Trends in Social Sciences and Humanities Research
ISSN: 2959-9938
DOI: 10.61784/tsshr3011

ANALYZING REGIONAL ECONOMIC INFLUENCING FACTORS BASED ON DIFFERENT CONTRACTION METHODS

CaiYun Peng
School of Mathematics and Statistics, Guangxi Normal University, Guilin 541006, Guangxi, China.
Corresponding Email:pcyjying@163.com

Abstract: Currently, the biggest issue facing social development in China is the imbalance and lack of coordination in regional economic development, with multiple factors influencing regional economies. In order to scientifically measure these influencing factors, we established a multiple linear regression model based on data from 31 provinces nationwide in 2020. We fitted the model using four different solution methods – traditional OLS estimation, ridge regression, Lasso, and elastic net estimation. By comparing the fitting effects of different models, we further analyzed the merits and demerits of various contraction methods, and found that elastic net exhibits excellent performance in terms of prediction accuracy and model simplicity. Lastly, based on the results from elastic net estimation, we conducted an analysis of the factors affecting regional economic development.
Keywords: Regional Economic Development; OLS; Ridge Regression; Lasso; Elastic Net

1 INTRODUCTION

In recent years, China has synergistically leveraged technological innovation and market value realization, providing robust support for economic development. Studying the relationship between technology and regional economic growth holds significant practical importance[1]. Factors influencing regional economic disparities in China, as identified by Wang Anqi[2] and Cao Haibo[3], include geographic location, capital factors, technological progress, degree of economic openness, and policy institutions. Zhu Xuexin et al. utilized a broad Cobb-Douglas production function to examine the contributing factors to national economic growth and the contribution rates of different indicators of technological progress[4]. Wang Wei et al. conducted empirical research on the impact of technological innovation capability on regional economic development using a comprehensive evaluation method[5]. Wang Xu et al. focused on the quantitative factors of the input-output process of technological innovation in their research[6]. Song Meizhe and Li Mengsu found spatial gradient distribution patterns such as "low in the west, high in the east" and "low inland, high coastal" in their study of the coupling coordination of higher education, technological innovation, and economic development[7]. Zhang Zhiruo et al. studied the spatiotemporal coupling characteristics of technology finance and regional economic development in different regions[8]. Hong Mingyong employed econometric analysis to investigate the relationship between technological innovation and regional economic growth nationwide[9]. Foreign scholar Lukas Hogenschurz researched the impact of science, technology, and innovation on economic and social benefits, emphasizing the importance of converting knowledge and innovation outcomes into societal and economic entities[10].
In above, building upon existing research, this study selects ten relevant factors influencing regional economic development—namely, education, technology, policy institutions, trade openness, urbanization rate, labor input, geographical location, environment, among others—to construct a multiple linear regression model. Various methods for solving multiple linear regression models are employed to select the optimal model. Subsequently, a detailed analysis of each factor is conducted, and corresponding recommendations are proposed to relevant departments.
Common estimation methods for multiple regression models after regularization include ridge regression and Lasso. Both of them reduce the risk of overfitting by adding a penalty term to the loss function and reduce the information redundancy caused by multicollinearity between independent variables[11]. Elastic net regression is a combination of ridge regression and Lasso regression, which can solve some limitations of ridge regression and Lasso regression[12].
Using traditional Ordinary Least Squares (OLS) estimation initially to fit, test, correct, and evaluate the research object, followed by separate applications of Ridge Regression, Lasso, and Elastic Net, is advantageous for finding the best model and assessing the research object. Improving methods for solving multiple linear regression models like OLS, Ridge Regression, and Lasso can effectively enhance model prediction accuracy, prediction efficiency, and model simplicity[13]. Scaling the original Elastic Net further allows achieving optimal results. For each fixed parameter, the Elastic Net problem can be efficiently solved using algorithms designed for solving the Lasso problem. Compared to solving the Lasso problem under equivalent circumstances, Elastic Net also reduces computational speed [12].

2 PREPARATORY KNOWLEDGE

Before establishing the model, it is necessary to introduce relevant data and models.

2.1 Selection of Indicators

© By the Author(s) 2024, under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).

Based on references, most scholars agree that factors such as human capital, education, resource and environmental quality, geographic location, technological progress, degree of economic openness, and policy institutions significantly influence regional economic development. Different scholars select various indicators to measure regional economic development levels across different aspects. Sturm Thomas argues that no single rational theory can provide a satisfactory explanation for every aspect [10]. Therefore, this paper synthesizes several references and selects 10 factors from the following aspects to analyze the factors influencing regional economic disparities. The selected indicators are denoted as $x_1, x_2, \cdots, x_{10}$ respectively, and are summarized in tabular form as follows(Table 1).

Table 1 Summary of Indicators and their Symbol Explanations

Primary indicator	Secondary indicator	Symbol
Economic development level	Gross Regional Product (GRP)	y
Education	Local fiscal expenditure on education	x_1
technology	Number of domestic patent applications accepted	x_2
policy systems	Local government general budget expenditure	x_3
	Proportion of fiscal revenue to GDP	x_4
trade openness	Total import and export volume divided by GDP	x_5
urbanization level	Urbanization rate	x_6
labor input	Urban employed population	x_7
geographical location	Longitude	x_8
	Latitude	x_9
environment	Total wastewater discharge (ten thousand tons)	x_{10}

The data for this study is sourced from the "China Statistical Yearbook 2021" [14].

2.2 Multiple Regression Model

The factors that typically influence the dependent variable are more than one, so it is generally necessary to establish a multiple regression model. Its general model is as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon_i \,, \tag{1}$$

Where Y is the dependent variable, $X_1, X_2, \cdots, X_k$ is the independent variable, $\beta_0, \beta_2, \cdots, \beta_k$ represent the parameters of the independent variables, and ε_i is the error term.

In a typical linear regression model, p predictor variables $x_1, x_2, \cdots, x_p$ are first specified, and the response Y is generated by model(1). The ultimate goal of linear regression is to minimize the gap between predicted values of the model and actual values. The loss function for a typical multiple linear regression model is:

$$L(\beta) = \|y - X\beta\|^2 \tag{2}$$

The loss function is typically solved using OLS (Ordinary Least Squares) for parameter estimation. This involves finding values that minimize the sum of squares of differences between actual values and model-predicted values, serving as estimates. The objective function of the loss function at this point is:

$$\hat{\beta} = \arg\min_{\beta} \|y - X\beta\|^2 \tag{3}$$

However, OLS often performs poorly in prediction and explanation. Therefore, many scholars have proposed adding penalties to the loss function based on OLS.

2.3 Contraction Methods

2.3.1 Ridge regression

Ridge regression minimizes the sum of squared residuals while constrained by the L2 norm of the coefficients, which reduces the residual sum of squares and prevents coefficients from becoming too large. As a form of continuous shrinkage method, ridge regression achieves improved predictive performance through a bias-variance trade-off. However, ridge regression does not produce a parsimonious model because it always retains all predictor variables in the model.

The loss function for ridge regression estimation, which includes an L2 norm penalty, is:

$$\|y - X\beta\|^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \tag{4}$$

2.3.2 Lasso

Lasso is another estimation method that adds a penalty to least squares estimation by applying an L1 penalty to regression coefficients. Lasso performs both continuous shrinkage and automatic variable selection, enabling model simplification. The loss function of Lasso is as follows:

$$\|y - X\beta\|^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{5}$$

Ridge regression and Lasso have certain limitations when addressing certain problems. For example:

1) In the case where p>>n, due to the nature of convex optimization problems, Lasso can select at most n variables before saturating, hence it cannot handle p>>n scenarios.
2) If there is a group of variables with very high pairwise correlations, Lasso tends to arbitrarily select one variable from this group without specifically considering which one to choose.
3) In the typical scenario where n>p, if there are highly correlated predictor variables, empirical observations suggest that Ridge regression generally outperforms Lasso in terms of predictive performance.

2.3.3 Original Elastic Net

To address these issues, this paper considers a new regularization technique—Elastic Net. Similar to Lasso, Elastic Net can select groups of correlated variables. Simulation studies and practical data examples indicate that Elastic Net often outperforms Lasso in terms of prediction accuracy. Elastic Net combines the features of Lasso regression and Ridge regression by adding two penalties to the loss function of Ordinary Least Squares (OLS) estimation, namely:

$$L(\lambda_1, \lambda_2, \beta) = \|y - X\beta\|^2 + \lambda_2 \|\beta\|^2 + \lambda_1 |\beta|_1, \tag{6}$$

The paper transforms the Elastic Net problem, which involves two penalties, into a Lasso-type problem with a single penalty by a simple transformation. This transformation extends the sample size from n to n+p, overcoming the first limitation of Lasso. The transformed penalty function of Elastic Net after this transformation is:

$$L(\gamma, \beta^*) = \left\|y^* - X^*\beta^*\right\|^2 + \gamma \left|\beta^*\right|_1, \tag{7}$$

The Elastic Net estimation also overcomes some limitations of both Lasso regression and Ridge regression. However, for each fixed λ, it appears to induce twice the amount of shrinkage, which may not effectively reduce substantial variance and introduces unnecessary additional bias. Therefore, the paper aims to improve the predictive performance of Elastic Net by correcting this double shrinkage issue.

2.3.4 The Elastic Net

After scaling the coefficients of the original Elastic Net estimation through a proportional transformation, the corrected Elastic Net estimation maintains the variable selection properties of the original Elastic Net. This scaling is the simplest method to eliminate excessive shrinkage. The scaled Elastic Net estimator is given by:

$$\hat{\beta}(elastic \quad net) = (1+\lambda_2)\hat{\beta}(naive \quad elastic \quad net), \tag{8}$$

After scaling by $1+\lambda_2$, Elastic Net automatically achieves optimal balance between maximum and minimum and resolves the issue of excessive shrinkage tendencies in the original Elastic Net in regression problems.

3 SIMULATION EXAMPLE OF REGIONAL ECONOMIC INFLUENCING FACTORS ANALYSIS IN CHINA

Due to the inconsistent units among the research data, we standardized the data to remove the dimensional effects, and then established the following multiple linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_{10} x_{10i}, \tag{9}$$

3.1 Traditional OLS Regression

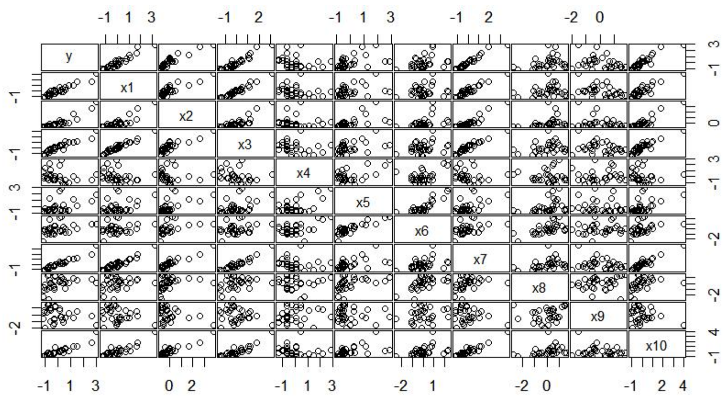

Figure 1 Scatter Plot of Variables in Regional Economics

Examining the correlations between variables reveals(Figure 1) that the dependent variable shows a clear linear relationship with some independent variables, while the linear relationship with other independent variables is not clear. There are also significant correlations among some independent variables. Therefore, the model we established may suffer from multicollinearity. To further assess whether the model exhibits multicollinearity, we calculated the variance inflation factors (VIFs) for each independent variable(Table 2).

Table 2 VIF Values for Each Independent Variable

Coefficients	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
VIF	45.127	13	38.524	2.837	9.385	6.606	48.195	2.051	1.5	14

From Table 2, it is evident that the VIF values for most variables are well above 10, indicating the presence of multicollinearity in the model. To assess the fitting and prediction effectiveness of the model, this study divides the observations into two parts, with 70% used as the training set and 30% as the test set.

3.1.1 Model estimation

Firstly, the fit using all variables in the ordinary least squares estimation was examined. The results indicate that only one independent variable is significant, with a model fit of 98.27%. However, the model shows signs of overfitting. It is necessary to address multicollinearity issues and revise the model accordingly.

3.1.2 Addressing multicollinearity

This study employs the most commonly used stepwise regression method to address the model. The results of stepwise regression analysis indicate that when using $x_2, x_3, x_4, x_6, x_{10}$ as coefficients in the regression equation, the minimum AIC value is -65.86. The remaining variables are examined for their impact on the model's fit. The fitting results indicate that all variables except x_{10} are significant. Therefore, variable x_{10} is excluded to optimize the model, and the model's fit is re-evaluated(Table 3).

Table 3 Refining the OLS Estimation Results

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.06210	0.04537	-1.369	0.191247
x_2	0.36036	0.08442	4.269	0.000673***
x_3	0.56689	0.08084	7.012	4.19e-06 ***
x_4	-0.32878	0.06378	-5.155	0.000118***
x_6	0.29429	0.07632	3.856	0.001555**

Signif.codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

At this point, the t-tests for the parameters of each selected independent variable in the model have all passed, indicating their significance. The model's F-test has also been successful, confirming its overall significance. The standard error of residuals is 0.1768, with an R-squared of 97.86%, indicating a good fit of the model.

3.1.3 Model validation

Next, the model is tested. Using the Breusch-Pagan test to check for heteroscedasticity, the calculation shows that the model does not exhibit heteroscedasticity. Further, plotting the residual plot and regression scatter plot(Figure 2) both indicate that the model fits well.

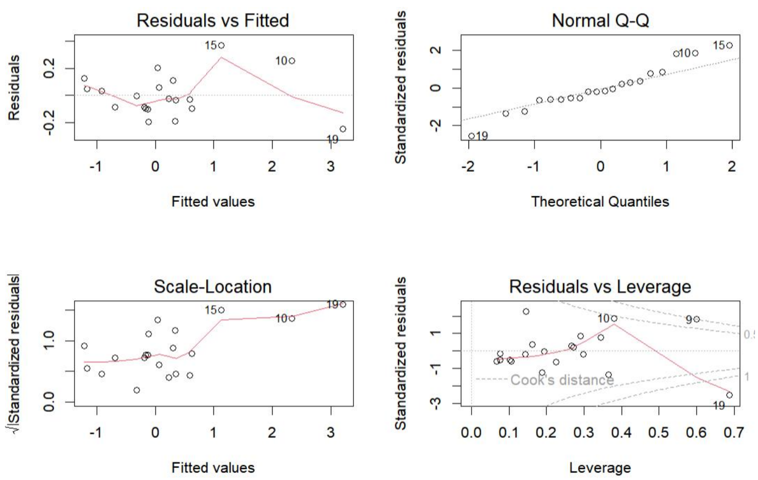

Figure 2 Residuals and Regression Plot of OLS Estimation

At this point, the root mean square error (RMSE) of the prediction sample is 0.377, with a coefficient of determination R-squared of 88.87%. The prediction errors for the sample are relatively large, but the overall fit is good. Therefore, under traditional OLS estimation, the regression equation obtained using standard methods is:

$$y = -0.062 + 0.36x_2 + 0.567x_3 - 0.329x_4 + 0.29x_6 \tag{10}$$

Among these, variables representing technological capability, policy institutions, and urbanization level remain in the model. The coefficient of the variable representing the ratio of fiscal revenue to GDP is negative, indicating an inverse relationship with the dependent variable, contrary to reality. The omitted education expenditure forms the basis for cultivating highly skilled technology personnel, and the labor force is a crucial factor influencing economic development. These variables are indispensable factors in measuring regional economic development. Therefore, despite the good model fit and prediction error of the multiple linear regression model based on traditional OLS estimation methods, its effectiveness and reliability are low for sparse models with small sample sizes.

3.2 Ridge Regression

Ridge regression is a biased estimation regression method specifically used for analyzing collinear data. Its essence lies in shrinking the coefficients of variables with less importance to the dependent variable towards zero, rather than completely excluding them from the model. This trade-off sacrifices some information and precision to obtain regression coefficients that are more realistic and reliable, thus improving upon least squares estimation.

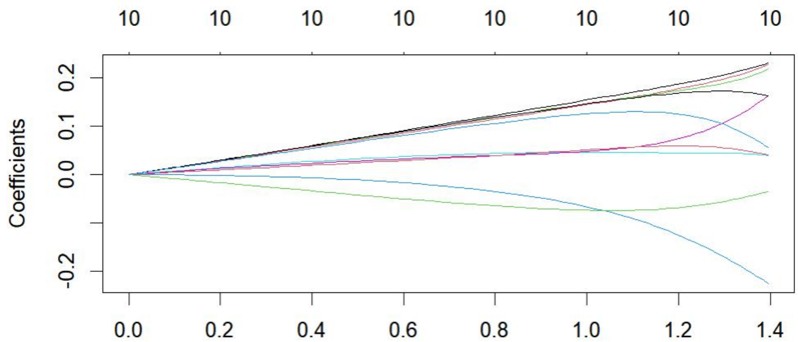

Figure 3 Ridge Trace Plot

From the ridge trace plot(Figure 3), it is observed that the magnitude of curve changes increases with the change in λ, and the curve tends to become unstable. It is challenging to directly determine the appropriate ridge parameter λ value from this plot. Therefore, we opt for ten-fold cross-validation to select the optimal ridge parameter value(Figure 4).

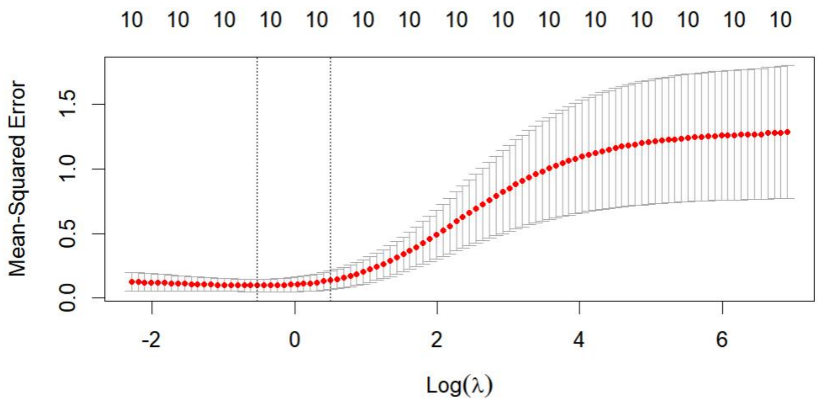

Figure 4 Ridge Regression Ten-fold Cross-Validation Plot

The vertical lines (in Figure 4) correspond to two λ values of 0.5924 and 1.6485. Under the condition of choosing the minimum λ value, the coefficients of the model variables are as follows(Table 4):

Table 4 Ridge Regression Estimation Results

Coefficients	Intercept	x_1	x_2	x_3	x_4	x_5
Estimate	0.01572446	0.18040776	0.1703356	0.16814676	-0.11132384	0.04487903
Coefficients	x_6	x_7	x_8	x_9	x_{10}	
Estimate	0.06605877	0.16508513	0.05887699	-0.071032	0.12825863	

Upon calculation, the root mean square error (RMSE) and coefficient of determination R-squared for predicting sample data are 0.1707 and 94.12%, respectively. After testing, the significance of each parameter in the ridge regression has been confirmed. Therefore, the fitted regression equation under ridge regression is:

$$y = 0.02 + 0.18x_1 + 0.17x_2 + 0.17x_3 - 0.11x_4 + 0.04x_5 + 0.07x_6 + 0.17x_7 + 0.06x_8 - 0.07x_9 + 0.13x_{10} \quad (11)$$

where except for x_4 and x_9, the coefficients of the remaining independent variables are positive. Based on the coefficients of the model's independent variables, the root mean square error of prediction, and the coefficient of determination, ridge regression demonstrates higher predictive accuracy compared to traditional OLS estimation. However, when all independent variables are included, the coefficient of determination is lower than that of OLS estimation.

3.3 Lasso

Lasso can shrink some regression coefficients, meaning it forces the sum of the absolute values of the coefficients to be less than a fixed value and directly sets some coefficients of independent variables to zero. Therefore, it possesses the characteristic of promoting model parsimony.

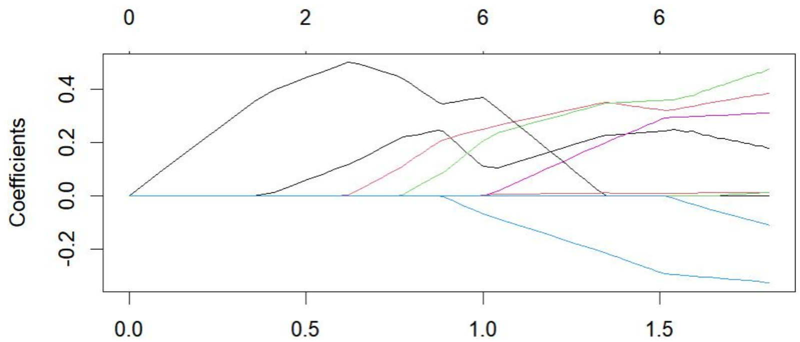

Figure 5 Lasso

The above figure(Figure 5) shows the coefficients changing with the parameter λ under Lasso estimation. As λ increases, the variables gradually enter the model. Similarly, ten-fold cross-validation is used to select the λ value with the minimum cross-validation error.Lasso Ten-fold Cross-Validation Plot can be seen in Figure 6.

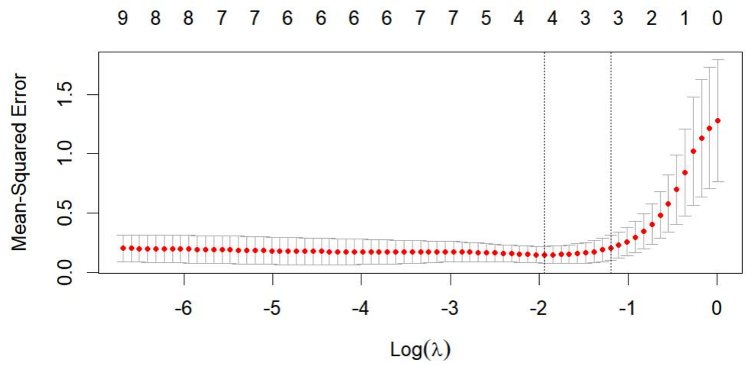

Figure 6 Lasso Ten-fold Cross-Validation Plot

Under the condition of selecting the minimum λ value 0.1434, the coefficients of the model's independent variables are as follows:

Table 5 Lasso Regression Estimation Results

Coefficients	Intercept	x_1	x_2	x_3	x_4	x_5
Estimate	0.06432890	0.23477557	0.16895035	0.05188976	—	—
Coefficients	x_6	x_7	x_8	x_9	x_{10}	
Estimate	—	0.38232146	—	—	—	

After calculation, the root mean square error (RMSE) and coefficient of determination R-squared for predicting the sample by the model are 0.248 and 92.56%, respectively. Following validation, the significance of selected parameters in Lasso regression has been confirmed.
Therefore, the regression equation fitted under Lasso regression is:

$$y = 0.06 + 0.23x_1 + 0.17x_2 + 0.05x_3 + 0.38x_7 \tag{12}$$

where all coefficients of the selected independent variables are positive. The included variables represent education, technology, policy institutions, and labor input, while variables representing the proportion of fiscal revenue to GDP for policy institutions, the ratio of import and export volume to GDP for trade openness, urbanization rate for urbanization level, longitude and latitude for geographical location, and total wastewater discharge for environmental conditions were excluded. The selection and exclusion of these variables align with practical considerations.
In terms of model effectiveness, Lasso regression outperforms OLS estimation; in terms of model simplicity, Lasso outperforms ridge regression. However, in terms of predictive accuracy and model fit, Lasso estimation is inferior to

both OLS and ridge regression. This phenomenon can be explained by the fact that OLS estimation minimizes the difference between model predictions and actual values, thereby achieving the highest R2 in terms of model fit. Ridge regression includes all independent variables in the fitting model, retaining more information than Lasso estimation, and thus outperforms Lasso in terms of predictive accuracy and model fit.

3.4 Elastic Net

Elastic Net is an estimation method that combines the regularization techniques of Ridge Regression and Lasso, leveraging the strengths of both methods.

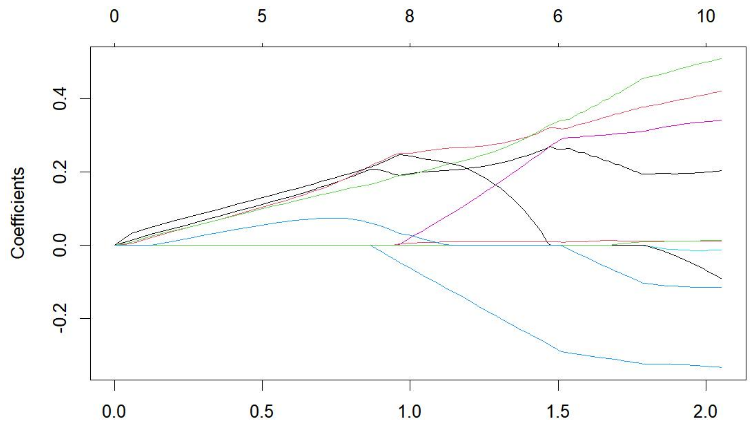

Figure 7 Elastic Net

Similar to Lasso estimation, individual variables (in Figure 7) gradually enter the model as their coefficients increase, thereby validating Elastic Net's advantage in variable selection. The selection of the optimal value is based on ten-fold cross-validation, choosing the value with the smallest cross-validation error(Figure 8). Following this, under the condition of selecting the minimum value, calculate the coefficients of the variables in the model(Table 5). Elastic Net Regression Estimation Results can be seen in Table 6.

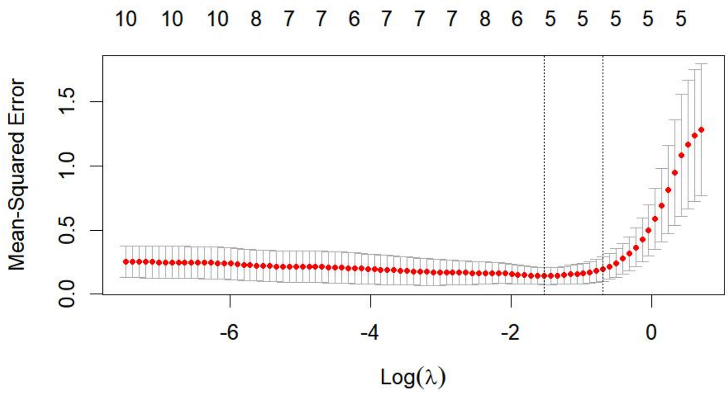

Figure 8 Elastic Net Ten-fold Cross-Validation Plot

Table 6 Elastic Net Regression Estimation Results

Coefficients	Intercept	x_1	x_2	x_3	x_4	x_5
Estimate	0.06028641	0.20452320	0.21156278	0.16460067	—	—
Coefficients	x_6	x_7	x_8	x_9	x_{10}	
Estimate	—	0.21113873	—	—	0.06343535	

Based on the calculation, the root mean squared error (RMSE) and the coefficient of determination R-squared for predicting the sample data are 0.248 and 92.56%, respectively. Upon inspection, the significance of the selected parameters under Lasso has been verified. Therefore, the regression equation fitted under Elastic Net is:

$$y = 0.06 + 0.2x_1 + 0.21x_2 + 0.16x_3 + 0.21x_7 + 0.06x_{10}, \quad (13)$$

where all coefficients of the selected predictor variables are positive. The selected variables are the same as those chosen in the Lasso estimation, representing education, technology, policy institutions, labor input, and additionally, total wastewater emissions representing environmental conditions. Other variables have been excluded. The selection

and exclusion of these variables are consistent with practical considerations. From the perspective of model effectiveness, Elastic Net outperforms OLS estimation; in terms of model parsimony, Elastic Net is superior to Ridge Regression; and regarding prediction accuracy and model fit, Elastic Net surpasses several other models.
Summary of simulated data from different estimation methods, including their model fitting and parameter selection, along with computed prediction mean square errors on test data, are as follows:

Table 7 Summary of Regression Results under Different Methods

	R^2	Parameter selection	RMSE	Variable selection
OLS	97.86%	—	0.1768	(2,3,4,6)
Ridge Regression	95.12%	$\lambda = 0.5924$	0. 1707	All
Lasso	92.56%	$\lambda = 0.1434$	0.2480	(1,2,3,7)
Elastic Net	95.02%	$\lambda = 0.2169$	0.1631	(1,2,3,7,10)

From Table 7, it is evident that Elastic Net demonstrates superior predictive accuracy compared to several other models. It exhibits greater sparsity than Ridge Regression, better effectiveness than OLS, and superior model fit compared to both Ridge Regression and Lasso, achieving an R-squared of 95.02%.
Specifically, OLS achieves an R-squared of 97.86% with a root mean square error of 0.1768. OLS estimation does not perform variable selection and may suffer from multicollinearity when not adjusted in the model. Ridge Regression and Lasso are improvements upon OLS, each with their own strengths, while Elastic Net combines the advantages of both Ridge Regression and Lasso. Therefore, this paper further analyzes the regression results under Elastic Net estimation.

3.5 Analysis of Elastic Net Regression Results

From equation (15), it is evident that the regression equation retains several significant factors influencing regional economic impact: education, technology, policy institutions, labor input, and environment. In theory, high levels of education, enhanced technological innovation capabilities, robust human capital, policy support, increased investment in labor assets, and environmental improvement all contribute to regional economic development. However, given the differences in regional capabilities in technological innovation, human capital, overall social fixed asset investment, and regional trade openness across the country, the actual impact of technological innovation, human capital, fixed asset investment, and trade openness on economic promotion varies.
Different regions have varying realities. In the production process of using technology as a factor of production, the benefits and contribution rates of each factor to economic development are different. Since all variables have been standardized during the data processing and analysis stages, the coefficients of each variable in the regression equation can to a considerable extent represent their contribution to the economy. Since the coefficients of each variable are positive, it indicates that these variables promote economic development. The contributions of education investment, technological strength, policy institutions, labor input, and environment to China's regional economic development are 20%, 21%, 16%, 21%, and 6%, respectively.

4 CONCLUSION AND RECOMMENDATIONS

4.1 Conclusion

This study established a multiple linear regression model to analyze factors influencing regional economic impact. The fitting, testing, model refinement, and evaluation were initially conducted using traditional least squares estimation. Subsequently, Ridge Regression, Lasso Regression, and Elastic Net were employed to identify the best fitting model. Finally, the regression equation derived from Elastic Net estimation was selected to analyze factors affecting regional economic development.
Comparing the four different model solving methods, Elastic Net was found to significantly enhance model fit, prediction accuracy, efficiency, and simplicity. Traditional OLS estimation often performs poorly when the number of predictors is much larger than the sample size or in scenarios of sparse models with many variables and small samples. Elastic Net effectively addresses these challenges by expanding the dimensionality of the sample and automating variable selection, transforming scenarios where the number of predictors greatly exceeds the sample size into standard multiple regression models. Scaling the original Elastic Net further optimizes results towards extremum solutions. Efficient algorithms solving the Lasso problem are utilized for each fixed parameter in the Elastic Net problem. Compared to Lasso under similar conditions, Elastic Net also reduces computational speed.
The final selected regression equation retained several factors with significant impact on regional economic development: education, technology, policy institutions, labor input, and environment. Less relevant factors were excluded from the model. The respective contributions of the selected variables to regional economic impact were determined as follows: education 20%, technology 21%, policy institutions 16%, labor input 21%, and environment 6%.

4.2 Recommendations

Based on the empirical findings of this study, the following recommendations can be provided:

4.2.1 Developing precision education-driven strategies for underdeveloped regions

For educationally disadvantaged regions like Tibet and Qinghai, it is essential to systematically and incrementally enhance technological innovation capabilities and develop precision education-driven strategies. Additionally, strategies aimed at increasing educational contributions in these areas should not solely focus on increasing educational investment. Instead, they should leverage opportunities for collaboration with other regions and universities, strengthen the sharing of scientific and research information among regions, enhance the transfer of technology and research outcomes to enterprises, and ensure that these achievements are applied in production and daily life [15].

4.2.2 Comprehensive Promotion of Regional Technological Innovation and Coordinated Economic Development

The government should comprehensively consider various situations regarding the strength of technological innovation capabilities and the level of economic development. It should accelerate achieving "prosperous regions helping less developed regions" in terms of technological innovation capabilities among different provinces, thereby avoiding further widening of spatial disparities nationwide[16]. Meanwhile, regions with strong technological innovation capabilities such as Guangdong Province should pay more attention to sustainable development of the environment and resources. Regions with weaker technological innovation capabilities can focus on integrating distinctive resources with technological innovation to stimulate the development of tourism, agriculture, and aquaculture industries, thereby promoting synchronized development of technology and the economy.

4.2.3 Increase Policy Support to Promote Labor Mobility

Labor mobility often depends on policy incentives and support. Workers move between regions, sectors, and employment statuses seeking higher wages. Increasing policy support can encourage individuals with higher and moderate educational levels to move to underdeveloped areas, thereby supporting regional development. Similarly, retaining low-skilled labor locally also contributes significantly to regional development.

COMPETING INTERESTS

The author have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Liu Qiongfang, Liu Chao. Research on the Dynamic Impact of Scientific and Technological Innovation and Industrial Structure on Regional Economy: An Empirical Study Based on GVAR Model. Times Finance, 2021(13): 33-35.

[2] Wang Anqi. Analysis of differences and influencing factors of high-quality regional economic development in China. Zhongnan University of Economics and Law, 2021.

[3] Cao Haibo. Analysis of regional economic growth differences and their influencing factors in China. Jilin:Jilin University, 2012.

[4] Zhu Xuexin, Fang Jianwen, Zhang Bin. The Impact of Scientific and Technological Innovation on China's Economic Development: An Empirical Study Based on Panel Data. Journal of Soochow University (Philosophy and Social Science), 2007(4): 18-21.

[5] Wang Wei, Xu Yilun, Yang Wenjuan, Tao Xu. The impact of scientific and technological innovation capability on regional economic development. Journal of Shandong Normal University (Natural Science Edition), 2017, 32(4): 126-136.

[6] Wang Xu, Chen Rong, Li Mingbao. Research on the Impact of Scientific and Technological Innovation on Regional Economy: An Empirical Analysis Based on Inter-provincial Panel Data. Industrial Technology and Economics, 2018, 37(9): 39-44.

[7] Song Meizhe, Li Mengsu. Measurement of the coupling and coordination relationship between higher education, scientific and technological innovation and economic development and its influencing factors. Modern Education Management, 2019(3): 19-25.

[8] Zhang Zhiruo. Research on the coupling relationship between science and technology finance and regional economic development in China. SCIENTIA GEOGRAPHICA SINICA, 2020, 40(5): 751-759.

[9] Hogenschurz Lukas, Acs Z. J. Strategies for the Development of Science, Technology and Innovation as a Public and Social Good. Journal of Entrepreneurship and Innovation in Emerging Economies, 2021, 7(1): 1-7.

[10] Sturm Thomas. Scientific innovation. Theoria: An International Journal for Theory, History and Foundations of Science, 2019, 34(3): 321-341.

[11] Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2006, 68(1): 49-67.

[12] Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology), 2005, 67(2): 301-320.

[13] Zou H. The adaptive lasso and its oracle properties. Journal of the American statistical association, 2006, 101(476): 1418-1429.

[14] National Bureau of Statistics of the People's Republic of China. China Statistical Yearbook from 2017 to 2021. Beijing:China Statistics Press.

[15] Chen Yong. Research on the influencing factors and solutions of regional economic development imbalance under the background of regional integration development in the Yangtze River Delta. Journal of Business Economics, 2022(02): 14-16.
[16] Fan Jie, Liu Hanchu. The Influence and Adaptation of Scientific and Technological Innovation Driven on the Change of China's Regional Development Pattern during the 13th Five-Year Plan Period. Economic Geography, 2016, 36(1): 1-9.