第8卷 第3期
1995年10月
盐城工业专科学校学报
Journal of Yancheng College of Technology
Vol. 8 No. 3
Oct. 1995

应用马尔可夫模型预测跨世纪人才结构

罗家仁　　王　玮
（南京审计学院，南京，210029）

摘　要　本文在统计（1989—1994）南京市某银行系统职员各学历层次的人才数的基础上，应用马氏模型理论，进行分析、预测，并推出修正方法。

关键词　马尔可夫过程　转移概率矩阵　主观概率法

1　马尔可夫随机过程简介

马尔可夫过程是一种动态随机数学模型，它是指某一系统未来时刻的状态只与现在时刻状态有关，而与过去历史状态无直接关系（即“无后效性”）。当时间与状态都取非负整数（以 E 表示）的马氏过程称为马氏链。

对于一个马氏链，称 $P_{ij}(n)$ 为系统由状态 i 经过几个时间间隔转移到状态 j 的转移概率，称 $P(n)=[P_{ij}(n)]$ 为马氏链的 n 步转移矩阵，当 $n=1$ 时，记 $P(1)=P=(P_{ij})$ 为马氏链一步转移矩阵。

马氏链转移矩阵具有下列性质：

(1) 一步转移矩阵 P 有 $0\leqslant P_{ij}\leqslant 1$，且 $\sum_j P_{ij}=1(i,j\in E)$ 成立。

(2) 若系统互斥的状态有 m 个，设 $S(0)$ 为系统的初始状态向量，$S(n)$ 为系统经过 n 次转移后的状态向量，由柯尔莫哥夫——切普曼定理可推得，$S(n+1)=S(n)\cdot P$，且在 P 保持不变条件下，有 $S(n+1)=S(0)\cdot P^{n+1}$ 成立。

2　应用实例

2.1　原始资料与分析

我们对南京市某银行系统职员（简称“银职”）6 个年度（1989—1994）的人才数按中专及高、初中生，大专生，本科生，研究生四个学历层次统计归类成表 1。由表 1 计算出各学历人数在总数中所占的比例得人才比例结构表 2。

表 1　某银职学历层次分析统计

年份	中专及高初中生	大专生	本科生	研究生	合计（人数）
1989	3285	516	41	0	3842
1990	3269	602	51	1	3923
1991	3166	785	81	2	4034
1992	3048	1030	111	4	4193
1993	2979	1261	136	8	4384
1994	2876	1415	156	12	4459

表 2　某银职学历层次人才比例

年份	中专及高初中生	大专生	本科生	研究生	合计（人数）
1989	0.8550	0.1343	0.0107	0	1.000
1990	0.8333	0.1535	0.0130	0.0002	1.000
1991	0.7848	0.1946	0.0201	0.0005	1.000
1992	0.7269	0.2456	0.0265	0.0010	1.000
1993	0.6796	0.2876	0.0310	0.0018	1.000
1994	0.6450	0.3173	0.0350	0.0027	1.000

从表 1、表 2 之中可以看出，(1) 现阶段“银职”人才总数中，高学历者尤其是研究生、本科生的

比例极低，而中专、高初中等低学历者占绝对优势。(2)这种不合需求的人才结构已发生初步的变化：高学历层次的比例在逐年增加，低学历层次比例在逐年减少。笔者认为其原因不外乎下述两个因素：(1)部分职员通过各种途径(如培训、成人高校、电大、函大、自考等)提高了文化水平，由低学历转为高学历；(2)新增加的职员中高学历者相对增多。

可以预料，随着国民经济建设的持续稳步增长，在今后若干年内，"银职"人才队伍不但在数量上将有所增长，而且在结构上将进一步变化，高学历比例有可能占绝对优势。

2.2 定量预测与分析

用马尔可夫模型对2000年后的跨世纪人才结构进行定量预测与分析，其步骤如下：

(1)确定系统状态 我们把"银职"的人才队伍看成一个系统，其上述四个层次每年的人数占总数中的比例视为一组"状态"，且用四维向量 $S(n)=(S_1(n),S_2(n),S_3(n),S_4(n))$ 表示，令 $n=t-1989$，t 为年份，一年为一期，由表2可知，各期状态向量为

$S(0)=(0.8550,\quad 0.1343,\quad 0.0107,\quad 0\quad)$

$S(1)=(0.8333,\quad 0.1535,\quad 0.0130,\quad 0.0002)$

$S(2)=(0.7848,\quad 0.1946,\quad 0.0201,\quad 0.0005)$

$S(3)=(0.7269,\quad 0.2456,\quad 0.0265,\quad 0.0010)$

$S(4)=(0.6796,\quad 0.2876,\quad 0.0310,\quad 0.0018)$

$S(5)=(0.6450,\quad 0.3173,\quad 0.0350,\quad 0.0027)$

当"银职"人才比例结构从一种状态变为另一种状态时，称之为"状态转移"，用转移概率 P_{ij} 与转移矩阵来描述。

(2)确定假设 ①预测期间，状态向量的维数保持不变；②转移矩阵逐年保持不变；③状态转移仅受前一时期的影响。

(3)确定转移矩阵 本文试使用解线性方程组和主观概率法相结合的方法。设所求一步转移矩阵为

$$P=\begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix}$$

利用公式 $S(n+1)=S(n)\cdot P$ 及 $\sum_{j=1}^{4}P_{ij}=1(i=1,2,3,4)$ 的方程组

$$\begin{cases} P_{11}+P_{12}+P_{13}+P_{14}=1 \\ P_{21}+P_{22}+P_{23}+P_{24}=1 \\ P_{31}+P_{32}+P_{33}+P_{34}=1 \\ P_{41}+P_{42}+P_{43}+P_{44}=1 \\ S_1(n)P_{11}+S_2(n)P_{21}+S_3(n)P_{31}+S_4(n)P_{41}=S_1(n+1) \\ S_1(n)P_{12}+S_2(n)P_{22}+S_3(n)P_{32}+S_4(n)P_{42}=S_2(n+1) \\ S_1(n)P_{13}+S_2(n)P_{23}+S_3(n)P_{33}+S_4(n)P_{43}=S_3(n+1) \\ S_1(n)P_{14}+S_2(n)P_{24}+S_3(n)P_{34}+S_4(n)P_{44}=S_4(n+1) \end{cases}$$

因线性方程组所含方程的个数少于未知量的个数，所以方程组有无穷多组解，下面要选出一组

满足 $0\leqslant P_{ij}\leqslant 1$ 的解。由于跨世纪人才结构变化只能是高学历者比例增大，低学历比例者减少，反之不然，故为简便，采取主观概率法选取 $P_{21}=0,P_{31}=0,P_{41}=0,P_{32}=0,P_{42}=0,P_{43}=0$，另外，假定低层次者在进一步状态转移中只能提高一个层次，这样便认定 $P_{13}=0,P_{14}=0,P_{24}=0$。于是，所求矩阵简化为

$$P=\begin{bmatrix} P_{11} & P_{12} & 0 & 0 \\ 0 & P_{22} & P_{23} & 0 \\ 0 & 0 & P_{33} & P_{34} \\ 0 & 0 & 0 & P_{44} \end{bmatrix}$$

相应的方程也简化为

$$P=\begin{cases} P_{11}+P_{12}=1 \\ P_{22}+P_{23}=1 \\ P_{33}+P_{34}=1 \\ P_{44}=1 \\ S_1(n)P_{11}=S_1(n+1) \\ S_1(n)P_{12}+S_2(n)P_{22}=S_2(n+1) \\ S_2(n)P_{23}+S_3(n)P_{33}=S_3(n+1) \\ S_3(n)P_{34}+S_4(n)P_{44}=S_4(n+1) \end{cases}$$

解之，得

$$P_{n}=\frac{S_1(n+1)}{S_1(n)},\quad P_{12}=1-P_{11},\quad P_{44}=1$$

$$P_{34}=\frac{S_4(n+1)-S_4(n)}{S_3(n)},\quad P_{33}=1-F_{34}$$

$$P_{23}=\frac{S_3(n+1)-S_3(n)P_{33}}{S_2(n)},\quad P_{22}=1-P_{23}$$

对 $n=0,1,2,3,4$，将得到5个可能不同值的 $P_{11}(n),P_{34}(n),P_{23}(n)$，为了能满足上述假设(2)，各作 P_{11},P_{34},P_{23}的平均值，即

$$P_{11}=\frac{1}{5}\sum_{n=0}^{4}P_{11}(n)=\frac{1}{5}\sum_{n=0}^{4}\frac{S_1(n+1)}{S_1(n)},$$

$$P_{34}=\frac{1}{5}\sum_{n=0}^{4}P_{34}(n)=\frac{1}{5}\sum_{n=0}^{4}\frac{S_4(n+1)-S_4(n)}{S_3(n)},$$

$$P_{23}=\frac{1}{5}\sum_{n=0}^{4}P_{23}(n)=\frac{1}{5}\sum_{n=0}^{4}\frac{S_3(n+1)-S_3(n)P_{33}}{S_2(n)}$$

具体算得结果 $P_{11}=0.945,P_{34}=0.025,P_{23}=0.028$。于是所求的一步转移矩阵为

$$P=\begin{bmatrix} 0.945 & 0.055 & 0 & 0 \\ 0 & 0.972 & 0.028 & 0 \\ 0 & 0 & 0.975 & 0.025 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(4)预测计算　以1994年的状态向量 $S(5)$为起点，运用公式 $S(n+1)=S(n)P$，就可计算得1995年至2000年后的各年预测值，列成表3。

表3 1995—2000年后各年预测值

年份	中专及高初中生	大专生	本科生	研究生
1995	0.6095	0.3439	0.0430	0.0036
1996	0.5760	0.3677	0.0516	0.0047
1997	0.5443	0.3891	0.0606	0.0060
1998	0.5144	0.4081	0.0700	0.0075
1999	0.4861	0.4250	0.0797	0.0092
2000	0.4594	0.4398	0.0896	0.0112
2001	0.4341	0.4528	0.0997	0.0134
2002	0.4102	0.4640	0.1099	0.0159
2003	0.3876	0.4736	0.1201	0.0187
…	…	…	…	…

表4 预测结果

年份	中专及高初中生	大专生	本科生	研究生
1995	0.6127	0.3423	0.0415	0.0035
1996	0.5821	0.3650	0.0484	0.0045
1997	0.5530	0.3857	0.0556	0.0057
1998	0.5254	0.4045	0.0631	0.0070
1999	0.4991	0.4215	0.0709	0.0085
2000	0.4741	0.4368	0.0789	0.0102
2001	0.4504	0.4505	0.0870	0.0121
2002	0.4279	0.4626	0.0953	0.0142
2003	0.4065	0.4733	0.1037	0.0165
…	…	…	…	…

(5)预测结果分析和修正 实际应用中，往往需要通过不断调整初始状态和转移矩阵，最后得出一个较为客观的预测结论。

由于我们选取了 $P_{21}=0,P_{31}=0,P_{32}=0,P_{41}=0,P_{42}=0,P_{43}=0$，这就可能使 P_{23},P_{33},P_{34} 之值偏大，从而可能导致 $S_3(n),S_4(n)$ 的预测值偏大，$S_1(n)$ 的预测值偏小，为此，再次用主观概率法对 P 进行修正。使

$$P_{11}=\frac{1}{4}\left[\sum_{n=0}^{4}P_{11}(n)-\min\{P_{11}(n),n=0,1,2,3,4\}\right]$$

$$P_{23}=\frac{1}{4}\left[\sum_{n=0}^{4}P_{23}(n)-\max\{P_{23}(n),n=0,1,2,3,4\}\right]$$

$$P_{34}=\frac{1}{4}\left[\sum_{n=0}^{4}P_{34}(n)-\max\{P_{34}(n),n=0,1,2,3,4\}\right]$$

P_{44} 仍保持不变。

从而，修正后的转移矩阵为

$$P=\begin{bmatrix}0.9500 & 0.050 & 0 & 0\\ 0 & 0.977 & 0.023 & 0\\ 0 & 0 & 0.0976 & 0.024\\ 0 & 0 & 0 & 1\end{bmatrix}$$

相应的预测结果为表4。

从表4预测结果可知，跨入21世纪后，该银行系统人才结构将发生质的变化，低学历层次将从1989年的绝对主体地位降为从属地位，代之以大专学历为主，本科生、研究生学历呈上升趋势的高层次结构，这是符合建设一支素质优良，数量适宜，结构优化，充满活力，具有国际视野，广博学识的金融队伍所需要的。

(上接第127页)

参考文献

1 大学数学法. 机械工业教育编辑部

2 奚从清，方立明. 大学生逆反心理浅析. 心理学. B4，1987，2

3 高金衡. 话说高等数学的讲课. 东南大学高等教育学报. 1995，1