

DOI: 10.11992/tis.201708033

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180403.1433.004.html>

## 快速双非凸回归算法及其电力数据预测应用

王锋华<sup>1</sup>, 成敬周<sup>1</sup>, 文凡<sup>2</sup>

(1. 国网浙江省电力公司, 浙江 杭州 310000; 2. 国网浙江省电力公司 经济技术研究院, 浙江 杭州 310000)

**摘 要:** 为适应产能输出、运营效益等电力数据预测应用, 文中提出一种快速双非凸回归 (double nonconvex regression, DNR) 预测算法。首先, 将经典稀疏编码分类技术解释为预测回归模型, 并划分为训练阶段和测试阶段, 使之适合标量预测应用; 其次, 针对经典 Lasso 模型存在的稀疏性不足以及噪声拟合单一问题, 该算法通过  $l_p$  范数约束逼近原始稀疏编码问题的误差重构项和系数正则项, 具有更为灵活的模型形式和应用范围。最后, 通过交替方向乘子框架实现了重构系数的优化升级策略。为确保 ADMM 优化子问题具有快速解, 提出一种改进的迭代阈值规则用于更新非凸  $l_p$  约束项, 解决了原始算法陷入的局部最优问题。在电力企业实际运行产出和运营指标数据上的实验结果表明, DNR 在预测效果和预测效率上均优于经典的支持向量机、BP 神经网络以及非凸约束预测方法。

**关键词:** 交替方向乘子法; 电力数据预测;  $l_p$  范数约束; 迭代阈值方法

**中图分类号:** TP18; TM715 **文献标志码:** A **文章编号:** 1673-4785(2018)04-0665-08

中文引用格式: 王锋华, 成敬周, 文凡. 快速双非凸回归算法及其电力数据预测应用[J]. 智能系统学报, 2018, 13(4): 665-672.

英文引用格式: WANG Fenghua, CHENG Jingzhou, WEN Fan. Fast double nonconvex regression algorithm for forecast of electric power data[J]. CAAI transactions on intelligent systems, 2018, 13(4): 665-672.

## Fast double nonconvex regression algorithm for forecast of electric power data

WANG Fenghua<sup>1</sup>, CHENG Jingzhou<sup>1</sup>, WEN Fan<sup>2</sup>

(1. State Grid Zhejiang Electric Power Company, Hangzhou 310000, China; 2. Economic Research Institute, State Grid Zhejiang Electric Power Company, Hangzhou 310000, China)

**Abstract:** In this paper, we propose a new forecasting algorithm called double nonconvex regression (DNR) for the fast forecast of electricity power data such as the outputs of production ability and operational benefit. First, we reinterpret the typical sparse coding classification method as a regression model for forecasting, and further divide the model into training and testing phases to fit scalar-quantity forecasts. Next, we transform the constraints of representation residuals and coefficient regularization into a nonconvex  $l_p$  norm for better approximation and broader application. Lastly, we adopt the alternating direction method of multipliers algorithm to optimize the formulated forecast problem. To achieve a fast update rule for  $l_p$  norm constrained subproblems, we propose a new iterative threshold method that avoids the local minimum issue. Compared with typical methods such as the SVM, BP neural network, and nonconvex regularization methods, the proposed algorithm achieves surprisingly good experimental results for electricity power data.

**Keywords:** alternating direction method of multiplier (ADMM); forecast of electric power data;  $l_p$  norm constraint; iterative threshold method

收稿日期: 2017-08-31. 网络出版日期: 2018-04-03.

基金项目: 国家电网浙江省电力公司科技项目 (5211JY15001V);  
国家电网公司科技项目 (5211011600RJ).

通信作者: 王锋华. E-mail: [wangfenghua0627@126.com](mailto:wangfenghua0627@126.com).

电网是加快工业进步、提高居民生活质量、保持社会稳定健康发展的基础。因此, 电网的发展具有重要的意义。电网企业的责任是确保安

全、经济、清洁和可持续的能源供应,从而为社会、经济的健康发展,人民生活水平的逐步提高保驾护航。其中,产能输出和运营指标管理作为电网企业经营活动的中枢,是企业发展的重中之重,能对企业整体运营状况、管控经营、监控战略产生实效。此外,通过挖掘数据资产的潜在价值,不仅能提升企业各部门的专业管理能力,而且还能促进企业安全、有序、健康、高效地运营,对决策制定发挥重要的辅助作用。然而,确定和下达经营决策的合理性取决于预测这些指标未来变化情况的准确性,由于预测电网企业运行运营指标是面向未来的,且影响指标变化的因素较多,因此预测过程具有很大的随机性和不确定性。结合电网企业的实际情况设计合适的预测方法,将直接关系到预测实施的精度。

目前,常用的传统预测方法有指数平滑法(exponential smoothing, ES)<sup>[1]</sup>、线性回归分析法(linear regression analysis, LRA)<sup>[2]</sup>、时间序列法(time series method, TS)<sup>[3]</sup>等,ES认为时间序列的态势具有稳定性或规则性,所以时间序列可被合理地顺势推延,且它认为最近的过去态势,在某种程度上会持续到未来,所以将较大的权数放在最近的资料上。LRA利用称为线性回归方程的最小平方函数对一个或多个自变量和因变量之间的关系进行建模的一种回归分析。TS则通过编制和分析时间序列,根据时间序列所反映出来的发展过程、方向和趋势,进行类推或延伸,借以预测下一段时间或以后若干年内可能达到的水平。虽然这些传统方法在预测应用中略有成效,但是它们预测对象单一、过度依赖历史数据,且无法考虑到未来预测过程中存在的不确定性因素。

因此,针对以上问题,一些学者尝试将神经网络应用到预测中去,使得预测系统具有一定的智能信息处理能力,取得了不错的预测效果。但是,由于神经网络采用的是经验风险最小化原则,容易陷入局部极小点且收敛速度慢,这极大地限制了该方法在实际过程中的应用。基于支持向量机的预测模型<sup>[4]</sup>很好地克服了神经网络的这一缺点,SVM采用结构风险最小化原则,整个求解过程转化为凸二次规划问题,能得到全局最优解。但是,由于SVM是借助二次规划来求解支持向量,而求解二次规划将涉及 $m$ 阶矩阵的计算( $m$ 为样本的个数),当 $m$ 数目很大时该矩阵的存储和计算将耗费大量的机器内存和运算时间。因此,其不适用于大规模训练样本。

为此,本文提出一种双非凸回归(double non-

convex regression, DNR)算法用于标量电力数据预测。该方法首先将稀疏编码技术<sup>[5]</sup>转化为回归预测应用,再采用 $l_p$ 范数替换原始的重构误差 $l_2$ 范数以及表示系数 $l_1$ 范数等约束,以获得更为灵活的模型目标泛函形式。最后,利用交替方向乘子法(alternating direction method of multipliers, ADMM)<sup>[6]</sup>优化求解目标函数;其中,为获得快速的子问题优化效果,提出一种改进的迭代阈值方法用于求解 $l_p$ 约束子问题,保证全局最优解并可实现并行实施方案。在电网企业运行运营指标真实数据上的实验结果表明该方法切实可行,且具有很高的预测精度。

## 1 双非凸回归算法

稀疏编码技术<sup>[5]</sup>已在模式识别领域得到广泛的应用,其通过误差平方最小化和稀疏性范数约束搜寻目标量的最佳逼近系数。原始的稀疏回归模型可以描述为

$$\min \|x - A\alpha\|_2^2 + \lambda \|\alpha\|_0 \quad (1)$$

式中: $x \in \mathbb{R}^{m \times 1}$ 、 $A \in \mathbb{R}^{m \times n}$ 和 $\alpha \in \mathbb{R}^{n \times 1}$ 分别表示查询向量、观测矩阵和表示系数,伪范数 $\|\cdot\|_0$ 用于计算系数矢量 $\alpha$ 的非零值个数。在分类问题中,式(1)的具体意义为:观测数据 $A$ 包含 $n$ 个 $m$ 维特征样本,未知样本 $x$ 通过表示系数 $\alpha$ 进行重构,要求 $\alpha$ 具有有限个非零元素且对应于同类观测样本。然而,在电力数据预测(或称为回归)问题中,常见的数据如电力产能、运营利润等都是标量值,而影响其具体取值的特征量为矢量形式。因此,本文将式(1)转化为回归问题,即 $x \in \mathbb{R}^{m \times 1}$ 表示 $m$ 个已知目标样本, $A \in \mathbb{R}^{m \times n}$ 表示 $m$ 个观测特征样本,每个样本的维数为 $n$ ,而系数 $\alpha \in \mathbb{R}^{n \times 1}$ 则表示每个特征子元素的重构贡献度。通过训练样本计算得到最优的 $\alpha$ 系数,可用于预测任意未知特征样本 $a$ 为 $y = a^T \alpha$ 。

遗憾的是,受 $l_0$ 伪范数约束影响,最小化式(1)是一个NP难问题,仅能在有限样本集中运行实施<sup>[7]</sup>。一般将之调整为Lasso问题,即通过 $l_1$ 范数替换 $l_0$ 范数

$$\min \|x - A\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (2)$$

已有理论表明,在一定的不连贯条件下<sup>[7]</sup>, $l_1$ 最小化问题很大概率等价于 $l_0$ 最小化问题。而且, $l_1$ 范数是最逼近于 $l_0$ 范数的凸约束形式,有利于模型的优化求解。

此外,式(1)和式(2)中所采用的 $l_2$ 范数最小化重构误差仅适合于高斯分布噪声环境<sup>[8-9]</sup>。为适应特定的样本或特征干扰,常采用等价于拉普

拉斯分布的  $l_1$  范数约束进行误差最小化<sup>[5]</sup>, 即

$$\min \|\mathbf{x} - \mathbf{A}\boldsymbol{\alpha}\|_1 + \lambda \|\boldsymbol{\alpha}\|_1 \quad (3)$$

然而, 当矩阵  $\mathbf{A}$  的不连贯条件无法满足, 或重构误差并不适应于常规的高斯分布或拉普拉斯分布时, 模型 (3) 的求解结果往往是次优的。针对此问题, 本文采用  $l_p$  范数 ( $p \in (0, 1]$ ) 替换式 (3) 中的  $l_1$  范数约束, 即

$$\min \|\mathbf{x} - \mathbf{A}\boldsymbol{\alpha}\|_q + \lambda \|\boldsymbol{\alpha}\|_p^p \quad (4)$$

为便于区分, 在重构误差中以符号  $q$  表示  $l_p$  范数约束。式 (4) 为本文所提回归模型的目标函数, 对误差项和正则项都采用非凸函数约束, 因此称之为双非凸回归算法 (double nonconvex regression, DNR), 其中对  $\boldsymbol{\alpha}$  约束  $l_p$  范数较  $l_1$  更接近于  $l_0$ , 使得重构系数  $\boldsymbol{\alpha}$  具有更强的稀疏性; 对误差约束  $l_p$  范数使之更贴近于椭圆分布<sup>[10]</sup>, 具有较高斯分布和拉普拉斯分布更为宽泛的适用范围。

## 2 模型优化求解

针对常规的单  $l_p$  约束稀疏回归问题, 迭代重加权最小二乘 (IRLS)<sup>[11]</sup>、迭代重加权  $l_1$  最小化 (IRL1)<sup>[12]</sup>、交替方向乘子<sup>[6]</sup>以及迭代阈值收缩 (IST)<sup>[13]</sup>等求解算法都得到了成功应用。然而, 对于式 (4) 所示双非凸约束模型, 所述求解算法都无法直接应用。如果强行将迭代重加权型算法扩展为双加权凸约束形式, 所得解也非常容易陷入局部最优; ADMM 算法能够进行有效地多变量拆分, 但是要求各子优化问题具有闭式解或快速求解策略; IST 具有高效的收敛性能, 但前提需将目标函数中的观测矩阵  $\mathbf{A}$  变换至正交形式。结合所述分析, 本节采用 ADMM 算法和 IST 算法融合策略进行目标函数式 (4) 的求解优化。首先, 通过 ADMM 变量拆分, 获得部分子问题的闭式解; 其次, 提出改进的迭代阈值算法进行非凸子问题优化; 最后, 给出完整的模型求解算法并分析其运算复杂度。

### 2.1 ADMM 变量拆分

考虑到 DNR 模型包含两个非凸  $l_p$  范数约束, 需要引入两个辅助变量用于问题简化, 将式 (4) 转化为等价约束优化式

$$\begin{aligned} \min & \|\mathbf{x} - \mathbf{A}\boldsymbol{\beta}\|_q + \lambda \|\boldsymbol{\beta}\|_p^p \\ \text{s.t. } & \mathbf{x} - \mathbf{A}\boldsymbol{\alpha} = \mathbf{e}, \boldsymbol{\alpha} = \boldsymbol{\beta} \end{aligned} \quad (5)$$

并得到其增广拉格朗日形式

$$\begin{aligned} L_{\mu_e, \mu_\beta}(\mathbf{e}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = & \|\mathbf{e}\|_q^q + \lambda \|\boldsymbol{\beta}\|_p^p + \frac{\mu_e}{2} \|\mathbf{x} - \mathbf{A}\boldsymbol{\alpha} - \mathbf{e}\|_2^2 + \\ & \frac{\mu_\beta}{2} \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_2^2 + \gamma_e^T (\mathbf{x} - \mathbf{A}\boldsymbol{\alpha} - \mathbf{e}) + \gamma_\beta^T (\boldsymbol{\alpha} - \boldsymbol{\beta}) \end{aligned} \quad (6)$$

式中:  $\mu_e, \mu_\beta > 0$  为惩罚参数,  $\gamma_e$  和  $\gamma_\beta$  为拉格朗日乘

子系数。根据 ADMM 变量分离规则, 式 (6) 包含以下迭代步骤:

1) 固定  $\boldsymbol{\alpha}$  和  $\gamma_\beta$ ,  $\boldsymbol{\beta}_{k+1}$  的更新子问题为

$$\begin{aligned} \boldsymbol{\beta}_{k+1} = \arg \min_{\boldsymbol{\beta}} L_{\mu_\beta}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma_\beta) = \\ \frac{\lambda}{\mu_\beta} \|\boldsymbol{\beta}\|_p^p + \frac{1}{2} \|\boldsymbol{\beta} - \boldsymbol{\alpha} - \frac{\gamma_\beta}{\mu_\beta}\|_2^2 \end{aligned} \quad (7)$$

2) 固定  $\boldsymbol{\alpha}$  和  $\gamma_e$ ,  $\mathbf{e}_{k+1}$  的更新子问题为

$$\begin{aligned} \mathbf{e}_{k+1} = \arg \min_{\mathbf{e}} L_{\mu_e}(\mathbf{e}, \boldsymbol{\alpha}, \gamma_e) = \\ \frac{1}{\mu_e} \|\mathbf{e}\|_q^q + \frac{1}{2} \|\mathbf{e} - \mathbf{x} + \mathbf{A}\boldsymbol{\alpha} - \frac{\gamma_e}{\mu_e}\|_2^2 \end{aligned} \quad (8)$$

3) 固定  $\mathbf{e}$ 、 $\boldsymbol{\beta}$ 、 $\gamma_e$  和  $\gamma_\beta$ ,  $\boldsymbol{\alpha}_{k+1}$  的更新子问题为

$$\begin{aligned} \boldsymbol{\alpha}_{k+1} = \arg \min_{\boldsymbol{\alpha}} L_{\mu_e, \mu_\beta}(\mathbf{e}, \boldsymbol{\beta}, \gamma_e, \gamma_\beta) = \\ \frac{\mu_e}{2} \|\mathbf{A}\boldsymbol{\alpha} - \mathbf{x} + \mathbf{e} - \frac{\gamma_e}{\mu_e}\|_2^2 + \frac{\mu_\beta}{2} \|\boldsymbol{\alpha} - \boldsymbol{\beta} + \frac{\gamma_\beta}{\mu_\beta}\|_2^2 \end{aligned} \quad (9)$$

4) 根据计算所得的  $\boldsymbol{\beta}$ 、 $\mathbf{e}$  和  $\boldsymbol{\alpha}_k$ , 更新  $\gamma_e$  和  $\gamma_\beta$

$$\begin{aligned} \gamma_e^{k+1} &= \gamma_e^k + \mu_e (\mathbf{x} - \mathbf{A}\boldsymbol{\alpha} - \mathbf{e}) \\ \gamma_\beta^{k+1} &= \gamma_\beta^k + \mu_\beta (\boldsymbol{\alpha} - \boldsymbol{\beta}) \end{aligned} \quad (10)$$

上述迭代步骤中, 式 (10) 是 ADMM 固有的乘子升级规则。式 (9) 通过微分后可得  $\boldsymbol{\alpha}_{k+1}$  的解析解:

$$\boldsymbol{\alpha}^{k+1} = \mathbf{C}(\mu_e \mathbf{A}^T (\mathbf{x} - \mathbf{e}) + \mathbf{A}^T \gamma_e + \mu_\beta \boldsymbol{\beta} - \gamma_\beta) \quad (11)$$

式中: 在给定  $\mu_e$  和  $\mu_\beta$  的前提下, 逆算子  $\mathbf{C} = (\mu_e \mathbf{A}^T \mathbf{A} + \mu_\beta \mathbf{I})^{-1}$  是常量, 可提前计算并缓存, 加速算法的求解效率。因此, 非凸子问题 (7) 和 (8) 是求解式 (6) 的关键步骤。考虑到  $p$  次  $l_p$  范数约束  $\|\mathbf{x}\|_p^p = \sum x_i^p$  的可叠加性, 式 (7) 和式 (8) 得以分解成独立且并行可解的标量子问题:

$$\min f(\delta) = \frac{1}{2} \|\delta - \sigma\|_2^2 + \lambda \|\delta\|_p^p \quad (12)$$

当  $p=1$  时, 可由经典的软阈值算法<sup>[14]</sup>进行有效求解。针对本文的非凸情况 ( $0 < p < 1$ ), IRLS、IRL1、IST 等求解算法都存在局部次优解的缺陷。如图 1 所示, 当  $\sigma=0.9$ ,  $p=0.2$  且  $\lambda=1$  时, IRLS、IRL1 和 IST 都陷入了局部最小值。为解决该问题, 本文提出一种改进的阈值迭代方法, 在保证高效求解的同时能够获得全局最优值。

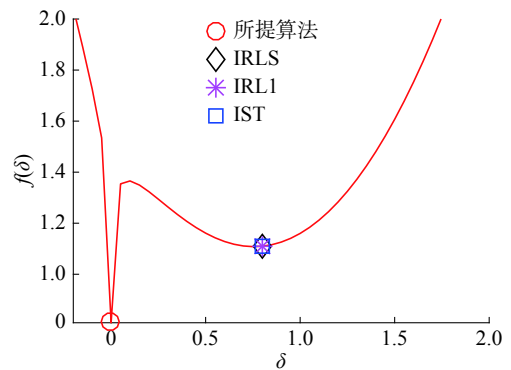


图1 几种算法对典型非凸问题式 (12) 的最优解

Fig. 1 Several algorithms for the optimal solution to the typical nonconvex problem in formula (12)



## 2.2 改进的迭代阈值优化算法

根据式(12)的对称性以及阈值收缩规则<sup>[13]</sup>, 当 $\sigma > 0$ 时, 最优解范围为 $[0, \sigma]$ ; 当 $\sigma < 0$ 时, 最优解范围为 $[\sigma, 0]$ 。不失一般性, 本节仅考虑 $\sigma > 0$ 的情形。设 $p=0.6, \lambda=2.5$ , 图2给出了不同 $\sigma$ 值下的 $f(\delta)$ 最优解情况。由图2中可见,  $f(\delta)$ 的最小值取决于某临界 $\sigma$ 值 $\tau_\sigma$ , 当 $\sigma < \tau_\sigma$ 时,  $\min f(\delta)$ 位于 $\delta=0$ ; 当 $\sigma \geq \tau_\sigma$ 时,  $\min f(\delta)$ 位于 $\delta > 0$ 的某个点。因此, 求解式(12)的核心由 $\tau_\sigma$ 和 $\delta$ 两个关键值确定。

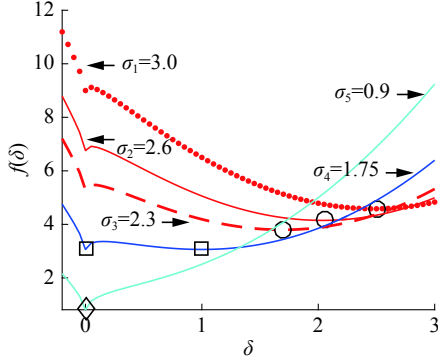


图2 不同 $\sigma$ 值下非凸问题 $f(\delta)$ 的最优解

Fig. 2 The optimal solution of non-convex problem  $f(\delta)$  under different  $\sigma$  values

式(12)的一阶和二阶微分别为

$$f'(\delta) = \delta - \sigma + \lambda p \delta^{p-1} \quad (13)$$

$$f''(\delta) = 1 + \lambda p(p-1)\delta^{p-2} \quad (14)$$

设 $f''(\delta^{(\lambda, p)})=0$ , 可得 $\delta^{(\lambda, p)}=(\lambda p(1-p))^{1/(2-p)}$ 。结合图2可知, 当 $\delta \in (0, \delta^{(\lambda, p)})$ 时,  $f(\delta)$ 是凹函数; 当 $\delta \in (\delta^{(\lambda, p)}, +\infty)$ 时, 则 $f(\delta)$ 是凸函数。进一步, 为保证 $f(\delta)$ 在 $(\delta^{(\lambda, p)}, +\infty)$ 具有最小值, 需满足 $f(\delta^{(\lambda, p)}) \leq 0$ , 文献[13]令 $f(\delta^{(\lambda, p)})=0$ 并计算出 $\tau_\sigma^{\text{IST}}$ 用于迭代阈值求解。然而, 该阈值设法存在问题, 如图1所示, IST计算所得的解满足上述所有规则, 且 $\sigma \geq \tau_\sigma^{\text{IST}}$ 时保证

$$\delta^* - \sigma + \lambda p \delta^{*p-1} = 0 \quad (15)$$

在 $(\delta^{(\lambda, p)}, +\infty)$ 中具有唯一的最小值。然而,  $f(\delta^*)$ 具体取值依然高于 $f(0)$ 。

从图2可见, 存在特定的 $\tau_\sigma$ 使得 $f(\delta^*)=f(0)$ , 当 $\sigma < \tau_\sigma$ 时,  $\delta=0$ 为 $f(\delta)$ 的最小值; 当 $\sigma \geq \tau_\sigma$ 时,  $f(\delta)$ 最小值在 $\delta > 0$ 的某个位置。因此, 正确的阈值 $\tau_\sigma$ 和 $\delta^*$ 计算公式应该为

$$\frac{1}{2}(\delta^* - \tau_\sigma)^2 + \lambda(\delta^*)^p = \frac{1}{2}(\tau_\sigma)^2 \quad (16)$$

$$\delta^* - \tau_\sigma + \lambda p(\delta^*)^{p-1} = 0 \quad (17)$$

将式(17)中的 $\tau_\sigma$ 值代入式(16)可得

$$\delta^{*p}(2\lambda(1-p) - (\delta^*)^{2-p}) = 0 \quad (18)$$

其最优解 $\delta^* \in (\delta^{(\lambda, p)}, +\infty)$ 为 $\delta^*=(2\lambda p(1-p))^{1/(2-p)}$ , 并可进一步计算出 $\tau_\sigma$ 为

$$\tau_\sigma = [2\lambda(1-p)]^{\frac{1}{2-p}} + \lambda p[2\lambda(1-p)]^{\frac{p-1}{2-p}} \quad (19)$$

根据式(15)和式(19), 所提的迭代阈值规则如算法1描述。算法1主要更改了阈值计算策略, 基本步骤与文献[13]类似, 当 $J=2$ 时能够获得令人满意的收敛结果。

结合算法1与ADMM优化框架, 完整的DNR优化步骤按式(7)~(10)循环进行, 具体的收敛条件按文献[6]设定。值得注意的是, 式(7)与式(8)由算法1并行计算实施, 其计算复杂度仅为 $O(n)$ , 而式(9)的计算复杂度在逆算子缓存的前提下为 $O(\max(n^2, nm))$ 。假设ADMM迭代次数为 $t$ , 则完整的算法复杂度为 $O(tn\max(n, m))$ , 远远优于IRLS、IRL1等算法的 $O(n^3)$ 。

### 算法1 改进的迭代阈值规则

输入 参数 $\sigma, \lambda, p, J$ ;

输出  $\delta^*$ 。

1) 按式(19)计算 $\tau_\sigma$ 值;

2) 如 $|\sigma| < \tau_\sigma$ ; 则令 $\delta^*=0$ ;

3)  $f$ 反之, 令 $k=0, \delta^k=\sigma$ ;

4) for  $k=1, 2, \dots, J$

5)  $\delta^{k+1} = |\sigma| - \lambda p(\delta^k)^{p-1}$ ;

6) End

7)  $\delta^* = \text{sgn}(\sigma)\delta^J$ ;

## 3 实验分析

分别对电力企业运行中的全负荷电能输出(兆瓦时)以及运营指标中的月利润总额进行预测。首先对电力企业的产能输出和运营数据进行实证分析, 然后将所提算法与经典的SVM<sup>[15]</sup>、BP神经网络<sup>[16]</sup>和非凸非光滑约束NNR方法<sup>[17]</sup>进行精度对比。

### 3.1 电能输出预测

通过某电力企业复循环动力装置(包括两个燃气涡轮, 一个汽轮机以及两个热回收系统)6年运行数据作为预测样本, 共含该企业全负荷运行674天所产生的9568个采样点, 样本特征包括环境温度(AT)、大气压力(AP)、相对湿度(RH)、排汽压力(V)4个维度。随机选择{10%, 20%, 30%, 40%, 50%}个采集数据作为训练样本, 其余作为测试数据。实验精度由绝对误差均值(MAE)和均方误差(RMSE)两者表示, 其计算式分别为

$$\text{MAE} = (|p_1 - r_1| + |p_2 - r_2| + \dots + |p_n - r_n|)/n \quad (20)$$

$$\text{RMSE} = \sqrt{((p_1 - r_1)^2 + (p_2 - r_2)^2 + \dots + (p_n - r_n)^2)/n} \quad (21)$$

式中: $p$ 和 $r$ 分别为预测值和真实值,  $n$ 为测试样本总量。

表 1 给出了 SVM、BP 神经网络、NNR 以及 DNR 三种对比算法在不同训练样本量下的预测精度对比, 其中 DNR 的参数值  $p=q=1$ 。从表 1 可见, DNR 算法在不同的训练数下都具有最低的误差均值和均方误差值, 展示了更为优秀的预测精度。而且, DNR 算法在 30%~50% 训练样本量下的精度非常接近, MAE 基本稳定在 4.95 左右, 而 BP 神经网络、SVM 和 NNR 在不同样本量下的预测值跨度相对较大, 说明 DNR 具有更高的算法稳定性, 对输入训练样本量要求更小。同时, DNR 计算所得的 AT、AP、RH、 $V$  这 4 个特征表示系数

绝对值分别为 0.767、0.085、0.102、0.426, 即 4 个特征的预测贡献度依次为  $AT>V>RH\approx AP$ , 与文献 [17] 的理论分析结果吻合。此外, SVM、BP、NNR、DNR 3 种算法在 50% 训练量下的完整预测时间分别为 2.91 s、0.48 s、0.32 s 和 0.06 s, 可见 DNR 具有明显更高的运行效率。最后, 表 2 和表 3 分别给出了 DNR 算法在不同  $p$ 、 $q$  值下的预测精度。从中可见, 随着  $p$ 、 $q$  值的优选变化, DNR 的预测精度得以进一步提升, 验证了非凸约束的优越性, 且最优值处于  $p$ 、 $q \in [0.5, 0.8]$ , 与文献 [18] 的理论结论吻合。

表 1 电能输出预测精度对比

Table 1 Comparison of prediction accuracy of energy output

训练数/%	SVM		BP		NNR		DNR	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
10	14.95	16.88	15.01	17.08	18.09	16.21	7.18	9.71
20	9.77	10.91	10.56	11.70	13.30	10.92	5.79	6.67
30	7.21	8.70	7.23	8.53	7.04	9.68	5.01	5.84
40	6.31	7.51	6.10	7.80	6.96	8.09	4.95	5.61
50	5.62	6.56	5.224	7.04	5.89	7.04	4.92	5.46

表 2 DNR 算法不同  $q$  值下的电能输出预测精度

Table 2 Predicting accuracy of energy output in different  $q$  of DNR algorithm

$q$	0.1	0.3	0.5	0.7	0.9	1.0
MAE	5.02	4.80	4.68	4.70	4.92	4.92
RMSE	5.90	5.58	5.18	5.20	5.46	5.46

表 3 DNR 算法不同  $p$  值下的电能输出预测精度

Table 3 DNR algorithm power output prediction accuracy in different  $p$

$p$	0.1	0.3	0.5	0.7	0.9	1.0
MAE	4.97	4.76	4.60	4.56	4.80	4.92
RMSE	5.87	5.43	5.12	5.10	5.38	5.46

3.2 运营数据预测

选取某电网企业自 2013 年 1 月—2014 年 12 月期间的流动资产周转率(次), 购电成本(万元), 可控费用(万元), 货币资金, 主营业务利润

率, 单位资产售电量, 每万元电网资产运行维护成本等真实数据作为训练样本集, 2015 年 1 月—2015 年 12 月期间的运营指标数据作为测试样本集。表 4 为部分训练样本集数据。

表 4 部分训练样本集数据

Table 4 Partial training sample data set

时间	流动资产周转率/次	可控费用/万元	货币资金/万元	单位资产售电/(万千瓦时·亿元 <sup>-1</sup> )
2013/12/1	8.05	2 162 425.116 0	735 383.013 5	14.69
2013/11/1	6.77	1 947 716.502 0	1 168 359.916 0	13.74
2013/10/1	6.23	1 632 149.723 0	1 120 286.637 0	12.54
2013/9/1	5.92	1 627 894.927 0	870 841.852 8	11.35
2013/8/1	5.32	1 357 263.579 0	704 199.691 1	10.11
2013/7/1	4.47	1 140 268.459 0	791 858.709 1	8.50
2013/6/1	3.75	1 015 995.566 0	701 248.501 6	0.13
2013/5/1	2.90	811 738.916 7	862 714.700 0	0.12
2013/4/1	2.36	611 711.335 4	781 078.839 1	0.13
2013/3/1	1.64	425 317.398 7	994 395.365 0	0.11
2013/2/1	1.15	328 205.746 8	799 377.916 8	0.08
2013/1/1	0.73	161 987.012 5	697 097.046 9	0.07

实验中通过 DNR、SVM、BP 神经网络和 NNR 这 3 种预测方法对 2013 年 1 月—2014 年 12 月连续 2 年的月利润总额进行拟合, 对 2015 年 1 月—2015 年 12 月 1 年的月利润总额进行预测, 并比较三者的预测精度, 其结果如图 3~6 以及表 5 所示。

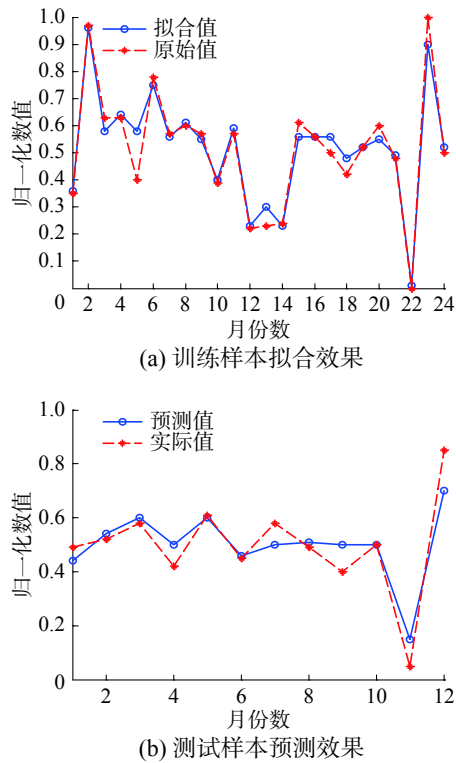


图 3 DNR 对训练样本的拟合效果及对测试样本的预测效果对比

Fig. 3 Comparison of the fitting effect of DNR on training samples and the prediction effect of test samples

由图 3 可知, DNR 对 24 个月数据的拟合值基本贴合原始值的走势, 表现出优秀的拟合能力。图 4 中 SVM 的拟合能力一般, 特别是第 1 个月~第 11 个月的拟合值与原始值相差较大。图 5 中 BP 神经网络对数据的拟合值符合原始值的一般走向, 只是在数值上存在一定程度上的等比例缩小。由图 6 可以看出 NNR 的拟合和预测误差较大。而对于 DNR、SVM 和 BP 神经网络的预测效果, 通过图 3~6 的预测对比可知 DNR 更为接近地预测出了 2015 年 1 月—2015 年 10 月的月利润总额。虽然其对 11 月~12 月这两个月的数据预测不甚理想, 但整体利润趋势与实际值吻合。相较而言, BP 神经网络对这两个月的数据预测较精准, 但 1 月~4 月的预测值却与实际值相差甚远。SVM 的预测值与实际值虽大致在同一数值层上, 但整体预测值离精确点相去较远。此外, NNR 虽与实际值走势相似, 但存在多个严重偏离真实值的预测值。

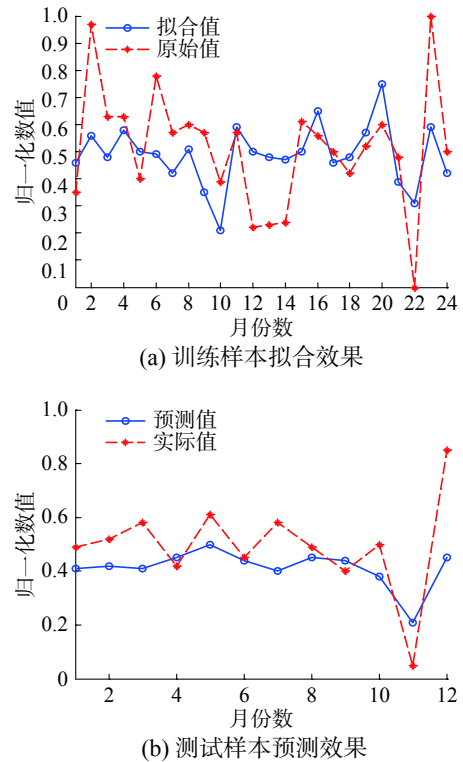


图 4 SVM 对训练样本的拟合效果及对测试样本的预测效果对比

Fig. 4 Comparison of test samples and training samples' fitting effect by SVM

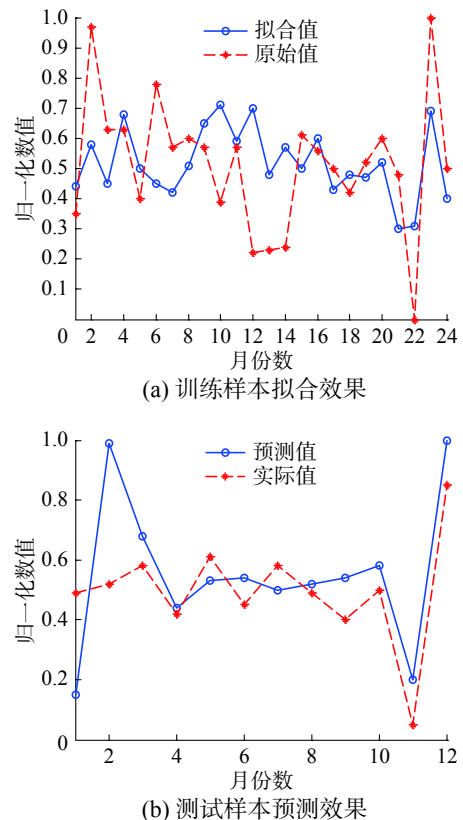


图 5 BP 神经网络对训练样本的拟合效果及对测试样本的预测效果对比

Fig. 5 Comparison of test samples and training samples' fitting effect by BP neural network

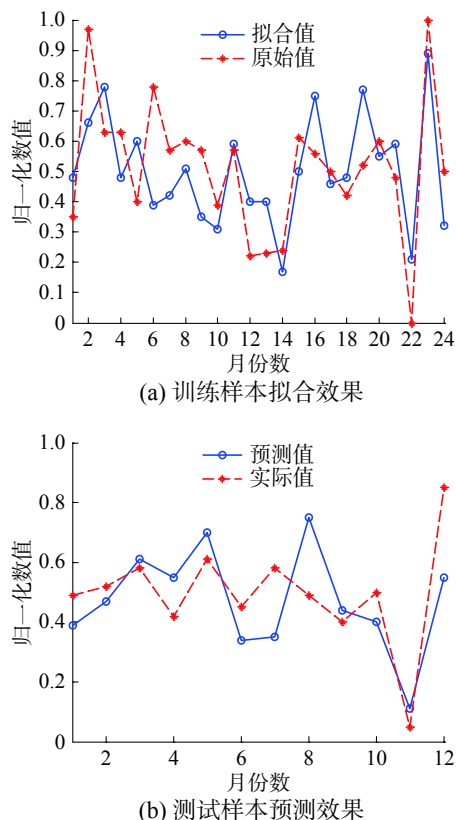


图6 NNR对训练样本的拟合效果及对测试样本的预测效果对比

Fig. 6 Comparison of test samples and training samples's fitting effect by NNR

表5 DNR、SVM、BP和NNR在运营数据中的预测对比

Table 5 Predictive comparison of DNR, SVM, BP and NNR in operational data

方法	MAE		RMSE		运行时间/s
	训练	测试	训练	测试	
DNR	0.082	0.103	0.20	0.20	0.017
SVM	0.196	0.287	0.29	0.36	0.86
BP	0.104	0.182	0.21	0.30	0.94
NNR	0.159	0.203	0.28	0.32	0.89

综上所述, DNR和BP神经网络在运营数据中对训练样本的拟合效果优于SVM。虽然DNR和BP神经网络的拟合效果接近,但对于测试样本的预测结果对比图中明显可以看出, DNR的预测结果最佳。NNR因为存在多个高偏离度的预测点而次于SVM。BP神经网络的预测效果最差。

此外,由表5可知,不论训练样本还是测试样本, DNR的MAE和RMSE值均小于SVM BP神经网络和NNR的误差均值和均方差,再次验证了图3~6的拟合效果和预测效果,并且DNR的运行时间(单位: s)也远远少于SVM、BP神经网络和NNR。

## 4 结束语

电网企业运行和运营数据预测是一个极为复杂的课题,数据采集过程中仪器老化产生的测量误差、人工疏忽导致的漏检误标等因素使得预测过程具有很大的随机性和不确定性。本文提出一种称为非凸回归的预测算法,改进了经典稀疏回归法中的模型约束形式,对重构误差和稀疏系数引入 $l_p(0 < p \leq 1)$ 正则化项约束,使之包含更为稀疏的目标项并具有更为灵活的扩展应用能力。通过交替方向乘子法对该回归模型进行求解,并对其中的子问题提出一种新的阈值优化规则,确保目标函数具有快速的非凸优化求解能力。实验结果表明,与支持向量机BP神经网络和非凸约束算法NNR相比,本文所提方法具有较高的预测精度和更好的预测效果,且运行效率高。

## 参考文献:

- [1] TAYLOR J W. Multi-item sales forecasting with total and split exponential smoothing[J]. Journal of the operational research society, 2011, 62(3): 555-563.
- [2] 彭敏, 张泰玮, 黄佳佳, 等. 基于回归模型与谱聚类的微博突发话题检测方法[J]. 计算机工程, 2015, 41(12): 176-181.
- [3] PENG Min, ZHANG Taiwei, HUANG Jiajia, et al. Microblog sudden topic detection method based on regression models and spectral clustering[J]. Computer engineering, 2015, 41(12): 176-181.
- [4] ARYA F K, ZHANG Lan. Time series analysis of water quality parameters at Stillaguamish River using order series method[J]. Stochastic environmental research and risk assessment, 2015, 29(1): 227-239.
- [5] 丁宏飞, 李演洪, 刘博, 等. 基于BP神经网络与SVM的快速路行程时间组合预测研究[J]. 计算机应用研究, 2016, 33(10): 2929-2932, 2936.
- [6] DING Hongfei, LI Yanhong, LIU Bo, et al. Expressway's travel time prediction based on combined BP neural network and support vector machine approach[J]. Application research of computers, 2016, 33(10): 2929-2932, 2936.
- [7] ZHENG Jianwei, YANG Ping, CHEN Shengyong, et al. Iterative re-constrained group sparse face recognition with adaptive weights learning[J]. IEEE transactions on image processing, 2017, 26(5): 2408-2423.
- [8] CHEN Liang, SUN Defeng, TOH K C. A note on the convergence of ADMM for linearly constrained convex optimization problems[J]. Computational optimization and applications, 2017, 66(2): 327-343.
- [9] CUI Zhuoxu, FAN Qibin. A nonconvex nonsmooth regu-



- larization method for compressed sensing and low rank matrix completion[J]. Digital signal processing, 2017, 62: 101–111.
- [8] YANG Meng, ZHANG Lei, YANG Jian, et al. Regularized robust coding for face recognition[J]. IEEE transactions on image processing, 2013, 22(5): 1753–1766.
- [9] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE transactions on pattern analysis and machine intelligence, 2009, 31(2): 210–227.
- [10] LUO Lei, YANG Jian, QIAN Jianjun, et al. Robust image regression based on the extended matrix variate power exponential distribution of dependent noise[J]. IEEE transactions on neural networks and learning systems, 2017, 28(9): 2168–2182.
- [11] 郑建伟, 黄琼芳, 杨平, 等. 特征加权组稀疏判别投影分析算法[J]. 自动化学报, 2016, 42(5): 746–759.
- ZHENG Jianwei, HUANG Qiongfang, YANG Ping, et al. Feature weighted group sparse discriminative projection algorithm[J]. Acta automatica sinica, 2016, 42(5): 746–759.
- [12] CANDÈS E J, WAKIN M B, BOYD S P. Enhancing sparsity by reweighted  $\ell_1$  minimization[J]. Journal of Fourier analysis and applications, 2008, 14(5): 877–905.
- [13] SHE Yiyuan. Thresholding-based iterative selection procedures for model selection and shrinkage[J]. Electronic journal of statistics, 2009, 3: 384–415.
- [14] CAI Jianfeng, CANDÈS E J, SHEN Zouwei. A singular value thresholding algorithm for matrix completion[J]. SIAM journal on optimization, 2010, 20(4): 1956–1982.
- [15] CHAËBANE N. A novel auto-regressive fractionally integrated moving average-least-squares support vector machine model for electricity spot prices prediction[J]. Journal of applied statistics, 2014, 41(3): 635–651.
- [16] LIU Ke, GUO Wenyang, SHEN Xiaoliu, et al. Research on the forecast model of electricity power industry loan based on GA-BP neural network[J]. Energy procedia, 2012, 14: 1918–1924.
- [17] TÜFEKCI P. Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods[J]. International journal of electrical power and energy systems, 2014, 60: 126–140.
- [18] ZHANG Yong, YE Wanzhou, ZHANG Jianjun. Sparse signal recovery by accelerated  $l_q$  ( $0 < q < 1$ ) thresholding algorithm[J]. International journal of computer mathematics, 2017, 94(12): 2481–2491.

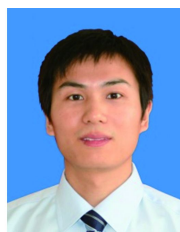
### 作者简介:



王锋华, 男, 1977 年生, 硕士研究生, 主要研究方向为电网数据融合和处理分析技术。牵头国家电网公司科技项目 5 项。出版专著 2 部。



成敬周, 男, 1980 年生, 博士研究生, 主要研究方向为电力系统交直流动态系统、电网数据挖掘与分析应用技术。参与国家电网公司科技项目 4 项, 发表学术论文 10 余篇。



文凡, 男, 1982 年生, 硕士研究生, 主要研究方向为电力系统自动化、电网数据分析技术。参与国家电网公司科技项目 5 项, 发表学术论文 10 余篇。