

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221848689>

Comparing patterns of component loadings: Principal Component Analysis (PCA) versus Independent Component Analysis (ICA) in analyzing multivariate non-normal data

Article in Behavior Research Methods · February 2012

DOI: 10.3758/s13428-012-0193-1 · Source: PubMed

CITATIONS

8

READS

196

2 authors, including:



Se-Kang Kim

Fordham University

62 PUBLICATIONS 460 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Profile analysis and correspondence analysis [View project](#)

Comparing patterns of component loadings: Principal Component Analysis (PCA) versus Independent Component Analysis (ICA) in analyzing multivariate non-normal data

Donghoh Kim · Se-Kang Kim

Published online: 21 February 2012
© Psychonomic Society, Inc. 2012

Abstract Principal component analysis identifies uncorrelated components from correlated variables, and a few of these uncorrelated components usually account for most of the information in the input variables. Researchers interpret each component as a separate entity representing a latent trait or profile in a population. However, the components are guaranteed to be independent and uncorrelated only when the multivariate normality of the variables is assumed. If the normality assumption does not hold, components are guaranteed to be uncorrelated, but not independent. If the independence assumption is violated, each component cannot be uniquely interpreted because of contamination by other components. Therefore, in the present study, we introduced independent component analysis, whose components are uncorrelated and independent even when the multivariate normality assumption is violated, and each component carries unique information.

Keywords Independent Component Analysis (ICA) · Principal Component Analysis (PCA) · Multivariate normality · Profile analysis

The main purpose of principle component analysis (PCA) is to transform correlated metric variables into a much smaller number of uncorrelated variables called *principal components* (PCs), which contain most of the information from the

original variables. PCA is often used as a preliminary analysis to determine the number of factors. However, in recent profile analysis studies, unrotated PCs or multidimensional scaling dimensions have been used for latent profiles in a population, rather than latent variables or factors (Davison, Gasser, & Ding, 1996; Davison, Kim, & Close, 2009; Frisby & Kim, 2008; Kim, Davison, & Frisby, 2007). In the present study, we took the second path for PCs and viewed unrotated components as latent profiles. Most social science researchers (e.g., psychologists or educational researchers) consider uncorrelated PCs to be independent identities (e.g., latent variables or latent profiles), and interpret them accordingly. However, uncorrelated PCs carry independence properties only when the multivariate normality assumption is met. If the normality assumption is violated, uncorrelated PCs are not necessarily independent. This fact implies that information included in one PC might be shared with information in other PCs. If the PCs are not independent, although they are uncorrelated, a portion of a trait from one component is shared with a trait in the other component, and as a result, each component loading does not carry a unique effect in a given dimension.

We frequently deal with observations from skewed variables. In such cases, the multivariate normality assumption does not hold, and PCs estimated from the data are not independent. The dependent PCs, even uncorrelated, cannot be automatically assumed to represent an exclusive trait for each component, and interpretation of the dependent PCs is contaminated. To circumvent the dependency of components of PCA, we introduced independent component analysis (ICA), which estimates components as statistically independent as possible. Through an analysis of a real example, we demonstrated the difference of component loadings between ICA and PCA by illustrating their component loading profiles. We hope that the ICA procedure

D. Kim (✉)
Department of Applied Mathematics, Sejong University,
Seoul 143-747, Korea
e-mail: donghoh.kim@gmail.com

S.-K. Kim
Department of Psychology, Fordham University,
Bronx, NY, USA
e-mail: sekim@fordham.edu

helps researchers to interpret the uncontaminated underlying structure of the data in terms of the most independent components for non-normal data.

Method

PCA transforms n observations with T random variables $\mathbf{X} = (X_1, \dots, X_t, \dots, X_T)$, $X_t = (X_{1t}, \dots, X_{it}, \dots, X_{nt})'$ to PCs through a linear combination of loadings. That is, the matrix of PCs \mathbf{Y} is

$$\mathbf{Y} = \mathbf{X}\mathbf{F},$$

where \mathbf{F} is a $T \times T$ PC loading matrix. The PC loading matrix consists of the eigenvectors of the covariance matrix of \mathbf{X} . The transformation based on the covariance matrix structure assures the uncorrelatedness of PCs and, furthermore, independence of PCs under the multivariate normality assumption held.

The ICA was introduced in the early 1980s motivated by the *cocktail-party problem* or *blind source separation*. Suppose that three separate (independent) people are speaking simultaneously at a cocktail party, and three recording machines record the mixture of people speaking through the microphone. Each person's conversation, say S_1 , S_2 , and S_3 , are directly unobservable, but recordable through microphone. The cocktail-party problem is simply stated as how to recover the original separate/independent sources S_1 , S_2 , and S_3 from the recorded (mixed or contaminated)

variables X_1 , X_2 , and X_3 . The ICA seeks a $T \times T$ unmixing matrix \mathbf{W} so that ICs, $\mathbf{S} = \mathbf{X}\mathbf{W}$, are as statistically independent as possible. If the distribution of the recorded variables follows the multivariate normal distribution, the covariance matrix will play an important role to recover the original sources. For the multivariate normal distribution, the mean and covariance structure determine the characteristics of the variables at hand, and the linear transformation with the eigenvectors of covariance matrix will produce the uncorrelated—and at the same time independent—components, and PCA and ICA results will be similar. However, as we can observe in the cocktail-party problem, it is not certain in reality that the multivariate normality of the observed variables holds. In such a case, covariance structure cannot fully explain the behavior of the observed variables. The motivations for using PCA and ICA might be different, but PCA and ICA have common aspects and applications, such as both methods can be used as data reducing methods and as tools to identify latent structures (e.g., latent profiles or factors) of the observed variables.

The methods for deriving the unmixing matrix consist of two steps: (a) specifying the criterion to measure the independence of ICs, and (b) optimizing the statistical independence criterion. Several methodologies have been proposed according to the independence criterion, such as likelihood (Pham, Garrat, & Jutten, 1992), mutual information (Comon, 1994), and so on. The mutual information $I(X_1, \dots, X_T)$ measures the dependence among T random variables X_1, \dots, X_T as $I(X_1, \dots, X_T) = \sum_{k=1}^T H(X_k) - H(\mathbf{X})$, $\mathbf{X} = (X_1, \dots, X_T)'$, where the differential entropy H of a

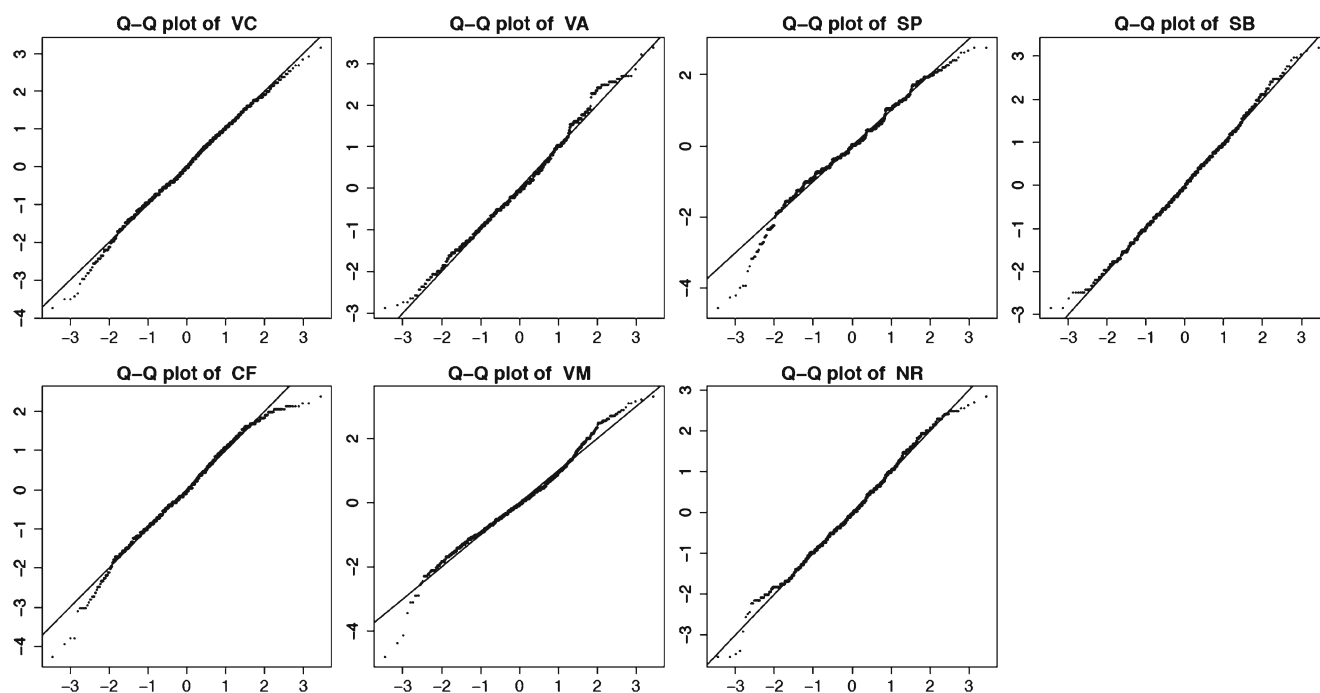


Fig. 1 Q–Q plot of each subtest

Table 1 Dimension profiles by independent component analysis and principal component analysis

	Independent Component Analysis			Principal Component Analysis		
	Component 1	Component 2	Component 3	Component 1	Component 2	Component 3
VC	-0.5943	0.2921	-0.0485	-0.4194	0.3287	-0.1701
VA	-0.3841	0.1299	0.1203	-0.4148	0.1615	0.0222
SP	0.3314	0.2053	0.7945	-0.3295	0.2497	0.7512
SB	-0.8739	0.0999	-0.4691	-0.3590	0.1152	-0.5881
CF	-0.3817	0.1377	0.1412	-0.4287	0.1710	0.0407
VM	0.0357	-0.7237	0.2002	-0.3170	-0.7381	0.1933
NR	-0.3641	-0.4603	-0.0926	-0.3609	-0.4649	-0.1467
Variance	3.320	0.831	0.763	3.320	0.831	0.763

random variable (or random vector) Y with density p is defined as $H(Y) = -\int p(y) \log p(y) dy$. The mutual information is zero if and only if the random variables are statistically independent. Thus, ICs can be derived by minimizing the mutual information among the components. To optimize the independence criterion, a stochastic gradient descent algorithm or fixed-point iteration algorithm has been adapted (see Hyvärinen, Kauhunen, & Oja, 2001 for mathematical details). Like the loading matrix of PCA, the unmixing matrix W measures the prominence of the observed variables constructing the components. That is, as the absolute value of the elements of the unmixing matrix increases, the corresponding variable has a strong effect on the components.

Although PCA and ICA pursue different approaches to extract components, we have tried to make a connection between PCs and ICs through the loading matrix F of PCA and the unmixing matrix W of ICA. There exists the matrix T such that $FT = W$. Therefore, ICs can be viewed and explained as the rotation of PCs, which produces uncorrelated and independent components.

Data analysis and results

We provided the possible applications of the ICA procedure by analyzing the norm sample of the Woodcock–Johnson III (WJ–III) Tests of Cognitive Ability (Woodcock, McGrew, & Mather, 2001). For illustration, the seven standardized cognitive subtest scores are analyzed by both PCA and ICA: VC (verbal comprehension), VA (visual-auditory learning), SP (spatial relations), SB (sound blessing), CF (concept formation), VM (visual matching), and NR (numbers reversed). After listwise deletion (of missing) on the seven cognitive subtest scores of 8,782 individuals, the sample size became 3,825. After excluding participants younger than 15 and older than 65, the sample was reduced to 1,767.

Figure 1 illustrates the Q-Q plots of the seven subtests. Distributions of most subtests are skewed. Normality does not hold for marginal distributions and thus the multivariate normality assumption does not hold. Therefore the observed PCs are not necessarily independent, and we need to be cautious when interpreting the PC loading profiles (which are represented by columns of the loading matrix F).

We conducted ICA and PCA for WJ–III Tests data and investigated the patterns of the first three profiles and independence of components. For the analysis, each variable was standardized. The independent component analysis was implemented by the R package fastICA (Marchini, Heaton, & Ripley, 2010).¹

Table 1 and Fig. 2 illustrate the loading profiles of ICA and PCA, respectively. The solid line represents the ICA loading profile, and the dotted line represents the PCA loading profile.

From the comparison analysis of PCA and ICA, we made several observations with statistical implications. The loading profiles of ICA produced both uncorrelated and independent components, whereas the profiles of PCA gave only uncorrelated components, not independent components, since multivariate normality did not hold in our example data. To test the independence of ICs and PCs respectively, the χ^2 test was applied to each pair of components. First, each component was discretized into three groups; then, we constructed 3×3 cross tables and ran χ^2 independence test. Table 2 shows the p value of the independence test. Except for component 2 and component 3, the p values of PCs are much smaller than the large p values of ICs, which implies the strong dependence of PCs but independence of ICs. Each IC loading profile provides unique information for each dimension, but PC loading profiles do not. As shown

¹ Available at the Comprehensive R Archive Network (CRAN <http://CRAN.r-project.org/>)

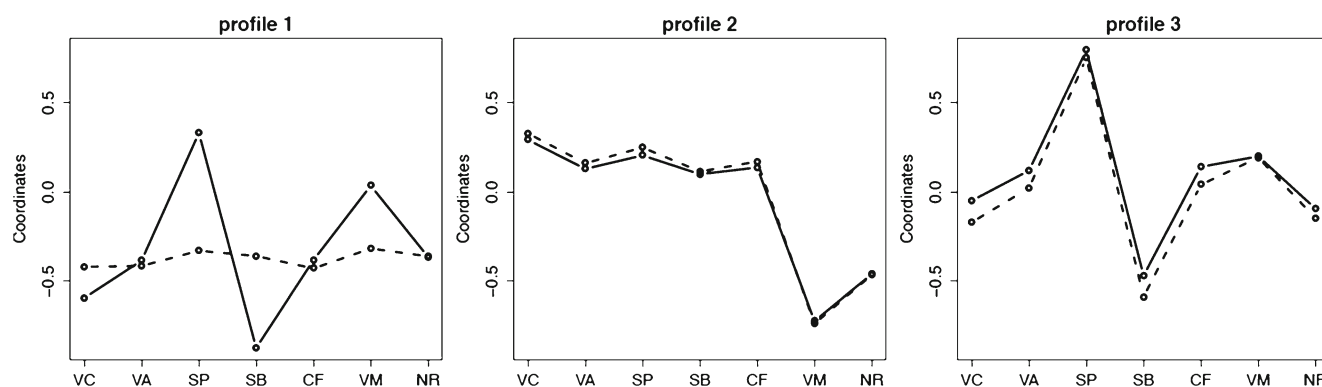


Fig. 2 Plot of dimension profiles by independent component analysis (*solid line*) and principal component analysis (*dotted line*)

in Fig. 2, the profile patterns of first PC and the first IC are quite different. As seen in Table 2, the first PC is not independent with the second and the third PCs. In other words, the first PC is contaminated with other PCs in its content characteristics, and the first PC is interpreted as a general factor overlapped with other components in its contents. However, the first IC presents its own uniqueness independent of other components. Of course, the second and third ICs are independent of each other, and the second and third PCs are as well, as shown in Table 2. Therefore, the second and third ICs and PCs are similar in their patterns (in Fig. 2). In short, the first IC has its uniqueness in its content characteristics, but the first PC does not.

Regarding data analysis aspects, the results of PCA and ICA can be interpreted as follows. In the first component profile, the magnitudes and patterns of loadings on the variables by ICA and PCA are quite different. The (unrotated) first PC profile is virtually flat, since it represents a general factor (or g) that is not independent of other subsequent components that signify group factors. However, the first IC profile has peaks of SP and VM and a valley of sound blessing. The profile may be labeled as “high visual relation versus low sound blessing,” and this profile is assumed to be independent of the other component profiles (as shown in Table 2).

The loading profile patterns for the second and third component were similar, and PC2 and PC3 profiles were indepen-

dent, as were IC2 and IC3 profiles (see Table 2). For the second component profile, the most IC loadings were smaller than the PC loadings, whereas for the third component profile, some PC loadings were smaller or larger than the IC loadings.

If researchers interpret each of the PCs (that are estimated from multivariate non-normal data) as a separate/unique entity across different dimensions, their interpretation is biased, since the components are not independent, although they are uncorrelated.

Since the criterion for choosing components for interpretation in ICA is not clear, we have tried to make a connection between PCA and ICA results. We view the (unmixing) IC loading matrix \mathbf{W} as a transformation of the PC loading matrix \mathbf{F} as follows:

- The component loadings are columns of \mathbf{F} or columns of \mathbf{W} , where for PCs, $\mathbf{Y} = \mathbf{X}\mathbf{F}$ and for ICs, $\mathbf{S} = \mathbf{X}\mathbf{W}$.
- The transform matrix \mathbf{T} can be found to connect the component loading matrix \mathbf{F} of PCA and the loadings matrix \mathbf{W} of ICA. That is, $\mathbf{F}\mathbf{T} = \mathbf{W}$.
- For the case of our example data, \mathbf{T} is

$$\mathbf{T} = \begin{pmatrix} 0.8967 & 0.0567 & -0.2135 \\ -0.1977 & 0.9513 & 0.0672 \\ 0.8999 & -0.0181 & 0.9413 \end{pmatrix}.$$

Table 2 p value of independence test for each pair of components

	Component 1 and Component 2	Component 1 and Component 3	Component 2 and Component 3
Independent component analysis	0.7087	0.7928	0.1950
Principal component analysis	0.0107	0.0004	0.1246

Summary and discussion

When researchers want to have uncorrelated and independent components for their studies, we recommend that they first check the multivariate normality of their data. If the multivariate normality assumption is met, the researchers can conduct PCA, and PCs will be uncorrelated and independent. However, if normality is violated, it will benefit the researchers to conduct both ICA with PCA. First, conduct PCA. PCA helps to determine dimensionality (or number of components) according to researcher's purpose or the characteristics of data at hand. Then, conduct ICA with the same number of components determined by the initial PCA. Note that for illustrative purpose, we included a three-component solution and investigated whether each component (from ICA and PCA) was independent of the other components. By sequential analysis of ICA, the researchers can identify ICs that correspond to the PC loading profile patterns as shown in Fig. 2.; then, they can interpret ICs either as latent variables or latent profiles.

Author note This research was supported for D. Kim by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0076223 and 2009-0086944). Thanks to Kevin McGrew at Institute for Applied Psychometrics for the data.

References

- Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing*, 36, 287–314.
- Davison, M. L., Gasser, M., & Ding, S. (1996). Identifying major profile patterns in a population: An exploratory study of WAIS and GATB pattern. *Psychological Assessment*, 8, 26–31.
- Davison, M. L., Kim, S.-K., & Close, C. (2009). Factor analytic modeling of within person variation in score profiles. *Multivariate Behavioral Research*, 44, 668–687.
- Frisby, C. L., & Kim, S.-K. (2008). Using profile analysis via multidimensional scaling (PAMS) to identify core profiles from the WMS-III. *Psychological Assessment*, 20, 1–9.
- Hyvärinen, A., Kauhunen, J., & Oja, E. (2001). *Independent component analysis*. New York: John Wiley & Sons.
- Kim, S.-K., Davison, M. L., & Frisby, C. L. (2007). Confirmatory factor analysis and profile analysis via multidimensional scaling. *Multivariate Behavioral Research*, 42, 1–32.
- Marchini, J. L., Heaton, C., & Ripley, B. D. (2010). *fastICA: FastICA Algorithms to perform ICA and Projection Pursuit*. Retrieved from <http://cran.rproject.org/web/packages/fastICA/index.html>
- Pham, D.-T., Garrat, P., & Jutten, C. (1992). Separation of a mixture of independent sources through a maximum likelihood approach. *Proceedings of European Signal Processing Conference*, 771–774.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca: Riverside Publishing.