Credit Decision-Making Problem of Small, Medium and Micro Enterprises Based on Logic Equation

Zihan Wang, Guangze Xu, Peilin Luo
College of Maths, Hangzhou Normal University, Hangzhou 311121, China.

Abstract: In this paper, the bank to small, medium and micro enterprises credit problems for multidimensional decision-making research, quantitative analysis of the credit risk of enterprises, the establishment of a model to give the bank's corresponding credit strategy, and the model has been tested, analyzed and promoted. In order to solve the problem, we establish a Logistic regression model to quantitatively evaluate the credit risk of enterprises, and use the gradient descent algorithm to help solve the problem, and finally decide whether to lend to a certain enterprise for the bank. Then through the establishment of multi-objective planning model, help the bank to decide how much to lend to each enterprise, how much interest rate, from which the bank in the annual total credit fixed to each enterprise credit strategy. Finally, we test, evaluate and extend the model. The nonlinear regression model based on gradient descent algorithm has high accuracy in solving the dichotomy problem, and the multi-objective programming provides ideas for banks to formulate specific credit strategies.

Keywords: Nonlinear regression; Gradient descent; Multi-objective programming; Corporate credit

1. Restatement and analysis of the problem

1.1 Restatement of issues

In this case, the loan amount of the bank to the lending enterprise is 10,000 yuan, the annual interest rate is 4%~15%, and the loan term is one year. According to the actual and basic data information, we need to establish a mathematical model to study the credit strategy for small, medium and micro enterprises, mainly to solve the following problems:

This question requires quantitative analysis of the credit risks of 123 enterprises, and the establishment of mathematical models to give the credit strategies of the banks in question for these enterprises when the loan amount is fixed.

1.2 Problem Analysis

As for the problem, we need to conduct quantitative analysis on the credit risk of enterprises, because banks generally decide whether to provide loans or not by the strength of enterprises and whether the supply and demand relationship is stable, and the interest rate preference is determined by the level of enterprise reputation and credit risk. Therefore, we select the enterprise effective invoice rate, profit rate, corporate revenue, number of transactions and the growth rate of total profit as the five data characteristics of the Logistic regression model as the indicators in this question, and establish a regression model to quantitatively analyze credit risk. On this basis, we established a multi-objective programming model. Considering the profit of the bank and the reputation of the customer enterprise as the target, combining with the reality and regression model to establish constraints, using software programming to calculate the bank to the enterprise credit strategy.

2. Model establishment, solution and analysis

2.1 Credit decision model based on Logistic equation

We take monthly valid invoice rate, average monthly profit margin, average monthly company revenue, monthly average number of partners and average monthly profit growth rate as five indicators of the regression model to reflect the stability of enterprise supply and demand relationship, enterprise strength, enterprise scale, enterprise reputation and enterprise development prospect respectively.

2.1.1 Establishment of risk quantification model based on nonlinear regression

In the quantitative analysis of credit risk, we use Logistic model to estimate whether an enterprise will default, that is, how risky the bank may be to lend to an enterprise, and take this as the standard of whether the bank will lend to the company.

The nonlinear regression model is established as follows:

1. Decision function

Assume that the decision function is:

$$\theta^T X = [\theta_0 + \theta_1 + \cdots + \theta_n] \cdot [1 + x + \cdots + x^{n-1}]^T$$

2. Assume the sigmoid function

The sigmoid function is introduced to map the function value of the decision function to the probability of a certain type, and its expression is:

$$h_\theta(X) = \frac{1}{1 + e^{-\theta^T X}}$$

This function maps the decision boundary function to, and is expressed as the probability that x is of some class.

$$h_\theta(X) = \begin{cases} > 0.5, \theta^T X > 0 \\ 0.5, \theta^T X = 0 \\ < 0.5, \theta^T X < 0 \end{cases}$$

3. Loss function Cross entropy is used as loss function

$$cost(h_\theta(x), y) = -log(h_\theta(x)) - (1 - y)log(1 - h_\theta(x))$$

So

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} cost(h_\theta(x^{(i)}), y^{(i)})$$

In order to obtain the minimum loss function, gradient descent algorithm is used to update parameters. Calculate the gradient of the loss function with respect to parameters, and the result is:

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m} cost(h_\theta(x^{(i)}), y^{(i)})$$

Therefore, the parameter update formula is:

$$\theta^{new} = \theta^{old} - \alpha\frac{1}{m}\sum_{i=1}^{m} cost(h_\theta(x^{(i)}), y^{(i)})\, x_j^{(i)}$$

2.1.3 Solving the nonlinear regression model

First, the characteristic data are processed: since the four indicators have different dimensions, they need to be standardized into the Logistic model for calculation, and the calculation formula is as follows:

$$x^* = \frac{X - E(X)}{\sqrt{D(X)}}$$

Coefficient of five parameters revenue, profit margin, number of partners, effective invoice rate and average growth rate of total profit, the monthly average negative invoice numbers are respectively named $x_1, x_2, x_3, x_4, x_5, x_6$, labeled P, where

$$p_i = \begin{cases} 1, \text{Firm i does not default} \\ 0, \text{Firm i do default} \end{cases}$$

The final model is established as follows:

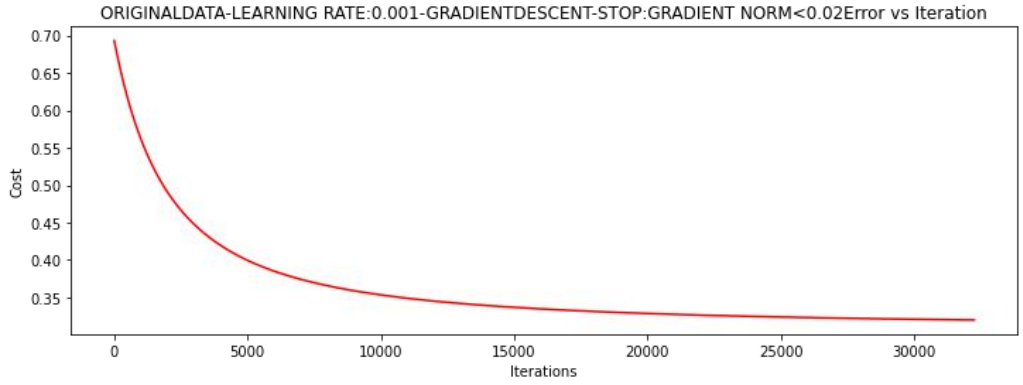
$$p = \frac{1}{1 + e^{-z}}$$

$$z = 2.280 + 0.858 * \mathrm{x}_2 + 1.436 * \mathrm{x}_3 + 1.010 * \mathrm{x}_4$$

2.1.4 Model optimization and testing

The loss rate function in the iterative solution process of the model is shown as follows:

Figure 1 Loss rate function diagram



It can be seen from the image that the final loss of the model is no more than 35%, indicating that the result is convergent.

To sum up, the accuracy of the logistic regression model is good, which reflects the rationality of our model.

2.2 Credit decision model based on multi-objective programming

2.2.1 Model establishment

Based on the quantitative analysis of the bank's credit risk of each enterprise, in order to further formulate the credit strategy for each enterprise, a multi-objective programming model is established with the three objectives of maximum bank income, minimum bank risk and maximum total loan.

According to the annual loan interest rate g_i, loan amount k_i and loan term of one year, the maximum target equation of bank income can be obtained:

$$\max z_1 = \sum_{i=1}^{n} g_i k_i$$

Obtain the objective function with the least risk of bank loss:

$$E(X) = \sum k_i(1 - p_i) - k_i q_i p_i$$

$$E(X^2) = \sum k_i^2 g_i^2 p_i + k_i^2(1 - p_i)$$

$$D(X) = \sum E(X^2) - E^2(X)$$

$$\min z_2 = D(X)$$

Maximum objective function of total loan:

$$\max z_3 = \sum_{i=1}^{n} k_i$$

Since the average customer churn rate of each credit rating cannot exceed the maximum customer churn rate of the current credit rating, the relationship among customer churn rate, credit rating and annual loan interest rate can be fitted according to the data. Rstdio software was used to fit the equation:

$$q_i = -0.072628 - 0.022621 y_i + 7.447852 g_i$$

According to the actual situation of the bank, namely, the loan limit of the lending enterprise is 100,000 to 1,000,000 yuan, the annual interest rate is 4% to 15%, the loan term is one year, the constraint condition of the total loan amount of the bank and each credit grade, the average customer loss rate of the enterprise cannot exceed the maximum customer loss rate of the current credit grade, the following objective function is established:

$$\max z_1 = \sum_{i=1}^{n} g_i k_i$$

$$\min z_2 = D(X)$$

$$\max z_3 = \sum_{i=1}^{n} k_i$$

2.2.2 Solution of multi-objective programming model

By linear weighting method, the three objective functions are weighted, and the multi-objective equation is transformed into a single objective equation:

$$\min z_4 = -0.4z_1 + 0.4z_2 - 0.2z_3$$

$$\text{s.t.}\begin{cases} \sum_{j=1}^{n_1} \frac{q_j^1}{n_1} \le d_1 \\ \sum_{j=1}^{n_2} \frac{q_j^2}{n_2} \le d_2 \\ \sum_{j=1}^{n_3} \frac{q_j^3}{n_3} \le d_3 \\ \sum_{i=1}^{n} k_i \le c \\ 0.04 \le g_i \le 0.15, i = 1,2,\cdots,n \\ 10 \le k_i \le 100, i = 1,2,\cdots,n \end{cases}$$

Where, q_j^irepresents the customer churn rate of the j^{th} enterprise in the i^{th} credit rating, $n_i$represents the total number of enterprises belonging to the i^{th} credit rating, and $d_i$represents the customer churn rate with the largest credit rating.

Use Matlab software to solve

You can get the annual interest rate on all corporate loans

3. Model innovation and promotion

Our logistic regression model is based on gradient descent algorithm established relevant knowledge, the method of the model and thinking in solving a variety of categories or index changes in the number of problems, has certain prediction function, apply to the classification decision making problems, promoting significance, such as credit evaluation, prediction product revenue, measure the success of the marketing and predict earthquakes, etc. Multi-objective programming model plays an important role in solving comprehensive evaluation problems, such as enterprise production, raw material ordering and merchant selection.

References

[1] Cui ML, Bai L, Lu SY, et al. Journal of Experimental Science and Technology, 201,19(6):8-12.

[2] Tang R. Research on credit risk assessment of smes based on Logistic Model [D]. Beijing: Tsinghua University, 2016.

[3] Shi L. Application of MATLAB in optimization model solving [J]. Journal of chifeng university (natural science edition),2012,28(8): 6-7.