

中国能源需求的模型选择

易金超 厦门大学经济学院

[摘要]主要想探讨一下1953 - 2006年来中国能源需求的变化情况,对原始数据进行对数差分和单位根检验后,运用ARIMA模型对能源需求进行拟合与预测,最后选择一个比较好的拟合模型来显示能源需求的变化。

[关键词]能源需求 单位根检验 ARIMA模型

一、文献回顾

随着中国对能源需求量的不断增长,对能源的关注也更加迫切。国内外很多学者对影响能源需求的影响因素(林伯强,2003;陈江生,2005)、能源需求预测(林伯强,2001;卢二坡,2005)以及预测方法(梁巧梅,魏一鸣,范英等,2004;卢二坡,2006)做了大量的研究。他们根据历史数据运用计量方法,特别是ARMA模型(汪建均,胡宗义,2006)、协整理论(林伯强,2001、2006;林伯强、魏巍贤等,2007)与GARCH模型(邹艳芬,陆宇海,2006)以及虚拟变量法、投入产出法(梁巧梅,魏一鸣,范英等,2004),对影响能源需求的影响因素以及预测方法做了有益的研究。林(2001)利用JOHANSEN AND JUSELIUS(JJ)法以及ECM模型讨论了能源总需求、GDP、价格与经济结构(重工业)之间的关系,指出能源价格、收入是能源需求的重要决定因素,同时经济结构也有重要影响;其后又分别对电力需求(2006)以及煤炭需求(2007)的收入弹性、价格弹性、结构弹性进行了比较详细的研究。但是他们大多在一个比较短的时间段,或者加入了别的因素以及考虑某个方面的需求拟合(比如煤炭)。在上述研究成果的基础上,我想对1953 - 2006年的能源需求进行一个比较长期的拟合,单纯从序列自身出发利用时间序列模型进行拟合。

二、数据选取与分析方法说明

本文数据是根据中经网经济统计数据库公布的数据进行分析,主要选取的是能源消耗总量(用TCOE表示见图1,单位为万吨标准煤),从1953年到2006年,总共53个观察值。我们先对总的能源需求数据进行对数(见图2),然后进行差分(见图3)。由于我们建模必须是序列平稳,所有我们用ADF以及PP检验对对数序列以及对数差分序列的平稳性检验。

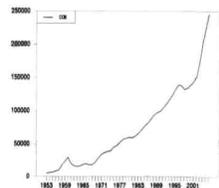


图1: 能源需求原序列

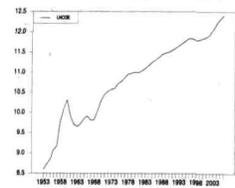


图2: 经过对数后的序列

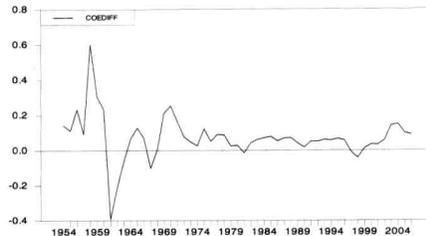


图3: 对数差分后的序列

我们分别讨论了三种类型:零均值无趋势项模型;有均值无趋势项模型;有趋势项模型。我们知道ADF检验有一个重要假设:

$$\text{var}(\varepsilon_t) = \sigma_\varepsilon^2$$

也就是说ADF检验主要适用于方差齐性的情况,它对于异方差以及自相关性的平稳性检验效果不是很理想。Phillips以及Perron(1988)对ADF检验进行了修正,提出了PP检验用于对存在异方差情况下的平稳性检验。我们在两个表格中,整理了在三种情况下的t值以及对应的概率值,从中我们可以发现不论是ADF还是PP检验,能源需求原序列对数显示的都是存在一个单位根,对数序列非平稳(在5%的范围内)。对数序列经过一阶差分后,无论是ADF检验还是PP检验,都发现可以拒绝原假设,即序列不存在单位根,表现平稳。所有我们可以对对数差分序列进行建模,但是我们要主要一个情况就是平稳序列是否为白噪音序列。如果是白噪音序列,那我们就没有建模的必要性了。由于小样本数据,根据Box和Ljung在Q统计量的基础上推导出的LB统计量:

$$LB = T(T+2) \sum_{k=1}^n \left(\frac{\hat{\rho}_k^2}{T-k} \right) \chi^2(n)$$

其中T为序列观测期数,n为指定的延迟期数,Box和Ljung推出LB统计量近似服从自由度为n的卡方分布。利用SAS软件分析我们可以得出,在滞后6期其LB值的概率值为0.0027,在滞后12期后其LB值为0.0031,在5%的范围内,我们可以对对数差分序列不是白噪音序列,即我们可以利用ARIMA模型进行平稳建模。

表1: 原始数据对数之后的单位根检验

类型	ADF test		PP test		结论(在5%的范围内)
	t-Statistic	Prob.*	t-Statistic	Prob.*	
类型1: 无均值无趋势项	2.06283	0.9898	3.03978	0.9992	存在单位根,序列不平稳
类型2: 有均值无趋势项	-1.51084	0.5202	-1.75779	0.3970	存在单位根,序列不平稳

类型3: 有均值有趋势项	-2.16462	0.4975	-3.08803	0.1197	存在单位根,序列不平稳
--------------	----------	--------	----------	--------	-------------

表2: 对数一阶差分后的单位根检验

类型	ADF test		PP test		结论(在5%的范围内)
	t-Statistic	Prob.*	t-Statistic	Prob.*	
类型1: 无均值无趋势项	-4.384295	0.0000	-3.68673	0.0004	不存在单位根,序列平稳
类型2: 有均值无趋势项	-4.168556	0.0018	-4.177929	0.0017	不存在单位根,序列平稳
类型3: 有均值有趋势项	-4.196360	0.0088	-4.245584	0.0076	不存在单位根,序列平稳

三、实证研究结果

在对数差分序列进行了平稳性检验后,我们可以根据序列的自相关系数以及偏自相关系数,确定模型的阶数,再运用ARIMA(求和自回归移动平均)模型进行数据拟合。为了更好的比较拟合的模型的可靠性,我们再这里选取1953 - 1997年的数据进行拟合,我们想留下一些观察值来评价模型的优劣性。本例中我们根据序列的自相关系数图以及偏自相关系数图发展了很多选择模型进行比较,从理论我们知道R值,AIC信息准则,SBC信息准则都可以为我们分辨选择那个模型更理想。

一般而言,ARIMA(p,d,q)模型是以下形式:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, t \neq s \\ E x_t \varepsilon_t = 0, \forall t < T \end{cases}$$

我们根据SAS拟合的自相关系数以及偏自相关系数图,发现自相关系数以及偏自相关系数都是滞后1阶与3阶在两倍的标准差之外,又是我们这里选择的比较模型中还用了疏系数模型,具体的模型选择可以通过下表来说明:

表三: 模型比较图

模型	参数估计	Pr> t 参数是否显著(5%)	残差是否白噪音(Pr >chisq)	AIC 准则	SBC 准则
	MU	0.0328*	6	-60.1982	
	0.07099	<.0001*	0.1804	-51.2773	

ARIMA(2,1,2)	MA1,1	<.0001*	12		
	-0.98824	0.0798	0.4959		
	MA1,2	0.9104	18		
	-0.73060		0.8923		
	AR1,1		24		
	-0.40705		0.9806		
	AR1,2				
	-0.02538				
JMA(1,(1,3))	MU	<.0001*	6	-62.086	
	0.07122	0.0032*	0.3347	-56.7334	
	MA1,1	<.0001*	12		
	-0.37575		0.4268		
	MA1,2		18		
	0.54307		0.7591		
ARI(1,3,1)	MU	0.0002*	6	-61.7673	
	0.07227	0.0015*	0.5117	-56.4148	
	AR1,1	0.0019*	12		
	0.43419		0.4384		
	AR1,2		18		
	-0.42548		0.7628		
ARIMA(1,3,1)	MU	<.0001*	6	-62.7535	
	0.07036	0.9360	0.5746	-53.8326	
	MA1,1	0.0178*	12		
	-0.01766	0.0195*	0.8313		
	MA1,2	0.6303	18		
	0.54065		0.9839		
ARIMA(3,1,1)	AR1,1	0.51802	24		
	AR1,2		0.9984		
	-0.10754				
	MU	0.07214	0.0001*	6	-60.4634
	MA1,1	0.0142*	0.0778	12	-55.1108
	-0.39374	0.0020*		18	
ARIMA(1,1,3)	AR1,1		0.1573		
	-0.48373		0.3854		
			24		
			0.6733		
	MA1,1	0.53639	0.0008*	6	-60.1477
	AR1,1	0.78824	<.0001*	0.3989	-56.5793
			12		
			18		
			0.9287		
			24		
			0.9855		
	JMA(1,3)	MU	0.07192	0.0073*	6
MA1,1			0.0012*	0.4839	-55.621
-0.51982			0.0519(*)	12	
MA1,2			0.0283*	0.6766	
-0.32690				18	
MA1,3		0.34059		0.9475	

				24	
				0.9928	
ARI(3,1)	MU	0.07208	0.0004*	6	-60.2518
	AR1,1	0.39612	0.0076*	0.4498	-53.115
	AR1,2	0.10204	0.5095	12	
	AR1,3		0.0021*	0.3294	
	-0.46358			18	
				0.7008	
				24	
				0.9187	

2011	12.5003	0.5210	11.4791	13.5215
------	---------	--------	---------	---------

注：*表示在5%范围内系数显著，残差白噪音检验分别选取了滞后6、12、18、24的值。

观察上面选择得模型，好几个模型都经过了系数得显著性检验，而且残差是白噪音。通过比较参数及其显著性，以及信息准则，我们最后选择了ARIMA(1,1,3)这个模型：

$$\Delta \ln coe_t = 0.78824 \Delta \ln coe_{t-1} + \epsilon_t - 0.53639 \epsilon_{t-3} \quad (1)$$

为了检验模型的有效性，我们选取了5个(1998年-2002年)没有参与拟合的对数数据，与拟合模型的估计值进行比较，假设样本区间是：T=n,n+1,...,n+h.记实际值为 y_t ,拟合值为 \hat{y}_t 我们可以得到均方根误差(MSE)，平均绝对误差(MAE)，平均相对误差(MPE)以及Theil不等系数(U),根据定义我们可以求得它们的值：

$$MSE(\sqrt{\frac{1}{h+1} \sum_{t=n}^{n+h} (y_t - \hat{y}_t)^2}) : 0.083112$$

$$MAE(\frac{1}{h+1} \sum_{t=n}^{n+h} |y_t - \hat{y}_t|) : 0.06687$$

$$MPE(\frac{1}{h+1} \sum_{t=n}^{n+h} \frac{|y_t - \hat{y}_t|}{y_t}) : 0.005627$$

$$U(MSE / \sqrt{\frac{1}{h+1} \sum_{t=n}^{n+h} y_t^2 + \frac{1}{h+1} \sum_{t=n}^{n+h} \hat{y}_t^2}) : 0.003516$$

MAE与MSE受因变量量纲影响，MPE与Theil不等系数是不受量纲影响的相对指标，通过对评价指标的分析，我们可以得出结论，选用ARIMA(1,1,3)能够较好的拟合能源需求模型。

在进行了模型有效性检验后，我们再来对所有的数据运用ARIMA(1,1,3)进行拟合，对53个观察值进行重新拟合后，我们得到了一个新的估计模型：

$$\Delta \ln coe_t = 0.80751 \Delta \ln coe_{t-1} + \epsilon_t - 0.53732 \epsilon_{t-3} \quad (2)$$

$$AIC = -81.2026 \quad SBC = -77.2620$$

新加进数据后再拟合，我们可以发现方程(1)与方程(2)没有发现结构性的变化。利用方程(2)，我们可以预测对数序列的5期拟合值。其结果如下：

	Obs	Forecast	Std Error	95% Confidence Limits
2007	12.4671	0.1104	12.2507	12.6835
2008	12.5052	0.2281	12.0592	12.9522
2009	12.5032	0.3546	11.8081	13.1983
2010	12.5016	0.4480	11.6275	13.3757

四、经济解释及结论

在第三部分我们讨论了我国能源需求的模型选择，从中我们可以得出一些有益的结论。

我们通过模型比较好的拟合了尽50多年的能源需求，拟合效果比较好；并且我们还进行了能源对数的预测。但是我们发现文中还有很多不足的地方，需要以后进行改进。比如，再预测中，只是对原序列的对数进行预测，如果指数化后，误差很大，如何减少预测误差是值得考虑的一个方面；还有就是能源需求是否是一个非线性序列，如果用非线性方法来拟合，是不是效果更好；对中国能源需求的拟合是用总体数据来分析，还是分别考虑石油、天然气、煤炭、水电的拟合，再综合起来考虑呢？那种拟合效果更好呢？这些是笔者以后努力的方向。

参考文献

- [1]陈平，中国粮价与通货膨胀关系(1987-1999)[J]. 经济学季刊,2002(4)
- [2]林伯强，中国能源需求的计量分析[J]. 统计研究,2001(10)
- [3]林伯强，中国电力发展：提高电价和限电的经济影响[J]. 经济研究,2006(5)
- [4]林伯强，魏巍贤、李丕东. 中国长期煤炭需求：影响与政策选择[J]. 经济研究, 2007(2)
- [5]卢二坡，我国能源需求预测模型研究[J]. 统计与决策, 2005(10)
- [6]汪建均，胡宗义，ARMA模型在我国电力需求预测中的应用[J]. 经济数学, 2006(3)
- [7]Walter Enders, 应用计量经济学时间序列分析[M]. 高等教育出版社，北京：2006
- [8]于俊年，计量经济学软件Eviews的使用[M]. 对外经济贸易大学出版社，北京：2006

作者简介：易金超(1986-)，江西宜春人，厦门大学经济学院金融系学生。