

马尔可夫链在地价预测模型中的应用研究

刘俊¹, 田崇新², 张小燕³, 郑光辉¹

(1. 南京师范大学地理科学学院, 210097, 江苏, 南京)

(2. 江苏金宁达不动产评估咨询有限公司, 210036, 江苏, 南京)

(3. 江苏易图通信息工程技术有限公司, 210036, 江苏, 南京)

[摘要] 在地价的预警预报研究过程中, 过去的理论主要存在3点不足, (1)往往没有建立科学完善的数学模型, 也没有用多种数学模型相互验证预测结论; (2)没有充分利用现有的信息, 例如地价监测体系的信息. (3)在实际应用中也存在预测地价水平与城市实际地价水平、预测走势与实际变化有偏差的问题. 解决这些问题通过以下方法: 充分利用现有的准确的地价监测体系的信息, 并在此基础上建立时间序列模型、空间分布预测模型及马尔可夫链地价预测模型; 以不动产周期理论和专家经验知识库检验经过数学模型测算的结论. 马尔可夫链地价预测模型是概率分布式的预测模型, 它专门用于分析系统中不确定性因素的概率分布, 经过检验可以认为这个模型适用于对地价变化进行科学的概率预测.

[关键词] 土地价格, 马尔可夫链, 预测模型

[中图分类号] F301.3, **[文献标识码]** B, **[文章编号]** 1001-4616(2005)03-0121-06

Research of Markov Chain Application in Land Price Warning Model

Liu Jun¹, Tian Chongxin², Zhang Xiaoyan³, Zheng Guanghui¹

(1. School of Geographical Science, Nanjing Normal University, 210097, Nanjing, China)

(2. Jiangsu Jin Ning Da Real-estate Evaluation and Consultation Corporation, 210036, Nanjing, China)

(3. Jiangsu E-map Information Engineering Corporation, 210036, Nanjing, China)

Abstract: There are two main disadvantages during the process of researching the land price warning system. Firstly, the integrated mathematical model hasn't been set up during the construction of the system, and the real estate cycle theory hasn't been introduced. Secondly, the current information hasn't been well utilized, such as the information about dynamic monitoring system of land price. When it is applied, there still exists the deviation between the forecasted land value and the actual land value, as well as between the forecasted trend and the real change. This paper gives some solutions as follows: (1) make full use of the existing accurate information of land price dynamic monitoring system, and set up time-series models and forecast models of spatial distribution. (2) verify the conclusions by the real estate cycle theory and expertise database. Compared with some traditional time-series forecasting methods, Markov chain models are better at analyzing and forecasting the volatility and trend of land price, which provide an efficient approach for us to comprehend the market movements.

To apply the theory of Markov chain to establishing mathematical model of land price's fluctuation and anticipation, we propose the theory of the cycle of land-price's motion and maximization of administrator's and our government's gain. Examples proves that this model and its application are feasible.

Key words: land-price, Markov chain, dynamic monitoring system

0 引言

城市地价是动态变化的, 尤其是在我国经济高速增长的时期, 各个城市的土地价格变化都比较大, 所

收稿日期: 2004-09-15.

基金项目: 国土资源部: 国土资源大调查子项目.

作者简介: 刘俊, 女, 1981—, 硕士研究生, 主要从事人文地理、区域规划、资源评价的学习与研究. E-mail: liujun1002@163.com

万方数据

以地价变化趋势和变动程度的预测就越困难。基于这一原因,本文在过去城市地价预测方法的基础上提出了应用城市地价动态监测信息建立数学模型预测的思路,试图解决这一问题。

在地理事件的预测中,被预测对象所经历的过程中各个阶段(或时点)的状态和状态之间的转移概率是最为关键的。马尔可夫链预测的基本方法就是利用状态之间的转移概率矩阵预测事件发生的状态及其发展变化趋势。马尔可夫链预测法的基本要求是状态转移概率矩阵必须具有一定的稳定性。因此,必须具有足够多的统计数据,才能保证预测的精度与准确性。换句话说,马尔可夫链预测模型必须建立在大量的统计数据的基础之上。这一点也是运用马尔可夫链预测方法预测地理事件的一个最为基本的条件^[1]。我国城市地价动态监测体系的建立是实现全国城市地价动态监测的基础和首要工作,在建立长期、稳定运行的地价监测体系的过程中获取了大量的地价统计数据,达到了马尔可夫链预测法应用要求,在此基础上,运用马尔可夫链预测法进行地价预测,掌握城市自身目前的地价水平和今后的变化趋势,实现对地价的全面监测、分析和预测,有助于今后基准地价的快速更新和政府对于土地资源的宏观控制。

1 地价的几个特点

土地是国家重要资产,在整个社会经济中具有特殊的重要性。因此,土地的利用和地价受到国家的明显控制,市场对地价的自由调节作用被限定在一定的程度之内。由于地产所具有的特殊自然性质和经济性质,其价格也具有与一般商品不同的特点^[2]。

地价具有明显的地域性^[2]。由于土地位置的固定性,在地区性市场之间,地价很难形成统一的市场价格,具有明显的地区性特征。另外,在同一地区城市内,土地的位置对土地收益具有突出的影响作用。

由于地价的上述特点,要预测地价的变化趋势并给出发生某种变化程度的概率就变得非常困难,而掌握城市地价的发展和变化,既可以为房地产开发商服务、旨在引导社会投资方向,还可以通过对各类用地的平均价格水平及其地价变化趋势的分析,控制城市土地的供求关系、调节城市土地的供需矛盾^[3],为政府的地产管理政策的制定提供科学的参考依据^[4],所以,引入新的预测模型是十分必要的。

在过去的研究过程中,一般认为地价预测包括两个方面:一是对城市地价总体水平的预测,它是一个时间序列的变量。二是对同一时点不同区位条件的地价空间分布预测,它是一个空间序列的变量。基于这样的认识,建立了两类地价预测模型,包括时间序列预测模型和空间分布预测模型。

2 时间序列预测模型

目前,国内主要有两种时间序列预测模型。

2.1 第一种模型

利用地价监测信息,取得过去年份的地价指数,以统计分析的趋势外推模型为基础,根据过去年份实际地价变化的时间序列数据,预测未来时期的地价变动指数。基本计算公式为:

$$P_t = f(p_{t-1} \cdots p_{t-k}).$$

式中, P_t 为预测期地价或预测期地价指数; $p_{t-1} \cdots p_{t-k}$ 为过去年份实际地价或地价指数。

举例:

设 $X_t, X_{t-1}, X_{t-2}, \cdots, X_{t-x}$ 为预测期第 $t+1$ 年以前某街区各年份的地价监测点平均租金,建立回归方程:

$$V_{t+1} = f(X_t).$$

式中, V_{t+1} 为第 $t+1$ 年某街区地价监测点平均租金预测值。

该回归模型可能是线性的,也可能是非线性的,视散点图分布来确定它的函数形式^[5]。

再由某街区历年地价监测点平均租金 X_t 与单位面积土地纯收益 I_t 两组时间数列,建立对应关系的回归模型:

$$I_t = A' + B'X_t.$$

因此,预测期第 $t+1$ 年该街区单位面积土地纯收益为

$$I_{t+1} = A' + B'X_{t+1}.$$

2.2 第二种模型

以统计分析的因果关系模型为基础,建立城市地价总水平与影响地价的社会、经济、行政等因素的相互关系模型。

关模型,包括多元线性回归模型^[7]:

$$P_{t+1} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n.$$

或多元非线性回归模型:

$$P_{t+1} = f(x_1, x_2, x_3, \cdots, x_n).$$

3 空间分布预测模型

空间分布预测模型用于测算不同用途地价及土地收益的空间变化,为政府和土地管理部门进行城市土地优化利用和存量土地调整提供决策依据^[7].

3.1 指数回归模型

经过线性模型与非线性模型、多因素模型与单因素模型等反复试算、比较,认为以土地质量指数(或综合定级分)为自变量的单因素指数模型回归拟合性最好^[8],数学表达式如下:

3.1.1 《城镇土地估价规程》推荐的指数模型形式为:

$$I_N = A_0(1 + R)^X = 2.21 * (1 + 0.11)^X.$$

式中, A_0 为回归常数; R 为地价(土地收益)级差系数; X 为单元土地质量指数; I_N 为单元平均土地纯收益(地租).

3.1.2 以 e 为底的指数模型形式为:

$$I_N = A_0e^{bx} = 2.21e^{0.1017x}.$$

式中: b 为回归系数; X 为单元土地综合定级分; I_N 为单元平均土地纯收益(地租)^[9].

4 马尔可夫链(Markov)预测模型

马尔可夫(Markov)预测法,是一种预测事件发生的概率的方法.它是基于马尔可夫链,根据事件的目前状况预测其将来各个时刻(或时期)变动状况的一种预测方法.马尔可夫预测法是对地理事件进行预测的基本方法,它是地理预测中常用的重要方法之一^[1].

众所周知,地价常常因为受到很多不确定性因素的影响,从而使得其涨落呈现出不确定性.这些不确定性因素多是随机的,如供需关系、土地使用者对不动产市场的主观预测及其心理期望等.用随机过程中的马尔可夫链理论建立地价预测的数学模型,并以此来对地价的时空变化特征和趋势进行定量分析,这样可以在一定程度上引导投资者的投资方向,规避风险,避免盲目和非理性的行为,也可以作为政府决策者的一种辅助手段,使其能在面对变幻莫测的不动产市场时采取科学的政策策略,以引导不动产市场的健康发展.

4.1 地价的状态转移概率与状态转移概率矩阵的计算

在地价的发展变化过程中,从某一种地价状态出发,下一时刻转移到其它地价状态的可能性,称为地价状态转移概率.根据条件概率的定义,由地价状态 E_i 转为地价状态 E_j 的状态转移概率 $P(E_i \rightarrow E_j)$ 就是条件概率 $P(E_j/E_i)$,即

$$P(E_i \rightarrow E_j) = P(E_j/E_i) = P_{ij}. \quad (1)$$

地价状态转移概率矩阵假定某一区域被预测的地价有 E_1, E_2, \cdots, E_n ,共 n 个可能的状态.记 P_{ij} 为从地价状态 E_i 转为地价状态 E_j 的状态转移概率,作矩阵:

$$\begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix} \quad (2)$$

则称 P 为地价的状态转移概率矩阵^[10].

如果被预测的某一城市地价目前处于状态 E_i ,那么在下一个时刻,它可能由状态 E_i 转向 $E_1, E_2, \cdots, E_i, \cdots, E_n$ 中的任一个状态.所以 P_{ij} 满足条件:

万方数据

$$\begin{cases} 0 \leq P_{ij} \leq 1 \quad (i, j = 1, 2, \cdots, n) \\ \sum_{j=1}^n P_{ij} = 1 \quad (i = 1, 2, \cdots, n). \end{cases} \quad (3)$$

一般地,将满足条件(3)的任何矩阵都称为随机矩阵,或概率矩阵. 不难证明,如果 P 为概率矩阵,则对任何数 $m > 0$,矩阵 P_m 都是概率矩阵. 如果 P 为概率矩阵,而且存在整数 $m > 0$,使得概率矩阵 P_m 中诸元素皆非零,则称 P 为标准概率矩阵^[11]. 可以证明,如果 P 为标准概率矩阵,则存在非零向量 $\alpha = [x_1, x_2, \cdots, x_n]$,而且 x_i 满足 $0 \leq x_i \leq 1$ 及 $\sum_{i=1}^n x_i = 1$,使得:

$$\alpha P = \alpha \quad (4)$$

这样的向量 α 称为平衡向量,或终极向量.

计算地价状态转移概率矩阵 P ,就是要求每个地价状态转移到其它任何一个状态的转移概率 $P_{ij}(i, j = 1, 2, \cdots, n)$. 为了求出每一个 P_{ij} ,采用频率近似概率的思想来加以计算^[12]. 考虑某城市地价变化的三个状态,即“上涨”、“不变”和“下跌”. 记 E_1 为“上涨”状态, E_2 为“不变”状态, E_3 为“下跌”状态. 表 1 给出了该城市 2002 年 1 月 ~ 2004 年 12 月期间地价状态变化情况.

表 1 某城市地价变化的状态转移情况

月份	2002. 1	2002. 2	2002. 3	2002. 4	2002. 5	2002. 6	2002. 7	2002. 8	2002. 9
序号	1	2	3	4	5	6	7	8	9
状态	E_1	E_1	E_2	E_3	E_2	E_1	E_3	E_2	E_1
月份	2002. 10	2002. 11	2002. 12	2003. 1	2003. 2	2003. 3	2003. 4	2003. 5	2003. 6
序号	11	12	13	14	15	16	17	18	19
状态	E_3	E_1	E_2	E_3	E_1	E_2	E_1	E_3	E_3
月份	2003. 7	2003. 8	2003. 9	2003. 10	2003. 11	2003. 12	2004. 1	2004. 2	2004. 3
序号	21	22	23	24	25	26	27	28	29
状态	E_3	E_3	E_2	E_1	E_1	E_3	E_2	E_2	E_1
月份	2004. 4	2004. 5	2004. 6	2004. 7	2004. 8	2004. 9	2004. 10	2004. 11	2004. 12
序号	21	22	23	24	25	26	27	28	29
状态	E_1	E_3	E_2	E_1	E_1	E_2	E_2	E_3	E_1

以下计算该城市地价变化的状态转移概率矩阵. 从表 1 中可知,在 15 个从 E_1 出发(转移出去)的状态转移中,有 3 个是从 E_1 转移到 E_1 的(即 $1 \rightarrow 2, 24 \rightarrow 25, 34 \rightarrow 35$),有 7 个是从 E_1 转移到 E_2 的(即 $2 \rightarrow 3, 9 \rightarrow 10, 12 \rightarrow 13, 15 \rightarrow 16, 29 \rightarrow 30, 35 \rightarrow 36, 39 \rightarrow 40$),有 5 个是从 E_1 转移到 E_3 的(即 $6 \rightarrow 7, 17 \rightarrow 18, 20 \rightarrow 21, 25 \rightarrow 26, 31 \rightarrow 32$). 故:

$$\begin{aligned} P_{11} &= P(E_1 \rightarrow E_1) = P(E_1 | E_1) = 5/13 = 0.200\ 0 \\ P_{12} &= P(E_1 \rightarrow E_2) = P(E_2 | E_1) = 7/15 = 0.466\ 7 \end{aligned}$$

按照上述同样的办法计算可以得到:

$$\begin{aligned} P_{21} &= P(E_2 \rightarrow E_1) = P(E_1 | E_2) = 7/13 = 0.538\ 5 \\ P_{22} &= P(E_2 \rightarrow E_2) = P(E_2 | E_2) = 2/13 = 0.153\ 8 \\ P_{23} &= P(E_2 \rightarrow E_3) = P(E_3 | E_2) = 4/13 = 0.307\ 7 \\ P_{31} &= P(E_3 \rightarrow E_1) = P(E_1 | E_3) = 4/11 = 0.363\ 6 \\ P_{32} &= P(E_3 \rightarrow E_2) = P(E_2 | E_3) = 5/11 = 0.454\ 5 \\ P_{33} &= P(E_3 \rightarrow E_3) = P(E_3 | E_3) = 2/11 = 0.181\ 8 \end{aligned}$$

所以,该城市地价变化的状态转移概率矩阵为:

$$P = \begin{pmatrix} 0.200\ 0 & 0.466\ 7 & 0.333\ 3 \\ 0.538\ 5 & 0.153\ 8 & 0.307\ 7 \\ 0.363\ 6 & 0.454\ 5 & 0.181\ 8 \end{pmatrix} \quad (5)$$

4.2 应用马尔可夫链对城市地价进行预测

首先定义状态概率为 $\pi_j(k)$. $\pi_j(k)$ 表示事件在初始($k = 0$)状态为已知的条件下,经过 k 次状态转移后,在第 k 个时刻(时期)处于状态 E_j 的概率. 根据概率的性质,显然有:

万元数据

$$\sum_{j=1}^N \pi_j(k) = 1. \quad (6)$$

从初始状态开始,经过 k 次状态转移后到达状态 E_j 这一状态转移过程,可以看作是首先经过 $(k-1)$ 次状态转移后到达状态 $E_i (i=1,2,\dots,n)$,然后再由 E_i 经过一次状态转移到达状态 E_j . 根据马尔可夫过程的无后效性及 Bayes 条件概率公式,有

$$\pi_j(k) = \sum_{i=1}^n \pi_i(k-1) P_{ij} \quad (j=1,2,\dots,n). \quad (7)$$

若记行向量 $\pi(k) = [\pi_1(k), \pi_2(k), \dots, \pi_n(k)]$,则由(7)式可得逐次计算状态概率的递推公式:

$$\begin{cases} \pi(1) = \pi(0)P \\ \pi(2) = \pi(1)P = \pi(0)P^2 \\ \dots\dots\dots \\ \pi(k) = \pi(k-1)P = \dots = \pi(0)P^k \end{cases} \quad (8)$$

(8)式中, $\pi(0) = [\pi_1(0), \pi_2(0), \dots, \pi_n(0)]$ 为初始状态概率向量.

4.2.1 第 k 个时刻(时期)的状态概率预测

在前例中,如果将 2004 年 12 月的城市地价变化状态记为 $\pi(0) = [0,1,0]$ (因为 2004 年 12 月处于“不变”状态),则将地价状态转移概率矩阵(5)式及 $\pi(0)$ 代入递推公式(8),就可以求得 2005 年 1 月 ~ 2005 年 12 月可能出现的各种地价变化状态的概率(见表 2).

表 2 2005 年 1 月 ~ 2005 年 12 月地价变化状态概率预测值

月份	2005. 1			2005. 2			2005. 3		
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3
状态	0.538 5	0.152 8	0.307 7	0.302 4	0.414 8	0.283 7	0.386 4	0.333 4	0.279 9
月份	2005. 4			2005. 5			2005. 6		
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3
状态	0.367 7	0.350 9	0.279 9	0.364 7	0.353 2	0.279 9	0.365 6	0.352 4	0.279 9
月份	2005. 7			2005. 8			2005. 9		
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3
状态	0.367 7	0.350 9	0.279 9	0.364 7	0.353 2	0.279 9	0.365 6	0.352 4	0.279 9
月份	2005. 10			2005. 11			2005. 12		
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3
状态	0.365 3	0.352 5	0.279 9	0.365 3	0.352 5	0.279 9	0.365 3	0.352 5	0.279 9

4.2.2 地价的终极状态概率预测

经过无穷多次状态转移后所得到的状态概率称为终极状态概率,或称平衡状态概率. 如果记终极状态概率向量为 $\pi = [\pi_1, \pi_2, \dots, \pi_n]$,则

$$\pi_i = \lim_{k \rightarrow \infty} \pi_i(k) \quad (i=1,2,\dots,n). \quad (9)$$

即:

$$\begin{aligned} \pi &= [\lim_{k \rightarrow \infty} \pi_1(k), \lim_{k \rightarrow \infty} \pi_2(k), \dots, \lim_{k \rightarrow \infty} \pi_n(k)] \\ &= \lim_{k \rightarrow \infty} \pi(k) \end{aligned} \quad (10)$$

按照极限的定义可知:

$$\lim_{k \rightarrow \infty} \pi(k) = \lim_{k \rightarrow \infty} \pi(k+1) = \pi. \quad (11)$$

将(11)式代入马尔可夫预测模型的递推公式(8)得

$$\lim_{k \rightarrow \infty} \pi(k+1) = P \times \lim_{k \rightarrow \infty} \pi(k),$$

即:

$$\pi = \pi P. \quad (12)$$

这样,就得到了终极状态概率应满足的条件:

$$(1) \pi = \pi P.$$

$$(2) 0 \leq \pi_i \leq 1 (i=1,2,\dots,n).$$

$$(3) \sum_{i=1}^n \pi_i = 1.$$

以上条件(2)与(3)是状态概率的要求,其中,条件(2)表示,在无穷多次状态转移后,事件必处在 n 个状态中的任意一个;条件(1)就是用来计算终极状态概率的公式.终极状态概率是用来预测马尔可夫过程在遥远的未来会出现什么趋势的重要信息.如前例关于某城市地价变化状态概率的预测中,设终极状态的概率为 $\pi = [\pi_1, \pi_2, \pi_3]$,则

$$[\pi_1, \pi_2, \pi_3] = [\pi_1, \pi_2, \pi_3] \begin{pmatrix} 0.2000 & 0.4667 & 0.3333 \\ 0.5385 & 0.1538 & 0.3077 \\ 0.3636 & 0.4545 & 0.1818 \end{pmatrix};$$

即:

$$\begin{cases} \pi_1 = 0.2000\pi_1 + 0.5385\pi_2 + 0.3636\pi_3 \\ \pi_2 = 0.4667\pi_1 + 0.1538\pi_2 + 0.4545\pi_3 \\ \pi_3 = 0.3333\pi_1 + 0.3077\pi_2 + 0.1818\pi_3 \end{cases} \quad (13)$$

求解方程组(13)得: $\pi_1 = 0.3653, \pi_2 = 0.3525, \pi_3 = 0.2799$.

这说明,该城市地价的变化,在无穷多次状态转移后,“上涨”和“不变”状态出现的概率都将大于“下跌”状态出现的概率.

5 结论

决定地价高低的主要因素是土地提供地租量的多少.地盘必须由经营者、投资者通过经营、租赁、买卖、投资等方式得以实现.地盘是土地经营者超额利润的组成部分.因此,影响土地提供地租量、影响地产市场行为的所有因素,都直接或间接地影响地价,是地价评估所必须考虑的.概括地讲,这些因素大致可归成两大类:影响土地供求的因素及影响宗地地租量的因素.供求关系使地价产生波动,是影响地价的外在因素;地租量的影响因素是地价的内在性因素,表示一个地区、一宗土地的收益水平.然而定量化的精确确定上述影响因素对地价的变化所起的作用目前还存在一定的难度.

应用马尔可夫链理论建立起地价运动数学模型,应用结果与实际情况吻合较好.这说明了该模型应用的可行性.与其它地价预测模型相比,马尔可夫链预测模型在预测地价的同时还可以预测达到某一地价水平的平均时间,这是其它预测模型难以实现的.

[参考文献]

- [1] 徐建华主编. 现代地理学中的数学方法(第二版)[M]. 北京: 高等教育出版社, 2004.
- [2] 鹿心社主编. 中国地产估价手册[M]. 北京: 改革出版社, 1993.
- [3] 刘卫东著. 土地系统研究的理论与实践[M]. 北京: 海洋出版社, 1993.
- [4] James Valente, ShanShan Wu, Alan Gelfand, Sirmans C F. Apartment Rent Prediction Using Spatial Modeling[M]. Connecticut: University of Connecticut Press, 2004.
- [5] 李玉英著. 地产价格评估和实证与创新[M]. 北京: 中国财政经济出版社, 1999.
- [6] 胡存智等. 土地估价理论与方法[M]. 北京: 改革出版社, 1995.
- [7] 蔡兵备, 欧阳安蛟. 城市地价评估方法发展与创新[M]. 北京: 社会科学文献出版社, 2002.
- [8] 刘卫东等著. 城市土地价格调查、评价及动态监测[M]. 北京: 科学出版社, 2002.
- [9] 张洪主编. 城市土地评价理论与方法[M]. 北京: 经济管理出版社, 2002.
- [10] Dynik E B, Yushkevich A A. Markov Processes Theorems and Problems [M]. New York: Plenum Press, 1996.
- [11] 张宇山, 廖芹. 马尔可夫链在股市分析中的若干应用[J]. 华南理工大学学报(自然科学版), 2003, 31(7): 74—77.
- [12] 颜荣芳. 股票市场预测的随机过程模型[J]. 西北师范大学学报(自然科学版), 1999, 35(3): 44—46.

[责任编辑: 陆炳新]