

# 基于灰色-马尔可夫模型的房地产周期研究

吴璟,刘洪玉

(清华大学 房地产研究所,北京 100084)

**摘要:**文章针对我国房地产周期研究中存在的内在机理复杂、数据信息不足等问题,提出了利用灰色-马尔可夫模型进行房地产周期分析和预测的想法,即利用 GM(1,1)模型估计长期趋势成分,利用马尔可夫链预测模型估计周期性波动成分,通过两者的结合完成对房地产周期的拟合和预测。实证表明,该模型能够成为当前我国房地产周期研究的一项有效工具。

**关键词:**房地产周期;灰色系统;马尔可夫链

中图分类号:F293.3

文献标识码:A

文章编号:1002-6487(2009)06-0032-03

我国房地产市场的独特性和数据基础的薄弱大大增加了房地产周期研究的难度。一方面,受发展阶段和社会背景的影响,我国房地产市场的运行规律与国外现有研究成果存在较大差距,许多在国外研究中大量使用的周期研究模型难以直接套用,同时针对我国房地产市场特殊性的理论研究又尚存在诸多不足。另一方面,目前我国房地产领域数据采集、统计和披露制度相对滞后,缺乏高质量、长时间序列基础数据的支持,进一步制约了许多复杂定量模型的应用。

针对上述情况,本文提出利用一维灰色动态模型和马尔可夫链预测模型相结合的灰色-马尔可夫模型进行房地产周期识别和预测的思路。灰色系统和马尔可夫链是系统分析领域广泛应用的两种方法,尽管其基本思路存在很大差异,但都是直接从数据入手进行分析,均具有所需信息少、精度高的特点。同时,两种方法具有一定的互补性,使得其相互结合后形成的灰色-马尔可夫模型能够进一步提高估计和预测的精度,近来已经在一些复杂系统研究中得到了应用。

## 1 房地产周期研究的理论模型

### 1.1 周期衡量指标的选择

房地产市场呈现出的周期性变动需要通过一定的量化指标来刻画。根据英国皇家测量师协会提出的定义,如果用一种标准的经济变量来表示房地产景气循环,应当是所有物业总收益率的周期性波动。但由于数据可得性问题,实际研究中学者们很少使用类似于总收益率这样的指标,而通常以更为容易获得的统计指标来代替。例如,我国统计部门在建立国房景气指数时,采用了房地产开发投资等8项统计指标;而在研究领域,根据刘洪玉等的总结,最经常使用的周期分析指标包括商品住房价格等7项<sup>[1]</sup>。

近年来我国学者在进行房地产周期研究时,多选择了综合采用多个指标的方法。本文选择基于单指标建立房地产周期研究模型,而这无疑需要选择能够最准确、最及时反映市场景气情况的指标。经过比选,本文选择了商品住房销售价格。这主要基于几方面原因:首先,价格是反映市场运行状况

的最重要指标,在房地产市场也同样如此,在房地产市场出现周期变化时,价格的波动是其最直观、最重要的指示器;其次,住房市场(尤其是新建商品住房市场)目前在我国房地产市场中占据了最为重要的地位,其波动和变化的影响也最为显著,当前受到的关注程度最高;最后,从数据可得性出发,住房价格统计在我国起步较早,具有较好的数据基础。

### 1.2 灰色-马尔可夫模型的总体思路

根据目前研究中的一般共识,房地产周期被视为围绕特定长期趋势发生的周期性往复循环过程。基于此,某一时刻的商品住房价格可以认为由长期趋势成分 PT, 周期性波动成分和随机成分 PR 三部分组成,即:

$$P=PT+PF+PR \quad (1)$$

其中,PF 是房地产市场周期研究的主要分析对象。而要对 PF 进行分析,就需要从住房价格的总体变化 P 中消除其他部分的影响,特别是长期趋势成分 PT 的影响。因此,房地产周期分析中一般需要首先对 PT 进行估计,由此得到对 PF 的间接估计值,在此基础上再利用特定模型对 PF 的该估计值进行直接拟合,发掘其变化规律,进而实现判别周期状态、预测未来趋势等目的。

估计长期趋势成分 PT 时,一种方法是从其内在机理入手,例如借助误差修正模型等时间序列分析方法探寻住房价格与宏观经济基本面、住房市场供求等因素的长期均衡关系,另一种方法则是直接从数据入手,通过指数、多项式等形式的函数进行拟合。本文采用一维灰色动态模型 GM(1,1)来估计 PT 成分。从理论角度看,我国商品住房价格还存在诸多不明确的影响因素,各种因素的影响幅度和影响机制也存在一定争论,而利用灰色系统思想实现的 GM(1,1)模型则避开了对价格变化理论机制的分析,直接从数据入手进行剖析,更可能得到理想的结果;从计算角度看,GM(1,1)将时间序列转化为微分方程,并最终得到指数曲线形式的模型的解,适用于对长期变化趋势的估计,具有较高的拟合精度。

对周期性波动成分 PF 的估计则引入了马尔可夫链预测模型。相比长期趋势成分 PT,PF 成分的变化机理更为复杂,甚至可能存在相当程度的非理性因素,其理论分析难度较

基金项目:北京市 2007 年重点调研课题

大。而直接从数据入手进行模型拟合时,传统的指数曲线形式、多项式形式模型,乃至于此前介绍的 GM(1,1)模型,在一定区间内均具有单调性,因此难以适用于存在剧烈波动的 PF 成分的估计。而马尔可夫链预测模型由于其无后效性的特点,十分适合于构建波动性较大的动态系统,成为实现对 PF 成分直接估计的一种较为理想的选择。

综合上述分析,回避对住房价格变化机制的理论分析,而直接从剖析数列规律入手,并发挥 GM(1,1)模型和马尔可夫链预测模型的优势,分别用于对长期趋势成分 PT 和周期性波动成分 PF 的直接估计,这就是本文所提出的基于灰色-马尔可夫模型进行房地产周期研究的总体思路。

### 1.3 长期趋势成分的估计:GM(1,1)模型的应用

对于住房价格原始时间序列:

$$P^{(0)}(t) = \{P^{(0)}(1), P^{(0)}(2), \dots, P^{(0)}(n)\} \quad (2)$$

进行一次累加,得到生成数列:

$$P^{(1)}(t) = \{P^{(1)}(1), P^{(1)}(2), \dots, P^{(1)}(n)\} \quad (3)$$

其中:  $P^{(1)}(k) = \sum_{i=1}^k P^{(0)}(i); k=1, 2, \dots, n。$

针对式(3),可以建立一阶线性微分方程:

$$\frac{dP^{(1)}(t)}{dt} + \alpha P^{(1)}(t) = \mu \quad (4)$$

式(4)即被称为 GM(1,1)模型,该式中的待估参数向量为:

$$[\alpha, \mu]^T = (B^T B)^{-1} B^T Y_N \quad (5)$$

其中:

$$B = \begin{bmatrix} -\frac{1}{2}[P^{(1)}(1)+P^{(1)}(2)], 1 \\ -\frac{1}{2}[P^{(1)}(2)+P^{(1)}(3)], 1 \\ \dots\dots\dots \\ -\frac{1}{2}[P^{(1)}(n-1)+P^{(1)}(n)], 1 \end{bmatrix}, Y_N = [P^{(0)}(2)+P^{(0)}(3), \dots, P^{(0)}(n)]^T$$

将其代入式(4),再进行微分方程求解,得到:

$$\hat{P}^{(1)}(k) = [P^{(0)}(1) - \frac{\hat{\mu}}{\hat{\alpha}}] e^{-\hat{\alpha}k} + \frac{\hat{\mu}}{\hat{\alpha}}, k \in N \quad (6)$$

利用式(6)得出对生成序列(3)的估计值后,就可以再通过迭减运算,得出原始数列(2)的估计值,该估计值即等价于对长期趋势成分 PT 的估计值,记为 PT。在此基础上,忽略高阶小量的随机波动,可以得到 PF 成分的间接估计值:

$$PF = P - PT \quad (7)$$

及其波动率的间接估计值:

$$\hat{\theta} = PF / PT \quad (8)$$

利用 PF 或  $\hat{\theta}$ ,就可以对房地产周期进行描述,并直观判断当时在周期中所处的阶段。

### 1.4 周期性波动成分的估计:马尔可夫链预测模型的应用

根据马尔可夫链预测模型的基本原理,系统达到某一状态的概率只和近期状态有关,在一定时期后马尔可夫过程逐渐趋于稳定状态而与原始条件无关。这一特性称为“无后效性”,即第 n 次结果只取决于第 n-1 次的结果,以此类推。因此马尔可夫链预测模型实际上是将一个系统的状态和状态转移定量表示的数学模型<sup>[9]</sup>。具体而言,有条件概率:

$$p_{ij}(n) = p\{X_n = j | X_{n-1} = i\} \quad (9)$$

表示系统在 n-1 时刻处于 i 状态的情况下,在 n 时刻系

统转移到 j 状态的条件概率。系统中所有可能的状态转移  $P_{ij}$  就构成了马尔可夫链的一步转移概率矩阵。

根据上述思路,针对波动率的间接估计值  $\hat{\theta}$ ,可以将其划分为若干动态的状态区间,并根据落入各状态区间的点及其在下一时期内的状态计算马尔可夫链的一步转移概率矩阵,从而对每一时期预测其下一时期落在各状态区间中的概率,达到对波动率(或者由此推导 PF 成分,下同)进行直接估计的目的,并进一步对其未来走势进行预测。

## 2 房地产周期识别和预测模型的应用

为了验证上述灰色-马尔可夫模型在房地产周期研究中的可行性和准确性,下面利用北京市数据对该模型进行实际应用。

### 2.1 数据

如前所述,本文选择商品住房价格作为反映房地产周期的指标。而就北京市而言,目前质量较高、时间序列较长的价格指标是“中房指数”中的北京住房指数。但其编制方法在 2005 年第二季度进行了较大调整,且未发布更新后的历史序列,因此目前可使用的最长序列为 1997 年第一季度至 2004 年第四季度。为此,本文选用 1997 年第一季度至 2004 年第四季度数据进行模型的拟合,并据此预测 2005 年第一季度走势,再将该预测值与实际值进行比较。对于各季度指数值,利用 CPI 调整消除通货膨胀影响,均统一按 1997 年价格计算。同时为了在不改变数据的前提下减小后续建模的计算量,对各指数均采用其自然对数形式。进行上述调整后,得到的实际走势如图 1 中实线所示。

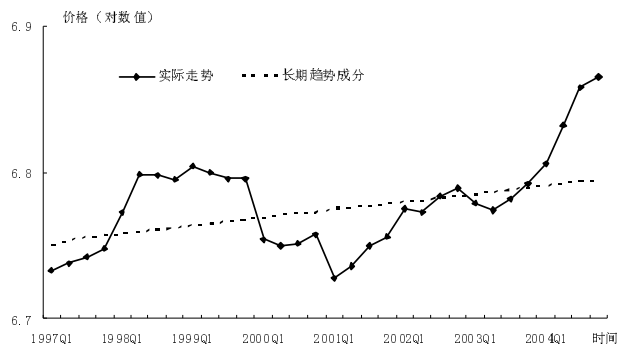


图 1 北京市商品住房价格的实际走势和长期趋势成分估计结果(1997~2004)

### 2.2 长期趋势成分的估计

根据文中介绍的方法,基于原始序列叠加得到生成数列  $P^{(0)}(t)$ (式(3)),并建立式(4)所示的线性偏微分方程。在此基础上,参照式(5),利用 Matlab 进行编程计算后,估计得到生成序列为:

$$\hat{P}^{(1)}(t) = 33768e^{0.002(t-1)} - 33761.5 \quad (10)$$

利用式(10)进行叠减计算,即可得到 PT 成分在各期内的估计值,如图 1 中的虚线所示。对结果进行后验差检验后得到 p 统计量为 0.71, c 统计量为 0.58,均超过 GM(1,1)模型拟合精度的三级标准。检验结果表明该模型是可以接受的,但同时也说明除 GM(1,1)模型所反映出的价格变化的长期趋势外,各期价格还存在着显著的波动,而这正是景气循环周期的体现。

2.3 周期性波动成分的间接估计和直接估计

得到长期趋势性成分的估计值后,就可以利用式(7)或式(8),得到绝对水平或波动率形式的对 PF 成分的间接估计值。这里采用较为直观的波动率形式,即式(8),得到的波动率间接估计结果如图 2 中实线所示。

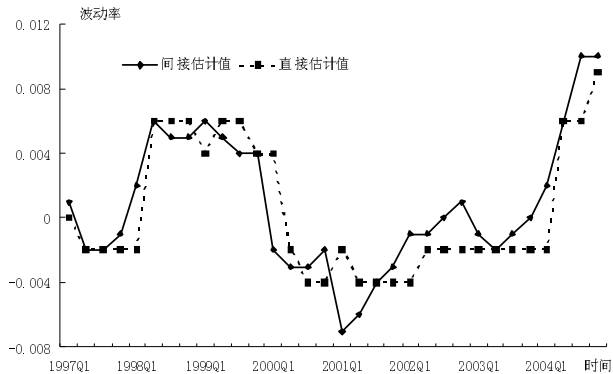


图 2 北京市商品住房价格中周期性波动成分的间接和直接估计结果(1997~2004)

从图 2 中可以看出,1997 年北京住房价格开始迅速攀升,于 1997、1998 年达到波峰,并在较长时间内保持高位运行。此后住房价格在 2000 年起转向下滑,并在 2001 年达到波谷,但此后迅速开始回升。特别是在经历了 2002 至 2003 年的一段调整期后,从 2003 年下半年起住房价格出现大幅度增长。至 2004 年第四季度此次上涨已经远远超过了前一次波峰,且仍处于扩张期,尚未出现下滑的趋势。

将图 2 与北京市房地产市场近年来的发展历程进行对比,可以发现两者是十分吻合的。1998 年国务院 23 号文出台后,延续多年的福利制分房制度逐步停止,消费者对商品住房的需求量猛增,同时市场对住房价格的预期也因此出现了较大幅度的提高。这些因素的综合作用使得住房价格迅速提高,并在 1998 年至 1999 年间保持高位运行。但这段时间内住房建设量也相应的随之提高,这些新建住房从 1999 年下半年开始投入市场,使得市场中的住房供应量迅速扩大,有效地阻碍了住房价格的进一步提高。因此,2000 年后住房价格转入波谷阶段。但基于北京市整体经济运行形势良好和居民收入不断提高这一大背景,波谷阶段并没有延续太长时间。从 2003 年开始,在土地价格提高等因素的作用下,北京房地产市场又进入新一轮快速增长期。

利用马尔可夫链预测模型对上述间接估计结果进行直接拟合。首先,根据波动率的变化范围,将其划分为 8 个动态区间,如表 1 所示。

表 1 马尔可夫链预测模型中动态区间的划分

动态区间	S1	S2	S3	S4
范围	(-0.009,-0.005]	(-0.005,-0.003]	(-0.003,-0.001]	(-0.001,0.001]
中点取值	-0.007	-0.004	-0.002	0.000
动态区间	S5	S6	S7	S8
范围	(0.001,0.003]	(0.003,0.005]	(0.005,0.007]	(0.007,0.013]
中点取值	0.002	0.004	0.008	0.011

按该动态区间划分,得到一步转移概率矩阵为:

$$P_{ij} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/4 & 1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 1/7 & 1/7 & 3/7 & 1/7 & 1/7 & 0 & 0 & 0 \\ 0 & 0 & 2/4 & 1/4 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2/2 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2/5 & 2/5 & 1/5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/1 \end{pmatrix} \quad (11)$$

利用式(11)所示的一步转移概率矩阵,得到对各时期波动率的直接估计值,如图 2 中虚线所示。可以发现,直接估计值能够较好地拟合各时期内的价格波动率,同时后验差检验也表明,c 统计量和 p 统计量都超过了一级精度的要求

2.4 走势预测

GM(1,1)模型和马尔可夫链预测模型均具有较强的外推性,可以应用于对未来走势的预测。与上述拟合分析过程相似,在进行预测时,同样利用 GM(1,1)模型预测未来的长期趋势成分,利用马尔可夫链预测未来的周期性波动成分(波动率形式),最后将两部分预测结果合并,得到最终预测值。

根据这一思路,利用式(10)和式(11)所示的 GM(1,1)模型和马尔可夫链预测模型拟合结果,对 2005 年第一季度的走势进行预测,如表 2 所示。

表 2 2005 年第一季度北京商品住房价格预测结果(对数形式)

时间	长期趋势成分预测值	波动率预测值	价格预测值
2005Q1	6.796	0.011	6.805

将预测值还原为指数形式,并利用 CPI 修正还原为按当时价格计算后得到 2005 年第一季度北京市商品住房价格的预测值为 1005 点,呈继续上升势头。而根据中房北京指数的统计结果,其实际值为 1012 点。预测的走势基本正确,但增长幅度略有低估。

3 结论

GM(1,1)模型和马尔可夫链预测模型是系统分析方法中常用的两种模型,GM(1,1)模型能够较好地拟合系统的长期变化趋势,而马尔可夫链则可以对短期内的剧烈波动做出较为准确的拟合。利用这两种模型的特点,本文提出以 GM(1,1)拟合房地产经济发展中的长期趋势,以马尔可夫链模型拟合围绕长期趋势的上下波动,并将两者的结果相互结合的房地产周期分析方法。实证结果表明,该模型能够较好地符合实际情况,有助于准确把握房地产市场所处的发展阶段,并对近期走势做出预测,从而为房地产周期研究提供了一种有效的工具。

参考文献:

- [1]刘洪玉,张红.房地产业与社会经济[M].北京:清华大学出版社,2006.
- [2]王露璐,邓丽,刘先涛.股票投资价值灰色马尔可夫预测[J].商业研究,2004,(1).
- [3]钟昌宝.基于灰色-马尔柯夫模型预测房地产价格[J].统计与决策,2005,(2).
- [4]RICS. Understanding the Property Cycle[M].London:RICS,1994.
- [5]徐国祥.统计指数理论及应用[M].北京:中国统计出版社,2004.
- [6]翁少群,张红.基于模糊模式识别理论的中国房地产市场发展阶段判别研究[J].土木工程学报,2004,(5).
- [7]朱骏,张红.基于时间序列与人工神经网络的房地产周期识别[J].清华大学学报(自然科学版),2006,(6).
- [8]邓聚龙.灰色系统基本方法[M].武汉:华中理工大学出版社,2005.
- [9]盛骤,谢士干,潘承毅.概率论与数理统计[M].北京:高等教育出版社,1989.
- [10]中国指数研究院.中国房地产指数系统:理论与实践[M].北京:经济管理出版社,2005.

(责任编辑/亦 民)