

Analysis and Prediction of Influencing Factors of Fiscal Revenue in Wenzhou Based on Data Mining Technology

Liangliang Zhuang, Huilin Huang*, Tong Wu

College of Mathematics and Information Science, Wenzhou University, Wenzhou Zhejiang
Email: *huilin_huang@wzu.edu.cn

Received: Nov. 23rd, 2017; accepted: Dec. 7th, 2017; published: Dec. 14th, 2017

Abstract

This paper established a Lasso model, which selects four key factors affecting the financial revenue of Wenzhou, such as the total wages of staff and workers, per capita consumption expenditure of urban residents, GDP and tax. On this basis, we also get the forecast values of above four factors especially from 2016 to 2018 by using GM (1, 1) model. Finally, we give the forecast values of financial revenue of Wenzhou from 2016 to 2018 by building an appropriate BP neural network model.

Keywords

Lasso Regression, GM (1, 1), BP Neural Network, Financial Revenue

基于数据挖掘技术的温州市财政收入的 影响因素分析及预测

庄亮亮, 黄辉林*, 吴 统

温州大学数学与信息科学学院, 浙江 温州
Email: *huilin_huang@wzu.edu.cn

收稿日期: 2017年11月23日; 录用日期: 2017年12月7日; 发布日期: 2017年12月14日

摘 要

本文首先建立了Lasso变量选择模型, 挑选出了影响温州市财政收入的四个关键性因素, 即在岗职工工
*通讯作者。

资总额、城镇居民人均消费性支出、地区生产总值和税收。在此基础上,分别针对这四个影响因素建立GM(1,1)模型,得到它们2016~2018年的预测值。最后,我们通过训练合适的BP神经网络模型,给出了温州市财政收入在2016~2018年的预测值。

关键词

Lasso回归, GM(1,1), BP神经网络, 财政收入

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

地方财政收入是区域国民经济的综合反映,也是市场经济国家的政府进行宏观调控的基础。随着我国社会主义市场经济体制的初步建立,特别是我国加入世界贸易组织,地方财政收入的分析与预测等问题越来越被社会各级领导所重视。地方财政收入分析及预测建模研究,不仅涉及区域经济发展、财税体制改革规则等方面的理论与实践,而且对有关部门制定合理的财政预算和财政政策,强化地方财政收入的监督管理,促进国民经济快速、健康、协调发展均有十分重要的实际应用价值。因此对地方财政收入进行预测,不仅是必要的,而且也是可能的。科学、合理的预测地方财政收入,对于克服年度地方预算收支规模的随意性和盲目性,正确处理地方财政收入与经济的相互关系具有十分重要的指导意义。温州作为最具有中国特色的民营经济体,其经济发展在全国经济中的地位举足轻重。

众所周知,从2015年1月1日起我国开始施行的新《预算法》强调各级政府必须建立跨年度预算平衡机制。基于宏观经济预测以及财政政策以及相关变动,对多年的财政收入与支出进行合理预测一直是我们在财政管理实践中需要解决的问题,准确的预测可以有效提高政府预算编制质量和财政管理效率。对温州政府而言,及时对温州市的财政收入进行合理有效的预测,不仅可以有力贯彻“依法治税”的精神,还可以有效帮助我市解决预决算偏离度过大的问题。目前,温州市在财政收入规模和结构等方面还存在不断完善发展的空间。我们拟尝试在假定现行政策制度不变的情况下,通过引入当前先进的数据挖掘技术,建立科学合理的地方财政收入预测模型,并结合温州市的实际数据对温州市“十三五”期间的财政收入进行预测。

我国很多学者对财政收入的预测已经进行了很多研究。丁文斌(2003) [1]以北京市的月度数据为基础,采用计量经济学中的协整分析方法对北京市地方财政收入与GDP的协整关系进行了详细分析。韦邦荣、杨玉生(2006) [2]利用协整理论对中国财政收入与GDP之间的关系进行实证研究,研究表明中国财政收入与GDP之间存在长期均衡的协整关系和短期动态调整机制,并且二者之间相互促进。梁蕾(2009) [3]利用我国1952~2007年的GDP和财政收入数据对二者之间的协整关系进行了实证研究。郭秀、路勇(2004) [4]详细介绍了如何建立一种地方财政收入的预测模型。张伦俊(1999) [5]通过三种时间序列、线性回归、自回归模型三种方式对税收收入进行预测,并比较了三种方法的预测精度,在此基础上对原模型进行修改和改进。程毛林、张伦俊(2005) [6]通过两种多元非线性回归模型的预测方法研究了全国财政收入与三次产业增加值之间的关系。白萍(2005) [7]通过对财政收入与三次产业增加值、社会从业人员以及收入水平等自变量之间的多元回归,利用岭回归方法解决其中存在的异方差、多重共线性等问题。孙元、吕宁(2007) [8]针对浙江省样本数据,利用传统的时间序列、多元线性回归和基于学习理论的SVM理论,分别对该

省一般预算收入和主要税种建立单项预测模型,最后应用组合预测将上述方法综合,对浙江省十一五期间的财政收入做出最终预测。范敏、石为人等人(2008) [9]基于地方财政收入受到多因素影响和经济系统非线性的性质,通过将灰色预测模型和神经网络模型结合起来的方法进行实证分析,实证结果显示该方法在地方财政预测中是可行的。崔志坤、朱秀变(2010) [10]通过分析影响财政收入的因素,运用简单的线性回归方程和 AR 模型,对我国近期和中长期财政收入进行了预测,并通过预测的财政收入中长期数据和中国社会发展现状,得到一些结论和启示。聂少林、杜卓等人(2011) [11]基于提高预测精度的考虑,将时间序列模型、收入结构模型、月度预测模型和线性回归模型进行组合,对辽宁省样本数据进行了实证分析。

综上所述可以看出,我国财政收入预测方面的研究成果较为丰富,近年来各地方政府也越来越重视建立各省、市地方财政收入预测模型,但是由于研究起步晚,研究成果还处在初级阶段,各种研究的实用性和系统性还有待进一步发展。研究内容和研究所侧重的角度也有很大的拓展空间。特别地,我国很多学者已经在财政收入的影响因素分析方面(如刘欣竹等(2012) [12]、王辉等(2015) [13]、许林(2006) [14])进行了很多研究,但他们大多数或者建立财政收入与影响因素之间的多元线性回归模型,要么对数据取对数以后建立相应的多元线性回归模型。多元线性回归模型通常使用普通的最小二乘法对回归模型的系数进行估计,预测变量的选取则采用逐步回归。事实上,不论是普通最小二乘法还是逐步回归,都有其不足之处。他们一般都局限于局部最优解而不是全局最优解。如果预测变量过多,子集选择的计算过程具有不可实行性,且子集选择具有内在的不连续性,从而导致子集选择极度多变。近几十年来,现代统计技术不断发展和完善,对新的数据运用新的研究方法来考察地方财政收入的影响因素是很有必要的。我们拟在已有研究的基础上应用 Lasso 变量选择方法建立影响温州市财政收入的因素选择模型。Tibshirani (1996) [15]所提出的 Lasso 算法则是一种能够实现指标集合精简的估计方法。利用该方法我们可以找出温州市财政收入的关键性影响因素。其次,本课题将针对温州市财政收入的关键性影响因素分别建立单影响因素的灰色预测模型 GM (1, 1) (刘思峰等人(2014))。最后,我们还将建立基于神经网络技术的温州市地方财政收入的分析预测模型,从而对温州市 2016~2018 年的财政总收入进行合理预测。

2. 变量选取和数据初探

2.1. 变量选取

在综合已有研究工作的基础上,结合经济理论对财政收入的诠释,通过查询温州市统计局统计年鉴,本文初步选择了 1994~2015 年度以下各因素为温州市财政收入(y)的解释变量,详见表 1。

x_1 : 在岗职工工资总额(亿元),指一定时期内职工的工资水平,可以用来反映社会分配情况。

x_2 : 社会消费品零售总额(亿元),反映了社会整体消费状况。通常人们的消费需求扩大会引起税收以及经济系统中其他方面发生变化,从而拉动财政收入的增长。

x_3 : 城镇居民人均可支配收入(元),城镇居民人均可支配收入越高就表示人们能用于平常生活中的收入越高,日常消费越大,财政收入的增加值也就越多。

x_4 : 城镇居民人均消费性支出(元),城镇居民人均消费支出越大说明人们的消费水平越高,消费的越多,财政收入也会相应增加。

x_5 : 年末总人口(万人),人口和经济的联系是双向的,在其余因素不变的情况下,地方总人口数越多,人均财政收入越少。

x_6 : 全社会固定资产投资额(亿元),是社会固定资产再生产的主要手段,能调整经济结构,增强经济实力,在一定程度上影响着财政收入。

x_7 : 地区生产总值(亿元),等于各行业增加值的总和,反映一个地区的经济发展水平,通常地区经

Table 1. Financial revenue and related data of Wenzhou
表 1. 温州市财政收入及其相关数据

Year	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	y
1994	25.53	136.38	5625	3942	692.4	91.43	295.87	30.63	18.82	0.54	21.68
1995	30.83	221.21	7507	5265	697.89	128.82	401.66	41.42	24.28	0.56	26.49
1996	37.29	274.89	8277	5868	704.37	157.02	507.05	47.37	30.01	0.56	32.20
1997	42.78	314.63	9034	6524	708.35	178.11	601.85	51.68	36.11	0.60	38.71
1998	48.08	345.04	8968	6311	718.04	201.20	672.06	53.13	42.36	0.62	45.99
1999	54.61	373.12	10,339	7317	721.62	234.66	729.07	52.37	53.08	0.67	55.15
2000	61.30	397.91	12,051	8194	736.32	265.25	822.02	53.21	69.74	0.69	73.87
2001	76.50	438.45	13,200	8999	738.81	328.66	924.30	56.19	90.92	0.73	96.11
2002	88.34	491.09	14,591	11,612	739.12	387.75	1052.35	54.49	115.97	0.75	126.26
2003	110.74	522.54	16,035	12,619	742.28	448.60	1212.49	55.67	136.79	0.75	151.77
2004	146.62	600.83	17,727	14,212	746.19	507.32	1388.91	62.62	163.87	0.78	182.44
2005	170.81	682.94	19,805	15,822	750.28	542.11	1590.82	64.89	182.92	0.76	204.92
2006	210.00	783.91	21,716	16,473	756.48	645.55	1826.92	65.51	215.76	0.75	241.09
2007	253.44	914.77	20,922	14,947	764.57	737.03	2146.62	70.82	277.41	0.79	293.26
2008	296.81	1096.21	22,851	16,564	771.99	758.44	2407.46	75.54	322.13	0.82	339.78
2009	336.72	1264.72	24,467	17,825	779.11	837.78	2520.51	80.31	344.10	0.88	360.72
2010	379.51	1498.10	27,981	20,295	786.8	925.98	2918.82	93.69	386.76	0.89	411.43
2011	431.78	1790.68	31,749	23,337	798.36	1751.52	3407.97	107.88	251.45	0.93	485.62
2012	484.11	1929.29	34,820	23,975	800.21	2357.11	3670.56	114.22	267.68	0.99	517.89
2013	506.72	2136.38	37,852	25,624	807.24	2950.00	4024.50	114.02	295.31	1.04	565.63
2014	554.32	2410.36	40,510	27,186	813.69	3053.35	4303.05	117.94	326.75	1.06	611.03
2015	594.40	2677.73	44,026	29,438	811.21	3456.39	4619.84	129.57	349.56	1.22	675.95

济越发达，生产总值就越高，财政收入就越多。

x_8 ：第一产业总值(亿元)，随着我国产业结构的转型升级，第一产业总值占生产总值的比例减小，对财政收入的影响日益减弱。

x_9 ：税收(亿元)，税收是财政收入的主要来源。

x_{10} ：第三产业与第二产业产值比，表示温州市的现有产业结构，产业结构的优化会刺激财政收入的增长。

2.2. 数据描述性分析

首先对数据进行描述分析，了解数据的总体特征，结果如表 2 所示。由表 2 我们发现，城镇居民人均可支配收入 x_3 、城镇居民人均消费性支出 x_4 、全社会固定资产投资额 x_6 、地区生产总值 x_7 标准差都在 770 以上，说明各年数值间差异性非常大。财政收入的均值和标准差为 252.64 和 212.34，这说明温州市各年份财政收入平均值 252.64 亿元，且各年份财政收入存在较大差异。

2.3. 相关分析

相关系数是用来反映两变量间线性相关关系。表 3 是 10 个解释变量与被解释变量财政收入之间的 Pearson 相关系数矩阵，由表 3 可知各影响因素与财政收入存在较高的正相关关系。

2.4. 多重共线性检验

通过传统最小二乘法建立多元线性回归模型，计算各个解释变量的方差膨胀因子 $VIF_i = 1/(1 - R_i^2)$ ，见表 4。通常情况下，如果 VIF 值大于 100，则认为有严重的多重共线性。表 4 中的结果表明各个变量之间都具有较为严重的多重共线性现象。

Table 2. Descriptive statistics of the main variables

表 2. 主要变量的描述性统计

变量	Min	Max	Mean	SD
x_1	25.53	594.40	224.60	191.81
x_2	136.38	2677.73	968.24	772.68
x_3	5625.00	44026.00	20456.95	11464.85
x_4	3942.00	29438.00	14652.23	7762.69
x_5	692.40	813.69	753.88	38.25
x_6	91.43	3456.39	952.00	1046.76
x_7	295.87	4619.84	1911.12	1377.08
x_8	30.63	129.57	72.42	28.09
x_9	18.82	386.76	181.90	125.69
x_{10}	0.54	1.22	0.79	0.18
y	21.68	675.95	252.64	212.34

Table 3. Pearson correlation coefficient matrix

表 3. Pearson 相关系数矩阵

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	y
x_1	1.00	0.99	0.99	0.98	0.96	0.94	1.00	0.98	0.89	0.96	1.00
x_2	0.99	1.00	0.99	0.97	0.94	0.97	0.99	0.99	0.84	0.97	0.99
x_3	0.99	0.99	1.00	0.99	0.97	0.95	0.99	0.98	0.88	0.98	0.99
x_4	0.98	0.97	0.99	1.00	0.98	0.91	0.98	0.97	0.90	0.97	0.99
x_5	0.96	0.94	0.97	0.98	1.00	0.87	0.97	0.96	0.92	0.96	0.97
x_6	0.94	0.97	0.95	0.91	0.87	1.00	0.94	0.95	0.70	0.93	0.94
x_7	1.00	0.99	0.99	0.98	0.97	0.94	1.00	0.99	0.89	0.97	1.00
x_8	0.98	0.99	0.98	0.97	0.96	0.95	0.99	1.00	0.84	0.96	0.98
x_9	0.89	0.84	0.88	0.90	0.92	0.70	0.89	0.84	1.00	0.86	0.90
x_{10}	0.96	0.97	0.98	0.97	0.96	0.93	0.97	0.96	0.86	1.00	0.97
y	1.00	0.99	0.99	0.99	0.97	0.94	1.00	0.98	0.90	0.97	1.00

Table 4. Variance expansion factor of each variable
表 4. 各个变量方差膨胀因子

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
VIF	614	391	1504	513	119	348	2132	132	102	78

3. 模型建立和实证分析

3.1. Lasso 回归模型建立

Lasso 回归模型(Trevor Hastie *et al.* (2008) [16], P. 68)是目前比较流行的高维数据分析方法之一,它可以很好的克服多重共线性问题,并且还可以同时进行变量选择和参数估计,其参数估计方程如下:

$$\beta^{\text{lasso}} = \arg \min_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

利用 R 语言中的 `glmnet` 包中的 `glmnet()` 函数,我们通过交叉验证选出 Lasso 回归的最佳 λ 值等于 2.553376,如图 1 所示。

我们将样本的 3/4 作为训练样本,剩余的 1/4 样本作为测试样本。通过交叉验证选取最优的参数 $\lambda = 2.553$,此时我们可以获得 Lasso 回归的参数估计结果以及测试均方误差 MSE 和测试均方根误差 RMSE 如表 5 最后一列所示。我们将 Lasso 回归的结果与传统的最小二乘方法、向前逐步回归以及向后逐步回归方法的参数估计结果进行了对比,通过观察比较这些方法所获得的测试均方误差(MSE)和测试均方根误差(RMSE)我们发现,Lasso 回归的测试 MSE = 110.216,其 RMSE = 10.498,这两个数值相比于最小二乘、向前逐步回归以及向后逐步回归所获得的 MSE 和 RMSE 要小得多,这说明 Lasso 回归在多重共线性非常严重时,除了是一种非常好的变量选择方法外,其所建立的方程预测效果相比于这三种方法更是具有得天独厚的优势。同样,我们还可以将 Lasso 回归结果和岭回归结果进行对比,通过交叉验证,我们选择最优岭回归参数 $\lambda = 18.437$,此时所获得的岭回归模型的测试均方误差 MSE = 118.906,而 RMSE = 10.904。这表明针对温州市 1994~2015 年的财政收入及其影响因素数据而言,Lasso 回归的预测效果略胜于岭回归的预测效果。

由表 5 中最后一列结果,我们可以获得温州市财政收入(y)的 Lasso 回归方程如下:

$$\hat{y} = -32.644 + 0.315x_1 + 0.002x_4 + 0.092x_7 + 0.054x_9 \quad (2)$$

由此,我们得到温州市财政收入的 4 个主要影响因素,它们分别是: x_1 ——在岗职工工资总额(亿元), x_4 ——城镇居民人均消费性支出(元), x_7 ——地区生产总值(亿元), x_9 ——税收(亿元)。

3.2. 建立灰色系统模型 GM (1, 1) (详细介绍可参考刘思峰等人(2014) [17])

我们拟通过 GM (1, 1)模型预测四个关键性因素 2016~2018 年的值。

灰色预测,是指对系统行为特征值的发展变化进行的预测,对既含有已知信息又含有不确定信息的系统进行的预测,也就是对在一定范围内变化的、与时间序列有关的灰过程进行预测。尽管灰过程中所显示的现象是随机的、杂乱无章的,但毕竟是有序的、有界的,因此得到的数据集合具备潜在的规律。灰色预测是利用这种规律建立灰色模型对灰色系统进行预测。目前使用最广泛的灰色预测模型就是关于数列预测的一个变量、一阶微分的 GM (1, 1)模型。它是基于随机的原始时间序列,经按时间累加后所形成的新的时间序列呈现的规律可用一阶线性微分方程的解来逼近。经证明,经一阶线性微分方程的解逼近所揭示的原始时间序列呈指数变化规律。因此,当原始时间序列隐含着指数变化规律时,灰色模型 GM (1, 1)的预测是非常成功的。

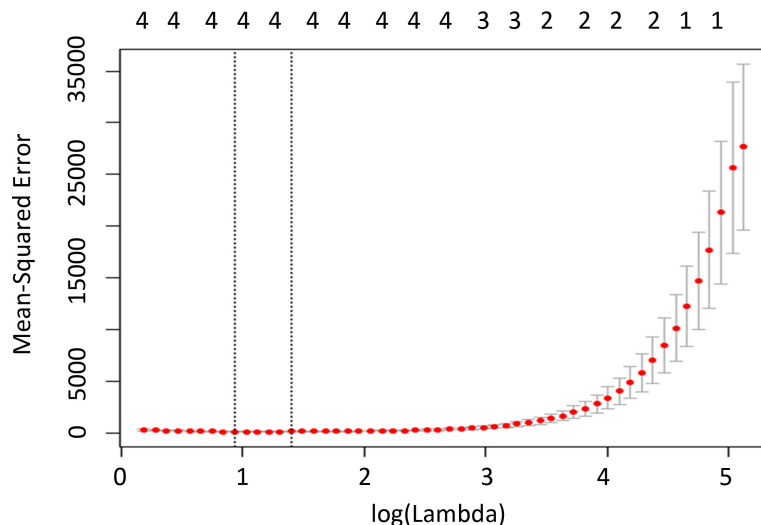


Figure 1. Cross-validation
图 1. 交叉验证

Table 5. Parameter estimation and error estimation results of five different methods
表 5. 五种不同方法的参数估计以及误差估计结果

	最小二乘	岭回 $\lambda = 18.437$	向前逐步	向后逐步	Lasso 回归 $\lambda = 2.553$
Intercept	-137.765	-266.701	34.198	34.198	-32.644
x_1	-0.017	0.184			0.315
x_2	0.091	0.038	0.083	0.083	
x_3	-0.021	0.002			
x_4	0.004	0.003	0.002	0.002	0.002
x_5	0.309	0.281			
x_6	0.016	0.030			
x_7	0.137	0.022	0.151	0.151	0.092
x_8	-2.358	0.394	-2.480	-2.480	
x_9	0.055	0.206			0.054
x_{10}	-76.335	27.305			
MSE	570.016	118.906	554.354	554.354	110.216
RMSE	23.875	10.904	23.545	23.545	10.498

为方便理解起见，我们对 GM (1, 1)先做个简短的介绍。令 $X^{(0)}$ 为 GM (1, 1)建模序列，

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$$

$X^{(1)}$ 为 $X^{(0)}$ 的 1-AGO 序列，

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)),$$

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$$

令 $Z^{(1)}$ 为 $X^{(1)}$ 的紧邻均值(MEAN)生成序列

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)),$$

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

则 GM (1, 1)的灰微分方程模型为

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (3)$$

式中 a 称为发展系数, b 为灰色作用量。设 $\hat{\alpha}$ 为待估参数向量, 即 $\hat{\alpha} = (a, b)^T$, 则灰微分方程(3)的最小二乘估计参数列满足

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_n$$

其中

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

称

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (4)$$

为灰色微分方程 $x^{(0)}(k) + az^{(1)}(k) = b$ 的白化方程, 也叫影子方程。

如上所述, 则有

a) 白化方程 $\frac{dx^{(1)}}{dt} + ax^{(1)} = b$ 的解也称时间响应函数为

$$\hat{x}^{(1)}(t) = \left(x^{(1)}(0) - \frac{b}{a} \right) e^{-at} + \frac{b}{a}$$

b) GM (1, 1)灰色微分方程 $x^{(0)}(k) + az^{(1)}(k) = b$ 的时间响应序列为

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(0) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}, \quad k=1, 2, \dots, n$$

c) 取 $x^{(1)}(0) = x^{(0)}(1)$, 则

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}, \quad k=1, 2, \dots, n$$

d) 还原值

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$$

上式即为预测方程。

下面我们建立灰色系统模型 GM (1, 1), 通过 R 语言编程, 计算在岗职工工资总额 x_1 、城镇居民人均消费性支出 x_4 、地区生产总值 x_7 、税收 x_9 等因素的 2016~2018 年的预测值, 并对其预测精度给予评级, 结果如表 6。

表6的结果显示,我们通过灰色系统模型GM(1,1)得到近三年2016~2018年度在岗职工工资总额 x_1 、城镇居民人均消费性支出 x_4 、地区生产总值 x_7 和税收 x_9 等4个关键因素的预测值,同时由于它们后验差比值检验C值均小于0.35,故而我们给予这些因素预测精度评估等级为好。

3.3. 温州财政收入神经网络预测模型

BP神经网络是一种按误差逆传播算法训练的多层前馈网络,能实现任何复杂非线性映射。本文我们综合考虑温州市预测财政收入的各种因素的影响,以便反映各种因素的综合效应。我们将结合Lasso变量选择的结果以及四个关键性影响因素在岗职工工资总额 x_1 、城镇居民人均消费性支出 x_4 、地区生产总值 x_7 和税收 x_9 的2016~2018年的预测值,训练合适的BP神经网络模型,我们拟在充分考虑各个关键因素的历史信息的综合效应后得到温州市财政收入2016~2018年的预测值。我们将BP神经网络的参数设置如下:误差精度设为 10^{-7} ,学习次数为10,000次,神经元个数为Lasso变量选择的变量个数4个。通过R语言编程,我们可以获得温州市财政收入2016~2018年的预测值分别为763.05646亿,769.90083亿,769.52872亿元。从图2我们可以看出,BP神经网络在预测温州市财政收入时具有非常好的效果,其真实值曲线和预测值曲线几乎重合。

4. 结论和启示

本文建立了Lasso变量选择模型、GM(1,1)以及BP神经网络模型的组合预测模型。一方面,我们通过Lasso回归模型方程式(2)深入刻画了在岗职工工资总额、城镇居民人均消费性支出、地区生产总值以及税收等4个因素与温州市财政收入的量化关系,由此我们得出的结论如下:1)温州市在岗职工工资总额对温州市财政收入具有较大的正向刺激作用,它每增加一个亿,温州市财政收入将增加3150万,所以温

Table 6. The forecast value and prediction accuracy of the financial revenue in Wenzhou

表6. 灰色预测模型温州财政收入相关因素预测值及预测精度表

变量	x_1	x_4	x_7	x_9
2016年预测值	917.468	33204.87	5823.009	487.850
2017年预测值	1047.187	35968.44	6517.164	531.278
2018年预测值	1195.247	38962.02	7294.068	578.573
后验差比值检验C值	0.063	0.063	0.091	0.237
预测精度等级	好	好	好	好

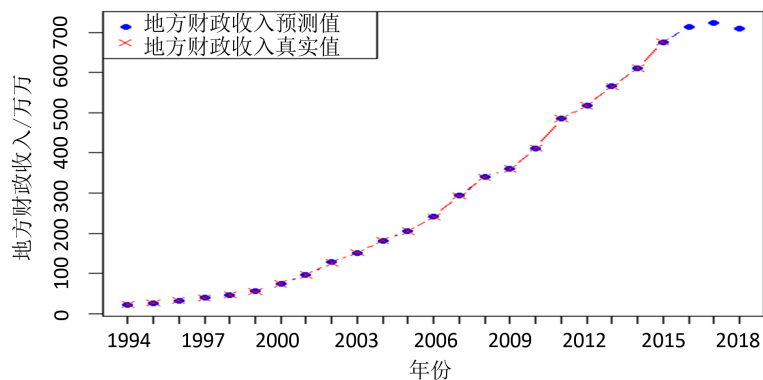


Figure 2. Comparison of real value and neural network predictive value of local fiscal revenue in Wenzhou

图2. 温州地方财政收入真实值和神经网络预测值对比图

温州市政府应努力吸引高端人才, 努力创造更多的就业岗位, 为高精尖人才创造好的工作和生活环境, 鼓励企业为高精尖人才提供好的待遇, 让温州本土的人才愿意留下来, 同时还能吸引外面的人才走进温州, 为温州的经济建设添砖加瓦, 让温州市的经济发展有人可用, 有才可依; 2) 城镇居民人均消费性支出也可正向刺激温州市财政收入的增加, 政府可适度因地制宜, 逐步扩大温州市内需, 以消费内需刺激经济增长, 逐步引导温州市的经济走出疲态从而形成一个良性循环; 3) 温州市 GDP 和税收收入也是影响温州市财政收入增长的正向因子, 由方程式(2)我们可以看出温州市地区生产总值每增加一个亿, 温州市财政收入将增加 920 万, 而温州市税收每增加一个亿, 温州市财政收入将增加 540 万, 温州市政府应制定平衡稳定温州市经济发展的方针政策和策略, 必要的话可适时推进相关税制改革。另一方面, 我们通过建立 GM (1, 1)和 BP 神经网络两个模型的组合预测模型, 给出了温州市财政收入在 2016~2018 年的预测值, 这可为温州市政府编制未来的财政支出预算提供决策依据。

基金项目

该工作得到了浙江省自然科学基金(No.LY17A010013)、温州市科技局软科学项目(No.R20160005)、国家自然科学基金(No.11201344)、温州大学 2017 大学生创新创业项目(No.DC2017043)的大力支持, 对此表示诚挚感谢。

参考文献 (References)

- [1] 丁文斌. 北京市地方财政收入与 GDP 的协整关系分析[J]. 北京统计, 2003(8): 48-51.
- [2] 韦邦荣, 杨玉生. 中国财政收入与 GDP 之间关系的协整分析与误差修正模型研究[J]. 统计与信息论坛, 2006(1): 49-53.
- [3] 梁蕾. 我国财政收入与 GDP 之间协整关系的分析[J]. 华北电力大学学报(社会科学版), 2009(1): 62-64.
- [4] 郭秀, 路勇. 构建一种地方财政收入的预测模型[J]. 价值工程, 2004, 23(3): 98-100.
- [5] 张伦俊. 税收预测模型的拟合与分析[J]. 数理统计与管理, 1999(1): 5-8.
- [6] 程毛林, 张伦俊. 多元非线性经济预测模型的建立方法[J]. 统计与决策, 2005(5): 20-21.
- [7] 白萍. 影响我国财政收入的多元回归线性模型[J]. 统计与决策, 2005(5): 92-94.
- [8] 孙元, 吕宁. 地方财政一般预算收入预测模型及实证分析[J]. 数量经济技术经济研究, 2007(1): 38-45.
- [9] 范敏, 石为人, 梁勇林, 等. 组合预测模型在地方财政收入预测中的应用[J]. 重庆大学学报, 2008, 31(5): 536-540.
- [10] 崔志坤, 朱秀变. 中国近期及中期财政收入预测分析[J]. 统计与决策, 2010(11): 112-115.
- [11] 聂少林, 杜卓, 於良, 等. 辽宁地方财政收入现状及增长预测[J]. 地方财政研究, 2011(2): 43-47.
- [12] 刘心竹, 李昊, 刘青青, 等. 我国地方财政收入影响因素的实证分析[J]. 中国集体经济, 2012(9): 84-85.
- [13] 王辉, 宋翔, 庄园. 影响我国财政收入的主要因素实证分析[J]. 新经济, 2015(10): 16-17.
- [14] 许林. 地方财政收入的影响因素分析——以重庆市为例的实证研究[J]. 特区经济, 2006(09): 277-278.
- [15] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B*, **58**, 267-288.
- [16] Hastie, T., Tibshirani, R. and Friedman. F. (2008) *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*. 2nd Edition, Springer, New York.
- [17] 刘思峰, 曾波, 刘解放, 等. GM (1, 1)模型的几种基本形式及其适用范围研究[J]. 系统工程与电子技术, 2014, 36(3): 501-508.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2169-2556，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：ass@hanspub.org