Applied Soft Computing 129 (2022) 109640

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Time series prediction via elastic net regularization integrating partial autocorrelation

Yanya Xing [a], Dongxi Li [b,*], Chenlong Li [c]

[a] College of Economics and Management, Taiyuan University of Technology, Taiyuan, 030024, China
[b] College of Data Science, Taiyuan University of Technology, Taiyuan, 030024, China
[c] College of Mathematics, Taiyuan University of Technology, Taiyuan, 030024, China

ARTICLE INFO

Article history:
Received 2 December 2021
Received in revised form 3 September 2022
Accepted 8 September 2022
Available online 22 September 2022

Keywords:
Elastic net
Variable selection
Partial autocorrelation coefficient
Time series prediction

ABSTRACT

In this paper, we propose a new elastic net regularization integrating partial autocorrelation coefficients (AEN-PAC) model for time series prediction. This model solves the inaccuracy of variable selection and parameter estimation caused by ignoring the dependence of time series in the adaptive elastic net. The proposed AEN-PAC model adds the partial autocorrelation coefficient to the penalty term of the adaptive elastic net, so that the influence of time on the data series can be well explained. Further, we prove a theorem to demonstrate that our method encourages grouping effects. Then, we convert the optimization problem of the proposed AEN-PAC model into an adaptive lasso model and propose an effective algorithm to solve it. Finally, we conduct a simulation study and empirical analysis on two time series sets. Simulation study shows that the proposed AEN-PAC model selects variable more correctly, compared with other models including Adaptive Elastic Net(AEN), Adaptive Lasso(AL), Elastic Net(EN), and Lasso. In addition, from the perspective of parameter estimation, the parameters estimated by our new model are closer to the real model. For Alibaba stock data and Nike stock data, the prediction errors are 7.07172 and 3.94916 respectively, which are smaller than other models. The results indicating that the proposed AEN-PAC model performs better in time series prediction.
© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Time series data is a series of numerical values distributed in time. Common time series data in life include stock price, advertisement data, temperature change, personal health data, etc. At present, with growing rapidly of the amount of time series data, it has become a hot topic to analyze time series data efficiently and generate business value [1,2]. There are many time series analysis methods in the literature used for time series data prediction. A large part of these methods is based on the autoregressive AR(p) model and its variants [3–5]. The advantages of these models are that they consider the characteristics of randomness, linear volatility, and nonlinear volatility with time in the time series data. Consequently, they can intuitively describe the dynamic changes of time series data.

Generally, in the AR(p) model fitting, the conventional parameter estimation methods such as the least squares estimation and the maximum likelihood estimation have the disadvantage that the variables cannot be adaptively selected. To achieve the purpose of variable selection, the lasso method proposed by Tibshirani [6] is considered in the AR(p) model [7,8]. Specifically, Wang, Li and Tsai [7] and Nardi and Rinaldo [8] conduct data simulation and empirical research on the lasso method, and find that the lasso method has certain applicability in parameter estimation and order selection of time series models. Lasso is easy to delete features that have great influence on dependent variables, which affects the accuracy of variable selection [9–11]. Therefore researchers use the adaptive lasso methods [12], elastic net [13,14] and adaptive elastic net [15] in the regression model fitting to improve prediction accuracy and model interpretability [12,14–20]. The elastic net method achieves the purpose of ridge regression for the selection of important features, but it gives the same penalty to different parameters, and the non-zero parameters obtained in this way will have a large deviation [19]. The adaptive elastic net adds a small penalty to large parameters and a large penalty to small parameters, which can reduce the deviation of non-zero parameter estimates [15]. However, the (adaptive) elastic net method does not take into account the influence of the time in modeling the time sequence and leads to an inaccuracy prediction for time series data.

In this paper, we propose a new elastic net regularization integrating partial autocorrelation coefficients model (AEN-PAC)

* Corresponding author.
E-mail address: dxli0426@126.com (D. Li).

https://doi.org/10.1016/j.asoc.2022.109640
1568-4946/© 2022 Elsevier B.V. All rights reserved.

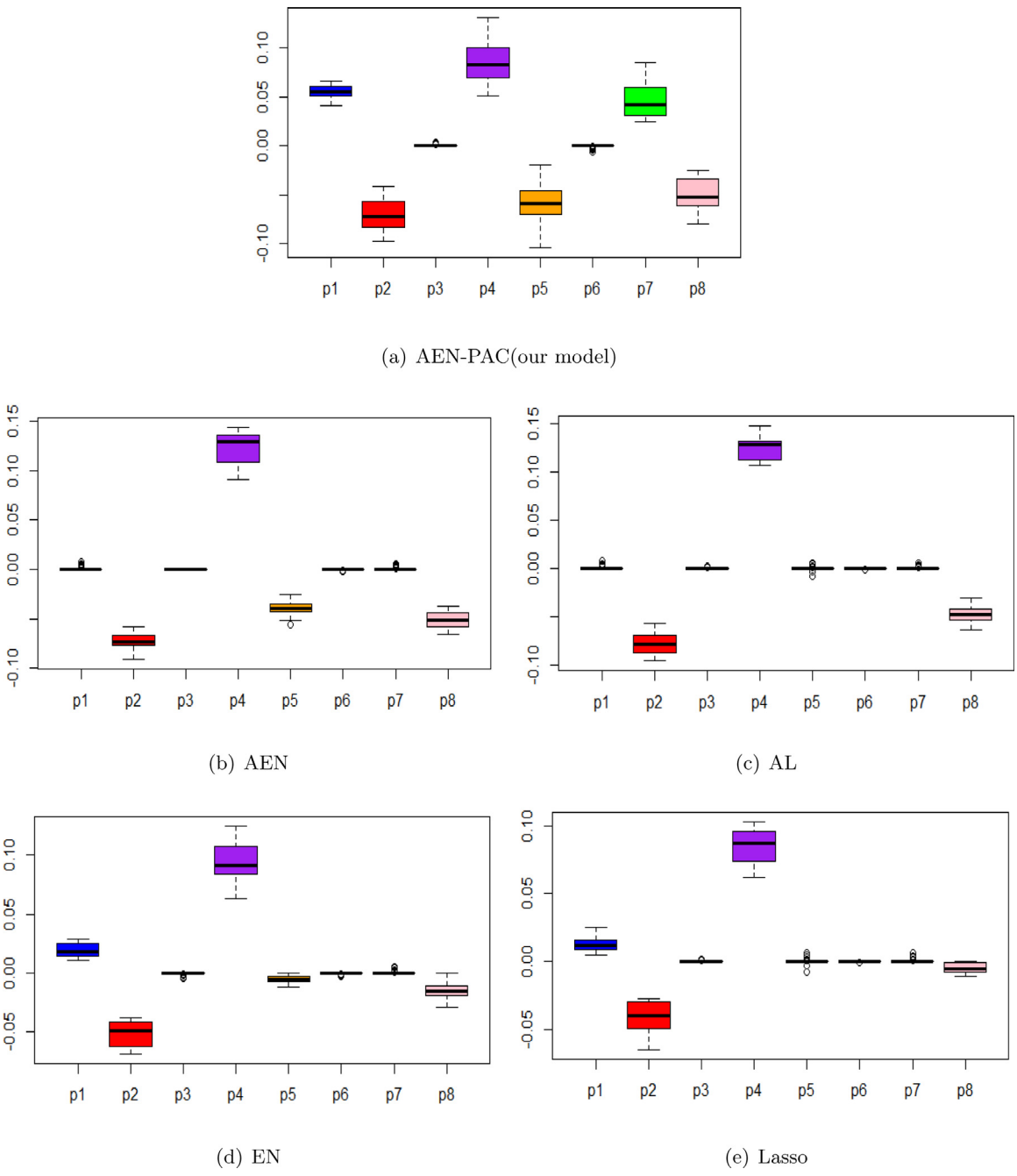


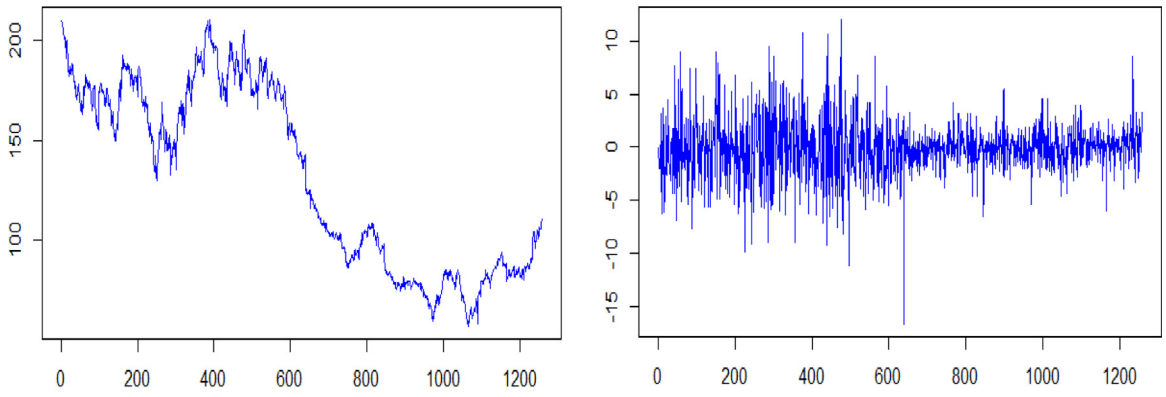
Fig. 3. Box plot of each parameter value for the AEN-PAC model (a), the AEN model (b), the AL model (c), the EN model (d) and the Lasso model (e) when the experiment was repeated 200 times.

3.2. Alibaba stock data

To test the effectiveness of the proposed AEN-PAC model in time series prediction, we performed experiment on Alibaba stock data. We selects Alibaba's daily closing price for a total of 1259 days from December 19, 2014 to December 19, 2019 as sample data. Since the training set and testing set are not directly defined, we use the stock data of the last 30 days as the testing set and the stock data of the previous 1170 days as the training set.

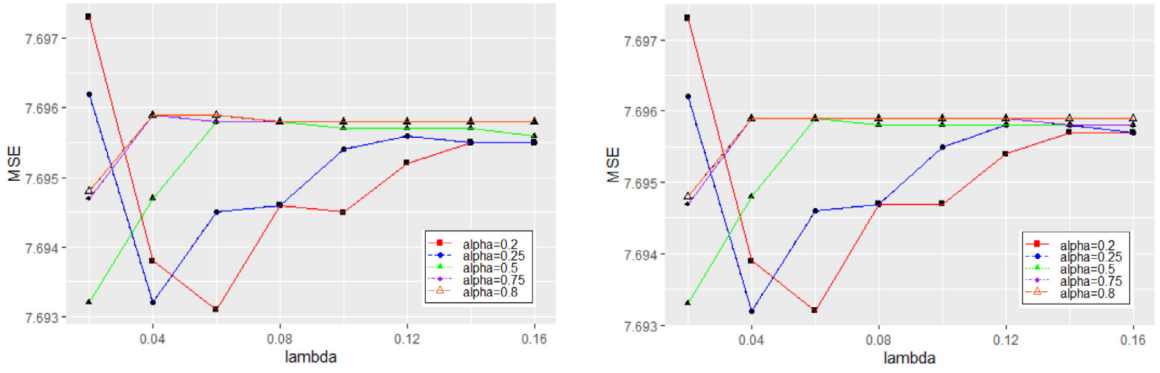
We draw a time chart of the daily closing price of Alibaba from December 19, 2014 to December 19, 2019 for a total of 1259 days (as shown in Fig. 4(a)). It can be seen that the observed value series shows a clear downward trend. For this observation series, ADF unit root test method is used to perform stationarity test, and $P = 0.57566$, which can be regarded as a non-stationary series. Calculating the first-order difference of Alibaba stock data, we obtain the observation data as shown in Fig. 4(b). It can be seen that there is no obvious upward and downward trend in the first-order difference. We use the ADF unit root test method to test its stationarity. Its $P = 2.23866e - 10$, and the model order is 17. Therefore, we will use the first-order difference of Alibaba's closing price to build the model.

Similarly to section 4.1, we first tuned the three hyper-parameters λ, α and η by 5-fold cross validation method. λ takes values from $\{0.02, 0.04, 0.06, 0.08, 0.10, 0.12, 0.14, 0.16\}$ and $\alpha \in \{0.2, 0.25, 0.5, 0.75, 0.8\}$ and $\eta \in \{0.01, 0.05\}$. Fig. 5 shows that the minimum of 5-fold cross validation MSE was obtained when



(a) The daily closing price of Alibaba (b) The first-order difference

Fig. 4. (a)Time chart of daily closing price of Alibaba stock and (b)the first difference of Alibaba stock data on daily closing price.



(a) $\eta = 0.01$ (b) $\eta = 0.05$

Fig. 5. The 5-fold cross validation MSE of the adaptive elastic net with partial autocorrelation coefficients for three parameters on Alibaba data.

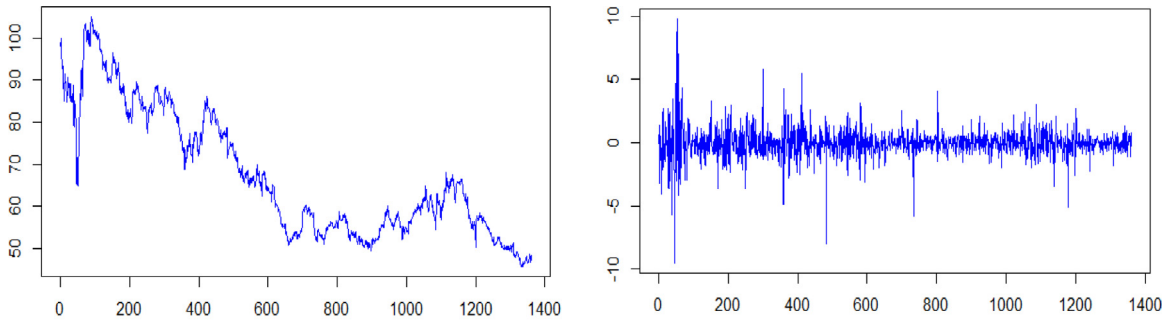
Table 4
Results of variable selection and parameter estimation on Alibaba stock data.

Model	Variable selection and parameter estimation
AEN-PAC (our model)	$y_t = 0.028y_{t-1} - 0.037y_{t-2} - 0.043y_{t-3} + 0.056y_{t-4} - 0.079y_{t-5} + 0.049y_{t-7} - 0.066y_{t-8} - 0.021y_{t-10} - 0.067y_{t-11} + 0.026y_{t-13} + 0.040y_{t-14} + 0.064y_{t-16} + 0.083y_{t-17}$
AEN	$y_t = -0.036y_{t-2} - 0.045y_{t-3} + 0.057y_{t-4} - 0.077y_{t-5} + 0.051y_{t-7} - 0.065y_{t-8} - 0.068y_{t-11} + 0.027y_{t-13} + 0.040y_{t-14} + 0.062y_{t-16} + 0.086y_{t-17}$
AL	$y_t = -0.040y_{t-2} + 0.054y_{t-4} - 0.076y_{t-5} + 0.047y_{t-7} - 0.064y_{t-8} - 0.066y_{t-11} + 0.045y_{t-14} + 0.060y_{t-16} + 0.086y_{t-17}$
EN	$y_t = 0.021y_{t-1} - 0.010y_{t-2} - 0.002y_{t-3} + 0.018y_{t-4} - 0.045y_{t-5} + 0.014y_{t-7} - 0.029y_{t-8} - 0.031y_{t-11} + 0.003y_{t-13} + 0.012y_{t-14} + 0.031y_{t-16} + 0.056y_{t-17}$
Lasso	$y_t = 0.014y_{t-1} - 0.003y_{t-2} + 0.010y_{t-4} - 0.038y_{t-5} + 0.007y_{t-7} - 0.022y_{t-8} - 0.024y_{t-11} + 0.006y_{t-14} + 0.025y_{t-16} + 0.051y_{t-17}$

$\lambda = 0.06$ and $\alpha = 0.2$. However, the least MSE of 7.6931 in Fig. 5(a) is less than 7.6932 in Fig. 5(b). Therefor, we take $\lambda = 0.06$, $\alpha = 0.2$ and $\eta = 0.01$ in the experiment.

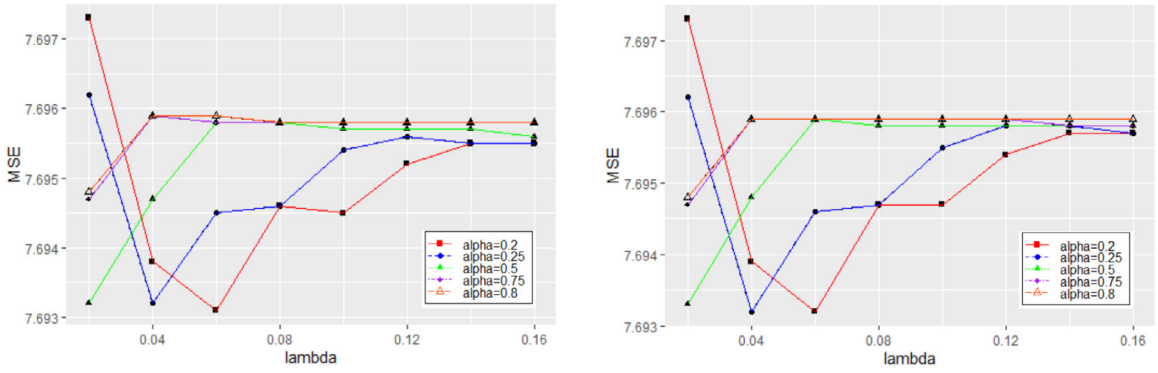
We use AEN-PAC, AEN, AL, EN and Lasso models to perform variable selection and parameter estimation on Alibaba stock data. The results obtained are shown in Table 4. It can be seen from the tables that the number of variables selected by the five models is 13, 11, 9, 12 and 10 respectively. Although we proposed AEN-PAC model selects more variables, according to simulation studies, AEN, AL, EN and lasso models are easy to delete important variables.

We respectively calculate the fitting error and prediction error of the five models in Table 4 on Alibaba stock data, and the results are shown in Table 5. The fitting error and prediction error of AEN-PAC model are 7.11024 and 7.07172 respectively. It can be clearly seen that the two errors of the AEN-PAC model are the smallest. This shows that our model has advantages in variable selection and parameter estimation of time series data. Therefore,



(a) The daily closing price of Nike (b) The first-order difference

Fig. 6. (a)Time chart of daily closing price of Nike stock and (b)the first difference of Nike stock data on daily closing price.



(a) $\eta = 0.01$ (b) $\eta = 0.05$

Fig. 7. The 5-fold cross validation MSE of the adaptive elastic net with partial autocorrelation coefficients for three parameters on Nike data.

Table 5
Results of error calculation.

Model	Fitting errors	Prediction errors
AEN-PAC(our model)	7.11024	7.07172
AEN	7.29708	7.25377
AL	7.31776	7.37908
EN	7.37329	7.39953
Lasso	7.41261	7.43351

using the AEN-PAC model to predict future timing value is more accurate.

3.3. Nike stock data

We also do experiment with the closing price of Nike stock from January 1, 2015 to May 31, 2020. As in Section 3.2, we take the data from the first 30 days as the testing set and the remaining 1170 days as the training set. Fig. 6(a) is the daily closing price sequence chart and Fig. 6(b) is the first-order difference sequence diagram of daily closing price. The ADF root test of the first-order difference data shows that $P = 9.607e - 27$, and the model order is 8. Therefore, we use the first-order difference of the daily closing price of Nike stock to establish the regression model.

Similarly to Sections 3.1 and 3.2, we tuned the three hyper-parameters λ, α and η by 5-fold cross validation method before conducting the contrast experiment. λ also takes values from $\{0.02, 0.04, 0.06, 0.08, 0.10, 0.12, 0.14, 0.16\}$ and $\alpha \in \{0.2, 0.25, 0.5, 0.75, 0.8\}$ and $\eta \in \{0.01, 0.05\}$. Fig. 7 shows that the minimum of 5-fold cross validation MSE was obtained when $\eta = 0.01, \lambda \in \{0.06, 0.08\}$ and $\alpha \in \{0.2, 0.25\}$. Therefor, we take $\alpha = 0.2$ or 0.5, one value of $\{0.06, 0.08\}$ for λ and $\eta = 0.01$ in the experiment.

We use AEN-PAC, AEN, AL, EN and Lasso models to select variables and estimate parameters with the first-order difference data of Nike stock. The results are shown in Tables 6 and 7. It can be seen that the number of variables selected by the five models are 7, 6, 4, 6 and 4 respectively. And the fitting error and prediction error of AEN-PAC model are 1.48203 and 3.94916 respectively. So the new model proposed in this paper has the smallest fitting error and prediction error. This once again shows that using the AEN-PAC model to predict future timing value is more accurate.

4. Conclusion

This paper proposes a novel algorithm: an adaptive elastic net regularization integrating partial autocorrelation coefficients. It solves the problems that the elastic net does not consider the sensitivity to time sequence and ignores the importance of lag order to model fitting when making time series prediction. In addition, we also developed a fast solution algorithm, which converts a complex elastic net model into an adaptive lasso model and then solves it. Simulation study shows that the parameters estimated by our new model are closer to the real model from the perspective of parameter estimation. For Alibaba stock data and Nike stock data, the prediction errors are 7.07172 and 3.94916 respectively, which are smaller than other models. The results

Table 6
Results of variable selection and parameter estimation on Nike stock data.

Model	Variable selection and parameter estimation
AEN-PAC (our model)	$y_t = 0.039y_{t-2} + 0.017y_{t-3} - 0.097y_{t-4} + 0.038y_{t-5} - 0.066y_{t-6} + 0.106y_{t-7} - 0.152y_{t-8}$
AEN	$y_t = 0.038y_{t-2} - 0.097y_{t-4} + 0.039y_{t-5} - 0.066y_{t-6} + 0.105y_{t-7} - 0.152y_{t-8}$
AL	$y_t = -0.083y_{t-4} - 0.057y_{t-6} + 0.098y_{t-7} - 0.145y_{t-8}$
EN	$y_t = 0.005y_{t-2} - 0.057y_{t-4} + 0.004y_{t-5} - 0.036y_{t-6} + 0.071y_{t-7} - 0.114y_{t-8}$
Lasso	$y_t = -0.043y_{t-4} - 0.025y_{t-6} + 0.059y_{t-7} - 0.101y_{t-8}$

Table 7
Results of error calculation.

Model	Fitting errors	Prediction errors
AEN-PAC(our model)	1.48203	3.94916
AEN	1.48247	4.08423
AL	1.48793	4.25075
EN	1.49505	4.26817
Lasso	1.50323	4.48233

indicating that the proposed AEN-PAC model performs better in time series prediction.

In general, Lasso and AL model are easier to delete important variables than AEN-PAC model. This is because Lasso only adds $\ell 1$-penalty term, which makes the coefficients of some features smaller, and even some coefficients with smaller absolute values directly become 0. Elastic net adds $\ell 2$-penalty term on the basis of lasso, so the deletion of important variables is avoided. However, EN and AEN do not consider the correlation of time series, so it is easy to delete the lag order related to the current time. Therefore, the AEN-PAC model we proposed has great advantages in the variable selection of time series data. Besides, the prediction ability of AEN-PAC model is better than that of AEN, AL, EN and Lasso models. Because the AEN-PAC model can guarantee the minimum prediction error, which can make the investors avoid the risk and has great commercial value.

CRediT authorship contribution statement

Yanya Xing: Data curation, Writing – original draft. **Dongxi Li:** Conceptualization, Supervision, Writing – review & editing. **Chenlong Li:** Methodology, Investigation, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11571009), and Applied Basic Research Programs of Shanxi Province (Grant No. 201901D111086).

References

[1] R. Sikarwar, M. Appalaraju, The impact of stock market performance on economic growth in India, Asian J. Res. Banking Finance 8 (2018) 49–57.
[2] L. Wang, L. Liu, Long-range correlation and predictability of Chinese stock price, Physica A 549 (2020) 124384.
[3] Z. Zhou, M. Gao, Q. Liu, H. Xiao, Forecasting stock price movements with multiple data sources: Evidence from stock market in China, J. R. Stat. Soc. 542 (2020) 123389.
[4] Y. Dong, S. Li, X. Gong, Time series analysis: An application of ARIMA model in stock price forecasting, in: Proceedings of the 2017 International Conference on Innovations in Economic Management and Social Science, 2017.
[5] G. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing 50 (2003) 159–175.
[6] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B Stat. Methodol. 58 (1996) 267–288.
[7] H. Wang, G. Li, C.L. Tsai, Regression coefficient and autoregressive order shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B Stat. Methodol. 69 (2007) 63–78.
[8] Y. Nardi, A. Rinaldo, Autoregressive process modeling via the lasso procedure, J. Multivariate Anal. 102 (2011) 528–549.
[9] Y. Liu, P. Zeng, L. Lin, Generalized $\ell 1$-penalized quantile regression with linear constraints, Comput. Statist. Data Anal. 142 (2020).
[10] Y. Zhao, H. Bondell, Solution paths for the generalized lasso with applications to spatially varying coefficients regression, Comput. Statist. Data Anal. 142 (2020).
[11] H. Shi, S. Liu, J. Chen, X. Li, Predicting drug-target interactions using lasso with random forest based on evolutionary information and chemical structure, Genomics 111 (2019) 1839–1852.
[12] J. Li, Q. Lu, Y. Wen, Multi-kernel linear mixed model with adaptive lasso for prediction analysis on high-dimensional multi-omics data, Bioinformatics 36 (2020) 1785–1794.
[13] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. R. Stat. Soc. Ser. B Stat. Methodol. 67 (2005) 301–320.
[14] F. Zhang, K. Sun, X. Wu, A novel variable selection algorithm for multi-layer perceptron with elastic net, Neurocomputing 361 (2019) 110–118.
[15] H. Zou, H.H. Zhang, On the adaptive elastic net with a diverging number of parameters, Ann. Statist. 37 (2009) 1751–1933.
[16] K.B. Gregory, D. Wang, C.S. McMahan, Adaptive elastic net for group testing, Biometrics 75 (2019) 13–23.
[17] Y. Wang, X.G. Yang, Y. Lu, Informative gene selection for microarray classification via adaptive elastic net with conditional mutual information, Appl. Math. Model. 71 (2019) 286–297.
[18] X. Xin, J. Hu, L. Liu, On the oracle property of a generalized adaptive elastic-net for multivariate linear regression with a diverging number of parameters, J. Multivariate Anal. 162 (2017) 16–31.
[19] J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its oracle properties, J. Amer. Statist. Assoc. 96 (2011) 1348–1360.
[20] H. Zou, The adaptive lasso and its oracle properties, J. Amer. Statist. Assoc. 101 (2006) 1418–1429.
[21] J. Friedman, T. Hastie, H. Hofling, R. Tibshirani, Pathwise coordinate optimization, Ann. Appl. Stat. 1 (2007) 302–332.