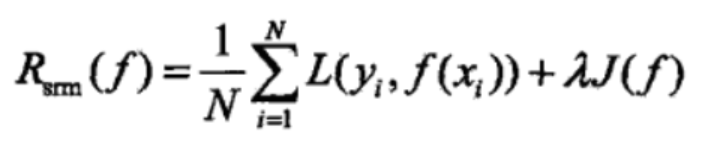
1. （必填）自己提出的问题的理解（罗列全部）：
2. **提出的问题1：**

采用结构化风险最小策略时，如何考虑模型的复杂度表示？

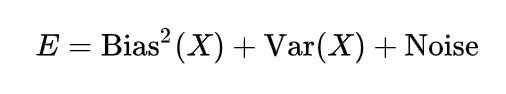
讨论后的理解：

结构风险定义为：

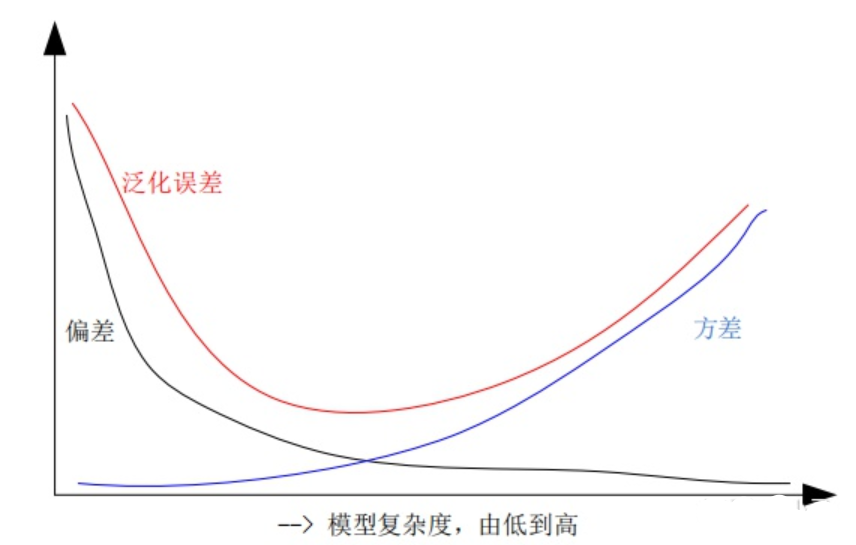


其中J(f)为模型的复杂度，是定义在假设空间上的泛函。模型f越复杂，复杂度J(f)就越大。也就是说，复杂度表示了对复杂模型的惩罚。

根据奥卡姆剃刀定律，“如无必要，勿增实体”，即认为相对简单的模型泛化能力更好。而模型泛化能力强、泛化误差小，即表示模型推广能力强，通俗理解就是在训练集中训练得到的优秀模型能够很好的适用于实际测试数据，而不仅仅是减少训练误差或者测试误差。泛化误差定义如下：



其中，E表示泛化误差，Bias代表偏差，Var代表方差，Noise代表噪声，如图所示为泛化误差与偏差和方差的关系：



随着训练程度加深，模型复杂度会增加，偏差减少，方差增大，而泛化误差呈现 U 型变化，对于一个“好的系统”通常要求误差小，正则化的作用即为适当的控制模型复杂度，从而使得泛化误差曲线取最小值。

1. **提出的问题2：**

如何考虑权衡经验风险和模型复杂度的参数？

讨论后的理解：

正则化：模型复杂度的单调递增函数，用来权衡经验风险和模型复杂度。正则化是假设模型服从先验概率，为模型添加先验。一般用L1和L2范数做正则化项。

L1正则：拉普拉斯先验，参数更新时使|w|≈0的参数向0靠近，将部分参数置为0，降低复杂度

L2正则：高斯先验，使得w变小加剧，而更小的参数意味着模型的复杂度更低。某一特征的权重过大时，容易依赖这个特征(L2正则可限制参数过大)。过拟合时，函数波动大、导数大、说明参数大(同上)

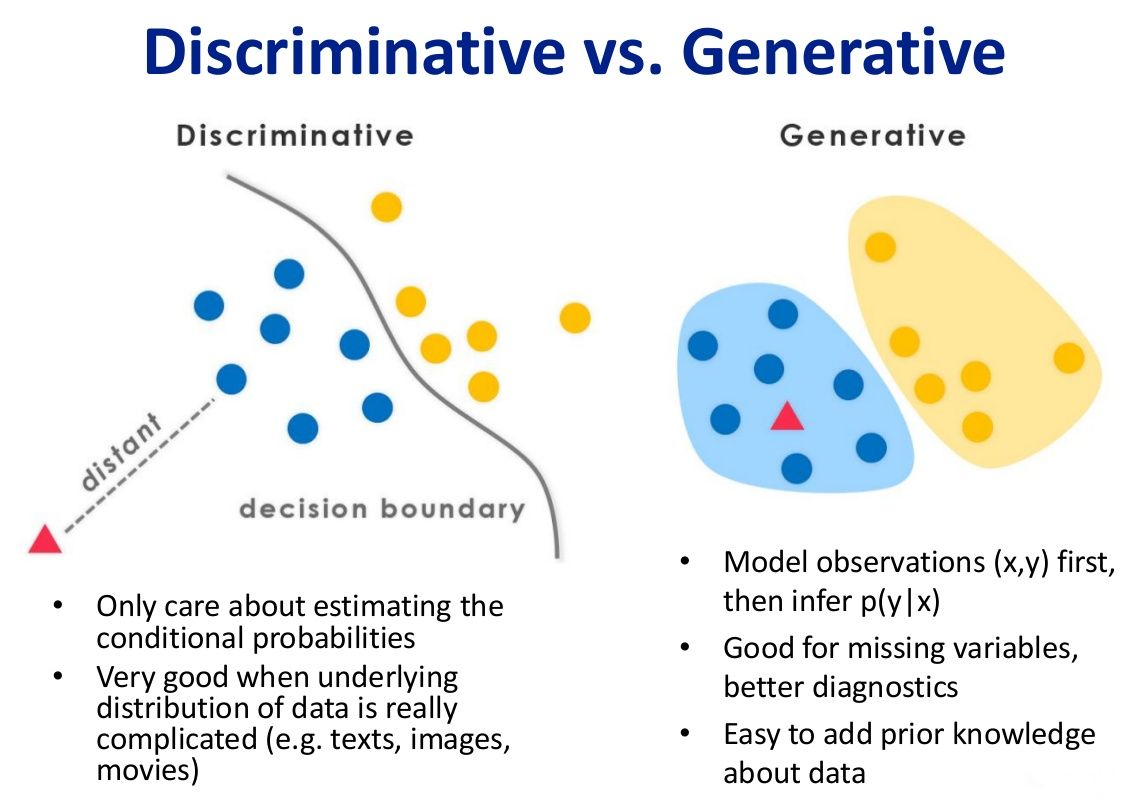
惩罚参数λ越大，表明先验分布协方差越小，偏差越大，模型越稳定，即：加入正则项是在偏差bias与方差variance之间做平衡。

1. （必填）别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. **问题3**：

生成方法和判别方法分别适用于什么情况？

自己的理解：

下图清晰地展示了生成模型和判别模型的区别：

机器学习本质上来说是从属性X预测Y的过程，即求得概率P(X|Y);

对于判别模型来说，求得P(X|Y)，对于未见实例X，根据P(X|Y)求得标记Y，即可以直接判别出来。如上图左半边部分所示，判别模型得到边界，如线性回归模型、支持向量机SVM等都是判别式模型，模型的特点都是输入属性X可以直接得到Y。

生成模型而生成式模型求得P(Y,X)，对于未见示例X，要求出X与不同标记之间的联合概率分布，然后大的获胜，如上图右边所示，与判别模型的区别在于生成模型没有边界，。机器学习中朴素贝叶斯模型、隐马尔可夫模型HMM等都是生成式模型，对于输入X，需要求出多个联合概率，然后选择较大的作为预测结果。

判别式模型应用场景：要确定一个羊是山羊还是绵羊，用判别模型的方法是从历史数据中学习到模型，然后通过提取这只羊的特征来预测出这只羊分别属于两种羊额的概率。

生成式模型应用场景：利用生成模型是根据山羊的特征首先学习出一个山羊的模型，然后根据绵羊的特征学习出一个绵羊的模型，然后从这只羊中提取特征，分别放到两个模型中测试，选择概率更大的模型类别作为结果。

1. **问题4：**

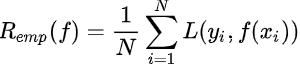
为什么根据大数定理，当样本容量趋于无穷时，经验风险趋于期望风险？

自己的理解：

期望风险公式如下：



经验风险公式如下：



大数定理是指某个随机事件在单次试验中可能发生也可能不发生，但在大量重复实验中往往呈现出明显的规律性，即该随机事件发生的频率会向某个常数值收敛，该常数值即为该事件发生的概率。换句话说当样本数据无限大时，样本均值趋于总体均值。经验风险是根据样本来计算，而期望风险是根据总体数据来计算。因此当样本容量趋近于无穷时，经验风险趋近于期望风险。

1. **问题5：**

如何证明条件概率分布模型选择对数损失函数就可以通过经验风险最小化推导极大似然估计？

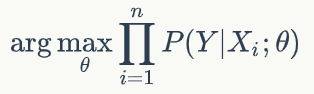
用数学描述：

模型为条件概率分布模型：

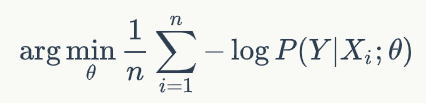
损失函数（策略）为对数损失函数，即 

证明经验风险最小化等价于极大似然估计。

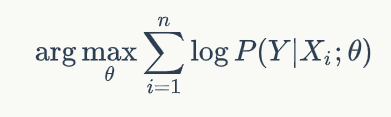
极大似然估计的形式（最大化似然函数）：



经验风险最小化形式（最小化经验风险）：



经验函数最小化形式等价于（去掉常数，通过负号修改 为 ）



而上面这个式子又等价于极大似然估计的对数似然形式。

所以经验风险最小化等价于极大似然估计。

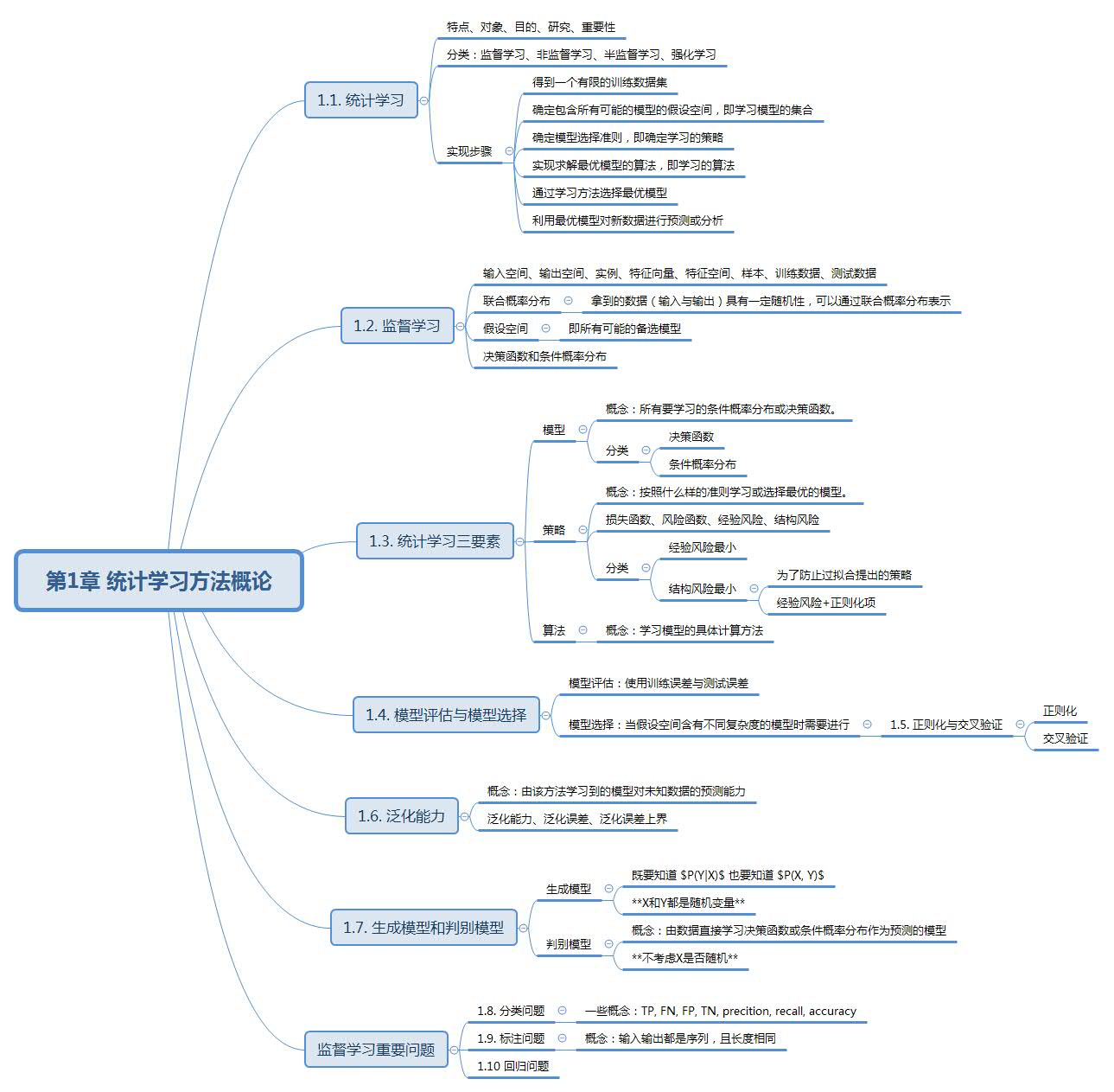
1. （必填）读书计划

1、本周完成的内容章节：统计学习方法（第一章）

2、下周计划：统计学习方法（第二章）

四、（选做）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）

1、读书摘要及理解（选做）



2、伪代码的具体实现(选做)