Rotation in Inference and Description

Yasuharu Okamoto
Faculty of Integrated Arts and Social Sciences
Japan Women's University

Introduction

Rotation in inference from and description of data are discussed in this paper. Fig. 1 shows that data are generated by latent variables and information contained in the data is represented by components, the number of which is less than the dimensionality of the data. Both in inference and in description, we can determine the optimum spaces with respect to the data, but to identify the axes we need some criteria outside of statistics. In identifying the axes, which have substantial meanings, rotation is used. First, consider rotation in inference.

Rotation in Inference

Here I suppose that data are generated by a factor analytic model. I set the following linear model:

$$\mathbf{x} = \mathbf{f}\mathbf{A}' + \mathbf{res} \tag{1}$$

where $\mathbf{x} = \left[x_1 \cdots x_p\right]$ is a p-dimensional random vector of generated data, $\mathbf{f} = \left[f_1 \cdots f_q\right]$ is a q-dimensional random vector of latent variables, $\mathbf{A}$ a p×q matrix of factor pattern and $\mathbf{res} = \left[e_1 \cdots e_p\right]$ a p-dimensional random vector of residuals. $\mathbf{res}$ is assumed to be independent of $\mathbf{f}$.

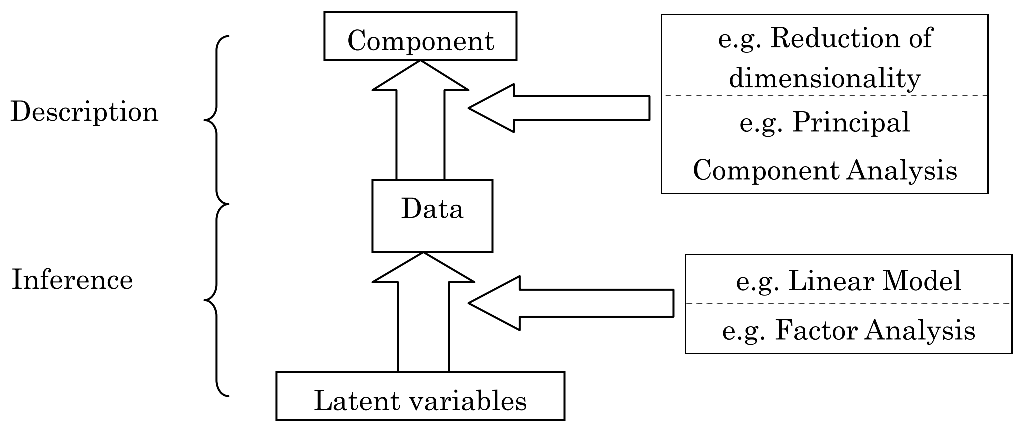

Figure 1. Inference and Description in Data Analysis.

Eq. (1) can be rewritten as follows,

$$\mathbf{x} = (\mathbf{fT})(\mathbf{T}^{-1}\mathbf{A}') + \mathbf{res} \qquad (2)$$

Rotation matrix $\mathbf{T}$ in eq. (2) means that we can identify only the space $S(\mathbf{f})$ spanned by $\mathbf{f}$. Many procedures have been proposed to identify $S(\mathbf{f})$.

From the model (1), we have a correlation matrix $\mathbf{R}$ of $\mathbf{x}$ as follows

$$\mathbf{R} = E(\mathbf{x}'\mathbf{x}) = \mathbf{A} \cdot E(\mathbf{f}'\mathbf{f}) \cdot \mathbf{A}' + E\left((\mathbf{res})' (\mathbf{res}) \right)$$

If the assumption

$$E\left((\mathbf{res})' (\mathbf{res}) \right) = \text{a diagonal matrix,}$$

which is needed to reduce the number of free parameters but seems to be implausible, is adopted, we can determine $\mathbf{A}$ so as to minimize the differences between corresponding off-diagonal elements of $\mathbf{R}$ and $\mathbf{AA}'$ with the convention $E(\mathbf{f}'\mathbf{f}) = \mathbf{I}$.

But, without the assumption

$$E\left((\mathbf{res})' (\mathbf{res}) \right) = \text{a diagonal matrix,}$$

more direct approach aims at minimizing the difference between $\mathbf{x}$ and $\mathbf{fA}'$. When the N observed data of $\mathbf{x}$ is denoted as an N×p matrix $\mathbf{X}$ and the corresponding values of the latent variables $\mathbf{f}$ by an N×q matrix $\mathbf{F}$, this direct approach (see Okamoto, 2005) estimates $\mathbf{F}$ and $\mathbf{A}$ by minimizing

$$\|\mathbf{X} - \mathbf{FA}'\|^2 \qquad (3)$$

But eq. (2) shows that by criterion (3), we can only determine the space $S(\mathbf{F})$. To find substantial axes, we need information outside of statistics

Rotation in Description

Next, let's consider rotation in description. Information contained in data $\mathbf{X}$ can be summarized in various ways, one of which is reduction of dimensionality of $\mathbf{X}$. The simplest method to see a low dimensional shape of the distribution of data $\mathbf{X}$ is to see a shadow of $\mathbf{X}$ on an appropriate space of lower dimensionality than that of $\mathbf{X}$. Mathematically, this can be done by the orthogonal projection $\mathbf{P}$, which results in the maximum dispersion of the projected data $\mathbf{XP}$ (Okamoto, 2006).

The orthogonal projection $\mathbf{P}$ determines just the space $\mathbf{XP}$, the shadow of $\mathbf{X}$ by $\mathbf{P}$. Coordinates of a point in $\mathbf{XP}$ are given by any basis $\{\mathbf{b}_1, \cdots, \mathbf{b}_q\}$. When $\mathbf{b}_i$'s are orthogonal, the matrix of coordinates are given by

$$\mathbf{XP}[\mathbf{b}_1 \cdots \mathbf{b}_q]$$

Any other basis $\{\mathbf{v}_1 \cdots \mathbf{v}_q\}$ can be related to $\{\mathbf{b}_1, \cdots, \mathbf{b}_q\}$ by some rotation matrix $\mathbf{T}$ as follows,

$$[\mathbf{v}_1 \cdots \mathbf{v}_q]' = \mathbf{T}[\mathbf{b}_1 \cdots \mathbf{b}_q]' \qquad (4)$$

By eq. (4), we can find substantial basis, which is determined by criterion outside of statistics.

Conclusion

If we see only mathematical calculations needed to minimize eq. (3) and those to find the optimum orthogonal projection $\mathbf{P}$, they are the same ones, which are based on singular value decomposition of $\mathbf{X}$, and called by the same name Principal Component Analysis. But, from the standpoint of substantial science, they are quite different. Minimization of eq. (3) seeks the best inference about the latent space, and the orthogonally projected data $\mathbf{XP}$ is optimum description of $\mathbf{X}$ in the space of lower dimensionality than that of $\mathbf{X}$. Inference and description should be distinguished. Inference is based on some model, some aspects of which are estimated from the data $\mathbf{X}$. On the other hand, description is reduction of data, which summarizes some aspects of the data $\mathbf{X}$. Description does not need any model, but inference is always linked to some model. Hence, the method to minimize eq. (3) and that to seek optimum projection $\mathbf{P}$ should be called by different names (Okamoto, 2005, 2006). If the method to optimize orthogonally projected data $\mathbf{XP}$ is called principal component analysis, which describe information in the data by reduced dimensionality, the other one to minimize eq. (3) should be called factor analysis, by which we infer the latent space, In both cases, rotations play the role to find substantial directions in the mathematically determined space, where the initial axes are derived to find the space by criterions set by technical necessity.

References

Okamoto, Y. (2005). On the special type of principal component analysis: It should belong to factor analysis. *Japan Women's University Journal*, 16, 43-50.

Okamoto, Y. (2006). A justification of rotation in principal component analysis: Projective viewpoint of PCA. *Japan Women's University Journal*, 17, 59-71.