1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：对泛化误差上界的理解？

讨论后的理解：泛化误差是检验训练后的模型具不具备代表性，用偏差（bias）和方差（variance）来描述。如果一种方法学习的模型比另外一种方法学习的模型拥有更小的泛化误差上界，那么这种方法就更有效。泛化误差上界是样本容量的函数，当样本容量增加时，泛化误差上界趋于0；泛化误差上界也是假设空间容量的函数，假设空间容量越大，模型越难学，泛化误差上界就越大。

1. 提出的问题2：为什么回归学习中最常用的损失函数是平方损失函数？评估几种常见的损失函数。

讨论后的理解：因为平方损失函数梯度随着误差增大或减小，收敛效果好。

其他几种损失函数：

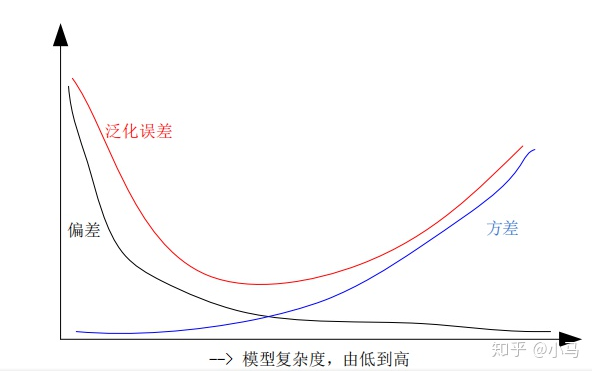
1）平方绝对值误差：对异常值具有较好鲁棒性；

1. Huber损失：对异常值更加鲁棒，在最优点附近由于调整为MSE，梯度更新会随着误差减小而减小，有利于收敛；
2. Log-Cosh损失：具有huber损失具备的所有优点，二阶处处可微；
3. （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）：
4. 问题3：课后习题1.2,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？

自己的理解：最小化经验风险，即最大化∑Ni=1logP(yi|xi)，最大化∏Ni=1P(yi|xi)即可，这个结果就是需要推到的极大似然估计。

1. 问题4：采用结构化风险最小策略时，如何考虑模型的复杂度表示？

自己的理解：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. 问题5：P23正则化项如何对应于模型的先验概率，为什么前边说模型越复杂正则化值越大，后面又说可以假设复杂的模型有较小的先验概率

自己的理解：因为正则化可以看作是对模型复杂度的量化指标，模型越复杂，正则 化要限制的参数空间就越大。复杂模型对变量之间的关系建立了更多约束，相比于简单 模型，更加难以满足，先验更低。两者都反映了正则化对应于模型的先验分布

1. （必填）读书计划
2. 本周完成的内容章节：<机器学习方法>第一章

下周计划：<机器学习方法>第二章