

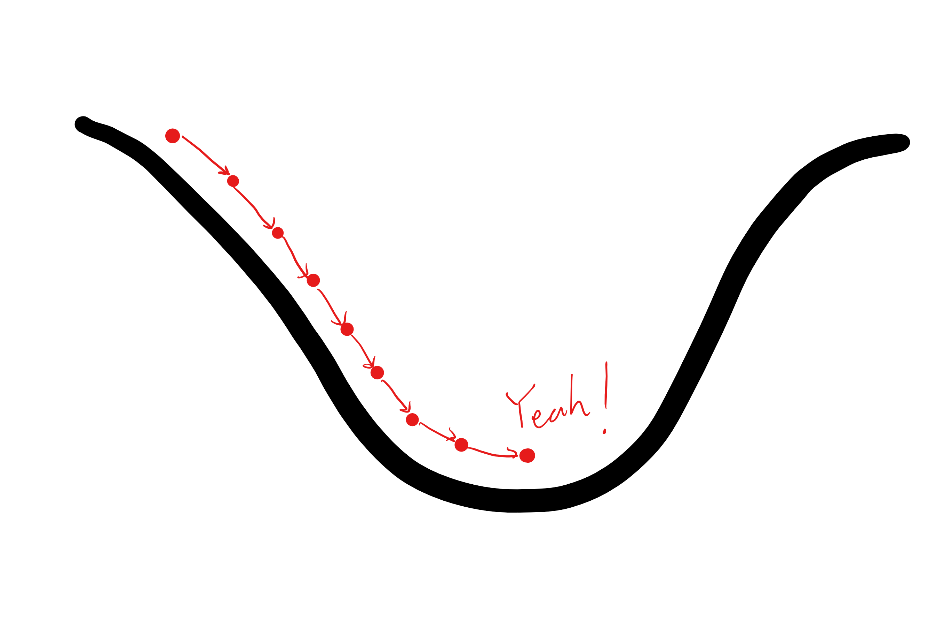
**1.微积分**

**1-1.极限，导数，微分**

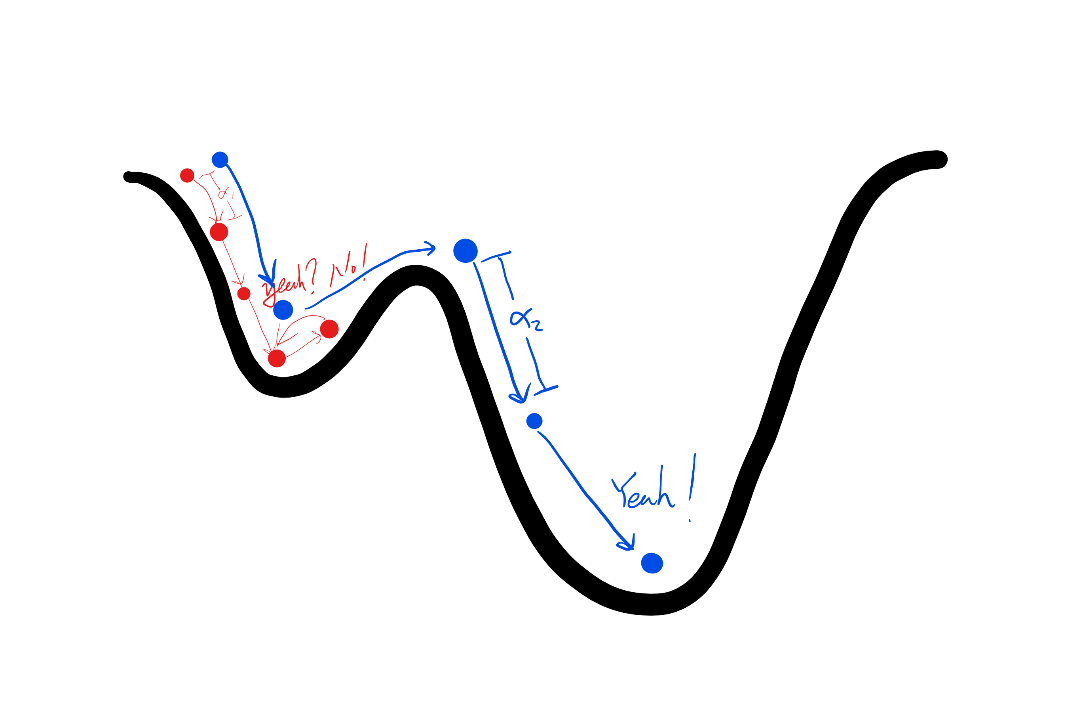
**1-1-1.梯度下降**

梯度类似于微分，微分是针对单变量，而梯度是针对向量。

因为直接获得最优解的计算量过大，所以通过均方误差的分析，采用迭代的方式来一步一步获得局部最优解，这就是梯度下降法。

梯度下降法通过电脑擅长的方式——迭代，极大的减少了求梯度的计算量，但是对于一些较为复杂的函数而言，梯度下降法获得的更多是局部最优解而不是全局最优解。

其中α是学习率，由右图我们可以看出，蓝方获得了更优的解，而红方停在了局部最优解。

由此可见，使局部最优解接近全局最优解的方式便是调节学习率。

**红点和蓝点实际上就代表了**[**反向传播算法**](https://link.jianshu.com/?t=https%3A%2F%2Fen.wikipedia.org%2Fwiki%2FBackpropagation)**，下山的路径其实就代表着算法中一直在寻找的参数最优解，山上当前点的最陡峭的方向实际上就是代价函数在这一点的梯度方向，场景中观测最陡峭方向所用的工具就是**[**微分**](https://link.jianshu.com/?t=https%3A%2F%2Fen.wikipedia.org%2Fwiki%2FDerivative)**。在下一步要走多长就是由我们算法中的学习率α所定义。**

梯度下降法本身具有弹性，可以容忍不完善的数据，如果我们不能完美地描述函数，或者我们意外的走错了一步，都不至于前功尽弃。

其中误差计算通常为均方误差，即（目标值 – 实际值）的平方，多采用均方误差的原因有以下几点：

1. 使用误差的平方，我们可以很容易使用代数计算出梯度下降的斜率。
2. 误差函数平滑且连续，这使梯度下降法很好的发挥了作用——没有发生间断，也没有发生跳跃。
3. 越接近最小值，梯度就越小，这意味着，如果我们使用这个函数调节步长，超调的风险就会变得较小。

**1-1-2.BP神经网络**

从结构上讲，BP网络具有输入层、隐藏层和输出层；从本质上讲，BP算法就是以均方误差为目标函数、采用梯度下降法来计算目标函数的最小值。

设第i层和第j层之间的权重为w(i,j)，对当前层权重反向传播修正权重的公式为：

**Δw(i,j) = α ····· 误差 · output(j) · (1 - output(j)) · output(i)**

通过反向传播修正链上的权重来改善输出情况。

**1-1-3.最大似然估计的求解**

**向量微积分**

**泰勒展开式**

xgboost

**离散型和连续型变量**

**2.线性代数**

**2-1矩阵论**

**2-1-1线性回归**

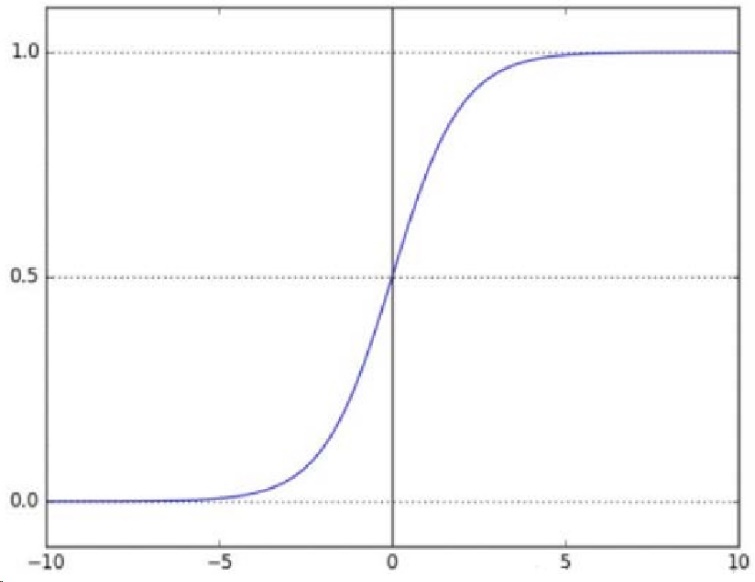
线性回归是利用数理统计中回归分析，来确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法，运用十分广泛。表达形式为y = wx+b，w为误差服从均值为0的正态分布。

回归分析中，只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示，这种回归分析称为一元线性回归分析。如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间是线性关系，则称为多元线性回归分析。

线性回归模型通常由对数据用最小二乘法来拟合。

**2-1-2逻辑回归**

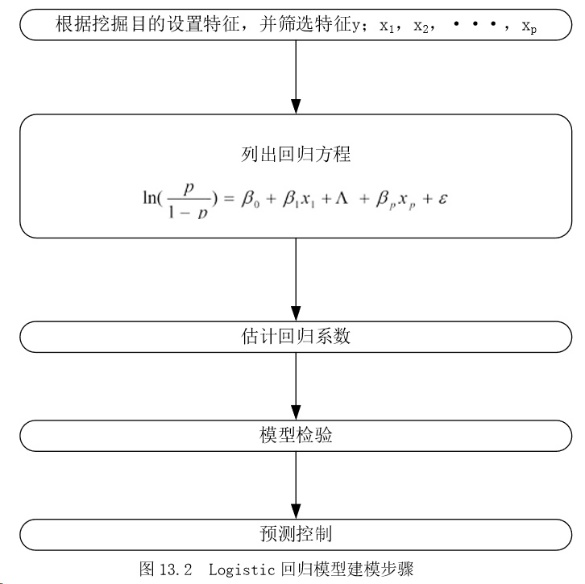
对数几率回归又称逻辑回归（logistic regression），将结果限定于一定范围之间进行输出，例如sigmoid函数，softmax函数等。



逻辑回归通常用于二分类问题，当遇到多分类问题时常用softmax函数。

逻辑回归函数往往会使数值相对小的数据对模型的影响变得更小，数值相对大的数据对模型的影响变得更大，且逻辑回归最终得到的是一个几率，而这更加适合非固定模式的机器学习对象。

逻辑回归建模过程如下：



基于树的随机森林

提升算法

支持向量机

神经网络

**特征值和特征向量**

主成分分析法PCA

谱聚类

奇异值分解SVD

**范数和距离**

正则化

**应用领域**

数据集模型拟合

图像处理

分类编码

潜在语义分析

推荐系统

深度学习

**3.信息论**

**信息熵**

贝叶斯学习

**相对熵**

ID3算法 GAN模型

**交叉熵**

LR分类 SVM多分类 逻辑回归 神经网络

**距离常量**

聚类算法 推荐系统

**4.动力学分析**

微分方程

非线性动力系统

**5.概率**

**基本类型**

先验概率 后验概率 条件概率-深度学习 全概率-贝叶斯公式 联合概率 似然概率

**概率分布**

离散概率分布

0-1分布几何分布 二项分布 泊松分布 超几何分布

连续概率分布

均匀分布 高斯分布-混合高斯模型GMM 指数分布

**判别模型与生成模型**

判别模型

线性回归 逻辑回归 SVM 条件随机场 神经网络 最进邻算法

生成模型

朴素贝叶斯模型 隐式马尔可夫模型 高斯混合模型 有限波兹曼机

**决策函数和条件概率**

决策函数

线性回归 支持向量机 神经网络

条件概率

贝叶斯分类

**最大似然估计**

算法调参和求参

**贝叶斯定理**

朴素贝叶斯算法

**图论**

神经网络中的搜索

**5-8树论**

**5-8-1决策树**

决策树（decision tree）是常见的机器学习方法，顾名思义，决策树是基于树的结构来进行决策的。

决策树学习的目的是为了产生一棵泛化能力强的决策树，基本流程遵循“分而治之”。我们希望决策树的分支结点所包含的样本尽可能属于同一类别，即节点的纯度高，而衡量纯度的指标就是信息熵，信息熵越小，纯度越高。

随机森林 深度森林

**6.统计**

**统计和假设检验**

线性回归 逻辑回归

**均值**

多元高斯分布

**方差**

主成分分析法