac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - lpha \frac{1}{10} \sum_i^{i+9} (h_{ heta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \ heta_j := heta_j - heta_j + h$$