机器学习的数学基础

1. 基本概念
   1. 梯度下降（正则项防止过拟合）

1.1.1 Gradient descent algorithm，

目标：已知{x,y}求最适合的系数组

建立cost方程 ,使得

基本步骤：

1. 有初始值
2. 持续更新
3. 达到J函数最小值

实现方法：

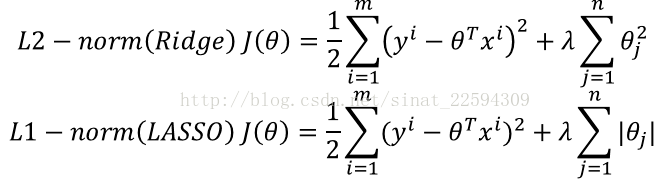
设定函数J，学习率α，使得如下计算后收敛

重复计算

1.1.2 相关概念

稀疏化模型：选择特征 ，L1范数

规则化范数：L2范数防止过拟合，L1范数选择特征



回归问题总结：<https://blog.csdn.net/gumpeng/article/details/51191376>

多变量梯度下降：<https://blog.csdn.net/pipisorry/article/details/43529845>

* 1. BP神经网络

Backpropagation algorithm

激活函数 sigmoid function

先正向传播

反向误差（输入层不需要误差，所以没有这项）至收敛

梯度检验:?

随机初始化参数：？

相关概念：<https://blog.csdn.net/u012052268/article/details/78705097?utm_source=blogxgwz0>

https://blog.csdn.net/weifenglin1997/article/details/77609598?utm\_source=blogxg

* 1. 最大似然估计的求解

the Principle of Maximum Likelihood （MLL）

最小二乘法回归与高斯分布的最大似然函数估计等价

* + 1. 实现步骤：

1. 写出似然函数
2. 取对数并求导为零，得到似然方程
3. 解似然方程
   * 1. EM算法
     2. jensen 不等式

凸函数f(x)下

* + 1. Expectation Maximization

隐含变量，列对数似然函数

For

算期望

解极大似然函数

end

* + 1. 相关概念

<https://blog.csdn.net/MosBest/article/details/52163246?utm_source=blogxgwz9>

<https://www.cnblogs.com/jerrylead/archive/2011/04/06/2006936.html>

<https://www.cnblogs.com/jerrylead/archive/2011/04/06/2006936.html>