第 38 卷第 2 期 重庆工商大学学报(自然科学版) 2021 年 4 月
Vol. 38 No. 2 J Chongqing Technol & Business Univ(Nat Sci Ed) Apr. 2021

doi:10.16055/j.issn.1672-058X.2021.0002.007

基于时变权重的区间时间序列组合预测模型构造*

高思凡
(安徽大学 经济学院,合肥 230601)

摘　要:时变权重组合预测模型可以有效反映各预测方法在各时刻点上的预测值对组合预测结果的影响,并可以提高预测精度,基于此,针对区间数时间序列构造组合预测模型的问题,提出一类构造区间数时变权重方法;该方法主要是在 3 种实数的时变权重求解方法基础上构造相应的 3 种区间数时变权重,并利用所得出的时变权重构造区间时间序列组合预测模型;为验证该模型的准确性,应用具体算例对模型加以实现,并通过区间预测误差度量指标验证该类模型的预测结果;结果发现:基于 3 种时变权重求解方法下的区间时间序列组合预测模型均优于各单个方法的预测结果,且基于最优化赋权法得到的时变权重区间组合预测结果的精确度较另外两种有所提高。

关键词:时变权重;区间值;组合预测
中图分类号:O212.4　　**文献标志码**:A　　**文章编号**:1672-058X(2021)02-0040-08

0 引　言

在经济社会中,对同一指标数列通常可以采取多种单个预测方法实现建模预测过程,但如果选择的预测方法不一样,往往会出现差异性的表现,其预测精度也会不尽相同。若简单的将预测误差大小作为衡量该方法是否适合该指标数列的标准,即仅仅保留预测精度高的单个预测方法,对于预测误差大、精度低的方法删除或抛弃,可能会造成有用信息的丢失。对于如何避免预测中有效信息的丢失,以及如何更加高效地吸取每种单个预测模型的优秀特质,Bates 和 Granger[1] 在 1969 年率先给出了解决方案,即通过创建组合预测模型达到充分结合每种单个预测模型所能得到的有效结果。组合预测模型是通过加权的形式将组合中的每种单个方法结合起来,以达到充分利用每种单个预测模型预测值的目的。自从组合预测的观念出现后,国内外许多学者相继提出了多种求解准则下的组合预测模型,其中 Clemen[2] 提出了以约束最小二乘方法为准则建立模型以提高预测结果的准确性;孙李红等[3] 提出以相关系数为准则并通过加权几何平均的形式建立模型。这些文献均验证了组合预测的结果具有更高的稳定性和更高的预测精度。

在组合预测中,不同的权重会带来不同的预测结果,因此权重的确定在组合预测至关重要。组合预测主要分为固定加权和变加权组合预测,方法的不同主要依赖于权重值求解方法的不同。在早期的研究中,组合预测中的权重大多为固定权重形式,也就是说,根据单个预测方法的不同给出不同的权系数,但是同一单个预测方法在所有时刻点上的权系数是相同的。然而现实生活中,同一个单个预测方法在所有时刻点上的表现不可能一直保持一致,其精度会随着时间的变化而变化,所以早期的这种固

收稿日期:2020-03-29;**修回日期**:2020-04-22.
* **基金项目**:安徽大学 2019 年大学生创新创业训练计划项目(S201910357351).
作者简介:高思凡(1996—),女,安徽合肥人,硕士研究生,从事组合预测研究.

定加权形式存在较为明显的缺陷。针对这个问题,后续的研究中逐渐开始出现了针对每种单个预测方法在每个时刻点的表现以构造变权重组合预测模型的讨论。其中张宇菲等[4]提出以面积和重心为预测精度并结合 IOWA 算子构造模糊变权的组合预测;张超等[5]提出基于 IGOELC-OWGA 算子建立最优化的组合预测模型。这些研究中,虽然利用了每种单个预测方法在每个时刻点上差异性的表现,但最终组合预测值的权重仍然是单一的数,无法反映各时刻点上的权系数值的变化特征。针对这个问题,相继出现了关于构造时刻非负权系数组合预测模型的讨论,这样就可以得到各时刻点上的每种单个预测方法的权重取值。关于此类时变权重的构造问题,相关学者提出了多种时变权重的确定方法。其中殷春武等[6]提出利用关联度和关联系数确定定权和变权的方法;张鹏等[7]在文章中提到利用预测误差平方和倒数法和二次非线性规划法确定定权和变权的方法。

近年来,关于区间数时间序列组合预测问题的讨论得到学术界的广泛关注,文献[8-11]给出了多种不同条件准则下的区间数时间序列变权组合预测模型的构造方法。但是将时变权重组合预测模型的构造方法与区间时间序列相结合的文献却鲜少出现,因此本文将在过往文献的基础上,将实数的时变权重构造方法与区间数结合起来,得到基于时变权重的区间时间序列组合预测模型并通过具体算例数据说明所提出的时变权重区间时间序列组合预测模型的准确性和可行性。

1　预备知识

1.1　区间数和区间时间序列

定义 1[3]　称 $a=[a^-,a^+]=(c;r)$ 为实轴上的一个区间数,其中 $c=(a^-+a^+)/2$ 为区间数的中点,$r=(a^+-a^-)/2$ 为区间数的半径。若 $a^-\geqslant 0$,则 a 被称为非负区间数,设所有非负区间数的集合为 Ω,其中 $a=[a^-,a^+]$,$b=[b^-,b^+]\in\Omega$ 为任意两个非负区间数,则区间数满足如下相关运算:

加法运算:$a+b=[a^-+b^-,a^++b^+]$

减法运算:$a-b=[a^--b^-,a^+-b^-]$

乘法运算:$ka=k[a^-,a^+]=[ka^-,ka^+]$

定义 2[3]　将一组按时间顺序排列的区间数 $\{[y_t^-,y_t^+]\}=\{[y_1^-,y_1^+],[y_2^-,y_2^+],\cdots,[y_n^-,y_n^+]\}$ 称为区间时间序列,其中 y_t^- 为区间最小值,称为区间下限,y_t^+ 为区间最大值,称为区间上限,$t=1,2,\cdots,n$。

1.2　区间时间序列预测模型的构造

传统的单个预测方法主要针对实数时间序列而言,针对区间值时间序列建立单个预测模型的大多文献表明,若想要对区间值数列进行预测,需要在预测之前对区间值数列进行转化。根据定义 1 中描述的区间中点和区间半径的概念,可将定义 2 中区间数时间序列拆分成区间中点序列 $\{c_t\}=(y_t^-+y_t^+)/2$ 和区间半径序列 $\{r_t\}=\{(y_t^+-y_t^-)/2\}$,由此区间数列被转换为相应的实数序列;进而对得到的两个实数序列 $\{c_t\}$ 和 $\{r_t\}$ 建立某种单个预测模型并预测,可以得到相应时刻点相应的预测结果 $\{\hat{c}_t\}$ 和 $\{\hat{r}_t\}$;最后对中点预测数列 $\{\hat{c}_t\}$ 和半径预测数列 $\{\hat{r}_t\}$ 做逆转换,即 $\hat{y}_t^-=\hat{c}_t-\hat{r}_t$ 和 $\hat{y}_t^+=\hat{c}_t+\hat{r}_t$,由此便可得到某种单个预测模型下的预测区间 $\{\hat{y}_t^-,\hat{y}_t^+\}$,其中 $\hat{y}_t^-$ 为区间预测值的最小值即下限,$\hat{y}_t^+$ 为区间预测值的最大值即上限。

1.3　时变权重组合预测模型

时变权重意味着组合预测模型每种单个预测方法在不同时刻点具有不同的权重,权重的取值不仅依赖单个预测方法的不同,同时也依赖于每种单个预测方法在每个时刻点上具体预测效果的差异。时变权重组合预测模型的具体定义如定义 3 所示:

定义 3[7]　假设某项指标时间序列 $\{Y_t,t=1,2,\cdots,n\}$ 存在 m 种单个预测方法可以对该序列实现预测,其中 y_{it} 示第 i 种单个预测方法在第 t 时刻的预测结果,每种单个预测方法在第 t 时刻的权系数可以记为向量 $\boldsymbol{w}_t=(w_{1t},w_{2t},\cdots,w_{mt})^{\mathrm{T}}$,且权重满足 $\sum_{i=1}^{m}w_{it}=1$,$w_{it}\geqslant 0$,$i=1,2,\cdots,m$,由此可建立时变权重组合预测模型,如式(1)所示:

$$\hat{y}_t=w_{1t}\hat{y}_{1t}+w_{2t}\hat{y}_{2t}+\cdots+w_{mt}\hat{y}_{mt}\tag{1}$$

其中:$\hat{y}_t$ 表示第 t 时刻的组合预测加权后结果。由式(1)可以发现不同时刻点上的组合结果依赖于不同时刻点的权重和不同单个方法拟合结果。

1.4　实数时变权重的求解方法

组合预测时变权重的确定依赖于每种单个预测

方法在各时刻点上的表现,过往文献中构造的组合预测变权重系数虽同样考虑了该问题,但最终得到的权系数还是归结为单一的数,无法反应各时刻点上权系数的变化特征,因此针对该问题,相关学者陆续提出不同的构造时刻非负权重求解方法,目前确定此类时变权重的方法有很多,其中较为常用的确定方法有[6-7]误差平方和倒数法、关联系数法和最优化赋权法等。

1.4.1 预测误差平方和倒数法[7]

预测误差平方和倒数法是根据每种单个预测模型的预测值与原始序列的观测值之间的误差,确定预测模型的组合权值。也即预测误差较大的单个预测模型组合权值较小,预测误差较小的单个预测模型组合权值较大。预测误差平方和倒数法时变权重确定公式如(2):

$$w_{it}=e_{it}^{-1}/\sum_{i=1}^{m}e_{it}^{-1} \tag{2}$$

其中:$e_{it}=(y_t-y_{it})^2$,表示第 i 种单个预测模型在第 t 时刻的误差平方和。显然,e_{it}越大,则其倒数 e_{it}^{-1}越小,从而在 $\sum_{i=1}^{m}e_{it}^{-1}$ 中所占的比例越小,则相应地在组合中所占的权重随之变小;反之,若 e_{it}越小,则其倒数 e_{it}^{-1}越大,从而在各 $\sum_{i=1}^{m}e_{it}^{-1}$ 中所占的比例越大,则相应地在组合中所占的权重随之变大。显然,由上述方法建立的每个时刻点上的权系数 $w_{it}\geqslant 0$ 且 $\sum_{i=1}^{m}w_{it}=1$。

1.4.2 关联系数法[6]

关联系数法是通过计算灰色关联度进一步确定每种单个预测方法在组合中所占的权重,即预测序列相关性低的方法在组合中所占的权系数小,反之预测序列相关性高的方法在组合中所占的权系数大。关联系数法时变权重确定公式如(3)所示:

$$w_{it}=r_{it}/\sum_{i=1}^{m}r_{it} \tag{3}$$

其中:r_{it}为各时刻点的灰色关联度,计算如公式(4)所示:

$$r_{it}=(\Delta_1+\rho\Delta_2)/(\Delta_3+\rho\Delta_2)$$
$$\begin{cases}\Delta_1=\min\limits_{i}\min\limits_{t}|y_t-\hat{y}_{it}|\\ \Delta_2=\max\limits_{i}\max\limits_{t}|y_t-\hat{y}_{it}|\\ \Delta_3=|y-\hat{y}_{it}|\end{cases} \tag{4}$$

其中:ρ 为分辨系数,$\rho\in[0,1]$,通常情况下 $\rho=0.5$;Δ_1 表示每种单个预测方法的预测结果与观测结果误差绝对值的两级最小值;Δ_2 表示每种单个预测方法的预测结果与观测结果误差绝对值的两级最大值;Δ_3 表示第 i 种单个预测方法的预测结果与观测结果之间的误差绝对值。与预测误差平方和倒数法确定的权重结果一样,根据关联系数法确定的权重同样满足所建立的组合预测模型各时刻点的权重 $w_{it}\geqslant 0$ 且 $\sum_{i=1}^{m}w_{it}=1$。

1.4.3 最优化赋权法

在组合预测模型中,权系数可以分为最优权系数和非最优权系数,其中不论是预测误差平方和倒数法还是关联系数法均为非最优权系数确定方法。通常情况下,根据最优权系数求得的组合预测结果较非最优权系数求得的结果具有更高的预测精度。因此,为了进一步提高最终的预测精度,权系数的选取应当以构造最优权系数为优先目标。通过最优化赋权法,即构造关于权系数求解的最优化模型,从而可以求得相关的最优权系数。通常情况下,最优权重的获得可以有以误差平方和最小为基准和以绝对值误差和最小为基准来求解,前者通过构造非线性公式,后者则通过构造线性公式。本文主要介绍基于绝对值误差和最小准则获得组合预测模型时变权重的方法。最优化赋权法确定时变最优权重公式如(5)所示:

$$\min S(w_t)=\left|\sum_{i=1}^{m}w_{it}e_{it}\right|=|w_{1t}e_{1t}+\cdots+w_{mt}e_{mt}|$$
$$\text{s.t}\begin{cases}\sum_{i=1}^{m}w_{it}=1\\ w_{it}\geqslant 0,i=1,2,\cdots,m\end{cases} \tag{5}$$

其中:$e_{it}=y_t-\hat{y}_{it}$为第 i 种单个预测方法的预测误差,由此得到的$\{w_{it},i=1,2,\cdots,m\}$为时变最优权重。

1.5 区间时变权重的求解方法

将上述实数时间序列的时变权系数的求解方法应用到区间数时间序列当中,在具体应用中需要考虑将区间值时间序列转换为区间中点数列$\{c_t\}$和区间半径数列$\{r_t\}$,即$\{[y_t^-,y_t^+]\}=\{[c_t,r_t]\}$,$t=1,2,\cdots,n$。

1.5.1 区间预测误差平方和倒数法

将实数序列的预测误差平方和倒数法应用到区

间数时间序列中,需要定义相应的区间预测误差平方和。定义 $e_i=(c_t-\hat{c}_i)^2+(r_t-\hat{r}_i)^2$ 为第 i 种预测区间结果相对于观测值区间的误差平方,区间的误差平方是由预测区间中点和原始区间中点之间的误差平方与预测区间半径和原始区间半径之间的误差平方构成的,该误差不仅可以度量预测区间较原始序列的位置偏离程度,同时可以度量预测区间的长度较原始序列的偏离程度。得到区间预测误差平方后,可得到相应的区间预测误差平方和倒数法确定权重,如式(6)所示:

$$w_{it}=e_{it}^{-1}/\sum_{i=1}^{m}e_{it}^{-1} \tag{6}$$

其中:$e_{it}=(c_t-\hat{c}_{it})^2+(r_t-\hat{r}_{it})^2$ 为第 i 种单个预测方法在第 t 时刻的区间预测误差平方。

1.5.2　区间关联系数法

在确定区间关联系数前,需要给出相应的区间关联度,文献[10]给出了区间灰色关联度的定义,其利用观测值区间数与预测值区间数之间的差异距离代替实值序列中观测值与预测值之间的绝对误差。根据文献[10]中的定义,利用观测值区间数与预测值区间数在第 t 时刻的差异距离 d_{it} 来替代式(4)中的 $|y_t-\hat{y}_{it}|$。区间关联度的计算如式(7)所示:

$$r_{it}=(\Delta_1+\rho\Delta_2)/(\Delta_3+\rho\Delta_2)$$
$$\begin{cases}\Delta_1=\min\limits_i\min\limits_t d_{it}\\ \Delta_2=\max\limits_i\max\limits_t d_{it}\\ \Delta_3=d_{it}\end{cases} \tag{7}$$

其中:$d_{it}=\sqrt{\dfrac{(y_{it}^{-}-y_t^{-})^2+(y_{it}^{+}-y_t^{+})^2}{2}}$ 为第 i 种单个预测区间 $[y_{it}^{-},y_{it}^{+}]$ 与观测值区间 $[y_t^{-},y_t^{+}]$ 在第 t 时刻之间的差异距离,ρ 为分辨系数,$\rho\in[0,1]$,通常情况下 $\rho=0.5$。Δ_1 表示差异距离的两级最小值;Δ_2 表示差异距离的两级最大值;Δ_3 表示差异距离。

根据得到的区间关联度,进一步可以确定区间系数法下的时变权重,如式(8)所示:

$$w_{it}=r_{it}/\sum_{i=1}^{m}r_{it} \tag{8}$$

其中 r_{it} 为各时刻点的灰色关联度。

显然,上述两种区间时变权重求解方法的组合预测权重不仅满足权重非负性要求还满足每种单个预测方法在同一时刻下权重和为 1 的要求。

1.5.3　区间最优化赋权法

在对区间数时间序列组合预测模型进行最优化赋权时,需要将区间数时间序列拆分成区间中点序列和区间半径序列,并以绝对值误差和最小为准则分别建立区间组合中点各时刻点的绝对值误差和 $C(w_t)$ 和区间半径各时刻点的绝对值误差和 $R(w_t)$,区间最优化赋权法确定最优时变权重公式如(9)所示:

$$\min C(w_t)=\left|\sum_{i=1}^{m}w_{it}e_{it}^{c}\right|$$
$$\min R(w_t)=\left|\sum_{i=1}^{m}w_{it}e_{it}^{c}\right|$$
$$\text{s.t}\begin{cases}\sum\limits_{i=1}^{m}w_{it}=1\\ w_{it}\geqslant 0,i=1,2,\cdots,m\end{cases} \tag{9}$$

其中:$e_{it}^{c}=c_t-\hat{c}_{it}$ 和 $e_{it}^{r}=r_t-r_{it}$ 分别为第 i 种单个预测方法在第 t 时刻的中点数列误差和半径数列误差。上述区间组合预测模型是一个多目标线性规划问题,不易于最终最优时变权重的求解,因此为求解该模型,需要引入偏好系数 q,这样便可实现多目标向单目标的转换,转换之后得到如下最优化模型,具体公式如(10)所示:

$$\min W(w_t)=qC(w_t^{c})+(1-q)R(w_t^{r})$$
$$\text{s.t}\begin{cases}\sum\limits_{i=1}^{m}w_{it}=1\\ w_{it}\geqslant 0,i=1,2,\cdots,m\end{cases} \tag{10}$$

其中:q 为偏好系数,$q\in[0,1]$。当 $q\in[0,0.5]$ 时,说明区间数径序列对区间预测值的重要性高于区间中点序列;若 $q\in[0.5,1]$,则说明区间中点数列对区间预测值的的重要性高于区间半径数列;一般情况下取 $q=0.5$,说明中点序列和半径序列对最终得到的结果具有同等的重要性。

2　区间时间序列预测误差度量公式

为检验所提及的基于 3 种区间时变权重求解方法建立的时变权重区间组合预测模型的准确性,根据预测效果评价的一般原则并针对区间值时间序

列,选取如下区间时间序列预测误差度量指标体系,对最终的每种区间单个预测结果和每种区间组合预测结果的准确性进行衡量[11]:

平均区间位置误差平方和:

$$F_{\mathrm{MSEP}} = \sum_{t=1}^{n} (c_t - \hat{c}_t)^2 / n$$

平均区间长度误差平方和:

$$F_{\mathrm{MSEL}} = \sum_{t=1}^{n} (r_t - \hat{r}_t)^2 / n$$

平均区间误差平方和:

$$F_{\mathrm{MSEI}} = F_{\mathrm{MSEP}} + F_{\mathrm{MSEL}}$$

平均区间相对误差和:

$$F_{\mathrm{MRIE}} = \sum_{t=1}^{n} \left(\frac{|c_t - \hat{c}_t|}{r_t + \hat{r}_t} \right) / n$$

3 实证分析

为了检验所提出的基于时变权重区间时间序列组合预测模型的可行性与预测结果的准确性,应用文献[8]中的数据进行具体实证数据测算。表 1 给出了观测区间和 3 种单个预测区间的具体数值。

根据 3 种单个预测方法各时刻点的拟合区间值,分别应用区间预测误差平方和倒数法、区间关联系数法和区间最优化赋权法确定组合中每种单个预测方法在每个时刻点上的权系数,并利用求出的每种单个预测方法在每个时刻点的权系数,结合 3 种单个预测模型在每个时刻点的拟合区间值,得到区间组合拟合值。其中在区间最优化赋权法中,偏好系数取值不同,将对后续的权重产生影响,从而影响最终组预测结果。因此,在具体算例中,还将进一步讨论不同的偏好系数下时变权重的数值变化情况以及最终组合预测结果的优劣性。最终求得的 3 种时变权重方法下的区间时变权重和区间组合预测具体结果如表 2 和表 3 所示,其中表 2 中最优化赋权法结果是在偏好系数为 0.5 时求得。

表 1 观测区间和 3 种单项方法预测区间

Table 1 The observation interval and three monomial methods to predict the interval

观测区间数		单个方法 1 预测区间数		单个方法 2 预测区间数		单个方法 3 预测区间数	
$[y_t^-, y_t^+]$	$[c_t, r_t]$	$[y_{1t}^-, y_{1t}^+]$	$[c_{1t}, r_{1t}]$	$[y_{2t}^-, y_{2t}^+]$	$[c_{2t}, r_{2t}]$	$[y_{3t}^-, y_{3t}^+]$	$[c_{3t}, r_{3t}]$
[3,4]	(3.5,0.5)	[2,4.5]	(3.7,1.3)	[3.6,5.4]	(4.5,0.9)	[3,3.6]	(3.3,0.3)
[5,5.6]	(5.3,0.3)	[2,2.6]	(4.1,1.9)	[5,6]	(5.5,0.5)	[4,5.2]	(4.6,0.6)
[4,6]	(5,1)	[3,8]	(5.5,2.5)	[5.2,7]	(6.1,0.9)	[4.3,5.1]	(4.7,0.4)
[6,10]	(8,2)	[4.6,11]	(7.8,3.2)	[6.8,11.6]	(9.2,2.4)	[6.1,7.3]	(6.7,0.6)
[6.6,8.8]	(7.7,1.1)	[6,12.4]	(9.2,3.2)	[8,9.6]	(8.8,0.8)	[7,8]	(7.5,0.5)
[9,11]	(10,1)	[7,15]	(11,4)	[9.6,12]	(10.8,0.2)	[9.1,9.9]	(9.5,0.4)

表 2 基于 3 种时变权重求解方法下的区间时变权重

Table 2 Interval time-varying weights based on three methods for solving time-varying weights

区间预测误差平方和倒数法			区间关联系数法			区间最优化赋权法		
单个方法 1	单个方法 2	单个方法 3	单个方法 1	单个方法 2	单个方法 3	单个方法 1	单个方法 2	单个方法 3
0.099 1	0.058 1	0.842 7	0.277 0	0.230 1	0.492 9	0.125 0	0.125 0	0.750 0
0.017 3	0.863 6	0.119 1	0.153 3	0.531 8	0.314 9	0.001 3	0.778 5	0.220 2
0.116 2	0.238 1	0.645 6	0.240 7	0.316 3	0.443 0	0.252 4	0.139 9	0.607 7
0.429 1	0.396 9	0.174 0	0.374 2	0.363 4	0.262 4	0.256 6	0.407 1	0.336 3
0.043 9	0.225 0	0.731 1	0.172 6	0.332 8	0.494 6	0.218 1	0.037 2	0.744 7
0.031 2	0.458 1	0.510 7	0.145 2	0.419 7	0.435 0	0.109 2	0.258 6	0.632 2

表 3　3 种区间时变权重下的区间组合预测拟合值

Table 3　Interval combination prediction fitting values under three kinds of interval time-varying weights

区间预测误差平方和倒数法		区间关联系数法		区间最优化赋权法	
$[y_{4t}^-, y_{4t}^+]$	$[c_{4t}, r_{4t}]$	$[y_{5t}^-, y_{5t}^+]$	$[c_{5t}, r_{5t}]$	$[y_{6t}^-, y_{6t}^+]$	$[c_{6t}, r_{6t}]$
[2.98,3.84]	(4.41,0.43)	[2.97,4.40]	(3.69,0.72)	[3.00,4.00]	(3.50,0.50)
[4.83,5.90]	(5.37,0.54)	[4.26,5.75]	(5.00,0.75)	[4.78,5.82]	(5.30,0.52)
[4.36,5.89]	(5.13,0.76)	[4.27,6.40]	(5.34,1.06)	[4.10,6.10]	(5.10,1.00)
[5.73,10.59]	(8.16,2.43)	[5.79,10.25]	(8.02,2.23)	[6.00,10.00]	(8.00,2.00)
[7.18,8.55]	(7.87,0.69)	[7.16,9.29]	(8.23,1.07)	[6.82,9.02]	(7.92,1.10)
[9.26,11.02]	(10.14,0.88)	[9.00,11.52]	(10.26,1.26)	[9.00,11.00]	(10.00,1.00)

根据表 2 中的结果可以发现,3 种时变权重求解方法下的区间时变权重在每种单个预测方法上具有相同的变化趋势。如在 $t=2$ 时刻,根据区间预测误差平方和倒数法求得的每种单个预测方法的权系数,单个方法 2 具有最大的权重,其次为单个方法 3,单个方法 1 的权重最小。对比区间关联系数法和区间最优化赋权法在该时刻的每种单个预测模型的权重可以发现,同样是单个方法 2 具有最大的权系数值,单个方法 3 的权系数值次之,单个方法 1 的权系数值最小。由此可以反向证明 3 种区间时变权重求解方法的准确性。进一步对比发现,相较于另外两种方法下的时变权重,区间关联系数法确定的各单个预测模型的区间时变权重之间差异较小,不能很好地区分各时刻点每种单个预测结果所能反映的有效信息,可能进一步影响最后的组合预测效果。

为检验所建立模型是否有效,利用区间时间序列预测误差度量指标对 3 种单个方法以及各时变权重求解方法下的预测结果进行分析,表 4 给出了每种单个方法以及每种组合时变权系数求解方法下的区间预测准确性测度结果。

表 4　不同时变权重求解方法下的预测有效度比较

Table 4　Comparison of prediction validity under different time-varying weighting methods

方法	单个预测方法 1	单个预测方法 2	单个预测方法 3	区间预测误差平方倒数法	区间关联系数法	区间最优化赋权法
F_{MESP}	0.863 7	0.923 3	0.433 3	0.017 3	0.097 1	0.009 6
F_{MSEL}	3.383 3	0.083 3	0.528 3	0.081 2	0.061 5	0.008 3
F_{MSEI}	4.220 0	1.006 7	0.961 7	0.098 6	0.158 6	0.017 9
F_{MRIE}	0.231 1	0.459 8	0.370 7	0.076 2	0.160 9	0.024 7

从表 4 中各度量指标的的结果不难发现,基于 3 种时变权重求解方法下的区间组合预测结果的各误差度量值均小于每种单个方法,证明了该类模型的预测准确性。通过进一步比较 3 种时变权重求解方法下的区间组合预测结果的预测误差,可以发现,基于区间关联系数法构造时变权重得到的区间组合预测结果较另外两种方法下预测结果的各项误差评价指标均较大,验证了依据表 2 中每种单个预测方法的时变权系数结果对于最后的区间组合预测效果影响的猜想。即在组合预测中,若不同的单个预测方法在各时刻点权重的差异不明显,将无法很好地区分和利用每种单个预测方法的有效信息,从而影响最后的组合预测结果。另外,还可以发现,基于区间最优化赋权法构造时变权重得到的区间组合预测结果较另外两种区间组合预测结果,预测精度有所提高,进一步证明了基于最优化赋权法得到的最优时变权重后的组合预测结果较基于非最优化赋权求得的结果具有更小的预测误差。

由上述结果可以发现,区间最优化赋权法最终得到的结果是所有预测方法中最优的,但不同的偏好系数将影响在此方法下得到的权重,进一步影响最终的预测效果。因此,为追求更加精准的预测结果,将设置不同的偏好系数值并得到相应的预测结果效度评价结果,具体结果如表 5 所示。

表 5 不同偏好系数下区间最优化赋权法组合预测有效度比较

Table 5 Comparison of the validity of the combined forecasting method of interval optimization weighting method under different preference coefficients

	$q=0.1$	$q=0.2$	$q=0.3$	$q=0.4$	$q=0.5$	$q=0.6$	$q=0.7$	$q=0.8$	$q=0.9$
F_{MESP}	0.060 9	0.045 2	0.014 6	0.027 8	0.009 6	0.005 7	0.000 0	0.000 0	0.000 0
F_{MSEL}	0.008 2	0.008 3	0.008 3	0.008 3	0.008 3	0.008 9	0.026 2	0.029 3	0.034 8
F_{MSEI}	0.069 1	0.053 5	0.022 9	0.036 1	0.017 9	0.014 6	0.026 2	0.029 3	0.034 8
F_{MRIE}	0.065 9	0.053 8	0.032 3	0.040 0	0.024 7	0.014 0	0.000 0	0.000 0	0.000 0

由表 5 中的结果首先可以发现,不同偏好系数下结果的有效度除 $q=0.1$ 和 $q=0.2$ 的 F_{MSEP} 较大外,其余结果较其余预测方法精度更高;其次,比较不同偏好系数下的评价结果发现,就 F_{MRIE} 而言,q 取大于 0.5 的值时,结果更优,但总体而言各个结果之间精度相差不大。

4 结 论

主要讨论了基于时变权重的区间值时间序列组合预测问题,利用 3 种实数时变权重求解方法构造相应的区间数时变权重的求解方法。为检验通过该方法得到的组合模型的准确性,通过具体算例求出基于 3 种时变权重求解方法下的每种单个预测模型在每个时刻点上的权系数,并利用区间值时间序列预测误差指标计算并度量最终的区间组合预测结果的准确性。比较各项误差度量结果可知,基于时变权重构造的区间组合预测模型较单个预测方法的误差度量结果有非常明显地降低;并且通过比较 3 种时变权重求解方法下的区间组合预测值的准确性发现:首先是基于区间关联系数法构造时变权系数得到的结果较另外两种方法下得到的结果,各项误差指标度量值均较大,进一步验证并说明了若单个预测模型在各时刻点权重的差异不明显,将无法很好地区分和利用单个预测方法在每个时刻点提供的有效信息,从而影响最终的组合预测结果;其次是基于最优化赋权法构造最优时变权重的区间组合预测模型结果较非最优化赋权法得到非最优时变权重构造的模型结果具有更小的预测误差和更高的准确度;最终,通过以上结论可以进一步说明基于时变权重的区间时间序列组合预测模型的可行性和准确性。因此,提出的方法建立的组合模型不仅可以很好地反映各时刻点上权重的变化特征而且可以提高预测结果的准确性。

参考文献(References):

[1] BATES J M, GRANGER C W J. The Combination of Forecasts[J]. Journal of the Operational Research Society,1969,20(4):451—468

[2] CLEMEN R T. Liner Constrains and the Efficiency of Combined Forecasts [J]. Journal of Forecasting, 1986, 5(1):31—38

[3] 孙李红,沈继红. 基于相关系数的加权几何平均组合预测模型的性质[J]. 系统工程理论与实践,2009, 29(9):84—91

SUN L H, SHEN J H. Properties of Weighted Geometric Mean Combination Forecasting Model Based on Correlation Coefficient [J]. System Engineering Theory and Practice,2009,29 (9):84—91(in Chinese)

[4] 张宇菲,陈华友. 基于面积和重心预测精度的 IOWA 算子的模糊变权组合预测模型[J]. 统计与决策, 2018,34(14):24—28

ZHANG Y F, CHEN H Y. Fuzzy Variable Weight Combination Prediction Model of IOWA Operator Based on Area and Center of Gravity Prediction Accuracy[J]. Statistics and Decision, 2018, 34(14): 24—28 (in Chinese)

[5] 张超,袁宏俊. 基于 IGOWLC-OWGA 算子最优组合预测模型[J]. 统计与决策,2019,35(8):73—76

ZHANG C, YUAN H J. An Optimal Combination Forecasting Model Based on IGOWLC-OWGA Operators[J]. Statistics and Decision, 2019, 35(8): 73—76 (in Chinese)

[6] 殷春武. 基于灰色关联度的第三产业发展趋势组合预测模型[J]. 统计与决策,2013(13):17—20

YIN C W. Combination Forecasting Model for the Development Trend of the Tertiary Industry Based on Grey Correlation[J]. Statistics and Decision,2013 (13):17—20(in Chinese)

[7] 张鹏. 组合预测中变权与定权的应用比较[J]. 统计与决策,2018,34(17):80—82

ZHANG P. Comparison of the Application of Variable Weight and Fixed Weight in Combination Forecasting[J].

Statistics and Decision, 2018, 34 (17): 80—82 (in Chinese)

[8] 袁宏俊,韦晨珺娃,钟梅. 基于联系数贴近度的区间型组合预测模型及其有效性[J]. 统计与信息论坛, 2017,32(6):31—37
YUAN H J, WEI C X W, ZHONG M. Interval-type Combined Forecasting Model Based on Contact Number Closeness and Its Effectiveness[J]. Statistics and Information Forum,2017,32 (6):31—37(in Chinese)

[9] 吴潜. 基于中点及半径的区间数组合预测模型性质研究[J]. 统计与决策,2017(4):36—40
WU Q. Study on the Properties of Interval Number Combination Prediction Model Based on Midpoint and Radius[J]. Statistics and Decision,2017 (4):36—40(in Chinese)

[10] 胡纪纲,芮源,袁宏俊. 基于区间关联度的 IOWGA 算子的区间组合预测[J]. 统计与决策,2016(12):19—22
HUJ G, RUI Y, YUAN H J. Interval Combination Forecasting of IOWGA Operators Based on Interval Correlation Degree[J]. Statistics and Decision,2016 (12): 19—22(in Chinese)

[11] 陶志富,刘金培,朱家明,等. 区间值时间序列预测效果测度研究[J]. 模糊系统与数学, 2018, 32(4): 135—144
TAO Z F, LIU J P, ZHU J M, et al. Research on the Measurement of Interval Value Time Series Prediction Effect[J]. Fuzzy Systems and Mathematics,2018,32(4): 135—144(in Chinese)

Construction of Interval Time Series Combined Prediction Model Based on Time-varying Weight

GAO Si-fan

(School of Economics, Anhui University, Hefei 230601, China)

Abstract: For the current time-varying weighted combination prediction model, it can effectively reflect the influence of the prediction value of each prediction method at each time point on the combination prediction result, and it can effectively improve the accuracy of the prediction, as well as based on the problem of constructing the combination prediction model of the interval number time series, a new method for constructing time-varying weights of interval numbers is proposed. This method is mainly based on three real-time time-varying weights solving methods to construct the corresponding three time-varying interval weights, and uses the obtained time-varying weights to construct interval time series combined prediction models. In order to verify the accuracy of the model, the model is implemented through specific examples and the prediction results of this type of model are measured by interval prediction error metrics. It was found that the interval time series combination prediction models based on three time-varying weights solving methods were better than the prediction results of each single method, and the accuracy of the time-varying weight interval combination prediction results obtained based on the optimization weighting method was better than the other two.

Key words: time-varying weight; interval value; combined prediction

责任编辑:李翠薇

引用本文/Cite this paper:
高思凡. 基于时变权重的区间时间序列组合预测模型构造[J]. 重庆工商大学学报(自然科学版),2021,38(2):40—47
GAO S F. Construction of Interval Time Series Combined Prediction Model Based on Time-varying Weight [J]. Journal of Chongqing Technology and Business University (Natural Science Edition),2021,38(2):40—47