

标么值加权平均组合变量生成法用于中长期 电力负荷预测

张筱慧

(中国农业大学信息与电气工程学院, 北京 100083)

摘要: 将标么值概念引入电力负荷预测, 将预测中使用的部分自变量做标准化处理, 可以消除变量取值单位对回归系数的影响, 同时寻找变量间及变量在不同地区间的相似性。利用变量的加权线性组合生成新变量, 利用相似性对一些变量的发展规律作出估计, 使偏最小二乘回归方法能够应用于我国农村电网。

关键词: 中长期电力负荷预测; 组合变量; 偏最小二乘回归

Per-unit value weighted average combined variable generated method for mid-long term power load forecasting

ZHANG Xiao-hui

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: The concept of per unit value is introduced into power load forecasting. Some independent variables used in forecasting is standardized to eliminate the effect of different unit variables to regress coefficients. At the same time, the comparability between different variables and variables in different areas are studied. New variables are created using the weighted linear combination of original variables, and the variable progress rule is estimated based on the comparability, and partial least-square regression can be used in power network of our country.

Key words: mid-long term power load forecast; combined variables; partial least-squares regression

中图分类号: TM715 文献标识码: A 文章编号: 1674-3415(2009)12-0040-04

0 引言

在电力负荷预测中应用偏最小二乘回归模型的最大限制是原始数据的获取, 特别是相关因素的预测值的获取。如果获得的相关因素预测值不可靠, 由此计算得到的电力负荷预测值也是不可靠的。尤其是在我国农村电网, 获得足够的可信的相关因素数据是比较困难的, 甚至找不到可靠的数据, 导致回归预测出现困难。当某些变量的变化规律具有相似性时, 本文提出将标么值概念引入电力负荷预测, 利用变量的加权线性组合生成新变量, 可以用于对一些变量的发展规律作出估计; 可以为获得可信的相关因素数据提供一种有效的方法; 可以使得偏最小二乘回归模型能够更有效地应用于中长期电力负荷预测。

基金项目: 国家科技支撑计划子课题“农村电网高可靠性供电及供电用电节电关键技术研究”资助项目(2006BAJ04B06-001)

1 自变量变化规律存在相似性

电力、电量与国民经济总值、工业总产值、乡镇企业产值等经济数据高度相关, 与人均用电量高度相关, 与某些产业的电力、电量高度相关, 部分地区与人口、耕地面积、供电面积等因素弱相关甚至不相关。

GDP 或工农业总产值的增长趋势在一定地域范围内具有相似的增长趋势和增长速度, 说明经济增长存在广泛性和可对比性。例如河北省北部地区的工农业总产值与全国工农业总产值、华北地区工农业总产值及河北省工农业总产值都具有相似性; 其工业产值的增长趋势与全国工业总产值、华北地区工业总产值及河北省工业总产值都具有相似性。以2002年各自的产值为标么基准值, 将各地域年度产值化为标么值, 如图1所示, 可以直观地看到这种相似性。根据国家统计局数据, 将1991年至2003年的13年间国民经济总值以1991年的数值为标么基准值, 将各地域年度值化为标么值, 其增长趋势

如图 2 所示,也可以直观地看到这种相似性。

利用这种经济发展的相似性,可以将进行预测计算的县及以下电网覆盖地区与周围其他地区、所属地级市、所在省、经济发展相似地区等进行对比,特别是附近有示范县时,示范县的发展过程具有很好的对比借鉴作用。

2 标么值加权平均组合变量生成方法

某一考察因素在不同的地域将得到不同的序列值。它们的绝对数值必然存在相当大的差异,似乎没有什么可比性。当部分序列的年增长率趋势和幅度具有相似性时,它们便具有一定的共性。将标么值概念引入电力负荷预测,发掘数据序列间的共性,并引入非参数回归思想,通过寻找匹配变量,采用标么值加权平均组合方法生成一个新的序列,用来描述被考察因素的宏观共性。

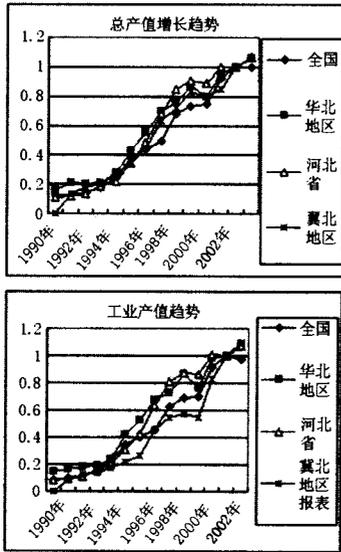


图 1 总产值及工业产值趋势对比

Fig.1 Contrast of GNP and industrial production trend

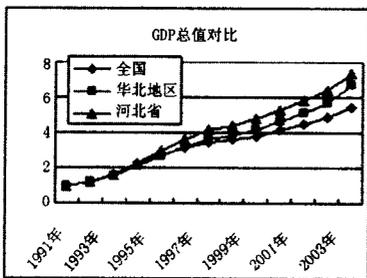


图 2 GDP 增长趋势对比

Fig.2 Contrast of the growth trend on GDP

设某考察因素 x 在 1, 2, 3, ..., m 个地域都获得了 n 年的观测值,除了预测地域,其他 $m-1$ 个

地域还获得了第 $n+1$ 年、 $n+2$ 年... $n+l$ 年的观测或预测值,即获得了 1 个长度为 n 的序列和 $m-1$ 个长度为 $n+l$ 的序列,记为 $x = [x_{ij}]$, $i=1,2,\dots,m$; $j=1,2,\dots,n,\dots,n+l$ 。

以 x_{i1} 作为标么基准值,将 x 化为标么值 X ,即 $X = [X_{ij}]$, $i=1,2,\dots,m$; $j=1,2,\dots,n+l$

$$X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \vdots & \vdots & & \vdots \\ x_{1n+l} & x_{2n+l} & \dots & x_{mn+l} \end{bmatrix} \cdot \begin{bmatrix} 1/x_{11} \\ 1/x_{21} \\ \vdots \\ 1/x_{m1} \end{bmatrix} \quad (1)$$

设其中第 k 个地域是预测地域,则令 $x_0=[X_k]$,即 $x_0=X_{kj}$, $i=1,2,\dots,m$; $j=1,2,\dots,n$; 则可利用式(2)计算关联系数:

$$\varepsilon_{ij} = \frac{\min_i \min_j |x_{0j} - X_{ij}| + \rho \max_i \max_j |x_{0j} - X_{ij}|}{|x_{0j} - X_{ij}| + \rho \max_i \max_j |x_{0j} - X_{ij}|} \quad (2)$$

其中: ρ 为分辨系数,为 0~1 之间的数,一般取 0.5。

再利用式(3)计算关联度

$$\gamma_i = \frac{1}{n \sum_{j=1}^n \varepsilon_{ij}} \quad (3)$$

如果某地域考察因素的关联度小于一定数值(一般可取 0.85),说明该地域的这个因素与预测地域的同一因素共性不大,没有组合意义,应从计算序列中剔除(也可以利用散点图观察,直接剔除不相似的地域)。

利用式(4)计算组合权重:

$$w_i = \frac{\gamma_i}{\sum_{i=1}^{m'} \gamma_i} \quad (4)$$

式中: m' 为考察因素特性与预测地域相似的地域数目。

将 m' 个考察因素加权平均,得到综合因素标么值:

$$\hat{X}_j = \sum_{i=1}^{m'} w_i \times X_{ij} \quad j=1,2,\dots,n+l, j \neq k \quad (5)$$

综合因素标么值就是新生成的标么值加权平均组合变量，它既可以作为一个独立的相关因素使用，也可以将其第 $n+1$ 年、 $n+2$ 年、 \dots 、 $n+l$ 年的数值作为该项考察因素第 $n+1$ 年、 $n+2$ 年、 \dots 、 $n+l$

年的标么预测值。

3 算例分析

以河北省北部地区用电量预测为例，1995 至 2003 年间全国、华北地区、河北省、河北省北部地区的工农业总产值数据如表 1 所示。则生成的工农业总产值加权平均标么值，亦如表 1 所示，工业产值数据也依此处理。

表 1 全国、华北地区、河北省、河北省北部地区及标么加权平均的工农业总产值
Tab.1 The per-unit total production value of industry and agriculture on Countrywide, North China, Hebei province and northern Hebei

	全国	华北地区	河北省	河北省北部地区	加权平均标么值
1995 年	584 960 668	66 674 571	28 652 603	9 192 744	0.374 936 896
1996 年	677 122 155	85 802 571	40 838 893	11 906 130	0.529 885 951
1997 年	769 283 642	108 029 797	57 022 595	16 781 331	0.596 277 341
1998 年	1 046 178 188	117 631 460	71 258 800	19 210 233	0.712 043 494
1999 年	1 129 667 243	135 732 655	76 320 699	22 051 410	0.774 423 311
2000 年	1 147 661 911	124 126 005	75 105 681	21 115 515	0.773 400 076
2001 年	1 418 506 178	148 671 843	83 905 025	22 483 528	0.904 063 482
2002 年	1 541 210 555	154 916 519	84 221 578	26 547 031	1
2003 年	1 535 659 823	331 665 715	89 951 748	28 321 828	1.084 568 636

表 2 标么加权平均工农业总产值工业总产值预测值以及用电量实际值和预测值 (亿千瓦时)
Tab.2 The per-unit total production value and production value of industry, the actual and predicted power consumption

河北省北部地区 43 个县的总和					河北省北部地区 A 县				
	总产值	工业产值	实际电量	预测电量		总产值	工业产值	实际电量	预测电量
2004 年	1.235	1.276	201.86	191.97	2004 年	1.565	1.743	4.8	4.64
2005 年	1.411	1.464	236.91	227.91	2005 年	1.721	2.004	5.54	5.16
2006 年	1.618	1.686	271.26	253.631	2006 年	1.893	2.305	6.15	5.85
2007 年	1.856	1.942	333.87	309.83	2007 年	2.083	2.651	6.91	6.52
河北省北部地区 B 县					河北省北部地区 C 县				
	总产值	工业产值	实际电量	预测电量		总产值	工业产值	实际电量	预测电量
2004 年	1.349	1.501	6.45	6.19	2004 年	1.263	1.345	5.71	5.61
2005 年	1.592	1.846	7.62	7.12	2005 年	1.49	1.654	7.28	6.73
2006 年	1.878	2.27	9.12	8.95	2006 年	1.758	2.034	8.99	8.51
2007 年	2.216	2.792	10.96	10.27	2007 年	2.075	2.502	11.18	10.68

计算河北省北部地区各县的 1995 至 2003 年工农业总产值、工业产值的加权平均标么值，并根据河北省及河北省北部地区对县及以下经济发展的预计，推出 2004 至 2007 年的加权平均标么值数值，如表 2 所示。利用 1995 至 2007 年的工农业总产值、工业产值的加权平均标么值及其他相关因素序列，

采用偏最小二乘法进行用电量预测，预测结果如表 2 所示，预测值与实际值平均误差小于 5%。利用修正灰色预测方法再对残差进行预测计算，对 2004 至 2007 年预测结果进行残差修正，可获得更接近实际值、残差分布更合理的预测效果。2008 至 2012 年预测值也被河北省北部地区电力相关部门认可。

算例验证了本文组合变量生成方法用于电力负荷预测的有效性。

4 结论

偏最小二乘回归方法是比较可靠的电力负荷预测算法,但在我国农村电网中,能够获取数据的自变量数量少、数据缺失现象比较常见,有时一些相关因素的预测非常困难,致使偏最小二乘回归方法难以获得广泛应用。

(1) 本文将标么值概念引入电力负荷预测,使变量间及变量在不同地区间的相似性得以显现,为利用非参数回归思想提供了数据依据;

(2) 本文采用非参数回归思想,利用同一变量在多个区域的数据序列群生成标么值加权平均组合变量,利用该组合变量弥补数据缺失,并利用该组合变量对一些相关因素的发展规律、增长速率等作出估计,为局部地区的经济发展趋势判断提供了一种预估方法;

(3) 本文提出的标么值加权平均组合变量生成方法能够有效地改善农村电网中相关因素数量少、数据缺失、预测困难等难题,为回归类方法在农村电网电力负荷预测中的应用提供了数据获取的有效手段;使偏最小二乘回归方法在我国农村电网获得广泛应用成为可能。

参考文献

- [1] 蒋惠凤,何有世,杨伟雄. 基于偏最小二乘回归的中长期电力负荷预测[J]. 电力系统及其自动化学报, 2007, 19(5):110-113.

JIANG Hui-feng, HE You-shi, YANG Wei-xiong. Medium and Long Term Load Forecasting Based on Partial Least-Squares Regression[J]. Proceedings of the Chinese Society of Universities for Electric Power System and Automation, 2007, 19(5):110-113.

- [2] 许璞. 基于 Web 的电力中长期负荷预测系统的研究与开发[硕士学位论文][D]. 北京: 中国农业大学, 2006. Xü Pu. Study and Develop on Medium -long Term Load Forecasting System Based on Web, Thesis[S]. Beijing: China Agricultural University, 2006.
- [3] 虞瑄, 程浩忠, 王旭, 等. 基于相关分析的中长期电力负荷综合预测方法[J]. 继电器, 2005, 33(5):49-52.
- [4] 康重庆, 夏清, 刘梅. 电力系统负荷预测 [M]. 北京: 中国电力出版社, 2007. KANG Chong-qing, XIA Qing, LIU Mei. Power System Load Forecasting[M]. Beijing: China Electric Power Press, 2007.
- [5] 国家发展和改革委员会, 国家电网公司. 电力需求侧管理工作指南[M]. 北京: 中国电力出版社 2007. National Development and Reform Commission, State Grid Cooperation of China, Power Demand Side Management Manual[M]. Beijing: China Electric Power Press, 2007.

收稿日期: 2008-07-25; 修回日期: 2008-10-10

作者简介:

张筱慧(1962-), 女, 副教授, 主要从事电力系统及其自动化、电力系统负荷预测等方面的科研和教学工作。E-mail: zxx8366@sina.com

世界首个±800 kV 直流工程进入非常时刻

日前, 南方电网“十一五”西电东送重点工程, 世界首个±800 kV 直流输电工程云南至广东特高压直流示范工程进入调试启动阶段。4月21日, 云广直流工程启动验收委员会第一次会议召开, 对云广直流工程建设、调试、投产作出安排部署。

云广直流工程启动验收委员会主任委员、公司副总经理王久玲强调, 要多与主设备制造商沟通协调, 督促其加大人力物力投入, 解决好有关技术问题, 严把质量关。紧密依靠各地党委、政府部门, 加大力度解决施工受阻问题。超前做好运行方式的编制等工作, 确保调试投产过程中南方电网系统的稳定。编制好调试规程规范, 为形成特高压直流调试的国际标准创造条件。不断总结经验, 切实总结好规划、设计、制造、调试及投产过程中各方面的问题、经验和智慧, 形成公司的知识和财富。

云广特高压直流工程的建设, 不仅可以满足小湾、金安桥等大型水电站送出和“十一五”云电送粤的需要, 也为今后西南地区水电大规模、远距离送出积累实践经验, 促进能源资源在更大范围内实现优化配置。目前, 该工程换流站电气安装总体完成66%, 直流线路组塔完成98.7%, 架线施工完成94.1%。按照调试启动计划, 4月28日广东穗东换流站交流场启动带电; 5月中旬, 云南楚雄换流站交流场启动带电; 5月下旬, 两端换流站开始站系统调试; 6月下旬开始极二带负荷系统调试; 7月30日前, 极二±400 kV设备投入试运行; 12月30日前, 极二±800 kV设备投入试运行。