第 23 卷第 4 期
2006 年 8 月

华 东 交 通 大 学 学 报
Journal of East China Jiaotong University

Vol.23 No.4
Aug., 2006

文章编号:1005-0523(2006)04-0051-04

用马氏链方法预测全国年发电量趋势

何 鑫,宋平岗,官二勇

(华东交通大学 电气与电子工程学院,江西 南昌 330013)

摘要:马尔科夫链是一种随机过程的数学模型,利用 1986 年到 2004 年的年发电量统计数据,采用马氏链方法从理论上对今后 2 年的年发电量趋势进行了预报.结果表明,按照目前增长速度,2006 年发电量将突破 25000 亿千瓦时.

关 键 词:发电量;马氏链;产量预测;转移矩阵

中图分类号:TM60,O211.62 **文献标识码**:A

1 引言

随着我国经济的迅猛发展和国家电力体制改革的不断深入,各大主要工业,如钢铁、冶炼、矿山等特大型企业对电力的需求日益增加,电力的使用已渗透到社会经济、生活的各个领域.由于电力具有便于转换能源形式,能高度集中和无限划分,清洁干净和易于控制,可大规模生产和远距离输送等特性,使电力发展和应用的程度,即一个国家的电气化程度,成了衡量其社会现代化水平高低以及物质文明和精神文明高低的重要标志之一.但电力作为一种商品,也将面对能源消费市场的激烈竞争.分析、预测电力在能源市场的占有率的发展趋势,可以明确电力在能源消费市场的地位,也可以估计煤炭、石油、天然气等可替代产品的发展变化速度,从而有效地帮助电力工业的各部门主动地采取措施提高市场占有率,达到其产品在消费市场的优势.由于受总的装机容量的限制,我国的电力供应与实际的工业需求存在着很严重的缺口.为了满足工业需求,"十一五"期间,国家增大了对电力项目的投资和宏观调控,但是对具体调控的依据还缺少相应的理论指导.

本文提出了应用马尔科夫链的方法,借助于对往年发电量的数据统计,来科学的预测未来两年国家对电量的需求.希望该方法对国家制订立项规划和宏观调控提供一定的理论依据和参考.

2 马尔科夫链基本原理[1-4]

2.1 马尔科夫过程与马尔科夫链

马尔科夫过程是具有无后效性的随机过程,这是马氏链最基本的一条性质.后效性是指:当过程在 t_m 时刻所处的状态为已知时,过程在大于 t_m 的时刻 t 所处的状态的概率特性只与过程在 t_m 时刻所处的状态有关,而与过程在 t_m 时刻以前的状态无关.

马尔科夫链:时间离散,状态离散的马尔科夫过程,简称马尔科夫链.

定义:设随机序列 $\{x(n), n=1,2,...\}$ 的离散状态空间为 E.若对于任意 m 个非负整数 $n_1, n_2, ..., n_m, 0\leq n_1<n_2<\cdots<n_m$)和任意自然数 k,以及任意 $i_1, i_2, \cdots, i_m, j\in E$,满足:

$$p\{x(n_m+k)=j|x(n_1)=i_1, x(n_2)=i_2, \cdots, x(n_m)=i_2, \cdots, x(n_m)=i_m\}=p(x(n_m+k)=j|x(n_m=i_m\} \quad (1)$$

则称随机序列 $\{x(n), n=1,2,\cdots\}$ 为马尔科夫链.

2.2 转移概率

设系统的状态有 n 个,系统在 t_m 时间处于状态 i,在下一时间 t_{m+1} 时转为状态 j 的概率为 p_{ij},则称 p_{ij} 为一步转移概率.在这里 p_{ij} 只与它在 t_m 时刻所处的状态 i 以及 t_{m+1} 时刻所处的状态 j 有关,而与 t_m 以前的任何历史状态无关.将这些 p_{ij} 依序排列,即构成一步转移概率矩阵 P:

收稿日期:2005-12-20

作者简介:何 鑫(1982-),男,江西人,硕士研究生,研究方向是电力电子技术及应用.

$$P=(p_{ij})=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{2}$$

一步转移矩阵有两个重要性质：1，矩阵中每个元素 P_{ij} 均为非负的，即 $p_{ij}\geq 0$，2，矩阵中每行元素相加其和为1，即．若系统在 t_m 时间处于状态 i，经过 K 次转移后在 t_{m+k} 时间处于状态 j 的概率为 p_{ij}，则 $p_{ij}(k)$ 称为 k 步转移概率，记作 $p(k)$．可证 $x(k+1)=x(k)^{*}p$．这样就可以根据现在的已知量来科学的预测未来未知量．

3 用马尔科夫链对发电量进行预报

某一时间的发电量状态，仅仅依赖于前一时间的发电量状态，这样就把发电量状态作为马尔科夫链看待．把发电量序列看作一随机过程，时间序列是 $t_1, t_2, \cdots, t_n\cdots$，与其对应的发电量状态是 $X_{t_1}, X_{t_2}, \cdots, X_{t_n}, \cdots$．在 $t_2>t_1$ 时，X_{t_2} 只与前一时间 t_1 的状态 X_{t_1} 有关．也就是说，在已知“现在”的条件下，“将来”的状态与“过去”的状态无直接关系，即所谓“无后效性”．

发电量序列的变化，从 t_1 时间的状态 S_1 只能推演在 t_2 时间处于状态 S_2 的概率分布．将发电量分为若干状态，记为 $S_1, S_2, \cdots, S_n$．产量序列中状态 S_i 经过 m 步转移到 S_j 的概率为 m 步转移概率 $p_{ij}^{(m)}$，m 步转移矩阵为 $P^{(m)}=(P_{ij}^{(m)})_{n\times n}$ $(m=1,2,3,\cdots)$．已知 X_t 的产量分布，则可推知 X_{t+m} 的产量分布为 $X_t\cdot P^m$．概率转移矩阵通常由频率转移矩阵取代[5]，计算公式为：

$$P_{ij}^{(m)}=\frac{M_{ij}}{M_i}(m=1,2,\cdots) \tag{3}$$

其中，M_i 为状态 S_i 出现的次数；M_{ij} 为状态 S_i 经过 m 步转移到状态 S_j 的次数．根据以上原理，为了便于状态的划分，采用了发电量增量百分比序列作为马氏链，统计出各种状态的转移概率，这样就可利用前一年的发电量预测下一年的发电量增量．

(1) 对发电量的原始数据[6]进行处理，得到增量序列，见表1

表1 相对年增量百分比(马氏链)状态表

年份	序号	实际发电量/亿千瓦时	增量百分比/%	年份	序号	实际发电量/亿千瓦时	增量百分比/%
1986	0	4496	—	1996	10	10650	7.1
1987	1	4973	10.6	1997	11	11198	5.1
1988	2	5451	9.6	1998	12	11431	2.1
1989	3	5847	7.3	1999	13	12176	6.5
1990	4	6213	6.3	2000	14	13510	11
1991	5	6775	9.0	2001	15	14706	8.9
1992	6	7542	11.3	2002	16	16548	12.5
1993	7	8384	11.2	2003	17	19080	15.3
1994	8	9138	9.0	2004	18	21870	14.6
1995	9	9942	8.8	—	—	—	—

(2) 对发电量增量百分比序列进行状态划分，见表2

表2 增量百分比序列取值状态

年增量百分比	0～4%	4%～6%	6%～8%	8%～10%	10%～12%	12%～14%	14%～16%
状态	*A*	*B*	*C*	*D*	*E*	*F*	*G*

(3) 根据划分状态，1986－2004年增量百分比构成时间序列的一个实际状态序列见表3

表3 相对年增量百分比(马氏链)状态表

年份	序号	状态	年份	序号	状态
1986	0	—	1996	10	*C*
1987	1	*E*	1997	11	*B*
1988	2	*D*	1998	12	*A*
1989	3	*C*	1999	13	*C*
1990	4	*C*	2000	14	*E*
1991	5	*D*	2001	15	*D*
1992	6	*E*	2002	16	*F*
1993	7	*E*	2003	17	*G*
1994	8	*D*	2004	18	*G*
1995	9	*D*			

(4) 根据表3实际状态用公式(1)计算1－8阶转移概率，构成1－8阶转移矩阵．

$$P^{(1)}=\begin{bmatrix}0&0&1&0&0&0&0\\1&0&0&0&0&0&0\\0&1/4&1/4&1/4&1/4&0&0\\0&0&2/5&1/5&1/5&1/5&0\\0&0&0&3/4&1/4&0&0\\0&0&0&0&0&0&1\\0&0&0&0&0&0&1\end{bmatrix}\quad P^{(2)}=\begin{bmatrix}0&0&0&0&1&0&0\\0&0&1&0&0&0&0\\1/4&0&0&1/2&1/4&0&0\\0&1/5&2/5&0&1/5&0&1/5\\0&0&1/4&1/2&0&1/4&0\\0&0&0&0&0&0&1\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}$$

$$P^{(3)}=\begin{bmatrix}0&0&0&1&0&0&0\\0&0&0&0&1&0&0\\0&0&1/4&0&1/2&1/4&0\\1/5&1/5&0&2/5&0&0&1/5\\0&0&1/2&1/4&0&0&1/4\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1.7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}\quad P^{(4)}=\begin{bmatrix}0&0&0&0&0&1&0\\0&0&0&1&0&0&0\\0&0&0&1/4&1/2&0&1/4\\1/4&0&1/4&1/4&1/4&0&0\\0&1/4&1/4&1/4&0&0&1/4\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}$$

$$P^{(5)}=\begin{bmatrix}0&0&0&0&0&0&1\\0&0&0&0&0&1&0\\0&0&0&3/4&0&0&1/4\\0&0&1/2&0&1/2&0&0\\1/3&1/3&0&0&1/3&0&0\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1.7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}\quad P^{(6)}=\begin{bmatrix}0&0&0&0&0&0&1\\0&0&0&1&0&0&1\\0&0&1/3&1/3&0&1/3&0\\0&1/4&0&1/2&1/4&0&0\\1/3&0&1/3&0&1/3&0&0\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}$$

$$P^{(7)}=\begin{bmatrix}1/7&1/7&1/7&1/7&1/7&1/7&1/7\\0&0&0&0&0&0&1\\0&0&1/3&1/3&0&0&1/3\\1/4&0&0&1/2&0&1/4&0\\0&0&1/3&1/3&1/3&0&0\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}\quad P^{(8)}=\begin{bmatrix}1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/3&1/3&0&0&0&0&1/3\\0&0&1/2&0&0&1/4&1/4\\0&0&0&2/3&1/3&0&0\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\\1/7&1/7&1/7&1/7&1/7&1/7&1/7\end{bmatrix}$$

(5) 利用各转移矩阵,编制 2005 年发电量增量百分比的概率分布表,见表 4.

表 4　2005 年发电量增量百分比概率分布表

起始年	起始	转移	状态概率分布						
序号	状态	步数	*A*	*B*	*C*	*D*	*E*	*F*	*G*
18	*G*	1	0	0	0	0	0	0	1
17	*G*	2	1/7	1/7	1/7	1/7	1/7	1/7	1/7
16	*F*	3	1/7	1/7	1/7	1/7	1/7	1/7	1/7
15	*D*	4	1/4	0	1/4	1/4	1/4	0	0
14	*E*	5	1/3	1/3	0	0	1/3	0	0
13	*C*	6	0	0	1/3	1/3	0	1/3	0
12	*A*	7	1/7	1/7	1/7	1/7	1/7	1/7	1/7
11	*B*	8	1/7	1/7	1/7	1/7	1/7	1/7	1/7
	合计		97/84	19/21	97/84	97/84	97/84	19/21	11/7
	平均概率分布		0.144	0.113	0.144	0.144	0.144	0.113	0.196

表 4 的编制方法:表 4 中年序号为 18 的第一行起始状态为 G,转移步数为 1,则在一阶转移矩阵中取第 7 行;年序号为 16 的第三行起始状态为 F,转移步数为 3,则在三阶转移矩阵中取第 6 行,依次类推,然后求和,最后用合计栏中的数字除 8(8 为表中所有的转移步数的种数),即得到平均概率分布栏.表 4 的说明:起始年序号为 18 的第一行中右边的

状态概率分布栏,即为由 2004 年的增量状态通过一步转移得到的 2005 年增量百分比的概率分布.第二行即为由 2003 年的状态通过二步转移得到的 2005 年增量百分比的概率分布,依次类推.最后一行得到 2005 年增量百分比的平均概率分布.

用离散随机变量 X 取代 2005 年发电增量百分比,其中 X 的概率分布见表 5

表 5 X 的概率分布

取值	0.02	0.05	0.07	0.09	0.11	0.13	0.15
概率	0.144	0.114	0.144	0.144	0.144	0.114	0.196

估计期望值 $E=0.02\times0.144+0.05\times0.114+0.07\times0.144+0.09\times0.144+0.11\times0.144+0.13\times0.114+0.15\times0.196=0.0917$,即得到 2005 年发电量的增量百分比的期望值为 9.2%.

同理,可以求得 2006 的发电量增量百分比.

表 6 2006 年发电量增量百分比概率分布表

取值	0.02	0.05	0.07	0.09	0.11	0.13	0.15
概率	0.129	0.082	0.248	0.129	0.201	0.082	0.129

于是可以得到 2006 年发电量百分比增量估计期望值为 8.4%.

表 7 发电量预测值

年份	实际发电量/亿千瓦时	增量百分比/%
2005	23882	9.2
2006	25984	8.8

4 结论

本文利用 1986 年到 2004 年的年发电量统计数据,采用马氏链方法从理论上对今后 2 年的年发电量趋势进行了预报,并利用马尔科夫模型计算出 2005 年和 2006 年全国发电量的预测值,可见我国 2005 年发电量将递增 9% 以上,预计在 2006 年我国年发电总量将突破 25000 亿千瓦时.该模型计算简便、应用广泛.预测准确,是做各种预测分析的一种较好的方法,值得我们继续研究和应用.

参考文献:

[1] 孙荣恒.随机过程及其应用[M].北京:清华大学出版社,2003.

[2] 吕一兵,彭惠明,张 超,等.用马氏链方法预测塔里木油田近期石油年产量趋势[J].西安石油大学学报(自然科学版),2004, 19(4):77—79.

[3] 王梓坤.生灭过程与马尔科夫链[M].北京:科学出版社,1980.

[4] 胡迪鹤.可数状态的马尔科夫过程论[M].武汉:武汉大学出版社,1983.

[5] 胡迪鹤.随机过程论[M].武汉:武汉大学出版社,2000.

[6] 中国统计局.中国统计年鉴[M].北京:中国统计出版社,2005.

Forecasting the Trend of Electricity Production in the Year Future by Using Markov Chain

HE Xin, SONG Ping-gang, GUAN Er-yong

(School of Electronical and Electrical Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Markov chain is a kind of mathematical model of random process, and it is widely applied in various kinds of forecasting technologies. Based on the data of electricity production from 1986 to 2004, the trend of the electricity production is forecasted in the following two years through Markov Chain. The predicted electricity production of in 2006 breaks through 6 million tons

Key words: electricity production; markov chain; production prediction; transition matrix