马尔科夫模型理论与其实际应用调研报告

曾舒立 PB19000200

1907 年，俄国数学家马尔科夫提出了马尔科夫过程，随后产生了马尔科夫链理论并进一步发展成为了隐马尔科夫模型（Hidden Markov Model），由于马尔科夫链和隐马尔科夫模型具有良好的数学基础作为支撑，拥有其他方法没有的优良性，至今仍是国内外学研究的热点。

本篇调研报告分为三个部分，第一部分将对马尔科夫模型的理论进行分析，介绍马尔科夫模型的基本概念，并对马尔科夫链模型的构造和模型的基本算法进行了总结与分析；第二部分介绍马尔科夫模型在某些具体问题上应用的方法；第三部分为个人对本次调研的收获与反思。

1 马尔科夫模型理论分析

1.1 马尔科夫过程

当一个系统（或过程）在某一具体时刻t_1所处的状态为已知条件，系统（或过程）在时刻$t > t_1$所处状态的条件概率分布与时刻t_1之间的状态都无关，而只与t_1时刻有关。换句话说，就是在已知系统（或过程）"现在"状态的条件下，"将来"与"过去"是相互独立的，其"将来"状态不依靠"过去"的状态。通常把这种性质称为马尔科夫性或无后效性。

马尔科夫性用概率分布函数来描述，可以表述为这样的形式：设随机过程$\{X(t), t \in T\}$的状态空间为 S。如果对于时间 t 的任意时刻，这里$t_1 < t_2 <\cdots< t_n, n \geq 3, t_i \in T$，在条件$X(t_i) = x_i, x_i \in S, i = 1,2,\cdots n-1$的情况下，$X(t_n)$的条件概率密度如果恰好等于$X(t_{n-1}) = x_{n-1}$时$X(t_n)$的条件概率分布函数，即为

$$P\{X(t_n) \leq x_n | X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_{n-1}) = x_{n-1}\} = P\{X(t_{n-1}) = x_{n-1}\}, x_n \in R$$

这时则称随机过程$\{X(t), t \in T\}$具有马尔科夫性或称为具有无后效性，进一步称这个过程为马尔科夫过程。

1.2 马尔科夫链

定义 1 设随机过程$\{X(t), t \in T\}, T = 0,1,2,\cdots$，状态空间$S = \{0,1,2 \cdots\}$。如果对任意正整数$m, n, p$及任意非负整数$j_m > j_{m-1} >\cdots> j_2 > j_1 (n > j_m)$以及$i_{n+p}, i_n, i_{j_m}, \cdots, i_{j_2}, i_{j_1}$有$P\left\{X(n+p) = i_{n+p} | X(n) = i_n, X(j_m) = i_{j_m}, \cdots, X(j_2) = i_{j_2}, X(j_1) = i_{j_1}\right\} = P\left\{X(n+p) = i_{n+p} | X(n) = i_n\right\}$成立，则称$X_T$为马尔科夫链。

定义 2 对于条件概率$P\{X(n+h) = j | X(n) = i\}$，即系统在$n$时处于状态$i$的条件下，（经过$h$步）在时刻$n+h$转移到状态$j$的条件概率，记为$p_{ij}(n, n+h)$，通常也简记为$p_{ij}^h(n)$，称为马尔科夫链的$h$步转移概率。特别地，当$h = 1$时，此时通常记$p_{ij}^1(n) = p_{ij}(n) = p_{ij}$并称$p_{ij}$为马尔科夫链的转移概率。当$p_{ij}(n, n+h) = p_{ij}(h)$时，称该转移概率具有平稳性，也称此时的马尔科夫链是齐次的或时齐的。目前在实际的应用中主要是对齐次的（时齐的）马尔科夫链的研究。

定义 3 马尔科夫链的转移矩阵

当$P^{(h)}$是由转移概率$p_{ij}^h(n)$组成的矩阵，且状态空间$S = \{0,1,2 \cdots\}$，即此时

$$P^{(h)}=\begin{bmatrix}p_{11}^{(h)}(n) & \cdots & p_{1m}^{(h)}(n)\\ \vdots & \ddots & \vdots \\ p_{m1}^{(h)}(n) & \cdots & p_{mm}^{(h)}(n)\end{bmatrix}$$

称为马尔科夫链的h步转移矩阵。此时当$h=1$时，可以写出马尔科夫链的一步转移矩阵P为

$$P=\begin{bmatrix}p_{11} & \cdots & p_{1n}\\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn}\end{bmatrix}$$

1.3 HMM 的产生与理论分析

1.3.1 隐马尔科夫模型的基本结构

隐马尔科夫模型是在马尔科夫链的基础上进一步发展而来的，也可以讲是对马尔科夫链理论的进一步扩充和提高。在马尔科夫链模型中，主要用到的是在有限的状态空间中，在确定了初始的状态概率分布后，然后运用转移矩阵来预测下一个状态的情况。而隐马尔科夫模型则人为的“规定”，在模型中的状态空间中的各个具体的状态是无法得到的，也就是“隐”（隐藏）的，只有这些状态所表现出来的现象是可以观测的（通常称为观测量）。例如，天气的情况可以分为“晴天”、“多云”、“下雨”这三个状态，而预测这三个状态出现的情况，则是根据观察到的海边海藻的表现状况，例如干的、稍干的、潮湿的、湿透的四种情况。把海藻的表现状况作为观测量，来进行预测，最终得到天气的具体状况。隐马尔科夫模型其实也可以看做是一个特殊的随机过程，由两部分组成：马尔科夫链和一般的随机过程。其中，马尔科夫链是用来表述系统初始概率分布和状态的转移（用状态转移矩阵来描述），一般的随机过程是用来表示（隐藏的）状态与实际的观测序列之间的关系，通过观测值的概率来进行描述。

1.3.2 隐马尔科夫模型的组成要素

通常情况下，一个标准的隐马尔科夫模型有一个五元组表示，即为$\lambda=(N,M,\pi,A,B)$，其中：

（1）用X来表示模型中的状态。状态的数量用N表示，状态之间存在着内在的联系。

（2）用Y来表示所有状态可能存在的观测序列。观测序列中观测值的数量用M表示。

（3）用$\pi=\{\pi_i\}, i=1,2,\cdots N$表示状态空间的初始概率分布，具体含义为第一个状态$q_1$取状态空间$X$中哪一个状态的概率。

（4）用$A=\left[a_{ij}\right]_{N\times N}$来表示状态转移矩阵，其中的$a_{ij}$表示在时刻$t-1$时，模型的状态为$x_i$，在时刻$t$时模型的状态转移到$x_j$。

（5）用$B=\left[b_j(i)\right]_{M\times N}$来表示观测序列的概率矩阵，其中的$b_j(i)$（也可以写为$b_{ij}$）表示状态$x_i$的条件下出现观测序列$Y_j$的概率。

一般情况下，通常把隐马尔科夫模型的五元组$\lambda=(N,M,\pi,A,B)$，将模型的状态数和观测序列的观测值数量忽略，而将模型简记为$\lambda=(\pi,A,B)$。

1.3.3 隐马尔科夫模型主要研究的问题

第一类问题：评价问题

在已知确定的隐马尔科夫模型$\lambda=(\pi,A,B)$和具体的观测序列$Y=$

$(y_1, y_2, \cdots, y_m)$的前提下，求出该模型产生出该观测序列的概率$p(Y|\lambda)$。具体的应用思想是在已知很多确定的因马尔科夫模型$\lambda_1, \lambda_2, \cdots$条件下，根据给定的观测序列$Y = (y_1, y_2, \cdots, y_m)$，来计算$p(Y|\lambda_i), i = 1,2, \cdots$的值，然后依据$p(Y|\lambda_i)$的最大值来确定最符合这个观测序列的隐马尔科夫模型，其实也就是模式匹配的一类问题。例如，这类问题成功的应用于语音信号的识别问题之中。

第二类问题：编码问题

在给定具体的隐马尔科夫模型$\lambda = (\pi, A, B)$和某一个确定的观测序列$Y = (y_1, y_2, \cdots, y_N)$，,通过观测序列来确定对应于该观测序列后面的状态序列，这就是编码问题的核心思想。解决该问题的主要依据是取产生该观测序列使得$p(Y|\lambda)$最大所对应的一组状态序列$X = (X_1, X_2, \cdots, X_N)$，即为实际中产生该观测序列所对应的实际状态序列。

第三类问题：学习问题

1.3.4 HMM 的基本算法分析

算法 1 评价问题的算法

在此类问题中，虽然观测序列Y是确定的，但是对应与该观测序列的状态序列$X = (X_1, X_2, \cdots, X_N)$却是不确定的，这是因为$x_i, i = 1,2, \cdots, N$的每一个值，都可以通过一定的概率出现给定的观测序列的值。对于任何一组状态序列$x_1, x_2, \cdots, x_N$产生观测序列$Y = (y_1, y_2, \cdots, y_N)$，都是遵循下面的联合概率分布的：

$$\pi_{x_1} b_{x_1}(y_1) a_{x_1 x_2} b_{x_2}(y_2) a_{x_2 x_3} \cdots a_{x_{n-1} x_n} b_{x_n}(x_N)$$

其中π_{x_1}是状态x_1的初始分布概率，$b_{x_1}(y_1)$是模型在x_1状态下产生观测值y_1的概率，$a_{x_1 x_2}$是模型由状态x_1转移到状态x_2的概率。通常采用下面的前向算法和后向算法。

算法 1.1 前向算法

这里引入前向因子$\alpha_t(i)$，令$\alpha_t(i) = p(y_1, y_2, \cdots, y_t, q_t = x_i|\lambda)$，表示在给定的隐马尔科夫模型$\lambda$下，整个系统在到达$t$时刻，观测序列为$y_1 y_2 \cdots y_t$，且$t$时刻的状态为$x_i$时的部分概率。则算法的具体描述如下：

Step1 输入：初始化，$\alpha_t(i) = \pi_i b_i(Y_1)$，其中$1 \le i \le N$

Step2 处理过程：进行递推处理，从前向后，逐步递推

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(Y_{t+1})$$

Step3 输出：终止条件，当$t = T$时，算法结束，

$$p(Y|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

对前向算法分析可得，该算法时间复杂度为$O(N^2 T)$。

算法 1.2 后向算法

这里引入后向因子$\beta_t(i)$，令$\beta_t(i) = p(y_t, y_{t+1}, \cdots, y_T, q_t = x_i|\lambda)$，表示在时刻$t$,系统的状态为$x_i$时，由时刻$t+1$到整个观测序列结束即输出序列为$y_t, y_{t+1}, \cdots, y_T$的部分概率。算法的具体描述如下：

Step1 输入：初始化，$\beta_t(i) = 1$，其中$1 \le i \le N$

Step2 处理过程：从后向前进行递推

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(Y_{t+1}) \beta_{t+1}(j)$$

Step3 输出：终止条件

$$p(Y|\lambda) = \sum_{i=1}^{N} \beta_1(i)$$

算法 2 处理解码问题的算法

在给定具体的隐马尔科夫模型$\lambda=(\pi,A,B)$和某一个确定的观测序列$Y=(y_1,y_2,\cdots,y_N)$，根据在那个状态序列下，该观测序列出现的概率最大的原则下，来确定出这个具体的状态序列$X=(X_1,X_2,\cdots,X_N)$(即编码。这类算法通常被叫做 Viterbi 算法，该算法采用动态规划的算法思想，复杂度为$O(K^2L)$，其中K和L分别为状态个数和序列长度。这里令$\delta_t(i)=\max_{x_1x_2,\cdots x_N} p(x_1,x_2,\cdots,x_{t-1},x_t=i,y_1,y_2,\cdots,y_N|\lambda)$，并引用$\phi_t(i)$来作为状态标记。Viterbi 算法的最终目的就是找到在最后的 T 时刻，找到最大的$\delta_T(i)$所表示的那个具体的状态序列。具体的算法如下：

Step1 输入：初始化$\delta_1(i)=\pi_i b_i(Y_1)$，$\phi_1(i)=0$，其中$1\le i\le N$

Step2 递推：

$$\delta_t(j)=\max_{1\le i\le N}\left[\delta_{t-1}(j)a_{ij}\right]b_j(Y_j),\phi_t(j)=\underset{1\le i\le N}{\operatorname{argmax}}\left[\delta_{t-1}(i)a_{ij}\right]$$

Step3 输出：到达T时刻时，算法终止得到状态序列最优总概率：$p^*=\max_{1\le i\le N}[\delta_T(i)]$，确定序列$Y_T^*=\underset{1\le i\le N}{\operatorname{argmax}}[\delta_T(i)]$，然后根据$Y_t^*=\phi_{t+1}(Y_{t+1}^*),t=T-1,T-2,\cdots 1$来确定最终的状态序列。最后可以得到最优的状态序列为$Y_T^*=\{y_1^*,y_2^*,\cdots y_T^*\}$

2 马尔科夫模型的应用

2.1 马尔科夫模型在 WEB 上的应用

由于网页加载速度过慢，人们往往要忍受长时间的等待，而网页缓存技术和网页预取技术可以有效减少人们的等待时间。但是，获取用户的访问行为方式是一个难题，不过，将马尔科夫模型应用于预测用户访问方式成为了一个很好的方法。具体方法为：根据服务器上的日志目录，统计每个网页 URL 的请求次数，然后根据页面之间的切换频率来构造转移矩阵，生成马尔科夫模型，并根据用户实际某一阶段的网页浏览顺序来预测下一个时间段的用户网页的访问趋势，从而主动地进行网页的预取，缓存处理，这样就能提高网页的访问效率，减少平均时延。

2.2 马尔科夫模型在股票预测方面的应用

在股票价格方面，可以利用马尔科夫链对股票的价格（指数）的走势进行预测。按照股票价格具体的上涨幅度和下降幅度，将股票价格分为下跌、持平、上涨三个状态，将前一阶段的某一支或某几支股票的价格走势作为训练集来构建马尔科夫模型，根据股票的三个状态出现的频率构造转移矩阵，将具体的某一个交易日的股票价格作为初始的概率分布，从而来进行未来的股票价格的走势的预测。通过马尔科夫理论的平稳方程可以验证模型的预测正确性，得出了在排除主观因素的影响外，马尔科夫模型在对具体股票价格近期走势的分析，其预测的可信度很高。

2.3 马尔科夫模型在教育方面的应用

马尔科夫模型应用在教师的教学评价、师资队伍建设和度量学生的问题解决能力等教育方面。为了客观的评价教师的教学水平和所达到的效果，马尔科夫模型也应用到其中，首先把教师的一个长期的教学过程分成若干的小过程，记为不同的时刻且把考生的成绩按具体的分数分成优、良、中、及格、不及格等及格状态。然后统计学生从一时刻

到下一时刻成绩的转移次数，从而构造转移矩阵，再根据稳态方程求出初始变量构造模型。为了评价不同教师的教学效果，可以确定一个评价标准。马尔科夫模型在测量学生问题的解决能力方面有着具体的应用。该模型根据教育心理学和数学教育心理学的研究，将学生解决问题的能力分为分析、设计方案、求解实施、验证四个状态，构造了马尔科夫模型，并根据得到的实际结果，分析问题的源头，采取具体的补救措施，进一步提高学生解决问题的能力。

3 收获与反思

马尔科夫模型是一种作用十分强大并且贴合实际的模型，它的实际意义在于，对于那些具有马尔科夫性且转移概率具有平稳性的实际过程，马尔科夫模型或隐马尔科夫模型可以很好的进行建模，这对于算法的实现有着重要的帮助。构造马尔科夫模型的关键在于先发现过程的具体状态，通过利用转移概率来对未来进行预测。在第二部分几种实际应用中可以发现，往往能建立马尔科夫模型的实际问题，总是有几种确定的状态，我们要找到这些状态并罗列出来。并且往往这些实际过程都是无后效性的，即“将来”与“过去”是相互独立的。

当模型中的状态空间的各个具体状态无法观察得到时，我们就可以建立隐马尔科夫模型，将这些状态量用其他可观测的观测量代替，再建立观测量与原来无法测得（隐藏）的状态量的关系（用状态转移矩阵来描述），成功建立隐马尔科夫模型。总而言之，隐马尔科夫模型就是马尔科夫模型在碰到一些实际问题时无法建模时可以采用的方法，去掉隐马尔科夫模型的外表，其本质就是马尔科夫模型。

本次调研使我了解了马尔科夫过程、马尔科夫链、马尔科夫模型、隐马尔科夫模型以及一些应用与具体的算法，学习到许多具体的知识，同时，在调研过程中意识自己深深的不足，比如报告中关于隐马尔科夫模型解决学习问题的方法没有给出，是因为参考文献有关这个问题的内容超出了我的知识储备。另外，不仅仅调研了书本上的知识，还学会了如何去找文献、看文献、写调研报告，这些都是本次调研的价值体现。

参考文献：

【1】何成刚.马尔科夫模型预测方法的研究及其应用【D】.安徽：安徽大学，2011：1-19.