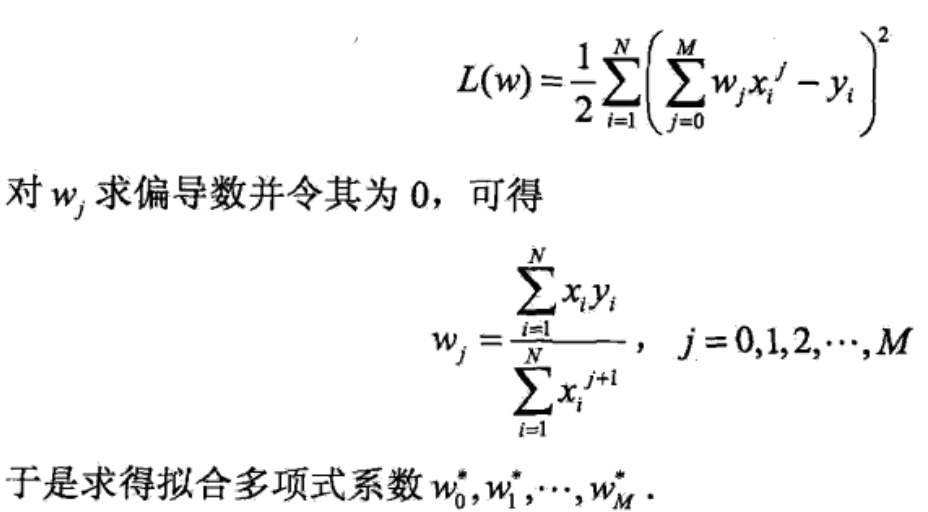
1. 自己提出的问题的理解（罗列全部）：
2. 提出的问题1：这个地方是如何去求偏导的？对外层求导时内层是否还保留有wj的其它项？  
   

讨论后的理解：此处可能需要用到其他的复杂知识，暂且认为是举例说明模型选择。

1. 提出的问题2：从贝叶斯估计的角度来看，正则化项对应于模型的先验概率，这句话如何理解？  
   讨论后的理解：先验概率可以说是根据以往的经验得到的概率，是一个假设，复杂的模型被采用的可能性小。
2. 别人提出的问题的理解（选择几个问题罗列，并给出理解）：
3. 问题3：如何理解罚项的定义？

自己的理解：正则化属于泛函，体现模型的复杂度，是从一个函数到一个值域的映射，调整风险结构。

1. 问题4：很好的解释已知数据的评判标准是什么？

自己的理解：是泛化损失函数，相同的泛化误差下，先验概率越小越好。

1. （必填）读书计划
2. 本周完成的内容章节：《统计学习方法》第一章
3. 下周计划：《统计学习方法》第二章

四、读书摘要及理解或伪代码的具体实现（读书摘要、伪代码的具体实现代码等可以写到这个部分）

读书摘要及理解：

1. 统计学习方法概论

三要素：模型、策略和算法

# 统计学习

基本假设是同类数据具有一定的统计规律性。

**假设空间：**模型属于由输入空间到输出空间的映射的集合

# 监督学习

**输入空间、输出空间：**输入和输出所有可能取值的集合。

每一个具体的输入是一个**实例**，通常由**特征向量**表示。所有特征向量存在的空间称为**特征空间**。

输入向量与输出向量均为变量序列的预测问题称为**标注问题**。

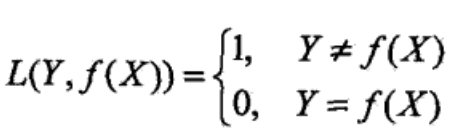
# 统计学习的三要素

## 模型

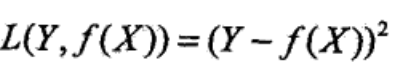
模型的假设空间包含所有可能的条件概率分布或决策函数。

## 策略

1. 损失函数和风险函数
2. 0-1损失函数



1. 平方损失函数



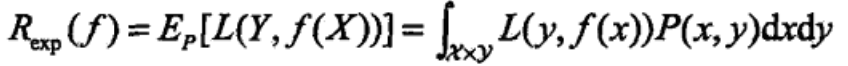
1. 绝对值损失函数



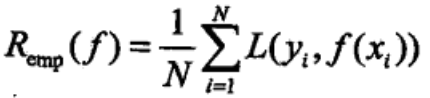
1. 对数损失函数（对数似然损失函数）



损失函数的期望，理论上是模型f（X）关于联合分布P（X）的平均意义下的损失，称为风险函数或期望风险。

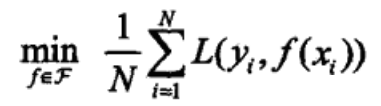


模型f（X）是模型关于训练样本集的平均损失称为经验风险或者经验损失。



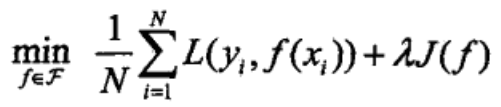
1. 经验风险最小化与结构风险最小化

经验风险最小化（ERM）：



样本容量很小时，经验风险最小化学习的效果未必很好，会产生过拟合现象，而结构风险最小化能防止过拟合。

结构经验最小化（SRM）：

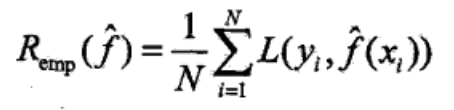


J（f）为模型复杂度，是定义在假设空间上的泛函。

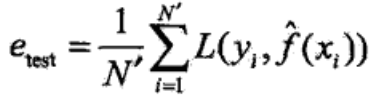
## 算法

# 模型估计与模型评测

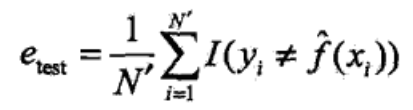
**训练误差：**



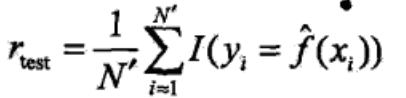
**测试误差：**



当损失函数是0-1损失时，测试误差为误差率：



准确率：



显然：



对未知数据的预测能力称为**泛化能力。**

# 正则化与交叉验证

## 正则化项

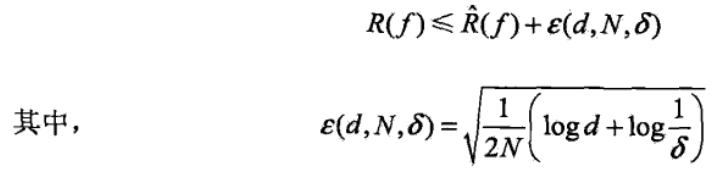
**正则化项**一般是模型复杂度的单调递增函数，符合奥卡姆剃刀原理，对应模型的先验概率。

## 验证集

验证集一般用于模型的选择

# 泛化能力

泛化误差上界：是样本容量的函数，容量增加，泛化上界趋于0；是假设空间容量的函数，容量越大，模型越难，泛化上界越大。



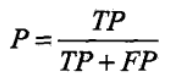
# 生成模型与判别模型

**生成方法：**学习联合概率分布P（X，Y）。

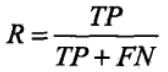
**判别方法：**学习决策函数或者条件概率分布P（X|Y）

# 分类问题

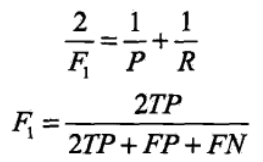
精准率：



召回率：



F1值（P和R的调和平均）：



# 标注问题

# 输入是一个观测序列，输出是一个标记序列或状态序列。

# 回归问题

分为一元回归和多元回归等，有线性回归和非线性回归。