



## 法律声明

本课件包括演示文稿、示例、代码、题库、视频和声音等内容，深度之眼和讲师拥有完全知识产权；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或者机构不得盗版、复制、仿造其中的创意和内容，我们保留一切通过法律手段追究违反者的权利。

### 课程详情请咨询

- 微信公众号：深度之眼
- 客服微信号：deepshare0920



公众号



微信

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料





deepshare.net

深度之眼

# 西瓜书公式推导

导师: Sm1les (Datawhale南瓜书项目负责人)

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料



# 隐马尔可夫模型公式推导

Derivation of Hidden Markov Model

---



# 本节大纲

Outline

---



deepshare.net

深度之眼

先修内容：西瓜书14.1、统计学习方法第10章

1. 隐马尔可夫模型简介

2. 概率计算问题

3. 学习问题

4. 预测问题

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料

# 预测问题

Prediction problem

---



deepshare.net

深度之眼

1. 近似算法

2. 维特比算法

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料

# 预测问题

Prediction problem



deepshare.net

深度之眼

近似算法：

近似算法思想：在每个时刻 $t$ 选择在该时刻最有可能出现的状态  $i_t^*$ ，从而得到一个状态序列  $I^* = (i_1^*, i_2^*, \dots, i_T^*)$ ，将它作为预测的结果。具体算法如下：

给定隐马尔可夫模型 $\lambda$ 和观测序列 $O$ ，在时刻 $t$ 处于状态  $q_i$  的概率  $\gamma_t(i)$  是

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^N \alpha_t(j)\beta_t(j)}$$

在每一时刻 $t$ 最有可能的状态  $i_t^*$  是

$$i_t^* = \arg \max_{1 \leq i \leq N} [\gamma_t(i)], \quad t = 1, 2, \dots, T$$

从而得到状态序列  $I^* = (i_1^*, i_2^*, \dots, i_T^*)$

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料

# 预测问题

Prediction problem



维特比算法：

维特比算法实际是用动态规划解隐马尔可夫模型预测问题，即用动态规划求概率最大路径，这时一条路径对应着一个状态序列。具体算法如下：

定义在时刻t状态为  $q_i$  的所有单个路径  $(i_1, i_2, \dots, i_t)$  中概率最大值为

$$\delta_t(i) = \max_{i_1, i_2, \dots, i_{t-1}} P(o_1, \dots, o_t, i_1, \dots, i_{t-1}, i_t = q_i), \quad i = 1, 2, \dots, N$$

由上述定义可知：

$$\delta_1(i) = \pi_i b_{io_1}$$

$$\delta_2(i) = \max_{1 \leq j \leq N} [\delta_1(j) a_{ji}] b_{io_2}$$

$$\delta_3(i) = \max_{1 \leq j \leq N} [\delta_2(j) a_{ji}] b_{io_3}$$

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料

# 预测问题

Prediction problem



deepshare.net

深度之眼

依次此类推可得如下递推公式：

$$\delta_t(i) = \max_{1 \leq j \leq N} [\delta_{t-1}(j) a_{ji}] b_{io_t}$$

定义在时刻t状态为  $q_i$  的所有单个路径  $(i_1, i_2, \dots, i_{t-1}, i_t)$  中概率最大的路径的第t-1个结点为

$$\psi_t(i) = \arg \max_{1 \leq j \leq N} [\delta_{t-1}(j) a_{ji}]$$

因此，取  $i_T^* = \arg \max_i [\delta_T(i)]$ ，则  $i_{T-1}^* = \psi_T(i_T^*)$ ,  $i_{T-2}^* = \psi_{T-1}(i_{T-1}^*)$ , ...,  $i_1^* = \psi_2(i_2^*)$

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料



例10.3 (盒子和球模型) 假设有3个盒子，每个盒子里面都装有红白两种颜色的球，盒子里的红白球数如下表所示：

盒子	1	2	3
红球数	5	4	7
白球数	5	6	3

按照下面的方法抽球，产生一个球的颜色观测序列：首先以0.2、0.4、0.4的概率从1、2、3号盒子中选取一个盒子，从这个盒子里随机抽出1个球，记录其颜色后放回，接着按以下概率选取下一个盒子

	1	2	3
1	0.5	0.2	0.3
2	0.3	0.5	0.2
3	0.2	0.3	0.5

确定转移的盒子后，再从盒子里随机抽出1个球，记录其颜色后放回，如此下去，重复3次，最终得到的观测序列为 $O=\{\text{红}, \text{白}, \text{红}\}$ ，记选取的盒子序列为状态序列，试求最优状态序列，即最优路径  $I^* = (i_1^*, i_2^*, i_3^*)$

解：该例子是一个典型的隐马尔可夫模型，由题意可得隐马尔可夫模型的参数（三要素）为：

$$A = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \quad B = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix} \quad \pi = (0.2, 0.4, 0.4)^T$$

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料

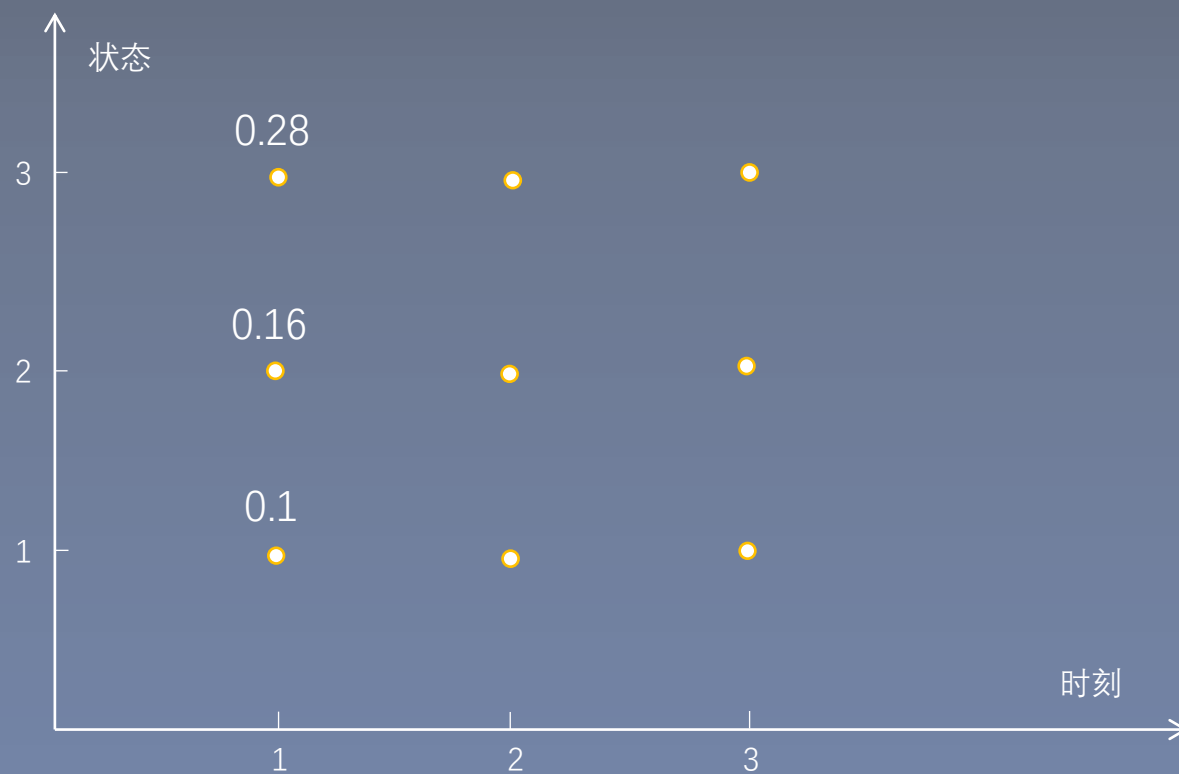
$$A = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \quad B = \begin{matrix} & \text{红} & \text{白} \\ \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix} \end{matrix} \quad \pi = (0.2, 0.4, 0.4)^T \quad O = \{\text{红}, \text{白}, \text{红}\}$$

按照维特比算法我们可以进行如下计算：

$$\delta_1(1) = \pi_1 b_{1o_1} = 0.2 \times 0.5 = 0.1, \quad \psi_1(1) = 0$$

$$\delta_1(2) = \pi_2 b_{2o_1} = 0.4 \times 0.4 = 0.16, \quad \psi_1(2) = 0$$

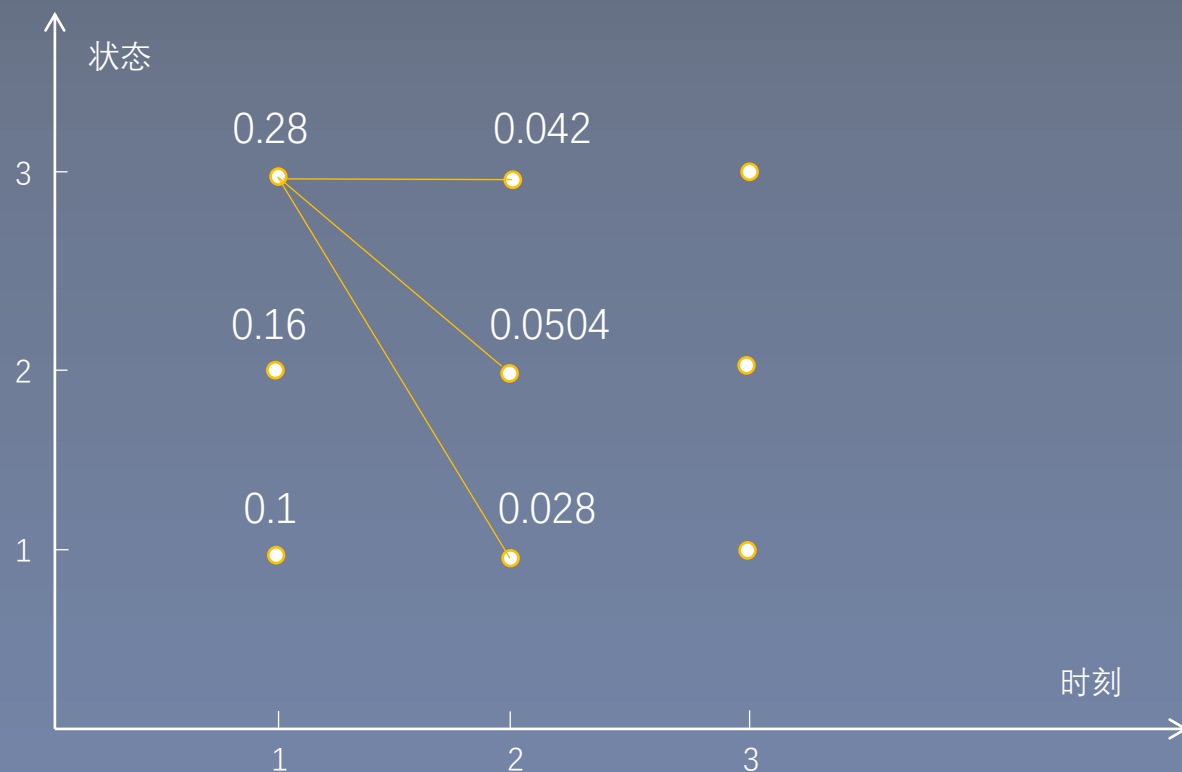
$$\delta_1(3) = \pi_3 b_{3o_1} = 0.4 \times 0.7 = 0.28, \quad \psi_1(3) = 0$$



$$\delta_2(1) = \max_{1 \leq j \leq 3} [\delta_1(j)a_{j1}]b_{1o_2} = \max \begin{Bmatrix} 0.1 \times 0.5 = 0.05 \\ 0.16 \times 0.3 = 0.048 \\ 0.28 \times 0.2 = 0.056 \end{Bmatrix} \times 0.5 = 0.028, \quad \psi_2(1) = \arg \max_{1 \leq j \leq 3} [\delta_1(j)a_{j1}] = 3$$

$$\delta_2(2) = \max_{1 \leq j \leq 3} [\delta_1(j)a_{j2}]b_{2o_2} = \max \begin{Bmatrix} 0.1 \times 0.2 = 0.02 \\ 0.16 \times 0.5 = 0.08 \\ 0.28 \times 0.3 = 0.084 \end{Bmatrix} \times 0.6 = 0.0504, \quad \psi_2(2) = \arg \max_{1 \leq j \leq 3} [\delta_1(j)a_{j2}] = 3$$

$$\delta_2(3) = \max_{1 \leq j \leq 3} [\delta_1(j)a_{j3}]b_{3o_2} = \max \begin{Bmatrix} 0.1 \times 0.3 = 0.03 \\ 0.16 \times 0.2 = 0.032 \\ 0.28 \times 0.5 = 0.14 \end{Bmatrix} \times 0.3 = 0.042, \quad \psi_2(3) = \arg \max_{1 \leq j \leq 3} [\delta_1(j)a_{j3}] = 3$$

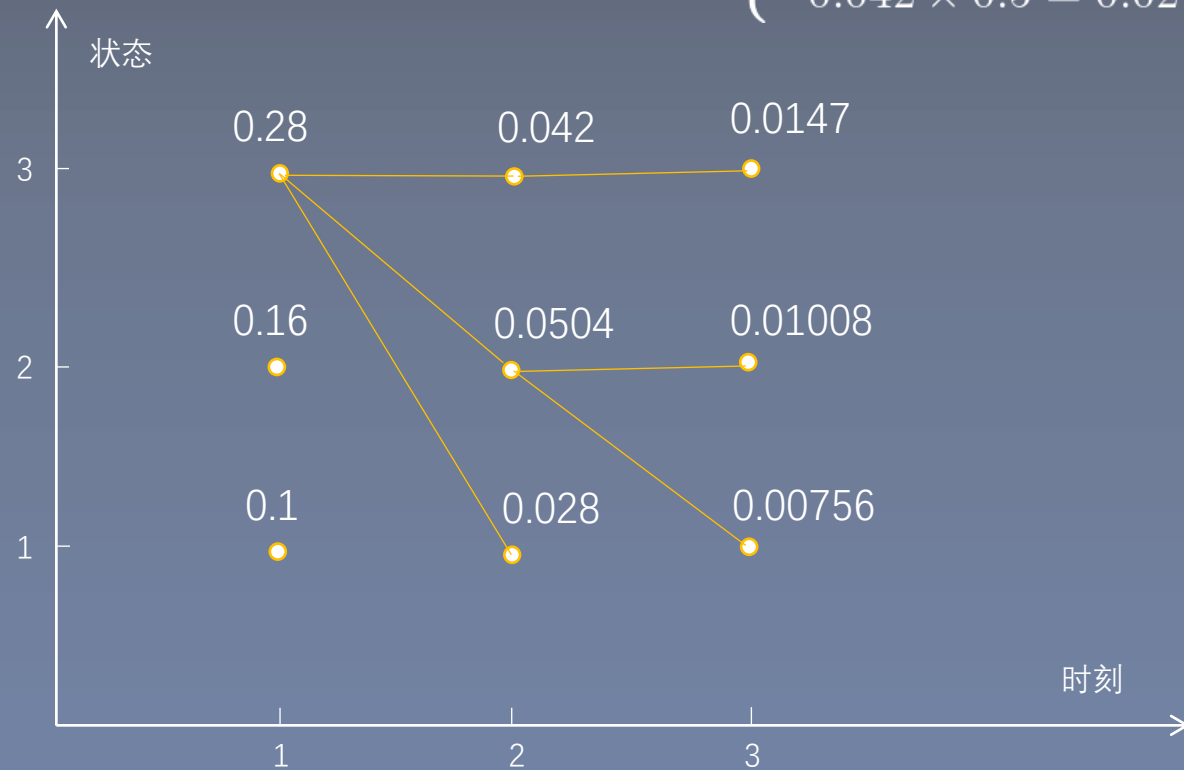




$$\delta_3(1) = \max_{1 \leq j \leq 3} [\delta_2(j)a_{j1}]b_{1o_3} = \max \begin{Bmatrix} 0.028 \times 0.5 = 0.014 \\ 0.0504 \times 0.3 = 0.01512 \\ 0.042 \times 0.2 = 0.0084 \end{Bmatrix} \times 0.5 = 0.00756, \quad \psi_3(1) = \arg \max_{1 \leq j \leq 3} [\delta_2(j)a_{j1}] = 2$$

$$\delta_3(2) = \max_{1 \leq j \leq 3} [\delta_2(j)a_{j2}]b_{2o_3} = \max \begin{Bmatrix} 0.028 \times 0.2 = 0.0056 \\ 0.0504 \times 0.5 = 0.0252 \\ 0.042 \times 0.3 = 0.0126 \end{Bmatrix} \times 0.4 = 0.01008, \quad \psi_3(2) = \arg \max_{1 \leq j \leq 3} [\delta_2(j)a_{j2}] = 2$$

$$\delta_3(3) = \max_{1 \leq j \leq 3} [\delta_2(j)a_{j3}]b_{3o_3} = \max \begin{Bmatrix} 0.028 \times 0.3 = 0.0084 \\ 0.0504 \times 0.2 = 0.01008 \\ 0.042 \times 0.5 = 0.021 \end{Bmatrix} \times 0.7 = 0.0147, \quad \psi_3(3) = \arg \max_{1 \leq j \leq 3} [\delta_2(j)a_{j3}] = 3$$



所以，最优状态序列为：

$$i_3^* = \arg \max_i [\delta_3(i)] = 3$$

$$i_2^* = \psi_3(i_3^*) = \psi_3(3) = 3$$

$$i_1^* = \psi_2(i_2^*) = \psi_2(3) = 3$$

# —— 结 语 ——

在这次课程中，我们学习了隐马尔可夫模型的  
预测问题

**感谢大家的收听，再见。**



关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料



**deepshare.net**

深度之眼

联系我们：

电话：18001992849

邮箱：[service@deepshare.net](mailto:service@deepshare.net)

QQ：2677693114



公众号



客服微信

关注公众号深度之眼，后台回复资料，获取AI必学书籍及完整实战学习资料