Research and Empirical Analysis on the Optimal Combination Mechanism of Credit Scorecard Based on QUBO Model

Zixian Wu

School of Mathematics and Statistics, Guangxi Normal University, Guilin 541006, China.

Abstract: This paper selects relevant specific data for empirical analysis. By randomly crawling the background network data of a bank and according to the current relevant national financial policies and plans, 100 reasonable scorecards are formulated. Then, the QUBO model is innovatively introduced to optimize the combination problem of bank credit scorecard: The integer programming operational problems are constructed, and the variables in them are transformed into 0-1 binary decision variables. The data information is represented by the mathematical expression of Hamiltonian energy, and the QUBO matrix in the standard form is obtained after the transformation of constraints. Finally, the solver in the OPTI Toolbox of matlab is used to solve the problems. Find the scorecard that maximizes the final income and its corresponding unique threshold, and finally calculate the maximum benefit value.

The QUBO credit card optimization model established in this paper is highly applicable, covers a wide range of industries, and has accurate prediction value, which maximizes the risk of bank financial collection loss, and provides a feasible and effective scheme for the financial industry.

Keywords: Hamiltonian Energy; QUBO Model; Integer Combination Programming; Quadratic Binary Optimization; Credit Score Card

1. Introduction

Credit score card refers to the bank credit card or related loans and other businesses, before the customer credit, need to pass various audit rules to assess the customer's credit rating, after passing the assessment of the customer can obtain credit or loan qualification. The rule review process actually scores customers after one or more combinations of rules.High loan pass rate will bring more interest income, but it will also lead to higher bad debt rate and increase the risk of capital loss.

2. Assumptions and regulations

In order to objectively test the validity of the model, firstly, a bank was randomly selected in our country, its internal data were mined and extracted, and 100 credit scoring card rules were properly adjusted and established, and the data were stored in the background attachment. Each credit score card corresponds to 10 thresholds, ranging from threshold 1 to threshold 10. And each threshold value corresponds to different pass rate and bad debt rate respectively. Now it is necessary to find one credit score card out of 100 and its corresponding threshold to maximize the final income.

In order to describe the actual transaction situation and carry out empirical research and analysis of the model in this paper, this paper stipulates the following hypotheses with relevant definitions:

Hypothesis 1: The loan fund is 1,000,000 yuan, and the bank loan interest income rate is 8%

Hypothesis 2: Bad debt loss is: loan funds × total pass rate × total bad debt rate

Assumption 3: The bank's final revenue is determined solely by the different credit score cards and their thresholds.

Rule 1: The final income of a bank is defined as: Final income = loan interest income - bad debt loss

Rule 2: Each threshold for each credit score card can only be selected once. The variables in the above definition can be adjusted according to the actual situation and have no impact on the overall calculation process.

3. Research ideas

We first consider the establishment of loan interest objective function, which is transformed into quadratic polynomial, in order to establish QUBO model to solve. Then, using the constraints, it is transformed into a Hamiltonian operator, that is, a penalty function. The final operation function is determined by combining the objective function with the penalty function. Finally, instructions are submitted to the D-wave quantum computer through the machine language program to obtain the corresponding credit score card and its threshold value.

4. QUBO basic model of credit score card

QUBO model is an unconstrained quadratic binary optimization model, which is the most widely used optimization model in quantum computing at present, and is also a model used to solve combinatorial optimization problems. The equation of QUBO model can be expressed by matrix, and the information can be input to solve the optimal value through various software tools.

The basic QUBO form is:

$$min/max\, y = x^T Q x$$

Q is the standard QUBO matrix, $\boldsymbol{x}$ is a vector of binary variables, each of which has a value of $\{0, 1\}$, and the goal of the function is to find the minimum or maximum value of ***y.***

5. Construct QUBO Hamiltonian energy expression

5.1 Determine the objective function

In order to establish a mathematical model to select a credit score card and its corresponding threshold to maximize the final income[5], we first construct an objective function and optimize the corresponding parameters of the function to maximize the value of the function to obtain the relationship between the final income and other relevant factors that we care about[6]

Let t_i be the pass rate of the j-th threshold in the i-th credit score card, and h_{ij} be the bad debt rate of the j-th threshold in the i-th credit score card.

From loan ：

interest income = loan funds × interest income rate × total pass rate × (1- total bad debt rate),

we can get the expression of loan interest income M as follows:

$$M = L \times s \times t_{ij} \times (1 - h_{ij}) \quad (1)$$

$$i = 1,2, \ldots\ldots 100$$

$$j = 1,2, \ldots\ldots 10$$

From bad debt loss = loan funds × total pass rate × total bad debt rate,

we get the expression of bad debt loss R as follows:

$$R = L \times t_{ij} \times h_{ij} \quad (2)$$

And the final income = loan interest income-bad debt income, so we determine the objective function as follows:

$$f = L \times s \times t_{ij} \times (1 - h_{ij}) - L \times t_{ij} \times h_{ij} \quad (3)$$

In order to establish the basic model of QUBO to optimize this kind of operational problems, we need to convert all variables into binary form, so that quantum computers can be used to help solve them.

5.2 Convert binary variables into binary form

First of all, we define the selection method of credit score cards and thresholds, classify credit score cards in binary (selected and unselected), and classify the thresholds of each credit score card in binary (selected and unselected):

To express it in binary form, we use the value of x_i to indicate whether the i-th credit score card is selected (1≤ i ≤100). If $x_i = 1$,it means that the i-th credit score card is selected; If $x_i = 0$, it means that the i-th credit score card has not been selected.

The value of y_j is used to indicate whether the j-th threshold is selected (1≤ j ≤10). The value method is the same as above. Therefore, x_{ij} can also be used to indicate whether the j-th threshold of the i-th credit score card is selected at the same time.

Next, we transform the objective function into a quadratic polynomial, so as to describe the relationship between the final income and the choice of credit score card and its threshold more intuitively. The quadratic polynomial needs to consider both loan income and bad debt loss, and punish the choice of credit score card and its threshold[7].

The expression of loan interest income M obtained in question 1is:

$$M = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times s \times t_{ij}x_{ij} \times \left(1 - h_{ij}x_{ij}\right)\right] \quad (4)$$

$$\mathrm{i} = 1,2, \ldots\ldots 100$$

$$\mathrm{j} = 1,2, \ldots\ldots 10$$

The expression of bad debt loss R is:

$$R = \sum_{i=1}^{100}\sum_{j=1}^{10} L \times t_{ij}x_{ij} \times h_{ij}x_{ij} \quad (5)$$

The maximum value of final income:

$$Max \;\; w = M - R$$

$$w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times s \times t_{ij}x_{ij} \times \left(1 - h_{ij}x_{ij}\right) - L \times t_{ij}x_{ij} \times h_{ij}x_{ij}\right] \quad (6)$$

Further expanding the formula (6), we can get the following formula:

$$w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times s \times t_{ij}x_{ij} - L \times s \times t_{ij}h_{ij}{x_{ij}}^2 - L \times t_{ij} \times h_{ij}{x_{ij}}^2\right]$$

$$w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times s \times t_{ij}x_{ij} - (L \times s + L) \times t_{ij} \times h_{ij}{x_{ij}}^2\right]$$

$$w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times \left(s \times t_{ij}x_{ij} - (s+1)t_{ij} \times h_{ij}{x_{ij}}^2\right)\right]$$

Therefore, the objective function is:

$$\max w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times \left(s \times t_{ij}x_{ij} - (s+1)t_{ij} \times h_{ij}{x_{ij}}^2\right)\right] \quad (7)$$

$$\begin{cases} \sum_{i=1}^{100} x_{ij} = 1 & (j = 1,2\ldots\ldots 10) \\ x_{ij} \in \{0,1\} & (i = 1,2, \ldots\ldots 100; j = 1,2\ldots\ldots 10) \end{cases} \quad (8)$$

Since x_{ij} = 0 or 1, there is always $x_{ij} = {x_{ij}}^2$, and we can transform the objective function into the following quadratic polynomial[8]:

$$\max w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times \left(s \times t_{ij}{x_{ij}}^2 - (s+1)t_{ij} \times h_{ij}{x_{ij}}^2\right)\right] \quad (9)$$

Equivalent to:

$$\max w = \sum_{i=1}^{100}\sum_{j=1}^{10}\left[L \times \left(s - (s+1)h_{ij}\right)t_{ij}{x_{ij}}^2\right]$$

5.3 Determine the constraint conditions

Because it is stipulated that the credit scorecard can only select one of the 100 scorecards, and it is stipulated that only one corresponding threshold can be selected, we get the following restrictions:

$$\begin{cases} \sum_{i=1}^{100} x_{ij} = 1 (j = 1 \ldots 10) \\ x_{ij} \in \{0,1\} \quad i = 1,2, \ldots\ldots 100\,;\; j = 1,2, \ldots\ldots 10 \\ \sum_{i=1}^{100} x_i = 1 \\ \sum_{j=1}^{10} y_j = 1 \end{cases} \quad (10)$$

We transform the above constraints into a penalty part, and constantly adjust the weight value of the penalty coefficient to bring it into the formula for solution, in which the expression of the penalty part is:

$$\lambda \sum_{i=1}^{100}\left(\sum_{j=1}^{10} x_{ij} - 1\right)^2 + \lambda\left(\sum_{i=1}^{100} x_i - 1\right)^2 + \lambda\left(\sum_{j=1}^{10} y_j - 1\right)^2 \quad (11)$$

6. Solve the final optimization overall objective function.

Since what we require is the maximum value of the final income $w = f + \lambda g$ .we can get the formula of the final income after subtracting the penalty part from the previous objective function expression. Because of $x_{ij} \in \{0,1\}$, there is always $x_i = {x_i}^2$, $y_j = {y_j}^2$, $x_{ij} = {x_{ij}}^2$ .According to the characteristics of binary, we can convert these parts into quadratic form, which is convenient for extracting coefficients into QUBO matrix.

The following is the formula of final income:

$$\max \mathrm{w} =$$

$$\sum_{i=1}^{100}\sum_{j=1}^{10}\left[L\times\left(s\times t_{ij}x_{ij}{}^{2}-(s+1)t_{ij}\times h_{ij}x_{ij}{}^{2}\right)\right]-\lambda\left[\sum_{i=1}^{100}\left(\sum_{j=1}^{10}x_{ij}-1\right)^{2}+\left(\sum_{i=1}^{100}x_{i}-1\right)^{2}+\left(\sum_{j=1}^{10}y_{j}-1\right)^{2}\right] \quad (12)$$

By using matrix augmented expression form, we can get:

$$\max w = X^{T}QX + c$$

Among them, in the QUBO model, the q matrix is a constant symmetric matrix of 1110×1110 and c is a constant. To sum up, according to our established QUBO model, we use python to define the problem, carry out sampling operation, then define w and then output relevant sampling results to describe the independent matrix of the penalty part. For the penalty coefficient, when it tends to infinity, the result tends to the original solution. In order to improve the solving efficiency, we take the penalty coefficient as 100,000, and finally run to get the final result: when $x_{ij}=x_{49\,1}$ that is, when $x_i = x_{49}$, $y_j = y_1$, When the final income is the maximum, that is, when we take the 49th credit score card and its first threshold, we get the maximum final income, which is 61172.

7. Summary

In this paper, the objective function is transformed into a quadratic form, and the QUBO model is established through appropriate treatment, and then it is transformed into a binomial QUBO matrix. Based on the innovative formula given in this paper, the results are more in line with the actual situation, and the final income of banks under different credit score cards and their thresholds can be selected reasonably and the maximum value of final income can be obtained.

The constraint conditions contained in the current situation of market problems are transformed into Hamiltonian operators, which can be added to the objective function, so that the function and its constraints can be integrated and solved in one expression. Moreover, using quantum computer to solve the problems can reduce the time, which greatly provides a new feasible scheme for the intelligent computing and processing system in the industry.

References

[1] Sun X, Wang X, Liu L, et al. Optical-response properties in hybrid optomechanical systems with quadratic coupling[J]. Journal of Physics B: Atomic, Molecular and Optical Physics, 2018, 51(4).

[2] Kos P, Prosen T. Time-dependent correlation functions in open quadratic fermionic systems[J]. Journal of Statistical Mechanics: Theory and Experiment, 2017, 2017(12).

[3] Nadiki H M, Tavassoly M. The influence of thermal photons on the dynamics of a quantum optomechanics system with quadratic cavity–membrane couplings[J]. Annals of Physics, 2017,386.

[4] Ranjbar M, Effati S, Miri S. An artificial neural network for solving quadratic zero-one programming problems[J]. Neurocomputing, 2016, 235.

[5] Guilherme BF, Rinaldo A. Spatial dependence in credit risk and its improvement in credit scoring[J]. European Journal of Operational Research, 2016, 249(2).

[6] Zhang Q, Wang J, Lu AQ. et al. An improved SMO algorithm for financial credit risk assessment – Evidence from China's banking[J], Neurocomputing, 2018, 272.

[7] Ma S. Equivariant Gauss sum of finite quadratic forms[J]. Forum Mathematicum, 2018, 30(4).

[8] Luo J, Tian Y, Yan X. Clustering via fuzzy one-class quadratic surface support vector machine[J]. Soft Computing, 2017, 21(19).