

中国能源消费的时间序列分析

胡俊娟

(浙江科技学院 理学院,杭州 310023)

摘要: 对1978—2010年中国能源消费总量进行时间序列分析,采用时序分析的ARIMA模型和Autoregressive模型对其拟合建模,并对这两类模型的拟合效果和预测效果进行了比较。模型的残差检验和参数显著性检验结果表明模型是适用的,同时表明中国能源消费总量存在短期自相关性,并在短期内还将保持较快的增长速度。

关键词: ARIMA模型; Autoregressive模型; 中国能源消费

中图分类号: O212; F222

文献标志码: A

文章编号: 1671-8798(2013)03-0164-04

Time series analysis of Chinese energy consumption

HU Junjuan

(School of Sciences, Zhejiang University of Science and Technology, Hangzhou 310023, China)

Abstract: We use ARIMA model and Autoregressive model to fit the data of Chinese total energy consumption from 1978 to 2010 by time series analysis methods, and make a comparison between ARIMA model and Autoregressive model according to the effect of fitting and forecasting. The results of residual test and the significant test of parameter model show that the model is applicable. Fitting models indicate short term collection of national energy consumption, which have a rapid growth rate during the short period.

Key words: ARIMA model; Autoregressive model; Chinese energy consumption

随着能源短缺问题的日益突出,能源消费问题备受关注^[1-4]。研究者对中国能源的内部结构用各种方法进行了分析,文献[4]中作者采用了ARIMA模型对中国能源消费进行了预测分析。本研究采用时间序列分析中的ARIMA模型和Autoregressive模型来对能源消费数据进行分析,以此揭示中国能源消费的趋势和特点,对中国的能源消费进行分析预测。

1 ARIMA模型和Autoregressive模型理论简介

1.1 ARIMA模型

ARIMA模型^[5-6]被众多学者广泛应用于各种类型的时间序列数据的分析和处理,是精度较高的短

收稿日期: 2012-12-04

作者简介: 胡俊娟(1979—),女,浙江省兰溪人,讲师,硕士,主要从事数据挖掘和数据分析研究。

期预测方法之一。它是依据变量本身的变化规律,对数据进行分析。该模型结构为:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\epsilon_t \\ E(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \sigma_\epsilon^2, E(\epsilon_t \epsilon_s) = 0, s \neq t \\ E(x_s \epsilon_t) = 0, \forall s < t \end{cases}$$

其中: $\Phi(B) = (1 - \varphi_1 B - \dots - \varphi_p B^p)$; $\Theta(B) = (1 + \theta_1 B + \dots + \theta_q B^q)$; $B^j x_t = x_{t-j}$; ∇ 为差分算子。

1.2 Autoregressive 模型

Autoregressive 模型是一种通过确定性因素分解方法,提取序列中主要的确定性信息以后,对残差序列相关性建立自回归的模型,因而, Autoregressive 模型又称为(残差)自回归模型,该模型构造为:

$$\begin{cases} x_t = T_t + S_t + \epsilon_t \\ \epsilon_t = \varphi_1 \epsilon_{t-1} + \dots + \varphi_p \epsilon_{t-p} + z_t \\ E(z_t) = 0, \text{Var}(z_t) = \sigma_z^2, E(z_t z_s) = 0, s \neq t \end{cases}$$

其中: T_t —趋势效应拟合; S_t —季节效应拟合。

2 实证

以中国能源消费数据为变量(记为 x_t , 单位是万t 标准煤),选取 1978—2010 年数据来分析(数据来源于中国统计年鉴)。原始数据图(图 1)有明显的指数上升趋势,是非平稳时序。为了避免数据存在异方差性,对它取自然对数(记为 $\ln x_t$)。

2.1 ARIMA 模型拟合数据

从图 2 可以看出,数据取对数之后有显著的线性递增趋势,所以首先对其进行一阶差分,差分后序列的白噪声检验结果如表 1 所示。

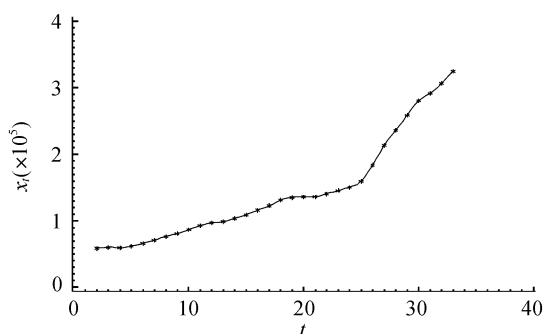


图 1 中国能源消费数据散点图

Fig. 1 Scatter diagram of Chinese energy consumption

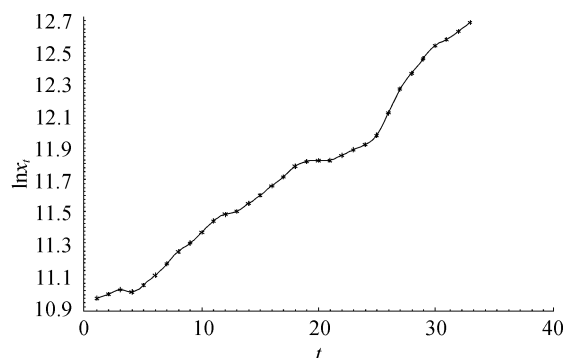


图 2 中国能源消费自然对数散点图

Fig. 2 Scatter diagram of logarithmic Chinese energy consumption

从白噪声检验(表 1)可以得出差分后序列不是白噪声序列,序列之间还存在着一定的相关关系,从而考察其自相关图(图 3)和偏自相关图(图 4)。

从图 3 和图 4 可以看出,样本自相关性具有拖尾属性,偏自相关性呈现截尾属性,所以用 ARIMA(1, 1, 0)模型拟合 $\{\nabla \ln x_t\}$, 得到拟合模型:

$$(1 - 0.676\ 01B)\nabla \ln x_t = 0.049\ 84 + \epsilon_t,$$

其中 ϵ_t 独立同分布,服从 $N(0, 0.000\ 697)$ 。

模型的残差检验和参数显著性检验结果分布如表 2 和表 3 所示,残差检验结果表明残差是白噪声序列,参数检验结果表明参数显著,模型可用。

表 1 差分序列 $\{\nabla \ln x\}$ 的白噪声检验

Table 1 White noise test of series of $\{\nabla \ln x\}$

延迟阶数	LB 统计量	P 值
6	29.59	0.000 5
12	32.77	<0.000 1

Autocorrelations																								
Lag	Covariance	Correlation	−1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	Std Error
0	73 371 456	1.000 00																	*****					0
1	48 530 302	0.661 43										.							*****					0.208 514
2	26 660 980	0.363 37										.							*****			.		0.285 519
3	24 655 633	0.336 04										.							*****			.		0.304 963
4	10 789 556	0.147 05										.							***			.		0.320 659
5	−1 178 875	−0.016 07										.									.			0.323 578
6	−8 662 511	−0.118 06										.					**				.			0.323 612
7	−18 027 298	−0.245 70										.				*****				.				0.325 480
8	−21 898 620	−0.298 46										.				*****				.				0.333 446

图 3 自相关函数图

Fig. 3 Diagram of autocorrelations

Partial Autocorrelations

Lag	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
1	0.661 43																					
2	-0.131 77																					
3	0.273 51																					
4	-0.342 68																					
5	0.080 15																					
6	-0.285 48																					
7	-0.007 82																					
8	-0.136 71																					

图 4 偏自相关函数图

Fig. 4 Diagram of partial autocorrelations

表 2 ARIMA(1,1,0)的白噪声检验

Table 2 White noise test of ARIMA(1,1,0) model

延迟阶数	LB 统计量	P 值
6	6.20	0.286 9
12	8.51	0.667 1
18	16.77	0.470 4
24	26.20	0.291 3

表 3 参数的显著性检验

Table 3 Parameter test of significance

参数	估计值	T 统计量	P 值
常数	0.049 84	0.049 84	0.000 5
φ_1	0.676 01	0.676 01	<0.000 1

2.2 用 Autoregressive 模型拟合数据

根据图 2 的对数数据,对趋势效应拟合线性回归模型,对残差序列 $\{\epsilon_t\}$ 进行自相关性检验,检验结果 Durbin-Watson 统计量值为 0.178 3,正自相关性的 $P < 0.000 1$,说明残差具有正自相关性。用向后逐步回归的结果是残差序列适合 AR(2)模型,则 Autoregressive 模型为:

$$\begin{cases} \ln x_t = 10.819 5 + 0.053 3t + \epsilon_t \\ \epsilon_t = 1.587 7\epsilon_{t-1} - 0.739 9\epsilon_{t-2} + z_t \\ E(z_t) = 0, \text{Var}(z_t) = 0.016 646, E(z_t z_s) = 0, s \neq t \end{cases}$$

对模型进行参数显著性检验,根据表 4 可知所有的参数显著,整个模型的拟合优度 R^2 为 0.998 1,说明模型的拟合效果很好。

3 拟合模型的比较

从拟合效果来看,比较表 5 中的信息量 AIC 和

表 4 参数的显著性检验

Table 4 Parameter test of significance

参数	估计值	T 统计量	P 值
常数	10.819 5	201.95	<0.000 1
t 的系数	0.053 3	19.25	<0.000 1
φ_1	1.587 7	-12.79	<0.000 1
φ_2	-0.739 9	5.76	<0.000 1

SBC 的值,用信息量最小原则,可以看出, Autoregressive 模型的拟合效果略好于 ARIMA(1,1,0)模型。从预测效果来看,比较表 6 中的绝对误差值,可以看出 ARIMA(1,1,0)模型比 Autoregressive 模

表 5 拟合效果比较

Table 5 Comparison of fitting results

模型	AIC	SBC
ARIMA(1,1,0)模型	-139.834	-136.903
Autoregressive 模型	-145.515 49	-139.529 46

型要好,可用于短期预测。用 ARIMA(1,1,0)模型得 2011 年的预测值是 343 419.9 亿 t,与国家统计局核算的 34.78 万亿 t 标准煤相比,相对偏差度不到 1.3%。2012 年的预测值为 362 308.6 亿 t,据中国能源研究会发布的《2012 中国能源发展报告》预计,2012 年中国一次能源消费总量约为 36.2 万亿 t 标准煤,模型预测值与报告估计值相对偏差度不到 1%。从拟合模型的形式看,模型都表示短期自相关性。这点表明如果前一两年的能耗高,通常后一年的也会较高。2012 年比 2011 年增长约 4%,中国能源消耗还将保持较高的增长趋势。模型就数据分析给出了间隔一期、二期相关,并没有考虑更长期相关的情况,所以长期是否相关有待进一步检验。短期相关性表明,间隔时间较短时,能源消耗的影响就大,所以要想改变高能耗的情况,需要尽早地制定相关的节能政策和方针。

表 6 预测效果比较

Table 6 Comparison of forecasting results

万 t

年份	原始序列值	ARIMA(1,1,0)模型		Autoregressive 模型	
		预测值	绝对误差	预测值	绝对误差
2006	258 676	256 680.2	-1 995.82	234 712.5	-23 963.5
2007	280 508	279 710	-798.006	247 570.1	-32 937.9
2008	291 448	301 124.8	9 676.793	261 132.1	-30 315.9
2009	306 647	303 953.1	-2 693.91	275 436.9	-31 210.1
2010	324 939	322 534.7	-2 404.29	290 525.4	-34 413.6
2011	约 347 800	343 419.9			
2012	约 362 000	362 308.6			

4 结 语

本研究对 1978—2010 年中国能源消费总量进行时间序列分析,从 ARIMA 模型和 Autoregressive 模型中可以看出,中国能源消费总量还呈现出较高的增长趋势、短期相关性,因而,要改变中国高能耗的情况,需要尽早地制定相关的节能政策和方针。

参考文献:

- [1] 姜磊,季民河. 技术进步、产业结构、能源消费结构与中国能源效率:基于岭回归的分析[J]. 当代经济管理,2011,33(5):13-16.
- [2] 许秀川,罗倩文. 中国经济增长与能源消费:基于内生结构模型的实证研究[J]. 统计与决策,2008(16):69-71.
- [3] 刘晶. 内蒙古能源与经济发展关系的实证研究[J]. 中外能源,2010(1):23-28.
- [4] 杨伟传. 中国能源消费的 ARIMA 模型预测分析[J]. 统计与决策,2009(11):71-72.
- [5] 王燕. 应用时间序列分析[M]. 2 版. 北京:中国人民大学出版社,2008.
- [6] Brockwell P J, Davis R A. Introduction to time series and forecasting[M]. New York: Springer,2010.