**梯度上升法**

**方向导数**

方向导数是指任意方向的变化率，在高等数学中，我们常对x或y求偏导，来获得在x或y轴上的变化率。根据导数定义

可以类似地写出方向导数的定义为

式中

偏导数和导数都能够看成是方向导数的一个特例，其思想都是用一小段函数的微小变化率来定义导数（偏导数）。

**梯度**

梯度的方向是方向导数最大的方向，也就是函数值变化得最快，最剧烈的方向。对各自变量进行求偏导，得出来的偏导数组成一个向量，该向量就是原函数的梯度，记作或。其中，为哈密顿算子。

例如：zr的偏导为

其梯度为。

梯度的意义在于它是函数值变化最快的方向，沿着梯度方向前进时，函数值增加得最快，能够最快地寻找到最大值。而沿着梯度反方向前进，函数值减小得最快，最容易获得最小值。梯度与方向导数的关系

* 梯度的方向是最大方向导数的方向
* 梯度的模是方向导数的最大值

**梯度下降与梯度上升**

梯度下降是指沿着梯度反方向进行迭代，可以用于求解目标函数的最小值。而相反的梯度上升法，则是沿梯度方向进行迭代，其可以用于求目标函数的最大值。

**梯度下降算法**

**基础概念**

* 步长：即算法在迭代过程中x变化的大小，也即沿梯度负方向前进的长度。
* 特征：指自变量，如就为
* 假设函数：根据先验条件，假设出来的函数。一般都是将特征用一个多项式来表示。如的假设函数就为
* 损失函数：该函数用于评估结果的好坏，损失函数值越小，则证明拟合效果越好，估计参数与最优参数的差值越小，反之则越差。常采用最小二乘法作为损失函数，有

**算法步骤**

1. 确定假设函数和损失函数

令假设函数为，则对应的损失函数为

1. 初始化，误差和步长
2. 计算当前损失函数的梯度
3. 计算梯度下降距离，即用步长乘以梯度
4. 若对于所有的，下降距离都小于误差，则算法停止，否则迭代继续
5. 更新，对应的迭代公式为

要注意的是，梯度下降算法在迭代的时候，对各参数的迭代并不是同步的。而是对一个参数求偏导后判断终止条件，若符合则终止迭代，否则继续迭代。也就是说会出现有些参数只需要少量迭代就能获得最优解，而有些参数需要多次迭代才能获得最优解。

**矩阵描述**

记，则假设函数可表示为

记则损失函数就可表示为