Personal Credit Assessment of Student Loans Based on Improved Analytic Hierarchy Process

Si-min YU, Wen-hui CHEN, Jie HE, Zi-hao ZHANG and Jun-cheng LI*

College of Mathematics and finance, Hunan University of Humanities, Science and Technology, Loudi 417000, China

*Corresponding Author

Keywords: Student loan, Analytic hierarchy process, Principal component analysis, Personal credit assessment.

Abstract. In order to establish a more rational and accurate personal credit scoring model for student loans, an improved Analytic Hierarchy Process (AHP) is presented in this paper. The proposed method first uses the Principal Component Analysis (PCA) to select the main components affecting the personal credit of student loans, and then uses the AHP to analyze the personal credit of college students. The results show that the proposed method is reasonable and effective that can provides a simpler and more accurate basis of decision for student loans.

Introduction

With the exposure of major media, contemporary college students have a certain demand for loans, which has caused bad consequences. Although the governments in some countries have adopted the relevant policies, student loans still exists and needs to be solved urgently. And if the establishment of personal credit evaluation model can provide a certain basis for each student loan platform, so how to establish a personal credit system model in line with contemporary college students has become a problem that many scholars continue to study.

For the problem of student loans, many scholars have carried out a variety of research. Zhao et al [1] used neural network GABP algorithm to evaluate the credit risk of national student loan; Han et al [2] used national data to analyze the influencing factors of Korean student loan default; Ionescu [3] established a dynamic model based on the repayment behavior of college graduates in higher education funded by the United States federal Student loan program; and so on.

On the other hand, some scholars have improved some aspects of AHP, but they have not compared the improved method with the conventional AHP, and compared the advantages and disadvantages of the improved method. The main purpose of this paper is to combine AHP with PCA, and use the improved method to assess the personal credit of student loans.

Personal Credit Assessment of Student Loans Based on AHP

AHP [4] is a structured technique based on mathematics and psychology for organizing and analyzing complex decisions. It proposes a multi-Objective comprehensive evaluation method and a hierarchical weight decision analysis method [5]. By using AHP, the decision maker decomposes the complex problem into several levels and several factors, calculates and compares the mathematical model between the factors, gets the weight of different schemes, and provides the basis for selecting the best decision [6].

Selection of the Indicators

According to the basic situation of college students and the problems to be studied, we need select some reasonable and appropriate indicators. Here, 5 first-level indicators and 15 second-level indicators are selected and shown in Table 1.

Keywords: Student loan, Analytic hierarchy process, Principal component analysis, Personal credit assessment.

Abstract. In order to establish a more rational and accurate personal credit scoring model for student loans, an improved Analytic Hierarchy Process (AHP) is presented in this paper. The proposed method first uses the Principal Component Analysis (PCA) to select the main components affecting the personal credit of student loans, and then uses the AHP to analyze the personal credit of college students. The results show that the proposed method is reasonable and effective that can provides a simpler and more accurate basis of decision for student loans.

Introduction

With the exposure of major media, contemporary college students have a certain demand for loans, which has caused bad consequences. Although the governments in some countries have adopted the relevant policies, student loans still exists and needs to be solved urgently. And if the establishment of personal credit evaluation model can provide a certain basis for each student loan platform, so how to establish a personal credit system model in line with contemporary college students has become a problem that many scholars continue to study.

For the problem of student loans, many scholars have carried out a variety of research. Zhao et al [1] used neural network GABP algorithm to evaluate the credit risk of national student loan; Han et al [2] used national data to analyze the influencing factors of Korean student loan default; Ionescu [3] established a dynamic model based on the repayment behavior of college graduates in higher education funded by the United States federal Student loan program; and so on.

On the other hand, some scholars have improved some aspects of AHP, but they have not compared the improved method with the conventional AHP, and compared the advantages and disadvantages of the improved method. The main purpose of this paper is to combine AHP with PCA, and use the improved method to assess the personal credit of student loans.

Personal Credit Assessment of Student Loans Based on AHP

AHP [4] is a structured technique based on mathematics and psychology for organizing and analyzing complex decisions. It proposes a multi-Objective comprehensive evaluation method and a hierarchical weight decision analysis method [5]. By using AHP, the decision maker decomposes the complex problem into several levels and several factors, calculates and compares the mathematical model between the factors, gets the weight of different schemes, and provides the basis for selecting the best decision [6].

Selection of the Indicators

According to the basic situation of college students and the problems to be studied, we need select some reasonable and appropriate indicators. Here, 5 first-level indicators and 15 second-level indicators are selected and shown in Table 1.
Table 1. The selected indicators.

<table>
<thead>
<tr>
<th>First-level indicators</th>
<th>Second-level indicators</th>
<th>First-level indicators</th>
<th>Second-level indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic information of students ($C_1$)</td>
<td>Household registration ($P_{11}$)</td>
<td>The environment in which students located ($C_4$)</td>
<td>Family situation ($P_{41}$)</td>
</tr>
<tr>
<td></td>
<td>Physical condition ($P_{12}$)</td>
<td></td>
<td>Nature of the school ($P_{42}$)</td>
</tr>
<tr>
<td>Consumption behavior of student ($C_2$)</td>
<td>Monthly consumption times ($P_{21}$)</td>
<td></td>
<td>Friend credit status ($P_{43}$)</td>
</tr>
<tr>
<td></td>
<td>Monthly consumption amount ($P_{22}$)</td>
<td></td>
<td>Comprehensive score ranking ($P_{51}$)</td>
</tr>
<tr>
<td>Morality of students ($C_3$)</td>
<td>Reward or punishment ($P_{31}$)</td>
<td>Situation of students in school ($C_5$)</td>
<td>Monthly living expenses ($P_{52}$)</td>
</tr>
<tr>
<td></td>
<td>The evaluation of people ($P_{32}$)</td>
<td></td>
<td>Monthly income ($P_{53}$)</td>
</tr>
<tr>
<td></td>
<td>Past lending situation ($P_{33}$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Construction of Judgment Matrix

We construct the judgment matrix of the 5 first-level indicators as follows,

\[
\begin{pmatrix}
1 & 1/5 & 1/3 & 1/2 & 1/7 \\
5 & 1 & 2 & 3 & 1/2 \\
3 & 1/2 & 1 & 2 & 1/3 \\
2 & 1/3 & 1/2 & 1 & 1/4 \\
7 & 2 & 3 & 4 & 1
\end{pmatrix}
\]

And the judgment matrixes of the 15 second-level indicators corresponding to different first-Level indicators are constructed as follows,

\[
\begin{pmatrix}
1 & 1/4 & 3 \\
4 & 1 & 5 \\
1/3 & 1/5 & 1
\end{pmatrix}
, \quad
\begin{pmatrix}
1 & 1/2 & 2 \\
2 & 1 & 3 \\
1/2 & 1/3 & 1
\end{pmatrix}
, \quad
\begin{pmatrix}
1 & 2 & 1/5 \\
1/2 & 1 & 1/4 \\
5 & 4 & 1
\end{pmatrix}
, \quad
\begin{pmatrix}
1 & 3 & 5 \\
1/3 & 1 & 2 \\
1/5 & 1/2 & 1
\end{pmatrix}
, \quad
\begin{pmatrix}
1 & 1/5 & 1/4 \\
5 & 1 & 2 \\
4 & 1/2 & 1
\end{pmatrix}
\]

With the dynamic development of thought, the model can continuously adjust and optimize the personal credit risk assessment process by changing the judgment matrix [7].

Consistency Test and Calculation of Index Weights

The judgment matrix is adjusted to satisfy the consistency, and then \( AW = \lambda_{\text{max}} W \) can be used to calculate the index weight.

Personal Credit Assessment of Student Loans Based on AHP with PCA

Since PCA can screen out the main factors, it is superior to choose the PCA in cases where there is no greater impact on the results, which also reduces computational cost. Therefore, we first use the PCA to filter out the main indicators containing the main information, and then apply AHP to calculate the weights of the selected indexes to obtain the final results.

Calculate the Correlation Coefficient Matrix

We set an appropriate questionnaire to obtain the data \( r_{ij} \) in the following matrix. Then the correlation coefficient matrix \( \rho = \left( \rho_{ij} \right) \) is calculated by

\[
\rho_{ij} = \frac{\sum_{k=1}^{p} (r_{ik} - \bar{r}_i)(r_{kj} - \bar{r}_j)}{\sqrt{\sum_{k=1}^{p} (r_{ik} - \bar{r}_i)^2 (r_{kj} - \bar{r}_j)^2}}.
\] (1)
where $\bar{r}_i$ is the average of line $i$ in matrix $R$.

**Calculate the Principal Components**

The eigenvalues of the matrix $R$ is calculated by $|\lambda I - R| = 0$. And the obtained eigenvalues can be arranged in order from large to small, viz. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. Then the contribution rate $con$ and the cumulative contribution rate $cum$ of the 15 second-level indicators can be calculated by

$$con = \frac{\lambda_i}{\sum_{k=1}^{p} \lambda_k} (i = 1, 2 \cdots p),$$

(2)

$$cum = \frac{\sum_{k=1}^{i} \lambda_k}{\sum_{k=1}^{p} \lambda_k} (i = 1, 2 \cdots p).$$

(3)

When the cumulative contribution rate of several eigenvalues reaches 80%, the main principal components $Z_i$ can express the information comprehensively and perfectly.

**Unitizing AHP**

The main indicators selected are re-divided into the primary and secondary indicators. And the AHP is used to analyze these indicators to comprehensively evaluate the personal credit information of college students [9].

**Example and Analysis**

We present an example to illustrate the effectiveness of the proposed method in this section.

**Obtaining the Scoring Data**

On the basis of the indicators, the corresponding questionnaire was set up and ten college students were randomly selected to fill in the questionnaire. The data obtained from the survey are graded on the basis of importance.

**Solution by AHP**

The weights of each indicator are obtained by using MATLAB to calculate the characteristic root of the judgment matrix. Then the total weight is obtained by multiplying the weight of the first-level indicators with the weight of the second-level indicators. Then, the primary weight and the total weight are sorted and shown the results in Table 2.

<table>
<thead>
<tr>
<th>First-level indicators</th>
<th>Weights of first-level indicators</th>
<th>Order number</th>
<th>Second-level indicators</th>
<th>Weights of Second-level indicators</th>
<th>Final weight</th>
<th>Final order number</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>0.0529</td>
<td>5</td>
<td>$P_{11}$</td>
<td>0.2255</td>
<td>0.0119</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{12}$</td>
<td>0.6738</td>
<td>0.0356</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{13}$</td>
<td>0.1007</td>
<td>0.0053</td>
<td>15</td>
</tr>
<tr>
<td>$C_2$</td>
<td>0.2659</td>
<td>2</td>
<td>$P_{21}$</td>
<td>0.2970</td>
<td>0.0790</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{22}$</td>
<td>0.5396</td>
<td>0.1435</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{23}$</td>
<td>0.1634</td>
<td>0.0434</td>
<td>7</td>
</tr>
<tr>
<td>$C_3$</td>
<td>0.1547</td>
<td>3</td>
<td>$P_{31}$</td>
<td>0.1865</td>
<td>0.0289</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{32}$</td>
<td>0.1265</td>
<td>0.0196</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{33}$</td>
<td>0.6870</td>
<td>0.1063</td>
<td>4</td>
</tr>
<tr>
<td>$C_4$</td>
<td>0.0942</td>
<td>4</td>
<td>$P_{41}$</td>
<td>0.6483</td>
<td>0.0611</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{42}$</td>
<td>0.2297</td>
<td>0.0216</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$P_{43}$</td>
<td>0.1220</td>
<td>0.0115</td>
<td>14</td>
</tr>
</tbody>
</table>
Solution by AHP Combined with PCA

The relative importance of each component can be evaluated by using the PCA [10]. Aiming at the second-level indicators, we use MATLAB to conduct principal component analysis on 15 indicators. The characteristic value, difference value, contribution rate and cumulative contribution rate is calculated respectively, which are shown in Table 3. Only values with cumulative contribution rate of 100% are selected in table 3.

Table 3. Data of the principal component analysis.

<table>
<thead>
<tr>
<th></th>
<th>Eigenvalue(eig)</th>
<th>Difference(fea)</th>
<th>Contribution(con)</th>
<th>Cumulative contribution rate(cum)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7957</td>
<td>1.1761</td>
<td>31.9711</td>
<td>31.9711</td>
<td></td>
</tr>
<tr>
<td>3.6196</td>
<td>1.3380</td>
<td>24.1306</td>
<td>56.1017</td>
<td></td>
</tr>
<tr>
<td>2.2816</td>
<td>0.6445</td>
<td>15.2109</td>
<td>71.3126</td>
<td></td>
</tr>
<tr>
<td>1.6372</td>
<td>0.4674</td>
<td>10.9144</td>
<td>82.2270</td>
<td></td>
</tr>
<tr>
<td>1.1698</td>
<td>0.3132</td>
<td>7.7986</td>
<td>90.0256</td>
<td></td>
</tr>
<tr>
<td>0.8566</td>
<td>0.4744</td>
<td>5.7105</td>
<td>95.7361</td>
<td></td>
</tr>
<tr>
<td>0.3822</td>
<td>0.1993</td>
<td>2.5478</td>
<td>98.2839</td>
<td></td>
</tr>
<tr>
<td>0.1829</td>
<td>0.1084</td>
<td>1.2193</td>
<td>99.5032</td>
<td></td>
</tr>
<tr>
<td>0.0745</td>
<td>0.0745</td>
<td>0.4968</td>
<td>100.0000</td>
<td></td>
</tr>
</tbody>
</table>

From Table 3, we can see that only four principal components are needed to achieve 82.2270% information integrity. Hence, the four principal components can basically replace the information quantity of all indicators. The load matrix $L$ can be calculated by $L = \sqrt{\text{eig} \times \text{fea}}$. Then, the obtained load matrix $L$ can be expressed as follows,

$$L = \begin{pmatrix}
    0.7769 & 0.4804 & 0.1561 & -0.0720 \\
    -0.1560 & 0.2082 & -0.0202 & 0.7560 \\
    0.3081 & 0.2461 & 0.4788 & 0.2633 \\
    0.6222 & 0.4777 & 0.4105 & 0.2383 \\
    0.4403 & -0.5702 & -0.6386 & -0.1377 \\
    -0.0869 & -0.4024 & 0.7830 & 0.2541 \\
    0.8853 & -0.3135 & -0.2083 & -0.1871 \\
    0.7227 & -0.3262 & 0.4136 & 0.0202 \\
    -0.0901 & 0.9395 & 0.0110 & -0.2382 \\
    0.8337 & 0.0268 & -0.2324 & 0.3089 \\
    0.4555 & 0.8064 & -0.1717 & -0.0841 \\
    0.0334 & -0.8285 & 0.4287 & -0.0770 \\
    0.8720 & -0.3652 & 0.0046 & -0.1333 \\
    0.8418 & -0.0099 & -0.2286 & 0.2964 \\
    0.2837 & 0.2292 & 0.5442 & -0.6805
\end{pmatrix}$$

It can be seen that the first principal component is $P_{51}, P_{31}, P_{32}, P_{41}, P_{51}, P_{52}$; the second principal component is $P_{33}, P_{42}$; the third principal component is $P_{23}$; the fourth principal component is $P_{12}$. Hence, we only need 10 indicators to replace the original 15 indicators for assessing the personal credit of college students synthetically by AHP.

Then the judgment matrix of each secondary indicator can be reset as follows,
and the selected 10 secondary indicators can be divided into three primary indicators which is named basic information of students, school status of students and consumption status of students respectively. Then, by the AHP, the second-level weight and the total weight can be obtained and shown the results in Table 4.

Table 4. Weight of each indicator.

<table>
<thead>
<tr>
<th>First-level indicators</th>
<th>First-level weights</th>
<th>Second-level indicators</th>
<th>Second-level weight</th>
<th>Final weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Basic information</td>
<td>0.0936</td>
<td>$P_{11}$</td>
<td>0.1243</td>
<td>0.0116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{12}$</td>
<td>0.3586</td>
<td>0.0336</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{41}$</td>
<td>0.5171</td>
<td>0.0484</td>
</tr>
<tr>
<td>Student school situation</td>
<td>0.2797</td>
<td>$P_{42}$</td>
<td>0.0952</td>
<td>0.0266</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{51}$</td>
<td>0.1904</td>
<td>0.0533</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{31}$</td>
<td>0.4775</td>
<td>0.1336</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{32}$</td>
<td>0.2369</td>
<td>0.0663</td>
</tr>
<tr>
<td>Student consumption situation</td>
<td>0.6267</td>
<td>$P_{52}$</td>
<td>0.1571</td>
<td>0.0985</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{23}$</td>
<td>0.5936</td>
<td>0.3720</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P_{33}$</td>
<td>0.2493</td>
<td>0.1562</td>
</tr>
</tbody>
</table>

The results of the survey of 10 college students are multiplied by the weights obtained by AHP and the weights obtained by AHP combined with PCA [11]. The ranking of each college student's personal credit assessment score can be obtained. Table 5 shows the ranking obtained by AHP combined with PCA (current ranking) and AHP (original ranking).

Table 5. Credit ranking of college students.

<table>
<thead>
<tr>
<th>Student’s serial Number</th>
<th>Current score</th>
<th>Current ranking</th>
<th>Original Score</th>
<th>Original ranking</th>
<th>Ranking changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. 1</td>
<td>4.1439</td>
<td>1</td>
<td>6.8364</td>
<td>5</td>
<td>-4</td>
</tr>
<tr>
<td>No. 2</td>
<td>3.9909</td>
<td>2</td>
<td>6.6515</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>No. 3</td>
<td>3.3423</td>
<td>6</td>
<td>5.4767</td>
<td>8</td>
<td>-2</td>
</tr>
<tr>
<td>No. 4</td>
<td>2.9786</td>
<td>9</td>
<td>4.0994</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>No. 5</td>
<td>4.0400</td>
<td>5</td>
<td>5.8552</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>No. 6</td>
<td>4.1601</td>
<td>3</td>
<td>6.2757</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No. 7</td>
<td>3.3087</td>
<td>7</td>
<td>5.4696</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>No. 8</td>
<td>3.4254</td>
<td>8</td>
<td>4.3167</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>No. 9</td>
<td>4.0969</td>
<td>4</td>
<td>5.9238</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>No. 10</td>
<td>1.8637</td>
<td>10</td>
<td>2.9927</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

From Table 5, we can see that the credit ranking of college students rarely changes. And an analysis of two high-ranking college students can be shown as follows,

(a) The second college student's personal credit ranking raises four. Comparing him to the original rank, we can find that although the second college student family situation, monthly living expenses of the two level two indicators score higher, his reward and punishment situation, other people's evaluation and other aspects than the first college students, which indicates that the second college students personal credibility has a problem.

(b) The third college student's personal credit ranking raises two. Comparing with the original rank, we can see that although the eighth college student has more living expenses each month, his past loan situation is not as good as the third college student, which indicates that his personal credibility is also problematic.

Therefore, it can be found that the results after principal component analysis are more reasonable
than the previous credit rankings of College students. It can be seen that the improved method not only reduces the data needed for analysis, but also improves the credibility of the results.

**Conclusions**

In this paper, we use AHP combined with PCA to reduce the number of indicators to simplify the calculation of personal credit assessment of college students, which makes the personal credit assessment of college students more convenient and convincing.

**Acknowledgement**

This work was supported by the Hunan Provincial Undergraduate Research Learning and Innovative Experimental Program of China (Xiang Jiao Tong [2018] No. 255_ No. 1008).

**References**


