

GM(1,1)扩展方法在经济增长预测中的应用及检验*

吴卓立 覃正 王魁恒 惠龙 丁洁

(西安交通大学管理学院 西安 710049)

(武汉大学信息管理学院)

摘要 介绍了一种 GM(1,1)模型与线性回归组合预测方法,使用该方法建立了西安市国内生产总值(GDP)预测模型,计算了西安市 2001~2003 年 GDP 总值,对比检验了单独使用 GM(1,1)模型和线性回归模型与该模型的精度。检验结果表明,该方法较好地克服了单独使用 GM(1,1)模型和线性回归模型的缺陷,在实际预测中取得了好的效果。

关键词 GM(1,1)模型 线性回归模型 国内生产总值预测 GDP

国民经济计划是国家宏观调控的导向性手段,如何科学合理地制定出国民经济发展目标,是政府工作中一个重要的问题,这要依靠科学的预测与决策方法。

在国民经济预测中,一般采用统计方法,而在大多数情况下需要大量数据,否则预测不够精确。灰色预测 GM(1,1)模型要求样本数据少且短期预测精度高,因此得到了广泛的应用。但是该模型主要适用于单一的指数增长的模型,对序列数据出现异常的情况很难加以考虑^[1]。将两种方法有机地结合起来产生的 GM(1,1)模型与线性回归组合预测方法可以克服 GM(1,1)模型与统计方法的缺点,在国民经济预测中取得满意的预测效果^[2]。

本文使用 GM(1,1)模型与线性回归组合预测方法,建立并检验了西安市地方生产总值预测模型。检验显示,使用该方法建立的模型比单独使用 GM(1,1)模型或线性回归模型的精度要高,在国民经济的预测中有一定实用价值。

1 GM(1,1)模型与线性回归组合预测方法

GM(1,1)模型与线性回归组合预测方法是根据灰色灾变预测中灾变日期的预测原理,在“跳变点”(跳变点的数据值相对于前连续的两个点来说显示出明显的非线性特点,即不在同一直线上)采用跳变预测函数预测其跳变值,在非跳变点采用线性回归预测。

具体步骤如下:

a. 按时间先后顺序列出样本序列:

$$x_{(0)} = (x_{(0)}(1), x_{(0)}(2), \dots, x_{(0)}(k)) \quad (1)$$

时间按顺序构成的序列:

$$Q_{(0)} = (Q_{(0)}(1), Q_{(0)}(2), \dots, Q_{(0)}(K)) \quad (2)$$

b. 根据历史数据画出折线图,找出灾变点,即这里的跳变点。根据灰色灾变原理,将数据中跳变点的跳变值构成数据序列:

$$x_{A(0)} = (x_{A(0)}(1), x_{A(0)}(2), \dots, x_{A(0)}(k_A)) \subset x_{(0)} \quad (3)$$

$x_{A(0)}$ 是跳变数据序列。

$$Q_{A(0)} = (Q_{A(0)}(1), Q_{A(0)}(2), \dots, Q_{A(0)}(k)) \subset Q_{(0)} \quad (4)$$

$Q_{A(0)}$ 是对应的跳变时间的序列。

c. 采用 GM(1,1) 模型分别计算出跳变时间和相应值的函数,并且预测出下一个或几个跳变时间和跳变值。

由 GM(1,1) 模型计算 $x_{A(0)}$ 和 $Q_{A(0)}$ 的跳变函数 $\hat{x}_{A(0)}(k+1)$ 和 $\hat{Q}_{A(0)}(k+1)$ 如下:

$$\begin{cases} \hat{x}_{A(1)}(k+1) = \left[x_{(0)}(0) - \frac{\mu_p}{\alpha_p} \right] e^{-\alpha_p k} + \frac{\mu_p}{\alpha_p} \end{cases} \quad (5)$$

$$\begin{cases} \hat{Q}_{A(1)}(k+1) = \left[Q_{(0)}(0) - \frac{\mu_p}{\alpha_p} \right] e^{-\alpha_p k} + \frac{\mu_p}{\alpha_p} \end{cases} \quad (6)$$

其中, α_p 为发展系数, μ_p 为灰作用量,它们是微分方程的参数。分别根据 $x_{A(0)}$ 和 $Q_{A(0)}$, 使用最小二乘估计得到 $\hat{x}_{A(0)}(k+1)$ 和 $\hat{Q}_{A(0)}(k+1)$ 的这两个参数值。

使用跳变函数预测跳变时间和跳变值分别为:

$$\begin{cases} \hat{x}_{A(0)}(k+1) = \hat{x}_{A(1)}(k+1) - \hat{x}_{A(1)}(k) \end{cases} \quad (7)$$

$$\begin{cases} \hat{Q}_{A(0)}(k+1) = \hat{Q}_{A(1)}(k+1) - \hat{Q}_{A(1)}(k) \end{cases} \quad (8)$$

d. 在不是跳变点的预测日期上使用线性回归模型。判别预测日期点属于哪个日期段的方法是:对任意相邻两个跳变点来说,可以认为前一个跳变点的上一个点与后一个跳变点的上一个点之间是线性关系。可以使用线性回归模型拟合。处于这一日期段之间的预测日期点通过线性回归方程计算出函数值。

使用的线性回归模型为:

$$y = ax + b \quad (9)$$

2 经济预测模型的建立与检验

2.1 西安市国内生产总值预测模型的建立 国内生产总值是衡量经济发展的一个重要指标,反映了社会经济系统发展的一种整体趋势,在政府的五年计划中有着重要的决策参考价值,也是下一个五年计划中制定目标的重要指标之一。本文以西安市国内生产总值为例,根据西安市 1985~2000 年国内生产总值数据,使用 GM(1,1)模型与线性回归组合预测方法建模并对模型做了精度检验。

基金项目:国家自然科学基金项目(编号:70471037);陕西省自然科学基金项目(编号:2004G02);西安市“十五”计划执行后评估研究项目(编号: XASJW11-521)。

作者简介:吴卓立,1982年生,男,西安交通大学管理学院硕士研究生;覃正,1958年生,男,西安交通大学管理学院教授,博士生导师。

表1 西安市1985~2000年国内生产总值 单位:亿元

序号	1	2	3	4	5	6	7	8
年份	1985	1986	1987	1988	1989	1990	1991	1992
值	57.58	65.78	80.16	99.22	109.38	116.51	136.14	164.85
序号	9	10	11	12	13	14	15	16
年份	1993	1994	1995	1996	1997	1998	1999	2000
值	229.56	289.82	330.35	406.95	488.82	525.8	576.39	643.26

(数据来源:《西安市统计年鉴》2003年版)

a. 根据西安市国内生产总值构建序列 $x_{(0)}$ 和 $Q_{(0)}$ 如下:

$$x_{(0)} = (x_{(0)}(1985), x_{(0)}(1986), \dots, x_{(0)}(2000))$$

$$= (57.58, 65.78, \dots, 643.26)$$

$$Q_{(0)} = (1985, 1986, \dots, 2000)。$$

为方便计算,以1985年为1,转换序列 $Q_{(0)} = (1, 2, \dots, 16)$ 。

b. 画出折线图(见图1),找出跳变点。使用EXCEL可以很方便地画出折线图和观察出跳变点。

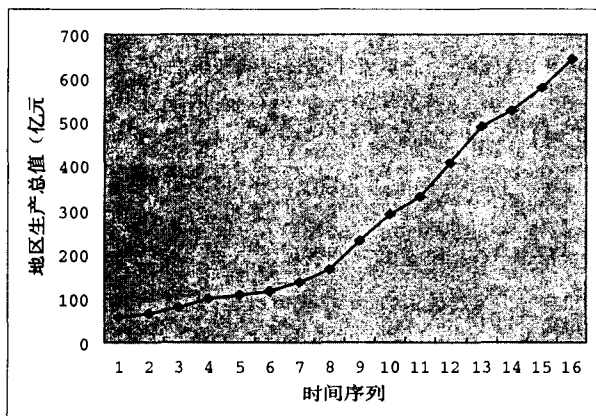


图1 西安市国内生产总值时间序列折线图

判别跳变点的原则:一是该数据点相对于前两个连续点显示出非线性特点明显的不在同一直线上。二是由于跳变点要根据前两个连续点而定,折线图的第一个点不作为跳变点处理。由图观察出跳变点序列及其时间序列如下:

$$x_{A(0)} = (x_{A(0)}(1), x_{A(0)}(2), x_{A(0)}(3), x_{A(0)}(4), x_{A(0)}(5)) \\ = (109.38, 136.14, 229.56, 406.95, 525.8)$$

$$Q_{A(0)} = (Q_{A(0)}(1), Q_{A(0)}(2), Q_{A(0)}(3), Q_{A(0)}(4), Q_{A(0)}(5)) \\ = (5, 7, 9, 12, 14)$$

c. 采用GM(1,1)模型分别计算出跳变值和相应时间函数的参数并预测下一个跳变点。由GM(1,1)模型计算得:

$\hat{x}_{A(1)}(k+1)$ 的 a_p 为 -0.4056 , μ_p 为 85.1746 , $Q_{A(1)}(k+1)$ 的 a_p 为 -0.2260 , μ_p 为 5.3018 , 从而得到预测方程如下:

$$\begin{cases} \hat{x}_{A(1)}(k+1) = 106.4882e^{0.4056k} & (10) \\ Q_{A(1)}(k+1) = 5.7568e^{0.2260k} & (11) \end{cases}$$

根据方程(10)、(11)预测下面两个跳变点。

$k=5$ 时,得:

$$\begin{cases} \hat{x}_{A(1)}(6) = 106.4882e^{0.4056 \times 5} = 809.25 & (12) \\ Q_{A(1)}(6) = 5.7568e^{0.2260 \times 5} = 17.822 \approx 18 & (13) \end{cases}$$

$k=6$ 时,得:

$$\begin{cases} \hat{x}_{A(1)}(7) = 106.4882e^{0.4056 \times 6} = 1214.05 & (14) \\ Q_{A(1)}(7) = 5.7568e^{0.2260 \times 6} = 22.34 \approx 22 & (15) \end{cases}$$

可知下两个跳变点分别是2002年($Q_{A(1)}(6) \approx 18$, 2002年对应 $Q_{(0)}$ 序列的18), 预测值为809.25亿元和2006年($Q_{A(1)}(7) \approx 22$, 2006年对应 $Q_{(0)}$ 序列的22), 预测值为1214.05亿元。

d. 计算跳变点之间的预测值。由表1对应的序号和使用线性回归模型的判别方法可知2001年属于[1997, 2001]区间, 使用该区间数据建立线性回归模型得:

$$y = 51.391x - 186.6 \quad (16)$$

其中, x 为 $Q_{(0)}$ 中该年的序号。

2001年,在 $Q_{(0)}$ 中对应 $x=17$,代入方程得该年的国内生产总值为687.047亿元。

同理,可知2003年属于[2001, 2005]区间,使用该区间的数据建模,即由点(17, 687.047)和(18, 809.25)得:

$$y = 122.2x - 1390.4 \quad (17)$$

2003年,在 $Q_{(0)}$ 中对应 $x=19$,代入方程得该年的国内生产总值为931.4亿元。

2.2 西安市国内生产总值预测模型的对比检验 下面就GM(1,1)模型与线性回归组合预测模型、线性回归模型以及GM(1,1)模型在该指标值的预测精度上进行了比较。

表2 三种预测方法比较(西安市国内生产总值2000~2002年)

年份	实际值	组合预测模型		线性回归模型		GM(1,1)模型	
		预测值	精度(%)	预测值	精度(%)	预测值	精度(%)
2001	733.85	687.047	93.19	613.902	80.46	700.04	95.17
2002	823.5	809.25	98.24	654.357	74.15	768.14	92.79
2003	941.6	931.4	98.90	694.812	64.48	842.87	88.29

表2中的预测精度显示,组合预测模型要比GM(1,1)模型在长期预测上精度要高,短期预测的表现也很不错。线性回归模型对这种类似指数型增长的指标,预测精度难以满足需要。可见,组合预测模型可以在西安市国内生产总值的预测上得到更为满意的效果。

3 结论

本文建立的西安市国内生产总值预测模型,可以很好地预测跳变点,从而得到分段的线性回归模型。检验结果表明,该模型可以比较准确地预测西安市国内生产总值,在实际应用中也取得满意的效果。实践证明,在国民经济其他指标的预测中也可以使用该方法建模,而且不需要大量数据就可以得到比较满意的预测效果。但在建模过程中,跳变点序列是靠人工观察图形得到的,可能由于个人主观因素而得到不尽相同的跳变点序列。

参考文献

- 1 许秀莉,罗健.一种新的组合灰色神经网络预测模型.厦门大学学报,2002;(3)
- 2 鲍一丹等.基于GM(1,1)模型和线性回归的组合预测新方法.系统工程理论与实践,2004;(3)
- 3 邓聚龙.灰预测与灰决策.武汉:华中理工大学出版社,2002
- 4 傅立.灰色系统理论及其应用.北京:科学技术文献出版社,1992
- 5 (美)布洛克威尔等著,田铮译.时间序列的理论与方法(第2版).北京:高等教育出版社,2001
- 6 陈业华.灰色灾变预测模型及其应用.北京航空航天大学学报,1998;(1)

(责编:暴王京)