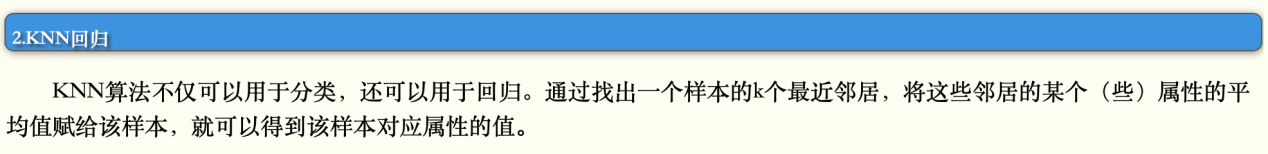
1. 自己提出的问题的理解（罗列全部）：
2. 提出的问题1：第三章开头说k近邻法是一种基本的分类与回归的方法，那如何将k近邻用于回归当中呢？是在找到k近邻之后求平均值这样的作法吗？

讨论后的理解：将邻居的平均值赋给该样本。



1. 提出的问题2：在最近邻的搜索算法中，当目标点和“当前最近点”之间的距离为半径构成的圆和另一子节点的区域相交时，会移动到另一个子节点当中递归地进行最近邻搜索。请问是怎么在另一子结点中搜索的？  
   讨论后的理解：用同样的操作来遍历另一个节点。
2. 提出的问题3：k近邻算法要求找到最近的k个最近邻，但书上的kd tree算法貌似只找了最近的一个实例点？请问如何寻找k个近邻呢？我能想到的一种方法是找到最近邻的一个之后将此点去掉再寻找下一个最近邻，请问还有什么效率更高的方法吗？  
   讨论后的理解：目前只能想到这一种方式，可能有其他的方式。
3. 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4：怎么理解超矩形区域的意思？

讨论后的理解：超矩形在二维里面是普通的矩形，三维里面是长方体。超矩形可以理解成垂直于每个坐标轴作一个切面形成的高维图形。

1. 问题5：k值的选择与数据有没有什么联系，除了交叉验证法多次计算选取最优k值之外有无别的方法？  
   讨论后的理解：与数据应该有关系，有一个经验范围，根据点的分布选取k值。
2. 问题6：为什么说用较小的k值，即较小的邻域中的训练实例进行预测，学习的近似误差会减小，反之是估计误差会减小？  
   讨论后的理解：近似误差是训练误差，估计误差是测试集的测试误差。k值越小，则形成了更多的单元，模型更复杂，所以近似误差更小，类似于过拟合，所以估计误差比较大。
3. 问题7：如何理解kd tree算法搜索的部分，尤其回退的环节？  
   讨论后的理解：通过回退才能算是遍历完了整个数据集，但对于通过半径判断后不相交的另一个子节点，不用再去搜索，因此算法才得以提升时间。
4. （必填）读书计划
5. 本周完成的内容章节：《统计学习方法》第二章和第三章
6. 下周计划：《统计学习方法》第四章

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：

# 感知机

旨在求出将训练数据进行线性划分的分离超平面。

## 2.1 感知机模型



## 2.2 感知机学习策略

### 2.2.1 数据集的线性可分性

能够将数据集的正实例点和负实例点完全正确地划分到超平面两侧的数据集为**线性可分数据集。**

### 2.2.2 感知机学习策略

损失函数的一个自然选择是误分类点的总数，但不可导，不易优化。

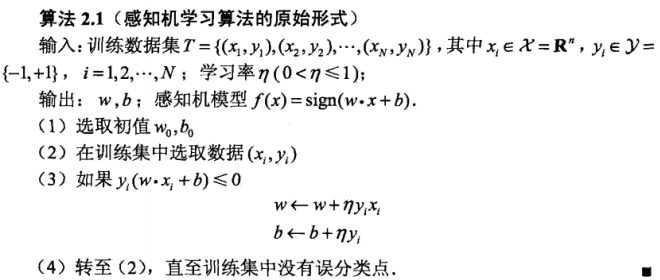
另一个损失函数是误分类点到超平面的总距离。



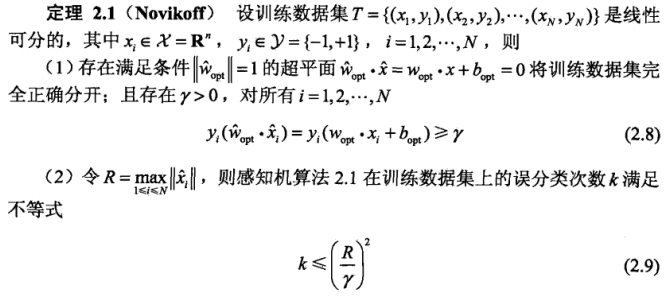
## 2.3 感知机学习算法

**随机梯度下降法**

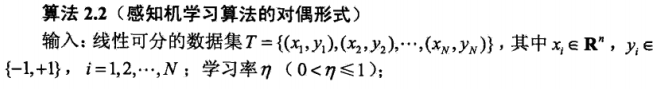
### 2.3.1 感知机学习算法的原始形式

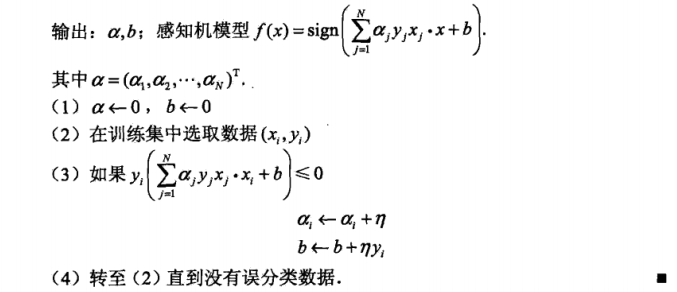


### 2.3.2 算法的收敛性



### 2.3.3 感知机学习算法的对偶形式

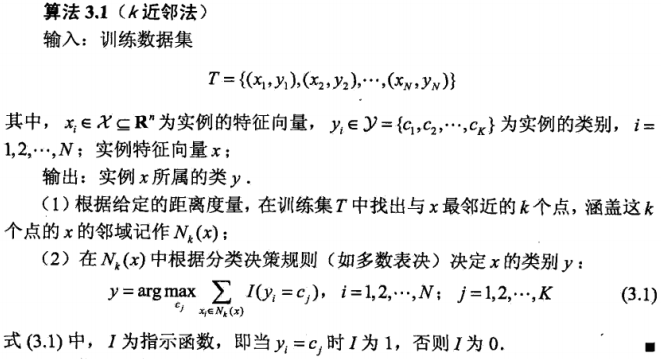




# k近邻法

k近邻法是一种基本分类与回归的方法。

## 3.1 k近邻法



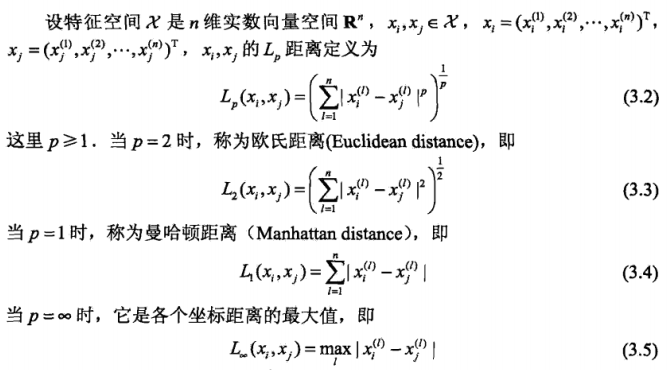
## 3.2 k 近邻模型

模型由三个基本要素——距离度量、k值的选择和分类决策规则决定。

### 3.2.1 模型

距离该点比其他点更近的所有点组成一个区域，叫做**单元。**

### 3.2.2 距离度量



### 3.2.3 k值的选择

k值的减少意味着整体模型变得复杂，容易发生过拟合。

k值一般取一个比较小的数值，通常采用交叉验证法来选取最优的k值。

### 3.2.4 分类决策规则

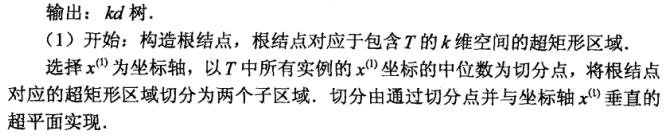
多数表决规则等价于经验风险最小化。

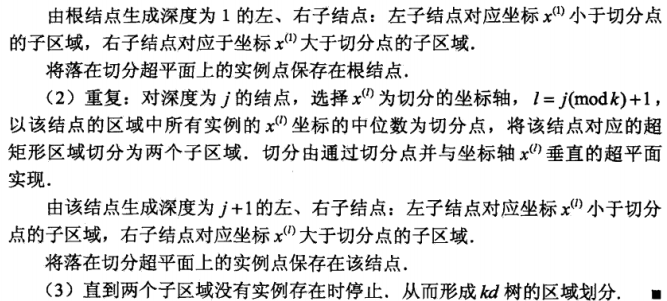
## 3.3 k近邻法的实现：kd树

### 3.3.1 构造kd树

Kd树是二叉树，表示对k维空间的一个划分。

依次选择坐标轴对空间切分，选择训练实例点在选定坐标轴的中位数为切分点，这样得到的kd树是平衡的。注意，平衡的kd树搜索时的效率未必是最优的。





### 3.3.2 搜索二叉树

如果父节点的另一个子节点的超矩形区域与超球体相交，那么在相交区域内寻找与目标点更近的实例点。

