

文章编号:1671-1637(2013)02-0092-07

基于 PCA 和 HMM 的汽车保有量预测方法

孙 璐^{1,2}, 郁 烨¹, 顾文钧²

(1. 东南大学 交通学院, 江苏 南京 210096; 2. 美国天主教大学 土木工程系, 华盛顿特区 20064)

摘 要:分析了常用的汽车保有量预测方法,提出了一种新的基于主成分分析和隐马尔可夫模型的汽车保有量预测方法。选取国民总收入、人均 GDP、人口总数量、城市化率、固定资产投资总额、进出口总额、城镇居民人均可支配收入、钢材产量、公路货运量、公路客运量、社会消费品零售总额 11 个指标作为汽车保有量的主要影响因素,运用主成分分析提取了主要影响因素的主成分。以提取的主成分与汽车保有量分别作为自变量、因变量,建立了回归分析模型。以汽车保有量回归预测值的年增长率为隐状态,以回归预测值与实际值的相对误差为可见信号,建立了隐马尔可夫模型,并对的汽车保有量回归预测值进行修正。分析结果表明:基于 1994~2008 年的中国汽车保有量及其主要影响因素的历史数据,应用提出的方法得到 2009、2010 年的汽车保有量修正值分别为 $6.220\ 96 \times 10^7$ 、 $7.825\ 12 \times 10^7$ veh;与 2009、2010 年实际汽车保有量比较,相对误差分别为 -0.95% 、 0.30% 。可见,基于主成分分析和隐马尔可夫模型的汽车保有量预测方法具有良好的预测精度,能够适用于短期预测。

关键词:汽车保有量预测;隐马尔可夫模型;主成分分析;回归分析;灰色预测

中图分类号:U491.14

文献标志码:A

Car ownership prediction method based on principal component analysis and hidden Markov model

SUN Lu^{1,2}, YU Ye¹, GU Wen-jun²

(1. School of Transportation, Southeast University, Nanjing 210096, Jiangsu, China; 2. Department of Civil Engineering, The Catholic University of America, Washington DC 20064, USA)

Abstract: The usual prediction methods of car ownership were analyzed, a new car ownership prediction method based on principal component analysis (PCA) and hidden Markov model (HMM) was put out. The 11 indexes including gross national income, per capita GDP, total population number, urbanization rate, total fixed asset investment, gross import and export, urban resident disposable income, steel output, highway passenger transport volume, highway freight transport volume, total retail sales of consumer goods were taken as the main influence factors of car ownership, and PCA was used to extract the principal components of main influence factors. The principal component and car ownership were taken as independent variable and dependent variable respectively, and the regression analysis model was set up. The annual growth rates of regression prediction values for car ownership were taken as hidden state, the relative errors between regression prediction values and actual values were taken as visible signal, the hidden Markov model was built, and the regression prediction values of car ownership were modified. Analysis result shows that based on car ownerships and the historical data of main

收稿日期:2012-09-17

基金项目:国家自然科学基金项目(51010044, U1134206)

作者简介:孙 璐(1972-),男,上海人,东南大学教授,长江学者,工学博士,从事道路交通安全研究。

influence factors in 1994-2008, the numbers of modified car ownership in 2009 and 2010 are $6.220\ 96 \times 10^7$ and $7.825\ 12 \times 10^7$ by using the proposed method. Compared with the actual values of car ownership in 2009 and 2010, relative errors are -0.95% and 0.30% respectively. So car ownership prediction method based on PCA and HMM has a high prediction accuracy and is suitable for short-term prediction. 9 tabs, 1 fig, 21 refs.

Key words: car ownership prediction; hidden Markov model; principal component analysis; regression analysis; grey prediction

Author resume: SUN Lu (1972-), male, professor, PhD, +86-25-83792619, sunl@cua.edu.

0 引言

随着中国经济水平的发展,国民收入水平的提高以及汽车价格的不断下降,汽车正逐步由奢侈消费品转变为一般消费品。至2010年末,中国汽车保有量达到了7 801.83万 veh,继续保持旺盛的增长态势。汽车工业产业是现代工业发展的一个主要推动力,汽车消费可以带动钢铁、石油相关行业的发展,带动国民经济整体水平的提高。与此同时,汽车保有量的快速攀升也带来了交通拥挤严重、污染增加、事故频繁、能耗严重、停车困难等问题。科学而准确地预测汽车保有量,对于道路交通规划、汽车工业发展规划以及汽车产业相关政策的制定都具有十分重要的意义。

所谓预测就是利用过去和当前的观测值对未来值进行估计,由此产生了多种预测方法^[1-3]。Dargay等将人均汽车保有量作为人均国民收入的函数,基于多个国家汽车保有量历史数据,建立了汽车保有量预测模型^[4];仲伟周等运用因子分析法和线性回归模型分析了影响中国汽车保有量的各种因素^[5];杨艳妮等基于支持向量机理论建立了汽车保有量滚动预测模型,并对北京市私家车保有量进行了预测^[6];曹晓飞等根据历年汽车保有量增长速度与GDP增长速度的比例关系,运用弹性系数法预测了未来几年的北京市汽车保有量^[7];何明等在主成分分析法的基础上,利用BP神经网络方法对汽车保有量进行了预测^[8];周骞等将神经网络应用于汽车保有量预测,建立了基于双BP神经网络的多相关因素汽车保有量预测模型^[9];朱海清等建立了路网服务水平影响下的路网容量计算模型,并对路网容量下各类汽车保有量进行了预测^[10];朱开永等利用灰色系统理论建立了汽车保有量预测模型^[11];朱灿等以ARIMA模型和Logistic曲线拟合模型作为前期预测模型,运用混合RBF神经网络构建了汽车保有量的非线性组合预测模型^[12];王琦等采用灰色

系统、多元回归、指数平滑、神经网络这4种预测方法对汽车保有量进行了预测,并运用熵值法确定各预测模型的加权系数,建立了组合预测模型^[13]。

上述研究对汽车保有量预测模型的建立和方法的应用作了许多开拓性的尝试,目前常用的汽车保有量预测方法包括时间序列法、弹性系数法、回归分析法、路网容量法、灰色模型法、人工神经网络法、支持向量机法以及组合预测法。时间序列法依据汽车保有量的历史数据,以时间为自变量建立预测模型,该方法简单易行,但不能反映事物的内在关系,也不能分析2个以上因素的相关关系,只适用于短期预测。弹性系数法着重考虑了一个主要因素对汽车保有量造成的影响,然而仅仅考虑一个因素难以全面地把握汽车保有量的变化情况,通常只用作粗略预测。回归分析法基于汽车保有量及其影响因素的历史数据进行回归分析,可找出汽车保有量与其影响因素之间的关系,进而依据各影响因素的未来发展水平对汽车保有量作出预测,然而该方法必须获得大量连续统计数据,在时间跨度较大时统计指标往往会存在统计口径不一等问题。路网容量法基于道路网络时空资源的限制推算出道路网络所能容纳的最大汽车数量,再结合各种车型的高峰小时出车率,进一步得出各种车型的汽车保有量,然而高峰小时出车率的确定往往较为困难。灰色模型法仅需要收集汽车保有量自身的历史数据,具有建模所需数据少与短期预测精度较高的优点,但是灰色模型要求汽车保有量的累加生成序列具有指数规律,而汽车保有量的历史数据未必具有这一特性。人工神经网络法具有高度的并行信息处理能力和良好的非线性映射能力,近年来被广泛应用于汽车保有量预测,然而该方法存在着学习过程容易陷入局部极值等问题。与人工神经网络法相比,支持向量机结构简单,泛化能力好,然而目前基于支持向量机理论开展汽车保有量预测的相关研究相对较少,这一方法尚待于进一步研究。组合预测法以适当的加权形式对不

同的单项预测模型进行组合,从而改善模型的拟合能力,提高模型的预测精度,但是组合预测法的预测性能在一定程度上仍依赖于各单项预测模型的预测性能,且如何合理确定各单项预测模型的加权系数也还有待于深入研究。

从上述分析可知,现有众多汽车保有量预测方法往往具有一定的局限性,因此,探索新的汽车保有量预测方法是研究汽车保有量预测问题的一个可行途径。隐马尔科夫模型是一种应用广泛的统计模型,目前已经在语音识别、图像处理、信号处理等领域获得了广泛应用。不同于现有的汽车保有量预测模型,本文提出一种汽车保有量预测方法,力图将隐马尔科夫模型引入到汽车保有量预测领域,先运用主成分分析提取汽车保有量众多主要影响因素的主成分,建立汽车保有量的回归分析模型,再运用隐马尔科夫模型分析汽车保有量回归预测值的年增长率与回归预测值的相对误差两者之间的关系,修正未来年的汽车保有量回归预测值,以提高预测精度。

表 1 主要影响因素

Tab. 1 Main influence factors

年份	汽车保有量/ 10 ⁴ veh	国民总收入/ 10 ⁸ 元	人均GDP/元	人口总量/ 10 ⁴ 人	城市化率/%	固定资产投资总额/ 10 ⁸ 元	进出口总额/ 10 ⁸ 元	城镇居民人均可支配收入/ 元	钢材产量/ 10 ⁴ t	公路货运量/ 10 ⁴ t	公路客运量/ 10 ⁴ 人	社会消费品零售总额/ 10 ⁸ 元
1994	941.95	48 108.5	4 044	119 850	28.51	17 042.1	20 381.9	3 496.2	8 428.00	894 914	953 940	18 622.9
1995	1 040.00	59 810.5	5 046	121 121	29.04	20 019.3	23 499.9	4 283.0	8 979.80	940 387	1 040 810	23 613.8
1996	1 100.08	70 142.5	5 846	122 389	30.48	22 913.5	24 133.8	4 838.9	9 338.02	983 860	1 122 110	28 360.2
1997	1 219.09	78 060.8	6 420	123 626	31.91	24 941.1	26 967.2	5 160.3	9 978.93	976 536	1 204 583	31 252.9
1998	1 319.30	83 024.3	6 796	124 761	33.35	28 406.2	26 849.7	5 425.1	10 737.80	976 004	1 257 332	33 378.1
1999	1 452.94	88 479.2	7 159	125 786	34.78	29 854.7	29 896.2	5 854.0	12 109.78	990 444	1 269 004	35 647.9
2000	1 608.91	98 000.5	7 858	126 743	36.22	32 917.7	39 273.2	6 280.0	13 146.00	1 038 813	1 347 392	39 105.7
2001	1 802.04	108 068.2	8 622	127 627	37.66	37 213.5	42 183.6	6 859.6	16 067.61	1 056 312	1 402 798	43 055.4
2002	2 053.17	119 095.7	9 398	128 453	39.09	43 499.9	51 378.2	7 702.8	19 251.59	1 116 324	1 475 257	48 135.9
2003	2 382.93	135 174.0	10 542	129 227	40.53	55 566.6	70 483.5	8 472.2	24 108.01	1 159 957	1 464 335	52 516.3
2004	2 693.71	159 586.7	12 336	129 988	41.76	70 477.4	95 539.1	9 421.6	31 975.72	1 244 990	1 624 526	59 501.0
2005	3 159.66	185 808.6	14 185	130 756	42.99	88 773.6	116 921.8	10 493.0	37 771.14	1 341 778	1 697 381	67 176.6
2006	3 697.35	215 904.4	16 500	131 448	44.34	109 998.2	140 974.0	11 759.5	46 893.36	1 466 347	1 860 487	76 410.0
2007	4 358.36	266 422.0	20 169	132 129	45.89	137 323.9	166 863.7	13 785.8	56 560.87	1 639 432	2 050 680	89 210.0
2008	5 099.61	316 030.3	23 708	132 802	46.99	172 828.4	179 921.5	15 780.8	60 460.29	1 916 759	2 682 114	108 487.7

另外,国家和地方的汽车产业政策(包括产业政策、产业政策、产业政策、产业政策等)对汽车保有量也有着明显的影响。汽车产业政策是国家根据国民经济发展的内在要求,调整产业结构和产业组织形式的政策措施。尽管汽车产业政策的变化明显影响到汽车保有量的增长情况,本文在预测时并没

1 汽车保有量影响因素分析

汽车保有量预测具有一定程度的不确定性,影响汽车保有量的因素较多,地区经济、人口、居民可支配收入、政策走向以及其他一些与汽车相关产业的变化都将影响汽车保有量的增长速度^[14]。选取不同的影响因素将获得不同的预测结果,因此,影响因素的选择非常重要。

综合考虑影响汽车保有量的诸多因素与数据的可得性,本文选取国民总收入、人均 GDP、人口总数量、城市化率、固定资产投资总额、进出口总额、城镇居民人均可支配收入、钢材产量、公路货运量、公路客运量、社会消费品零售总额等 11 个因素作为中国汽车保有量的主要影响因素进行分析,见表 1,数据来自于历年的中国统计年鉴。使用 SPSS 统计分析软件分别验证上述 11 个因素与汽车保有量的相关性,分析结果见表 2,最小的相关系数为 0.899,由此可见上述 11 个因素均与汽车保有量显著相关。

有将汽车产业政策考虑在内,主要原因有 2 点,一是如何量化汽车产业政策变化对汽车保有量增长的影响尚有待于深入研究,二是对未来的汽车产业政策进行预测也较为困难。本文方法只适用于在汽车产业政策相对稳定的情况下预测汽车保有量,而如何将汽车产业政策纳入到汽车保有量预测过程还有待

表 2 相关系数

Tab. 2 Correlation coefficients

变量	指标	相关系数
x_1	国民总收入/ 10^8 元	0.998
x_2	人均 GDP/元	0.988
x_3	人口总数量/ 10^4 人	0.899
x_4	城市化率/%	0.939
x_5	固定资产投资总额/ 10^8 元	0.993
x_6	进出口总额/ 10^8 元	0.992
x_7	城镇居民人均可支配收入/元	0.997
x_8	钢材产量/ 10^4 t	0.994
x_9	公路货运量/ 10^4 t	0.990
x_{10}	公路客运量/ 10^4 t	0.974
x_{11}	社会消费品零售总额/ 10^8 元	0.996

于进一步深入研究。

2 主成分分析

2.1 主成分分析

主成分分析 (Principal Component Analysis, PCA) 是一种数据压缩和特征提取的多变量统计分析技术,能够有效去除数据间的相关性,目的就是在保证数据损失尽可能小的前提下,经过线性变换舍弃小部分信息,以少数新的综合变量取代原始变量,要求主成分能够充分反映原始变量的信息,同时又互不相关,从而进行样本评价^[15-16]。

运用相关分析法,11 个因素最小的两两相关系数达到了 0.848,表明所选取的 11 个因素在信息表达上有很大部分的重叠,有必要进行主成分分析。进行主成分分析时,为了使 11 个因素之间具有可比性,应用 SPSS 统计分析软件对 11 个因素分别进行了标准化处理。国民总收入、人均 GDP、人口总数量、城市化率、固定资产投资总额、进出口总额、城镇居民人均可支配收入、钢材产量、公路货运量、公路客运量、社会消费品零售总额等 11 个因素的原始数据分别记为 x_1, x_2, \dots, x_{11} , 标准化之后的数据分别记为 X_1, X_2, \dots, X_{11} 。

使用 SPSS 统计分析软件分析 11 个因素的主成分,计算特征值、方差贡献率、累计方差贡献率,见表 3。主成分分析一般提取对应特征值大于 1 的主成分,同时保证累积贡献率不小于 85%。第 1 主成分的累积方差贡献率达到了 96.321%,能够充分反映原始数据的信息,且只有第 1 主成分对应特征值大于 1,故提取第 1 主成分来代表原来的 11 个影响因素。根据 11 个因素标准化数据可得到

表 3 方差贡献率

Tab. 3 Total variance explained

主成分	特征值	方差贡献率/%	累计方差贡献率/%
1	10.595	96.321	96.321
2	0.317	2.881	99.202
3	0.080	0.726	99.928
4	0.003	0.031	99.960
5	0.003	0.017	99.927
6	0.001	0.013	99.899
7	0.001	0.005	99.995
8	0.000	0.003	99.998
9	0.000	0.002	100.000
10	0.000	0.000	100.000
11	0.000	0.000	100.000

$$Y=0.306\ 3X_1+0.306\ 6X_2+0.282\ 6X_3+0.282\ 5X_4+0.303\ 5X_5+0.302\ 9X_6+0.307\ 2X_7+0.303\ 5X_8+0.303\ 2X_9+0.300\ 8X_{10}+0.306\ 9X_{11}$$

式中:Y 为所提取出的第 1 主成分,本文简称为主成分。

2.2 回归分析

汽车保有量与主成分的关系见图 1,从中可以看出汽车保有量与主成分近似具有线性关系。以提取出的主成分为自变量,以汽车保有量作为因变量进行回归分析,回归方程为

$$Z=2\ 261.945+396.456Y\quad R^2=0.986$$

式中:Z 为汽车保有量; R^2 为判定系数。

将主成分代入回归方程即可计算出相应年份的汽车保有量回归预测值。通过与本年度的汽车保有量实际值进行比较,可得出汽车保有量回归预测值与当年实际值的相对误差(以下简称为相对误差)。

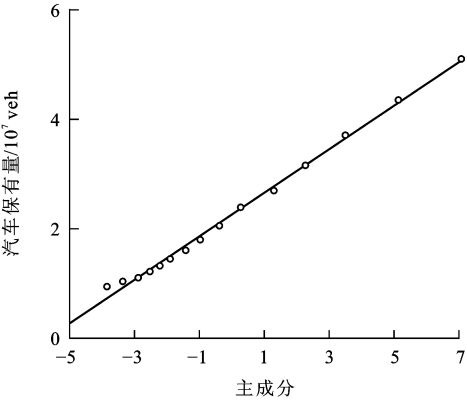


图 1 汽车保有量与主成分的关系

Fig. 1 Relationship between car ownership and principal component

由本年度的回归预测值与上一年度的回归预测值可得到本年度汽车保有量回归预测值的年增长率(以下简称回归预测值年增长率)。表 4 为 1994~2008 年汽车保有量回归预测值、相对误差与回归预测值的年增长率,除 1994、1995 年的相对误差较大以外,其他年份的相对误差较小。

表 4 汽车保有量的回归预测值

Tab. 4 Regression prediction values of car ownership

年份	主成分	回归预测值/ 10 ⁴ veh	相对误差/ %	回归预测值的 年增长率/%
1994	-3.851 8	734.88	-21.98	
1995	-3.362 9	928.70	-10.70	26.37
1996	-2.892 4	1 115.24	1.38	20.09
1997	-2.528 6	1 259.47	3.31	12.93
1998	-2.225 2	1 379.75	4.58	9.55
1999	-1.910 7	1 504.44	3.54	9.04
2000	-1.433 5	1 693.63	5.27	12.58
2001	-0.980 0	1 873.42	3.96	10.62
2002	-0.383 9	2 109.75	2.76	12.61
2003	0.278 3	2 372.28	-0.45	12.44
2004	1.303 5	2 778.73	3.16	17.13
2005	2.269 5	3 161.70	0.06	13.78
2006	3.506 1	3 651.96	-1.23	15.51
2007	5.127 8	4 294.89	-1.46	17.61
2008	7.083 6	5 070.28	-0.58	18.05

3 隐马尔可夫模型的建立与汽车保有量修正

3.1 建立隐马尔可夫模型

隐马尔可夫模型(Hidden Markov Model, HMM)是由马尔可夫链扩展而来的一种随机过程^[17-21],描述了一个双重随机过程。隐马尔可夫模型常用 5 元参数组 $\lambda=(S,O,A,B,\pi)$ 来描述,其中状态集合 $S=\{S_1,S_2,\dots,S_N\}$, N 为状态总数量;可见信号集合 $O=\{O_1,O_2,\dots,O_M\}$, M 为可见信号总数量; A 为状态转移概率; B 为信号发出概率; π 为初始状态概率。

基于上文的主成分分析和回归分析,可得出预测年份的汽车保有量回归预测值与相应的年增长率。如果能够基于回归预测年增长率与相对误差合理作出估计,就能够对回归预测值进行修正,进而提高其预测精度。鉴于此,本文运用隐马尔可夫模型分析回归预测值年增长率与相对误差之间的关系。

将表 4 中 1997~2008 年的回归预测值年增长率与相对误差分别作为隐状态与可见信号,建立离散隐马尔可夫模型。由于 1994、1995 年的相对误差较大,因此,舍去 1994~1996 年的相关数据。将回归预测值年增长率与相对误差分别按照百分位数分类,本文以 3 类为例进行划分,分类标准见表 5。按照从小到大的顺序,回归预测值年增长率被划分为 S_1 、 S_2 和 S_3 ,相对误差被划分为 O_1 、 O_2 和 O_3 。根据回归预测值年增长率、相对误差的具体分类,得到状态转移序列和可见信号序列,见表 6。隐马尔可夫模型的状态转移概率见表 7,可见信号的分布概率见表 8。

表 5 分类标准

Tab. 5 Classification standard

百分位数	回归预测值年增长率(隐状态)	相对误差(可见信号)
33.33	12.49	-0.28
66.67	14.93	3.46

表 6 状态转移序列和可见信号序列

Tab. 6 State transfer sequence and visible signal sequence

年份	隐状态	可见信号	年份	隐状态	可见信号
1997	2	2	2003	1	1
1998	1	3	2004	3	2
1999	1	3	2005	2	2
2000	2	3	2006	3	1
2001	1	3	2007	3	1
2002	2	2	2008	3	1

表 7 状态转移概率

Tab. 7 State transfer probability

状态	S_1	S_2	S_3
S_1	0.25	0.50	0.25
S_2	0.75	0.00	0.25
S_3	0.00	0.33	0.67

表 8 可见信号的分布概率

Tab. 8 Distribution probability of visible signals

状态	O_1	O_2	O_3
S_1	0.25	0.00	0.75
S_2	0.00	0.75	0.25
S_3	0.75	0.25	0.00

3.2 汽车保有量修正

为了准确预测未来几年的汽车保有量,本文采用灰色模型法预测未来几年的主成分值,进而得出未来几年的汽车保有量回归预测值与年增长率,基于隐马尔可夫模型估计出相对误差,进而对回归预测值进行修正。

采用灰色模型法进行主成分预测,构造 GM(1,1) 模型,则有

$$y^{(1)}(k) = \sum_{i=1}^k y^{(0)}(i)$$

式中: $y^{(0)}(i)$ 为原始时间数列的第 i 个观测值; $y^{(1)}(k)$ 为累加时间数列的第 k 个值。

应用 2005~2008 年的主成分数据建立预测方程,对 2009、2010 年的主成分进行预测,预测结果见表 9。依据主成分的灰色预测值,计算得出 2009、2010 年汽车保有量的回归预测值分别为 $6.163\ 11 \times 10^7$ 、 $7.752\ 35 \times 10^7$ veh,2009、2010 年回归预测值的年增长率分别为 21.55%、25.79%。可见,2009、2010 年的回归预测值的年增长率均处于第 3 类隐状态 S_3 ,该状态下最可能发出的可见信号为第 1 类可见信号 O_1 。以第 1 类可见信号的平均值 -0.93% 作为相对误差的估计值,进而对回归预测值进行修正。由此得出,2009、2010 年汽车保有量修正预测值分别为 $6.220\ 96 \times 10^7$ 、 $7.825\ 12 \times 10^7$ veh。2009、2010 年汽车保有量的实际值分别为 $6.280\ 61 \times 10^7$ 、 $7.801\ 83 \times 10^7$ veh,计算得出 2009、2010 年回归预测值的预测误差分别为 -1.87% 、 -0.63% ,修正预测值的预测误差分别为 -0.95% 、 0.30% 。从预测误差来看,修正预测值的预测精度优于回归预测值。基于主成分分析和隐马尔科夫模型的汽车保有量预测具有较高的预测精度,能够适用于短期预测。

表 9 预测结果
Tab.9 Prediction result

年份	2009	2010
主成分灰色预测值	9.840 1	13.848 7
汽车保有量回归预测值(10^7 veh)	6.163 11	7.752 35
回归预测值年增长率/%	21.55	25.79
汽车保有量修正预测值(10^7 veh)	6.220 96	7.825 12
汽车保有量实际值(10^7 veh)	6.280 61	7.801 83
回归预测值的预测误差/%	-1.87	-0.63
修正预测值的预测误差/%	-0.95	0.30

需要指出的是,上述隐马尔科夫模型是基于 1997~2008 年的回归预测值的年增长率与相对误差的波动情况而建立的。事实上,2009、2010 年汽车保有量回归在预测值的年增长率仍属于第 3 类隐状态 S_3 ,但实际已经大大超出了历年来民用汽车保有量年增长率的波动范围,产生这一现象的原因是隐马尔科夫模型是一种有限状态机理论,从统计年鉴获取的汽车保有量数据是按年进行统计的,数据

量相对较少,导致了隐马尔科夫模型训练不充分,导致该方法只适用短期(1 年或 2 年)预测。如果能够获得更详细的历史数据,例如按月统计的民用汽车保有量数据,该现象将得到明显改善。

4 结 语

鉴于目前将隐马尔可夫模型应用于汽车保有量预测的相关研究很少,本文在主成分分析和隐马尔可夫模型的基础上,建立了汽车保有量预测方法。本文方法在一定程度上反映了实际情况下汽车保有量与汽车保有量驱动因素之间的关系。在隐马尔科夫模型中,汽车保有量的实际增长率是汽车保有量增长驱动因素作用下的一个可能输出,但不是唯一的输出,而在回归预测模型中,汽车保有量与其增长驱动因素之间的关系则是一一对应的确定关系,可见,隐马尔可夫预测模型比回归预测模型更符合实际情况。隐马尔可夫模型是一种有限状态机理论,本文在进行实例验证时采用了按年统计的汽车保有量数据,其数据量相对较少,导致模型训练不够充分,其应用范围也存在着一定的局限性。此外,本文经验性地选取了隐马尔可夫模型的状态数量和可见信号数量(均取为 3),而如何合理确定隐马尔可夫模型的参数以更好地描述实际问题还有待进一步深入研究。

参 考 文 献 :

References :

- [1] 张益,陈淑燕,王炜. 短时交通量时间序列智能复合预测方法概述[J]. 公路交通科技,2006,23(8):139-142.
ZHANG Yi, CHEN Shu-yan, WANG Wei. Survey of traffic volume time series intelligent compound forecasting methods[J]. Journal of Highway and Transportation Research and Development, 2006, 23(8): 139-142. (in Chinese)
- [2] POTOGLU D, SUSILO Y. Comparison of car ownership models[J]. Transportation Research Record, 2008(2076): 97-105.
- [3] 陈远通. 汽车保有量预测技术方法及其应用[D]. 广州:华南理工大学,2010.
CHEN Yuan-tong. The methodology and implementation of car ownerships forecast technique[D]. Guangzhou: South China University of Technology, 2010. (in Chinese)
- [4] DARGAY J, GATELY D. Income's effect on car and vehicle ownership, worldwide: 1960-2015[J]. Transportation Research Part A: Policy and Practice, 1999, 33(2): 101-138.
- [5] 仲伟周,王军. 基于因子分析法的中国民用汽车保有量影响因素分析及其政策含义[J]. 统计与信息论坛,2008,23(11): 16-20.

- ZHONG Wei-zhou, WANG Jun. The analysis of the effective factor of China's civilian vehicle holdings and its policy implications based on the factor analysis[J]. Statistics and Information Forum, 2008, 23(11): 16-20. (in Chinese)
- [6] 杨艳妮, 陆化普. 基于支持向量机的城市私人汽车保有量滚动预测[J]. 公路工程, 2010, 35(5): 40-43.
- YANG Yan-ni, LU Hua-pu. An improved prediction model for the stock of urban private cars based on support vector machine[J]. Highway Engineering, 2010, 35(5): 40-43. (in Chinese)
- [7] 曹晓飞, 邵春福. 运用弹性系数法预测北京机动车保有量[J]. 道路交通与安全, 2008, 8(4): 31-34.
- CAO Xiao-fei, SHAO Chun-fu. Forecast of vehicle amount by elastic coefficient method in Beijing[J]. Road Traffic and Safety, 2008, 8(4): 31-34. (in Chinese)
- [8] 何 明, 过秀成. BP 神经网络主成分分析法在汽车保有量预测中的应用[J]. 交通与计算机, 2007, 25(4): 96-99.
- HE Ming, GUO Xiu-cheng. Application of BP neural network principal component analysis to forecasting the automobile population[J]. Computer and Communications, 2007, 25(4): 96-99. (in Chinese)
- [9] 周 骞, 杨东援. 基于多相关因素的汽车保有量预测神经网络方法[J]. 公路交通科技, 2001, 18(6): 71-74.
- ZHOU Qian, YANG Dong-yuan. Multi-relative factor forecast of vehicle population by neural network[J]. Journal of Highway and Transportation Research and Development, 2001, 18(6): 71-74. (in Chinese)
- [10] 朱海清, 陈学武. 路网服务水平影响下的机动车规模预测[J]. 交通运输工程与信息学报, 2005, 3(4): 80-85.
- ZHU Hai-qing, CHEN Xue-wu. Amount prediction of motor-driven vehicles under the influences of road network service level[J]. Journal of Transportation Engineering and Information, 2005, 3(4): 80-85. (in Chinese)
- [11] 朱开永, 周圣武, 娄可元, 等. 基于私家车保有量预测与调控的灰色模型研究[J]. 中国矿业大学学报, 2008, 37(6): 868-872.
- ZHU Kai-yong, ZHOU Sheng-wu, LOU Ke-yuan, et al. Research on grey model about forecast and control for total number of private cars[J]. Journal of China University of Mining and Technology, 2008, 37(6): 868-872. (in Chinese)
- [12] 朱 灿, 周和平, 钟璧樵. 基于混合 RBF 网络模型的汽车保有量组合预测[J]. 长沙理工大学学报: 自然科学版, 2011, 8(2): 13-16.
- ZHU Can, ZHOU He-ping, ZHONG Bi-qiang. Forecasting on vehicle population based on combined model with mixed neural network[J]. Journal of Changsha University of Science and Technology: Natural Science, 2011, 8(2): 13-16. (in Chinese)
- [13] 王 琦, 王花兰. 基于熵值法的城市汽车保有量组合预测[J]. 交通科技与经济, 2009, 11(6): 53-55.
- WANG Qi, WANG Hua-lan. Combination forecasting of vehicle population based on the entropy method[J]. Technology and Economy in Areas of Communications, 2009, 11(6): 53-55. (in Chinese)
- [14] 亓芳芳. 基于面板数据模型的中国民用汽车消费需求预测研究[D]. 合肥: 中国科学技术大学, 2010.
- QI Fang-fang. Forecasting research on the consumption demand of civilian vehicle in China based on panel data model[D]. Hefei: University of Science and Technology of China, 2010. (in Chinese)
- [15] 吴 明, 王莹莹, 冯 琪. 基于主成分分析的综合运输通道布局模型研究[J]. 公路交通科技, 2011, 28(1): 154-158.
- WU Ming, WANG Ying-ying, FENG Qi. Research on comprehensive transportation channel layout model based on main component analysis[J]. Journal of Highway and Transportation Research and Development, 2011, 28(1): 154-158. (in Chinese)
- [16] 张文峰, 史忠科. 主成分分析法在城市道路交通安全评价中的应用[J]. 计算机工程与应用, 2007, 43(36): 246-248.
- ZHANG Wen-feng, SHI Zhong-ke. Application of principal component analysis in evaluation of city road traffic safety[J]. Computer Engineering and Applications, 2007, 43(36): 246-248. (in Chinese)
- [17] RABINER L R. A tutorial on hidden Markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257-286.
- [18] RABINER L R, JUANG B H. An introduction to hidden Markov models[J]. IEEE ASSP Magazine, 1986, 3(1): 4-16.
- [19] 张庆生, 齐 勇, 侯 迪, 等. 基于隐马尔科夫模型的上下文感知活动计算[J]. 西安交通大学学报, 2006, 40(4): 398-401.
- ZHANG Qing-sheng, QI Yong, HOU Di, et al. Context-aware activity computing based on hidden Markov model[J]. Journal of Xi'an Jiaotong University, 2006, 40(4): 398-401. (in Chinese)
- [20] 吕颖钊, 贺拴海. 在役桥梁承载力模糊可靠性的马尔科夫预测[J]. 长安大学学报: 自然科学版, 2005, 25(4): 39-43.
- LU Ying-zhao, HE Shuan-hai. Prediction of Markov fuzzy reliability for existing bridge[J]. Journal of Chang'an University: Natural Science Edition, 2005, 25(4): 39-43. (in Chinese)
- [21] 张 珍. 基于隐马尔科夫模型的波动率预测[D]. 广州: 华南理工大学, 2012.
- ZHANG Zhen. Volatility forecast via hidden Markov models[D]. Guangzhou: South China University of Technology, 2012. (in Chinese)