Principal Component Analysis of Modal interval-valued Data with Constant Numerical Characteristics

WANG Huiwen[1,2], CHEN Meiling[1,2], LI Nan[1,2], WANG Lanhui[3]

1. *School of Economics and Management, Beihang University, Beijing 100191, China*
2. *Research Center of Complex Data Analysis，Beihang University，Beijing 100191，China*
3. *School of Economics and Management，Beijing Forestry University，Beijing 100083，China*

Abstract: Modal interval-valued data is one of the most important types of symbolic data and each unit of its matrix contains a histogram or a distribution function. In this paper, a new method through Principal Component Analysis of modal interval-valued data is discussed. This Principal Component Analysis (PCA) method aims to reduce the dimensions of a large dataset by reconstructing the covariance matrix. The fundamental elements of the covariance matrix such as mean, variance and the covariance and their definition method is important in Principal Component Analysis. Some of the current researches on Principal Component Analysis of modal interval-valued data have contributions to dimension reduction of modal interval-valued data by transforming the histogram-valued data into interval data. In other existing methods, the definition of mean is in distributive data form and the mean is an average level for all modal-valued data observations. However, data centralization based on the mean defined this way actually obtains the residual distribution. The result of Principal Component Analysis in accordance with the matrix of residual distributions may thus fail to present the essential variation of the original data accordingly. In this paper, we define numerical characteristics of modal interval-valued data as real constants which can make full use of information in histograms. Centralization in terms of constant numerical characteristics is to relocate the modal-valued variances as a whole to get original histograms whose gravity center is settled on the origin. Therefore, the Principal Component Analysis of modal interval-valued data with constant numerical characteristics based on the obtained covariance matrix is proposed. Simulation proves the effectiveness of the proposed method.

Key words: Modal interval-valued data; Principal Component Analysis; Constant Numerical Characteristics

1 introduction

Symbolic data analysis method is one of the most groundbreaking theoretical achievements in modern statistical data analysis field. Modal interval-valued data is one of the symbolic data types and each unit of a high-dimensional modal interval-valued data matrix contains a histogram or a distribution function.

This paper will focus on PCA of modal interval-valued data. The routine method applied on PCA on modal interval-valued data is transforming it into interval data by a certain transformation method. For instance, Rodriguez O., Diday E., Winsberg S.[1] (2000) and Sun Makosso Kallyth, Edwin Diday[2] (2010) ; thus it can be analyzed by Principal Component Analysis method of interval-valued data.

A more direct method for histogram PCA was presented by P. Nagabhushan and R. Pradeep Kumar[3] (2007). In the paper, they defined unit histogram, null histogram, and the basic arithmetic operations of histogram such as addition, subtraction, multiplication, division and proposed a histogram PCA. However, means of histogram variables based on the above method is in histogram form, and data centralization obtains the residual histograms consequently. The result of PCA histogram with the matrix of residual histograms may thus fail to represent the essential variation of the original data accordingly.

In this paper, we attempt to explore a new PCA method for modal interval-valued variables. Based on the method of numerical characteristics integral calculation on continuous random variables in probability theory, we firstly define constant numerical characteristics about modal interval-valued data, then implement them for dimension reduction modeling of model interval-valued dataset through PCA. The proposed method not only can make use of complete information in the histogram, but also may give a more reliable conclusion, since data centralization on the basis of the constant numerical characteristics. Furthermore, an approximate method to calculate the linear combination of modal interval-valued variables is given according to the algorithm of univariate histograms theory of histogram-valued data[4] (2006) combined with

Moore algebra[5] (1962) in interval data analysis. Thus the projecting original modal interval-valued data to principal axes can be realized.

This paper is structured as follows: Section 2 introduces several basic definitions about numerical characteristics of modal interval-valued data and the derivation process and calculation steps of PCA on modal interval-valued data; simulation is conducted in section 3 to validate the effectiveness of the proposed method; the last section gives out the summary.

2 Methodology

We consider a $n \times p$ data matrix $\mathbf{X}_{n\times p} = (x_{ij})_{n\times p}$, which is called modal interval-valued data, and whose elements are all random variables that follow a histogram or a distribution function.

Here $x_{ij} = \{I_{ij}, f_{ij}\}$ means the random variable ξ defined on the field of definitions I_{ij} with the density function $f_{ij}(\xi)$。For histogram data, the density function of histogram data $f_{ij}(\xi)$ would be denoted as follows,

$$f_{ij}(\xi) = \begin{cases} \dfrac{p_{ij}^k}{x_{ij}^{k+1} - x_{ij}^k}, x_{ij}^k \le \xi < x_{ij}^{k+1} (k = 1, 2, \cdots, K_{ij}) \\ 0 \quad ,\text{else} \end{cases}, i = 1, \cdots, n; j = 1, \cdots, p, \tag{1}$$

Subject to:

$$0 \le p_{ij}^k \le 1, \quad \sum_{k=1}^{K_{ij}} p_{ij}^k = 1 \cdot$$

where K_{ij} is the number of modalities of x_{ij}, and $I_{ij}^k = [x_{ij}^k, x_{ij}^{k+1})$ is the *j*th sub-interval of I_{ij}, which satisfies $I_{ij} = \bigcup_k I_{ij}^k = [x_{ij}^1, x_{ij}^{K_{ij}+1})$, p_{ij}^k is the frequency of I_{ij}^k. It is assumed that within each sub-interval I_{ij}^k, the random variable ξ is uniformly distributed across the sub-interval. Hence, for histogram data x_{ij} also can be denoted as follows:

$$x_{ij} = \{I_{ij}, f_{ij}\} = \{\xi \in [x_{ij}^k, x_{ij}^{k+1}), p_{ij}^k, k = 1, \cdots, K_{ij}\}, i = 1, \cdots, n; j = 1, \cdots, p.$$

2.1 The first moment, second moment and second order mixed moment

According to classical probability theory, the first moment, second moment and second order mixed moment of modal-valued variable can be defined as follows:

Definition 1. For a modal-valued variable $\mathbf{X}_j$, the *first moment* is given by

$$\mu_j = \mathrm{E}(\mathbf{X}_j) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}(x_{ij}), \tag{2}$$

Where the first moment of unit x_{ij} is defined as

$$\mathrm{E}(x_{ij}) = \int_{-\infty}^{+\infty} \xi \cdot f_{ij}(\xi) d\xi = \sum_{\mathrm{k}} \int_{x_{ij}^k}^{x_{ij}^{k+1}} \xi \cdot \frac{p_{ij}^k}{x_{ij}^{k+1} - x_{ij}^k} d\xi = \sum_{\mathrm{k}} \frac{1}{2} p_{ij}^k \left(x_{ij}^{k+1} + x_{ij}^k\right), \tag{3}$$

Accordingly, the *centralization* of x_{ij} is given by

$$y_{ij} = x_{ij} - \mathrm{E}\left(\mathbf{X}_j\right) = \{\xi \in [x_{ij}^k - \mathrm{E}\left(\mathbf{X}_j\right), x_{ij}^{k+1} - \mathrm{E}\left(\mathbf{X}_j\right)), p_{ij}^k, k = 1, \cdots, K_{ij}\}, i = 1, \cdots, n; j = 1, \cdots, p. \tag{4}$$

Definition 2. Given any two modal-valued variables $\mathbf{X}_j$ and $\mathbf{X}_k$ $(j \ne k)$, the *second order mixed moment* is defined as

$$\mathrm{E}\left(\mathbf{X}_j \cdot \mathbf{X}_k\right) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left(x_{ij}x_{ik}\right), \tag{5}$$

Where

$$\mathrm{E}\left(x_{ij}x_{ik}\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \xi\eta \cdot f_{ij}(\xi) f_{ik}(\eta) d\xi d\eta = \int_{-\infty}^{\infty} \xi \cdot f_{ij}(\xi) d\xi \cdot \int_{-\infty}^{\infty} \eta \cdot f_{ij}(\eta) d\eta = \mathrm{E}\left(x_{ij}\right) \cdot \mathrm{E}\left(x_{ik}\right), \tag{6}$$

Here x_{ij} and x_{ik} are supposed to be independent.

Definition 3. For any modal-valued variable $\mathbf{X}_j$, the *second moment* is defined by

$$\mathrm{E}\left(\mathbf{X}_j^2\right)=\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(x_{ij}^2\right), \tag{7}$$

Where

$$\mathrm{E}\left(x_{ij}^2\right)=\int_{-\infty}^{\infty}\xi^2\cdot f_{ij}(\xi)d\xi=\sum_k\int_{x_{ij}^k}^{x_{ij}^{k+1}}\xi^2\cdot\frac{p_{ij}^k}{x_{ij}^{k+1}-x_{ij}^k}d\xi=\sum_k\frac{1}{3}p_{ij}^k\left[\left(x_{ij}^{k+1}\right)^2+x_{ij}^{k+1}\cdot x_{ij}^k+\left(x_{ij}^k\right)^2\right]. \tag{8}$$

Therefore the covariance and variance of modal-valued variables $\mathbf{X}_j,\mathbf{X}_k$ are as follows:

$$\mathrm{Cov}\left(\mathbf{X}_j,\mathbf{X}_k\right)=\mathrm{E}\left(\left(\mathbf{X}_j-\mathrm{E}\left(\mathbf{X}_j\right)\right)\left(\mathbf{X}_k-\mathrm{E}\left(\mathbf{X}_k\right)\right)\right)=\mathrm{E}\left(\mathbf{X}_j\cdot\mathbf{X}_k\right)-\mathrm{E}\left(\mathbf{X}_j\right)\cdot\mathrm{E}\left(\mathbf{X}_k\right), \tag{9}$$

$$\mathrm{D}\left(\mathbf{X}_j\right)=\mathrm{E}\left(\left(\mathbf{X}_j-\mathrm{E}\left(\mathbf{X}_j\right)\right)^2\right)=\mathrm{E}\left(\mathbf{X}_j^2\right)-\mathrm{E}\left(\mathbf{X}_j\right)^2. \tag{10}$$

2.2 Linear combination algorithm of modal interval-valued variables

Linear combination algorithm for interval-valued variables has been introduced by Moore(1962), the definition can be expressed as follow:

Definition 4. Given p interval-valued variables $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$, all with n observations and real numbers $a_j\in\mathbb{R}(j=1,2,\cdots,p)$, each observation can be regarded as a **hypercube**. Define an interval-valued variable $\mathbf{Y}$ as a linear combination of $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$, viz.

$$\mathbf{Y}=\sum_{j=1}^{p}a_j\mathbf{X}_j=\left(\left[\underline{y}_1,\overline{y}_1\right],\left[\underline{y}_2,\overline{y}_2\right],\cdots,\left[\underline{y}_n,\overline{y}_n\right]\right)', \tag{11}$$

where $\underline{y}_i=\sum_{j=1}^{p}a_j\left[\tau\underline{x}_{ij}+(1-\tau)\overline{x}_{ij}\right]$ and $\overline{y}_i=\sum_{j=1}^{p}a_j\left[(1-\tau)\underline{x}_{ij}+\tau\overline{x}_{ij}\right]$, with $\tau=\begin{cases}0, & a_j\le 0\\ 1, & a_j>0\end{cases}$.

With Definition 4, an algorithm to calculate the linear combination of modal interval-valued variables can be presented. Given p modal interval-valued variables $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$, and real numbers $a_j\in\mathbb{R}(j=1,2,\cdots,p)$, an modal interval-valued variable $\mathbf{Y}$ as a linear combination of $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$ can be defined as follows,

$$\mathbf{Y}\approx\sum_{j=1}^{p}a_j\mathbf{X}_j, \tag{12}$$

where $\mathbf{Y}$ is a histogram vector，each element can be expressed as follows,

$y_i=\{I_i',f_i'\}=\{\xi\in[y_i^k,y_i^{k+1}),p_i^{k\prime};k=1,2,\cdots,K_i\}$, where $K_i=\max\{K_{ij}\,|i=1,\cdots,n;j=1,\cdots,p\}$。

For the ij th histogram x_{ij}, the number of modalities K_{ij}, each sub-interval has its density function as $p_{ij}^k,k=1,\cdots,K_{ij}$, thus the i th observation contains $\prod_{j=1}^{p}K_{ij}$ hypercubes in p dimension space. The density of each hypercube is the product of the density of corresponding sub-interval and be denoted as p_{iu}', $u=1,2,\cdots,\prod_{j=1}^{p}K_{ij}$. According to formula (11) in Definition 4, calculating linear combination to all the hypercubes, we can obtain I_{iu}', $u=1,2,\cdots,\prod_{j=1}^{p}K_{ij}$. Then we can get the maximum and the minimum values of I_{iu}', $u=1,2,\cdots,\prod_{j=1}^{p}K_{ij}$, denote as I_i'. The sub-interval $I_i^{k\prime}=[y_i^k,y_i^{k+1}),k=1,2,\cdots,K_i$ can be obtained by dividing I_i' into K_i parts. Projecting the above I_{iu}' $u=1,2,\cdots,\prod_{j=1}^{p}K_{ij}$ to $I_i^{k\prime}$, which is denoted as follows.

$$p_i^{k\prime}=\sum_u p_{iu}'\frac{\left\|I_{iu}'\cap I_i^{k\prime}\right\|}{\left\|I_{iu}'\right\|},u=1,2,\cdots,\prod_{j=1}^{p}K_{ij}. \tag{13}$$

According to the above mentioned steps, we can get the linear combination of all variables in the i th histogram observation.

2.3 The Algorithm

For convenience, suppose the modal interval-valued vectors are all centralized. With the above definitions, we begin to derive the Principal Component Analysis method for modal interval-valued data. Similar to the numeric case, the *k*th modal interval-valued data PC $\mathbf{Y}_k\,(k=1,2,\cdots,p)$ is a linear combination of $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$, i.e., $\mathbf{Y}_k=\mathbf{X}\mathbf{u}_k=\sum_{j=1}^{p}u_{kj}\mathbf{X}_j$, where $u_{kj}\in\mathbb{R}(j=1,2,\cdots,p)$, with the constraints of $\|\mathbf{u}_k\|=1$ and $\mathbf{u}_k'\mathbf{u}_l=0\ \ (l=1,2,\cdots,p,\ \ l\neq k)$. Also, the first m principal components $\mathbf{Y}_1,\mathbf{Y}_2,\cdots,\mathbf{Y}_p$ must maximize total variance to represent the original information carried by $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$. According to definitions proposed above, we have

$$
\begin{aligned}
\mathrm{D}(\mathbf{Y}_k) &= \mathrm{E}(\mathbf{Y}_k^2)\\
&= \mathrm{E}\left(\left(u_{k1}\mathbf{X}_1+u_{k2}\mathbf{X}_2+\cdots+u_{kp}\mathbf{X}_p\right)^2\right)\\
&= \left(u_{k1},u_{k2},\cdots,u_{kp}\right)\begin{pmatrix}\mathrm{E}(\mathbf{X}_1^2) & \mathrm{E}(\mathbf{X}_1\cdot\mathbf{X}_2) & \cdots & \mathrm{E}(\mathbf{X}_1\cdot\mathbf{X}_p)\\ \mathrm{E}(\mathbf{X}_2\cdot\mathbf{X}_1) & \mathrm{E}(\mathbf{X}_2^2) & \cdots & \mathrm{E}(\mathbf{X}_2\cdot\mathbf{X}_p)\\ \vdots & \vdots & \ddots & \vdots\\ \mathrm{E}(\mathbf{X}_p\cdot\mathbf{X}_1) & \mathrm{E}(\mathbf{X}_p\cdot\mathbf{X}_2) & \cdots & \mathrm{E}(\mathbf{X}_p^2)\end{pmatrix}\begin{pmatrix}u_{k1}\\ u_{k2}\\ \vdots\\ u_{kp}\end{pmatrix}\\
&= \mathbf{u}_k'\mathbf{V}\mathbf{u}_k, \qquad (14)
\end{aligned}
$$

where $\mathbf{V}$ represents the covariance matrix of $\mathbf{X}_1,\mathbf{X}_2,\cdots,\mathbf{X}_p$.

The following derivation is the same with the classical PCA that for numeric data, i.e., looking for m orthogonalised vectors $\mathbf{u}_1,\mathbf{u}_2,\cdots\mathbf{u}_p$ to achieve maximization of $\sum_{k=1}^{p}\mathrm{D}(\mathbf{Y}_k)$ with $\mathrm{D}(\mathbf{Y}_1)\geq\mathrm{D}(\mathbf{Y}_2)\geq\cdots\geq\mathrm{D}(\mathbf{Y}_p)$ by solving equations of $\mathbf{V}\mathbf{u}_p=\lambda_p\mathbf{u}_p\,(1\leq k\leq p)$. Thus, $\mathbf{u}_1,\mathbf{u}_2,\cdots\mathbf{u}_p$ are the orthonormal eigenvectors of $\mathbf{V}$, corresponding to the eigenvalues $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_p$. By algorithm of linear combination of modal-valued variables, see formula (12), we finally get the *k*th modal-valued PC $\mathbf{Y}_k=\mathbf{X}\mathbf{u}_k$.

3 Experimental Results of Synthetic Data Sets

This section we conduct a comparison between PCA of modal interval-valued data and PCA of numeric data. The histogram dataset will be generated in Monte-Carlo simulation method. And the numeric dataset corresponding to this histogram dataset will be obtained by differentiating I_{ij}, the field of definitions of the histogram. The comparison will focus on the eigenvalues and eigenvectors of the covariance matrix of histogram dataset and numeric dataset. It'll be concluded that the eigenvalues and eigenvectors are similar for the two types of data. The calculation results of numeric dataset tend to the histogram dataset's results when the number of differentiation in I_{ij} becomes lager.

3.1. Dataset

3.1.1 Histogram dataset

We implement Monte-Carlo simulation method to generate a 50×4 histogram dataset. For generality, we take all the number of modalities as three. Therefore, the ij histogram $x_{ij}=\{I_{ij},f_{ij}\}$ can be generated randomly by the two steps as follows.

1) Define the generation method of field of definitions: generate the center $x_{ij}^C\sim U[-5,5]$ randomly and the radius $x_{ij}^R\sim U[1,10]$, thus we can get the interval of histogram $I_{ij}=[x_{ij}^C-x_{ij}^R,x_{ij}^C+x_{ij}^R]$ and divide the interval I_{ij} equally into three parts, we obtain $I_{ij}^k=[x_{ij}^k,x_{ij}^{k+1}),k=1,\cdots,3$.
2) Define the generation method of density(frequency) function: generate three data $q_{ij}^1,q_{ij}^2,q_{ij}^3$, where

$q_{ij}^k \sim U[0,1], k=1,2,3$, then orthogonalize the data as follows:

$$Q_{ij} = \sum_{k=1}^{3} q_{ij}^k \text{ , } \quad p_{ij}^k = q_{ij}^k / Q_{ij}, k=1,2,3 .$$

Meantime, it satisfies the constraint $\sum_{k=1}^{3} p_{ij}^k = 1$ and we can get the density function of the ij th histogram in the dataset.

3.1.2 Numeric dataset

When we get the histogram dataset, for each histogram we get a numeric dataset by differentiating the field of definitions. That is for each histogram $x_{ij} = \{I_{ij}, f_{ij}\}$, to choose m positive integer numbers in the field of definition interval I_{ij} with dividing the kth sub-interval equally into $m \cdot p_{ij}^k (k=1,\cdots,K)$ parts, therefore the numeric dataset bide by approximately to the distribution of x_{ij}. We call m as the number of differentiation. The larger of the number of differentiation, the more similar of numeric dataset to the histogram. For example, for two dimension histogram,

$$(x,y) = (\{[2,3),0.2;[3,4),0.5;[4,5],0.3\}, \quad \{[5,6),0.2;[6,7),0.7;[7,8],0.1\}$$

Fig1 shows the conditions of m=20 and m=50 respectively. In the left chart there are 441 points while there are 2601 in the right chart.

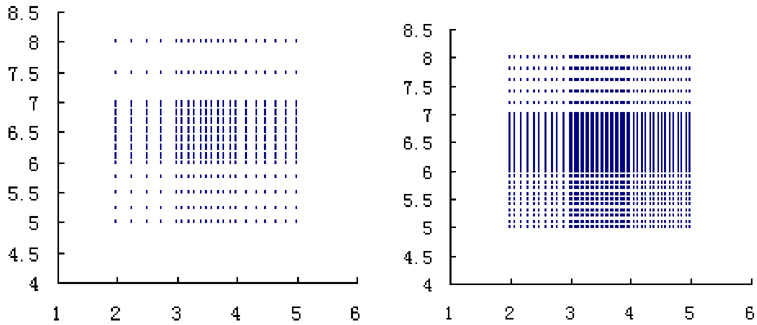

Fig1. The diagram of numeric dataset with the conditions of m =20 and m =50

Therefore, a p dimension histogram is expanded into a $(m+1)^p$ numeric dataset, on which classical multivariate analysis can be performed. It proves the method proposed in this paper is reasonable, if the calculation results obtained by PCA tend to the histogram dataset's while the number of differentiation is increasing

3.2 The simulation result of PCA of histogram data

For the 50×4 histogram dataset generated above, we calculate its eigenvalues and eigenvectors of its covariance matrix by PCA method; the eigenvalues are sorted descending as follows:

λ_1^*=23.5362; λ_2^*=20.4752 ; λ_3^*=19.3498 ; λ_4^*=15.0953.

And the corresponding eigenvectors are showed in table 1.

Table 1. Eigenvectors of corresponding eigenvalue

$\mathbf{u}_1^*$	$\mathbf{u}_2^*$	$\mathbf{u}_3^*$	$\mathbf{u}_4^*$
-0.6325	0.6589	-0.1626	0.3734
0.1177	0.4471	-0.4286	-0.7763
0.6725	0.2391	-0.4836	0.5067
0.3659	0.5557	0.7456	-0.0362

In the following part, we compare the eigenvalues and eigenvectors of histogram data and numeric data. Assume the number of differentiations are 10、20、30、40、50、70、100 and 200, we calculated the eigenvalues λ_2^m and eigenvectors $\mathbf{u}_j^m$ $j=1,2,\cdots,4; m=10,20,30,40,50,70,100,200$. For the eigenvalues, we take Absolute Error to compare the both, i.e.:

$$AE(\lambda_j, m) = \left|\lambda_j^m - \lambda_j^*\right|, j=1,\cdots,4; m=10,20,30,40,50,70,100,200 .$$

For the eigenvectors, we take the Absolute Cosine Value, which is defined as follows:

$$\mathrm{ACV}_j^m = \left| \frac{\mathbf{u}_j^{*\prime} \mathbf{u}_j^m}{\left\| \mathbf{u}_j^* \right\| \cdot \left\| \mathbf{u}_j^m \right\|} \right|, \quad m = 10, 20, 30, 40, 50, 70, 100, 200,\ j = 1, \cdots, 4$$

The horizontal axis in Fig2 denotes the number of differentiation. In the left chart of Fig2 the curves show the AE changes while the number of differentiation is increasing, we can see the AE is getting smaller while the number of differentiation is on the increase. The vertical axis of the right chart denotes ACV_j^m, the results shows the absolute error of ACV_j^m converges to 1 while the number of differentiation is increasing which illustrates the angle between the two compared vectors converges to 0. The obtained components are more alike when the similarity is high.

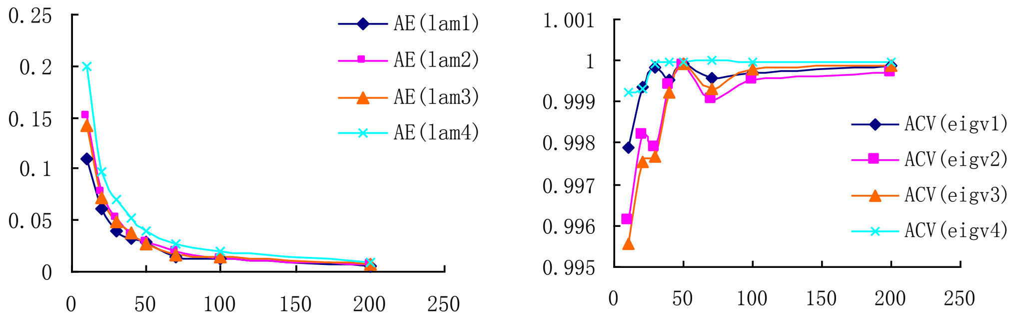

Fig2.The AE of eigenvalues and ACE of eigenvectors

Finally, we contrast the sample distribution of first principle component of numeric data and histogram data. The first principle component of histogram data is obtained by the linear combination of histogram as section 2.2, while the first principle component of numeric data is calculated the empirical distribution frequency. For saving computation, we take m=10, the sample size of number dataset is $(10+1)^4 \times 50$. Then the results of Two Independent Samples Kolmogorov-Smirnov test from the 50 samples show that both distributions are consistent.

4 Conclusion

This paper proposed a new PCA method based on numeric characteristic constant type for modal interval-valued variables, which is drawn on the method of numerical characteristics integral calculation on continuous random variables in probability theory, In the process of computation, the method not only adopts the complete information of histogram, but also make the characteristic analysis clear and the result is reasonable and precise. The simulation proves the rationality and effectiveness of the method.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 70821061, 70771004, 71031001).

References

[1] Rodriguez O., Diday E., Winsberg S. Generalization of principal components analysis to histogram data, Workshop of Symbolic Data Analysis, 4th Eur. Conf. Principles and Practice of Knowledge Discovery in Databases, Lyon, France, Sept. 2000.

[2] Kallyth S.M., Diday E., Symbolic PCA of Compositional Data, 19th International Conference on Computational Statistics, Paris - France, August 22-27, 2010.

[3] Nagabhsushan P., Kumar P., Principal Component Analysis of Histogram Data. Springer-Verglag Berlin Heidelberg, EdsISNN Part II LNCS 4492, 1012-1021, 2007.

[4] Billard L., Diday E., Descriptive Statistics for Interval-valued Observations in the Presence of Rules, Computational Statistics, 21:187-210, 2006.

[5] Moore Ramon E., Baker Kearfott R., Cloud Michael J., Introduction to Interval Analysis, 223 pages. Siam. ISBN 978-0-898716-69-6, 2009.