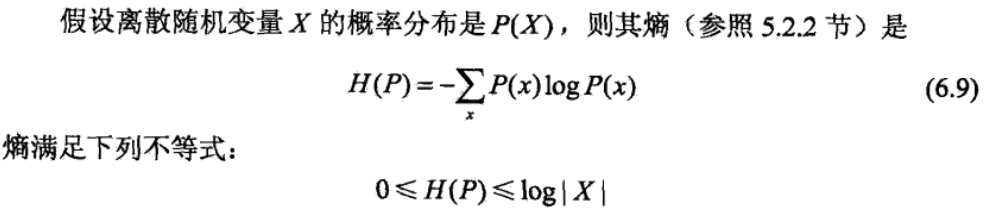
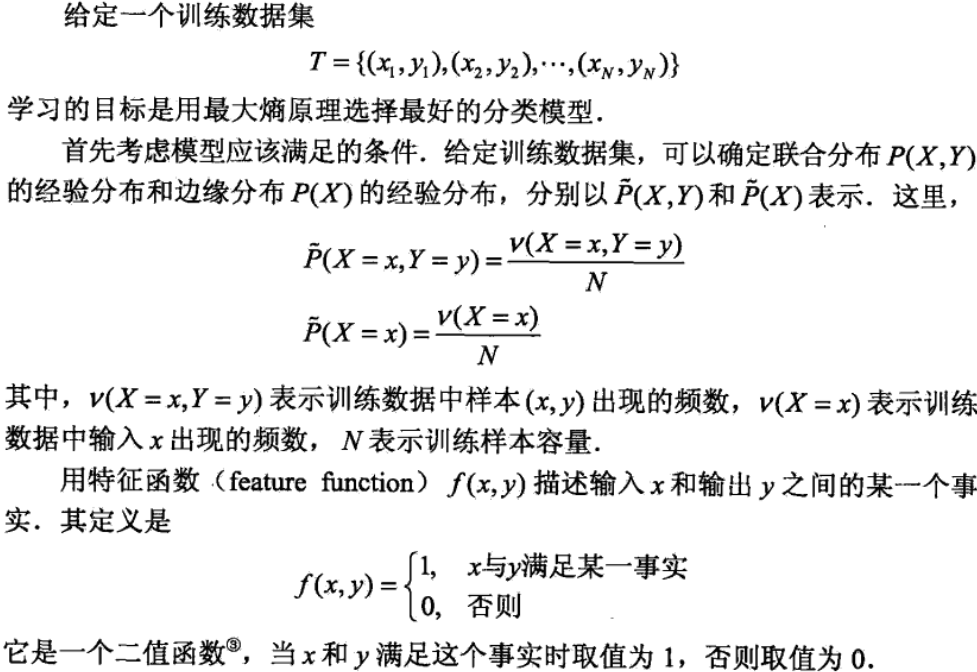
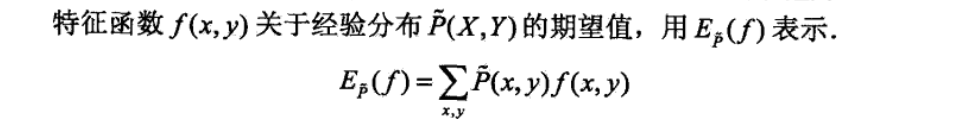
## 6最大熵

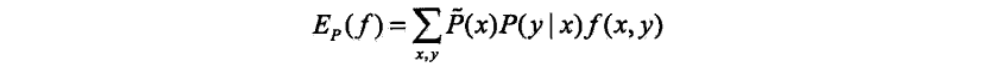
最大熵原理是概率模型学习的一个准则。最大熵原理认为，学习概率模型时，在所有可能的概率模型（分布）中，熵最大的模型是最好的模型。

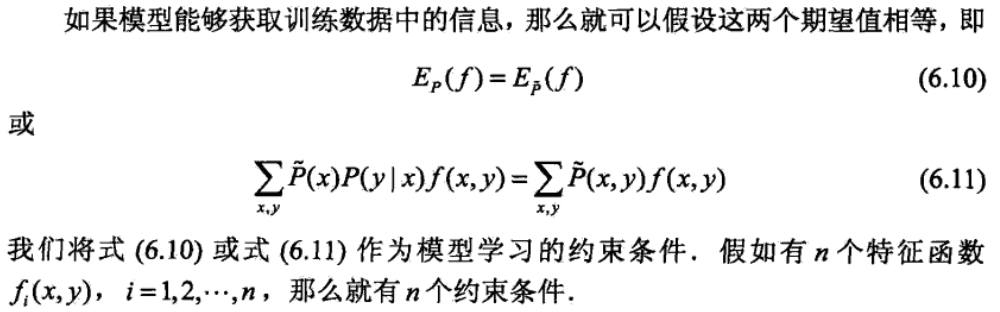




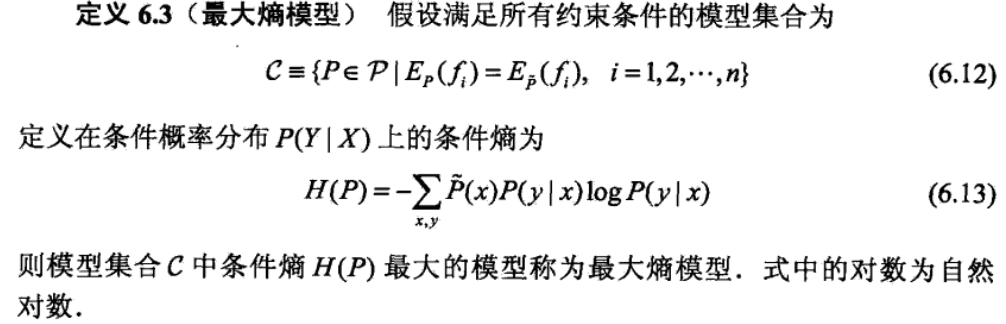


C:\Users\NKU\AppData\Local\Temp\mx3E952.png

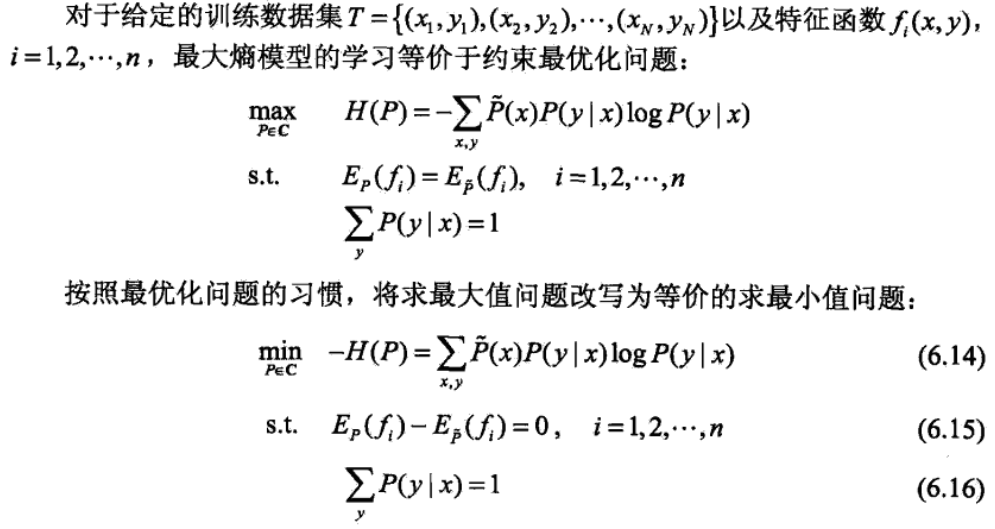


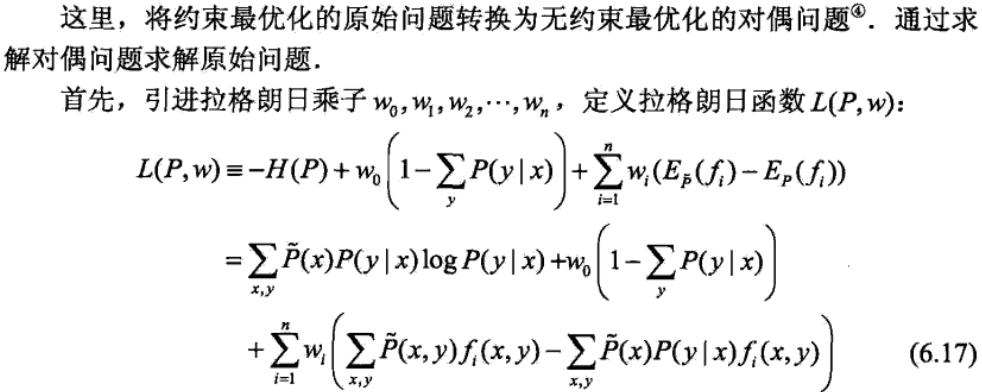


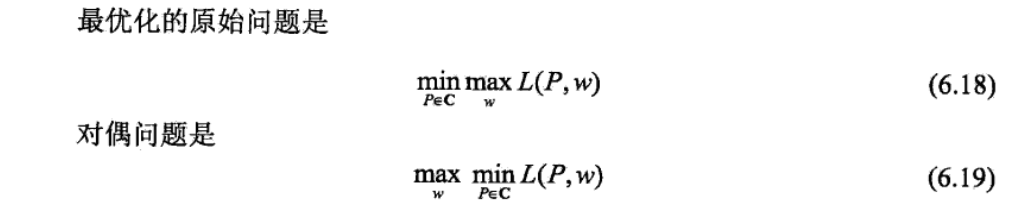
**最大熵模型的定义**



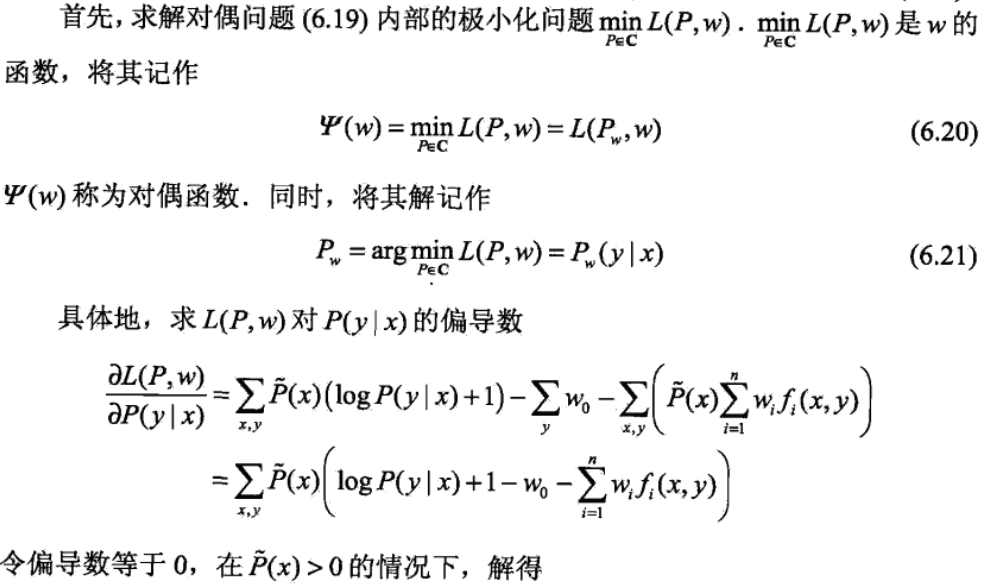
**最大熵模型的学习**

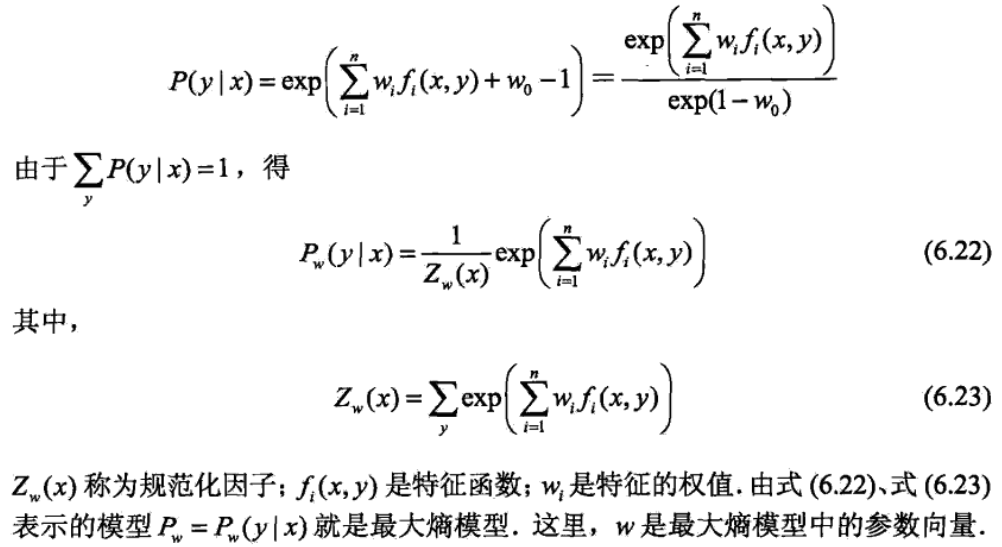


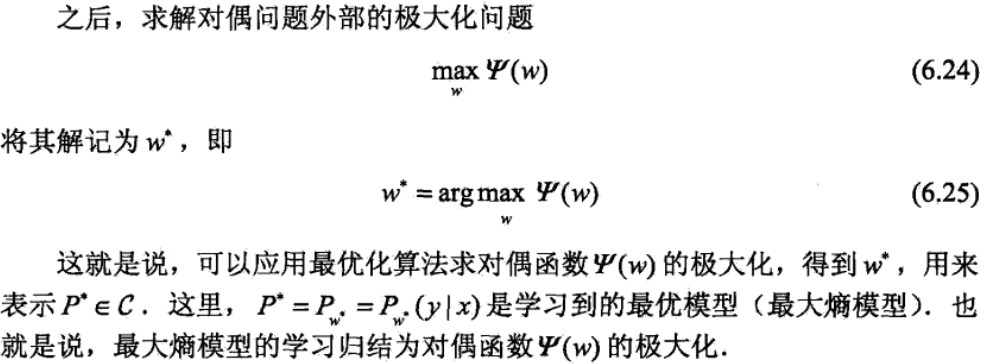




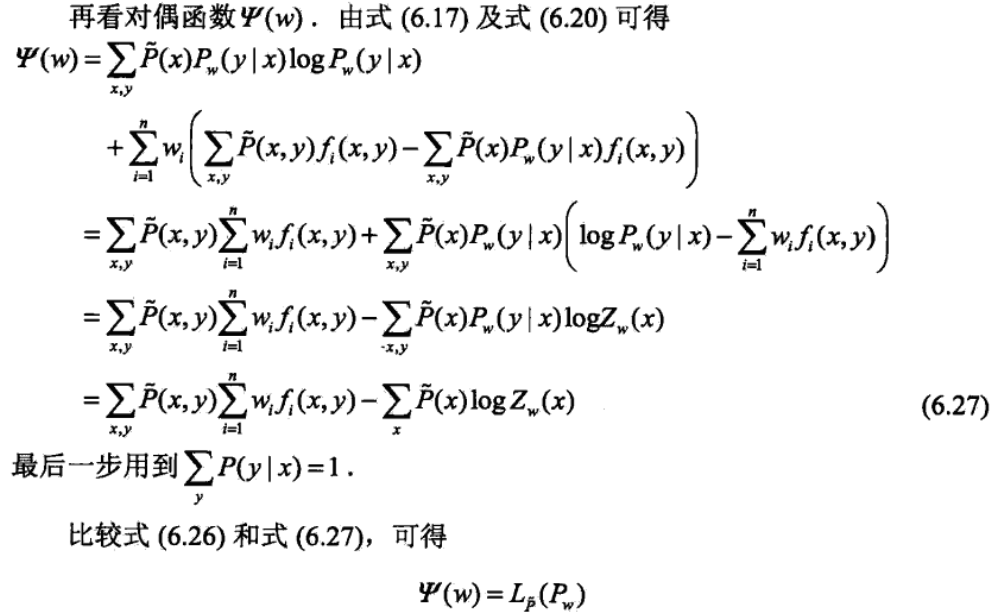
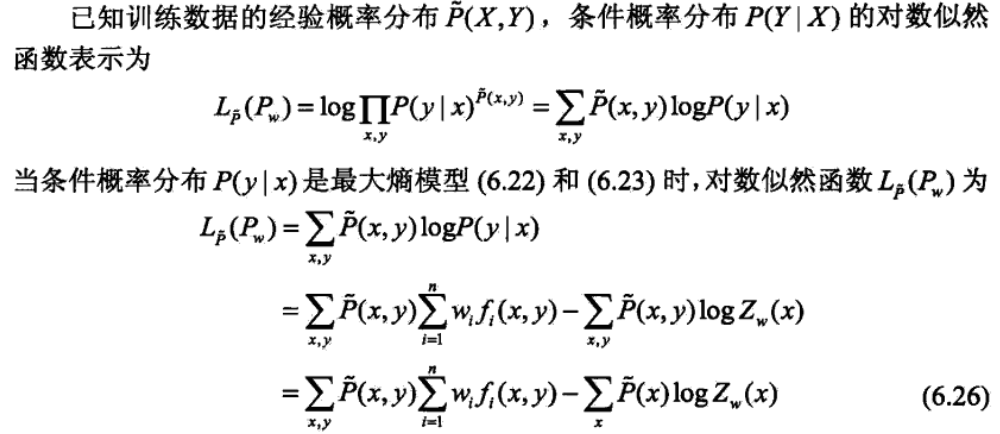
**模型的学习过程**







**对偶函数的极大化等价于最大熵模型的极大似然估计**



**学习策略**

逻辑斯谛回归模型、最大熵模型学习归结为以似然函数为目标函数的最优化问题，通常通过迭代算法求解。从最优化的观点看，这时的目标函数具有很好的性质。它是光滑的凸函数，因此多种最优化的方法都适用，保证能找到全局最优解。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。牛顿法或拟牛顿法一般收敛速度更快。

