Mini-batch gradient descent

- \rightarrow Batch gradient descent: Use all m examples in each iteration
- Stochastic gradient descent: Use 1 example in each iteration

Mini-batch gradient descent: Use b examples in each iteration

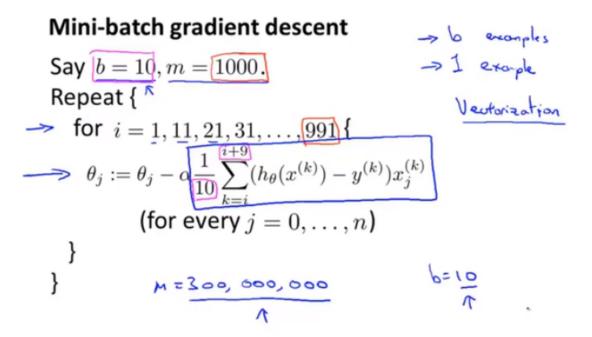
ni-batch gradient descent: Use b examples in each iteration
$$b = \text{mini-batch size}. \qquad b = 10.$$
Gret
$$b = 10 \text{ examples } (x^{(i)}, y^{(i)}), \dots (x^{(i+q)}, y^{(i+q)})$$

$$\Rightarrow 0_j := 0_j - d = 10$$

$$k = i + 10$$

$$i := i + 10$$

相当于每次选取b个数据,而不是所有数据,进行加和,梯度下降处理



优点: 便于实现向量化, 并行计算

缺点: 多了一个参数b需要确认