# 随机梯度下降法 Stochastic gradient descent

批梯度下降方法：

$$J_{train}(\theta)=\frac{1}{2m}\sum_{i+1}^{m}(h_\theta(x^{(i)})-y^{(i)})^2$$

$$Repeat\{$$

$$//\text{for every j=1, ..., n}$$

$$\theta_j:=\theta_j-\alpha\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)})-y^{(i)})x_j^{(i)}$$

$$\}$$

随机梯度下降法

1. 初始化打乱训练样本
2. 代码

$$Repeat\{$$

$$\text{for i := 1, ..., m}\{$$

$$//\text{for every j=0, ..., n}$$

$$\theta_j:=\theta_j-\alpha(h_\theta(x^{(i)}-y^{(i)})x_j^{(i)})$$

$$\}\}$$

- 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

$$
Repeat\{ \\
\text{//for every j=1, 11,21,31,..., 991}\\
\theta_j:=\theta_j-\alpha\frac{1}{10}\sum_{i}^{i+9}( h_\theta(x^{(i)})-y^{(i)})x_j^{(i)}\\
\}
$$

$$

$$Repeat\{$$

$$//\text{for every j=1, 11,21,31,..., 991}$$

$$\theta_j:=\theta_j-\alpha\frac{1}{10}\sum_{i}^{i+9}(h_\theta(x^{(i)})-y^{(i)})x_j^{(i)}$$

$$\}$$

$$Repeat\{$$

$$//\text{for every j=1, 11,21,31,..., 991}$$

$$\theta_j:=\theta_j-\alpha\frac{1}{10}\sum_{i}^{i+9}(h_\theta(x^{(i)})-y^{(i)})x_j^{(i)}$$

$$\}$$