

Analysis on Energy Conservation and Emission Reduction for Electricity Consumers Based on Principal Component Analysis

Xiaohui XU^{1, a}, Jinsong LIU^{2, b}, Yirong SU^{3, c}, Yangyang XI^{3, d}, shuanghu WANG^{3, e}

¹China Electric Power Research Institute, Renewable Energy Department, Nanjing, China

²State Grid Shanghai Municipal Electric Power Company, Shanghai, China

³State Grid Electric Power Research Institute, Nanjing, China

^aemail: xuxiaohui@163.com, ^bemail: jsli@163.com, ^cemail: syr@126.com, ^demail: xyy@163.com, ^eemail: wangshuanghu@163.com,

Keywords: principal component analysis; energy conservation; emission reduction; correlation analysis

Abstract. With the increasing concerns on energy shortage and environment protection, the energy conservation and emission reduction is becoming a more of concern for electricity consumers. In this paper, typical indexes are selected for reflecting electricity consumers' statues of energy shortage and environment protection. To simply the analysis on energy conservation and emission reduction for electricity consumers, the Principal Component Analysis is adopted to reduce the number of indexes whereas without original information about energy conservation and emission reduction loss. Finally, an example is presented to demonstrate the analysis process of the proposed method for the electricity consumers in an actual region.

Introduction

Energy conservation and emission reduction is throughout all aspects of the use of electricity. From the conversion, transmission, distribution of the energies to the terminal use, we should take effective measures from the aspects of economy, technology and so on, to reduce the waste of energy[1][2]. In 1995, the world energy council defined energy efficiency as reducing the energy input when providing equal energy service. Since then, the word energy conservation is replaced by energy efficiency more widely in the world. In essence, the both meaning is the same.

Electricity consumers are an important object of energy conservation and emission reduction. In the process of electricity consumers' using electricity with the resource consumption and pollutant emission, that can better reflect the status of the electricity consumers to use energy and internal energy utilization characteristics. From another point of view, the energy utilization features can also reflect the problems existing in the energy utilization of electricity consumers. If we can better grasp the existing problems of the energy utilization, which can be targeted to formulate the corresponding measures to eliminate these problems to improve the effect of energy conservation and emission reduction [3][4][5].

According to the above, evaluating the effect of energy conservation of the electricity consumers will involve a number of indexes. However, the information between the indexes may be highly overlapping and highly relevant, which brings many obstacles to the application of statistical methods. To resolve this problem, the simplest and most direct solution is to cut down the number of variables, but it will inevitably lead to problems such as loss of information and incomplete information. Principal Component Analysis is a method that can reduce the dimension of the variables effectively and not cause the loss of large amounts of information.

Principal Component Analysis

In mathematical sense, Principal Component Analysis is a method of data dimension reduction, the essence of which is to make large number of related indexes $X=[X_1, X_2, \dots, X_p]$ (such as p index) for recombination, to get a group of composite index F_m that has less number and is unrelated to each other to replace a large number of original indexes. Of course, in the process of the extraction of indexes, which can ensure that not only reflecting the original information represented by variable X_1, X_2, \dots, X_p in the maximum extent, but also the new indexes remain independent of each other (information not overlap).

Assume that a principal component index formed by the linear combination of the original variables is

$$F_1 = a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \quad (1)$$

If the principal component F_1 contains the largest amount of information, that is referred to as the first principal component, that should have the greatest variance in X_1, X_2, \dots, X_p of all linear combinations. According to this principle, we can combine the second principal component index F_2 . In order to reflect the original information effectively, F_1 's information will not appear in F_2 , namely the F_2 and F_1 remain independent, unrelated, the Covariance of which is $\text{Cov}(F_1, F_2)=0$ expressed in mathematical language, so F_2 and F_1 that is unrelated have the greatest variance in X_1, X_2, \dots, X_p of all linear combinations. So on we can construct the principal component F_1, F_2, \dots, F_m of the original variable index X_1, X_2, \dots, X_p ($m \leq p$)^[4].

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ F_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ \dots \\ F_m = a_{m1}X_1 + a_{m2}X_2 + \dots + a_{mp}X_p \end{cases} \quad (2)$$

In order to define all the principal component, we need to determine the expression of the principal components F_i ($i=1, 2, \dots, m$) about original variable X_j ($j=1, 2, \dots, p$), that is the coefficient a_{ij} ($i=1, 2, \dots, m; j=1, 2, \dots, p$). The covariance matrix of the original variables is

$$\sum = (s_{ij})_{p \times p} \quad (3)$$

Among them

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \quad (i, j = 1, 2, \dots, p) \quad (4)$$

From the above analysis we know that previous m larger principal components variance values can be represented by previous m larger characteristic root which corresponds to eigenvector that is the coefficient of principal component F_i expression. Assume that previous m larger eigenvalues are $\lambda_1 \geq \lambda_2 \geq \dots \lambda_m > 0$, a_i is the eigenvector that corresponds to eigenvalue, then it is about the coefficient of the original variables of principal component F_i . Therefore, the original variable of i principal component F_i is:

$$F_i = a_i X \quad (5)$$

The contribution rate of a principal component of the variance can be used to reflect the size of the information, if the amount of information is represented by the sum of variance of all the principal components. We define α_i as:

$$\alpha_i = \lambda_i / \sum_{i=1}^m \lambda_i \quad (6)$$

In general, when the accumulation of a set of principal component contribution rate is greater than or equal to 85%, this group of principal component can enough reflect the information of the original variable. According to this principle, we can determine the amount of principal component needed to reflect the information of original variable X . Suppose that the cumulative contribution rate $G(m)$ involving the variance(information) of m principal components F_1, F_2, \dots, F_m is

$$G(m) = \sum_{i=1}^m \lambda_i / \sum_{k=1}^p \lambda_k \quad (7)$$

When contribution rate $G(m)$ is greater than or equal to 85%, m is the number of principal components needed.

Principal component load reflects the correlation between principal component F_i and original variable X_j , l_{ij} ($i=1, 2, \dots, m; j=1, 2, \dots, p$) that is the load of original variable X_j ($j=1, 2, \dots, p$) on all principal component F_i ($i=1, 2, \dots, m$),

$$l(Z_i, X_j) = \sqrt{\lambda_i} a_{ij} (i=1, 2, \dots, m; j=1, 2, \dots, p) \quad (8)$$

On the basis of the principal component load matrix, we can judge the influence privacy closely related with the electricity consumers on energy conservation and emission reduction.

The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph).

Numerical example

After investigation and research, analyze the status of energy conservation and emission reduction of the electricity consumers by an area as an example. On the basis of the screening of the users' relevant data, the effective data of 8 users are obtained, as shown in Table 1.

Table 1 Relevant data for the index of energy conservation and emission reduction

User	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
1	0.786	0.521	0.882	0.734	0.337	0.859	0.38	7952.832	0.692
2	0.643	0.752	0.774	0.636	0.456	0.876	0.49	1422.803	0.561
3	0.832	0.765	0.802	0.537	0.327	0.821	0.14	1452.608	0.219
4	0.861	0.654	0.793	0.765	0.213	0.893	0.41	3915.901	0.288
5	1.342	1.078	0.869	0.813	0.159	0.905	1.33	1275.469	1.642
6	1.537	1.231	0.758	0.827	0.132	0.932	0.341	3622.376	0.297
7	1.635	1.522	0.691	0.646	0.439	0.851	4.136	13669.06	4.442
8	1.439	1.439	0.768	0.742	0.257	0.935	0.512	1317.54	1.873

The results are shown in table 3 after calculation. Due to the first, second and third principal components F_1 、 F_2 、 F_3 in which the ratio of the cumulative explained variance has exceeded 85%, so only the front three eigenvalues of R is given in table 2.

Table 2 Eigenvalues of correlation matrix of energy conservation and emission reduction

Eigenvalues of the serial number	1	2	3
Eigenvalue λ	8.161	2.249	0.893
Ratio of the cumulative variance	0.703	0.790	0.859
Eigenvalue	α_1	α_2	α_3
Model (*)	$F_1=\alpha_1 \mathbf{Y}^T$	$F_2=\alpha_2 \mathbf{Y}^T$	$F_3=\alpha_3 \mathbf{Y}^T$

The factor load matrix is obtained by using the SPSS software shown in Table 3.

Table 3 Load matrix of the energy conservation and emissions reduction factor

	Z_1	Z_2	Z_3
y_1	0.99	-0.03	0.043
y_2	0.93	0.25	-0.042
y_3	-0.9	0.54	0.41
y_4	0.95	0.11	-0.11
y_5	-0.5	0.85	0.03
y_6	0.98	0.02	-0.01
y_7	0.95	-0.01	0.01
Y_8	0.95	0.20	-0.05
y_9	0.95	0.19	-0.03

Results can be obtained from table 3, the front three principal component ($Z_1 \sim Z_3$) reflects all the original information basically. The second principal component is closely related to waste water, waste gas and solid waste emission, showing the factors that power user enterprise have impacts on the environment.

Summary

This paper presents a principal component analysis method that extracting a small amount of the principal components which can reflect original index, which can effectively reduce the variable dimension without causing a large amount of information related to energy conservation and emission reduction loss. On this basis, according to the characteristics of power users, selecting the representative index of energy conservation and emission reduction, which can reflect the effect of energy conservation and emission reduction, and using principal component analysis to analyze the main factors closely related to the energy conservation and emission reduction. Finally, an example is presented to analyze relevant data about energy conservation and emission reduction for some electricity consumers in an actual region, reducing waste emissions and improving level of production technology to strengthen the effect of energy conservation and emission reduction.

Acknowledgement

This work was financially supported by China National science and technology support program projects funded projects (2013BAA01B04).

References

- [1] WEN Fushuang, Synergistic dispatch of PEVs charging and wind power in Chinese regional power grids, Automation of Electric Power Systems, 2011, 35(14), pp.1-10.
- [2] LIU Xiaoxiao, LIAN Guohai, The research of electric car charging infrastructure commercialization operation mode, Hunan electric power, 2011, 31(1), pp.59-62.
- [3] XU Fan, YU Guoqin, GU Linfeng, et al. Tentative analysis of layout of electrical vehicle charging station, East China Electric Power, 2009, 37(10), pp.1678-1682.
- [4] XI Jianbo, WEN Yubin, LI Rui, Discussion on demand forecast method for EV charging facilities. Jiangxi electric power, 2010, 34(5):1-5.
- [5] CHEN Liangliang, ZHANG Beibei, ZHOU Bin, et al, Optimal allocation of carbon capture systems in power plants, Automation of Electric Power Systems, 2011, 35(7), pp.81-85.