《统计学习方法》第十一章读书笔记

一、自己提出的问题的理解：

1. 提出的问题1：不是很理解P225 11.32式的分子？按照书上的文字叙述，分子应该是两个向量中对应i的标记的概率的积，但公式中直接是两个向量的内积？对11,33式也有相同的疑问。

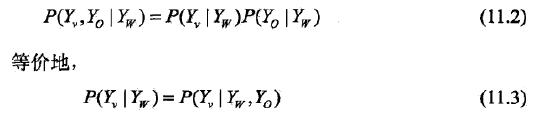
**讨论后的理解**：此处我自己的理解是分子应该是两个概率相乘，而不是两个向量，只有这样这两个公式才能解释得通。

二、别人提出的问题的理解：

2、问题2：为什么说成对的、局部的、全局的马尔可夫性是等价的？

**讨论后的理解：**三种定义是等价的，应该只是数学表示形式不同。从几何角度理解，都是在表达“如果两个节点没有边相连，则它们对应的随机变量发生的概率是互相独立的”这一件事。







三、读书计划

1. 本周完成的内容章节：第十一章
2. 下周计划：第十二章及前半本书的复习回顾

四、读书摘要及理解

条件随机场（conditional random fields，简称 CRF，或CRFs），是一种判别式概率模型，是随机场的一种，常用于标注或分析序列资料，如自然语言文字或是生物序列。条件随机场是条件概率分布模型 P(Y|X) ，表示的是给定一组输入随机变量 X 的条件下另一组输出随机变量 Y 的马尔可夫随机场，如同马尔可夫随机场。

条件随机场是一种判别式无向图模型，其问题为给定条件随机场，由输入序列对输出序列进行预测，也就是计算条件概率。为了方便起见，引入类似HMM的前向-后向算法。

定义表示序列位置的标记是时，在位置之前的部分标记序列的非规范化概率。注意，在这里不加入规范化因子是因为我们不想影响下一步的计算。

在11.2.4中，定义了在给定时，从转移到的非规范化概率，即：

由此，可以得到在位置的标记是时，在位置之前的部分标记序列的非规范化概率 的递推公式:

在起点处，我们定义：

假设可能的标记总数是, 则的取值就有个，我们用表示这m个值组成的前向向量：

可以用矩阵表示由形成的阶矩阵：

此时递推公式可以用矩阵来表示：

后向向量的定义同理。