

# 梯度下降法

---

$$f(x_{i+1}) = f(x_i) - f'(x_i) \Delta x$$

对应的，最小二乘的拟合实际上就是对误差函数做最小化，这时候自变量是需要拟合的参数

## 批梯度下降法(BGD)

---

[知乎解答](#)

就是在计算梯度的时候使用整个数据集的数据来计算(注意需要平均), 这样得到的梯度能精确的反应整个数据集的变化方向, 缺点是数据集很大的时候计算很慢

$$f_{\text{new}} = f_{\text{old}} - h \frac{1}{n} \sum_{i=1}^n f'_i$$

## 小批量梯度下降法

---