

doi:10.16055/j.issn.1672-058X.2022.0001.013

基于 IOWA 算子的我国居民消费水平 组合预测研究

李颖, 庄科俊

(安徽财经大学 统计与应用数学学院, 安徽 蚌埠 233030)

摘要:为了研究我国居民消费水平的变化趋势,将预测我国未来 4 年(2020—2023 年)的居民消费水平,选择的样本数据是从 2002—2019 年这 18 年间的我国居民消费水平,构造基于 ARIMA(2,1,1)模型、Holt-Winters 无季节模型以及多元回归模型的诱导有序加权算术平均(IOWA)算子组合预测模型,主要运用使偏差平方和最小的组合预测优化方法预测我国未来 4 年(2020—2023 年)的居民消费水平,并对诱导有序加权算术平均 IOWA 算子组合预测模型进行了有效性评价;以新发展格局的内循环和外循环为依据,寻找解释变量并建立多元回归模型;研究发现:IOWA 算子组合预测模型预测精度大于各单项预测模型,并且各项误差均显著低于单项预测模型,未来 4 年我国居民消费水平不会有大幅度波动,但有小幅度下降的趋势。

关键词:居民消费水平;IOWA 算子;组合预测模型

中图分类号:F224.0

文献标志码:A

文章编号:1672-058X(2022)01-0092-09

0 引言

目前,对于非负变权重组合预测模型在理论与应用两方面的研究比较丰富。在组合预测模型的非负权重研究方面,由 Bates 和 Granger^[1]首次提出组合预测模型的概念,通常将他们所提出的最小方差组合预测模型简称为 B-G 组合预测模型。但是,根据 B-G 组合预测模型所选取的最优权重可能会产生具有负分量的权系数向量,负的权重分量可能是由组合预测模型存在的多重共线性导致的,又由于对负权重经济含义的解释存在争议,因此,很多学者建议加上非负权重约束条件。如孔庆凯、唐小我^[2-4]等一批学者在组合预测模型规避负权重方面提出了简单平均加权法、均方差倒数加权法和二项

式组合预测方法等,推动了组合预测模型在非负权重领域的迅速发展。组合预测模型在变权重领域的研究方面,由于传统的组合预测模型赋予单个预测模型在所有观察期内都相同的权重,可能会出现单个模型在某一观察期预测精度高却在另一观察期预测精度低,从而导致预测精度不一致的问题。显然,如果给单项模型赋予固定权重,就会导致预测精度不高、预测误差较大等。自 21 世纪以来,很多学者开展了对变权重组合预测模型的深入研究。陈华友等^[5]发现有序加权平均(OWA)算子组合预测模型有很大的缺点,因此他引出诱导有序加权平均(IOWA)算子的概念,并将此概念应用于预测企业所得税,调查结果显示诱导有序加权平均(IOWA)算子能够增加组合预测的精度。在预测我国居民消费水平的研究方面,宋峰等^[6]以 GM(1,1)模型、线

收稿日期:2021-01-26;修回日期:2021-03-24.

基金项目:国家自然科学基金(12061033);安徽省高校自然科学研究重点项目(KJ2020A0016).

作者简介:李颖(1998—),女,安徽淮南人,硕士研究生,从事数量经济学研究.

性回归模型以及时间序列模型为基础,构建组合预测模型,预测 2010—2012 年安徽省城镇居民人均消费;侯甜甜等^[7]通过建立 ARIMA 模型预测我国 CPI 的变化趋势,并且以主成分分析法所确定的影响我国 CPI 的主要因素作为解释变量建立多元线性回归模型;张甜瑞^[8]通过对 2000-01—2019-08 的月度数据建立 ARIMA(2,1,2)×(1,1,1)模型,预测未来一年陕西省的 CPI,并且把最后 8 期数据作为参照用于检验模型的预测精度,得出陕西省的居民价格指数不会存在剧烈波动。

综上所述,对 IOWA 算子组合预测模型的理论与应用研究较为成熟,但是关于预测我国居民消费水平的论文大多采用一种预测模型,或是采用赋予单个模型固定权重的组合预测模型,大多使用单项预测模型的论文都无法对预测模型的有效性进行检验。因此,本文根据 ARIMA(2,1,1)模型、Holt-Winters 无季节模型和多元回归模型,主要是利用有序加权算术平均 IOWA 算子组合预测模型,预测中国在接下来的 4 年内(2020—2023 年)居民消费水平,并构建模型的有效性评价体系。在新发展格局下,以内循环与外循环为依据寻找“双循环”的解释变量构建多元线性模型,使得对未来几年的居民消费水平预测具有更大的参考价值与实际意义。

1 模型简介

1.1 IOWA 算子

定义 1(OWA 算子) 设 n 元函数 $W_{\text{OWA}}:R^n \rightarrow R$, $\mathbf{W}=(w_1, w_2, \dots, w_n)^T$ 是与 W_{OWA} 相对应的权系数向量, \mathbf{W} 满足 $\sum_{i=1}^n w_i = 1, (w_i \geq 0, i=1, 2, \dots, n)$ 。令 $W_{\text{OWA}}(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{i=1}^n w_i \beta_i$, 其中 β_i 是 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 按照从大到小排列的第 i 个数值, 称 W_{OWA} 为 n 元有序加权算术平均算子, 简称 OWA 算子。

定义 2(IOWA 算子) 设 $(\langle \alpha_1, \beta_1 \rangle, \langle \alpha_2, \beta_2 \rangle, \dots, \langle \alpha_n, \beta_n \rangle)$ 为 n 个二维数组, $\mathbf{W}=(w_1, w_2, \dots, w_n)^T$ 是与 W_{IOWA} 相对应的权系数向量, \mathbf{W} 满足 $\sum_{i=1}^n w_i = 1,$

$w_i \geq 0, i=1, 2, \dots, n$ 。令 $W_{\text{IOWA}}(\langle \alpha_1, \beta_1 \rangle, \langle \alpha_2, \beta_2 \rangle, \dots, \langle \alpha_n, \beta_n \rangle) = \sum_{i=1}^n w_i \beta_{\alpha\text{-index}(i)}$, 对诱导因子 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 进行降序排列, $\alpha\text{-index}(i)$ 表示诱导因子的排列顺序 i , 简称 IOWA 算子。

显然, IOWA 算子是对 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 诱导因子按照降序排序后所对应的 $(\beta_1, \beta_2, \dots, \beta_n)$ 值进行有顺序的加权平均, 因此赋予各单项模型在每一时点上的权重与诱导值的大小无关, 而与诱导值的相对大小即排序有关。

1.2 IOWA 组合预测模型

假设对同一个问题用 $m(m \geq 2)$ 种方法进行预测。令 x_t 表示在时间 t 的实际观测值, x_{it} 表示方法 i 在时间 t 的预测值, e_{it} 表示方法 i 在时间 t 的预测误差 $(e_{it} = x_t - x_{it}; i=1, 2, \dots, m; t=1, 2, \dots, T)$, w_i 表示方法 i 占组合预测模型的权重 $(i=1, 2, \dots, m; \sum_{i=1}^m w_i = 1)$, 则样本期内预测值及误差的计算公式分别为 $\hat{x}_t = \sum_{i=1}^m w_i x_{it}, e_t = x_t - \hat{x}_t = x_t - \sum_{i=1}^m w_i x_{it} = \sum_{i=1}^m w_i x_t - \sum_{i=1}^m w_i x_{it} = \sum_{i=1}^m w_i (x_t - x_{it}) = \sum_{i=1}^m w_i e_{it} (t=1, 2, \dots, T)$ 。

本文采用预测精度作为诱导因子, 如果诱导因子 $\alpha_{it}(i=1, 2, \dots, m; t=1, 2, \dots, T)$ 选择第 t 期用第 i 种预测方法的预测精度, 则 α_{it} 的表达式为

$$\alpha_{it} = \begin{cases} 1 - \left| \frac{x_t - \hat{x}_{it}}{x_t} \right|, & \left| \frac{x_t - \hat{x}_{it}}{x_t} \right| \leq 1 \\ 0, & \left| \frac{x_t - \hat{x}_{it}}{x_t} \right| \geq 1 \end{cases}$$

对诱导因子所生成的 m 个二维数组 $(\langle \alpha_{1t}, x_{1t} \rangle, \langle \alpha_{2t}, x_{2t} \rangle, \dots, \langle \alpha_{mt}, x_{mt} \rangle)$ 按照诱导因子的降序排列, 根据误差平方和最小准则求得各精度的权重系数向量为

$$\mathbf{W}=(w_1, w_2, \dots, w_m)^T$$

$$\left(\sum_{i=1}^m w_i = 1, w_i \geq 0, i=1, 2, \dots, m \right)$$

由预测精度 $\alpha_{1t}, \alpha_{2t}, \dots, \alpha_{mt}$ 所生成的 $x_{1t}, x_{2t}, \dots, x_{mt}$, $(t=1, 2, \dots, T)$ IOWA 算子组合预测模型预测诱导预测误差: $e_{\alpha\text{-index}(it)} = x_t - x_{\alpha\text{-index}(it)} (i=1, 2, \dots, m; t=1, 2, \dots, T)$ 。

令 $\mathbf{R}_m = (1, 1, \dots, 1)^T$ 为 m 维单位向量, 权系数向量 $\mathbf{W} = (w_1, w_2, \dots, w_m)^T$ 的约束条件为 $\mathbf{R}_m^T \mathbf{W} = 1, \mathbf{W} \geq 0$, 于是 IOWA 组合预测模型在时间 t 的预测误差为

$$\begin{aligned} x_t - \hat{x}_t &= x_t - \sum_{i=1}^m w_i x_{\alpha-index(i)} = \\ &= \sum_{i=1}^m w_i x_t - \sum_{i=1}^m w_i x_{\alpha-index(i)} = \\ &= \sum_{i=1}^m w_i (x_t - x_{\alpha-index(i)})^2 = \\ &= \sum_{i=1}^m w_i e_{\alpha-index(i)} \quad (t = 1, 2, \dots, T) \end{aligned}$$

模型的总预测误差平方和为

$$\begin{aligned} U &= \sum_{t=1}^T e_t^2 = \sum_{t=1}^T \left(x_t - \sum_{i=1}^m w_i x_{\alpha-index(i)} \right)^2 = \\ \hat{x}_t &= \sum_{i=1}^m \left(\sum_{i=1}^m w_i e_{\alpha-index(i)} \right)^2 = \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^m \sum_{j=1}^m w_i w_j \left(\sum_{t=1}^T e_{\alpha-index(i)} e_{\alpha-index(j)} \right) = \\ &= \sum_{i=1}^m \sum_{j=1}^m w_i w_j \bar{E}_{ij} \quad (\bar{E}_{ij} = \bar{E}_{ji} = \\ &= \sum_{t=1}^T e_{\alpha-index(i)} e_{\alpha-index(j)}, i, j = 1, 2, \dots, m) \end{aligned}$$

$\bar{\mathbf{E}} = (\bar{E}_{ij})_{m \times m}$ 为 m 阶 IOWA 的预测误差信息矩阵, 于是 $U = \mathbf{W}^T \bar{\mathbf{E}} \mathbf{W}$ 。故基于误差平方和最小的优化准则 IOWA 组合预测模型可以写成:

$$\min U = \mathbf{W}_m^T \bar{\mathbf{E}}_m \mathbf{W}_m, \text{ s. t. } \begin{cases} \mathbf{R}_m^T \mathbf{W} = 1 \\ \mathbf{W} \geq 0 \end{cases}$$

1.3 模型评价体系

建立 IOWA 算子组合预测模型评价体系所使用的误差和平均精度如表 1 所示。

表 1 误差评价体系

Table 1 Error evaluation system

名称	公式
平方和误差	$F_{SSE} = \sum_{t=1}^N (x_t - \hat{x}_t)^2 = \sum_{t=1}^N e_t^2$
平均绝对误差	$F_{MAE} = \frac{1}{N} \sum_{t=1}^N x_t - \hat{x}_t = \frac{1}{N} \sum_{t=1}^N e_t $
平均相对误差	$F_{MRE} = \frac{1}{N} \sum_{t=1}^N \left \frac{x_t - \hat{x}_t}{x_t} \right = \frac{1}{N} \sum_{t=1}^N \varepsilon_t $
均方根误差	$F_{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2} = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2}$
均方根相对误差	$F_{RMSRE} = \sqrt{\frac{1}{N} \sum_{t=1}^N \left(\frac{x_t - \hat{x}_t}{x_t} \right)^2} = \sqrt{\frac{1}{N} \sum_{t=1}^N \varepsilon_t^2}$
均方百分比误差	$F_{MSPE} = \frac{1}{N} \sqrt{\sum_{t=1}^N \left(\frac{x_t - \hat{x}_t}{x_t} \right)^2} = \frac{1}{N} \sqrt{\sum_{t=1}^N \varepsilon_t^2}$
平均精度	$F_{MA} = \frac{1}{N} \sum_{t=1}^N \left(1 - \left \frac{x_t - \hat{x}_t}{x_t} \right \right) = \frac{1}{N} \sum_{t=1}^N (1 - \varepsilon_t)$

2 实证分析

2.1 数据选择与数据来源

用人均消费支出 ($cons_t$, 元) 衡量被解释变量。选取 2002—2019 年我国人均消费衡量我国居民消费水平,变化趋势如图 1 所示。根据贾康与刘薇^[9]在加快形成新发展格局的政策建议中所提到的把握新发展格局的要领,把解释变量分为内循环变量与外循环变量,分别选择人均可支配收入 (dpi_t , 元) 和实际利用外商直接投资金额 (fdi_t , 万美元)。根据于佳茹^[10]对我国居民消费水平的研究中所采取的变量,选取控制变量分别是国内消费税 (tax_t , 亿元)、恩格尔系数 (egc_t , %)。对于数据缺失值采用邻近

值填充法。数据均来源于中国统计年鉴。

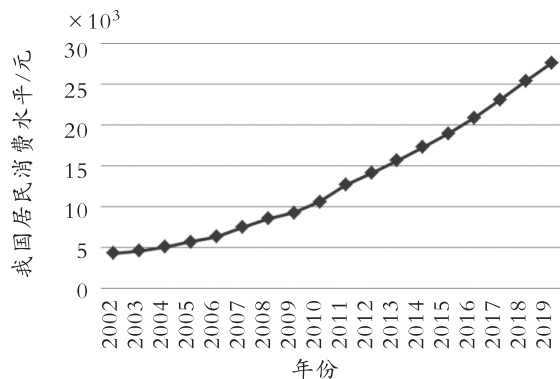


图 1 居民消费水平的变化趋势

Fig. 1 The variation trend of resident consumption level

2.2 变量的定义与描述性统计

对非平稳数列进行取对数以及取差分得到平稳数列(表 2)。

表 2 变量的选取与基本特征

Table 2 Variables and their basic characteristics

变量类别	简 写	变量定义	平均值	标准差	最小值	最大值
被解释变量	$\ln cons_t$	居民消费水平对数	4.045 3	0.268 4	3.629 0	4.440 3
内循环解释变量	$\ln dpi_t$	居民可支配收入对数	4.284 1	0.138 7	4.007 6	4.487 6
外循环解释变量	$\ln fdi_t$	实际利用外商直接投资金额取对数	6.972 2	0.148 2	6.722 2	7.140 3
控制变量	$d\ln tax_t$	国内消费税对数取一阶差分	0.063 5	0.060 0	-0.013 5	0.268 1
	$\ln egc_t$	恩格尔系数	1.556 0	0.061 5	1.459 0	1.627 9

2.3 单项预测模型

2.3.1 ARIMA 预测模型

将 1990 年到 2019 年间我国居民消费水平的年度数据作为样本,采用 ARIMA 模型对未来 4 年的人均消费进行预测,对 $\ln cons$ 做 ADF 检验,如表 3 所示,得到 1 阶差分时 P 值为 0.034 3,小于 0.05,所以 ARIMA(p, d, q) 中的 d 取 1。对 $\ln cons$ 做自相关与偏自相关分析可得:由偏自相关图知此模型的滑动平均项的项数为 2,由自相关图可知该模型的自回归项数为 1,故 ARIMA(p, d, q) 中的 p 值选 2, q 值选 1。而伴随概率 $P < 0.05$ 的 Q 统计量可以判断此序列不是白噪声序列。最终确定选择 ARIMA(2, 1, 1) 模型,用 $\ln cons$ 对 AR(1), AR(2), MA(1) 进行回归可知,该 ARIMA(2, 1, 1) 模型为

$$\ln cons_t = 3.704 7 + 1.887 5 \ln cons_{t-1} - 0.894 1 \ln cons_{t-2} + (6.946 5) \quad (22.296 67)$$

$$\varepsilon_t + 0.458 9 \varepsilon_{t-1} \quad (-10.801) \quad (2.401 56)$$

$$R^2 = 0.998 426 \quad \bar{R}^2 = 0.998 174$$

$$D.W. = 1.993 487 \quad F = 3 965.054$$

由 ARIMA(2, 1, 1) 模型的残差序列白噪声检验可知该序列是较为稳定的序列。而伴随概率 $P > 0.05$ 的最右侧的 Q 统计量可以表明是白噪声序列。查表得在 1% 的上下界中, $d_L = 1.01, d_U = 1.42, d_U = 1.42 \leq D.W. = 1.993 \leq 4 - d_U = 2.58$, 故该模型不存在自相关。根据该模型的 t 值、 P 值都显著,拟合优度较高以及残差序列为白噪声可知该模型拟合得很好。

2.3.2 Holt-Winters 无季节指数平滑模型

该模型是能够预测无四季变化但是有时间趋势的一种时间序列数据,Holt-Winters 无季节指数平滑模型如下:

$$a_t = \alpha y_t + (1 - \alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1}$$

其中, $\alpha, \beta \in (0, 1)$ 为阻尼因子; y_t 平滑后的 \hat{y}_t 为 $\hat{y}_{t+d} = a_t + b_t d$, a_t 表示截距, b_t 表示斜率。针对 Holt-Winters 无季节指数平滑模型,用 Stata 预测我国居民消费水平如表 3 所示。

2.3.3 多元回归模型

考虑我国居民消费水平和居民人均可支配收入、实际利用外商直接投资以及国内消费税和恩格尔系数都是时间序列,对非平稳数列取自然对数以及差分得到平稳数列,因此最终得到如下模型:

经检验:该模型的 $R^2 = 0.9944$,表明模型拟合得较好,杜宾值为 2.3528,查表得在 1% 的上下界中, $d_L = 0.71, d_U = 1.42, d_U = 1.42 \leq D.W. = 2.35 \leq 4 - d_U = 2.59$ 。

表 3 单项模型的预测值与预测精度

Table 3 Predicted values and prediction accuracies of single model

年 份	真实值/元	ARIMA 模型		H-W 无季节模型		多元回归模型	
		预测值/元	预测精度	预测值/元	预测精度	预测值/元	预测精度
2002	4 256	4 118.713	0.968	3 645.889	0.857	4 269.441	0.997
2003	4 542	4 619.783	0.983	5 053.417	0.887	4 549.723	0.998
2004	5 056	4 780.060	0.945	5 339.417	0.944	4 903.577	0.970
2005	5 671	5 709.909	0.993	5 853.417	0.968	5 466.747	0.964
2006	6 302	6 259.571	0.993	6 468.417	0.974	6 506.179	0.968
2007	7 434	6 936.825	0.933	7 099.417	0.955	7 485.831	0.993
2008	8 483	8 872.738	0.954	8 231.417	0.970	8 917.539	0.949
2009	9 226	9 318.843	0.990	9 280.417	0.994	9 122.671	0.989
2010	10 550	9 860.224	0.935	10 023.420	0.950	11 669.440	0.894
2011	12 646	12 209.007	0.965	11 347.420	0.897	12 353.983	0.977
2012	14 075	15 020.489	0.933	13 443.420	0.955	12 481.535	0.887
2013	15 615	14 931.590	0.956	14 872.420	0.952	16 528.788	0.941
2014	17 271	17 359.007	0.995	16 412.420	0.950	17 470.849	0.988
2015	18 929	18 702.390	0.988	18 068.420	0.955	18 968.151	0.998
2016	20 877	20 479.075	0.981	19 726.420	0.945	20 722.397	0.993
2017	23 070	22 774.234	0.987	21 674.420	0.940	23 023.111	0.998
2018	25 378	25 120.060	0.990	23 867.420	0.940	25 584.327	0.992
2019	27 563	27 469.950	0.997	26 175.420	0.950	26 442.651	0.959

根据上述杜宾值,可以排除自相关对该模型的影响,由该模型的 White Test 的 p 值为 0.6149,大于 0.05,可知该模型通过异方差检验,可以排除异方差对该模型的影响。因为大多数的 VIF 都大于 1 小于 10,不存在全大于 10 或者小于 1 的情况,故该模型不存在多重共线性。内生性检验的 $P =$

0.6183,且 $P > 0.05$,故可以排除内生性的影响。综上所述可知该模型拟合得较好。对该模型的 $\ln \text{dpi}$, $\ln \text{fdi}$, $\ln \text{egc}$ 分别建立时间序列模型,运用静态预测的方法得到 2020 年至 2023 年的预测值,对 $\ln \text{tax}$ 进行 Holt-Winters 预测,得到 2020 年至 2023 年的预测值,在此基础上预测我国居民消费水平未来 4 年

(2020—2023 年)的值。

2.4 组合预测模型

2.4.1 模型的构建

构造的 IOWA 算子组合预测模型偏差平方和最低,具体思路如下:首先,按照预测精度的高低将各单项预测模型在每一时期的预测值进行重新排列,

求出预测精度从高到低的误差大小;其次,构建误差分布表,求得模型的信息矩阵;最后,根据误差信息矩阵求得各精度预测值所占有的权重,并构建组合预测模型。

先按照预测精度的高低重新排列各单项预测模型在每一时期的预测值(表 4)。

表 4 按照精度由大到小排序的预测值与精度

Table 4 The predicted values and precision are sorted from most to least in order of precision

年 份	真实值/元	精度最高			精度次高			精度最低		
		预测值/元	误差/元	精度	预测值/元	误差/元	精度	预测值/元	误差/元	精度
2002	4 256	4 269.441	-13.441	0.997	4 118.713	137.287	0.968	3 645.889	610.111	0.857
2003	4 542	4 549.723	-7.723	0.998	4 619.783	-77.783	0.983	5 053.417	-511.417	0.887
2004	5 056	4 903.577	152.423	0.970	4 780.060	275.940	0.945	5 339.417	-283.417	0.944
2005	5 671	5 709.909	-38.909	0.993	5 853.417	-182.42	0.968	5 466.747	204.253	0.964
2006	6 302	6 259.571	42.429	0.993	6 468.417	-166.42	0.974	6 506.179	-204.179	0.968
2007	7 434	7 485.831	-51.831	0.993	7 099.417	334.583	0.955	6 936.825	497.175	0.933
2008	8 483	8 231.417	251.583	0.970	8 872.738	-389.74	0.954	8 917.539	-434.539	0.949
2009	9 226	9 280.417	-54.417	0.994	9 318.843	-92.843	0.990	9 122.671	103.329	0.989
2010	10 550	10 023.42	526.580	0.950	9 860.224	689.776	0.935	11 669.44	-1 119.44	0.894
2011	12 646	12 353.98	292.017	0.977	12 209.01	436.993	0.965	11 347.42	1 298.580	0.897
2012	14 075	13 443.42	631.580	0.955	15 020.49	-945.49	0.933	12 481.54	1 593.465	0.887
2013	15 615	14 931.59	683.410	0.956	14 872.42	742.580	0.952	16 528.79	-913.788	0.941
2014	17 271	17 359.01	-88.007	0.995	17 470.85	-199.85	0.988	16 412.42	858.580	0.950
2015	18 929	18 968.15	-39.151	0.998	18 702.39	226.610	0.988	18 068.42	860.580	0.955
2016	20 877	20 722.40	154.603	0.993	20 479.08	397.925	0.981	19 726.42	1 150.580	0.945
2017	23 070	23 023.11	46.889	0.998	22 774.23	295.766	0.987	21 674.42	1 395.580	0.940
2018	25 378	25 584.33	-206.33	0.992	25 120.06	257.940	0.990	23 867.42	1 510.580	0.940
2019	27 563	27 469.95	93.050	0.997	26 442.65	1 120.35	0.959	26 175.42	1 387.580	0.950

其次,构建误差分布表,求得模型的误差信息矩阵为

$$E = \begin{pmatrix} 1\ 410\ 839.401\ 4\ 466\ 810.049\ 7\ -81\ 625.086\ 4 \\ 466\ 810.049\ 7\ 4\ 205\ 428.156\ 7\ 816\ 560.612\ 3 \\ -81\ 625.086\ 4\ 816\ 560.612\ 3\ 16\ 514\ 258.530\ 5 \end{pmatrix}$$

以误差平方和最小为目标建立(IOWA)算子组合预测模型为

$$\begin{aligned} \min U &= W_3^T E W_3 \\ \text{s. t. } &\begin{cases} w_1 + w_2 + w_3 = 1 \\ w_1 \geq 0, w_2 \geq 0, w_3 \geq 0 \end{cases} \end{aligned}$$

即组合优化模型:

$$\begin{aligned} \min U &= 1\ 410\ 839.401\ 4w_1^2 + 933\ 620.099\ 4w_1w_2 - \\ &163\ 250.172\ 8w_1w_3 + 4\ 205\ 428.156\ 7w_2^2 + \\ &16\ 514\ 258.530\ 5w_3^2 + 1\ 633\ 121.224\ 6w_2w_3 \\ \text{s. t. } &\begin{cases} w_1 + w_2 + w_3 = 1 \\ w_1 \geq 0, w_2 \geq 0, w_3 \geq 0 \end{cases} \end{aligned}$$

利用 LINGO 求解可得:

$$w_1 = 0.759, w_2 = 0.176, w_3 = 0.065, U = 1\ 148\ 140$$

最后,根据误差信息矩阵求得各精度预测值所占有的权重,并构建组合预测模型。根据该预测模型算出 2002 年到 2019 年我国居民消费水平(表

5)。求得 IOWA 组合预测模型为

$$\hat{x}_t = 0.759x_{\alpha\text{-index}(1t)} + 0.176x_{\alpha\text{-index}(2t)} + 0.065x_{\alpha\text{-index}(3t)} \quad (t = 1, 2, \dots, 18)$$

表 5 组合预测模型的预测值

Table 5 The predicted values of the combined prediction model

年份	预测值/元	预测精度	年份	预测值/元	预测精度
2002	4 202.382	0.987	2011	12 263.05	0.970
2003	4 594.794	0.988	2012	13 658.46	0.970
2004	4 910.168	0.971	2013	15 024.99	0.962
2005	5 719.361	0.991	2014	17 317.16	0.997
2006	6 312.357	0.998	2015	18 862.90	0.997
2007	7 382.137	0.993	2016	20 614.83	0.987
2008	8 388.887	0.989	2017	22 891.64	0.992
2009	9 276.926	0.994	2018	25 391.02	0.999
2010	10 101.69	0.958	2019	27 205.00	0.987

2.4.2 模型评价

用误差平方和最小的 IOWA 组合预测模型分别计算各项误差并进行归一化处理(与各种误差最大值进行比较,表 6)。由表 6 可知,该组合预测模型的平方和误差只有 Holt - Winters 无季节模型的

8.95%,平均绝对误差只有 Holt-Winters 无季节模型的 25.63%,平均相对误差只有 Holt-Winters 无季节模型的 26.25%,均方根误差只有 Holt-Winters 无季节模型的 29.92%,均方根相对误差和均方百分比误差都只有 ARIMA(2,1,1)模型的 1.97%,平均精度高达 98.5%。因此,用此组合预测模型对未来 4 年我国居民消费水平进行预测会具有较高的预测精度。

2.4.3 预测结果

IOWA 组合模型对历史时期的预测是根据由实际值与预测值计算得出的预测精度进行诱导排序的,但未来的实际值并没有数据,故不能使用上文中求得的权重来预测未来 4 年我国居民的消费水平,故下面依据每一时期各单项预测模型在组合预测模型中所做出的贡献分配权重。对每一单项预测模型各期分配到的权重进行求和,用求得的各单项模型的权重之和除以 18 得到各单项预测模型对组合预测模型所做出的贡献。赋予 ARIMA 模型、Holt-Winters 模型以及多元回归模型的权重分别为 $W_1 = 0.332, W_2 = 0.244, W_3 = 0.424$ 来预测未来 4 年我国居民消费水平(表 7)。计算公式为

$$\hat{x}_t = 0.332x_{t\text{-ARIMA}} + 0.244x_{t\text{-HW}} + 0.424x_{t\text{-多元}} \quad (t = 19, 20, \dots, 22)$$

表 6 各项模型误差比较

Table 6 Error comparison of each model

误差度量/%		ARIMA 模型	H-W 无季节模型	多元回归模型	组合预测模型
平方和误差	F_{SSE}	2 919 915.59	12 823 660.040	6 386 950.461	1 148 143.525
归一化		22.77	100.00	49.81	8.95
平均绝对误差	F_{MAE}	314.837	708.731	380.977	181.638
归一化		44.42	100.00	53.75	25.63
平均相对误差	F_{MRE}	0.029	0.056	0.030	0.015
归一化		50.53	100.00	53.66	26.25
均方根误差	F_{RMSE}	402.762	844.053	595.677	252.558
归一化		47.72	100.00	70.57	29.92
均方根相对误差	F_{RMSRE}	0.989	0.065	0.045	0.019
归一化		100.00	6.56	4.52	1.97
均方百分比误差	F_{MSPE}	0.233	0.015	0.011	0.005
归一化		100.00	6.56	4.52	1.97
平均精度	F_{MA}	97.15	94.35	96.97	98.52

表7 各模型未来4年的预测值
Table 7 The predicted values of each model
in the next four years 单位:元

年份	ARIMA 模型	H-W无 季节模型	多元回 归模型	组合预 测模型
2020	26 393	28 360	26 974	27 119
2021	27 848	29 158	24 732	26 846
2022	30 102	29 955	22 627	26 897
2023	31 814	30 753	19 727	26 430

从预测结果来看,未来4年我国居民消费水平不会有大幅度的波动,将维持在26 000~27 000元的水平。

3 结 语

文章在开始处介绍了构建计算有序加权平均IOWA算子组合预测模型的相关理论内容,接着采用ARIMA(2,1,1)模型、Holt-Winters无季节模型和多元线性模型构建(IOWA)组合预测模型对未来4年的居民消费水平进行预测。其中多元线性回归模型把解释变量分为内循环变量与外循环变量,分别选择人均可支配收入和实际利用外商直接投资金额,控制变量选取的是国内消费税和恩格尔系数,用这4个变量来预测内循环与外循环对我国居民消费水平的冲击程度。

所构建的IOWA算子组合预测模型的步骤:首先,以各单个模型在每一年中的预测精度作为诱导因子,按照每年各单个模型的预测精度由大到小的顺序进行排序;其次,依据每一时期各单项预测模型在组合预测模型中的贡献分配权重,对每一单项预测模型各期分配到的权重进行求和,用上面求得各单项模型的权重之和简单平均得到各单项预测模型对组合预测模型的贡献;最后,基于上一步中求得的权重建立组合预测模型,并对该模型进行评价。

最后研究结果表明:IOWA算子组合预测模型是文章提出的4种预测模型中具有最高精度、最低误差的模型。未来4年我国居民消费水平不会有大幅度波动,但有小幅度下降的趋势。未来4年我国居民消费的预测值分别是27119.279,26 846.191,26 896.611,26 430.128(单位:元),变化幅度分别

为-1.610%,-1.007%,0.188%,-1.734%。我国居民人均消费有微微下降的趋势,造成这种现象的原因可能是逆全球化的趋势,也可能是受到全球新冠疫情的影响。具体造成我国居民消费水平不增反而微微下降的原因有待进一步的深入研究。消费是推动经济发展的重要力量,应积极推进消费的正常增长,完善经济持续发展计划。

参考文献(References):

- [1] BATES J M, GRANGER C W J. Combination of forecasts[J]. Operation Research Quarterly, 1969, 20(4):451—468.
- [2] 孔庆凯. 平均预测法的应用条件[J]. 预测,1985(5):18—21.
KONG Qing-kai. Application conditions of average prediction method [J]. Prediction,1985(5):18—21.
- [3] 唐小我,曹长修. 递归等权组合预测方法研究[J]. 电子科技大学学报,1992(5):545—550.
TANG Xiao-wo, CAO Chang-xiu. Research on recursive equal weight combination prediction method [J]. Journal of University of Electronic Science and Technology of China,1992(5):545—550.
- [4] 唐小我,王竹,曾勇. 几种正权重组合预测方法预测误差平方和上界的比较[J]. 数量经济技术经济研究,1994(2):17—22.
TANG Xiao-wo, WANG Zhu, ZENG Yong. Comparison of the upper bound of the square sum of prediction error of several positive weight combination forecasting methods[J]. Journal of Quantitative and Technical Economics,1994(2):17—22.
- [5] 陈华友,蔡正高. 诱导有序加权平均的组合预测模型及其应用[J]. 安徽大学学报(自然科学版),2005(1):1—6.
CHEN Hua-you, CAI Zheng-gao. Induced, ordered weighted average combined prediction model and its application [J]. Journal of Anhui University (Natural Science Edition),2005(1):1—6.
- [6] 宋峰,杨桂元. 安徽省城镇居民人均消费的组合预测[J]. 皖西学院学报,2011,27(2):10—12.
SONG Feng, YANG Gui-yuan. Study on the consumption of urban residents in anhui province[J]. Journal of West Anhui University,2011,27(2):10—12.
- [7] 侯甜甜,杨淑宁. 我国消费价格指数的走势预测及影

- 响因素分析[J]. 统计理论与实践, 2020(4):34—41.
HOU Tian-tian, YANG Shu-ning. The trend prediction of China's consumer price index and the analysis of influencing factors [J]. Statistical Theory & Practice, 2020(4):34—41.
- [8] 张甜瑞. 陕西省居民消费价格指数预测[J]. 合作经济与科技, 2020(7):67—69.
ZHANG Tian-rui. Consumer price index forecast of Shanxi province [J]. Cooperative Economy and Technology, 2020(7):67—69.
- [9] 贾康, 刘薇. 加快形成新发展格局的政策建议[N]. 经济日报, 2020-12-18:12.
JIA Kang, LIU Wei. Policy suggestions on accelerating the formation of new development pattern [N]. Economic Daily, 2020-12-18:12.
- [10] 于佳茹. 对我国居民消费水平及影响因素的实证分析[J]. 经贸实践, 2018(14):10—11.
YU Jia-ru. An empirical analysis of China's residents' consumption level and its influencing factors [J]. Economic and Trade Practice, 2018(14):10—11.

Combined Prediction Research on China's Residents Consumption Level Based on IOWA Operator

LI Ying, ZHUANG Ke-jun

(School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Anhui Bengbu 233030, China)

Abstract: In order to study the changing trend of China's residents consumption level to predict the residents consumption level in next four years (2020—2023), this paper selects the samples of China's residents consumption level in 18 years from 2002 to 2019, constructs induced ordered weighted arithmetic averaging (IOWA) operator combination forecast model based on ARIMA (2, 1, 1) model, Holt-Winters no season model and multivariate regression model, mainly uses the combined prediction optimization method with minimizing error square sum to predict the residents consumption level in next four years (2020—2023), evaluates the effectiveness of induced ordered weighted arithmetic averaging (IOWA) operator combination prediction model, searches explanatory variables and sets up multivariate regression model based on inner cycle and outer cycle of new development pattern. Research finds that the prediction accuracy of IOWA operator combination prediction model is higher than that of single prediction model and each error is significantly lower than that of the single prediction model. In the next four years, the consumption level of China's residents will not fluctuate largely but tend to decline in a small range.

Key words: resident consumption level; IOWA operator; combined prediction model

责任编辑:李翠薇

引用本文/Cite this paper:

李颖, 庄科俊. 基于 IOWA 算子的我国居民消费水平组合预测研究[J]. 重庆工商大学学报(自然科学版), 2022, 39(1): 92—100.

LI Ying, ZHUANG Ke-jun. Combined Prediction Research on China's Residents Consumption Level Based on IOWA Operator[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(1): 92—100.