

基于鸽群优化算法的支持向量机在电力需求总量预测中的应用

田书欣¹,周全²,程浩忠³,柳璐³,路亮²,江粟²

(1. 上海电力大学 电气工程学院,上海 200090;2. 国家电网公司西南分部,四川 成都 610041;

3. 上海交通大学 电气工程系 电力传输与功率变换控制教育部重点实验室,上海 200240)

摘要:电力需求总量的科学预测是经济转型阶段电力系统规划与运行的重要依据。引入融合分段二次Lagrange插值函数的新型灰色关联理论,从经济发展、产业结构、用电环境以及居民生活4个方面分析社会经济新常态指标与电力需求总量之间的相关程度,筛选出影响电力需求增长率波动的关键因素;进而以电力需求总量及相关因素为训练数据集,利用融合莱维飞行特征的改进鸽群优化算法对支持向量机的参数进行优化,建立具有最佳参数、强泛化能力的电力需求总量预测模型。基于我国某电网区域电力需求历史实测数据的算例结果表明,所建模型具有更好的优化效率和预测精度。

关键词:电力需求总量;新常态指标;鸽群优化算法;支持向量机;灰色关联

中图分类号:TM 715

文献标志码:A

DOI:10.16081/j.epae.202004016

0 引言

经济的快速增长有力地保障了社会安定繁荣和人民生活水平的提高。与此同时,经济增长所需的用电量也急剧增加,这使得电力需求与经济社会协调发展问题日益受到关注。电力需求变化趋势的准确预测,不仅可以确保规划期内该地区用电的供需平衡,还可以促进电力能源与经济、生态的可持续发展。电力需求包括电量类需求和负荷类需求两大类^[1]。针对电量类的电力需求,文献[2-3]对中长期尺度下的电量不确定性及周期性进行深入分析,基于构建灰色系统理论或核密度估计方法对电量的发展趋势进行中长期预测。文献[4-5]根据国民经济历史发展和城市电力需求情况,分析电力负荷发展过程中的产值单耗、电力弹性系数和人均用电量3项指标,并利用产值用电单耗法、电力弹性系数法及人均电量法,对全社会的中长期用电量提出对应预测方案。文献[6]通过使用人工神经网络对气候、电力需求样本数据进行训练,在此基础上形成气候与电力需求之间的典型关系曲线,从而实现对未来电力需求总量变化趋势的预测。文献[7-8]利用结构风险最小化原则代替传统的经验风险最小化,充分挖掘电力需求数据序列的内在规律,提出基于支持向量机SVM(Support Vector Machine)的负荷类电力需求预测模型。现有文献采用的电力需求预测模型各有优缺点,但随着我国经济发展进入转型阶段,

亟需捕捉经济新常态下电力需求特征的预测方法。

鉴于此,本文采用融合分段二次Lagrange插值函数的新型灰色关联理论对电力需求总量的各经济社会影响因素进行相关性分析,并将相关性较高的主要因素作为SVM训练样本的输入,同时利用改进的鸽群优化PIO(Pigeon-Inspired Optimization)算法对SVM参数进行优化,提高电力需求总量预测模型的精度和泛化感知能力。最后,以我国某电网区域年电力需求总量的数据序列作为实际算例,验证本文所提方法的可行性和有效性。

1 经济转型发展及电力需求总量关联分析

1.1 电力需求关联指标

经济新常态下电力需求总量增长率的振荡放缓的影响因素涉及多个方面。为了与我国经济结构变化相适应,以及与国际管理规定相协调相接轨,可从经济发展、产业结构、用电环境以及居民生活4个方面选取9个社会经济新常态指标对其与电力需求总量之间的关联性进行分析,如表1所示。

表1 经济转型与电力需求总量关联分析的原始指标

Table 1 Original indexes of correlation analysis between economic transition and total power demand

类别	指标	指标编号
经济发展	GDP	B ₁
	GDP年增长率	B ₂
产业结构	一产用电量	B ₃
	二产用电量	B ₄
	三产用电量	B ₅
用电环境	最大负荷日所在月份的平均气温	B ₆
	年最高气温	B ₇
	年最大负荷日的日最高气温	B ₈
居民生活	居民生活用电量	B ₉

收稿日期:2019-10-13;修回日期:2020-02-19

基金项目:国家电网公司管理咨询项目(SGXN000GHWT17-00001)

Project supported by the Management Consulting Project of State Grid Corporation of China(SGXN000GHWT1700001)

1.2 新型灰色关联分析模型

传统灰色关联分析方法是利用行为因子序列曲线几何形状的相似程度来量化分析其相关程度^[9],而忽略了波动序列曲线之间的接近性问题。实际上,相关性大小不仅与曲线相似程度紧密关联,而且与曲线之间的接近程度密切相关。

针对经济新常态下电力需求相关数据序列可能出现的振荡情况,本文引入分段二次 Lagrange 插值函数对相关序列数据进行逼近,并结合序列数据初值化处理方式,利用2组序列数据逼近函数所围图形面积来表征其灰色相似关联度和灰色相近关联度^[10]。具体分析步骤如下。

(1)采集电力需求总量与某一相关影响因素的2组统计数据序列 $X_1 = \{x_1(1), x_1(2), \dots, x_1(n)\}$ 和 $X_2 = \{x_2(1), x_2(2), \dots, x_2(n)\}$,其中 n 为采集的样本数。数据序列 X_1 和 X_2 的逼近函数分别为 $f_1(t)$ 、 $f_2(t)$ ($t \in [1, n]$)。常用的逼近函数有多项式函数等。鉴于灰色关联分析的准确度主要受到原始数据序列光滑离散性的限制,本文可采用具有良好光滑性和高逼近精度的分段二次 Lagrange 插值对原始数据序列进行预先处理,即对数据序列 X_i ($i = 1, 2$),任取相邻数据 $x_i(k-1)$ 、 $x_i(k)$ 、 $x_i(k+1)$,以 $[k-1, k+1]$ 为插值区间构造分段二次 Lagrange 函数:

$$f_{ik}(t) = x_i(k-1)l_{k-1}(t) + x_i(k)l_k(t) + x_i(k+1)l_{k+1}(t) \quad (1)$$

其中, $l_{k-1}(t) = (t-k)(t-k-1)/2$, $l_k(t) = -(t-k+1)(t-k-1)$, $l_{k+1}(t) = (t-k+1)(t-k)/2$, $k=2, 4, 6, \dots$, 且 $n \geq 3$ 。

结合式(1),推导得到数据序列 X_i ($i=1, 2$) 的分段二次 Lagrange 插值函数的形式如下。

若 $n \geq 3$, 且 n 为奇数, 则:

$$f_i(t) = \sum_{k=2}^{n-1} \delta_k f_{ik}(t) \quad i = 1, 2 \quad (2)$$

$$\delta_k = \begin{cases} 1 & t \in [k-1, k+1] \\ 0 & t \notin [k-1, k+1] \end{cases}$$

若 $n \geq 4$, 且 n 为偶数, 则:

$$f_i(t) = \sum_{k=2}^{n-2} \delta_k f_{ik}(t) + \delta_{f_{i(n-1)}}(t) \quad i = 1, 2 \quad (3)$$

$$\delta_k = \begin{cases} 1 & t \in [k-1, k+1] \\ 0 & t \notin [k-1, k+1] \end{cases}$$

$$\delta_{f_{i(n-1)}} = \begin{cases} 1 & t \in [n-1, n] \\ 0 & t \notin [n-1, n] \end{cases}$$

基于式(1)~(3)得到 X_i ($i=1, 2$) 在 $t \in [k-1, k+1]$ 时的分段二次 Lagrange 插值函数。

(2)如果对2组数据序列 X_1 和 X_2 进行关联度分析,需要首先对它们进行初值化处理,得到新的数据序列 $X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)\}$ 和 $X_2^{(0)} =$

$\{x_2^{(0)}(1), x_2^{(0)}(2), \dots, x_2^{(0)}(n)\}$,其中原始数据序列中的 $x_1(1)$ 和 $x_2(1)$ 均不为0, 则:

$$x_i^{(0)}(k) = x_i(k)/x_i(1) \quad k=1, 2, \dots, n; i=1, 2; n \geq 3 \quad (4)$$

相应地,对逼近函数进行初值化处理后得到无量纲的 $f_1^{(0)}(t)$ 和 $f_2^{(0)}(t)$, 即:

$$f_i^{(0)}(t) = f_i(t)/f_i(1) \quad t=1, 2, \dots, n; i=1, 2; n \geq 3 \quad (5)$$

(3)定义 α_{12} 为2组不同数据序列 X_1 和 X_2 的基于所围图形面积的灰色相似关联度:

$$\alpha_{12} = \frac{1}{1 + \Delta s_{12}} \quad (6)$$

$$\Delta s_{12} = \int_1^n |f_1^{(0)}(t) - f_2^{(0)}(t)| dt \approx \sum |f_1^{(0)}(t) - f_2^{(0)}(t)| \Delta t \quad (7)$$

其中, Δs_{12} 为2组不同数据序列的逼近函数无量纲处理后 $f_1^{(0)}(t)$ 和 $f_2^{(0)}(t)$ 对应曲线与直线 $t=1$ 和直线 $t=n$ 所围图形的面积。

α_{12} 反映了2个量纲不同的不同序列数据 $X_1^{(0)}$ 和 $X_2^{(0)}$ 对应曲线的相似程度。 α_{12} 的值越接近1, 即 Δs_{12} 的值越接近0, 则说明 $f_1^{(0)}(t)$ 和 $f_2^{(0)}(t)$ 重合的可能性越大, $f_1(t)$ 和 $f_2(t)$ 就越相似, $X_1^{(0)}$ 和 $X_2^{(0)}$ 以及对应的不同数据序列 X_1 和 X_2 的相似度也就越高。

定义 β_{12} 为2组不同数据序列 X_1 和 X_2 的基于所围图形面积的灰色相近关联度:

$$\beta_{12} = \frac{1}{1 + \Delta S_{12}} \quad (8)$$

$$f_i^{(1)}(t) = \frac{f_i(t)}{\max(f_i(t))} \quad t=1, 2, \dots, n; i=1, 2; n \geq 3 \quad (9)$$

$$\Delta S_{12} = \int_1^n |f_1^{(1)}(t) - f_2^{(1)}(t)| dt \approx \sum |f_1^{(1)}(t) - f_2^{(1)}(t)| \Delta t \quad (10)$$

其中, $\max(f_i(t))$ 表示数据序列逼近函数的最大值; ΔS_{12} 为不同数据序列无量纲化后的逼近函数 $f_1^{(1)}(t)$ 和 $f_2^{(1)}(t)$ 对应曲线与直线 $t=1$ 和直线 $t=n$ 所围图形的面积。

β_{12} 反映了2组数据序列 X_1 和 X_2 的相近程度, 其值越接近于1, 则 ΔS_{12} 越接近0, 说明 $f_1(t)$ 和 $f_2(t)$ 、 $X_1^{(0)}$ 和 $X_2^{(0)}$ 以及 X_1 和 X_2 均越来越相近。

本文提出的灰色相似关联度和灰色相近关联度能够多视角反映电力需求总量与各影响因素数据序列之间的相关性,避免关联性分析失效或失准的缺陷。

2 基于 PIO-SVM 模型的电力需求预测

2.1 SVM 模型

SVM 应用于电力需求预测,本质上是利用非线性变换将电力需求相关影响因素作为训练样本

数据集,通过构造最优决策函数,将非线性函数优化问题转换为高维特征线性函数的回归问题^[11]。本文中的SVM模型是在由灰色关联理论确定的电力需求关键影响因素基础上形成训练样本集 $D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l), \mathbf{x} \in \mathbf{R}^n, y \in \mathbf{R}\}$,其中 l 为样本点数。进而形成所求的线性函数:

$$f(\mathbf{x}) = \mathbf{w}\phi(\mathbf{x}_i) + b \quad (11)$$

其中, \mathbf{w} 为函数平面法向量; b 为函数平面的偏置; $\phi(\cdot)$ 为将输入向量映射到Hilbert空间的非线性变换。

相应的最优化问题的模型^[12]为:

$$\begin{cases} \min F = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{s.t. } y_i - \mathbf{w}\phi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i & i = 1, 2, \dots, l \\ \mathbf{w}\phi(\mathbf{x}_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (12)$$

其中, C 为正则化参数; ε 为估计精度; ξ_i, ξ_i^* 为在误差 ξ 约束下训练误差的限值。

结合拉格朗日乘子可得:

$$\begin{aligned} L = & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) + \\ & \sum_{i=1}^l \lambda_i (y_i - \mathbf{w}\phi(\mathbf{x}_i) - b - \varepsilon - \xi_i) + \\ & \sum_{i=1}^l \lambda_i^* (\mathbf{w}\phi(\mathbf{x}_i) + b - y_i - \varepsilon - \xi_i^*) - \\ & \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \end{aligned} \quad (13)$$

其中, $\lambda_i, \lambda_i^*, \eta_i, \eta_i^* \geq 0 (i = 1, 2, \dots, l)$ 。

则优化问题的对偶形式为:

$$\begin{cases} \max L = -\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\lambda_i - \lambda_i^*) (\lambda_j - \lambda_j^*) K(\mathbf{x}_i, \mathbf{x}_j) + \\ \quad \sum_{i=1}^l (\lambda_i - \lambda_i^*) y_i - \sum_{i=1}^l (\lambda_i + \lambda_i^*) \varepsilon \\ \text{s.t. } \sum_{i=1}^l (\lambda_i - \lambda_i^*) = 0 \\ 0 \leq \lambda_i, \lambda_i^* \leq C \quad i = 1, 2, \dots, l \end{cases} \quad (14)$$

其中, $K(\mathbf{x}_i, \mathbf{x}_j)$ 为核函数,其影响模型泛化能力,性能较优的常用核函数有径向基核函数和多项式核函数,若 $K(\mathbf{x}_i, \mathbf{x}_j) = K_{ij} = \varphi(\mathbf{x}_i) \varphi(\mathbf{x}_j) = \exp[-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / (2\sigma^2)]$,则为径向基核函数,若 $K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + 1)^d$,则为多项式核函数, $\mathbf{x}_i \cdot \mathbf{x}_j$ 为 \mathbf{x}_i 与 \mathbf{x}_j 的内积, σ 为核函数的宽度参数, d 为核函数的维度参数。

式(12)的矩阵形式为:

$$\begin{cases} \min \frac{1}{2} \boldsymbol{\lambda}^T \mathbf{H} \boldsymbol{\lambda} + \boldsymbol{\gamma}^T \boldsymbol{\lambda} \\ \text{s.t. } [1, 1, \dots, 1, -1, \dots, -1, -1] \boldsymbol{\lambda} = 0 \\ 0 \leq \lambda_i, \lambda_i^* \leq C \quad i = 1, 2, \dots, l \end{cases} \quad (15)$$

其中, $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_l, \lambda_1^*, \dots, \lambda_l^*]^T$ 为待优化变量; \mathbf{H} 和 $\boldsymbol{\gamma}$ 为系数阵,由式(16)确定。

$$\begin{cases} \mathbf{H} = \begin{bmatrix} \mathbf{K} & -\mathbf{K} \\ -\mathbf{K} & \mathbf{K} \end{bmatrix}, \mathbf{K} = (K_{ij})_{l \times l} \quad i, j = 1, 2, \dots, l \\ \boldsymbol{\gamma} = \begin{bmatrix} \varepsilon \mathbf{I} + \mathbf{Y} \\ \varepsilon \mathbf{I} - \mathbf{Y} \end{bmatrix} \end{cases} \quad (16)$$

其中, $\mathbf{Y} = [y_1, y_2, \dots, y_l]^T$; \mathbf{I} 为元素全为1的列向量。

于是得到回归方程:

$$\begin{aligned} f(\mathbf{x}) = & \left[\sum_{i=1}^l (\lambda_i - \lambda_i^*) \phi(\mathbf{x}_i) \right] \phi(\mathbf{x}) + b = \\ & \sum_{i=1}^l (\lambda_i - \lambda_i^*) (\phi(\mathbf{x}_i) \phi(\mathbf{x})) + b = \\ & \sum_{i=1}^l (\lambda_i - \lambda_i^*) K(\mathbf{x}_i, \mathbf{x}) + b \end{aligned} \quad (17)$$

SVM模型预测性能主要受到正则化参数 C 、核函数宽度参数 σ 或维数参数 d 的影响。为了获得更加精确的预测结果^[13-14],选择最优参数组合至关重要。

2.2 改进的PIO算法

PIO算法是近年来由模拟自然界中鸽子归巢行为而提出的一种群体智能优化算法^[15],其在众多领域中展现出收敛速度快、鲁棒性强^[16]的优越特性。本文在现有PIO算法的基础上增加鸽子在干扰下迷失方向或在陌生释放地点熟悉地形风貌过程中的随机飞行行为,增强该算法的全局搜索能力,从而形成改进的PIO算法,其基本原理如下。

首先,基于地磁场建立地图模型,在二维空间中用 z_i 表示第 i 只鸽子的位置,用 v_i 表示第 i 只鸽子的速度,则鸽子 k 的位置 \mathbf{Z}_k 和速度 \mathbf{V}_k 可表示为:

$$\begin{aligned} \mathbf{Z}_k &= [z_{k1}, z_{k2}, \dots, z_{km}] \\ \mathbf{V}_k &= [v_{k1}, v_{k2}, \dots, v_{km}] \end{aligned} \quad (18)$$

其中, $k = 1, 2, \dots, M_1$; M_1 为鸽群的种群数量; m 为优化对象的维数。

鸽子的位置和速度通过每次优化迭代进行更新。鸽子的当前速度由其上一代速度、上一代位置和最优化位置共同决定,相应更新公式如下:

$$\begin{cases} \mathbf{V}_i(t) = \mathbf{V}_i(t-1) e^{-Rt} + r_{an} (\mathbf{Z}_g - \mathbf{Z}_i(t-1)) \\ \mathbf{Z}_i(t) = \mathbf{Z}_i(t-1) + \mathbf{V}_i(t) \end{cases} \quad (19)$$

其中, t 为当前迭代次数; R 为地磁因子, $R \in [0, 1]$; \mathbf{Z}_g 为通过比较所有鸽子的位置得到的当前种群中全局最优位置; r_{an} 表示 $[0, 1]$ 之间的随机数。

针对鸽子磁场波动时可能出现的迷失方向或陌生地点的随机搜索情况,在现有PIO算法中引入莱维飞行特征对鸽子位置进行更新操作:

$$\mathbf{Z}_i(t) = \mathbf{Z}_i(t-1) + \frac{0.01 r_1 \mu}{|r_2|^{1/\theta}} \mathbf{Z}_i(t-1) \quad (20)$$

$$\mu = \left\{ \frac{\Gamma(1+\theta)\sin(\pi\theta/2)}{\Gamma[(1+\theta)/2]\theta \cdot 2^{(\theta-1)/2}} \right\}^{1/\theta} \quad (21)$$

其中, $\Gamma(\cdot)$ 表示服从 Γ 分布; r_1 和 r_2 为 $[0, 1]$ 之间随机数; θ 为常量。

当迭代次数达到所要求的最大值后, 停止在地图模型中鸽群的优化工作, 进入地标模型进行搜索。地标模型是模拟鸽子靠近目的地时更多依赖附近地标直接飞向目标的特性。因而在地标模型中, 离目的地较远的鸽子对目标附近的地标不熟悉, 不具有辨识路径的能力, 可直接淘汰舍去, 这将使每次迭代后鸽子的数量减少一半。目的地附近的鸽子将成为鸽群的中心, 带领其他鸽子飞向目标。根据此原则, 选取种群中心点 $Z_c(t)$ 作为剩余鸽群的飞行参考方向, 对个体的位置进行迭代更新:

$$Z_c(t) = \frac{\sum_{i=1}^{M_p(t)} Z_i(t) F(Z_i(t))}{M_p(t) \sum_{i=1}^{M_p(t)} F(Z_i(t))} \quad (22)$$

$$M_p(t) = \frac{M_p(t-1)}{2} \quad (23)$$

$$Z_i(t) = Z_i(t-1) + r_{an}(Z_c(t) - Z_i(t-1)) \quad (24)$$

其中, $Z_i(t)$ 为第 t 次迭代第 i 只鸽子的位置; M_p 为每次迭代中鸽子的数量; $F(\cdot)$ 为适应度函数, 对最小化问题, $F(Z_i(t)) = 1/(f_{\min}(Z_i(t)) + \varepsilon)$, $f_{\min}(Z_i(t))$ 为最小目标函数, 对最大化问题, $F(Z_i(t)) = f_{\max}(Z_i(t))$, $f_{\max}(Z_i(t))$ 为最大目标函数。

通过每次迭代舍去偏离目标的个体, 既确保了算法的较优信息又提高了算法迭代效率。循环迭代至最大次数, 结束地标模型中的算法工作。

2.3 电力需求总量预测模型

针对中长期电力需求预测跨度时间长、获取历史数据量有限的特点, 本文采用适宜小样本容量系统的 SVM 模型对电力需求总量变化趋势进行预测, 并引入基于 PIO 算法的参数优化技术来提高 SVM 预测中长期电力需求总量的精度和收敛速度。基于 PIO-SVM 的电力需求模型预测步骤如下。

(1) 初始化参数。根据实际需要设定模型的种群规模 M_1 、最大迭代次数 N_e 、地磁因子 R 以及个体鸽子的初始位置和速度等。个体鸽子的位置代表 SVM 模型的参数 C 、 σ 或 C 、 d 。

(2) 产生 SVM 训练数据集。利用融合分段二次 Lagrange 插值函数的新型灰色关联理论, 通过考虑相似关联度和相近关联度来多视角选取经济转型阶段影响电力需求总量的关键因素, 并将这些关键影响因素和电力需求总量的样本归一化数据组成 SVM 点对构成的训练样本集 D 。

(3) PIO-SVM 的参数优化。基于 SVM 的训练样本集, 以式(25)中训练误差作为适应度函数 f_e , 采用改进 PIO 算法, 选定 SVM 参数的初值, 并通过地图模型和地标模型的操作, 结合式(19)–(24)对鸽子位置和速度进行迭代更新, 根据适应度函数值选取当前单个鸽子最佳位置和当前种群整体最佳位置。

$$f_e = \frac{1}{l} \sum_{i=1}^l \left(\frac{y'_i - \hat{y}'_i}{y'_i} \right)^2 \quad (25)$$

其中, y'_i 为归一化后的训练集数据向量的第 i 个元素; \hat{y}'_i 为 SVM 模型的输出向量的第 i 个元素。

(4) 训练 SVM。根据步骤(3)中最小训练误差对应的最佳参数对 SVM 进行设置, 选择径向基核函数或多项式核函数训练一次 SVM, 并基于训练得到的模型对预测集点对样本进行回归预测和反归一化处理。

(5) 评价预测模型。评价 SVM 模型的预测精度, 若不符合要求, 转至步骤(3), 将更新后的正则化参数 C 、核函数宽度参数 σ 或维数参数 d 代入 SVM 模型, 对产生的新种群重新进行 SVM 训练, 并保存各次迭代中适应度最好的个体, 直到达到最大迭代次数为止。

(6) 输出预测结果。基于获取最优参数的 PIO-SVM 模型输出电力需求总量预测结果, 与实际值进行比较分析。

3 算例分析

本文将所提方法应用于我国某实际电网区域的电力需求总量分析。采用该地区 1990 至 2013 年的全社会整体年电力需求总量数据作为训练集, 以 2014 至 2018 年电力需求总量作为经济转型期间目标测试集, 对该阶段经济新常态型下电力需求总量变化趋势进行预测分析。图 1 给出了该地区 1990 至 2018 年的全社会整体年电力需求总量及其增长率的变化情况。

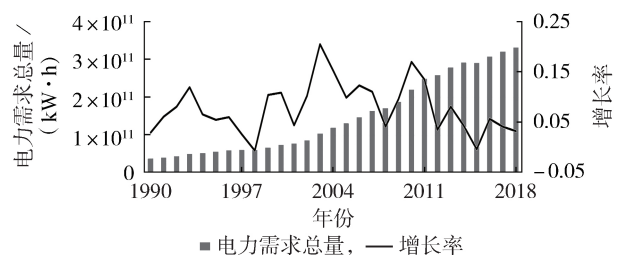


图1 1990至2018年某区域电力需求总量及增长率

Fig.1 Total power demand and growth rate in a region from 1990 to 2018

从图1可以看出, 1990到2018年该电网区域电力需求增长了8.5倍, 其中2001至2010年期间是该电网区域电力需求增长最快的时期。但是从2011

年开始,受到金融危机反复和蔓延以及我国经济进入转型调整阶段的影响,该电网区域经济发展走向增速减缓区,相应电力需求总量增长率有逐渐放缓的趋势。尤其是2015年,电力需求总量出现负增长,其部分原因是政府开始转变经济发展方式,将经济发展的重点从高能耗的制造业转到服务业。

3.1 电力需求总量与其影响因素的关联度分析

针对经济转型发展对电力需求的影响,采用表1中与社会经济转型形势紧密联系的经济发展、产业结构、用电环境以及居民生活4个方面的因素进行分析。同时为了更准确地分析这些因素与电力需求之间的关联性,利用分段二次 Lagrange 插值函数对经济新常态下原始电力需求总量数据进行逼近处理,使其逼近曲线较常规逼近函数曲线更加光滑贴近,从而解决原有灰色相关分析方法不能客观反映波动序列相似性与接近性的弊端。图2给出了多项式核函数、分段二次 Lagrange 插值函数拟合1990至2013年电力需求总量数据的情况。

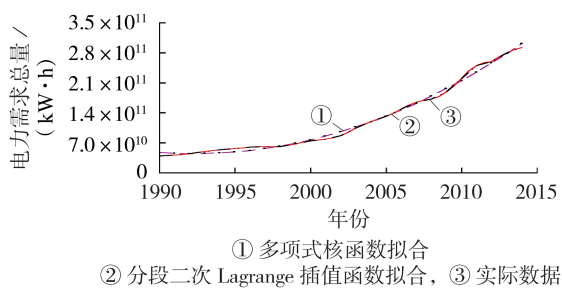


图2 电力需求总量数据逼近曲线的比较
Fig.2 Comparison of approximation curves for total power demand data

由图2可以看出,分段二次 Lagrange 插值函数能很好地跟踪原始数据的变化趋势,其贴近期度高,且能够保持逼近曲线较为光滑。因而,利用分段二次 Lagrange 插值函数对原始数据进行逼近处理后,能使随之建立的灰色相似关联度和灰色相近关联度模型具有较高的评估精度。鉴于电力需求总量与其他相关影响因素的数值和量纲均差别较大,需要首先对各相关指标进行无量纲化处理,然后结合式(1)—(10),采用融合分段二次 Lagrange 插值函数的新型灰色关联分析方法计算表1中各项指标 B_1 — B_9 与电力需求之间的灰色相似关联度和灰色相近关联度,结果如表2所示。结合式(9),指标 B_1 — B_9 无量纲化处理后的相关数据见附录中表A1。

结合表2,选取灰色相似关联度大于0.5、灰色相近关联度大于0.1的关键影响因素,即认为GDP、二产用电量和居民生活用电量与电力需求总量的灰色相似关联度和灰色相近关联度相对较高,说明这3项指标与电力需求总量的相关性较大。

表2 各项指标与电力需求之间的灰色关联度

Table 2 Gray correlations between each index and power demand

指标	灰色相似关联度	灰色相近关联度
B_1	0.5141	0.1070
B_2	0.2510	0.0219
B_3	0.1546	0.0203
B_4	0.6479	0.1312
B_5	0.4032	0.0314
B_6	0.0931	0.0226
B_7	0.1041	0.0236
B_8	0.1066	0.0235
B_9	0.5469	0.1336

3.2 电力需求总量预测及结果分析

基于表2,选用GDP、二产用电量和居民生活用电量作为SVM预测模型的输入变量训练样本集,对训练样本进行归一化,然后利用改进的PIO算法对SVM模型中的参数进行优化。其中,训练样本为1990至2013年电力需求数据,而将2014至2018年电力需求数据作为测试样本。为了验证融合PIO算法的SVM的预测效果,本文分别采用多元线性回归MLR(Multiple Linear Regression)、多项式核函数下基于粒子群优化算法的支持向量机(PKF-PSO-SVM)、径向基核函数下基于粒子群优化算法的支持向量机(RBF-PSO-SVM)^[17]、多项式核函数下基于PIO算法的支持向量机(PKF-PIO-SVM)、径向基核函数下基于PIO算法的支持向量机(RBF-PIO-SVM)对电力需求总量进行预测。设置种群数为60,迭代次数为100。表3给出了径向基核函数下粒子群优化算法和改进PIO算法寻优后的SVM正则化参数 C 和核函数宽度参数 σ 以及多项式核函数下粒子群优化算法和改进PIO算法寻优后的SVM正则化参数 C 和维数参数 d 。

表3 不同SVM模型的参数对比

Table 3 Comparison of parameters optimized by different SVM models

模型	C	σ	d
PKF-PSO-SVM	13.68	—	4.980
RBF-PSO-SVM	14.22	0.31	—
PKF-PIO-SVM	5.35	—	2.163
RBF-PIO-SVM	7.76	0.18	—

下面利用5种模型对该电网区域1990至2018年电力需求总量及其相关数据进行分析,预测结果见附录中表A2。基于表A2中预测结果得到对比曲线如图3所示。5种模型预测结果与实际值的平均相对误差和单年最大相对误差如表4所示。

对比图3和表4中预测结果可知:2种核函数下的PIO-SVM模型预测结果的平均相对误差和最大相对误差均小于2种核函数下的PSO-SVM模型,说

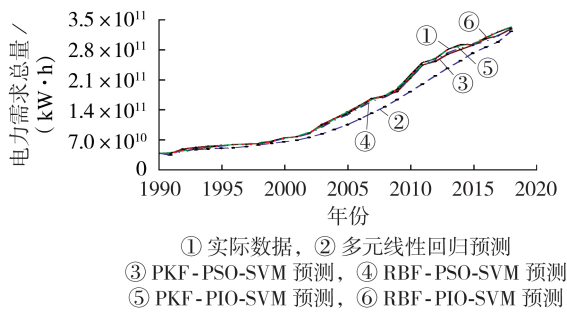


图3 不同模型预测结果对比

Fig.3 Comparison of forecasting results among different models

表4 不同电力需求模型的预测误差

Table 4 Forecasting errors of different power demand models

预测模型	平均相对误差 / %	最大相对误差 / %
多元线性回归	11.900	21.03
PKF-PSO-SVM	2.277	9.67
RBF-PSO-SVM	2.078	6.19
PKF-PIO-SVM	0.996	4.76
RBF-PIO-SVM	0.588	5.63

明应用寻优能力更优的PIO-SVM模型能获取更佳参数,解决年电力需求总量预测问题具有更高的精度;虽然在训练SVM阶段基于径向基核函数的PIO-SVM模型存在个别年份最大相对误差要高于基于多项式核函数的PIO-SVM模型的情况,但总体的平均相对误差较小,且在图3中其预测的电力消费总量曲线与实际曲线重合度更高,说明基于径向基核函数的PIO-SVM模型泛化能力更好。因此,基于灰色相似和相近关联性可以充分挖掘经济转型阶段与电力需求总量变化相关的关键因素,并结合改进PIO算法寻优能力强的特点选择最优参数,获得具有更高电力需求预测精度、更优泛化能力的SVM优化模型,进而通过跟踪感知经济转型阶段电力需求的变化趋势,为优化经济结构、调整电力产业布局提供理论支撑。

4 结论

随着经济增速的放缓和经济结构转型,我国电力需求增长率整体呈现振荡缓慢新常态,尤其是2015年,电力需求总量出现负增长,这使得传统的预测方法已经不能很好地适应经济转型时期电力需求总量中长期预测的需要。选取表征经济转型特征的GDP、二产用电量和居民生活用电指标与电力需求总量作为输入进行基于PIO算法的SVM预测模型训练,算例分析对比了PKF-PSO-SVM、RBF-PSO-SVM、PKF-PIO-SVM、RBF-PIO-SVM模型以及多元线性回归模型的预测结果与预测性能指标。可以得出,与粒子群优化算法、多项式核函数相比,采用径

向基核函数下PIO算法对模型参数进行寻优,可以使SVM的训练更加准确,从而形成寻优能力强、预测精度高、泛化能力优的电力需求总量预测新方法,该方法能够科学准确跟踪经济转型诱发增长率波动背景下未来电量类电力需求发展的变化趋势,为确保电量供需平衡、提高电力企业经营效益提供一定指导作用。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 康重庆,杨高峰,夏清. 电力需求的不确定性分析[J]. 电力系统自动化,2005,29(17):14-19,39.
KANG Chongqing, YANG Gaofeng, XIA Qing. Analysis of the uncertainty of electric power demand[J]. Automation of Electric Power Systems, 2005, 29(17): 14-19, 39.
- [2] 赵书强,胡利宁,田捷夫,等. 基于中长期风电光伏预测的多能源电力系统合约电量分解模型[J]. 电力自动化设备,2019,39(11):13-19.
ZHAO Shuqiang, HU Lining, TIAN Jiefu, et al. Contract power decomposition model of multi-energy power system based on mid-long term wind power and photovoltaic electricity forecasting[J]. Electric Power Automation Equipment, 2019, 39(11): 13-19.
- [3] 周宏,黄婷,戴韧,等. 几种灰色模型用于电力消费中期预测研究[J]. 电网技术,2000,24(7):49-54.
ZHOU Hong, HUANG Ting, DAI Ren, et al. Study on alternative grey models to forecast middle term electric power consumption[J]. Power System Technology, 2000, 24(7): 49-54.
- [4] 薛雪,石文星,孙天,等. 一种应用于智能电网的建筑电力负荷互动管理策略[J]. 电力自动化设备,2018,38(12):100-106.
XUE Xue, SHI Wenxing, SUN Tian, et al. Interactive building power load management strategy for smart grid[J]. Electric Power Automation Equipment, 2018, 38(12): 100-106.
- [5] 范德成,王韶华,张伟. 低碳经济目标下我国电力需求预测研究[J]. 电网技术,2012,36(7):19-25.
FAN Decheng, WANG Shaohua, ZHANG Wei. Research on prediction of China's electric power demand under low-carbon economy target[J]. Power System Technology, 2012, 36(7): 19-25.
- [6] SANTOS MARQUES M, SOARES F C, RUSSI J L. Modeling typical power demand curves using climatic data[J]. IEEE Latin America Transactions, 2016, 14(7):3278-3284.
- [7] 李瑾,刘金朋,王建军. 采用支持向量机和模拟退火算法的中长期负荷预测方法[J]. 中国电机工程学报,2011,31(16):63-66.
LI Jin, LIU Jinpeng, WANG Jianjun. Mid-long term load forecasting based on simulated annealing and SVM algorithm[J]. Proceedings of the CSEE, 2011, 31(16): 63-66.
- [8] 王东,史晓霞,尹交英. 不同核函数的支持向量机用于空调负荷预测的对比研究[J]. 电工技术学报,2015,30(增刊1):531-535.
WANG Dong, SHI Xiaoxia, YIN Jiaoying. Prediction on hourly load of air conditioning by RBF support vector machine [J]. Transactions of China Electrotechnical Society, 2015, 30 (Supplement 1): 531-535.
- [9] 王雁凌,吴梦凯,周子青,等. 基于改进灰色关联度的电力负荷影响因素量化分析模型[J]. 电网技术,2017,41(6):1772-1778.
WANG Yanling, WU Mengkai, ZHOU Ziqing, et al. Quantita-

- tive analysis model of power load influencing factors based on improved grey relational degree[J]. Power System Technology, 2017, 41(6):1772-1778.
- [10] 胡钊,徐则诚,金哲,等. 基于灰色关联模型的输电线路舞动预警方法[J]. 电力自动化设备, 2018, 38(2):117-122.
HU Po, XU Zecheng, JIN Zhe, et al. Early warning method of transmission line galloping based on grey relational model[J]. Electric Power Automation Equipment, 2018, 38(2):117-122.
- [11] 张慧,卢文冰,赵雄文,等. 基于最小二乘支持向量机和小波神经网络的电力线通信信道噪声建模研究[J]. 电工技术学报, 2018, 33(16):3879-3888.
ZHANG Hui, LU Wenbing, ZHAO Xiongwen, et al. Noise modeling for power line communication channel using the LS-SVM and wavelet neural networks[J]. Transactions of China Electrotechnical Society, 2018, 33(16):3879-3888.
- [12] 刘渝根,陈超. 基于人工蜂群算法优化支持向量机的接地网腐蚀速率预测模型[J]. 电力自动化设备, 2019, 39(5):182-186, 200.
LIU Yugen, CHEN Chao. Corrosion rate prediction model of grounding grid based on support vector machine optimized by artificial bee colony algorithm[J]. Electric Power Automation Equipment, 2019, 39(5):182-186, 200.
- [13] 李赢,舒乃秋. 基于模糊聚类和完全二叉树支持向量机的变压器故障诊断[J]. 电工技术学报, 2016, 31(4):64-70.
LI Ying, SHU Naiqiu. Transformer fault diagnosis based on fuzzy clustering and complete binary tree support vector machine[J]. Transactions of China Electrotechnical Society, 2016, 31(4):64-70.
- [14] 戴远航,陈磊,张玮灵,等. 基于多支持向量机综合的电力系统暂态稳定评估[J]. 中国电机工程学报, 2016, 36(5):1173-1180.
DAI Yuanhang, CHEN Lei, ZHANG Weiling, et al. Power system transient stability assessment based on multi-support vector machines[J]. Proceedings of the CSEE, 2016, 36(5):1173-1180.
- [15] DUAN H B, QIAO P X. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning[J]. International Journal of Intelligent Computing and Cybernetics, 2014, 7(1):24-37.
- [16] ZHANG S J, DUAN H B. Gaussian pigeon-inspired optimization approach to orbital spacecraft formation reconfiguration[J]. Chinese Journal of Aeronautics, 2015, 28(1):200-205.
- [17] 孙毅,石墨,单葆国,等. 基于粒子群优化支持向量机的电能替代潜力分析方法[J]. 电网技术, 2017, 41(6):1767-1771.
SUN Yi, SHI Mo, SHAN Baoguo, et al. Electric energy substitution potential analysis method based on particle swarm optimization support vector machine[J]. Power System Technology, 2017, 41(6):1767-1771.

作者简介:



田书欣

田书欣(1985—),男,安徽合肥人,讲师,博士,主要研究方向为电力系统规划(E-mail: tsx396@163.com);

周全(1973—),男,四川泸州人,高级工程师,硕士,主要研究方向为电网规划;

程浩忠(1962—),男,浙江东阳人,教授,博士,主要研究方向为电力系统规划、电压稳定、电能质量等。

(编辑 王锦秀)

Application of pigeon-inspired optimization algorithm based SVM in total power demand forecasting

TIAN Shuxin¹, ZHOU Quan², CHENG Haozhong³, LIU Lu³, LU Liang², JIANG Li²

(1. College of Electric Power Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

2. Southwest Branch of State Grid Corporation of China, Chengdu 610041, China;

3. Key Laboratory of Control of Power Transmission and Conversion, Ministry of Education, Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Scientific forecasting of total power demand is an important basis for power system planning and operation during the period of economy transition. A novel grey correlation theory combined with piecewise quadratic Lagrange interpolation function is introduced to analyze the correlation degree between social economic new normal indexes and total power demand from four aspects of economic development, industrial structure, power utilization environment and residential living, and the key factors that affecting the growth rate fluctuation of total power demand are screened. Further, the total power demand and its relevant factors are taken as training dataset, an improved pigeon-inspired optimization algorithm merging with Lévy flights is adopted to optimize the parameters of SVM (Support Vector Machine), and a forecasting model of total power demand with the best parameters and strong generalization capability is built. The case results based on the historical measured power demand data of a power grid region in China show that the proposed model has better optimization efficiency and forecasting accuracy.

Key words: total power demand; new normal index; pigeon-inspired optimization algorithm; support vector machines; grey correlation

附录：

表 A1 1990 至 2013 年某实际电网区域电力需求相关指标归一化数据
 Table A1 Normalized historical data of related indexes for total electricity consumption in a practical power grid region from 1990 to 2013

年份	指标归一化数据								
	B ₁	B ₂	B ₃	B ₄	B ₅	B ₆	B ₇	B ₈	B ₉
1990	0.031	0.468	0.761	0.292	0.131	0.778	0.679	0.679	0.168
1991	0.035	0.397	0.143	0.058	0.026	0.792	0.679	0.679	0.035
1992	0.041	0.501	0.159	0.059	0.027	0.792	0.679	0.679	0.038
1993	0.053	0.785	0.146	0.061	0.029	0.806	0.690	0.690	0.038
1994	0.071	1.000	0.139	0.062	0.032	0.806	0.690	0.690	0.044
1995	0.090	0.732	1.000	0.403	0.234	0.806	0.690	0.690	0.298
1996	0.105	0.495	0.984	0.397	0.245	0.806	0.690	0.690	0.333
1997	0.120	0.386	0.728	0.323	0.207	0.806	0.690	0.690	0.297
1998	0.128	0.200	0.700	0.341	0.253	0.819	0.702	0.702	0.349
1999	0.134	0.138	0.680	0.338	0.263	0.819	0.702	0.702	0.371
2000	0.145	0.220	0.556	0.303	0.244	0.819	0.702	0.702	0.305
2001	0.159	0.279	0.566	0.336	0.260	0.819	0.702	0.702	0.336
2002	0.176	0.316	0.648	0.346	0.273	0.833	0.714	0.714	0.353
2003	0.200	0.381	0.674	0.419	0.338	0.833	0.714	0.714	0.412
2004	0.239	0.550	0.673	0.492	0.420	0.861	0.833	0.833	0.458
2005	0.275	0.433	0.701	0.576	0.490	0.861	0.833	0.833	0.535
2006	0.319	0.458	0.762	0.635	0.514	0.889	0.952	0.952	0.590
2007	0.386	0.594	0.760	0.702	0.569	0.861	0.905	0.905	0.626
2008	0.465	0.587	0.751	0.739	0.645	0.722	0.929	0.905	0.726
2009	0.523	0.353	0.761	0.780	0.720	0.806	0.881	0.810	0.888
2010	0.634	0.606	0.779	0.879	0.784	0.806	0.929	0.833	0.870
2011	0.783	0.670	0.876	0.981	0.885	1.000	1.000	1.000	0.881
2012	0.891	0.390	0.943	1.000	1.000	0.972	0.952	0.952	1.000
2013	1.000	0.349	0.861	0.911	0.951	0.972	0.952	0.881	0.869

表 A2 电力需求模型预测值与实际值的对比

Table A2 Comparison between actual value and predicted value based on different power demand models

年份	实际值 / (GW · h)	预测值 / (GW · h)				
		MLR	PKF-PSO-SVM	RBF-PSO-SVM	PKF-PIO-SVM	RBF-PIO-SVM
1990	38 934	39 021	38 934	38 934	38 952	38 934
1991	41 336	35 023	39 009	40 856	41 161	39 010
1992	44 701	48 196	49 026	41 932	45 070	46 030
1993	50 056	47 546	52 640	48 831	49 846	50 640
1994	53 375	49 563	55 533	52 466	54 057	53 030
1995	56 316	50 802	57 792	55 546	57 315	55 790
1996	59 695	52 393	59 167	58 865	59 465	59 170
1997	61 271	55 191	60 775	60 790	60 947	60 780
1998	60 922	58 422	60 965	60 896	63 823	60 460
1999	67 319	61 938	66 002	66 034	67 602	66 900
2000	74 670	65 459	74 293	72 943	73 904	74 290
2001	78 022	69 706	77 686	77 006	78 485	77 690
2002	86 040	75 995	84 745	84 233	88 105	85 740
2003	103 724	84 318	106 470	99 826	101 968	103 470
2004	119 532	94 396	119 319	115 591	118 469	119 320
2005	131 466	105 888	135 294	128 291	134 484	131 290
2006	147 764	118 674	149 634	143 869	147 856	147 630
2007	164 173	131 965	166 085	160 112	163 469	164 080
2008	171 054	146 475	173 008	168 866	173 001	171 010
2009	187 611	163 027	181 607	183 862	190 133	187 610
2010	219 588	181 554	213 627	212 443	217 105	219 630
2011	249 224	199 662	243 306	241 868	245 552	249 310
2012	258 222	218 451	252 347	254 951	260 215	258 350
2013	279 286	236 655	268 854	272 419	278 870	279 450
2014	291 601	253 766	279 412	283 165	289 017	291 810
2015	290 830	270 896	289 985	290 297	292 980	291 080
2016	307 520	285 300	304 819	298 375	305 745	307 820
2017	320 420	296 980	311 763	313 151	320 815	320 760
2018	330 980	322 100	327 367	326 214	331 943	331 370