# Trend Vector Representation of Multiple Transition Frequency Tables by Regularized Homogeneity Analysis

Kohei Adachi
Koshien University, Hyogo 665-0006, Japan, e-mail: VYS03471@nifty.ne.jp

**Keywords**: Optimal scaling, Homogeneity analysis, Transition frequency tables

## Abstract

Individuals' choices of categories observed on two occasions are described by transition frequency tables. I present a regularized optimal scaling method to analyze a set of the tables obtained from multiple sources and represent a transition trend for each source as a vector. In the method, transition frequency tables for sources $k$ =1,…, $K$ are expressed as $\mathbf{F}_k = \mathbf{G}'_{k1}\mathbf{G}_{k2}$, where $\mathbf{G}_{kl}$ is an $n_k$-individuals by $m$-categories indicator matrix representing individuals' choices on occasion $l$ = 1, 2. I let $\mathbf{X}_{kl}$ ($n_k \times p$) contain $p$-vectors of individuals' scores on occasion $l$ and $\mathbf{Y}$ ($m \times p$) contain those of category scores which are invariant among sources and occasions. Further, I let $p$-vector $\mathbf{z}_k$ represent a trend for source $k$. With $\alpha > 0$ and $\mathbf{1}_k$ the $n_k$-vector of ones, a loss function is defined as

$$\phi(\mathbf{X}_k, \mathbf{Y}, \mathbf{z}_k) = \sum_{k=1}^{K} \left\{ \left\| \mathbf{X}_{k1} - \mathbf{G}_{k1}\mathbf{Y} \right\|^2 + \left\| \mathbf{X}_{k2} - \mathbf{G}_{k2}\mathbf{Y} \right\|^2 + \alpha \left\| (\mathbf{X}_{k2} - \mathbf{X}_{k1}) - \mathbf{1}_k \mathbf{z}'_k \right\|^2 \right\}. \quad (1)$$

Here, $\left\| \mathbf{X}_{kl} - \mathbf{G}_{kl}\mathbf{Y} \right\|^2$ expresses the loss of homogeneity between the scores of individuals and those of chosen categories, while penalty function $\left\| (\mathbf{X}_{k2} - \mathbf{X}_{k1}) - \mathbf{1}_k \mathbf{z}'_k \right\|^2$ represents the loss of homogeneity between trend vectors and the inter-occasion changes in individual scores. Under a normalization condition on $\mathbf{Y}$, (1) is minimized over $\mathbf{X}_{kl}$, $\mathbf{Y}$ and $\mathbf{z}_k$. The optimal $\mathbf{Y}$ and $\mathbf{z}_k$ are proved independent of the choice of $\alpha$. Graphical representation of the resulting $\mathbf{Y}$ and $\mathbf{z}_k$ allows us easily to grasp transition trends. Further, projecting the rows of $\mathbf{Y}$ onto $\mathbf{z}_k$ gives the unidimensional scales of categories useful for scrutinizing the trends. For illustration, the presented method is applied to mobility tables and confusion matrices. The relationships of the method to Adachi's (1997) approach and Zielman and Heiser's (1993) MDS model are considered.

## References

Adachi, K. (1997). Homogeneity analysis of transition matrices for spatially representing a transition trend. *Behaviormetrika*, **24**, 159-178.

Gifi, A. (1990). *Nonlinear multivariate analysis*. Chichester: Wiley.

Zielman, B., and Heiser, W. J. (1993). Analysis of asymmetry by a slide-vector. *Psychometrika*, **58**, 101-114.