1. 自己提出的问题的理解（罗列全部）：
2. 提出的问题1：这里的Π(x)怎么计算的？  
   7d941c852d82689871e2a96858a4fdb

讨论后的理解：这里就是P(Y=1|x)的一个记号，和前面的意思相同。

1. 提出的问题2：如何理解这里的fi？  
   6915dd8cf98233706f325b357f492e2  
   讨论后的理解：fi(x,y)相当于是问题的约束条件，例如例题中的P(A)+P(B)=3/10;Ep(fi)=Ep~(fi)相当于是对所有模型的集合的约束条件。
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4：在逻辑斯谛模型中，条件概率分布为什么选用了sigmoid函数来表示？

讨论后的理解：中部上升速度快，有利于分类。

1. 问题5：怎么理解最大熵原理中，为什么在所有可能的概率模型（分布）中，熵最大的模型是最好的模型？  
   讨论后的理解：选择熵最大的模型是一种选择随机变量统计特性最符合客观情况的准则。
2. 问题6：多项逻辑斯谛公式是怎么得来的？  
   讨论后的理解：第K项P(Y=K|x)会表示成1-前面k-1项。
3. （必填）读书计划
4. 本周完成的内容章节：《统计学习方法》第六章
5. 下周计划：《统计学习方法》第七章

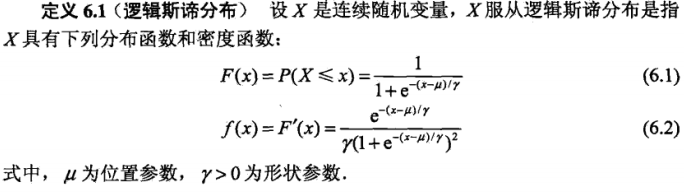
四、读书摘要及理解或伪代码的具体实现（读书摘要、伪代码的具体实现代码等可以写到这个部分）

读书摘要及理解：

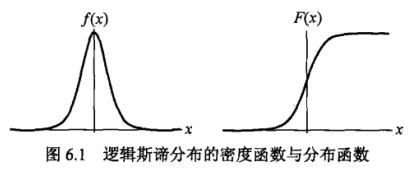
1. 逻辑斯蒂回归与最大熵模型

6.1 逻辑斯蒂回归模型

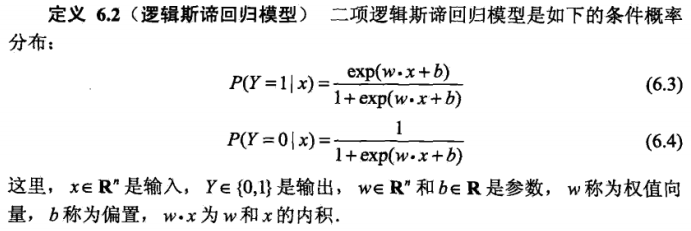
6.1.1 逻辑斯蒂分布



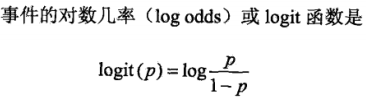
形状参数γ的值越小，曲线在中心附近增长得越快。



6.1.2 二项逻辑斯蒂回归模型



一个事件的**机率**是指该事件发生的概率与该事件不发生的概率的比值。

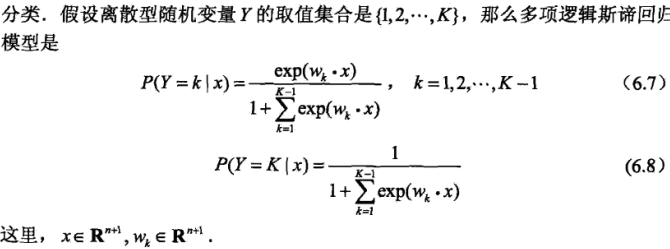


在逻辑斯蒂回归模型中，输出Y=1的对数几率是输入x的线性函数。

6.1.3 模型参数估计

逻辑斯蒂回归学习中通常采用的方法是梯度下降法及拟牛顿法。

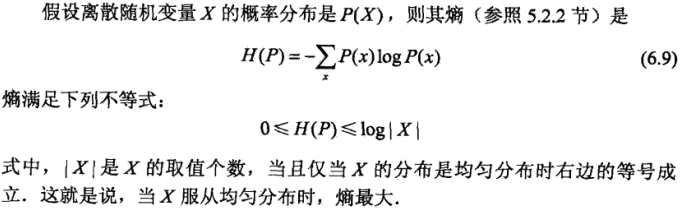
6.1.4 多项逻辑斯蒂回归



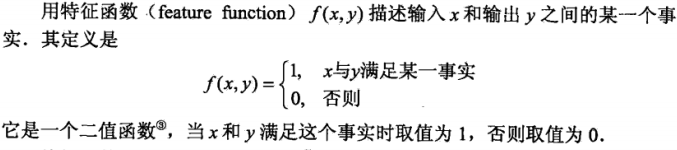
6.2 最大熵模型

6.2.1 最大熵原理

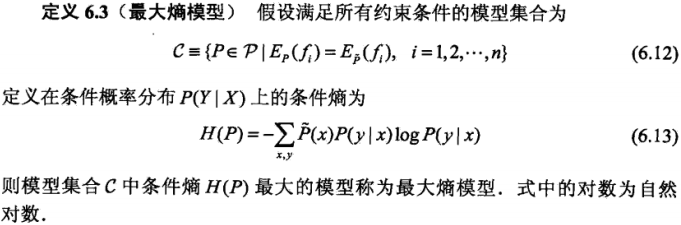
最大熵眼里认为，学习概率模型时，在所有可能的概率模型（分布）中，熵最大的模型时最好的模型。



6.2.2 最大熵模型的定义



如果模型能够获取的训练数据中的信息，那么就可以假设这两个期望值相等。



6.2.3 最大熵模型的学习

学习过程就是求解最大熵模型的过程。

6.2.4 最大似然估计

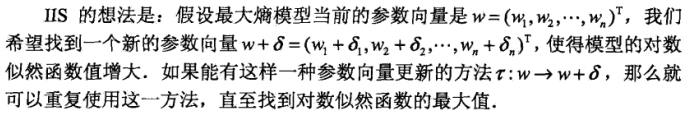
对偶函数的极大化等价于最大熵模型的极大似然估计。

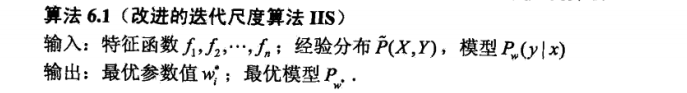
最大熵模型与逻辑斯蒂回归模型有类似的形式，他们又称为**对数线性模型**。

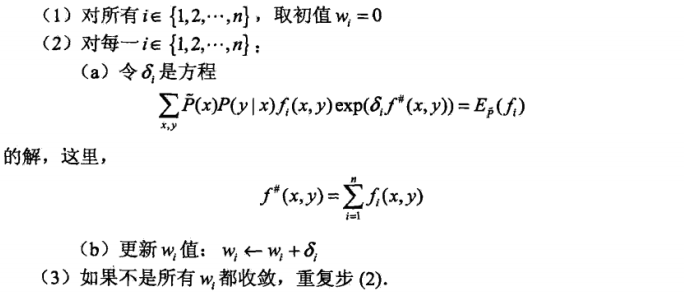
6.3 模型学习的最优化算法

常用的方法有改进的迭代尺度法、梯度下降法、牛顿法或拟牛顿法。牛顿法和拟牛顿法一般收敛速度更快。

6.3.1 改进的迭代尺度法







6.3.2 拟牛顿法

