

# 2024回忆版

2025年1月13日 16:46

Recorded by Tang Hoilap

## 1. 线性回归

- 损失函数求 $\omega$ 求偏导；解释损失函数的意义
- 参数更新式子；学习率偏大/偏小的后果
- 正则化项的作用

## 2. SVM

- 给出一组数据点和 $\omega, b$ ，写出超平面的表达式
- 求出所有数据点到超平面的距离，判断哪一个点是支持向量
- 判断移动某一个点后超平面方程是否会改变（不会，因为该点移动后的位置到超平面距离不变）
- 和线性核函数相比，高斯核函数的功能

## 3. 集成学习-Boosting

题干：有三个一维数据样本点（ $x_1=1, x_2=3, x_3=5$ ），对其进行boosting分类，使用桩分类器。//不用担心，题目有桩分类器的定义

- 考点：boosting的算法流程：求出分类边界；给错误分类的数据点加权
- Boosting能否完全正确分类测试样本？如果可以，计算迭代次数
- 在迭代过程中，损失是否会单调递减？

## 4. PCA

- 主成分的含义
- PCA两种理解角度
- 从最小化重建误差的角度求解主成分

## 5. EM

- 写出EM步方程
- 写出GMM模型中的Q表达式
- 已知模型参数 $\theta$ ，样本点 $x$ ，如何估计出所属的类 $z$
- 如果Q无法写出闭合表达式，写出两种解决办法

## 6. 推荐系统-协同滤波

- 用户-用户协同滤波的工作原理；求相似度的方法
- 如何解决冷启动问题
- 评价模型的指标