

基于 COWA 算子的相关性区间组合预测优化模型及其近似求解

周代娣¹, 陈华友¹, 张倩昀², 江立辉^{1,3}

(1. 安徽大学数学科学学院, 安徽合肥, 230601;

2. 安徽大学经济学院, 安徽合肥, 230601; 3. 合肥学院人工智能与大数据学院, 安徽合肥 230601)

摘要: 在 Holt 指数平滑模型、多层感知器 (MLP) 模型及支持向量机 (SVM) 模型三种区间预测方法的基础上, 通过引进 COWA 算子和相关系数的概念, 构建基于相关系数的区间型最优组合预测模型, 探讨了模型的若干性质. 同时利用相关系数作为测度单点预测方法在组合预测中的贡献指标, 结合作对策略中的 Shapley 值, 给出相应的最优组合预测模型的近似求解. 利用西德克萨斯中质原油现货价格数据验证了模型的可行性与有效性, 并分析了 COWA 算子中的态度参数的灵敏度.

关键词: 区间预测; Holt 指数平滑模型; MLP 模型; SVM 模型; COWA 算子; 相关系数; Shapley 值
中图分类号: O224 **文献标识码:** A doi: 10.3969/j.issn.0253-2778.2020.05.015

引用格式: 周代娣, 陈华友, 张倩昀, 等. 基于 COWA 算子的相关性区间组合预测优化模型及其近似求解[J]. 中国科学技术大学学报, 2020, 50(5): 673-681.

ZHOU Daidi, CHEN Huayou, ZHANG Qianyun, et al. Optimization model of correlation interval combination forecasting based on COWA operator and its approximate solutions[J]. Journal of University of Science and Technology of China, 2020, 50(5): 673-681.

Optimization model of correlation interval combination forecasting based on COWA operator and its approximate solutions

ZHOU Daidi¹, CHEN Huayou¹, ZHANG Qianyun², JIANG Lihui^{1,3}

(1. School of Mathematical Sciences, Anhui University, Hefei 230601, China;

2. School of Economics, Anhui University, Hefei 230601, China;

3. School of Artificial Intelligence and Big Data, Hefei University, Hefei 230601, China)

Abstract: On the basis of Holt exponential smoothing model, multi-layer perceptron (MLP) model and support vector machine (SVM) model, an interval optimal combination forecasting model was built by introducing the concept of COWA operator and correlation coefficient and some properties of the model were discussed. At the same time, the correlation coefficient was used as the contribution index of the single point forecasting method in the combined forecasting, and the Shapley value in the cooperative game was combined to give the approximate solution of the corresponding optimal combined forecasting model. The data of WTI spot price were used to demonstrate the feasibility and effectiveness of the model, and the attitude parameters on COWA operator was done with the sensitivity analysis.

Key words: interval prediction; Holt's exponential smoothing; MLP model; SVM model; COWA operator; correlation coefficient; Shapley value

收稿日期: 2019-06-10; 修回日期: 2019-07-14

基金项目: 国家自然科学基金(71871001, 71771001), 安徽省高校人文社科重点项目(SK2018A0605)资助.

作者简介: 周代娣, 女, 1996年生, 硕士. 研究方向: 预测与决策. E-mail: 2723754503@qq.com

通讯作者: 陈华友, 博士/教授. E-mail: huayouc@126.com

0 引言

在预测实践中,预测对象往往是复杂的系统,其发展趋势具有不同的特征,因此我们经常采用不同的单点预测方法对其进行预测.例如,指数平滑法实际上是一类加权移动平均预测方法.它是对过去观测结果进行加权平均组合,并分配了一个递减的权重.权重的大小体现了近期的观测结果包含的信息量大,应赋予相对较大权重.指数平滑法具有计算简单,适应性好的特点.其常见的一种扩展是 Holt 指数平滑模型,它带有一个额外的趋势分量,具有 2 个基本平滑公式和 1 个预测公式,模型简单但综合性强,在区间值时间序列预测和油井产量动态预测中得到广泛的应用^[1,2].人工神经网络具有处理线性问题或非线性问题^[3]的能力.应用最广泛的是具有两层前馈 MLP 网络,与最简单的神经网络单层感知器不同,它不仅由输入层和输出层组成,而且还增加了隐藏层.文献[4]和[5]分别利用多层感知器模型预测了物流总额和太阳能温室装置淡化水的产量.支持向量机是一类机器学习方法,它与传统机器学习理论不同之处在于基于结构风险最小化原理而非经验最小化原理^[6],常用于分类、回归、非线性函数逼近等问题的求解.文献[7]和[8]分别探讨了支持向量机在城市空气质量预警系统和中国高考成绩预测分析中的应用.

考虑到不同的单点预测方法能够从不同的角度反映预测对象的发展趋势.因此,自 Bates 和 Granger^[9]提出组合预测概念以来,组合预测方法受到广大预测学者的青睐,并取得了一些研究成果^[10-14].理论研究和预测实践表明组合预测能综合利用多个单点预测模型的结果,从而提高了预测精度,使得预测结果更为可靠.但组合预测最为关键的问题就是如何分配各单点预测方法的权重.

对于区间数据,文献[15]引进了区间组合预测.文献[16]则给出了区间数的运算法则.由于区间数的运算法则具有其本身的特点,如果利用传统组合预测方法进行处理,分别构建模型区间数的上下界的组合预测方法,不仅计算复杂,而且有可能在未来外推的预测区间产生上下界错位现象.因此,文献[17]提出了连续有序加权平均(continuous ordered weighted averaging, COWA)算子,将区间型数据集结成实数型数据进行考虑,降低了模糊性.

本文转换研究思路,采用 Holt 指数平滑模型、多层感知器(MLP)及支持向量机(SVM)三种区间

预测方法,利用 COWA 算子将区间型数据转化为实数型数据,同时引入相关系数作为测度单点预测方法在组合预测中的精度指标,构建基于相关系数的区间型最优组合预测模型,并探讨了模型解的若干性质.考虑到所建立的组合预测模型是一个非线性规划问题,求解相对复杂,文中利用合作对策理论中的 Shapley 值法来确定权重,给出相应的最优组合预测模型的近似求解,原油现货价格的预测实例验证了模型的可行性与有效性.

1 基本概念

定义 1.1^[16] 若 $a = [a^L, a^U] = \{x \mid a^L < x < a^U, a^L, a^U \in \mathbb{R}\}$, 则称 a 为一个区间数,其中 a^L , a^U 分别称为 a 的下界和上界.

特别地,若 $a^L = a^U$, 则 a 退化成一个实数.若 $0 < a^L < a^U$, 则称 a 为正区间数.

给定两个正区间数 $a = [a^L, a^U]$, $b = [b^L, b^U]$, $\lambda \in \mathbb{R}$, 则有运算法则如下:

$$\textcircled{1} \text{加法运算: } a + b = [a^L + b^L, a^U + b^U];$$

$$\textcircled{2} \text{减法运算: } a - b = [a^L - b^U, a^U - b^L].$$

$$\textcircled{3} \text{数乘运算: } \lambda a = \begin{cases} [\lambda a^L, \lambda a^U], & \lambda \geq 0; \\ [\lambda a^U, \lambda a^L], & \lambda < 0; \end{cases}$$

记 $m = (a^L + a^U) / 2$, $r = (a^U - a^L) / 2$, 则称 m 为区间的中心, r 为区间的半径.显然, $a^L = m - r$, $a^U = m + r$. 区间数 a 也可记作 $a = [a^L, a^U] = (m; r)$.

定义 1.2^[17] 设函数 $Q: [0, 1] \rightarrow [0, 1]$ 为基本单位区间单调(BUM)增函数,满足: $Q(0) = 0$, $Q(1) = 1$. 设 $[m, n]$ 为一区间数,若

$$F_Q([m, n]) = \int_0^1 \frac{dQ(x)}{dx} (n - x(n - m)) dx \quad (1)$$

称 F_Q 为 COWA 算子, F_Q 也可记为 COWA.

令 $\alpha = 1 - \int_0^1 x(dQ(x)/dx) dx$, 称 α 为 BUM 函数 Q 的态度参数,则式(1)可等价表示成

$$F_Q([m, n]) = COWA([m, n]) = (1 - \alpha)m + \alpha n \quad (2)$$

由式(2)知:COWA 算子 F_Q 通过态度参数 α 将区间数转化成实数,简化计算并降低了模糊性.

2 区间型预测模型

2.1 区间型 Holt 指数平滑模型

Holt 指数平滑模型对趋势直接进行平滑并对

原时间序列进行预测,灵活性大.需要考虑的是两个平滑参数以及初值的选取问题.

设某种指标序列的实际观察区间值为 $x_t = [x_t^L, x_t^U], t = 1, 2, \dots, N$. 设 α, β 为平滑系数, 满足 $0 < \alpha, \beta < 1, s_t = [s_t^L, s_t^U]$ 为第 t 期的区间修正平滑值, $b_t = [b_t^L, b_t^U]$ 为第 t 期的区间增量值, $t = 1, 2, \dots, N$.

为表述上的方便,以区间值的左右端点构成的二维向量仍然用原有的区间值记号表达,即 $x_t = [x_t^L, x_t^U], s_t = [s_t^L, s_t^U], b_t = [b_t^L, b_t^U]$ 均可以看成是一个二维的行向量, Holt 指数平滑模型的步骤如下:

Step 1 初始值设置. 本文设定初始值为

$$s_2 = [x_2^L, x_2^U], b_2 = [x_2^L - x_1^L, x_2^U - x_1^U].$$

Step 2 根据平滑公式计算区间修正平滑值与区间增量值,平滑公式如下:

$$\left. \begin{aligned} s_t^T &= Ax_t^T + (I - A)(s_{t-1}^T + b_{t-1}^T), \\ b_t^T &= B(s_t^T - s_{t-1}^T) + (I - B)b_{t-1}^T \end{aligned} \right\} \quad (3)$$

式中, T 为转置, I 为单位矩阵; A, B 为 2×2 阶平滑参数矩阵, 控制曲线的平滑度^[1]. 假定

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

其中 $0 \leq a_{ij}, b_{ij} \leq 1$. 则式(3)可进一步写成:

$$\begin{aligned} s_t^T &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_t^L \\ x_t^U \end{bmatrix} + \begin{bmatrix} 1 - a_{11} & -a_{12} \\ -a_{21} & 1 - a_{22} \end{bmatrix} \begin{bmatrix} (s_{t-1}^L + b_{t-1}^L) \\ (s_{t-1}^U + b_{t-1}^U) \end{bmatrix} = \\ & \begin{bmatrix} a_{11}x_t^L + (1 - a_{11})(s_{t-1}^L + b_{t-1}^L) + a_{12}(x_t^U - s_{t-1}^U - b_{t-1}^U) \\ a_{22}x_t^U + (1 - a_{22})(s_{t-1}^U + b_{t-1}^U) + a_{21}(x_t^L - s_{t-1}^L - b_{t-1}^L) \end{bmatrix}, \\ b_t^T &= \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} s_t^L - s_{t-1}^L \\ s_t^U - s_{t-1}^U \end{bmatrix} + \begin{bmatrix} 1 - b_{11} & -b_{12} \\ -b_{21} & 1 - b_{22} \end{bmatrix} \begin{bmatrix} b_{t-1}^L \\ b_{t-1}^U \end{bmatrix} = \\ & \begin{bmatrix} b_{11}(s_t^L - s_{t-1}^L) + (1 - b_{11})b_{t-1}^L + b_{12}(s_t^U - s_{t-1}^U - b_{t-1}^U) \\ b_{22}(s_t^U - s_{t-1}^U) + (1 - b_{22})b_{t-1}^U + b_{21}(s_t^L - s_{t-1}^L - b_{t-1}^L) \end{bmatrix} \end{aligned} \quad (4)$$

Step 3 以一步向前预测误差平方和达到最小原则求解参数. 定义区间型一步向前预测误差平方和为 e^2 :

$$\begin{aligned} e^2 &= \sum_{t=3}^N (x_t - \hat{x}_t)(x_t - \hat{x}_t)^T = \\ & \sum_{t=3}^N (x_t^L - s_{t-1}^L - b_{t-1}^L, x_t^U - s_{t-1}^U - b_{t-1}^U) \cdot \\ & \begin{pmatrix} x_t^L - s_{t-1}^L - b_{t-1}^L \\ x_t^U - s_{t-1}^U - b_{t-1}^U \end{pmatrix} = \\ & \sum_{t=3}^N (x_t^L - s_{t-1}^L - b_{t-1}^L)^2 + \sum_{t=3}^N (x_t^U - s_{t-1}^U - b_{t-1}^U)^2 \end{aligned} \quad (5)$$

故求解模型:

$$\begin{aligned} \min e^2 &= \sum_{t=3}^N (x_t^L - s_{t-1}^L - b_{t-1}^L)^2 + \\ & \sum_{t=3}^N (x_t^U - s_{t-1}^U - b_{t-1}^U)^2 \\ \text{s. t. } & \begin{cases} 0 \leq a_{ij}, b_{ij} \leq 1; \\ s_t^L + b_t^L \leq s_t^U + b_t^U, t = 3, 4, \dots, N \end{cases} \end{aligned} \quad (6)$$

Step 4 重复步骤 2, 3 直至预测结束. 利用 Holt 指数平滑模型, 向前 I 期的预测值 $\hat{x}_{t+I} = s_t +$

$b_t I, I$ 为外推预测期数.

2.2 区间型多层感知器(MLP)模型

多层感知器的原理是将输入的多个数据集映射到单一的输出的数据集上,其组成结构为感知单元组成输入层、多层计算节点的隐藏层、一层计算节点的输出层,是一种前馈神经网络模型.

在多层感知器模型中,输出值 x_t 与输入值 $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ 之间的关系如下:

$$x_t = w_0 + \sum_{j=1}^q w_j \cdot f(w_{0j} + \sum_{i=1}^p w_{ij} x_{t-i}) \quad (7)$$

式中, w_{0j}, w_0 分别表示输入层到隐藏层、隐藏层到输出层的偏置值, w_{ij} 表示输入层的第 i 个节点到隐藏层的第 j 个节点的权重, w_j 表示隐藏层到输出层的权重, p 为输入节点的个数, q 为隐藏节点的个数, f 为隐藏层的转换函数,其作用是将非线性引入到神经元的输出. 本文取 logistic 函数定义的 Sigmoid 非线性函数, $f(u) = 1/(1 + \exp(-u))$.

对于区间型数据,则构建含有两个前馈层的神经网络,其中输入节点数为 $2p$, 隐藏节点数为 q 以及 2 个输出节点,每个输出分别对应着预测区间的中心和半径,然后利用区间的中心和区间的半径来

表达形式,将其转化为区间的左右端点.

2.3 区间型 WOA-SVM 模型

支持向量机(SVM)是一类机器学习方法,与神经网络不同,它基于结构风险最小化,通过非线性映射对高维特征空间中的数据进行映射,并在空间中进行线性回归^[18]. 由于其在求解非线性问题方面的优良性能,被广泛应用于预测领域.

设输入向量 $x = (x_1, x_2, \dots, x_N)$, 将 x 通过 $\varphi: R^n \rightarrow H$ 映射到高维特征空间 H 中, 则回归方程为

$$f(x) = w \cdot \varphi(x) + b \quad (8)$$

式中, $w = (w_1, w_2, \dots, w_N)$ 为权向量, 符号 ' \cdot ' 为点乘, b 为误差项, $\varphi(x)$ 就是将样本数据映射到高维空间的映射函数.

对于非线性问题, 求解权向量和误差项的目标函数式可表示成如下数学规划问题:

$$\begin{aligned} \min Z = & \frac{1}{2} \|w\|^2 + C \left(\sum_{i=1}^N (\xi_i + \xi_i^*) \right) \\ \text{s. t. } & \begin{cases} y_i - (w \cdot \varphi(x) + b) < \varepsilon + \xi_i, \\ (w \cdot \varphi(x) + b) - y_i < \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, N, C > 0 \end{cases} \end{aligned} \quad (9)$$

式中, ε 为允许误差, C 为平衡模型复杂度和误差项的惩罚系数, ξ_i, ξ_i^* 为松弛变量, 当预测存在误差时都大于 0, 否则等于 0, y_i 为输入向量 x_i 的输出.

由于式(8)为非线性回归问题, 因此考虑引用核函数 $k(x_i, x)$. 本文取核函数为径向基函数 RBF, 即 $k(x_i, x) = \exp(-\|x - x_i\| / (2\sigma^2))$. 利用拉格朗日方法将式(9)转换成如下对偶问题(其中 α_i, α_i^* 为拉格朗日乘子):

$$\begin{aligned} \max L = & -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(x_i, x_j) + \\ & \sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) - \sum_{i=1}^N \varepsilon (\alpha_i + \alpha_i^*) \\ \text{s. t. } & \begin{cases} \sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0, \\ 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \dots, N \end{cases} \end{aligned} \quad (10)$$

则式(8)可表示成

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (11)$$

所以, 利用式(10)求出的 α, α^* 代入式(11)就可计算出预测值. 同时, 从式中可看到核函数参数 σ 和惩罚因子 C 是提高 SVM 算法泛化能力和分类性能的关键^[19]. 本文采用鲸鱼优化算法 (whale optimization algorithm, WOA) 来进行优化^[7].

针对区间型数据, 考虑将区间的中心序列与半径序列分别作为两组输入向量. 利用 WOA 算法优化 SVM 模型中的参数进行预测, 然后再转化为区间的左右端点.

3 基于 COWA 算子的相关性区间组合预测优化模型与性质研究

设某种指标序列的实际观察区间值为 $x_t = [x_t^L, x_t^U], t = 1, 2, \dots, N$, 有 m 种区间预测方法, 第 i 种区间预测方法在第 t 时刻的预测区间值序列为 $x_{it} = [x_{it}^L, x_{it}^U], i = 1, 2, \dots, m, t = 1, 2, \dots, N$.

定义 3.1 令

$$\hat{x}_t = [\hat{x}_t^L, \hat{x}_t^U] = \left[\sum_{i=1}^m l_i x_{it}^L, \sum_{i=1}^m l_i x_{it}^U \right] \quad (12)$$

则称 \hat{x}_t 为区间组合预测值, 其中 l_1, l_2, \dots, l_m 为 m 种区间预测方法的加权系数, 满足 $\sum_{i=1}^m l_i = 1, l_i \geq 0$.

定义 3.2 令

$$F_Q(\hat{x}_t) = \text{COWA}(\hat{x}_t), F_Q(x_{it}) = \text{COWA}(x_{it}) \quad (13)$$

则称 $F_Q(\hat{x}_t)$ 为基于 COWA 算子的区间组合预测值, 称 $F_Q(x_{it})$ 为基于 COWA 算子的区间预测值.

定义 3.3 令

$$\left. \begin{aligned} \bar{x} &= \frac{1}{N} \sum_{t=1}^N F_Q(x_t), \bar{\hat{x}} = \frac{1}{N} \sum_{t=1}^N F_Q(\hat{x}_t), \\ \bar{x}_i &= \frac{1}{N} \sum_{t=1}^N F_Q(x_{it}) \end{aligned} \right\} \quad (14)$$

则称 \bar{x} 为基于 COWA 算子的区间型实际观察值序列的均值, $\bar{\hat{x}}, \bar{x}_i$ 分别称为组合预测值序列和第 i 种区间预测值序列的均值.

定义 3.4 令

$$\begin{aligned} R_i &= \sum_{t=1}^N (F_Q(x_t) - \bar{x})(F_Q(x_{it}) - \bar{x}_i) / \\ & \left(\sqrt{\sum_{t=1}^N (F_Q(x_t) - \bar{x})^2} \sqrt{\sum_{t=1}^N (F_Q(x_{it}) - \bar{x}_i)^2} \right) \end{aligned} \quad (15)$$

$$\begin{aligned} R &= \sum_{t=1}^N (F_Q(x_t) - \bar{x})(F_Q(\hat{x}_t) - \bar{\hat{x}}) / \\ & \left(\sqrt{\sum_{t=1}^N (F_Q(x_t) - \bar{x})^2} \sqrt{\sum_{t=1}^N (F_Q(\hat{x}_t) - \bar{\hat{x}})^2} \right) \end{aligned} \quad (16)$$

则称 R_i 为基于 COWA 算子的第 i 种区间预测方法预测值序列与实际观察值区间序列的相关系数, 称 R 为基于 COWA 算子的区间型组合预测值序列与实际观察值区间序列相关系数.

命题 3.1 基于 COWA 算子区间组合预测值序列与基于 COWA 算子区间预测值序列满足如下关系:

$$F_Q(\hat{x}_t) = \sum_{i=1}^m l_i F_Q(x_{it}), \forall t = 1, 2, \dots, N \quad (17)$$

证明 由式(2)和(13)知

$$\begin{aligned} F_Q(\hat{x}_t) &= (1 - \alpha)\hat{x}_t^L + \alpha\hat{x}_t^U, \\ F_Q(x_{it}) &= (1 - \alpha)x_{it}^L + \alpha x_{it}^U \end{aligned} \quad (18)$$

由式(12)可得

$$\hat{x}_t^L = \sum_{i=1}^m l_i x_{it}^L, \hat{x}_t^U = \sum_{i=1}^m l_i x_{it}^U \quad (19)$$

结合式(18)和(19), 则有

$$\begin{aligned} F_Q(\hat{x}_t) &= (1 - \alpha) \sum_{i=1}^m l_i x_{it}^L + \alpha \sum_{i=1}^m l_i x_{it}^U = \\ &= \sum_{i=1}^m l_i ((1 - \alpha)x_{it}^L + \alpha x_{it}^U) = \sum_{i=1}^m l_i F_Q(x_{it}). \end{aligned}$$

证毕.

命题 3.2 基于 COWA 算子的区间型组合预测值序列的均值和第 i 种区间预测值序列的均值满足如下关系:

$$\bar{\hat{x}} = \sum_{i=1}^m l_i \bar{x}_i \quad (20)$$

证明 由式(14)和(17)知

$$\begin{aligned} \bar{\hat{x}} &= \frac{1}{N} \sum_{t=1}^N F_Q(\hat{x}_t) = \frac{1}{N} \sum_{t=1}^N \left(\sum_{i=1}^m l_i F_Q(x_{it}) \right) = \\ &= \sum_{i=1}^m l_i \left(\frac{1}{N} \sum_{t=1}^N F_Q(x_{it}) \right) = \sum_{i=1}^m l_i \bar{x}_i. \end{aligned}$$

证毕.

定义 3.5 称 $e_t = F_Q(x_t) - \bar{x}$ 为第 t 时刻基于 COWA 算子的实际观察值与其均值的离差, $e_{it} = F_Q(x_{it}) - \bar{x}_i$ 为相应的第 i 种区间预测的离差, $i = 1, 2, \dots, m, t = 1, 2, \dots, N$.

由式(17)、(20)和定义 3.5, 基于 COWA 算子的第 i 种区间预测方法预测值序列以及区间组合预测值序列与实际观察值区间序列相关系数式(15)和(16)可表示成

$$R_i = \sum_{t=1}^N e_t e_{it} / \left(\sqrt{\sum_{t=1}^N e_t^2} \sqrt{\sum_{t=1}^N (e_{it})^2} \right), \quad \left. \begin{aligned} & \\ & i = 1, 2, \dots, m \end{aligned} \right\} \quad (21)$$

$$R = \sum_{i=1}^m l_i \left(\sum_{t=1}^N e_t e_{it} \right) / \left(\sqrt{\sum_{t=1}^N e_t^2} \sqrt{\sum_{t=1}^N \left(\sum_{i=1}^m l_i e_{it} \right)^2} \right) \quad (22)$$

从式(22)可看出, 组合预测值序列与实际观察值序列基于 COWA 算子的区间型相关系数 R 是关于加权系数 l_1, l_2, \dots, l_m 的函数, 可记作 $R(l_1, l_2, \dots, l_m)$. 相关系数愈大说明预测值与实际值愈接近, 因此基于 COWA 算子的区间型组合预测模型可表示如下:

$$\begin{aligned} \max R(l_1, l_2, \dots, l_m) &= \\ & \sum_{i=1}^m l_i \left(\sum_{t=1}^N e_t e_{it} \right) / \left(\sqrt{\sum_{t=1}^N e_t^2} \sqrt{\sum_{t=1}^N \left(\sum_{i=1}^m l_i e_{it} \right)^2} \right) \\ \text{s. t. } & \begin{cases} \sum_{i=1}^m l_i = 1, \\ l_i \geq 0, i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (23)$$

记 $R_{\min} = \min_{1 \leq i \leq m} \{R_i\}$, $R_{\max} = \max_{1 \leq i \leq m} \{R_i\}$. 显然, R_{\min} 和 R_{\max} 分别表示基于 COWA 算子的区间预测值序列与实际观察值区间序列的相关系数的最小值和最大值. 从相关系数的角度来考察, R_{\min} 对应的区间预测方法是最差的, 而 R_{\max} 对应的区间预测方法是最好的.

定义 3.6 若 $R(l_1, l_2, \dots, l_m) < R_{\min}$, 则称基于 COWA 算子的区间型组合预测模型为劣性组合预测, 若 $R_{\min} \leq R(l_1, l_2, \dots, l_m) \leq R_{\max}$, 则称相应的组合预测模型为非劣性的, 若 $R(l_1, l_2, \dots, l_m) > R_{\max}$, 则称之为优性组合预测.

记 $L = (l_1, l_2, \dots, l_m)^T$ 为区间组合预测加权系数的列向量, $e_i = (e_{i1}, e_{i2}, \dots, e_{iN})^T$ 为第 i 种区间预测值的离差列向量.

定义 3.7 令 $e_{ij} = e_i^T e_j = \sum_{t=1}^N e_{it} e_{jt}, i, j = 1, 2, \dots, m$, 称 $E = (e_{ij})_{m \times m}$ 为区间组合预测协方差信息矩阵.

从定义 3.7 可以看出, $e_{ij} = e_i^T e_j = e_j^T e_i = e_{ji}$, 所以区间组合预测协方差信息矩阵 E 是对称矩阵.

引理 3.1 若 m 种区间预测方法的离差向量组 e_1, e_2, \dots, e_m 是线性无关的, 则区间组合预测协方差信息矩阵 E 是正定矩阵.

证明方法类似于文献[20]的定理 4.2.1, 限于篇幅, 故从略.

定理 3.1 若 m 种区间预测方法的离差向量组 e_1, e_2, \dots, e_m 是线性无关的, 则模型(22)的任一可行解对应的组合预测模型至少是非劣性组合.

证明 设 $L = (l_1, l_2, \dots, l_m)^\top$ 是模型(23)的一组可行解, 因为 $R_{\min} = \min_{1 \leq i \leq m} \{R_i\}$, 则有

$$R_i \geq R_{\min}, i = 1, 2, \dots, m-1 \quad (24)$$

由式(21)和(24)可得

$$\sum_{t=1}^N e_t e_{it} \geq R_{\min} \left(\sqrt{\sum_{t=1}^N e_t^2} \sqrt{\sum_{t=1}^N (e_{it})^2} \right), i = 1, 2, \dots, m \quad (25)$$

将式(25)代入到式(22)中, 并注意定义 3.7, 则有

$$\begin{aligned} R &\geq \sum_{i=1}^m l_i \left(R_{\min} \sqrt{\sum_{t=1}^N e_t^2} \sqrt{e_{ii}} \right) / \\ &\quad \left(\sqrt{\sum_{t=1}^N e_t^2} \sqrt{\sum_{t=1}^N \left(\sum_{i=1}^m l_i e_{it} \right)^2} \right) = \\ R_{\min} \sum_{i=1}^m l_i \sqrt{e_{ii}} / &\left(\sqrt{\sum_{i=1}^m \sum_{j=1}^m l_i l_j \sum_{t=1}^N e_{it} e_{jt}} \right) = \\ R_{\min} \sum_{i=1}^m l_i \sqrt{e_{ii}} / &\left(\sqrt{\sum_{i=1}^m \sum_{j=1}^m l_i l_j e_{ij}} \right) = \\ R_{\min} \sum_{i=1}^m l_i \sqrt{e_{ii}} / &(\sqrt{L^\top E L}) \quad (26) \end{aligned}$$

由于离差向量组 e_1, e_2, \dots, e_m 是线性无关的, 所以由引理 3.1 知矩阵 $E = (e_{ij})_{m \times m}$ 是正定矩阵, 则其任意的二阶主子式均大于 0, 即 $\begin{vmatrix} e_{ii} & e_{ij} \\ e_{ji} & e_{jj} \end{vmatrix} > 0$, 从而有 $e_{ii} e_{jj} > e_{ij}^2$, 此即

$$e_{ij} < \sqrt{e_{ii}} \sqrt{e_{jj}}, i, j = 1, 2, \dots, m \quad (27)$$

所以,

$$\begin{aligned} \sqrt{L^\top E L} &= \sqrt{\sum_{i=1}^m \sum_{j=1}^m l_i l_j e_{ij}} \leq \\ \sqrt{\sum_{i=1}^m \sum_{j=1}^m l_i l_j \sqrt{e_{ii}} \sqrt{e_{jj}}} &= \sum_{i=1}^m l_i \sqrt{e_{ii}} \quad (28) \end{aligned}$$

将式(28)代入到式(26)中, 则有 $R \geq R_{\min}$, 由定义 3.6 知结论成立. 证毕.

推论 3.1 若 m 种区间预测方法的离差向量组 e_1, e_2, \dots, e_m 是线性无关的, 则简单平均组合预测方法至少是非劣性组合.

4 非线性规划问题的近似求解

模型(23)是一非线性规划问题, 考虑到其求解的复杂性, 本文利用合作对策理论中的 Shapley 值法进行近似求解. 对策的局中人为 m 种区间预测方法, 为此引入如下概念:

定义 4.1 称 $P = \{1, 2, \dots, m\}$ 为 m 种区间预

测方法参与组合预测的合作对策的局中人集, $S \subset P, S = \{i_1, i_2, \dots, i_s\}$, 则称 S 为合作对策的联盟, 组合预测值序列与实际观察值序列基于 COWA 算子的区间型相关系数 $R(l_1, l_2, \dots, l_m)$ 称为合作对策的收益函数. 记为 $R(P) = R(l_1, l_2, \dots, l_m)$.

从定义 4.1 可以看出, 联盟 S 表示从 m 种区间预测方法中取 s 个区间预测方法参与组合预测. 则联盟 S 对应的收益函数记为

$$R(S) = R(l_{i_1}, l_{i_2}, \dots, l_{i_s}).$$

定义 4.2^[20] 若

$$\left. \begin{aligned} \varphi_i(v) &= \sum_{i \in S} \frac{(m - |S|)! (|S| - 1)!}{m!} \cdot \\ &[R(S) - R(S - \{i\})], \\ i &= 1, 2, \dots, m \end{aligned} \right\} \quad (29)$$

称 $\varphi_i(v)$ 为第 i 种区间预测方法的平均“贡献”, 其中 m 为组合预测方法总数, $|S|$ 为参与组合预测中的联盟 S 的个数.

定义 4.3^[20] 称 $\Phi(v) = (\varphi_1(v), \varphi_2(v), \dots, \varphi_m(v))$ 为合作 m 种区间预测方法对策的 Shapley 值.

Shapley 值是合作对策收益分配的一种方法. 显然, 某个区间预测方法对组合预测的贡献越大, 其对应的权重应该也越大, 为此构造了如下权重计算办法.

由文献[21]知: $\sum_{i=1}^m \varphi_i(v) = R(P)$, 令

$$l_i = \varphi_i(v) / R(P), i = 1, 2, \dots, m \quad (30)$$

则 l_i 为各种区间预测方法的权重. 显然, l_i 满足非负性和归一化的条件.

5 实例分析

为了说明所提模型的可行性, 本文从美国能源署网站截取了 2015-03-16~2019-03-18 时段西德克萨斯中质原油现货价格(WTI spot price)数据, 该数据以周为频率的, 单位为每桶的油价. 以一周内的原油现货价格最小值与最大值分别作为实际观察区间值的下界和上界. 2015-03-16~2018-10-29 时段共 190 个区间数据作为训练集, 求出模型参数, 2018-11-05~2019-03-18 时段共 20 个区间数据作为测试集, 进行预测. 以 Holt 指数平滑模型、MLP 神经网络以及 WOA-SVM 模型作为三种区间预测方法对实际区间值进行预测. 其中 MLP 神经网络的参数设置如下: 为防止过拟合, 划分训练、测试和验证数据的比例为 18 : 1 : 2. 在实验中, 以中心和半径

的前三期预测下一期,故输入层、输出层神经元个数分别为 6 和 2. 模型以实际值与预测值的均方误差达到最小为调参目标,多次实验选择隐藏层节点数为 12. 类似的,我们针对 WOA-SVM 模型进行相应的参数设置:最大迭代次数设置为 100;种群数量设置为 30;WOA 算法中对数螺旋形状的常量 b 设置为 1;SVM 算法中惩罚系数范围设置为 $[0.01, 6000]$,径向基函数 RBF 中参数 σ 范围设置为 $[0.01, 100]$. 以训练集的误差平方和为适应度函数,分别对原始区间对应的中心序列和半径序列利用 WOA-SVM 模型进行优化,可以获得相应区间预测值.

上述 Holt 指数平滑模型、MLP 神经网络以及 WOA-SVM 模型的三种区间预测的结果如图 1~3 所示,其中横坐标为 t/week ,纵坐标为 WTI spot prices/(USD/barrel).

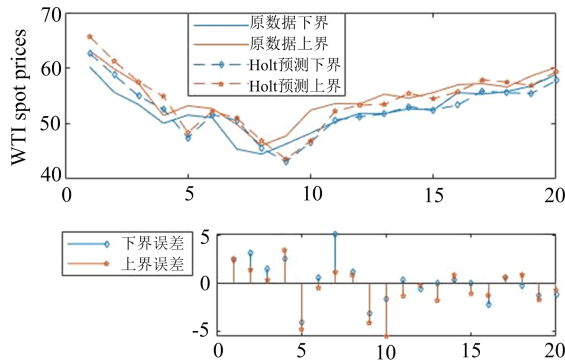


图 1 Holt 指数平滑区间预测

Fig. 1 Interval forecasting with Holt exponential smoothing

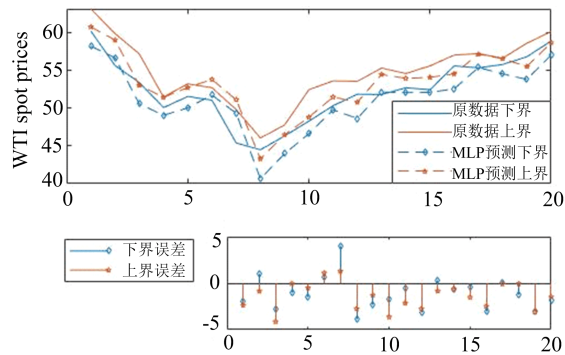


图 2 MLP 区间预测

Fig. 2 Interval forecasting with MLP

对上述区间型数据,若取 BUM 函数 $Q(x) = x^2$, 则态度参数 $\alpha = 1 - \int_0^1 x dQ(x) = 1/3$, 由式 (2)知

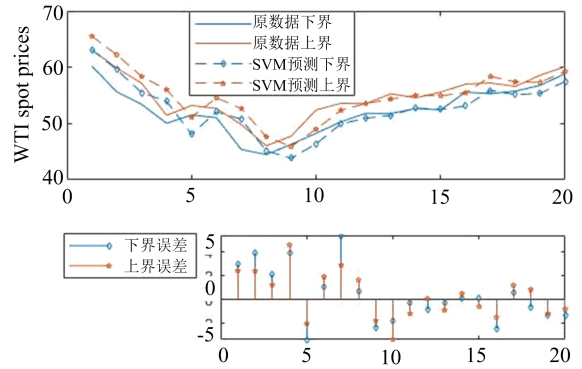


图 3 WOA-SVM 区间预测

Fig. 3 Interval forecasting with WOA-SVM

$$x_t = F_Q([x_t^L, x_t^U]) = \frac{2}{3}x_t^L + \frac{1}{3}x_t^U,$$

$$x_{it} = F_Q([x_{it}^L, x_{it}^U]) = \frac{2}{3}x_{it}^L + \frac{1}{3}x_{it}^U, i=1,2,3.$$

由式(15)可得三种区间预测方法对应的相关系数: $R_1 = 0.8992, R_2 = 0.9264, R_3 = 0.8858$.

一般而言,区间预测方法的误差大,则其在组合预测中对应的权重应该较小. 为此,引入均方误差倒数方法来计算组合预测的初始权重,其计算公式为

$$E_{ii} = \sum_{t=1}^N (x_t - x_{it})^2, l_i = E_{ii}^{-\frac{1}{2}} / (\sum_{i=1}^m E_{ii}^{-\frac{1}{2}}), i=1,2,3 \quad (31)$$

由式(31)可得初始权重:

$$L = (l_1, l_2, l_3) = (0.3335, 0.3440, 0.3226).$$

由式(16)可得合作对策的收益,则有: $R(P) = R\{1,2,3\} = 0.9229$.

同理,由式(31)可得联盟 $S = \{1,2\}$ 参与的组合预测方法对应的初始权重为

$$(l_1, l_2) = (E_{11}^{-\frac{1}{2}}, E_{22}^{-\frac{1}{2}}) / (\sum_{i=1}^2 E_{ii}^{-\frac{1}{2}}) = (0.4923, 0.5077).$$

所以由式(16)知可得 $R(\{1,2\}) = 0.9339$.

类似地有: $R(\{1,3\}) = 0.8943, R(\{2,3\}) = 0.9273$.

由式(29)知第一种区间预测方法的平均贡献为

$$\varphi_1(v) = \frac{1}{3}R(\{1\}) +$$

$$\frac{1}{6}[R(\{1,2\}) - R(\{2\})] +$$

$$\frac{1}{6}[R(\{1,3\}) - R(\{3\})] +$$

$$\frac{1}{3}[R(\{1,2,3\}) - R(\{2,3\})] = 0.3009.$$

同理可得： $\varphi_2(v) = 0.3311, \varphi_3(v) = 0.2909$.

由式(30)知,基于 Shapley 值的组合预测加权系数为

$$l_1 = 0.3260, l_2 = 0.3587, l_3 = 0.3153.$$

相比于图 1~图 3 这 3 种单项预测方法,图 4 给出了组合区间预测的效果图.图中 0 值线表示预测值和实际值相等的情形.从图 4 可以发现,组合区间预测的上下界误差距离 0 值的数据点比较多,从而组合预测与实际值更接近,可以直观地看出组合预测效果更好.

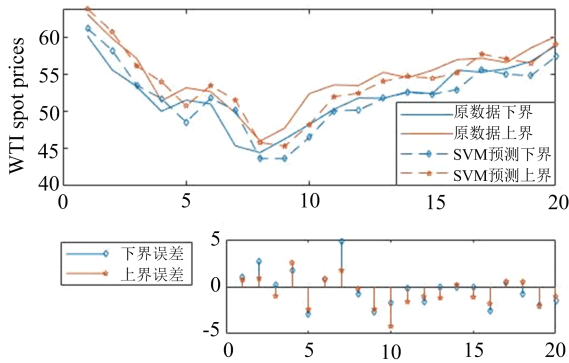


图 4 组合区间预测

Fig. 4 Combination interval forecasting

6 模型评价

为了说明预测模型的有效性,利用文献[22]及文献[23]的指标进行评价.

①平均区间中心位置误差平方和:

$$MSEP = \frac{1}{N} \sum_{t=1}^N (m_t - \hat{m}_t)^2;$$

②平均区间长度误差平方和:

$$MSEL = \frac{1}{N} \sum_{t=1}^N (r_t - \hat{r}_t)^2;$$

③连续区间预测误差平方和:

$$ISSE = \sum_{t=1}^N (F_Q(x_t) - F_Q(x_{it}))^2;$$

④连续区间均方差:

$$IMSE = \frac{1}{N} \sqrt{\sum_{t=1}^N (F_Q(x_t) - F_Q(x_{it}))^2}.$$

计算结果如表 1 所示.从表 1 可以看出,本文提出的区间组合预测模型除了 MSEL 这个指标,组合预测效果略微差之外,其余误差指标的值均小于各个区间预测方法的,表明本文提出的区间组合预测方法是有效可行的.

表 1 各项区间预测误差

Tab. 1 Each interval forecasting error

	MSEP	MSEL	ISSE	IMSE
Holt 指数平滑模型	4.3881	0.6213	86.4102	0.4648
MLP 模型	3.9643	0.3485	81.2255	0.4506
WOA-SVM 模型	4.3342	0.2855	92.3563	0.4805
组合预测	2.8328	0.3560	59.5631	0.3859

7 参数的灵敏度分析

以上所举例子中取 BUM 函数为 $Q(x) = x^2$, 此时态度参数为 $\alpha = 1/3$. 但是针对不同的态度参数,结果可能会有所不同,为探究态度参数对组合模型的权重系数及误差指标的影响情况,取态度参数 α 分别为 0, 0.1, 0.2, ..., 1.0, 参数的灵敏度分析结果分别如图 5 及图 6 所示.

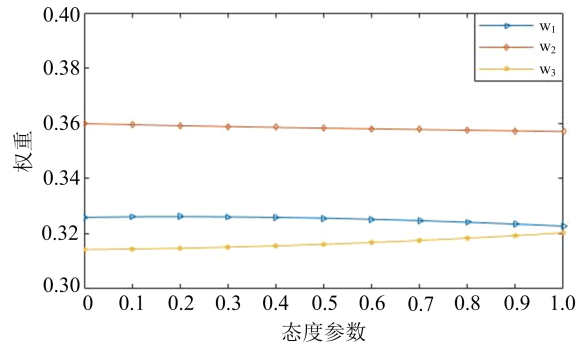


图 5 α 对 w 的影响

Fig. 5 The influence of α on w

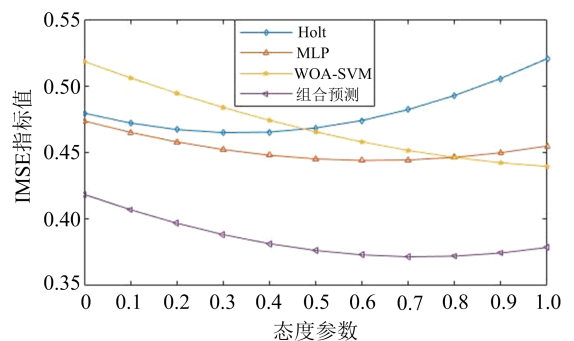


图 6 α 对 IMSE 指标的影响

Fig. 6 The influence of α on IMSE index

从图 5 可以看出:无论态度参数 α 取何值, w_2 的权重一直大于 w_1, w_3 , 说明 w_2 在组合预测中起主要作用, w_3 的权重随着 α 的逐渐增大而增大,当 $\alpha \in [0, 0.3]$ 时,权重 w_1 有所增大,而后随着 α 的逐渐增大而减小.

由图6可以看出:无论态度参数 α 取何值,组合预测的IMSE指标均小于其余三种区间预测值,说明该方法是合理有效的,特别当 $\alpha=0.7$ 左右时,组合预测的IMSE值达到最小。

8 结论

由于不确定环境的影响,对复杂系统的预测采用区间数信息,相对于实数信息可能更加合理地反映客观实际的情况。现有的文献对区间数的预测,通常考虑区间的左右端点的信息,分别构建组合预测模型,有可能在外推预测中出现区间上下界错位的现象,为此本文将连续有序加权平均算子引入到区间组合预测模型中,通过参数的设置将区间型数据转化为实数型数据,可以克服上述缺陷。另外,从统计学的角度考虑,相关系数是一种新的预测精度的度量准则,因此建立基于COWA算子的相关性区间组合预测优化模型具有重要的理论和现实意义。

本文构建了新的基于COWA算子的区间组合预测模型,探讨了模型的若干性质。考虑到模型计算的复杂性,文中将组合预测视为合作博弈。通过引入Shapley值的方法,探讨了新的组合预测模型的权重的近似计算方法,并结合西德克萨斯中质原油现货价格数据验证了模型的有效性。由于COWA算子的参数不确定性,如何合理确定最优参数,进一步提高组合预测的精度,并探讨基于COWA算子的区间组合预测模型的冗余预测方法的判定等性质^[24],有待于进一步研究。

参考文献(References)

- [1] MAIA A L S, CARVALHO F D A T D. Holt's exponential smoothing and neural network models for forecasting interval-valued time series [J]. *International Journal of Forecasting*, 2011, 27(3):740-759.
- [2] 尚文利,张立婷,李世超,等. 基于Holt指数平滑模型的油井产量动态预测[J]. *自动化与仪表*, 2018, 33(4): 68-70.
- [3] 常欣卓,杨开忠,李新,等. 基于非线性自回归神经网络的局部大气密度预测方法[J]. *中国科学技术大学学报*, 2017, 47(12):1015-1022.
- [4] 张乐,汪传旭. 基于GM(1,1)-MLP神经网络组合模型的物流总额预测[J]. *上海海事大学学报*, 2018, 39(4): 61-65.
- [5] ZAREI T, BEHYAD R. Predicting the water production of a solar seawater greenhouse desalination unit using multi-layer perceptron model [J]. *Solar Energy*, 2019, 177: 595-603.
- [6] 赵春晓. 基于支持向量机的混沌时间序列预测方法的研究[D]. 沈阳:东北大学, 2008.
- [7] XU Y, YANG W, WANG J. Air quality early-warning system for cities in China[J]. *Atmospheric Environment*, 2017, 148:239-257.
- [8] 张莉,卢星凝,陆从林,等. 支持向量机在高考成绩预测分析中的应用[J]. *中国科学技术大学学报*, 2017, 47(1):1-9.
- [9] BATES J M, GRANGER C W J. The combination of forecast[J]. *Operational Research Quarterly*, 1969, 20(4): 451-468.
- [10] 唐小我,马永开,曾勇,杨桂元. 现代组合预测和组合投资决策方法及应用研究[M]. 北京:科学出版社, 2003.
- [12] LAOUAFI A, MORDJAOUI M, HADDAD S, et al. Online electricity demand forecasting based on an effective forecast combination methodology [J]. *Electric Power Systems Research*, 2017, 148: 35-47.
- [13] XIONG T, LI C, BAO Y, et al. A combination method for interval forecasting of agricultural commodity futures prices [J]. *Knowledge-Based Systems*, 2015, 77: 92-102.
- [14] WANG J, HENG J, XIAO L, et al. Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting [J]. *Energy*, 2017, 125: 591-613.
- [15] 沈家骅,严振祥. 基于区间分析的组合预测系数确定方法[J]. *武汉理工大学学报(交通科学与工程版)*, 2006(6): 1077-1080.
- [16] SENGUPTA A, PAL T K. Fuzzy Preference Ordering of Interval Numbers in Decision Problems[M]. New York: Springer, 2009.
- [17] YAGER R R. OWA aggregation over a continuous interval argument with applications to decision making[J]. *IEEE Transactions on Systems, Man and Cybernetics: Part B*, 2004, 34(5): 1952-1963.
- [18] MUKHERJEE S, GIROSI F, OSUNA E. Nonlinear prediction of chaotic time series using support vector machines[C]// *Neural Networks for Signal Processing VII. Proceedings of the 1997 IEEE Signal Processing Society Workshop. IEEE*, 1997: 511-520.
- [19] 刘仁志,黄张裕,秦洁,等. 基于粒子群优化支持向量机的边坡稳定性预测[J]. *甘肃科学学报*, 2019, 31(1): 102-106.
- [20] 陈华友. 组合预测方法有效性理论及其应用[M]. 北京:科学出版社, 2008.
- [21] 王建华. 对策论[M]. 北京:清华大学出版社, 1986.
- [22] 陈华友,李翔,金磊,姚梦杰. 基于相关系数及IOWA算子的区间组合预测方法[J]. *统计与决策*, 2012(6): 83-86.
- [23] 朱家明,陈华友,周礼刚,等. 基于ICOFWA算子的连续区间模糊组合预测模型及其应用[J]. *模糊系统与数学*, 2016, 30(3):172-184.
- [24] 陈华友,侯定丕. 基于预测有效度的优性组合预测模型研究[J]. *中国科学技术大学学报*, 2002, 32(2):172-180.