======2020/07/13==============================================

1. 机器学习的分类
2. 机器学习的主要分类
3. 无监督学习
4. 无监督学习应用
5. 监督学习
6. 监督学习应用

======2020/07/13=======机器学习入门===============================

1. 机器学习分类
2. 有监督学习
3. 无监督学习
4. 强化学习
5. 无监督学习（只有输入的数据，没有对应的结果）

样本点的分组和聚类

应用：谷歌新闻

1. 机器学习
2. 监督学习（既有输入，也有输出）

包括 分类（离散取值） 和 回归（连续取值）

相似度学习

回归：线性回归

分类：K-NN，朴素贝叶斯，逻辑回归，SVM，决策树

1. 无监督

聚类：K-Means， 降维：SVD（奇异值分解），PCA（主成分分析）

1. 强化学习

遗传算法

1. 深度学习

CNN、RNN、生成对抗网络、感知机、DNN

1. 集成学习

学习器（类似混合方式），随机森林

1. 监督学习
2. 三要素：模型、策略、算法

模型（model）：总结数据的内在规律

策略（strategy）：选取最优模型的评价准则

算法（algorithm）：选取最优模型的具体方法

1. 监督学习实现步骤
2. 监督学习过程实力
3. 数据 -处理-> 初始模型 -训练-> 成熟模型 -预测-> 预测结果
4. 模型评估策略
5. 模型评估

1.1 训练集和测试集

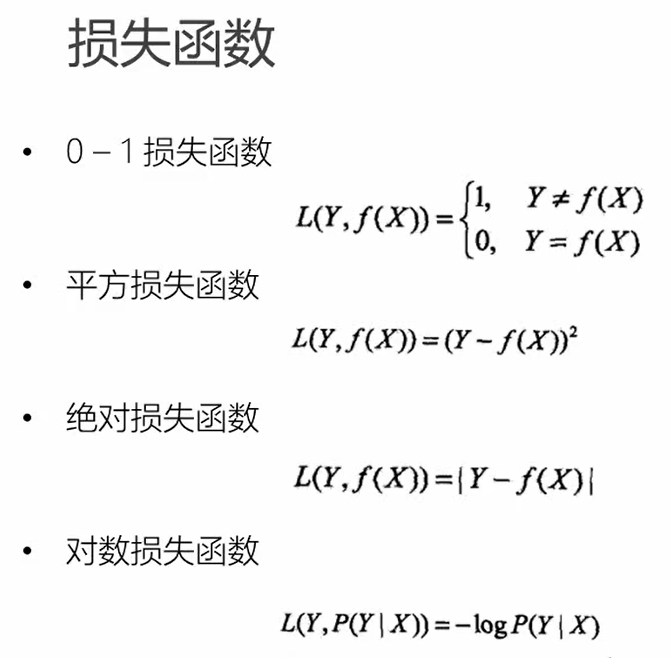
1.2 损失函数（误差）：用来衡量模型预测误差的大小，依据模型变化，损失函数是模型里面系数的函数。损失函数越小，效果越好

0-1损失函数

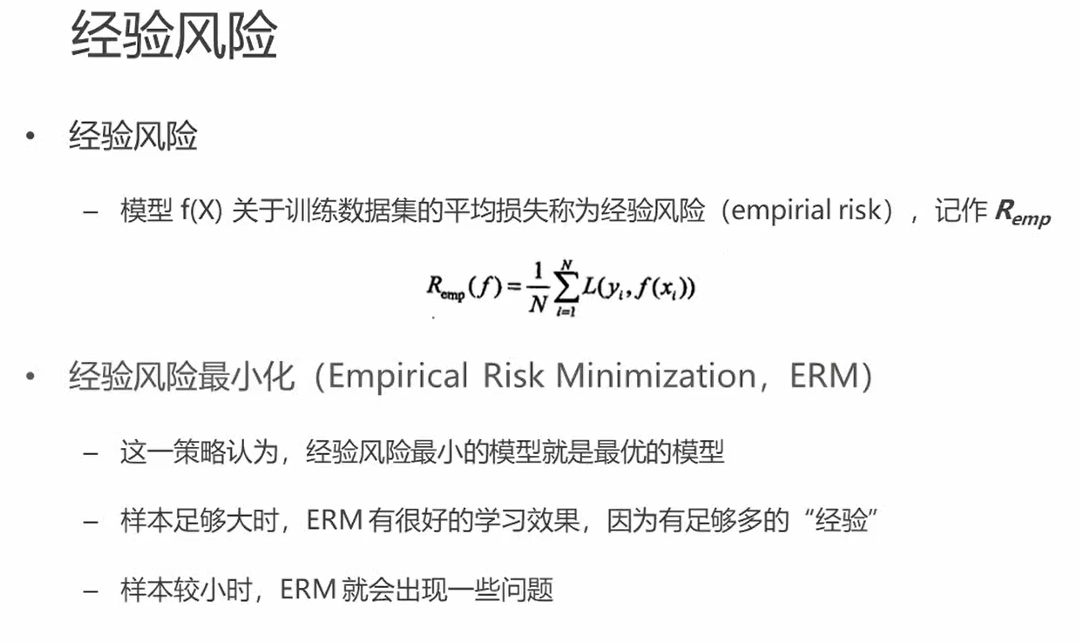
平方损失函数

绝对损失函数

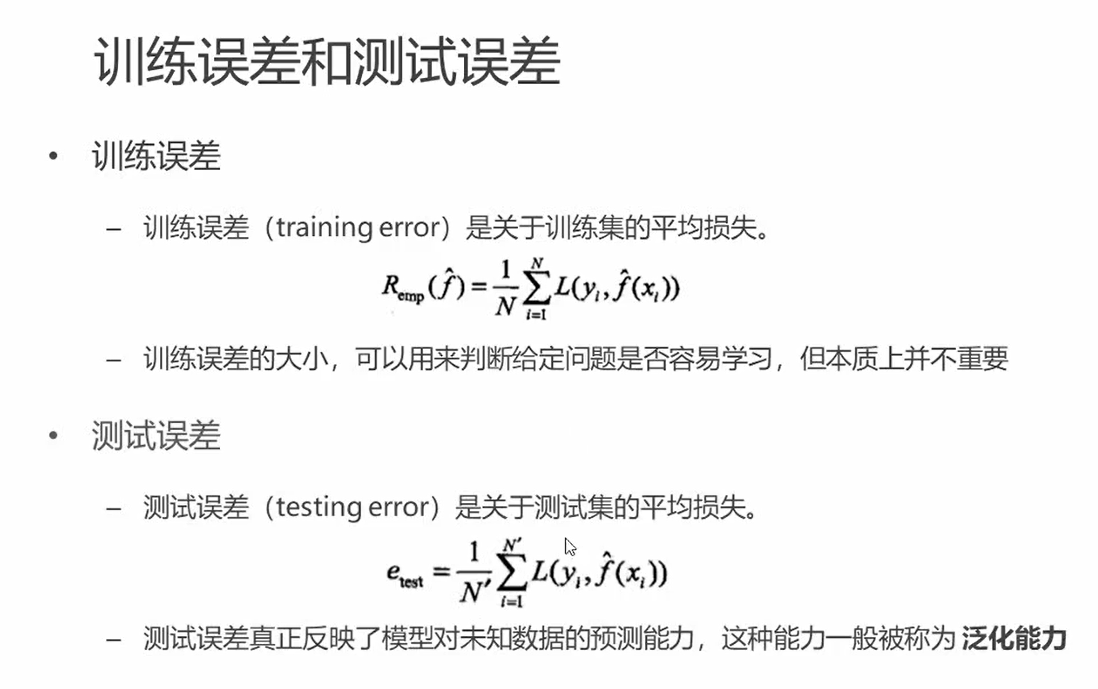
对数损失函数



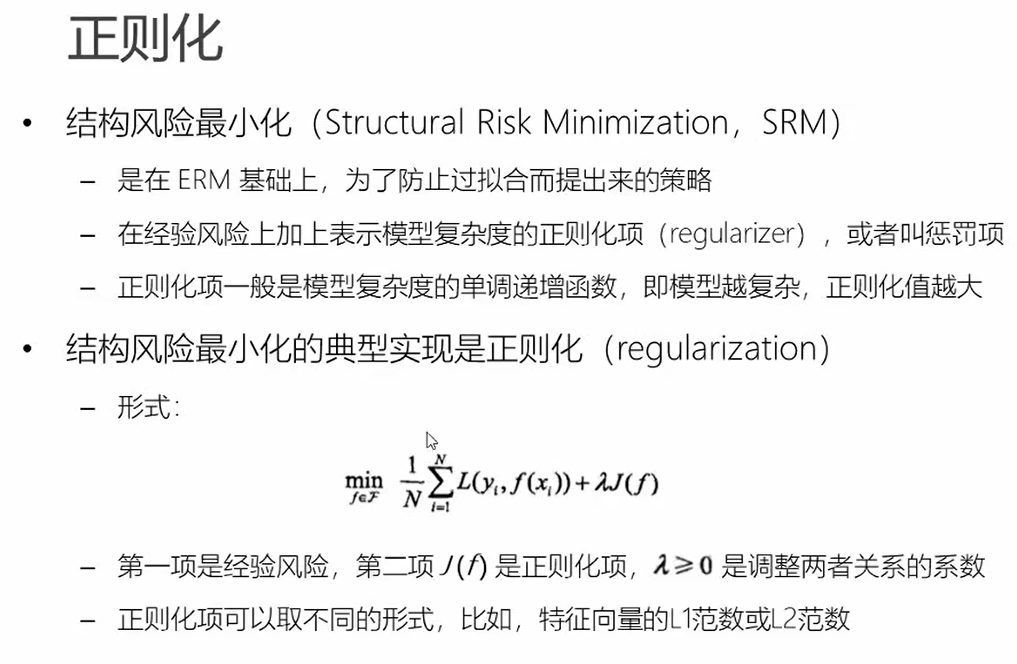
* 1. 经验风险，关于训练集的平均损失



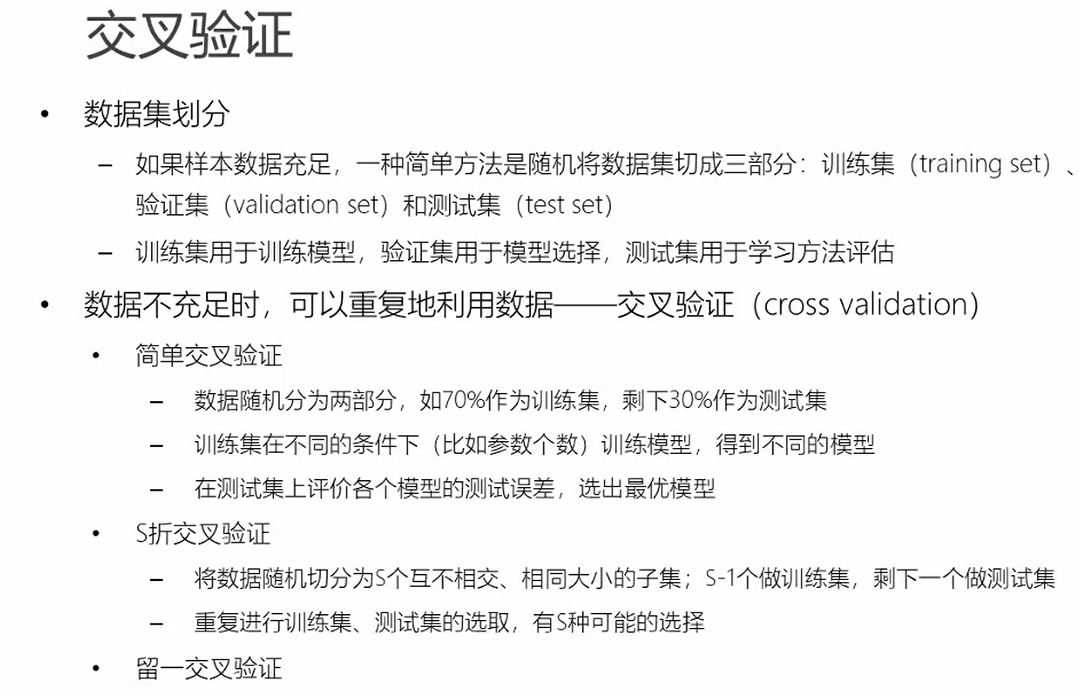
* 1. 训练误差和测试误差（泛化能力 ）



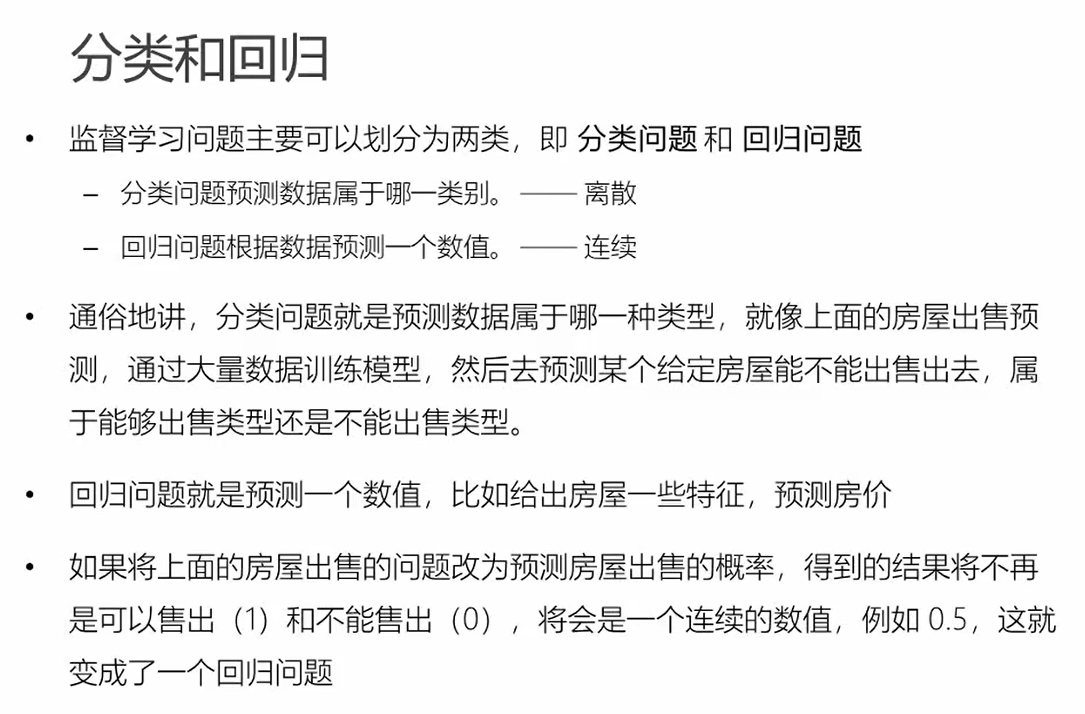
* 1. 过拟合和欠拟合
  2. 正则化



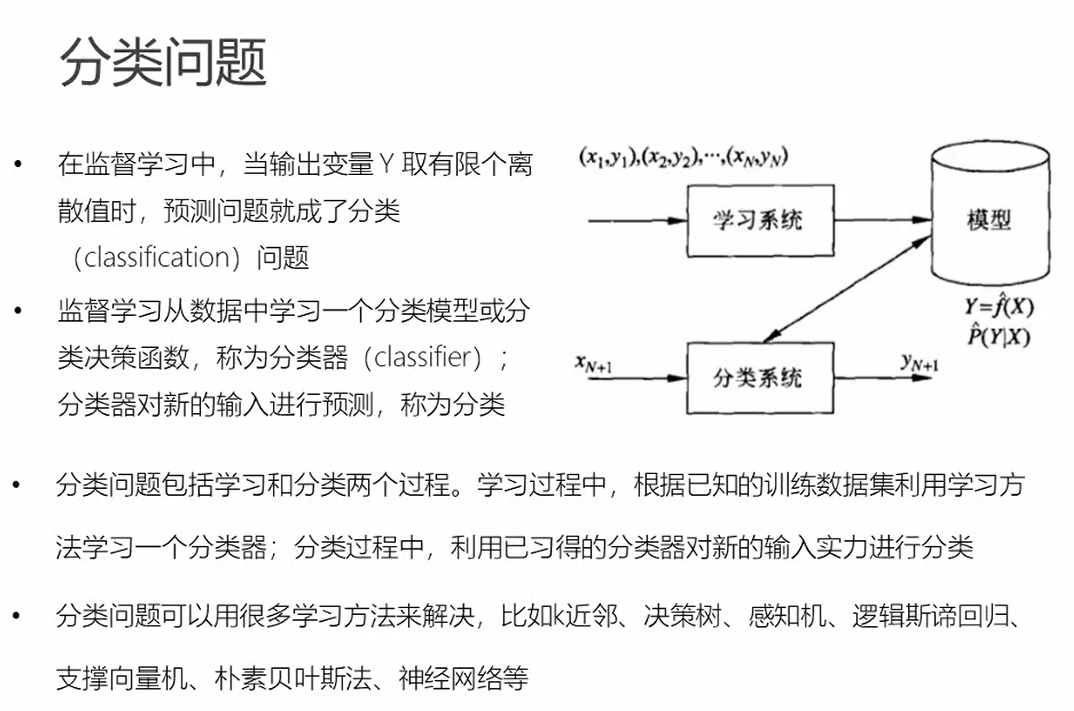
* 1. 交叉验证



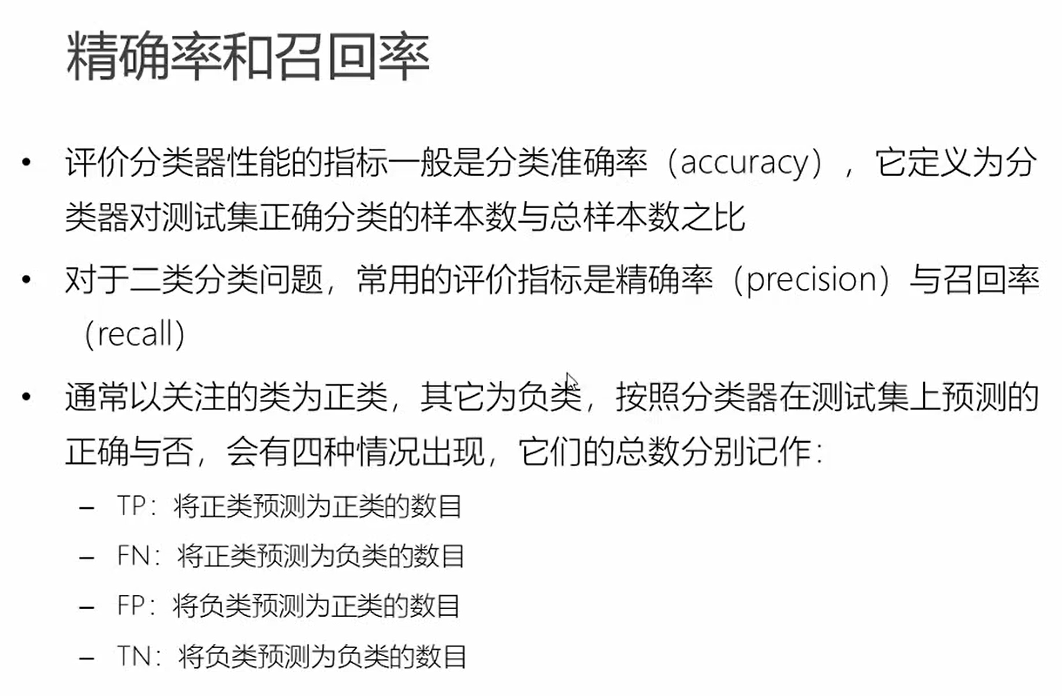
1. 监督学习的分类 ---- 分类和回归
2. 区别：最后输出结果Y离散或连续，离散为分类，连续为回归



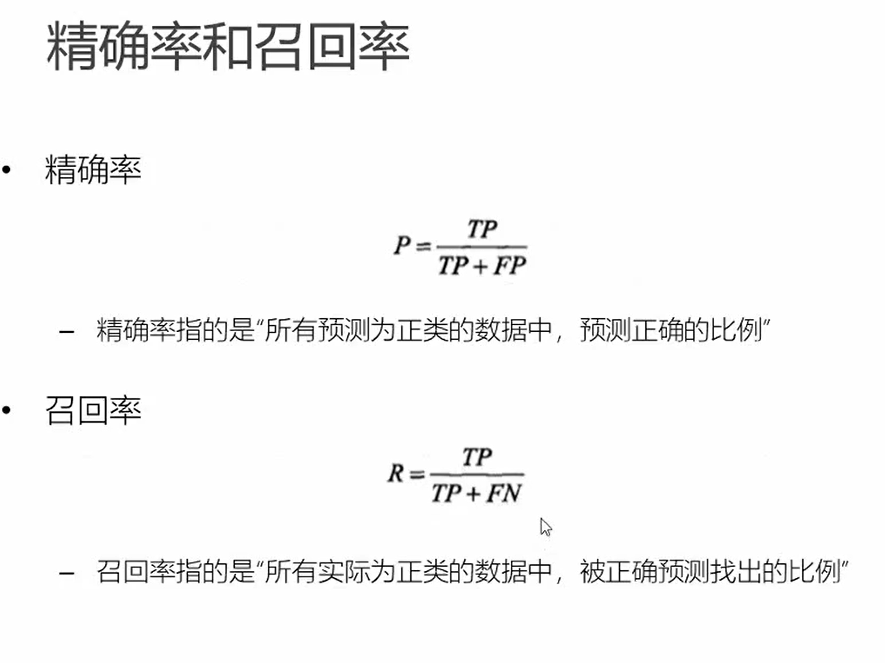
1. 分类：用分类器对新输入的数据进行预测



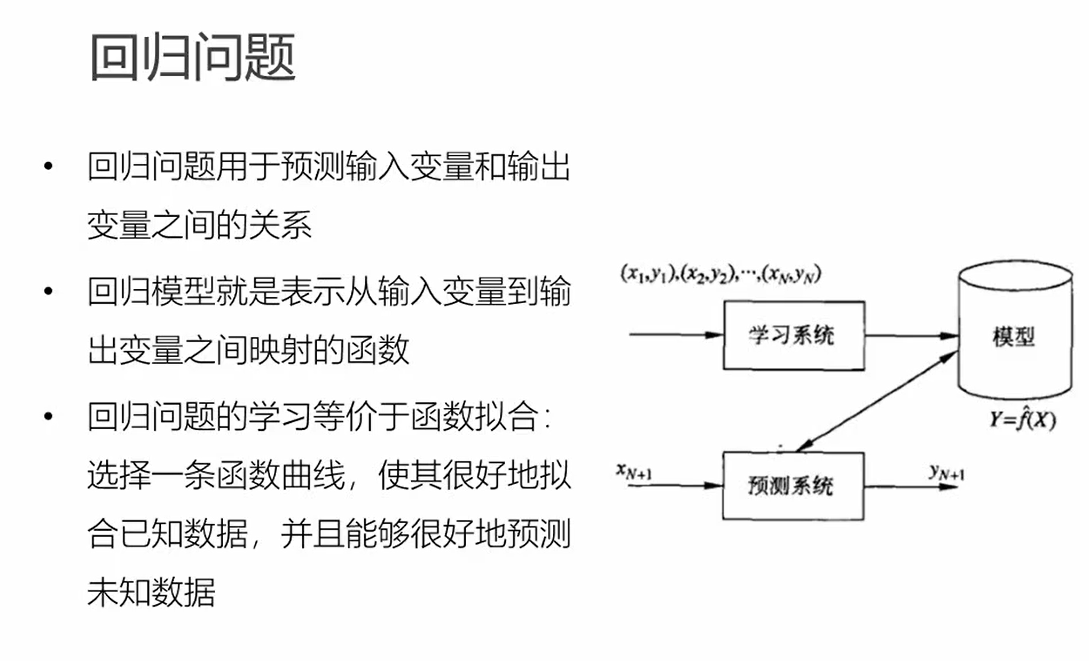
1. 评价指标 ---- 精确率和召回率



2x2的矩阵 ---- 混淆矩阵



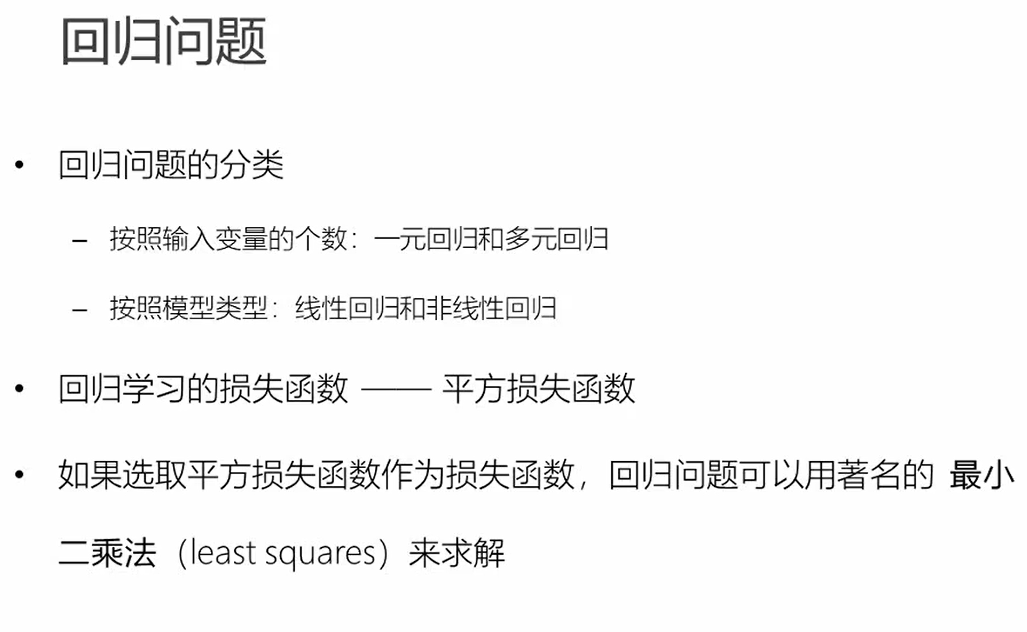
1. 回归



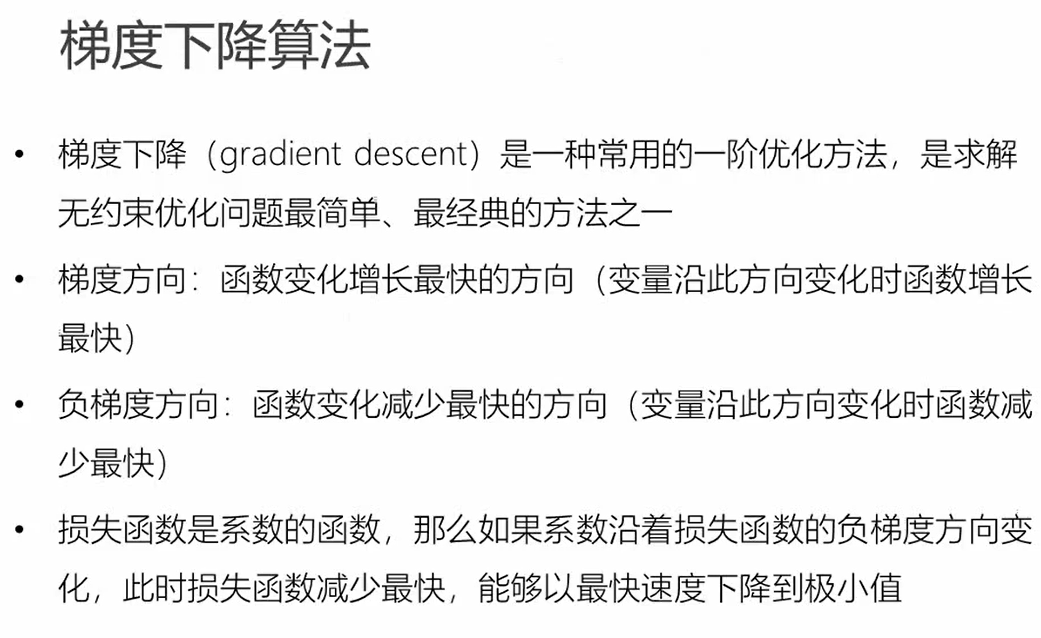
回归问题的分类

按照输入变量的个数

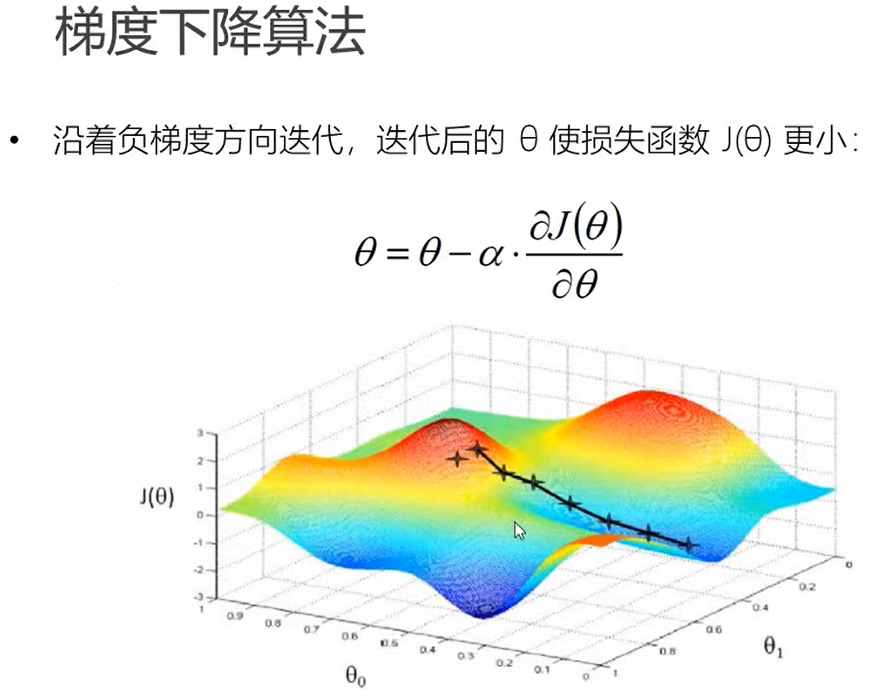
按照模型类型



1. 模型求解算法（学习算法）
   1. 梯度下降算法
   2. 牛顿法和拟牛顿法



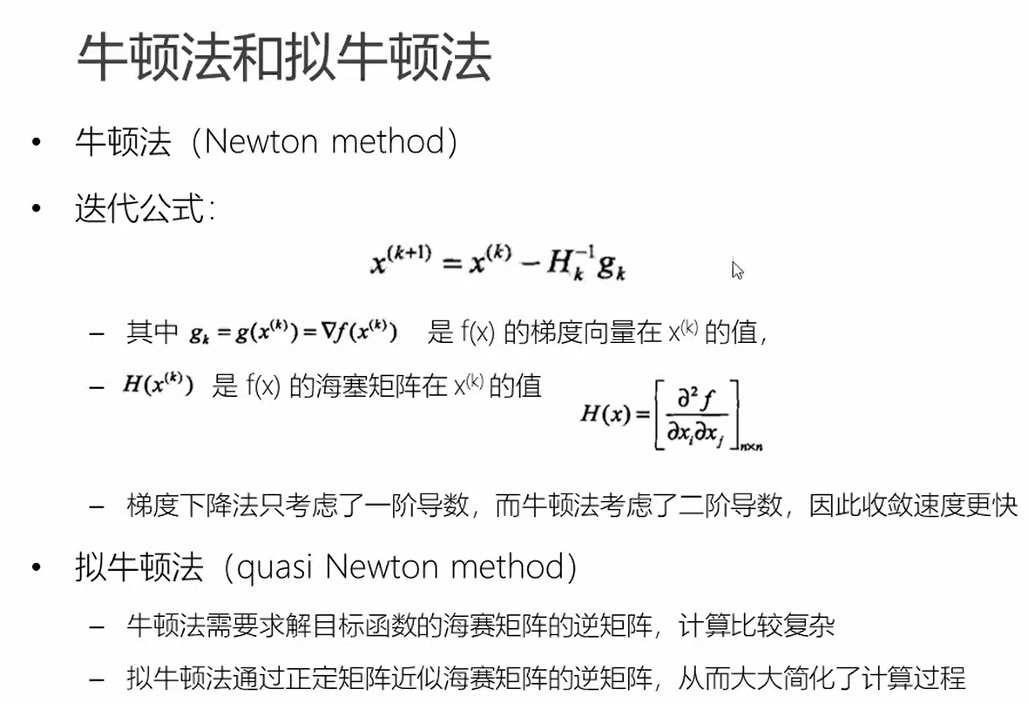
梯度是变化率最大的方向



下凸函数，可以求导梯度下降最小值

局部最优，全局最优

6.牛顿法和拟牛顿法



======2020/07/13====Python简介==================================

1. Python：解释性语言，交互性语言，面向对象
2. 环境Python3 Anaconda

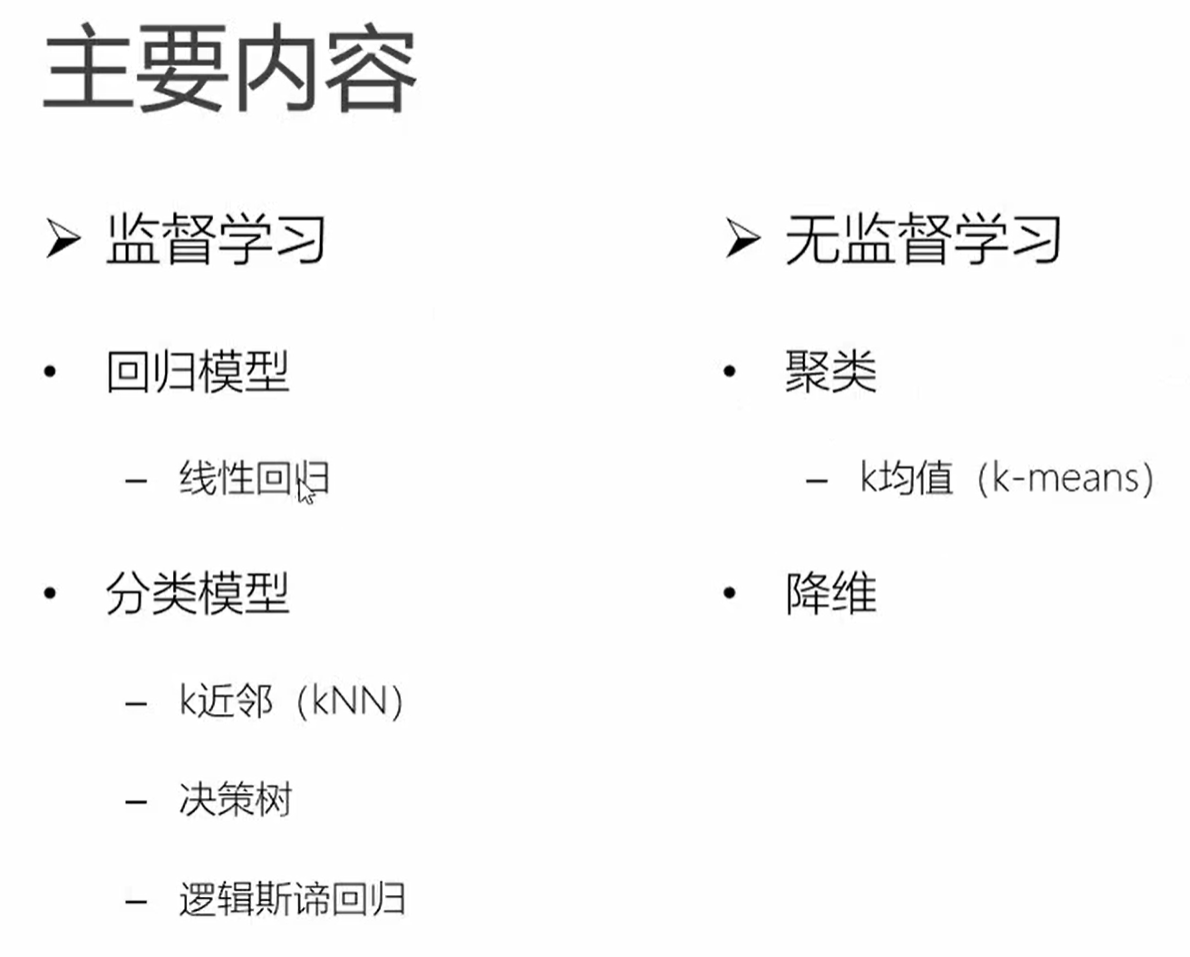
======2020/07/14====Python简介==================================

1. Python基础
   1. 严格缩进
   2. numpy数值计算；matplotlib画图；pandas数据分析，数学计算；sklearn机器学习；
   3. 编码；标识符，开头为字母或下划线；注释，单行#，多行”””或’’’；关键字和保留字；Python的引号；空行；print输出，默认换行；
   4. 基本数据类型

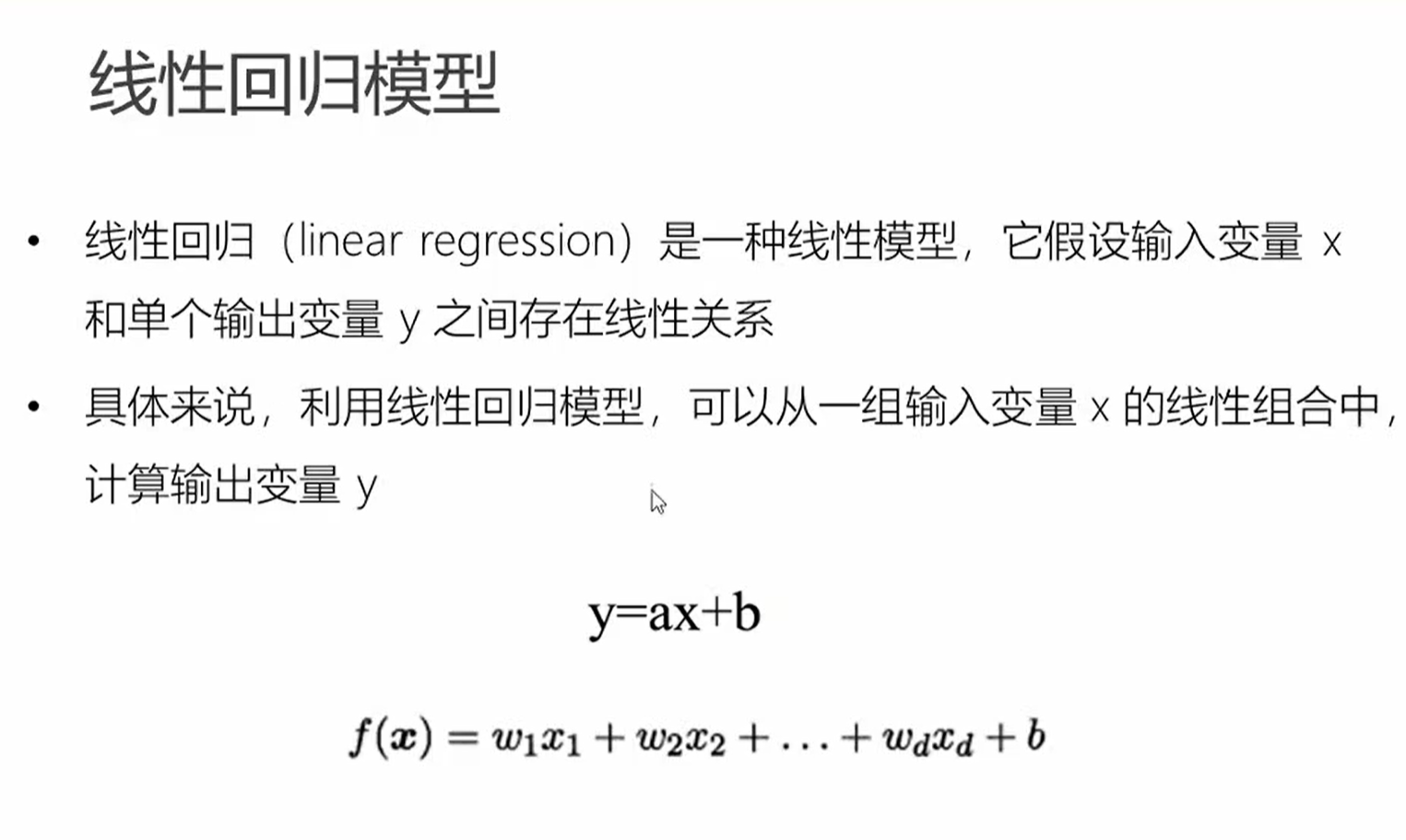
赋值；标准数据类型；

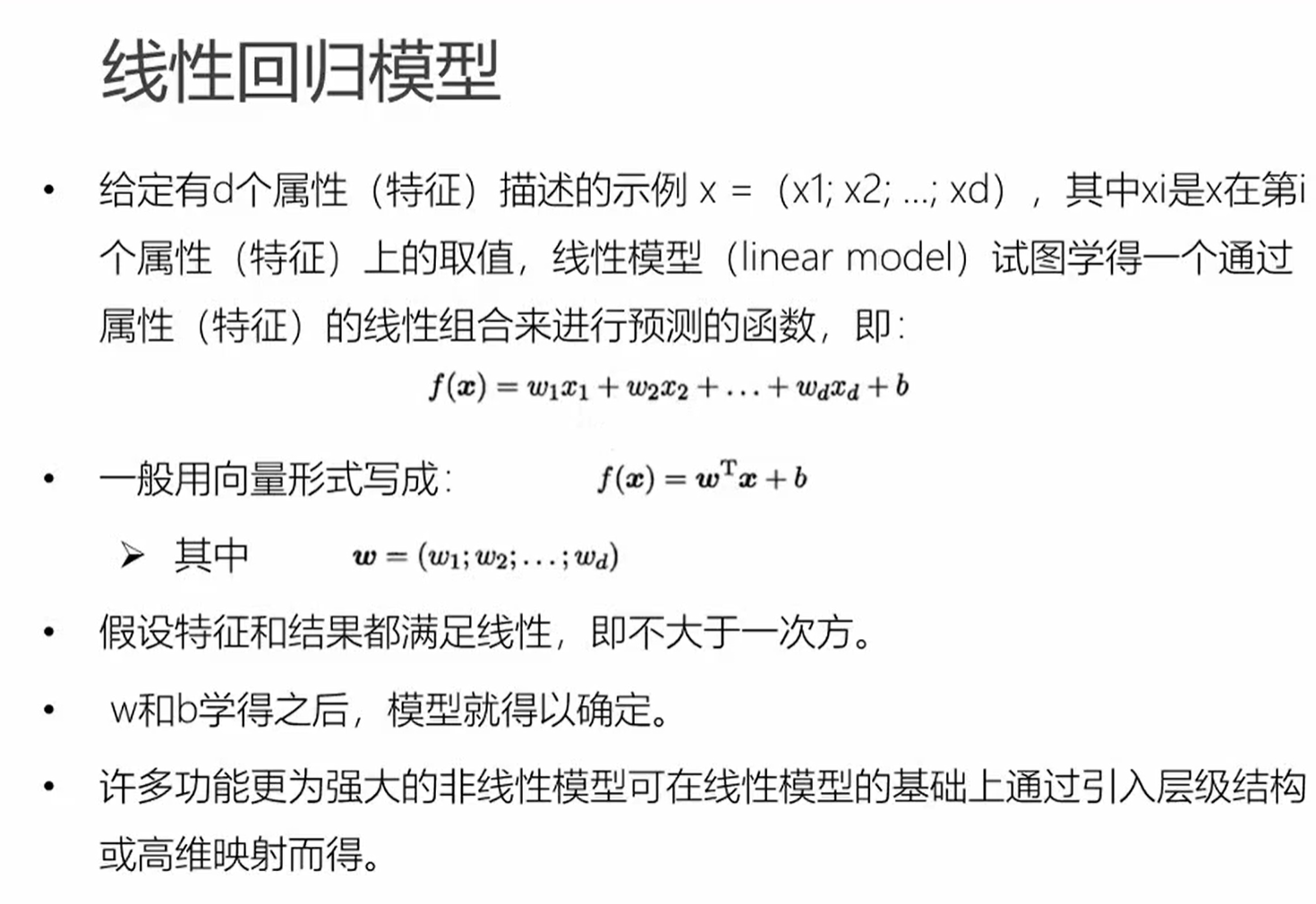
======2020/07/14====机器学习模型==================================

主要内容



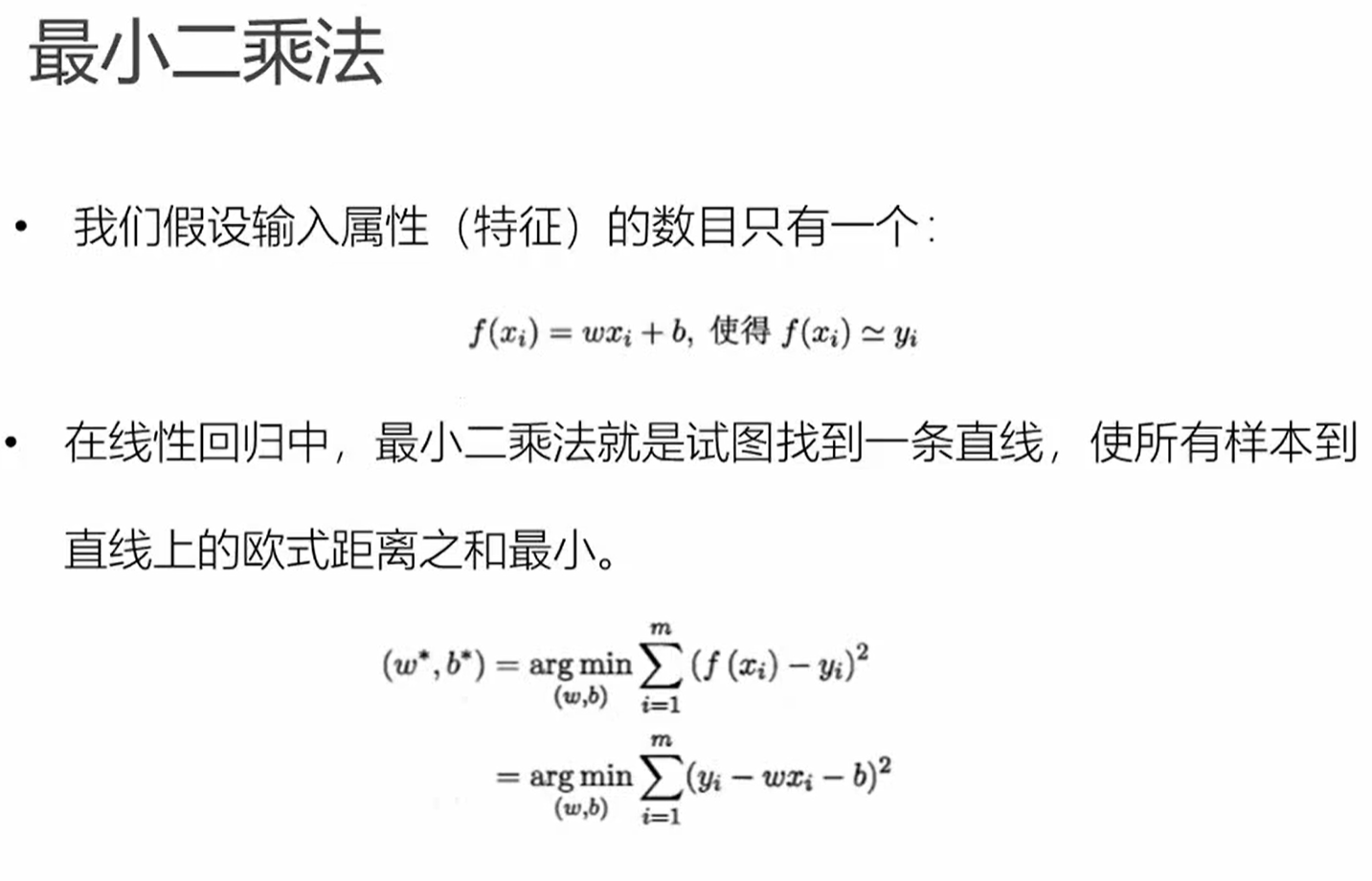
1. 监督学习 ---- 回归模型
   1. 线性回归模型
      1. 一元线性回归
      2. 多元线性回归
   2. 非线性回归模型
   3. 最小二乘法
2. 线性回归模型

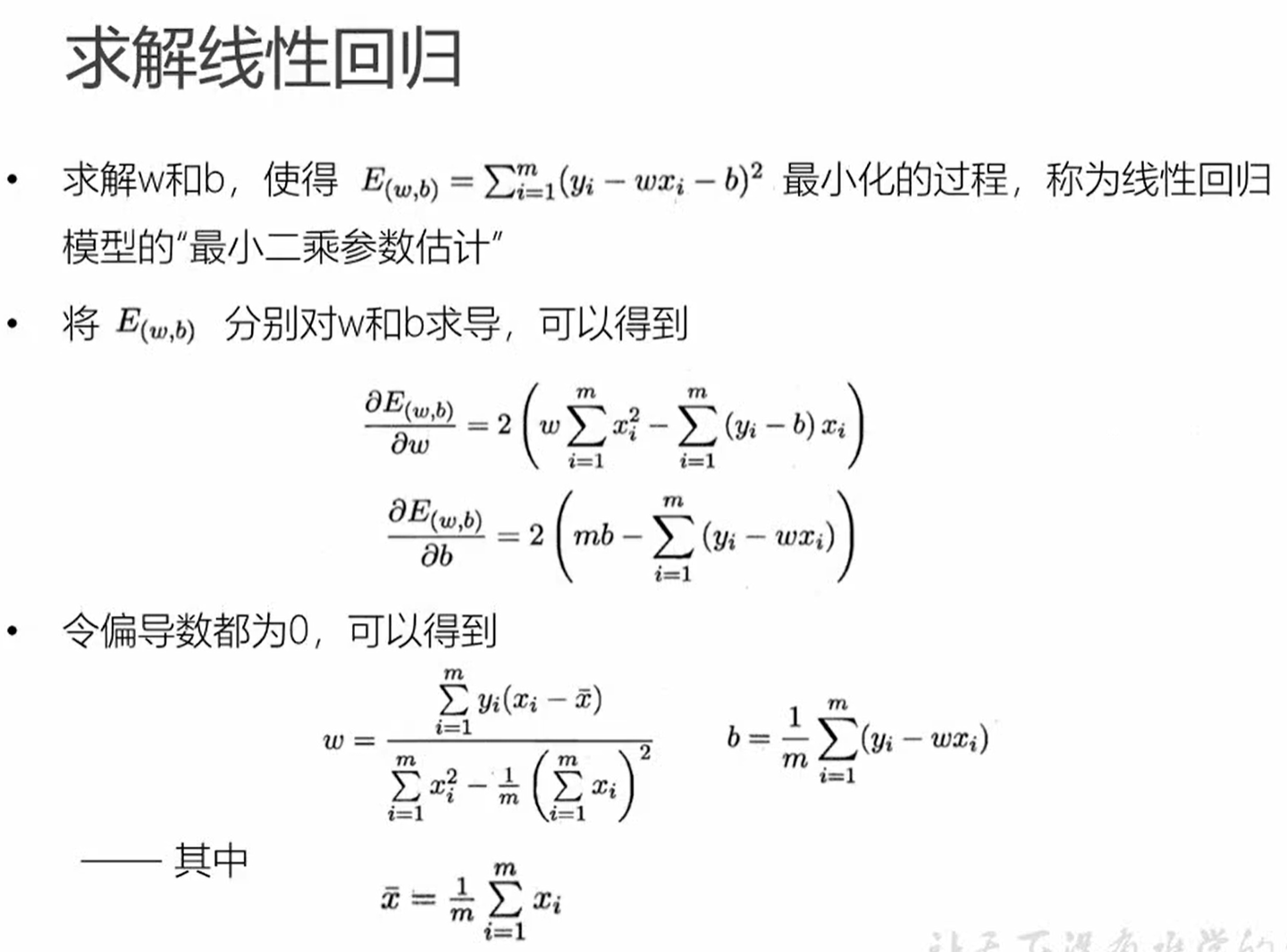




1. 最小二乘法







1. 简单线性回归（最小二乘法）