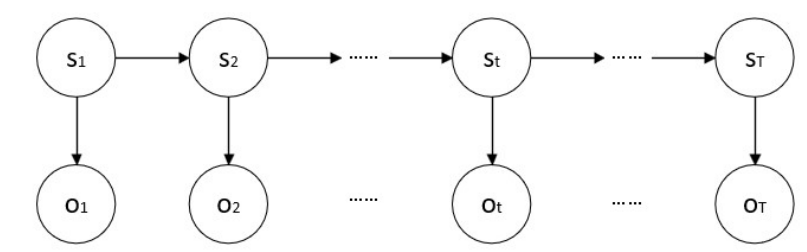
概率模型

# HMM-隐马尔科夫模型

具有马尔科夫属性的概率有向图



## 1.1 两个空间三组参数

状态空间：隐状态S的取值范围

观测空间：观测状态O的取值范围

发射概率：初始状态的概率分布，记做π

转移概率：不同状态之间的转移概率，可以用转移矩阵表示，记做a

输出概率：基于当前状态，不同输出的概率分布，记做b

模型参数λ = (a, b, π)

## 两个假设

1. 齐次假设：即马尔科夫假设
2. 观测独立性假设：观测值只取决于对应的状态值，与其他状态无关

## 三个问题

1. 概率计算：已知模型参数λ求某个观测序列O的概率
2. 预测：已知模型参数和观测序列O，求最有可能的隐状态序列
3. 学习：已知观测序列O(或者还要状态序列S)，对参数进行估计

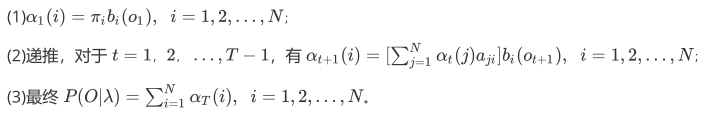
# 2、三个问题的计算方法

## 2.1 概率计算-前、后向算法

两种方法从本质上讲都是**动态规划**，他们可以分别用来求解p(O|λ)

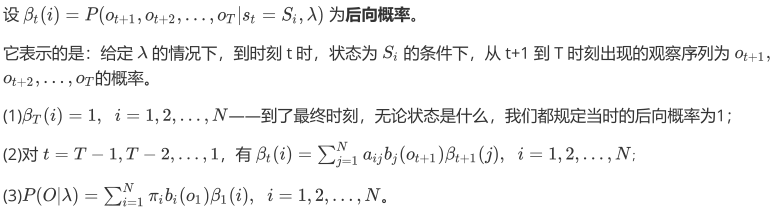
**前向算法：**





前向概率：下标表示时刻，i表示t时刻的状态，并且观测序列包括t时刻。

**后向算法：**



后向概率：下标表示时刻，i表示t时刻的状态，但是观测序列不包括t时刻。

**结合前后向算法：**



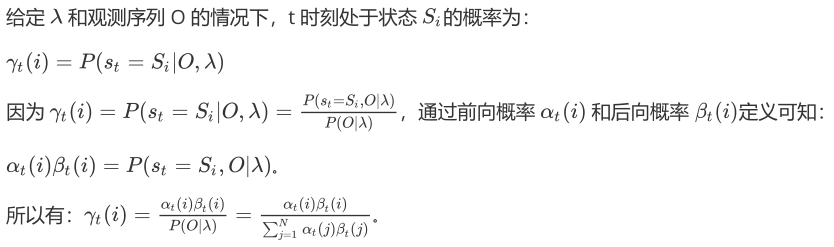
用三者其一都可以计算，而且复杂度是一样的，因为前后向算法都是动规。

## 2.2 预测算法

已知观测序列，求最有可能的状态序列

1. 拼凑法

求出每个时刻不同隐状态的概率分布，选取概率最大的状态作为S\*(i)，然后将不同时刻的状态拼在一块即可。缺点是：可能不是全局最优解。



分母的理解：加法公式P(A)=P(AB1)+P(AB2)+...+P(ABn)

1. 维特比算法

动规+回溯

## 2.3 学习算法

当只有观测序列O，用无监督方式，当有状态序列S和观测序列O，用监督方式.

1. 监督学习

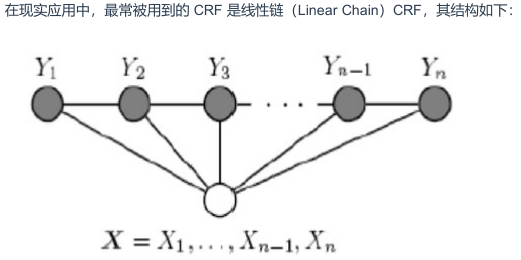
基于O/S序列来估计模型参数，采用统计的思想，用频数来逼近频率(包括发射概率、转移概率、输出概率)

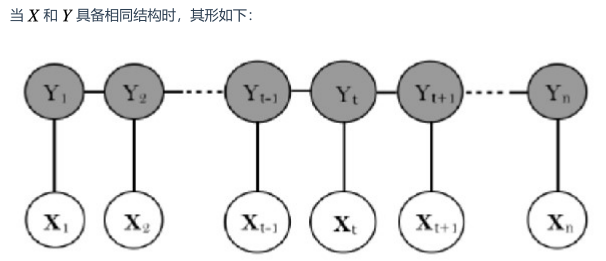
1. 无监督学习

Baum-Welch算法

条件随机场-CRF

# 1、定义



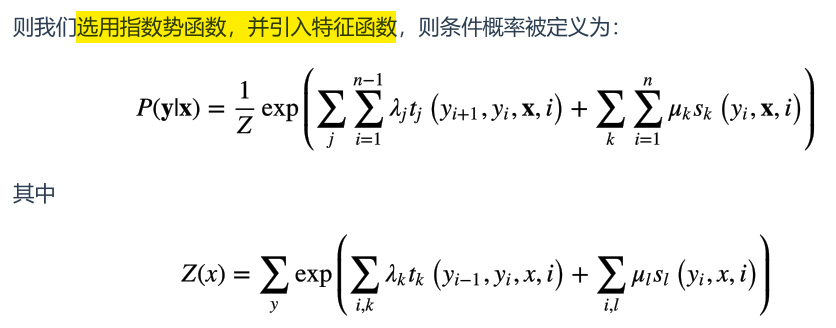


CRF也对应状态序列和观测序列。

**定义：**已知观测序列，求状态序列时，如果状态序列满足马尔科夫属性（只和直接关联的状态有关）即:



称P(Y|X)为条件随机场。

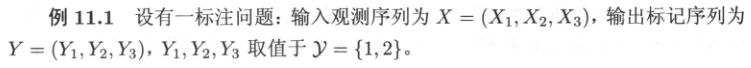


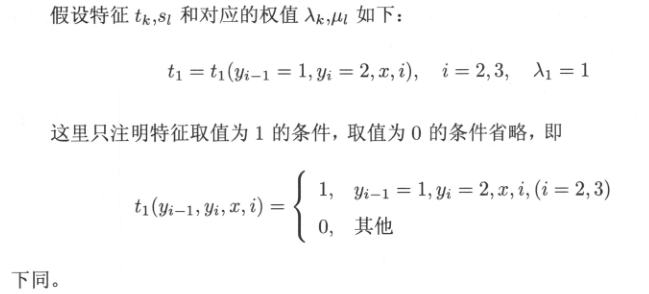
i表示不同的位置(或者时刻)，j表示不同的特征函数

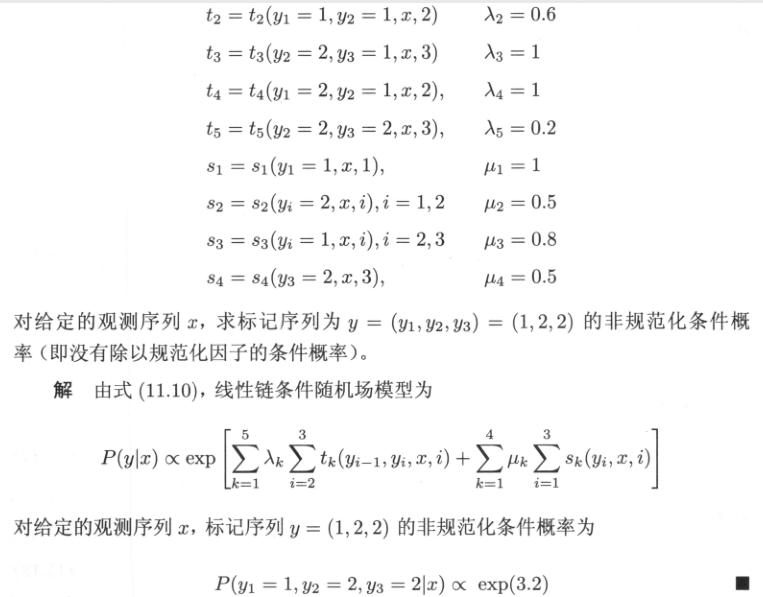
**特征函数**：刻画数据可能成立的经验特征。例如动词一般不连续出现，形容词后加名词。

λ和u是去权重函数、Z是为了归一化。

**例子：《统计学习方法》**







**首先将t1写为：**

**t1 = t1(y1=1, y2=2, x, 2)**

**t1’ = t1’(y2=1, y3=2, x, 3)**

然后这个检查y序列是否满足各个特征函数：t1和t5满足，s1、s2、s4满足，所以结果为3.2.

## 2、三个问题、解法

1、概率计算-前后向算法

2、预测-维特比算法

3、学习-改进的迭代尺度算法