

实验2. 隐马尔科夫模型实践

MG1733079, 杨佩成, 18362903155@163.com

2017 年 12 月 1 日

综述

隐藏马尔科夫模型（Hidden Markov Model）是一种著名的有向图模型，主要用于时序数据建模，在语音识别、自然语言处理等领域有广泛应用[3]。

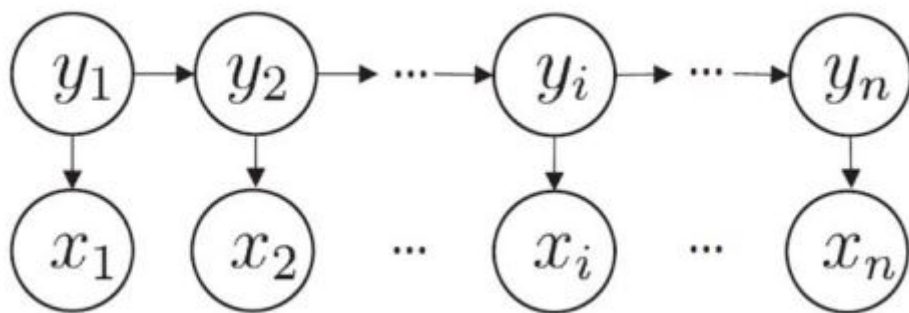


Figure 1: 隐马尔科夫模型的图结构

如Figure 1所示，隐马尔科夫模型中的变量可分为两组。第一组是状态变量 $\{y_1, y_2, \dots, y_n\}$ ，其中 $y_i \in \mathcal{Y}$ 表示第 i 时刻的系统状态。通常状态变量是隐藏的、不可被观测的，因此状态变量也叫隐变量。第二组是观测变量 $\{x_1, x_2, \dots, x_n\}$ ，其中 $x_i \in \mathcal{X}$ 表示第 i 时刻的观测值。

Figure 1中的箭头表示了变量间的依赖关系。在任意时刻，观测变量仅依赖于状态变量。同时， t 时刻的状态 y_t 仅依赖于 $t-1$ 时刻的状态 y_{t-1} 。基于这种依赖关系，所有变量的联合概率分布为

$$P(x_1, y_1, \dots, x_n, y_n) = P(y_1)P(x_1|y_1) \prod_{i=2}^n P(y_i|y_{i-1})P(x_i|y_i).$$

除了结构信息，要确定一个隐马尔科夫模型需要以下三组参数：

- 状态转移概率：模型在各个状态间转换的概率，通常记为矩阵 $\mathbf{A} = [a_{ij}]_{N \times N}$ ，其中

$$a_{ij} = P(y_{t+1} = s_j | y_t = s_i), \quad 1 \leq i, j \leq N,$$

表示在任意时刻 t ，若状态为 s_i ，则下一时刻状态为 s_j 的概率。

- 输出观测概率：模型根据当前状态获得各个观测值的概率，通常记为矩阵 $\mathbf{B} = [b_{ij}]_{N \times M}$ ，其中

$$b_{ij} = P(x_t = o_j | y_t = s_i), \quad 1 \leq i \leq N, 1 \leq j \leq M$$

表示在任意时刻 t ，若状态为 s_i ，则观测值 o_j 被观测的概率。

- 初始状态概率：模型在初始时刻各个状态出现的概率，通常记为 $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_N)$ ，其中

$$\pi_i = P(y_1 = s_i), \quad 1 \leq i \leq N$$

表示模型的初始状态为 s_i 的概率。

通过指定状态空间 \mathcal{Y} 、观测空间 \mathcal{X} 和上述三组参数，就可以确定一个隐马尔科夫模型。在本次实验中，我们使用HMM对股票数据进行分析 and 预测。对于单只股票，我们可以观察到的值可以是涨、跌、不涨不跌。我们假设股票的涨跌由内在的隐变量驱动，即牛市或熊市。

实验一.

实验一我们对训练好的HMM，利用维特比算法对模型进行推断。维特比算法是一种使用动态规划思想来寻找最有可能的隐状态序列的算法，维特比算法的基础可以概括为以下三点[2]：1、如果概率最大的路径经过篱笆网络的某点，则从开始点到该点的子路径也一定是从开始到该点路径中概率最大的。2、假定第 i 时刻有 k 个状态，从开始到 i 时刻的 k 个状态有 k 条最短路径，而最终的最短路径必然经过其中的一条。3、根据上述性质，我们在计算第 $i+1$ 状态的最短路径时，只需要考虑从开始到当前的 k 个状态值的最短路径和当前状态到第 $i+1$ 状态的最短路径即可。本次实验中，我们用来计算隐状态序列的算法如下。

Algorithm 1 VITBRBI

Input: initial probabilities $\boldsymbol{\pi}$, a sequence of observations \mathcal{X} , transition matrix \mathbf{A} , emission matrix \mathbf{B}

Output: the most likely hidden state sequence \mathcal{Y}

```
1: for each state  $i \in \{1, 2, \dots, K\}$  do
2:    $T_1[i, 1] \leftarrow \pi_i B_{ix_i}$ 
3:    $T_2[i, 1] \leftarrow 0$ 
4: end for
5: for each observation  $i = 2, 3, \dots, T$  do
6:   for each state  $j \in \{1, 2, \dots, K\}$  do
7:      $T_1[j, i] \leftarrow \max_k (T_1[k, i-1] \cdot A_{kj} \cdot B_{jx_i})$ 
8:      $T_2[j, i] \leftarrow \operatorname{argmax}_k ((T_1[k, i-1] \cdot A_{kj} \cdot B_{jx_i}))$ 
9:   end for
10: end for
11:  $\mathcal{Y}_T \leftarrow \operatorname{argmax}_k (T_1[k, T])$ 
12: for  $i \leftarrow T, T-1, \dots, 2$  do
13:    $\mathcal{Y}_{i-1} \leftarrow T_2[\mathcal{Y}_i, i]$ 
14: end for
15: return  $\mathcal{Y}$ 
```

实验二.

在参数未知并且只有参数序列的条件下, 我们可以使用Baum–Welch算法学习对应HMM的参数, Baum–Welch算法实际上是一种EM算法。在本次实验中我们要实现Baum–Welch算法中的前向算法。给定HMM的模型 $\theta = (\mathbf{A}, \mathbf{B}, \pi)$, 到 t 时刻观测序列为 x_1, x_2, \dots, x_t 且状态为 i 的概率为前向概率, 记为 $\alpha_i(t) = P(X_1 = x_1, \dots, X_t = x_t, Y_t = i | \theta)$, 我们可以通过下面的算法计算 $\alpha_i[1]$ 。

Algorithm 2 Forward Algorithm

Input: initial probabilities π , a sequence of observations \mathcal{X} , transition matrix \mathbf{A} , emission matrix \mathbf{B}

Output: forward probabilities α_i

```
1:  $\alpha_i(1) \leftarrow \pi_i B_{ix_1}$ 
2: for  $t=1, 2, \dots, T-1$  do
3:    $\alpha_i(t+1) \leftarrow B_{ix_{t+1}} \sum_{j=1}^N \alpha_j(t) A_{ji}$ 
4: end for
5: return  $\alpha_i$ 
```

实验三.

在实验三中, 我们需要实现Baum–Welch算法中的后向算法。后向算法是指计算给定观测序列计算其后向概率的算法。后向概率是指, 给定隐马尔可夫模型 $\theta = (\mathbf{A}, \mathbf{B}, \pi)$, 在时刻 t 状态为 i 的条件下, 从 $t+1$ 到 T 的观测序列为 $X_{t+1}, X_{t+2}, \dots, X_T$ 的概率。后向概率记为 $\beta_i(t) = P(X_{t+1} = x_{t+1}, \dots, X_T = x_T | Y_t = i, \theta)$ 。我们可以通过下面的算法计算 $\beta_i[1]$ 。

Algorithm 3 Backward Algorithm

Input: a sequence of observations \mathcal{X} , transition matrix \mathbf{A} , emission matrix \mathbf{B}

Output: backward probabilities θ_i

```
1:  $\beta_i(T) \leftarrow 1$ 
2: for  $t=T-1, T-2, \dots, 1$  do
3:    $\beta_i(t) = \sum_{j=1}^N \beta_j(t+1) A_{ij} B_{jx_{t+1}}$ 
4: end for
5: return  $\theta_i$ 
```

总结

在本次实验中, 我们学习了如何使用Baum–Welch算法从观测数据中训练HMM模型, 并通过维特比算法对隐状态序列进行推测。虽然大部分工作都是参考的维基百科中的内容, 但在实现的过程加深了对HMM的理解。在最终实验测试的脚本中, 测得准确率为64.7%。

参考文献

- [1] 李航. 统计学习方法. 清华大学出版社, 2012.
- [2] 吴军. 数学之美. 第2版. 人民邮电出版社, 2014.
- [3] 周志华. 机器学习. 清华大学出版社, 2016.