Journal of Modern Applied Statistical Methods

Volume 15 | Issue 1 Article 29

5-1-2016

# Generalized Singular Value Decomposition with Additive Components

Stan Lipovetsky
*GfK*, stan.lipovetsky@gfk.com

Journal of Modern Applied Statistical Methods
May 2016, Vol. 15, No. 1, 584-599.

ISSN 1538 − 9472

# Generalized Singular Value Decomposition With Additive Components

**Stan Lipovetsky**
GfK North America
Minneapolis, MN

The singular value decomposition (SVD) technique is extended to incorporate the additive components for approximation of a rectangular matrix by the outer products of vectors. While dual vectors of the regular SVD can be expressed one via linear transformation of the other, the modified SVD corresponds to the general linear transformation with the additive part. The method obtained can be related to the family of principal component and correspondence analyses, and can be reduced to an eigenproblem of a specific transformation of a data matrix. This technique is applied to constructing dual eigenvectors for data visualizing in a two dimensional space.

*Keywords:* singular value decomposition, general linear transformation, principal components, dual eigenvectors, perceptual mapping

## Introduction

Data visualization, also known as perceptual mapping, is a representation of multivariate observations in two dimensional plots. It can be performed with various techniques, particularly principal component analysis (PCA) and factor analysis (Dillon and Goldstein, 1984; Seber, 1984; Falissard, 1999), multidimensional scaling (Chambers et al., 1983; Cleveland, 1985), and singular value decomposition (SVD) techniques (Golub and van Loan, 1983). These methods are applied for bi-plotting when both observations and variables are represented in the same graph (Gabriel 1971; Gabriel and Odoroff, 1990). Perceptual mapping includes methods of correspondence analysis or dual scaling for categorical data (Nishisato, 1980; Greenacre, 1984; Carroll et al., 1986; Benzecri, 1992), canonical correlation analysis and MANOVA (Novak, 1995; Tishler and Lipovetsky, 1996), and many-way eigenvector analysis (Cooper, 1988; Carlier and Kroonenberg, 1996; Lipovetsky, 1994). Various additional techniques are also used for specific problems in PCA, SVD, and perceptual mapping (Kaciak

*Dr. Lipovetsky is a Senior Research Director in the Marketing Sciences. Email him at: stan.lipovetsky@gfk.com.*

and Louviere, 1990; Sinha and DeSarbo, 1998; Steenkamp et al., 1994; Shugan, 1987; Lipovetsky and Conklin, 2001, 2003, 2004, 2005; Lipovetsky, 2009).

Projection of multivariate observation in a two dimensional space is usually performed by PCA or SVD techniques. In a PCA approach, the variables are centered by their means (and often normalized by their standard deviations), and in an SVD approach the original matrices of data are processed. Both PCA and SVD solve some eigenproblems and, from the several main eigenvectors (those correspond to maximum eigenvalues), a couple of vectors are picked for use as plot axes. In PCA, those are usually the 1st and 2nd or 2nd and 3rd main eigenvectors. In SVD, especially when it is used for correspondence analysis, the 1st eigenvector represents the mean values or their proportions among the variables, so the 2nd and the 3rd eigenvectors are applied.

A general linear transformation, including additive part, to construct the dual vectors for approximation of a data matrix. It can be called the Generalized SVD, or GSVD. This technique can be seen as another member of the family of SVD and PCA methods. GSVD corresponds to the general linear transformation of the axes of a plot, including shift of the beginning of coordinates and change of scales' units.

Despite problems in interpretation of the perceptual maps, the visualization of data is needed and widely used in applied statistical analysis. As it is noted in Jones and Pearce (2000), "In America visually-oriented, quantitatively illiterate culture, images have a great deal of power, so if a picture is today worth a thousand words, it must be worth at least a billion numbers". It could be particularly true for multivariate data and its statistical consideration (see also Lipovetsky and Mandel, 2009).

## SVD and PCA

Consider a matrix approximation by an outer product of two dual vectors. Let **x** denote a data matrix of $m$ by $n$ order, with elements $x_{ij}$ of $i$th observations $i=(1,\ldots,m)$ by $j$th variables $j=(1,\ldots,n)$. A model of matrix approximation by two vectors is

$$x_{ij} = \lambda b_i a_j + \varepsilon_{ij} \ , \tag{1}$$

where $b_i$ and $a_j$ are elements of vectors **b** and **a** of $m$th and $n$th order, respectively, $\lambda$ is a normalizing parameter, and $[\varepsilon_{ij}]$ is a matrix of residuals. This equation in matrix form is

$$\mathbf{x} = \lambda \mathbf{b}\mathbf{a}' + \boldsymbol{\varepsilon} \ , \tag{2}$$

with $\mathbf{b}$ as a column-vector and $\mathbf{a}'$ as a row-vector (prime denotes transposition). The least squares (LS) procedure for determining the vectors $\mathbf{a}$ and $\mathbf{b}$ by minimizing deviations $\boldsymbol{\varepsilon}$, also known as Eckart-Young approximation (Eckart and Young, 1936; Lipovetsky and Tishler, 1994), is defined by the objective:

$$S = \|\boldsymbol{\varepsilon}\|^2 = \|\mathbf{x} - \lambda \mathbf{b}\mathbf{a}\|^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \lambda b_i a_j\right)^2 \to \min \ . \tag{3}$$

From the first-order conditions

$$\partial S / \partial b_i = 0, \quad \partial \mathrm{S} / \partial a_j = 0 \ , \tag{4}$$

a system of equations yields

$$\begin{cases} \sum_{j=1}^{n} x_{ij} a_j = \lambda b_i \sum_{j=1}^{n} a_j^2 \quad , \quad i = 1, \ldots, m \, , \\ \sum_{i=1}^{m} x_{ij} b_i = \lambda a_j \sum_{i=1}^{m} b_i^2 \quad , \quad j = 1, \ldots, n \, . \end{cases} \tag{5}$$

Normalizing vectors by their scalar norm equal one,

$$\sum_{j=1}^{n} a_j^2 = 1, \quad \sum_{i=1}^{m} b_i^2 = 1 \ , \tag{6}$$

we represent (5) in a matrix form as follows:

$$\mathbf{x}\mathbf{a} = \lambda \mathbf{b}, \quad \mathbf{x}'\mathbf{b} = \lambda \mathbf{a} \ . \tag{7}$$

Substituting one of the equations in (7) into the other, we get two problems:

$$(\mathbf{x}'\mathbf{x})\mathbf{a} = \lambda^2 \mathbf{a} \ , \tag{8}$$

that is the eigenproblem for variables aggregate, and

$$(\mathbf{x}\mathbf{x}')\mathbf{b} = \lambda^2 \mathbf{b} \ , \tag{9}$$

that is the dual eigenproblem for the observations aggregate. First eigenvectors **a** and **b** for maximum eigenvalue $\lambda^2$ in (7)-(8) solve the problem (1)-(2) for approximation of a matrix by dual vectors' outer product.

If **x** is a matrix of the standardized (centered and normalized) variables, then **x'x** coincides with the correlation matrix, and the problem (8) coincides with PCA. Indeed, in PCA we construct a linear combination of variables

$$\mathbf{y} = \mathbf{xa} \ , \tag{10}$$

and maximize its variance

$$\text{var}(\mathbf{y}) = \mathbf{y'y} = \mathbf{ax'xa} \to \max \ . \tag{11}$$

With normalizing restriction (6) for vector **a**, we represent (11) as a conditional objective

$$\mathbf{a}(\mathbf{x'x})\mathbf{a} + \mu(\mathbf{a'a} - 1) \to \max \ , \tag{12}$$

with a Lagrange term $\mu$. Maximizing (11) due to (4), we obtain the same SVD eigenproblem (8) with $\mu = \lambda^2$. Comparison of (7) and (10) shows that $\mathbf{y} = \lambda\mathbf{b}$, so the aggregator $y$ of PCA is equal to the dual SVD vector **b** normalized by $\lambda^2$. Similarly, in a dual PCA approach, we can consider a linear combination of observations

$$\mathbf{z} = \mathbf{x'b} \ , \tag{13}$$

and maximize its variance; that yields the SVD eigenproblem (9) and the relation $\mathbf{z} = \lambda\mathbf{a}$ between PCA aggregator **z** and SVD dual vector **a**.

The SVD (7) and PCA relations (10) and (13) show that every element of an eigenvector is a linear combination of the elements of the dual eigenvector. In some analogies, it corresponds to a linear model without intercept. Let us consider now a model with intercept, meaning more general linear relations between dual eigenvectors.

## Generalized SVD

In place of the model of matrix approximation by outer product of two vectors (1), let us take a general model with linear and mixed effects:

$$x_{ij} = \mu + b_i + a_j + \lambda b_i a_j + \varepsilon_{ij} \ . \tag{14}$$

As it is known in the fitting of two-way tables in the analysis of variance (Tukey, 1977; Emerson & Wong, 1985), such a model can be re-arranged to the multiplicative fit model with the additive constant

$$x_{ij} = \nu + \lambda \beta_i \alpha_j + \varepsilon_{ij} \ , \tag{15}$$

where the relations between the parameters used in (14) and in (15) are as follows:

$$\nu = \mu - \lambda^{-1}, \quad \beta_i = b_i + \lambda^{-1}, \quad \alpha_j = a_j + \lambda^{-1} \ . \tag{16}$$

The model (15) is the extension of the regular SVD model (1) by the constant item $\nu$. If the values of parameters $\nu$, $\lambda$, $\boldsymbol{\beta}$, and $\boldsymbol{\alpha}$ in the model (15) are estimated, then by (16) we obtain all the parameters $\mu$, $\lambda$, $\mathbf{b}$, and $\mathbf{a}$ of the model (14) as well.

Take the model (15) and use it in the objective (3); that becomes

$$S = \|\boldsymbol{\varepsilon}\|^2 = \|\mathbf{x} - \nu - \lambda \boldsymbol{\beta}\boldsymbol{\alpha}'\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \left( x_{ij} - \nu - \lambda \beta_i \alpha_j \right)^2 \rightarrow \min \ . \tag{17}$$

From the condition $\partial S / \partial \nu = 0$, we get the expression for the constant

$$\nu = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} - \lambda \left( \frac{1}{m} \sum_{i=1}^{m} \beta_i \right) \left( \frac{1}{n} \sum_{j=1}^{n} \alpha_j \right) \equiv \bar{\mathbf{x}} - \lambda \bar{\boldsymbol{\beta}} \bar{\boldsymbol{\alpha}} \ , \tag{18}$$

so it equals the mean value by all the data in matrix $\mathbf{x}$ minus the interaction part of the mean values of the $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ dual vectors. Substituting (18) into (15), we represent this model as

$$\tilde{x}_{ij} = \lambda \beta_i \alpha_j - \lambda \bar{\boldsymbol{\beta}} \bar{\boldsymbol{\alpha}} + \varepsilon_{ij} \tag{19}$$

with the matrix of deviation from the total mean defined as

$$\tilde{x}_{ij} = x_{ij} - \bar{\mathbf{x}} \ . \tag{20}$$

To reduce the model (19) to the regular SVD model (1), we impose the constraint that both mean values of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ equal zero:

$$\sum_{j=1}^{n} \alpha_j = 0, \quad \sum_{i=1}^{m} \beta_i = 0 \ . \tag{21}$$

Such restrictions are used in the analysis of variance with non-additive fits (Mandel, 1969, 1971) and in factor analysis (Gollob, 1968).

With help of the relations (18)-(21), we represent (17) as a conditional objective

$$\begin{aligned} S &= \|\boldsymbol{\varepsilon}\|^2 = \|\tilde{\mathbf{x}} - \lambda \boldsymbol{\beta}\boldsymbol{\alpha}'\|^2 + 2\rho\left(\boldsymbol{\beta}'\mathbf{e}_m\right) + 2\tau(\boldsymbol{\alpha}'\mathbf{e}_n) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} (x_{ij} - \lambda\beta_i\alpha_j)2 + 2\rho\sum_{i=1}^{m}\beta_i + 2\tau\sum_{j=1}^{n}\alpha_j \to \min \end{aligned} \tag{22}$$

where $\rho$ and $\tau$ are Lagrange multipliers for incorporating conditions (21) into the objective (17), and $\mathbf{e}_m$ and $\mathbf{e}_n$ denote the uniform vector-columns of order $m$ and $n$, respectively. From the first order conditions (4) applied to the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ in (22), a system of equations yields:

$$\begin{cases} \sum_{j=1}^{n} \tilde{x}_{ij}\alpha_j = \lambda\beta_i \sum_{j=1}^{n} \alpha_j^2 + \rho/\lambda \quad , \quad i = 1, \ldots, m \,, \\ \sum_{i=1}^{m} \tilde{x}_{ij}\beta_j = \lambda\alpha_j \sum_{i=1}^{m} \beta_i^2 + \tau/\lambda \quad , \quad j = 1, \ldots, \mathrm{n} \,. \end{cases} \tag{23}$$

Comparison with (5) shows that equation (23) corresponds to the general linear transformation of one vector to another with additive constants.

Multiplying the first group of equations (23) by $\beta_i$ and summing over the index $i$, or the second group of equations (23) by $\alpha_j$ and summing over $j$, and taking into account conditions (21), we get expression for the term $\lambda$:

$$\lambda = \frac{\sum_i^m \sum_j^n \tilde{x}_{ij}\beta_i\alpha_j}{\sum_i^m \beta_i^2 \sum_j^n \alpha_j^2} \ . \tag{24}$$

Summing the first group of equations (23) by index $i$, the second group of equations (23) by index $j$, and using conditions (21), we obtain expressions for the terms $\rho$ and $\tau$:

$$\rho = \lambda \sum_{j=1}^{n} \left( \frac{1}{m} \sum_{i=1}^{m} \tilde{x}_{ij} \right) \alpha_j \quad , \quad \tau = \lambda \sum_{i=1}^{m} \left( \frac{1}{n} \sum_{j=1}^{n} \tilde{x}_{ij} \right) \beta_i \ . \tag{25}$$

Substituting (25) into (23) yields the following system of equations:

$$\begin{cases} \sum_{j=1}^{n} \left( \tilde{x}_{ij} - \frac{1}{m} \sum_{i=1}^{m} \tilde{x}_{ij} \right) \alpha_j = \lambda \beta_i \sum_{j=1}^{n} \alpha_j^2 \quad , \quad i = 1, \ldots, m \, , \\ \sum_{i=1}^{m} \left( \tilde{x}_{ij} - \frac{1}{n} \sum_{j=1}^{n} \tilde{x}_{ij} \right) \beta_j = \lambda \alpha_j \sum_{i=1}^{m} \beta_i^2 \quad , \quad j = 1, \ldots, \mathrm{n} \, . \end{cases} \tag{26}$$

Using definition (20) we can re-write expressions in the parentheses in (26) via the original data matrix:

$$\begin{cases} \tilde{x}_{ij} - \frac{1}{m} \sum_{i=1}^{m} \tilde{x}_{ij} = x_{ij} - \bar{\mathbf{x}} - \frac{1}{m} \sum_{i=1}^{m} \left( x_{ij} - \bar{\mathbf{x}} \right) = x_{ij} - c_j \, , \\ \tilde{x}_{ij} - \frac{1}{n} \sum_{j=1}^{n} \tilde{x}_{ij} = x_{ij} - \bar{\mathbf{x}} - \frac{1}{n} \sum_{j=1}^{n} \left( x_{ij} - \bar{\mathbf{x}} \right) = x_{ij} - r_i \, , \end{cases} \tag{27}$$

where $c_j$ and $r_i$ denote, respectively, the column mean (average by rows) and row mean (average by columns) values of the matrix $[x_{ij}]$:

$$c_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij} \quad , \quad r_i = \frac{1}{n} \sum_{j=1}^{n} x_{ij} \ . \tag{28}$$

In a matrix $\mathbf{x}$, rows often correspond to observations and columns to attributes. Then in (27), the first matrix $[x_{ij} - c_j]$ consists of observations centered and the second matrix $[x_{ij} - r_i]$ consists of attributes centered. Denote these two matrices as $\mathbf{Y}$ and $\mathbf{Z}$, so in matrix form the expressions in (27) are

$$\mathbf{Y} = \mathbf{x} - \mathbf{e}_m \mathbf{c'} \quad , \quad \mathbf{Z} = \mathbf{x} - \mathbf{re'}_n \ , \tag{29}$$

where uniform vectors $\mathbf{e}$ are as in (22), $\mathbf{c}$ and $\mathbf{r}$ are vector-columns of the order $n$ and $m$, respectively, with the elements of (28), and the prime denoting transposition.

Using (27)-(29) we represent the system (26) in matrix notation:

$$\mathbf{Y\alpha} = \lambda(\boldsymbol{\alpha}'\boldsymbol{\alpha})\boldsymbol{\beta} \quad , \quad \mathbf{Z}'\boldsymbol{\beta} = \lambda(\boldsymbol{\beta}'\boldsymbol{\beta})\boldsymbol{\alpha} \ . \tag{30}$$

Normalizing vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ by scalar norm (6), we simplify (30) to the system

$$\mathbf{Y}\boldsymbol{\alpha} = \lambda\boldsymbol{\beta} \quad , \quad \mathbf{Z}'\boldsymbol{\beta} = \lambda\boldsymbol{\alpha} \ . \tag{31}$$

Substituting one of the equations in (31) into the other yields the problems

$$\left(\mathbf{Z}'\mathbf{Y}\right)\boldsymbol{\alpha} = \lambda^2\boldsymbol{\alpha} \quad , \quad \left(\mathbf{Y}\mathbf{Z}'\right)\boldsymbol{\beta} = \lambda^2\boldsymbol{\beta} \ . \tag{32}$$

Relations (31) and (32) define the generalized SVD method. The eigenproblems (32) yield the attributes' and observations' aggregates corresponded to generalizations of the regular SVD solutions (8)-(9). First eigenvectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ for maximum eigenvalue $\lambda^2$ in (32) solve the problem (17) or (22) for approximation of a matrix by outer product of vectors with additive component.

The GSVD problems (32) can be represented, using (29), more explicitly:

$$\begin{aligned} &\left(\left(\mathbf{x}' - \mathbf{e}_n\mathbf{r}'\right)\left(\mathbf{x} - \mathbf{e}_m\mathbf{c}'\right)\right)\boldsymbol{\alpha} = \lambda^2\boldsymbol{\alpha} \, , \\ &\left(\left(\mathbf{x} - \mathbf{e}_m\mathbf{c}'\right)\left(\mathbf{x}' - \mathbf{e}_n\mathbf{r}'\right)\right)\boldsymbol{\beta} = \lambda^2\boldsymbol{\beta} \, . \end{aligned} \, , \tag{33}$$

It then becomes clear that problems in (33) belong to the family of SVD and PCA techniques while differing from both of them. The regular SVD (7)-(9) operates with a matrix $\mathbf{x}$ and the same matrix transposed $\mathbf{x}'$, while in the GSVD (31)-(33) there actually are two matrices corresponded to data centering in different directions (29). Suppose matrix $\mathbf{x}$ consists of centered and normalized variables and $\mathbf{x}'\mathbf{x}$ coincides with the correlation matrix. In this case, due to (29), $\mathbf{Y} = \mathbf{x}$ but $\mathbf{Z} \neq \mathbf{Y}$, and the problems (32)-(33) cannot be reduced to PCA.

For numerical evaluations it could be convenient to express vector-columns $\mathbf{c}$ and $\mathbf{r}$ (28) in matrix notation:

$$\mathbf{c} = \frac{1}{m}\mathbf{x}'\mathbf{e}_m \quad , \quad \mathbf{r} = \frac{1}{n}\mathbf{x}\mathbf{e}_n \ . \tag{34}$$

Using (34), we represent the problems in (33) directly via the data matrix:

$$\begin{aligned} &\left(\left(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}'_n\right)\mathbf{x}'\left(\mathbf{I}_m - \frac{1}{m}\mathbf{e}_m\mathbf{e}'_m\right)\mathbf{x}\right)\boldsymbol{\alpha} = \lambda^2\boldsymbol{\alpha}, \\ &\left(\left(\mathbf{I}_m - \frac{1}{m}\mathbf{e}_m\mathbf{e}'_m\right)\mathbf{x}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{e}_n\mathbf{e}'_n\right)\mathbf{x}'\right)\boldsymbol{\beta} = \lambda^2\boldsymbol{\beta}, \end{aligned} \tag{35}$$

where $\mathbf{I}_n$ and $\mathbf{I}_m$ are identity matrices of the order $n$ and $m$, respectively. We see that numerical estimations for GSVD problems (35) are not much more difficult than for the regular SVD (8)-(9). It is sufficient to solve just one eigenproblem of the smaller dimension – usually the number of variables $n$ is less than the number of observations $m$, so we can solve the first problem of (35). The dual solution can then be found by the linear transformation (31).

Although an eigenvector in a separated eigenproblem is defined up to an arbitrary sign, the vectors of dual solutions have directions defined by their dual transformation (formula (7) for SVD, (10) and (13) for PCA, or (31) for GSVD). It means that sign could be changed to the opposite simultaneously in each pair of dual vectors.

## Numerical Examples

Consider a matrix of proportions from an actual marketing research project for six brands, *X1*, *X2*,…, *X6* and nine attributes, *A*, *B*, …, *I*, presented in Table 1.

**Table 1.** Example I: Matrix of proportions.

| | ***X1*** | ***X2*** | ***X3*** | ***X4*** | ***X5*** | ***X6*** |
|---|---|---|---|---|---|---|
| ***A*** | 0.249 | 0.166 | 0.317 | 0.119 | 0.059 | 0.077 |
| ***B*** | 0.266 | 0.283 | 0.345 | 0.293 | 0.154 | 0.228 |
| ***C*** | 0.217 | 0.203 | 0.279 | 0.249 | 0.148 | 0.214 |
| ***D*** | 0.127 | 0.051 | 0.129 | 0.231 | 0.054 | 0.206 |
| ***E*** | 0.211 | 0.222 | 0.335 | 0.211 | 0.153 | 0.235 |
| ***F*** | 0.278 | 0.246 | 0.31 | 0.285 | 0.181 | 0.26 |
| ***G*** | 0.266 | 0.289 | 0.438 | 0.378 | 0.257 | 0.426 |
| ***H*** | 0.259 | 0.263 | 0.349 | 0.257 | 0.164 | 0.265 |
| ***I*** | 0.246 | 0.302 | 0.444 | 0.167 | 0.222 | 0.232 |

**Table 2.** Example I: Pairs of dual vectors for SVD, PCA, and GSVD.

| | SVD-vectors | | PCA-vectors | | GSVD-vectors | |
|---|---|---|---|---|---|---|
| | ***a2*** | ***a3*** | ***a2*** | ***a3*** | ***$\alpha$1*** | ***$\alpha$2*** |
| *X1* | 0.22 | -0.61 | 0.22 | -0.59 | 0.21 | -0.70 |
| *X2* | 0.30 | 0.16 | 0.31 | -0.30 | 0.34 | -0.01 |
| *X3* | 0.47 | 0.10 | 0.48 | 0.26 | 0.51 | 0.36 |
| *X4* | -0.58 | -0.42 | -0.58 | -0.53 | -0.57 | -0.34 |
| *X5* | -0.02 | 0.53 | -0.01 | 0.27 | 0.01 | 0.22 |
| *X6* | -0.55 | 0.35 | -0.54 | 0.38 | -0.50 | 0.47 |
| | ***b2*** | ***b3*** | ***b2*** | ***b3*** | ***$\beta$1*** | ***$\beta$2*** |
| *A* | 0.54 | -0.56 | 0.53 | -0.14 | 0.48 | -0.46 |
| *B* | 0.04 | -0.27 | 0.04 | -0.56 | 0.04 | -0.25 |
| *C* | -0.09 | -0.13 | -0.09 | -0.08 | -0.10 | -0.13 |
| *D* | -0.54 | -0.35 | -0.56 | 0.28 | -0.61 | -0.19 |
| *E* | 0.07 | 0.12 | 0.07 | 0.30 | 0.07 | 0.20 |
| *F* | -0.11 | -0.19 | -0.11 | -0.39 | -0.10 | -0.24 |
| *G* | -0.39 | 0.40 | -0.38 | 0.19 | -0.32 | 0.57 |
| *H* | 0.01 | -0.04 | 0.02 | -0.13 | 0.03 | 0.01 |
| *I* | 0.48 | 0.51 | 0.49 | 0.53 | 0.51 | 0.49 |

The results obtained in SVD, PCA, and GSVD solutions are presented in Table 2, where all of the vectors are normalized by the totals of the squared elements equal one. For the SVD (7)-(9) and PCA (10)-(13) solutions, we take the second (**a2**, **b2**) and the third (**a3**, **b3**) pairs of dual vectors (because the first dimension of means is related to intensity rather than to structure). For the generalized solution (35) we take its first (**α1**, **β1**) and second (**α2**, **β2**) pairs of dual vectors (because due to the construction of this method we have already reduced the data to just the structure among the items in both directions).

Comparison among the vectors **a2** for SVD, **a2** for PCA, and **α1** for GSVD shows that they are very close to one another. Their dual vectors are also similar among themselves, although **β1** for GSVD solutions differs from the two other very close solutions **b2** for SVD and **b2** for PCA. The next solutions, **a3** for SVD, **a3** for PCA, and **β2** for GSVD, are already not so similar among themselves; they can differ by value and by sign, too. The same concerns their dual vectors **b3** for SVD,

**b3** for PCA, and **β2** for GSVD. However, the biggest coefficients are always of the same sign within each group of the vectors compared.

Another example presents real data for ten brands and fifteen attributes (those are: good value, natural, variety of flavors, I would pay more, good to serve to guests, reenergizes, innovative, brand I trust, cares for environment, socially responsible, served in selected restaurants, authentic, symbol of conviviality, high quality, my brand). The proportions by choices are given in Table 3.

The results of SVD, PCA, and GSVD solutions are presented in Table 4 and are constructed similarly to those described above in Table 2, with the second (**a2**, **b2**) and third (**a3**, **b3**) pairs of dual vectors for SVD and PCA and the first (**α1**, **β1**) and second (**α2**, **β2**) pairs of dual vectors for GSVD. Again, the vectors **a2** for SVD and PCA and **α1** for GSVD are very similar. The corresponding dual vectors are less similar among themselves. The next solutions, **a3** for SVD and PCA and **α2** for GSVD, differ more; the same concerns their dual vectors, **b3** for SVD and PCA and **β2** for GSVD. The reason for the differences is easy to explain by recalling that the total of the elements in any SVD vector does not equal zero while **b**-vectors in PCA have zero totals because of the centering of data in columns. In

**Table 3.** Example II: Matrix of proportions.

| | *X1* | *X2* | *X3* | *X4* | *X5* | *X6* | *X7* | *X8* | *X9* | *X10* |
|---|---|---|---|---|---|---|---|---|---|---|
| *A* | 0.166 | 0.194 | 0.198 | 0.085 | 0.108 | 0.036 | 0.099 | 0.086 | 0.119 | 0.093 |
| *B* | 0.250 | 0.282 | 0.259 | 0.071 | 0.085 | 0.068 | 0.102 | 0.085 | 0.083 | 0.068 |
| *C* | 0.143 | 0.168 | 0.281 | 0.078 | 0.131 | 0.040 | 0.072 | 0.108 | 0.077 | 0.075 |
| *D* | 0.196 | 0.214 | 0.241 | 0.052 | 0.063 | 0.037 | 0.049 | 0.044 | 0.045 | 0.044 |
| *E* | 0.281 | 0.332 | 0.319 | 0.082 | 0.110 | 0.055 | 0.097 | 0.099 | 0.078 | 0.063 |
| *F* | 0.213 | 0.260 | 0.265 | 0.090 | 0.110 | 0.056 | 0.086 | 0.088 | 0.081 | 0.065 |
| *G* | 0.164 | 0.213 | 0.239 | 0.069 | 0.094 | 0.058 | 0.067 | 0.086 | 0.047 | 0.051 |
| *H* | 0.285 | 0.335 | 0.326 | 0.094 | 0.123 | 0.055 | 0.111 | 0.112 | 0.095 | 0.078 |
| *I* | 0.145 | 0.191 | 0.209 | 0.058 | 0.071 | 0.048 | 0.081 | 0.079 | 0.060 | 0.048 |
| *J* | 0.166 | 0.222 | 0.218 | 0.056 | 0.077 | 0.049 | 0.073 | 0.072 | 0.059 | 0.045 |
| *K* | 0.256 | 0.304 | 0.260 | 0.049 | 0.046 | 0.033 | 0.056 | 0.058 | 0.040 | 0.037 |
| *L* | 0.270 | 0.166 | 0.163 | 0.045 | 0.047 | 0.035 | 0.033 | 0.032 | 0.026 | 0.029 |
| *M* | 0.171 | 0.187 | 0.193 | 0.055 | 0.063 | 0.043 | 0.050 | 0.054 | 0.043 | 0.041 |
| *N* | 0.282 | 0.349 | 0.332 | 0.091 | 0.107 | 0.057 | 0.088 | 0.100 | 0.080 | 0.067 |
| *O* | 0.170 | 0.213 | 0.205 | 0.058 | 0.083 | 0.034 | 0.061 | 0.074 | 0.057 | 0.058 |

GSVD, the solution for each vector equals zero due to the data centering in both directions. Which of the solutions is more interpretable and preferable depends on a nature of data in the project. If a decision maker is not satisfied with the results of classical SVD or PCA, it is possible to try the new approach of the GSVD.

**Table 4.** Example II: Pairs of dual vectors for SVD, PCA, and GSVD

| | SVD-vectors | | PCA-vectors | | GSVD-vectors | |
|---|---|---|---|---|---|---|
| | ***a2*** | ***a3*** | ***a2*** | ***a3*** | ***a1*** | ***a2*** |
| *X1* | -0.594 | 0.590 | -0.572 | 0.523 | -0.518 | 0.667 |
| *X2* | -0.240 | -0.420 | -0.157 | -0.217 | -0.607 | -0.252 |
| *X3* | 0.213 | -0.414 | 0.254 | -0.514 | -0.274 | -0.603 |
| *X4* | 0.196 | 0.174 | 0.204 | 0.175 | 0.170 | 0.085 |
| *X5* | 0.412 | 0.191 | 0.420 | 0.153 | 0.241 | -0.158 |
| *X6* | 0.051 | 0.022 | 0.050 | 0.004 | 0.166 | 0.251 |
| *X7* | 0.238 | 0.170 | 0.256 | 0.268 | 0.152 | 0.020 |
| *X8* | 0.326 | -0.015 | 0.342 | -0.001 | 0.188 | -0.147 |
| *X9* | 0.323 | 0.368 | 0.336 | 0.459 | 0.233 | 0.035 |
| *X10* | 0.259 | 0.269 | 0.260 | 0.281 | 0.248 | 0.102 |
| | ***b2*** | ***b3*** | ***b2*** | ***b3*** | ***b1*** | ***b2*** |
| *A* | 0.380 | 0.556 | 0.350 | 0.611 | 0.384 | 0.086 |
| *B* | -0.062 | 0.167 | -0.041 | 0.246 | -0.129 | 0.115 |
| *C* | 0.525 | -0.044 | 0.498 | -0.204 | 0.366 | -0.408 |
| *D* | -0.136 | -0.164 | -0.167 | -0.264 | -0.011 | 0.026 |
| *E* | -0.090 | -0.082 | -0.030 | -0.006 | -0.355 | -0.148 |
| *F* | 0.129 | 0.059 | 0.141 | 0.082 | 0.009 | -0.088 |
| *G* | 0.152 | -0.197 | 0.127 | -0.265 | 0.162 | -0.139 |
| *H* | 0.028 | 0.066 | 0.098 | 0.173 | -0.310 | -0.172 |
| *I* | 0.168 | -0.094 | 0.125 | -0.125 | 0.263 | -0.053 |
| *J* | 0.058 | -0.151 | 0.031 | -0.155 | 0.128 | -0.047 |
| *K* | -0.434 | -0.338 | -0.418 | -0.305 | -0.386 | 0.055 |
| *L* | -0.521 | 0.630 | -0.589 | 0.442 | -0.035 | 0.807 |
| *M* | -0.064 | -0.003 | -0.118 | -0.090 | 0.165 | 0.164 |
| *N* | -0.099 | -0.195 | -0.028 | -0.104 | -0.407 | -0.218 |
| *O* | 0.056 | -0.015 | 0.022 | -0.037 | 0.157 | 0.018 |

## Summary

A modified singular value decomposition technique is described as a general linear model with an additive part in the mutual representation of the dual vectors one via another. The dual transformation with additional parameters is more flexible to adjusting one direction of the data matrix to another, so the generalized SVD can better express the structure of the simultaneous positioning of a matrix's two directions in the two-dimensional plot. The GSVD technique is reduced to an eigenproblem of a specific structure of the product of two matrices, each centered in one of the directions of the data (35). This eigenproblem does not produce the eigenvectors of intensity that correspond to the first dual vectors in singular value decomposition or in principal component analysis. The GSVD eigenvectors, beginning from the first one, describe the structure of the relations among the directions in the data matrix. The generalized SVD yields a possibility to obtain an additional structure of items' locations on the perceptual map. Comparing maps of different techniques, a researcher can find the more interpretable perceptual map. The GSVD technique enriches the family of SVD and PCA methods for obtaining dual vectors, and can be useful for theoretical investigation and practical applications in numerous problems that use eigenvector or SVD analysis.

## References

Benzecri, J. P. (1992). Correspondence analysis handbook. New York: Marcel Dekker Inc.

Carlier, A. & Kroonenberg, P. M. (1996). Decompositions and biplots in three-way correspondence analysis. *Psychometrica, 61*(2), 355-373. doi: 10.1007/BF02294344

Carroll, J. D., Green, P. E., & Schaffer C. M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research*, *23*(3), 271-280. doi: 10.2307/3151485

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). Graphical methods for data analysis. Belmont, CA: Wadsworth.

Cleveland, W. S. (1985). The elements of graphing data. Monterey, CA: Wadsworth.

Cooper, L. G. (1988) Competitive maps: The structure underlying asymmetric cross elasticities. *Management Science, 34*(6), 707-723. doi: 10.1287/mnsc.34.6.707

Dillon, W. R. & Goldstein, M. (1984). Multivariate analysis, Methods and applications. New York: Wiley.

Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psycometrika*, *1*(2), 211-218. doi: 10.1007/BF02288367

Emerson, J. D. & Wong G. Y. (1985). Resistant nonadditive fits for two-way tables. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Exploring data tables, trends, and shapes* (67-124). New York: John Wiley and Sons.

Falissard, B. (1999). Focused principal component analysis: Looking at a correlation matrix with a particular interest in a given variable. *Journal of Computational and Graphical Statistics*, *8*(4), 906-912. doi: 10.1080/10618600.1999.10474855

Gabriel, K. R. (1971). The biplot graphical display of matrices with applications to principal component analysis. *Biometrika*, *58*(3), 453-467. doi: 10.1093/biomet/58.3.453

Gabriel, K. R. & Odoroff, C. L. (1990). Biplots in biomedical research. *Statistics in Medicine, 9*(5), 469-485. doi: 10.1002/sim.4780090502

Gollob, H. F. (1968). A statistical model which combines features of factor analysis and analysis of variance techniques. *Psychometrika*, *33*(1), 73-115. doi: 10.1007/BF02289676

Golub, G. H. & van Loan, C. F. (1983). Matrix computations. Baltimore: Johns Hopkins University Press.

Greenacre, M. J. (1984). Theory and application of correspondence analysis. London: Academic Press.

Jones, R. & Pearce, J. (2000). A postmodern view of fractions and the reciprocals of Fermat primes. *Mathematics Magazine*, *73*(2), 83-97. doi: 10.2307/2691078

Kaciak, E., & Louviere J. (1990). Multiple correspondence analysis of multiple choice experiment data. *Journal of Marketing Research*, *27*(4), 455-465. doi: 10.2307/3172630

Lipovetsky, S. (1994). Multimode data analysis for decision making. In G. H. Tzeng et al. (Eds.), *Multiple criteria decision making* (275-282). Springer-Verlag.

Lipovetsky, S. (2009). PCA and SVD with nonnegative loadings. *Pattern Recognition*, *42*(1), 68-76. doi: 10.1016/j.patcog.2008.06.025

Lipovetsky, S. & Conklin, M. (2001). Dual priority-antipriority Thurstone scales as AHP eigenvectors. *Engineering Simulation*, *18*(2001), 631-648.

Lipovetsky, S. & Conklin, M. (2003) Dual- and triple-mode matrix approximation and regression modeling. *Applied Stochastic Models in Business and Industry, 19*(4), 291-301. doi: 10.1002/asmb.503

Lipovetsky, S. & Conklin, M. (2004). Nonlinear Thurstone scaling via SVD and Gower plots. *International Journal of Operations and Quantitative Management, 10*(4), 259-273.

Lipovetsky, S. & Conklin, M. (2005). Singular value decomposition in additive, multiplicative, and logistic forms. *Pattern Recognition, 38*(7), 1099-1110. doi: 10.1016/j.patcog.2005.01.010

Lipovetsky, S. & Mandel, I. (2009). How art helps to understand statistics. *Model Assisted Statistics and Applications, 4*(4), 313-324. doi: 10.3233/MAS-2009-0135

Lipovetsky, S. & Tishler, A. (1994). Linear methods in multimode data analysis for decision making. *Computers and Operations Research, 21*(2), 169-183. doi: 10.1016/0305-0548(94)90050-7

Mandel J. (1969). A method for fitting empirical surfaces to physical or chemical data. *Technometrics, 11*(3), 411-429. doi: 10.1080/00401706.1969.10490704

Mandel J. (1971). A new analysis of variance model for non-additive data. *Technometrics*, *13*(1), 1-18. doi: 10.1080/00401706.1971.10488751

Nishisato, S. (1980). Analysis of categorical data: Dual scaling and its applications. Toronto: Toronto University Press.

Novak, T.P. (1995). MANOVAMAP: Graphical representation of MANOVA in marketing research. *Journal of Marketing Research, 32*(3), 357-374. doi: 10.2307/3151987

Seber, G. A. F. (1984). Multivariate observations. New York: Wiley.

Shugan, S.M. (1987). Estimating brand positioning maps using supermarket scanning data. *Journal of Marketing Research, 24*(1), 1-18. doi: 10.2307/3151749

Sinha, I., & DeSarbo W. S. (1998). An integrated approach toward the spatial modeling of perceived customer value. *Journal of Marketing Research, 35*(2), 236-249. doi: 10.2307/3151851

Steenkamp, J. B., van Trijp, H., & Ten Berge, J. (1994). Perceptual mapping based on idiosyncratic sets. *Journal of Marketing Research, 31*(1), 15-27. doi: 10.2307/3151943

Tishler, A. & Lipovetsky, S. (1996). Canonical correlation analysis for three data sets: A unified framework with applications to management. *Computers and Operations Research*, *23*(7), 667-679. doi: 10.1016/0305-0548(95)00067-4

Tukey, J. W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley.