# Principal Component Analysis (PCA) in Smart Growth Theory

## Qingqing Zhang

North China Electric Power University, Baoding 071003, China

rukawakaede723@163.com

**Keywords:** Smart growth index, Principal Component Analysis.

**Abstract.** Today, smart growth theory is playing an important role in urban development. In order to help implementing smart growth theories into city design around the world, we select Bendigo in Australia and Galway in Ireland as the objects of study. We define a smart growth index to evaluate the success of the smart growth of a city. The higher the index is, the higher the level of urban smart growth will be. To build an evaluation system, we select 10 indicators of 6 similar mid-sized cities in Australia and Ireland respectively according to ten principles for smart growth and three E's. To eliminate the correlation among different factors, we use Principal Component Analysis to calculate the principal components of collected data and their contribution rate, then linearly combine the value of them to get smart growth function. Put the indicators 'data of Bendigo and Galway into the smart growth function to calculate smart growth index. We find that Bendigo's smart growth index is higher, but it is lower than the average level in selected cities of Australia.

## 1. Introduction

Smart Growth theory [1], originated in the 1990's, advocate a people-oriented, healthy and harmonious development concept, and it has an increasingly important position in urban planning [2]. In order to help implement the s-mart growth theories into city design around the world, we have chosen Bendigo from Australia and Galway from Ireland, taking the ten principles of smart growth into account, to measure the success of the selected cities in the current development plan by using principal component analysis (PCA) to build a smart growth function and make a comprehensive evaluation of the data.

## 2. Principal Components Analysis(PCA)

In order to scientifically and accurately reflect the success rate of urban smart growth, we defined the smart growth index [3], which takes following aspects into account:
1. The urban population density
2. Natural population growth rate to measure the urban population size;
3. Per capita GDP;
4. Per capita disposable income;
5. Unemployment rate;
6. Work hours per per week per person;
7. House rents per week, to a certain extent, also reflects the city
Construction and land use conditions;
8. The number of bus stations around the community;
9. Residents walking proportion;
10. The percentage of undergraduate students aged 15 and over.
Considering these 10 indicators above, we chose twelve cities from Australia and Ireland and get the urban smart growth index [6][7] using Principal Components Analysis (PCA). We believe that the index is more comprehensive and objective to reflect the level of urban smart growth and has a certain reference value.

## 2.1 Data Standardization

We assume that there are m principal component analysis index variables: $X_1, X_2 \ldots\ldots X_m$, and n evaluation objects. $X_{ij}$ is the value of the index j of the evaluation object i. We convert each index into standardized indicators $\overline{X}_{ij}$ :

$$\overline{X}_{ij} = \frac{X_{ij} - \overline{X}_j}{S_j}, (i,j=1,2,...,m) \tag{1}$$

Where: $\overline{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$ , $S_j = \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij} - \overline{X}_j)^2$ ,(i,j=1,2,...,m)

## 2.2 Calculating the Correlation Coefficient Matrix.

The correlation coefficient matrix:

$$R = (r_{ij})_{m*m} \tag{2}$$

According to the existing definition of PCA:

Where: $r_{ij} = \dfrac{\sum_{k=1}^{n} \overline{X}_{ki} * \overline{X}_{kj}}{n-1}$ ,(i,j=1,2,...,m)

$r_{ii} = 1$, $r_{ij} = r_{ji}$, $r_{ij}$ is the correlation coefficient of the index i and index j.

## 2.3 Calculating the Eigenvalue and Eigenvector

Firstly,we calculated the eigenvalue $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m \geq 0$ of R and its eigenvector $u_1, u_2,..., u_m$ . where: $u_j = (u_{1j}, u_{2j},..., u_{nj},)^T$

Then we get m new index variables which is composed of eigenvector:

$$\begin{cases} y_1 = u_{11}\overline{X}_1 + u_{21}\overline{X}_2 + ... + u_{n1}\overline{X}_{n1} \\ y_2 = u_{12}\overline{X}_1 + u_{22}\overline{X}_2 + ... + u_{n2}\overline{X}_{n2} \\ \ldots \\ y_m = u_{1m}\overline{X}_1 + u_{2m}\overline{X}_2 + ... + u_{nm}\overline{X}_{nm} \end{cases} ,(i=1,2,...,m) \tag{3}$$

Where: $y_1$ is the first principal component, $y_2$ is the secend principal component, $y_m$ is the principal component m.

## 2.4 Select the Principal Component P (P ≤ M) and Calculate the Comprehensive Evaluation Value.

We calculated the contribution rate and cumulative contribution rate of the eigenvalue $\lambda_j$ , (j = 1, 2, ..., m).We call

$$b_j = \frac{\lambda_j}{\sum_{k=1}^{m} \lambda_k} , (j = 1, 2, ..., m) \tag{4}$$

The contribution rate of principal component,

$$\alpha_p = \frac{\sum_{k=1}^{p} \lambda_k}{\sum_{k=1}^{m} \lambda_k} \tag{5}$$

The cumulative contribution rate of $y_1, y_2,..., y_p$ .We chose the first p $y_1, y_2,..., y_p$ as p principal component when the $\alpha_p$ is close to 1 and use them to take the place of the m original index variables. Then we can analyze the principal components comprehensively.

When we were going to measure the success of the chosen cities, we defined Z as the smart growth index:

$$Z = \sum_{j=1}^{p} b_j y_j \tag{6}$$

## 3. Data Processing and Calculation

Since not all of the factors are positively correlated with the smart growth index, we need to do some data processing. And we can get the revised index matrix X.
We use MATLAB to standardize the 10 evaluation factors and find the correlation coefficient matrix. The cumulative contribution rates are show in Table.

Table 1: The cumulative contribution rate of the characteristic root

| serial number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| cumulative contribution rate | 0.5670 | 0.7134 | 0.8463 | 0.9408 | 0.9619 |

According to the data above, we can see that the cumulative contribution rate of the first five characteristic roots is more than 96%.The principal component analysis performs well. Next we select five principal components (cumulative contribution rate reach to 96.19%) to go head a comprehensive evaluation.

Calculate the first five eigenvectors of the characteristic root and we get five principal components: $y_1, y_2, ..., y_5$.

From the coefficients of each principal component we can get that: the first principal component mainly reflects the natural population growth rate, per capita disposable income, per capita GDP and unemployment ratethe second principal component mainly reflects per capita disposable income and the number of bus stations around the community the third principal component mainly reflects the the percentage undergraduate students aged fifteen and over the forth principal component mainly reflects the natural population growth rate and per capita GDP the fifth principal component mainly reflects per capita disposable income, per capita GDP and residents walking proportion.

Plugging the index standardized value $\overline{X}_{ij}$ of each object into (3) and regard the contribution rate of the five principal components as weight, our Principal Component Analyze Model is constructed:

$$Z = \sum_{j=1}^{p} b_j y_j \tag{6}$$

Substitute the five principal components of the sample into (6), we can obtain their smart growth evaluation results.
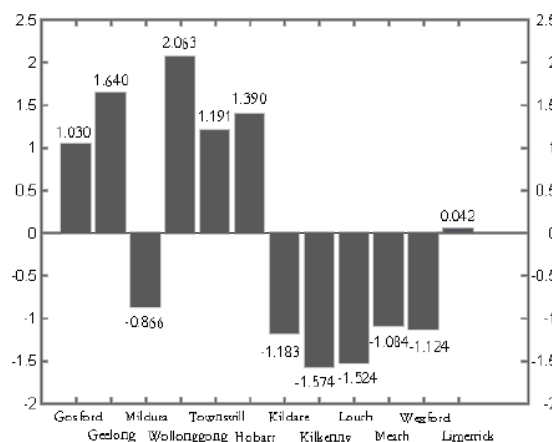


Figure 1: Smart growth index in the 12 cities from Australia and Ireland

By access to information about the development of cities in Australia and Ireland, We find that the overall urban development of Australia is better than that of Ireland, which is consistent with our conclusion by principal component analysis. We think the main reason for this situation lies in the

following points: In urban planning and construction, Australia has introduced a number of effective policies.

On one hand, by dividing different spheres to strengthen land intensive use and set different intensive standards in different circles, on the other hand, build high-density low-volume villa group in the suburbs to avoid the emergence of low density urban area in the United States. At the same time, a reasonable definition of urban development boundaries, encouraging high-density land development, giving priority to the development of rail transportation systems and other measures also greatly ensure the overall level of its urban development. In recent years, the construction and exploration of learning city has also improved the level of urban development in Australia.

However, it should be noted that, due to the very different development of countries, our current evaluation system only applies to cities in Australia and Ireland. For cities in other countries, we need to reselect the data to revise the parameters.

## 4.   Conclusion

We choose Bendigo in Australia and Galway in Ireland ,based on the ten principles for smart growth and select natural population growth rate, per capita GDP and other eight indicators ,use Principal Component Analysis to build smart growth function .We define the smart growth index to reflect the success rate of urban smart growth.

Using the functions we build, we analyze the Bendigo and Galway growth plans, and draw the following conclusions:

1. Of the two cities, Bendigo's smart growth index is higher than that of Galway, so we believe that Bendigo's current smart growth plan is better than Galway. This conclusion is consistent with the conclusion that Australia's overall urban development is better than that of Ireland;

2. In Australia, Bendigo's smart growth index is slightly lower than the average level of Australian city of the same size, so we believe that despite the current growth plan in Bendigo has made some achievements, but compared with other similar size cities in Australia there are still some Issues. They should strengthen exchanges with other cities when drawing up a plan.

3. In Ireland, Galway's smart growth index is slightly higher than the average level of Ireland city of the same size. Therefore, we believe that, there are some problems in Galway's current growth plan. But compared with other similar cities in the country, they have made some positive results. They may consider promoting their growth plan to do other cities in the country.

## 5.   References

[1]. Li Ming. Study on the Harmonious Community Construction and Evaluation Index System Based on the Principle of Smart Growth. Changan University, China. 2007.6.

[2]. Zuqi Ma.From Urban Sprawl to Smart Growth, the Change of Land Use in America. Urban Problems, 2007(10):86-90.

[3]. A Colquhoun. Building Sustainable Communities. Taiwan: SHU-TE University, 2005.

[4]. Haiming Lin.An Analysis of Ten Problems in the Application of Principal Component Analysis. Statistics and Decision, 2007:16-18.

[5]. Liuping Sun, Wuyong Qian. Improvement of Comprehensive Evaluation Method Based on Principal Component Analysis. Mathematics in Practice and Theory, 2009(18):15-20.

[6]. http://www.cso.ie/en/index.html.

[7]. http://www.abs.gov.au/browse?opendocument&ref=topBar.