

我国月度CPI的组合预测及分析

韩春蕾^{1,2},高婉君³

(1.滨州医学院 卫生管理学院,山东 烟台 264003;2.山东大学 经济学院,济南 250100;
3.西安交通大学 医学院,西安 710061)

摘要:文章利用我国2000年1月到2011年12月的月度CPI历史时间序列数据,分别运用季节性ARIMA模型和趋势外推法(抛物线模型)对CPI序列进行拟合,并进行两种模型的组合预测,然后对三种模型的预测精度分析比较。结果显示,组合预测模型的预测精度要优于单一模型的预测精度,可以在CPI的预测中应用和推广。

关键词:CPI;ARIMA模型;趋势外推法;组合预测

中图分类号:F201;C81 **文献标识码:**A **文章编号:**1002-6487(2014)01-0011-03

0 引言

消费者物价指数(CPI)是反映消费者购买商品价格水平的统计指标,表示与居民生活密切相关的商品和服务的价格变化趋势或幅度。CPI涉及经济发展,关乎社会和谐,维系人民生计,是管理者制定宏观经济政策、分析货币市场和债券市场及央行公开市场业务的重要参考依据,历来备受政府和民众关注。因此对CPI的精确预测具有重要的现实意义。

本文分别用ARIMA模型和趋势外推模型两种单项模型和组合预测模型对CPI进行分析与预测,并将预测精度进行对比,以期对未来CPI的走势有更精确地把握,从而对未来的经济政策以及宏观经济决策提出相关建议。

1 资料来源和方法

1.1 资料来源

通过查阅2001~2012年的《中国统计年鉴》,得到我国2000年1月到2011年12月的月度CPI历史数据,以1999年各月度为基期进行预处理。

1.2 研究方法

1.2.1 ARIMA模型

当序列存在季节性时,需进行季节差分,步长为季节周期的长度,参数D是为得到非季节性时间序列所进行的季节差分阶数。在进行季节性差分处理后可能仍存在季节性相关和偏相关,P为季节性偏自相关阶数,Q为季节性自相关阶数。最后得到的ARIMA(p,d,q)(P,D,Q)用以下方程表示:

$$\varphi_p(B)\Phi B^s \nabla^d \nabla^D x_t = \theta_q(B)\Theta B^s \mu_t \quad (1)$$

其中B为一般滞后算子, B^s 为季节性滞后算子,S为季节周期时长^[1]。

1.2.2 趋势外推法

(1)图形识别法。以时间t为横轴、观察值为纵轴,将时间序列数据绘制成散点图,观察选择较为合适的函数曲线模型。

(2)差分法。差分后非平稳序列可变为平稳序列。根据差分特性,可选择合适的趋势模型。常用的趋势预测模型有四种:即多项式模型($\hat{y}_t = b_0 + b_1t + b_2t^2 + \dots + b_nt^n$);指数曲线预测模型($\hat{y}_t = ae^{bt}$),修正指数曲线预测模型($\hat{y}_t = a + bc^t$);对数曲线预测模型($\hat{y}_t = a + b \ln t$);生长曲线预测模型(皮尔曲线 $\hat{y}_t = L/1 + ae^{-bt}$,龚珀兹曲线 $\hat{y}_t = ka^{bt}$)^[2]。

1.2.3 组合预测

本文选取常用的方差倒数法确定权重,表达式为:且

$$W_j = D_j^{-1} / \sum_{j=1}^m D_j^{-1} \text{ 且 } \sum_{j=1}^m W_j = 1, j = 1, 2, \dots, m. \quad (2)$$

式中 D_j 为第j个模型的误差平方和,即

$$D_j = \sum_{i=1}^m (Y_i - \hat{Y}_{ij})^2$$

组合预测的形式采用线性组合模型:

$$y_{0t} = w_1 y_{1t} + w_2 y_{2t} + \dots + w_n y_{nt} \quad (3)$$

式中, y_{0t} 为t期的组合预测值; $y_{1t} \dots y_{nt}$ 为n种不同单项预测模型在t期的预测值; $w_1 \dots w_n$ 为相应的n种组合权重, w_i 依据方差倒数法加以确定。

当n=2时: $w_1 = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2); w_2 = 1 - w_1$

式中, σ_i^2 (i=1,2)为第i种单项预测模型的残差方差。

基金项目:山东省社会科学规划研究项目(13DJJJ24);山东省统计科研重点研究课题(KT12131);山东省应用统计学会重点研究课题(2012yytj001);烟台市社会科学规划研究项目(2012-JJ-03)

作者简介:韩春蕾(1981-),女,山东莱西人,博士研究生,讲师,研究方向:数量经济。

高婉君(1987-),女,山西运城人,硕士,研究方向:统计预测。

当 $n > 2$ 时: $w_t = 1/Q_t$ ($t=1,2,\dots,n$), Q_t 为第 t 种单项预测模型的残差平方和。^[3]

1.2.4 预测精度

通常评价预测精度的方法有以下几种:

① 预测误差的方差(MSE)和标准差(SDE)

$$MSE = \frac{\sum_{i=1}^n e_i^2}{n} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

$$SDE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

② 平均相对误差(MPE)和平均相对误差绝对值(MAPE)

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} \quad (6)$$

因为MPE存在正负抵消的问题,因此,MAPE应用更为广泛,公式如下:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (7)$$

2 我国月度CPI预测分析

2.1 ARIMA模型预测

2.1.1 平稳性检验

对CPI和一阶差分序列(DCPI)绘制时序图(图1),可以看出原序列有明显的增长趋势,而差分序列基本平稳。对DCPI进行ADF单位根检验, $T = -10.2057$, 相伴概率 $P = 0.000$, 因此它是平稳的。

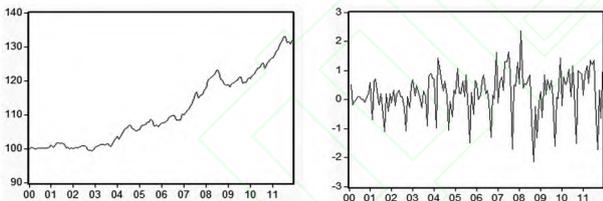


图1 CPI及DCPI序列的时序图

运用Eviews6.0软件实现相关图的绘制,一阶差分序列的相关图如图2^[4]:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.161	0.161	3.7812	0.052	
2	0.136	0.113	6.4902	0.039	
3	-0.026	0.066	6.5922	0.086	
4	-0.066	0.071	7.2367	0.124	
5	0.056	0.061	7.3417	0.196	
6	-0.117	0.119	9.4004	0.162	
7	0.043	0.065	9.6868	0.207	
8	-0.075	0.066	10.562	0.228	
9	-0.111	0.115	12.453	0.189	
10	-0.002	0.043	12.453	0.256	
11	0.106	0.154	14.225	0.221	
12	0.447	0.406	46.870	0.000	
13	0.034	0.136	46.053	0.000	
14	0.063	0.068	46.687	0.000	
15	-0.124	0.132	49.181	0.000	
16	-0.162	0.111	53.871	0.000	
17	-0.080	0.040	54.535	0.000	
18	-0.209	0.130	61.767	0.000	
19	-0.055	0.079	62.267	0.000	
20	-0.210	0.128	69.734	0.000	
21	-0.165	0.041	74.356	0.000	
22	0.034	0.134	74.556	0.000	
23	-0.037	0.137	74.789	0.000	
24	0.440	0.325	108.52	0.000	

图2 一阶差分序列DCPI的自相关图与偏自相关图

由图2可知,一阶差分序列平稳,但滞后12期的自相关系数和偏自相关系数分别是0.447和0.406,说明存在周期为12个月的季节波动。

2.1.2 模型定阶

建立类似ARIMA(2, 1, 1)(1, 1, 1)¹²季节模型,对各种

组合进行检验并进行对比,对比结果见表1:

表1 四个模型的相关检验结果

(p,q)	调整后的决定系数	AIC	SC
(2,1)	0.7727	1.4974	1.6154
(2,0)	0.7740	1.4772	1.5717
(1,1)	0.7571	1.5182	1.6121
(1,0)	0.7569	1.5017	1.5723

注:AIC值为赤池准则的特征值;SC为施瓦茨准则的特征值。

比较表1的各项结果,(p,q)取值为(2,0)时,调整后的决定系数(Adjusted R²)最大且AIC值与SC值最小。综上,选择模型ARIMA(2, 1, 0)(1, 1, 1)¹²作为最终的预测模型,即:

$$\nabla \nabla_{12} CPI = \frac{1 + 0.885800B^{12}}{(1 - 0.210662B - 0.164656B^2)(1 + 0.195275B^{12})} \epsilon_t$$

2.1.3 模型的检验

① 参数检验

表2 ARIMA模型的参数检验

变量	系数	标准误	T统计量	P值
AR(1)	0.2106	0.0819	2.5700	0.0115
AR(2)	0.1646	0.0776	2.1198	0.0362
SAR(12)	-0.1952	0.0758	-2.5731	0.0114
MA(12)	-0.8858	0.0001	-5217.5	0.0000

由表2可得,各参数的t检验值及相应的P值显示,各参数均有统计学意义, $DW = 2.0482 \approx 2$, $adj.R^2 = 0.7740$, 说明模型构建比较理想。

② 模型检验

残差的自相关系数和偏自相关系数检验显示,残差的12阶延迟已经不存在相关性;残差序列无相关性($P = 0.4412$); $MAPE = 0.3396\%$, 预测精度比较高。

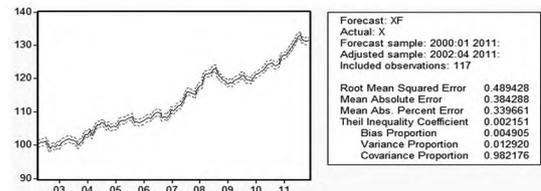


图3 ARIMA模型预测图

2.2 趋势外推法预测

我国月度CPI的时间序列时序图(图1)显示了时间序列呈现非线性趋势分布,提示采用趋势外推拟合。本文将时间 t 作为自变量,将我国月度CPI数据作为因变量进行趋势拟合。且通过计算可以得到,CPI序列的二阶差分大致相等,所以可以建立CPI序列的二次预测模型,即抛物线模型: $\hat{y}_t = b_0 + b_1 t + b_2 t^2$ 。

由图1可以看出原始序列具有明显的季节性,所以先对月度CPI序列进行季节调整(如图4),可以看出季节性得到消减。

对季节调整后的序列进行方程估计,得到预测模型: $CPI_{sa} = 99.21729884 + 0.021140129t + 0.0014605696t^2$

$adj.R^2 = 0.973264$, $F = 2603.845$, 方程和系数均通过显著性检验。 $MAPE = 1.1068\%$, 预测精度相对于ARIMA模型较低。

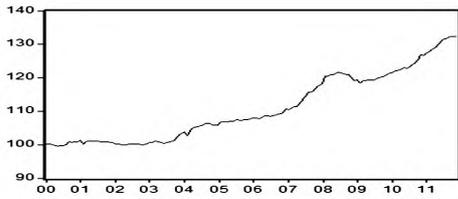


图4 季节调整后的时序图

表3 抛物线模型的参数检验

变量	系数	标准误	T统计量	P值
C	99.2173	0.4191	236.7357	0.0000
t	0.0211	0.0133	1.5842	0.0154
t ²	0.0015	8.91E-05	16.3839	0.0000

2.3 组合预测模型拟合

对比以上模型的预测效果可知,趋势外推拟合模型的预测误差较大,可能是因为影响CPI的因素复杂,不能简单的用时间变量t来解释,而ARIMA模型的预测精度较高。为提高预测精度,考虑建立组合预测模型如下:

$$y_{0t} = \omega_1 y_{1t} + \omega_2 y_{2t}$$

其中, y_{0t} 为组合预测模型的预测值; ω_1, ω_2 为单项预测模型的权重系数; y_{1t}, y_{2t} 分别为ARIMA(p,d,q)模型和抛物线模型的预测值。根据权重公式(2), 得 $\omega_1 = 0.6571, \omega_2 = 0.3429$, 则组合预测模型为:

$$y_{0t} = 0.6571y_{1t} + 0.3429y_{2t}$$

2.4 各模型拟合效果与预测精度评价

用各模型对我国2011年的月度CPI数据进行预测并与实际值比较,结果见表4和图5:

表4 2011年我国月度CPI实际值和预测值对比

月份	实际值	ARIMA模型		抛物线模型			组合预测			
		预测值	绝对误差	相对误差(%)	预测值	绝对误差	相对误差(%)	预测值	绝对误差	相对误差(%)
1	104.9	105.250	0.350	0.334	105.427	0.527	0.502	105.340	0.440	0.420
2	104.9	104.549	-0.351	-0.334	105.301	0.401	0.382	104.836	-0.064	-0.061
3	105.4	104.987	-0.413	-0.392	105.520	0.120	0.114	105.199	-0.201	-0.191
4	105.3	105.374	0.074	0.070	105.414	0.114	0.108	105.417	0.117	0.111
5	105.5	105.009	-0.491	-0.466	105.333	-0.167	-0.158	105.149	-0.351	-0.332
6	106.4	105.892	-0.508	-0.478	105.750	-0.650	-0.610	105.873	-0.527	-0.496
7	106.5	106.185	-0.315	-0.296	105.636	-0.864	-0.812	106.026	-0.474	-0.445
8	106.2	106.400	0.200	0.188	105.453	-0.747	-0.704	106.105	-0.095	-0.090
9	106.1	106.442	0.342	0.322	105.494	-0.606	-0.571	106.146	0.046	0.044
10	105.5	105.321	-0.179	-0.170	105.078	-0.422	-0.400	105.267	-0.233	-0.221
11	104.2	104.466	0.266	0.255	104.169	-0.031	-0.030	104.393	0.193	0.186
12	104.5	104.081	-0.419	-0.401	104.644	0.144	0.137	104.303	-0.197	-0.188

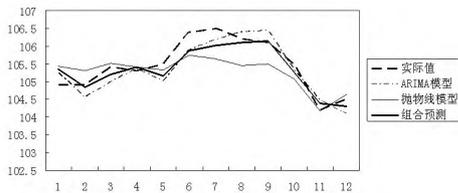


图5 三种模型预测值与实际值的比较

可见,组合模型的预测误差明显小于ARIMA模型和抛物线模型,且组合预测值较其他两种模型的预测值更接近实际观测值。三种模型预测精度比较如下:

表5 模型精度比较

模型	MSE	SDE	MAPE
ARIMA模型	0.1213	0.3483	0.0031
抛物线模型	0.2323	0.4820	0.0038
组合模型	0.0849	0.2914	0.0023

由表5可以看到组合预测模型的预测误差的方差,平均相对误差和平均相对误差绝对值均比ARIMA模型和抛物线模型小,说明组合预测模型的精度高于各单一模型。

2.5 外推预测

运用组合预测模型对我国2012年1月到12月的月度CPI进行预测,预测结果见表6,预测值非常接近实际值。

表6 2012年月度CPI的组合预测

月份	实际值	预测值
1	104.5	104.5844
2	103.2	104.3702
3	103.6	103.9274
4	103.4	103.8974
5	103	103.4501
6	102	102.7155
7	101.8	102.1657
8	102.0	102.2554
9		102.7496
10		102.8464
11		103.4195
12		103.0018

备注:截至定稿时,我国月度CPI的实际值只公布至2012年8月份。

3 结论

(1)CPI趋势预测的方法虽然很多,但是由于近几年来世界经济的不断发展与变化,全球经济一体化进程的不断深入,消费者物价水平受到各种不确定因素的影响。因此在建立模型的时候,无法将所有因素考虑进去,模型仅能使用历史数据所反映出来的基本信息,CPI预测的难度越来越大。本文运用组合预测模型对2012年的CPI进行了预测,说明组合预测可以应用到CPI的预测。

(2)组合预测模型综合了单项预测模型的优点,从而减少预测的系统误差,提升预测效果,提高预测精度。

(3)本文运用组合预测模型进行预测,得到我国2012年1到12月份的CPI预测值,与真实值接近,预测精度较高。

参考文献:

[1]王燕.应用时间序列分析[M].北京:中国人民大学出版社,2008.
 [2]徐国祥.统计预测与决策[M].上海:上海财经大学出版社,2008.
 [3]汪同三,张涛.组合预测:理论、方法及应用[M].北京:社会科学文献出版社,2008.
 [4]张晓峒. Eviews使用指南与案例(数量经济学应用系列)[M].北京:机械工业出版社,2007.

(责任编辑/亦 民)