PSYCHOMETRIKA—VOL. 67, NO. 2, 211–224
JUNE 2002

GENERALIZED CONSTRAINED MULTIPLE CORRESPONDENCE ANALYSIS

HEUNGSUN HWANG AND YOSHIO TAKANE

MCGILL UNIVERSITY

A comprehensive approach for imposing both row and column constraints on multivariate discrete data is proposed that may be called generalized constrained multiple correspondence analysis (GCMCA). In this method each set of discrete data is first decomposed into several submatrices according to its row and column constraints, and then multiple correspondence analysis (MCA) is applied to the decomposed submatrices to explore relationships among them. This method subsumes existing constrained and unconstrained MCA methods as special cases and also generalizes various kinds of linearly constrained correspondence analysis methods. An example is given to illustrate the proposed method.

Key words: multiple correspondence analysis, linear constraints, projection operators, generalized singular value decomposition.

1. Introduction

Multiple correspondence analysis (MCA) is a useful technique to examine relationships among more than two sets of discrete variables (Benzécri, 1973; Greenacre, 1984, Lebart, Morineau, & Warwick, 1984, Nishisato, 1980). MCA is primarily a descriptive method designed to assign scores to rows (representing the subjects) and the columns (representing the response categories of the discrete variables), yielding a graphical representation of the rows or the columns of the dummy-coded discrete variables. This graphical display may facilitate the understanding of the interdependency among the data sets.

In practice, each set of discrete variables is often accompanied by some additional information about the subjects and/or categories. For example, subjects' demographic information (e.g., age, gender, level of education, etc.) and their group membership may be available as auxiliary information. Any relationships among variables (e.g., no interaction between variables, equality among variables, group membership of variables, etc.) may also be a priori known. Such additional information can be incorporated in the form of linear constraints (Böckenholt & Böckenholt, 1990; Nishisato, 1984; Takane & Shibayama, 1991; Takane, Yanai, & Mayekawa, 1991; ter Braak, 1988; van Buuren & de Leeuw, 1992; Yanai, 1986; Yanai, 1998; Yanai & Maeda, 2000). By imposing row and/or column constraints we may obtain simpler interpretations since the data to be analyzed are already structured by the constraints (Böckenholt & Böckenholt, 1990). We may also explore relationships among multivariate discrete data sets from diverse perspectives, analyzing a variety of submatrices of each discrete data matrix prescribed by the constraints. In addition, if the constraints are consistent with the data, more reliable estimates of parameters are obtained.

In this paper, we propose a comprehensive approach for incorporating both row and column constraints into the discrete data sets under a single algebraic framework. This method may be called generalized constrained multiple correspondence analysis (GCMCA). GCMCA includes existing constrained MCA (e.g., Gifi, 1990; Nishisato, 1984; van Buuren & de Leeuw, 1992; Yanai, 1998; Yanai & Maeda, 2000) as special cases. Takane and Shibayama (1991) provided a comprehensive framework for incorporating linear constraints on a data matrix in the context of principal components analysis (PCA). We follow a similar approach here: In GCMCA, each data

Heungsun Hwang is now at Claes Fornell International Group. The work reported in this paper was supported by Grant A6394 from the Natural Sciences and Engineering Research Council of Canada to the second author.

Requests for reprints should be addressed to: Heungsun Hwang, Claes Fornell International Group, 625 Avis Drive, Ann Arbor, MI 48108. E-Mail: hhwang@mail.cfigroup.com.

0033-3123/2002-2/2000-0800-A $00.75/0

© 2002 The Psychometric Society

set is first decomposed into several submatrices according to its row and/or column information, and then MCA is applied to a set of the decomposed submatrices, selected from each data set, in order to explore associations among them. Technically, the former amounts to projection of the data matrix onto the space spanned by a matrix of linear constraints, while the latter involves the generalized singular value decomposition (GSVD) of a certain matrix.

The present paper is organized as follows: Section 2 describes the proposed method in detail. Decompositions of data matrices according to row and/or column constraints are discussed. Various types of constrained MCA methods are then presented, depending on which decomposed submatrix is analyzed. This section shows that our method consists of existing column-wise or row-wise constrained MCA methods as special cases. It turns out that our method includes multivariate extensions of various kinds of linearly constrained correspondence analysis methods. In section 3, an example is given to illustrate the proposed method. The final section briefly summarizes the previous sections and discusses further prospects of the proposed method.

2. The Method

2.1. *Decompositions of Data Matrices by Linear Constraints*

Let $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_K]$ denote an n by r superindicator matrix consisting of K indicator matrices, where $\mathbf{Z}_i$ is an n by r_i data matrix of dummy variables and $r = \sum_i r_i$. Usually matrix $\mathbf{Z}_i$ represents a set of subjects' responses to item i, whose rows correspond with subjects and columns with response categories in item i. Let $\mathbf{G} = [\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_K]$ denote an n by b supermatrix consisting of K row constraint matrices, where $\mathbf{G}_i$ is an $n \times b_i (\leq n)$ row constraint matrix corresponding to $\mathbf{Z}_i$ and $b = \sum_i b_i$. Matrix $\mathbf{G}_i$ may take a variety of forms. For example, it may be an n-component vector of ones or a matrix of dummy variables indicating subjects' group membership, or demographic information. Matrix $\mathbf{G}_i$ may also be a block diagonal matrix to reflect any groupings of rows of $\mathbf{Z}_i$. Let $\mathbf{H} = [\mathbf{H}_1 \oplus \mathbf{H}_2 \oplus \cdots \oplus \mathbf{H}_K]$ denote an r by c supermatrix consisting of K column constraint matrices ($\oplus$ denotes the direct sum), where $\mathbf{H}_i$ is an $r_i \times c_i (\leq r_i)$ column constraint matrix for $\mathbf{Z}_i$, and $c = \sum_i c_i$. Matrix $\mathbf{H}_i$ may be an r_i-component vector of ones or any matrix capturing relationships among columns of $\mathbf{Z}_i$ (e.g., a design matrix for pair comparisons, an additivity constraint matrix, an equality constraint matrix, etc.). Matrices $\mathbf{Z}$ and $\mathbf{G}$ are assumed to be columnwise-centered. (Notice that if $\mathbf{Z}$ is centered, $\mathbf{H}$ does not need to be so.) When no row and/or column constraint matrix is available for $\mathbf{Z}_i$, we may simply set $\mathbf{G}_i$ and $\mathbf{H}_i$ equal to the identity matrices of appropriate size.

Following Takane and Shibayama (1991), each data matrix, $\mathbf{Z}_i$, can be decomposed into four submatrices according to $\mathbf{G}_i$ and $\mathbf{H}_i$ as follows:

$$\begin{aligned}\mathbf{Z}_i &= (\mathbf{\Phi}_i + \mathbf{\Sigma}_i)\mathbf{Z}_i(\mathbf{\Psi}_i + \mathbf{\Omega}_i) \\ &= \mathbf{\Phi}_i\mathbf{Z}_i\mathbf{\Psi}_i + \mathbf{\Phi}_i\mathbf{Z}_i\mathbf{\Omega}_i + \mathbf{\Sigma}_i\mathbf{Z}_i\mathbf{\Psi}_i + \mathbf{\Sigma}_i\mathbf{Z}_i\mathbf{\Omega}_i, \end{aligned} \quad (1)$$

where

$$\begin{aligned}\mathbf{D}_i &= \mathbf{Z}_i'\mathbf{Z}_i, \\ \mathbf{\Phi}_i &= \mathbf{G}_i(\mathbf{G}_i'\mathbf{G}_i)^-\mathbf{G}_i', \\ \mathbf{\Sigma}_i &= \mathbf{I} - \mathbf{\Phi}_i, \\ \mathbf{\Psi}_i &= \mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{D}_i, \\ \mathbf{\Omega}_i &= \mathbf{I} - \mathbf{\Psi}_i.\end{aligned}$$

Here, $\mathbf{I}$ is the identity matrix. Matrix $\mathbf{\Phi}_i$ is the orthogonal projection operator (we call simply projector hereafter) onto the space spanned by the column vectors of $\mathbf{G}_i$, denoted by $\text{Sp}(\mathbf{G}_i)$, and matrix $\mathbf{\Sigma}_i$ is its orthogonal complement (the orthogonal projector onto $\text{Ker}(\mathbf{G}_i')$, where $\text{Ker}(\mathbf{A}')$

denotes the null space of $\mathbf{A}'$, that is, the set of vectors $\mathbf{m}$ such that $\mathbf{A}'\mathbf{m} = \mathbf{0}$). Matrix $\mathbf{\Psi}_i$ is the projector onto $\mathrm{Sp}(\mathbf{H}_i)$ along $\mathrm{Ker}(\mathbf{H}_i'\mathbf{D}_i)$, and matrix $\mathbf{\Omega}_i$ is the projector onto $\mathrm{Ker}(\mathbf{H}_i'\mathbf{D}_i)$ along $\mathrm{Sp}(\mathbf{H}_i)$. The $\mathbf{D}_i$ is called a metric matrix, which is nonnegative definite.

The decomposition of $\mathbf{Z}_i$ is unique, and each submatrix in (1) can be given a specific interpretation (Takane & Shibayama, 1991). The first term represents the portion of $\mathbf{Z}_i$ explained by both $\mathbf{G}_i$ and $\mathbf{H}_i$, the second term by $\mathbf{G}_i$, but not by $\mathbf{H}_i$, the third term by $\mathbf{H}_i$, but not by $\mathbf{G}_i$, and the last term by neither $\mathbf{G}_i$ nor $\mathbf{H}_i$. The four submatrices in (1) are either columnwise or rowwise orthogonal in their respective metric matrices. (The metric matrix for $\mathbf{\Phi}_i$ and $\mathbf{\Sigma}_i$ is an identity matrix.) This implies

$$\mathrm{SS}(\mathbf{Z}_i) = \mathrm{SS}(\mathbf{\Phi}_i\mathbf{Z}_i\mathbf{\Psi}_i) + \mathrm{SS}(\mathbf{\Phi}_i\mathbf{Z}_i\mathbf{\Omega}_i) + \mathrm{SS}(\mathbf{\Sigma}_i\mathbf{Z}_i\mathbf{\Psi}_i) + \mathrm{SS}(\mathbf{\Sigma}_i\mathbf{Z}_i\mathbf{\Omega}_i), \tag{2}$$

where $\mathrm{SS}(\mathbf{Y}) = \mathrm{tr}(\mathbf{Y}'\mathbf{Y})$ (see Takane & Shibayama, 1991). That is, the sum of squares of $\mathbf{Z}_i$ can be defined as the sum of sums of squares of the four decomposed submatrices in (1).

Then, the decompositions of all discrete data matrices can be expressed in terms of $\mathbf{Z}$ as follows.

$$\begin{aligned}\mathbf{Z} &= (\mathbf{\Phi} + \mathbf{\Sigma})\tilde{\mathbf{Z}}(\mathbf{\Psi} + \mathbf{\Omega}) \\ &= \mathbf{\Phi}\tilde{\mathbf{Z}}\mathbf{\Psi} + \mathbf{\Phi}\tilde{\mathbf{Z}}\mathbf{\Omega} + \mathbf{\Sigma}\tilde{\mathbf{Z}}\mathbf{\Psi} + \mathbf{\Sigma}\tilde{\mathbf{Z}}\mathbf{\Omega},\end{aligned} \tag{3}$$

where

$$\begin{aligned}\tilde{\mathbf{Z}} &= [\mathbf{Z}_1 \oplus \mathbf{Z}_2 \oplus \cdots \oplus \mathbf{Z}_K], \\ \mathbf{\Phi} &= [\mathbf{\Phi}_1, \mathbf{\Phi}_2, \ldots, \mathbf{\Phi}_K], \\ \mathbf{\Sigma} &= [\mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \ldots, \mathbf{\Sigma}_K], \\ \mathbf{\Psi} &= [\mathbf{\Psi}_1 \oplus \mathbf{\Psi}_2 \oplus \cdots \oplus \mathbf{\Psi}_K],\end{aligned}$$

and

$$\mathbf{\Omega} = [\mathbf{\Omega}_1 \oplus \mathbf{\Omega}_2 \oplus \cdots \oplus \mathbf{\Omega}_K].$$

In this paper, the supermatrices of the projectors (i.e., $\mathbf{\Phi}$, $\mathbf{\Sigma}$, $\mathbf{\Psi}$, and $\mathbf{\Omega}$) are assumed consisting of only the same form of the projectors of row or column constraint matrices. For instance, $\mathbf{\Phi}$ comprises the projectors onto $\mathrm{Sp}(\mathbf{G}_i)$ $(i = 1, \ldots, K)$, and the counterpart supermatrix, $\mathbf{\Sigma}$, contains their ortho-complements, and so on. However, we could also contemplate the situations in which the supermatrices include both forms of projectors. For example, suppose $K = 3$. Then $\mathbf{\Phi}$ may include $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$, and $\mathbf{\Sigma}_3$, and then $\mathbf{\Sigma} = [\mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \mathbf{\Phi}_3]$. Note that once the projector submatrices of $\mathbf{\Phi}$ are chosen, those of $\mathbf{\Sigma}$ are automatically decided as their orthogonal complement matrices. It is the case for $\mathbf{\Psi}$ and $\mathbf{\Omega}$. This allows the decompositions in (3) to remain orthogonal. The expression of (3) can therefore be quite versatile, by which various kinds of combinations among the decomposed submatrices selected from each data matrix can be derived for analysis.

The four terms in (3) correspond with four sets of K decomposed submatrices, given in (1). Notice however that the four terms can be derived only when both sets of row and column constraint matrices are available. We may, more generally, consider the following decomposed subsets:

$\mathbf{S}_1$	$\mathbf{S}_2$	$\mathbf{S}_3$	$\mathbf{S}_4$	$\mathbf{S}_5$	$\mathbf{S}_6$	$\mathbf{S}_7$	$\mathbf{S}_8$	$\mathbf{S}_9$
$\mathbf{Z}$	$\mathbf{\Phi}\tilde{\mathbf{Z}}$	$\mathbf{\Sigma}\tilde{\mathbf{Z}}$	$\mathbf{Z}\mathbf{\Psi}$	$\mathbf{Z}\mathbf{\Omega}$	$\mathbf{\Phi}\tilde{\mathbf{Z}}\mathbf{\Psi}$	$\mathbf{\Phi}\tilde{\mathbf{Z}}\mathbf{\Omega}$	$\mathbf{\Sigma}\tilde{\mathbf{Z}}\mathbf{\Psi}$	$\mathbf{\Sigma}\tilde{\mathbf{Z}}\mathbf{\Omega}$

(4)

In (4), $\mathbf{S}_1$ simply indicates a set of the unconstrained data matrices. The next two decomposed subsets, denoted by $\mathbf{S}_2$ and $\mathbf{S}_3$, can be derived when only $\mathbf{G}$ is incorporated. The $\mathbf{S}_2$ corresponds with the portions of $\mathbf{Z}$ that can be explained by $\mathbf{G}$ while $\mathbf{S}_3$ indicates the residuals after the effects due to $\mathbf{G}$ are eliminated. Similarly, $\mathbf{S}_4$ and $\mathbf{S}_5$ are derived when only $\mathbf{H}$ is incorporated. The last four decomposed subsets ($\mathbf{S}_6$ through $\mathbf{S}_9$) correspond with the four terms in the full decompositions in (3), when both $\mathbf{G}$ and $\mathbf{H}$ are incorporated.

In (3), it is assumed that each item has its corresponding column constraint matrix. It indicates that each column constraint matrix is used to specify relationships among the categories within the same item. However, it is also possible to consider a column constraint matrix that represents relationships among categories across different items. We can easily add the across-item column constraint matrix to (3). For instance, suppose an across-item column constraint matrix for items 1 and 2, denoted by $\mathbf{H}_{12}$, and the corresponding projectors, denoted by $\mathbf{\Psi}_{12}$ and $\mathbf{\Omega}_{12}$. Then, $\mathbf{\Psi}$ and $\mathbf{\Omega}$ in (3) are given as $\mathbf{\Psi} = [\mathbf{\Psi}_{12} \oplus \mathbf{\Psi}_3 \oplus \cdots \oplus \mathbf{\Psi}_K]$, and $\mathbf{\Omega} = [\mathbf{\Omega}_{12} \oplus \mathbf{\Omega}_3 \oplus \cdots \oplus \mathbf{\Omega}_K]$, with which the decompositions of $\mathbf{Z}$ are derived as in (3). This implies nothing but we combine $\mathbf{Z}_1$ and $\mathbf{Z}_2$ as a single data matrix, say $\mathbf{Z}_{12} = [\mathbf{Z}_1, \mathbf{Z}_2]$, and specify the relationships among the variables in $\mathbf{Z}_{12}$ by $\mathbf{H}_{12}$. In this case, we still use

$$\mathbf{D}_{12} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix}, \text{ where } \mathbf{D}_1 = \mathbf{Z}_1'\mathbf{Z}_1, \text{ and } \mathbf{D}_2 = \mathbf{Z}_2'\mathbf{Z}_2, \text{ and define}$$

$$\mathbf{\Psi}_{12} = \mathbf{H}_{12}(\mathbf{H}_{12}'\mathbf{D}_{12}\mathbf{H}_{12})^{-}\mathbf{H}_{12}'\mathbf{D}_{12}.$$

The zero-centroid property (that the coordinates of category points add up to zero within each item when weighted by the marginal frequencies of responses to the categories) is satisfied for the categories of the two items together.

The row and column constraints may be specified by either the reparametrization or the null-space method (Böckenholt & Takane, 1994; Takane et al., 1991). The former method specifies the space spanned by column vectors of a constraint matrix, while the latter specifies the ortho-complement space. In this paper, as (1) and (3) imply, all linear constraints are imposed by the reparametrization method, that is, the space spanned by a set of constraints is directly specified, onto which a data matrix is projected. However, it is sometimes easier to specify constraints in the null-space form (e.g., equality among categories). In this paper, the constraints expressed in the null-space form are transformed into the reparametrization form. The transformation is straightforward. For instance, assume an equality constraint matrix, say $\mathbf{L}'$, specified by the null-space method as follows

$$\mathbf{L}'\mathbf{u} = \mathbf{0}, \tag{5}$$

where $\mathbf{u}$ is a vector of parameters. If we assume that the first and the last elements of $\mathbf{u}$ are equal, for example, $\mathbf{L}'$ comes down to a vector whose first element is 1, the last element is -1, and the other elements are zeros. We may reparametrize (5) into

$$\mathbf{u} = \tilde{\mathbf{H}}\mathbf{u}^* \tag{6}$$

for some $\mathbf{u}^*$, where $\tilde{\mathbf{H}} = \mathbf{I} - \mathbf{L}(\mathbf{L}'\mathbf{L})^{-}\mathbf{L}'$. This implies that $\text{Ker}(\mathbf{L}') = \text{Sp}(\tilde{\mathbf{H}})$. (see Seber, 1984, pp. 403–405; Takane et al., 1991). The derived $\tilde{\mathbf{H}}$ can now be used for the decompositions of (3). In the example section, a column constraint representing equality of categories across items is specified in the null-space form, and it is turned into a reparametrized constraint.

2.2. *Multiple Correspondence Analysis of the Decomposed Submatrices*

Once $\mathbf{Z}$ is decomposed, MCA can be applied to any decomposed subset, given in (4), to explore associations among the decomposed data matrices in the subset. MCA amounts to the correspondence analysis (CA) of a decomposed subset or the CA of the Burt table of the decomposed subset. Computationally, thus, MCA comes down to the generalized singular value decomposition (GSVD) of the decomposed subset with certain metric matrices (Greenacre, 1984, pp. 138–143). The GSVD of matrix $\mathbf{W}$ with metric matrices $\mathbf{M}$ and $\mathbf{N}$ (assumed positive definite) is defined as

$$\mathbf{W} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}', \tag{7}$$

where $\mathbf{U}'\mathbf{M}\mathbf{U} = \mathbf{I}$ and $\mathbf{V}'\mathbf{N}\mathbf{V} = \mathbf{I}$, and $\boldsymbol{\Lambda}$ is diagonal and positive definite (e.g., Greenacre, 1984, Appendix A). It is denoted as $\text{GSVD}(\mathbf{W})_{M,N}$. The $\text{GSVD}(\mathbf{W})_{M,N}$ can be obtained as follows. Let $\mathbf{M} = \mathbf{T}\mathbf{T}'$ and $\mathbf{N} = \mathbf{F}\mathbf{F}'$ be arbitrary square root decompositions of $\mathbf{M}$ and $\mathbf{N}$, respectively. Let the ordinary singular value decomposition (SVD) of $\mathbf{T}'\mathbf{W}\mathbf{F}$ be denoted by

$$\mathbf{T}'\mathbf{W}\mathbf{F} = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}', \tag{8}$$

where $\mathbf{P}'\mathbf{P} = \mathbf{Q}'\mathbf{Q} = \mathbf{I}$. Then, the $\mathbf{U}$ and $\mathbf{V}$ in (7) are obtained by $\mathbf{U} = (\mathbf{T}')^{-1}\mathbf{P}$ and $\mathbf{V} = (\mathbf{F}')^{-1}\mathbf{Q}$, respectively, and $\boldsymbol{\Lambda} = \boldsymbol{\Delta}$. The GSVD of a matrix always exists (Takane & Shibayama, 1991). It is unique if $\mathbf{M}$ and $\mathbf{N}$ are nonsingular (as has been assumed). If they are not, we may use any g-inverses in the above formulas. It is convenient to use the Moore-Penrose inverses in the above formulas to obtain a unique solution.

The ordinary (unconstrained) MCA is the CA of $\mathbf{S}_1$. This amounts to calculating

$$\text{GSVD}(\mathbf{S}_1\boldsymbol{\Theta}_1^{-1})_{I,\Theta_1}, \tag{9}$$

where $\boldsymbol{\Theta}_1 = K\tilde{\mathbf{S}}_1'\tilde{\mathbf{S}}_1 = K[\mathbf{D}_1 \oplus \mathbf{D}_2 \oplus \cdots \oplus \mathbf{D}_K]$ (Greenacre, 1984, pp. 138–139). The $\tilde{\mathbf{S}}_1$ is a block diagonal matrix with each submatrix of $\mathbf{S}_1$ as elements. (In this case, therefore, $\tilde{\mathbf{S}}_1 = \tilde{\mathbf{Z}}$.) The $\boldsymbol{\Theta}_1$ is a diagonal matrix whose elements are row (or column) marginals of $\mathbf{S}_1'\mathbf{S}_1$, called the Burt table. The CA of the Burt table provides the same solutions as $\text{MCA}(\mathbf{S}_1)$ (Greenacre, pp. 140–143). The singular values obtained from the analysis of the Burt table are the squares of those obtained from (9). We denote (9) by $\text{MCA}(\mathbf{S}_1)$ here.

$\text{MCA}(\mathbf{S}_1)$ can be alternatively defined as K-set canonical correlation analysis or homogeneity analysis of $\mathbf{S}_1$ (e.g., Gifi, 1990; Yanai, 1998). This amounts to minimizing the following criterion

$$f = K^{-1}\sum_{i=1}^{K}\text{SS}(\mathbf{X} - \mathbf{Z}_i\mathbf{V}_i), \tag{10}$$

with respect to $\mathbf{V}_i$ for a given $\mathbf{X}$, where $\mathbf{X}$ is the matrix of object scores, subject to $\mathbf{X}'\mathbf{X} = \mathbf{I}$. In (10), the estimate of $\mathbf{V}_i$ is given by

$$\mathbf{V}_i = (\mathbf{Z}_i'\mathbf{Z}_i)^{-}\mathbf{Z}_i'\mathbf{X}. \tag{11}$$

Putting (11) in (10) leads to

$$f^* = K^{-1}\sum_{i=1}^{K}\text{SS}(\mathbf{X} - \boldsymbol{\Gamma}_i\mathbf{X}), \tag{12}$$

where $\boldsymbol{\Gamma}_i = \mathbf{Z}_i(\mathbf{Z}_i'\mathbf{Z}_i)^{-}\mathbf{Z}_i'$. Minimizing (12) thus reduces to maximizing

$$\text{tr}\left(\mathbf{X}'\left[\sum_{i=1}^{K}\boldsymbol{\Gamma}_i\right]\mathbf{X}\right), \tag{13}$$

with respect to $\mathbf{X}$. This amounts to calculating the eigenvalue decomposition of $\sum_{i=1}^{K}\boldsymbol{\Gamma}_i$ (Yanai, 1998), which in turn is equivalent to GSVD of $\mathbf{S}_1\boldsymbol{\Theta}_1^{-1}$ (Gifi, 1990, pp. 107–109).

We note that when $K = 2$, that is, $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2]$, $\text{MCA}(\mathbf{S}_1)$ reduces to ordinary CA (OCA) of $\mathbf{Z}_1'\mathbf{Z}_2$. In this case, the difference between MCA and OCA is in rescaling of principal coordinates (Greenacre, 1984, pp. 130–133; Gifi, 1990, pp. 272–273). The inertias or squared singular values obtained from OCA are related to those from MCA as follows: let ρ and λ denote inertias obtained from OCA and from MCA, respectively. Then,

$$\rho = 4\left(\lambda - \frac{1}{2}\right)^2 \tag{14}$$

(Greenacre, 1984, p. 131).

We may consider various types of constrained MCA methods by replacing $\mathbf{S}_1$ (and $\tilde{\mathbf{S}}_1$) in (9) with any other decomposed subsets of $\mathbf{Z}$. For example, we may consider MCA($\mathbf{S}_4$), which amounts to obtaining

$$\text{GSVD}(\mathbf{S}_4\mathbf{\Theta}_4^{-1})_{I,\Theta_4}, \tag{15}$$

where $\mathbf{\Theta}_4 = K\tilde{\mathbf{S}}_4'\tilde{\mathbf{S}}_4 = K(\mathbf{\Psi}_1'\mathbf{D}_1\mathbf{\Psi}_1 \oplus \cdots \oplus \mathbf{\Psi}_K'\mathbf{D}_K\mathbf{\Psi}_K)$. This analyzes the portions of data matrices accounted for by their corresponding column constraint matrices. This may be called constrained multiple correspondence analysis (CMCA). This is equivalent to constrained homogeneity analysis of dummy variables (Gifi, 1990, p. 332; van Buuren & de Leeuw, 1992), which amounts to minimizing

$$f = K^{-1}\sum_{i=1}^{K} \text{SS}(\mathbf{X} - \mathbf{Z}_i\mathbf{H}_i\mathbf{V}_i) = K^{-1}\sum_{i=1}^{K} \text{SS}(\mathbf{X} - \mathbf{Z}_i^*\mathbf{V}_i), \tag{16}$$

with respect to $\mathbf{V}_i$ for a given $\mathbf{X}$, where $\mathbf{Z}_i^* = \mathbf{Z}_i\mathbf{H}_i$. As in MCA($\mathbf{S}_1$), minimizing (16) amounts to obtaining the eigenvalue decomposition of $\sum_{i=1}^{K}\mathbf{\Gamma}_i^*$, where $\mathbf{\Gamma}_i^* = \mathbf{Z}_i^*(\mathbf{Z}_i^{*\prime}\mathbf{Z}_i^*)^-\mathbf{Z}_i^{*\prime}$. This is also equivalent to the eigenvalue decomposition of $\sum_{i=1}^{K}\mathbf{\Pi}_i$, where $\mathbf{\Pi}_i = \mathbf{Z}_i\mathbf{\Psi}_i(\mathbf{\Psi}_i'\mathbf{Z}_i'\mathbf{Z}_i\mathbf{\Psi}_i)^-\mathbf{\Psi}_i'\mathbf{Z}_i'$, which in turn is equivalent to calculating (15) (see Appendix 1 for the equivalence between $\mathbf{\Gamma}_i^*$ and $\mathbf{\Pi}_i$).

In addition, MCA($\mathbf{S}_4$) (or the constrained homogeneity analysis) is equivalent to generalized canonical correlation analysis with linear constraints (GCCAC) for sets of dummy variables (Yanai, 1998). The only difference is that in MCA($\mathbf{S}_4$) linear constraints are imposed by the reparametrization method whereas in GCCAC they are imposed by the null space method. One form of linear constraints can be transformed to the other form, as explained in the previous section. When $K = 2$, furthermore, MCA($\mathbf{S}_4$) subsumes canonical correspondence analysis (CCA) by ter Braak (1986), which amounts to the CA of $\mathbf{\Psi}_1'\mathbf{Z}_1'\mathbf{Z}_2\mathbf{\Psi}_2$ (Böckenholt & Takane, 1994; Takane & Hwang, 2000). It thus includes canonical analysis with linear constraints (CALC) by Böckenholt & Böckenholt (1990), which is equivalent to CCA (Takane et al., 1991).

In MCA($\mathbf{S}_4$), if $\mathbf{H}$ is used to impose equality of categories across items, there exists an alternative method, where the relevant columns of $\mathbf{Z}_i$ (i.e., the categories to be assumed equal across items) are summed, they are replaced by the sum, and a CA on the resultant data matrix is performed (Greenacre, 1984, p. 95; van Buuren & de Leeuw, 1992). However, this method is limited to imposing across-item equality constraints only.

We may also consider MCA($\mathbf{S}_3$), obtained by

$$\text{GSVD}(\mathbf{S}_3\mathbf{\Theta}_3^{-1})_{I,\Theta_3}, \tag{17}$$

where $\mathbf{\Theta}_3 = K\tilde{\mathbf{S}}_3'\tilde{\mathbf{S}}_3 = K(\mathbf{Z}_1'\mathbf{\Sigma}_1\mathbf{Z}_1 \oplus \cdots \oplus \mathbf{Z}_K'\mathbf{\Sigma}_K\mathbf{Z}_K)$. This is the MCA of the residuals obtained after the effects due to $\mathbf{G}$ are removed from $\mathbf{Z}$. If there is a single row constraint matrix common to all data matrices (i.e., $\mathbf{G}^* = \mathbf{G}_1 = \cdots = \mathbf{G}_K$), MCA($\mathbf{S}_3$) is equivalent to partial MCA, which amounts to minimizing

$$f = K^{-1}\sum_{i=1}^{K} \text{SS}(\mathbf{X} - \mathbf{\Sigma}^*\mathbf{Z}_i\mathbf{V}_i), \tag{18}$$

with respect to $\mathbf{V}_i$ (Yanai & Maeda, 2000), where $\mathbf{\Sigma}^* = \mathbf{I} - \mathbf{G}^*(\mathbf{G}^{*\prime}\mathbf{G}^*)^-\mathbf{G}^{*\prime}$. Partial MCA can be also viewed as a special case of Nishisato's (1984) forced classification (see Nishisato, 1994, p. 245). As $K = 2$, MCA($\mathbf{S}_3$) subsumes partial correspondence analysis (Daudin, 1980; Yanai, 1986) as a special case, which is the CA of $\mathbf{Z}_1'\mathbf{\Sigma}^*\mathbf{Z}_2$.

Partial MCA can be also performed in a different way: we replicate $\mathbf{G}^*$ over all $\mathbf{Z}_i$'s, and apply K-set canonical correlation analysis or homogeneity analysis to the enlarged data sets, for example, $\tilde{\mathbf{Z}}_i = [\mathbf{Z}_i, \mathbf{G}^*]$ (Gifi, 1990, p. 248; Yanai & Maeda, 2000). In this method, however, there are $g = \text{rank}(\mathbf{G}^*)$ trivial solutions, and the $(g+1)$-th or higher singular values and the corresponding singular vectors are only considered.

If there are K distinct row constraint matrices available (i.e., $\mathbf{G}_l \neq \mathbf{G}_{l'}$, where $l = 1, \ldots, K-1$ and $l' = l+1, \ldots, K$), MCA($\mathbf{S}_3$) may be called multi-partial MCA, inspired by the terminology of bipartial canonical correlation analysis (Timm & Carlson, 1976). This case reduces to bipartial correspondence analysis when $K = 2$, which is the CA of $\mathbf{Z}_1'\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{Z}_2$ (Takane & Hwang, 2000). Also, we may suppose that row constraint matrices are available only for some, say K^* $(< K)$, data matrices (i.e., $\mathbf{G}_t \neq \mathbf{G}_{t'}$, where $t = 1, \ldots, K^* - 1$ and $t' = t+1, \ldots, K^*$). If (18) is applied to the residual matrices obtained after removing the effects of $\mathbf{G}$ from K^* corresponding data matrices and the remaining $K - K^*$ unconstrained data matrices, this case may be called multi-part MCA, borrowing from the terminology of part canonical correlation analysis (Timm & Carlson, 1976). This reduces to part correspondence analysis (Takane & Hwang, 2000) for the pair of the unconstrained data matrix and the residual matrix after partialling out the effect of the row constraint matrix, which amounts to the CA of $\mathbf{Z}_1'\mathbf{\Sigma}_2\mathbf{Z}_2$ or $\mathbf{Z}_1'\mathbf{\Sigma}_1\mathbf{Z}_2$.

We may also apply MCA to $\mathbf{S}_8$, the portions of data matrices explained by their column constraint matrices, but not by their row constraint matrices. This is analogous to solving

$$\text{GSVD}(\mathbf{S}_8\mathbf{\Theta}_8^{-1})_{I,\mathbf{\Theta}_8}, \tag{19}$$

where $\mathbf{\Theta}_8 = K\tilde{\mathbf{S}}_8'\tilde{\mathbf{S}}_8 = K(\mathbf{\Psi}_1'\mathbf{Z}_1'\mathbf{\Sigma}_1\mathbf{Z}_1\mathbf{\Psi}_1 \oplus \cdots \oplus \mathbf{\Psi}_K'\mathbf{Z}_K'\mathbf{\Sigma}_K\mathbf{Z}_K\mathbf{\Psi}_K)$. This may include a variety of constrained MCA. For instance, suppose $\mathbf{H}_l \neq \mathbf{H}_{l'}$. Then, if $\mathbf{G}_l \neq \mathbf{G}_{l'}$, this case may be called multi-partial CMCA. This case subsumes bipartial CCA by Takane and Hwang (2000), which is the CA of $\mathbf{\Psi}_1'\mathbf{Z}_1'\mathbf{\Sigma}_1\mathbf{\Sigma}_2\mathbf{Z}_2\mathbf{\Psi}_2$. On the other hand, if $\mathbf{G}^* = \mathbf{G}_1 = \cdots = \mathbf{G}_K$, this may be called partial CMCA. This case is the multivariate extension of partial CCA (Takane & Hwang, 2000), which is the CA of $\mathbf{\Psi}_1'\mathbf{Z}_1'\mathbf{\Sigma}^*\mathbf{Z}_2\mathbf{\Psi}_2$. If $\mathbf{G}_t \neq \mathbf{G}_{t'}$, in addition, this case may be called multi-part CMCA.

To sum up, after $\mathbf{Z}$ is decomposed into several subsets by $\mathbf{G}$ and/or $\mathbf{H}$, we may apply MCA to any decomposed subsets of $\mathbf{Z}$ according to our empirical interests. This often yields new types of constrained MCA methods (e.g., MCA($\mathbf{S}_8$)), which may allow us to explore associations among variables from more diverse perspectives. Some of the constrained MCA may be regarded as multivariate extensions of existing linearly constrained correspondence analysis methods. In the next section, we demonstrate the use of (15), (17), and (19) with an example.

3. Example: The French Worker Data

The present example is part of the French Worker Survey (Adam, Bon, Capdevielle, & Mouriaux, 1970). The survey was conducted in July 1969 on a sample of French workers to explore the political attitudes and social behavior of the working class. From 70 survey items, we only used the four items reported in Le Roux and Rouanet (1998). Although each of the four items originally consisted of eight response categories, we only used four categories for each item, removing rarely chosen categories and/or combining similar categories. The four items and their four response categories are provided in Appendix 2. Categories in items 1 and 2 indicate different French trade unions. In both items, Category 1 (CGT) is known to be strongly linked with Communist party, Category 2 (CFDT) is loosely connected with the noncommunist-left, Category 3 (Autonomous) is inclined toward right wing, and Category 4 (Nonaffiliated) indicates no union to support (Item 1) or no union to belong to (Item 2). Categories in Item 3 correspond with presidential candidates from different political parties. Category 1 (Jacques Duclos) is the candidate from Communist party, Category 2 (Gaston Defferre) is from Socialist party, Category 3 (Alain Poher) is the candidate representing the political center, and Category 4 (Georges Pompidou) is the Gaullist candidate, who won the election. The Gaullist candidate speaks for right wing politically. Categories in Item 4 correspond with different political parties. Fifty four response patterns were observed and the sample size was 274.

We used the first three items as data matrices, indicating that $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3]$. We then chose the last item as a row constraint matrix, since we were interested in seeing how preferences for political parties were related with workers' responses to other items, particularly their voting behavior revealed by item 3. It was common to all $\mathbf{Z}_i$'s ($i = 1, 2, 3$), implying that $\mathbf{G}_1 = \mathbf{G}_2 = \mathbf{G}_3$. We also considered a column constraint matrix to represent equality of categories across items 1 and 2, denoted by $\mathbf{L}'_{12}$, as follows.

$$\mathbf{L}'_{12} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

The first row of $\mathbf{L}'_{12}$ indicates the equality of the first categories of items 1 and 2, the second row the equality of the second categories, the third row the equality of the third categories, and the last row the equality of the fourth categories. Such equality assumptions among the four categories in items 1 and 2 may make sense because the same categories were repeated in two similar (or comparable) items regarding trade unions, and similar interpretation of the categories may be possible (see van Buuren & de Leeuw, 1992). The $\mathbf{L}'_{12}$ was transformed into the reparametrized constraint matrix, denoted by $\tilde{\mathbf{H}}_{12}$, by the procedure described in section 2.1. On the other hand, we did not consider a column constraint matrix for item 3, since it was difficult to come up with any specific relationships among the candidates due to their politically distinct backgrounds. Therefore, we had $\mathbf{H} = [\tilde{\mathbf{H}}_{12} \oplus \mathbf{I}]$.

MCA was applied to the subsets obtained after $\mathbf{Z}$ was decomposed according to $\mathbf{G}$ and $\mathbf{H}$. Here we present four cases, MCA($\mathbf{S}_1$), MCA($\mathbf{S}_4$), MCA($\mathbf{S}_3$), and MCA($\mathbf{S}_8$), corresponding to the unconstrained MCA, CMCA, partial MCA, and partial CMCA, respectively. The inertias or squared singular values estimated from each case are presented in Table 1. In MCA, however, the total inertia (i.e., the sum of the inertias) is inflated since MCA amounts to fitting both diagonal

TABLE 1.
The inertias obtained from MCA($\mathbf{S}_1$), MCA($\mathbf{S}_4$), MCA($\mathbf{S}_3$), and MCA($\mathbf{S}_8$)

	MCA($\mathbf{S}_1$)	MCA($\mathbf{S}_4$)	MCA($\mathbf{S}_3$)	MCA($\mathbf{S}_8$)
	.6863	.6763	.6195	.6084
	.5911	.5631	.5122	.5044
	.4784	.4378	.4283	.4216
	.3557	.3332	.3582	.3313
	.2629	.2459	.3244	.3147
	.2114	.1979	.2836	.2793
	.1938	.0000	.2271	.0000
	.1335	.0000	.1521	.0000
	.0868	.0000	.0945	.0000
	.0000	.0000	.0000	.0000
	.0000	.0000	.0000	.0000
	.0000	.0000	.0000	.0000
Total	3.0000	2.4543	3.0000	2.4597

TABLE 2.
The adjusted inertias and their percentages of the total inertia in the parenthesis obtained from MCA(S_1), MCA(S_4), MCA(S_3), and MCA(S_8)

	MCA(S_1)		MCA(S_4)		MCA(S_3)		MCA(S_8)	
	.2803	(58.60)	.2646	(64.85)	.1843	(66.30)	.1703	(67.13)
	.1495	(31.26)	.1188	(29.13)	.0720	(26.90)	.0659	(25.97)
	.0473	(9.90)	.0246	(6.02)	.0203	(7.30)	.0175	(6.90)
	.0011	(0.24)			.0014	(0.50)		
Total	.4783		.4080		.2780		.2536	

and off-diagonal blocks of the Burt table, given by $\mathbf{S}_i'\mathbf{S}_i$. It leads that the proportions of the total inertia explained by inertias are underestimated. One way of dealing with this problem is to adjust the inertias according to Benzécri's (1979) formula, quoted in Greenacre (1984, p. 145). Let $\tilde{\lambda}$ denote the adjusted inertia corresponding to λ. Then, the formula is given by

$$\tilde{\lambda} = \left(\frac{K}{K-1}\right)^2 \left(\lambda - \frac{1}{K}\right)^2. \tag{20}$$

Note that $\tilde{\lambda}$ turns out to be equal to ρ in (14) when $K = 2$. Thus, we see that $\tilde{\lambda}$ is the inertia obtained from when the diagonal blocks of the Burt table are set to zero matrices. The above formula is applied only for the inertias greater than $1/K$. In this example, therefore, the inertias greater than $1/3$ were adjusted. The adjusted inertias estimated from each case are presented in Table 2, along with their percentages of the recalculated total inertia (i.e., the sum of the adjusted inertias).

In Table 2, the first two adjusted inertias explain about 90% or more than 90% of the adjusted total inertia in all cases. Thus a two-dimensional solution seems sufficient to account for the variations among item categories for all the cases.

The two-dimensional configuration for the estimated category points in MCA($\mathbf{S}_1$) is presented in Figure 1. (Note that the two-dimensional principal coordinates were rescaled according to their corresponding adjusted inertias.)

The estimated category points of Item 1 are labeled as cg1, cf1, au1, and na1, those of Item 2 as cg2, cf2, au2, and na2, and those of Item 3 as COM, SOC, CTR, and GAU. The order of the labels is equivalent to that of the categories in each item. In the figure, the right middle portion seems to represent a politically communist-left position, characterized by cg1, cg2, and COM. It thus may be safe to say that the French workers voting for a list sponsored by CGT or affiliated to CGT were more likely to vote for the communist candidate, Jacques Duclos (COM). On the other hand, the middle bottom portion seems to stand for a politically more rightist position, distinguished by au1, au2, GAU. It is found that a group of workers voting for an independent list or belonging to no union, represented by na1 and na2, seemed to show similar political attitudes to the rightist group of workers, and both groups of workers were more likely to vote for the Gaullist candidate, Georges Pompidou, labeled as GAU. The top left portion in the figure seems to express a noncommunist-left or socialist position, featured by cf1 and cf2. It may indicate that a group of French workers voting for or belonging to CFDT tended to vote for the candidate from Socialist party, Gaston Defferre, labeled as SOC. Also, the Socialist candidate may be found to be more similar to the politically central candidate, Alain Poher, labeled as CTR. It seems to make sense because socialism may be seen to be ideologically closer to the central-left.

Figure 2 displays the two-dimensional configuration of the estimated category points obtained from MCA($\mathbf{S}_4$). This case implies that $\mathbf{H} \neq \mathbf{I}$ and $\mathbf{G} = \mathbf{I}$, which is analogous to the columnwise constrained MCA. The response categories with the same category number in items 1 and 2 were assigned the same value, due to imposing the across-item constraint. In Figure 2,

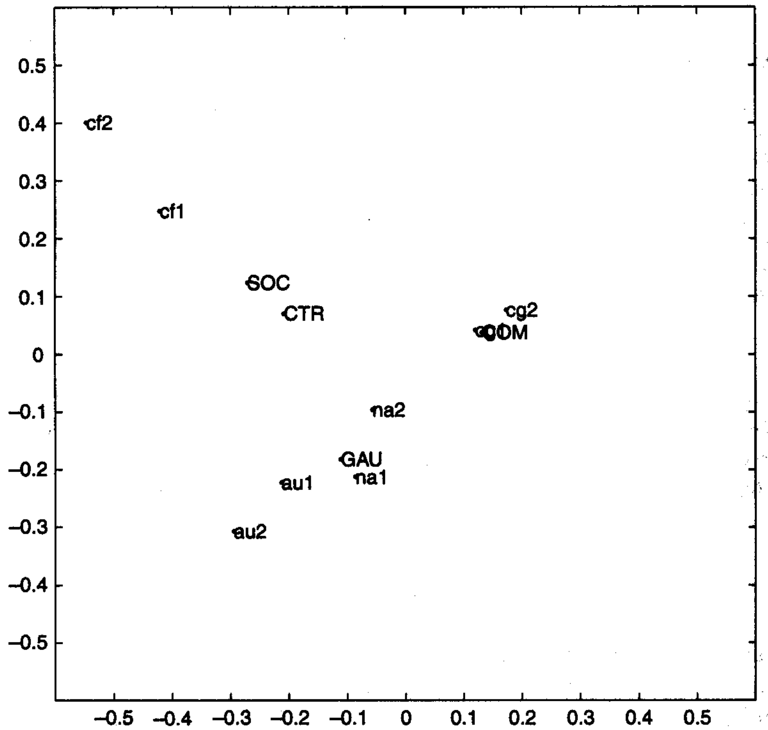

FIGURE 1.
The two-dimensional category point configuration of the French worker data derived from MCA($\mathbf{S}_1$).

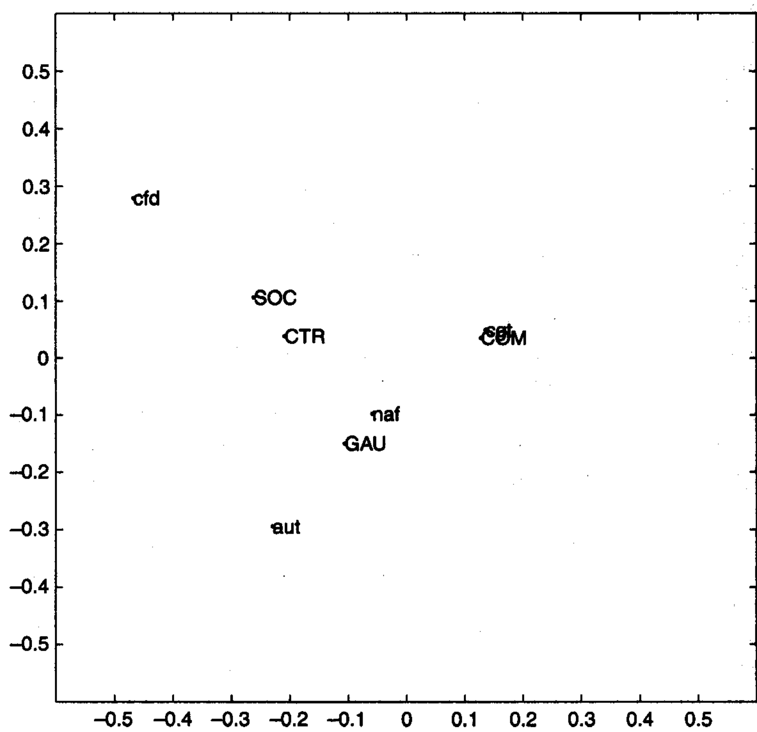

FIGURE 2.
The two-dimensional category point configuration of the French worker data derived from MCA($\mathbf{S}_4$).

the pairs of the same-valued categories are displayed under a single label, that is, cg1 and cg2 are combined as cgt, cf1 and cf2 as cfd, au1 and au2 as aut, and na1 and na2 as naf. We may say that the single-labeled category points stand for preference/support for different trade unions. Interpretations of the figure seems to be essentially the same as the unconstrained MCA. This may indicate that the equality restrictions do not dramatically change the solution. This may help us simplify our interpretations, significantly reducing the number of parameters. Although they are not presented to make the figure more concise, the 95% confidence regions (Ramsay, 1978) obtained by the bootstrap method (Efron, 1979) were either very similar to or smaller than those from the unconstrained MCA. This gives the additional piece of information that our assumptions regarding the column structure of the data seem to be reasonable.

MCA was also applied with $\mathbf{H} = \mathbf{I}$ and $\mathbf{G} \neq \mathbf{I}$ (i.e., MCA(S_3)). This case may be called partial MCA, since the effect of a common row constraint matrix was removed from **Z**. Figure 3 presents the two-dimensional display of the category points estimated under partial MCA.

In Figure 3, it seems to be striking that after eliminating the effects of item 4, asking preference for political parties, a group of French workers who voted for or joined CGT seemed to show similar political attitudes to a group of workers without any union to uphold or belong to, and both groups were more likely to vote for the Gaullist candidate, or they tended to support the Gaullist candidate as much as the Communist candidate. This implies the party preferences might be little related to true voting behavior of some groups of French workers.

Finally, we also examined the case of MCA(S_8), which may be called as partial CMCA. Figure 4 shows the two-dimensional configuration of the estimated category points obtained from partial CMCA.

In the display, the response categories belonging to the same category number in items 1 and 2 were assigned the same values due to **H**, and were labeled in the same way as Figure 2.

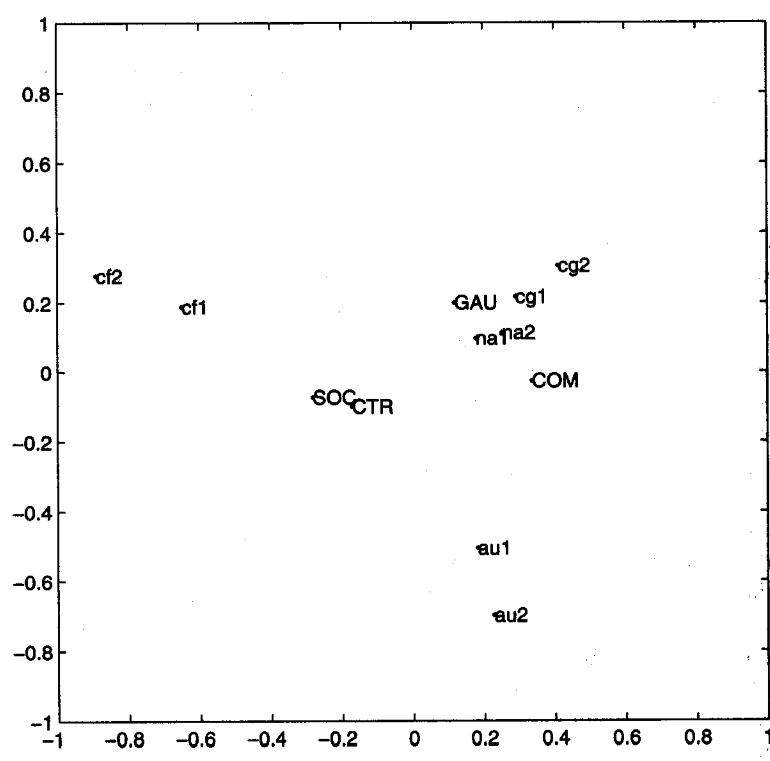

FIGURE 3.
The two-dimensional category point configuration of the French worker data derived from MCA(S_3).

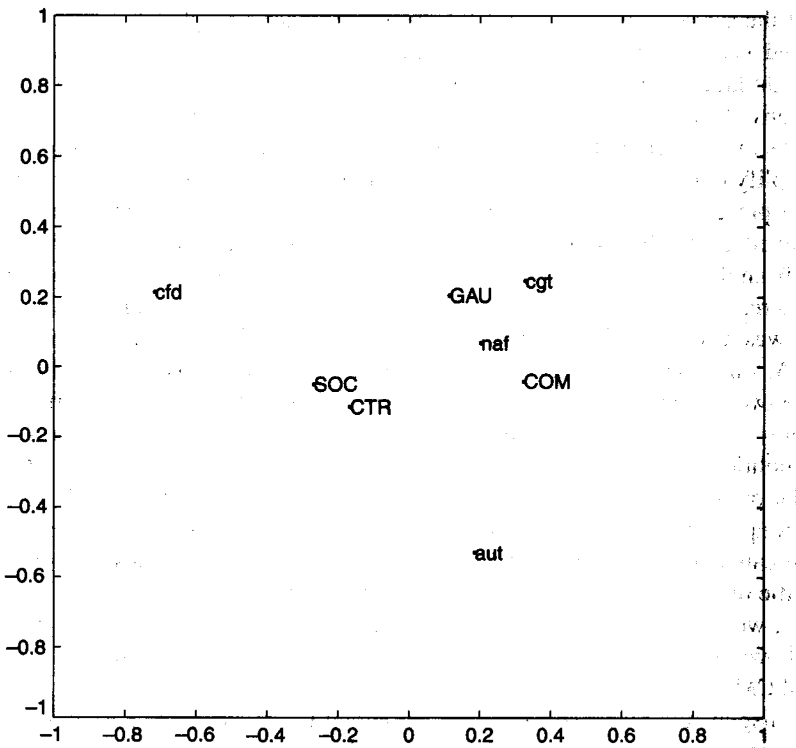

FIGURE 4.
The two-dimensional category point configuration of the French worker data derived from MCA($\mathbf{S}_8$).

Due to eliminating the effect of $\mathbf{G}$, also, the categories representing CGT and nonaffiliation (i.e., cgt and naf) are closely located. This may indicate that the political attitudes of the CGT workers were akin to those of the independent workers. This is consistent with the results of partial MCA.

4. Discussion

In this paper, we provide extensions of MCA by imposing linear row and column constraints in a unified way. They include existing constrained MCA as special cases, and also subsume multivariate versions of a variety of constrained correspondence analysis methods. The method enables us to easily incorporate prior knowledge or additional information into the analysis. Given in the exemplary application, this often leads to parsimonious interpretations of the data. This also allows us to explore the data structure from a broader range of perspectives. The prior knowledge or additional information may stand for a structural hypothesis concerning the data (e.g., equality of variables, etc.). By comparing the unconstrained and constrained solutions, the adequacy of the hypothesis can be empirically investigated.

In this paper, only the nine subsets, given in (4), are considered, which consist of the same form of the decomposed submatrices. However, our analysis does not need to be necessarily restricted to those nine subsets. In principle, we could create a number of additional subsets in various ways. For example, we may include more than one of the nine subsets as submatrices or we may choose certain submatrices from each $\mathbf{S}_i$, and concatenate them into a subset. Such supplementary analyses may not be directly derived from the orthogonal decompositions of the data matrix by row and/or column constraints, but from ad hoc combinations of the nine subsets or their submatrices according to our empirical interests. Nevertheless they may further widen the scope of GCMCA, allowing for studying more diverse aspects of the associations among the data sets. For instance, we may apply MCA to a new subset such as $\mathbf{S}_{10} = [\mathbf{S}_1, \mathbf{S}_2]$. From this we may examine the relationships between the part of **Z** structured by **H** and the residual part, which are usually analyzed separately.

As mentioned in section 2.2, MCA is viewed as a special case of K-set canonical correlation analysis, which uses K sets of dummy variables as data matrices, instead of continuous variables. Hence a similar approach can be readily applied to the continuous case. The constrained K-set canonical correlation analysis includes generalized constrained canonical correlation analysis (Takane & Hwang, 1998) as a special case. In practice, however, it is usually more difficult to find interesting examples in the continuous case than in the discrete case. We note that K-set canonical correlation analysis subsumes PCA (when each set consists of one continuous variable). This motivates extensions of PCA, where the effects of different constraints are incorporated into different variables in the main data, and various aspects of the data may be investigated. The resultant constrained PCA method subsumes constrained PCA of Takane and Shibayama (1991) as a special case. In the present method, only one set of linear constraints is incorporated in each side (row or column side) of a data matrix. However, it may be possible to impose different sets of constraints on different dimensions (DCDD). This type of constraints may, in some cases, yield more meaningful analyses since it often better captures the essence of original empirical hypotheses. Takane, Kiers, and de Leeuw (1995) presented the use of the DCDD type of constraints in PCA and classical multidimensional scaling. A similar approach may be adopted in MCA to incorporate such types of constraints.

Appendix A

We show that $\mathbf{\Gamma}_i^* = \mathbf{Z}_i^*(\mathbf{Z}_i^{*\prime}\mathbf{Z}_i^*)^-\mathbf{Z}_i^{*\prime}$ is equivalent to $\mathbf{\Pi}_i = \mathbf{Z}_i\mathbf{\Psi}_i(\mathbf{\Psi}_i'\mathbf{D}_i\mathbf{\Psi}_i)^-\mathbf{\Psi}_i'\mathbf{Z}_i'$. The following results are useful in the sequel.

$$\mathbf{\Psi}_i'\mathbf{D}_i\mathbf{\Psi}_i = \mathbf{\Psi}_i'\mathbf{D}_i = \mathbf{D}_i\mathbf{\Psi}_i, \tag{1A}$$

$$\mathbf{\Psi}_i \in \{\mathbf{\Psi}_i^-\}, \quad \text{and} \quad \mathbf{\Psi}_i' \in \{(\mathbf{\Psi}_i')^-\}, \tag{2A}$$

$$(\mathbf{\Psi}_i'\mathbf{D}_i)^- = \mathbf{D}_i^-\mathbf{\Psi}_i', \quad \text{and} \quad (\mathbf{D}_i\mathbf{\Psi}_i)^- = \mathbf{\Psi}_i\mathbf{D}_i^-. \tag{3A}$$

(1A) and (2A) are trivial. A proof of (3A) is given by Takane and Hwang (2000, Appendix 2).

Using the above results, it is easily shown that

$$\mathbf{\Pi}_i = \mathbf{Z}_i\mathbf{\Psi}_i(\mathbf{\Psi}_i'\mathbf{D}_i\mathbf{\Psi}_i)^-\mathbf{\Psi}_i'\mathbf{Z}_i' = \mathbf{Z}_i\mathbf{\Psi}_i\mathbf{D}_i^-\mathbf{\Psi}_i'\mathbf{Z}_i'. \tag{4A}$$

Then,

$$\begin{aligned}\mathbf{Z}_i\mathbf{\Psi}_i\mathbf{D}_i^-\mathbf{\Psi}_i'\mathbf{Z}_i' &= \mathbf{Z}_i\mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{D}_i\mathbf{D}_i^-\mathbf{D}_i\mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{Z}_i' \\ &= \mathbf{Z}_i\mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{Z}_i' \\ &= \mathbf{Z}_i\mathbf{H}_i(\mathbf{H}_i'\mathbf{D}_i\mathbf{H}_i)^-\mathbf{H}_i'\mathbf{Z}_i' = \mathbf{Z}_i^*(\mathbf{Z}_i^{*\prime}\mathbf{Z}_i^*)^-\mathbf{Z}_i^{*\prime}.\end{aligned} \tag{5A}$$

Appendix B

The four items used in Le Roux and Rouanet (1998) from the French Worker Survey (Adam et al., 1970).

1. In professional elections in your firm, would you rather vote for a list supported by:
   (a) CGT
   (b) CFDT
   (c) Autonomous
   (d) Nonaffiliated
2. At the present time, are you affiliated to a Union, and in the affirmative, which one:
   (a) CGT
   (b) CFDT

(c) Autonomous
(d) Not affiliated

3. On the last presidential election [1969], can you tell me the candidate for whom you have voted?
   (a) Jacques Duclos
   (b) Gaston Defferre
   (c) Alain Poher
   (d) Georges Pompidou
4. Which political party do you feel closest to, as a rule?
   (a) Communist [PCF]
   (b) Socialist [SFIO + PSU + FGDS]
   (c) RI
   (d) Gaullist [UNR]

References

Adam, G., Bon, F., Capdevielle, J., & Mouriaux, R. (1970). *L'ouvrier français en 1970* [The French workman in 1970]. Paris: FNSP.

Benzécri, J.P. (1973). *L'Analyse des données*. Vol. 2. L'Analyse des correspondances. Paris: Dunod.

Benzécri, J.P. (1979). Sur le calcul des taux d'inertia dans l'analyse d'un questionaire. Addendum et erratum à. *Cahiers de L'analyse des Données, 4*, 377–378.

Böckenholt, U., & Böckenholt, I. (1990). Canonical analysis of contingency tables with linear constraints. *Psychometrika, 55*, 633–639.

Böckenholt, U., & Takane, Y. (1994). Linear constraints in correspondence analysis. In M.J. Greenacre & J. Blasius (Eds.), *Correspondence analysis in social sciences* (pp. 112–127). London: Academic Press.

Daudin, J.J. (1980). Partial association measures and an application to qualitative regression. *Biometrika, 67*, 581–590.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics, 7*, 1–26.

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

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

Lebart, L., Morineau, A., & Warwick, K.M. (1984). *Multivariate descriptive statistical analysis*. New York, NY: Wiley.

Le Roux, B., & Rouanet, H. (1998). Interpreting axes in multiple correspondence analysis: Method of the contributions of points and deviations. In M.J. Greenacre & J. Blasius (Eds.), *Visualization of categorical data* (pp. 197–220). Chestnut Hill, MA: Academic Press.

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

Nishisato, S. (1984). Forced classification: A simple application of a quantitative technique. *Psychometrika, 49*, 25–36.

Nishisato, S. (1994). *Elements of dual scaling: An introduction to practical data analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ramsay, J.O. (1978). Confidence regions for multidimensional scaling analysis. *Psychometrika, 43*, 145–160.

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

Takane, Y., & Hwang, H. (2000). Generalized constrained canonical correlation analysis. Manuscript submitted for publication.

Takane, Y., Kiers, H., & de Leeuw, J. (1995). Component analysis with different sets of constraints on different dimensions. *Psychometrika, 60*, 259–280.

Takane, Y., & Shibayama, T. (1991). Principal component analysis with external information on both subjects and variables. *Psychometrika, 56*, 97–120.

Takane, Y., Yanai, H., & Mayekawa, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika, 56*, 667–684.

ter Braak, C.J.F. (1986). Canonical correspondence analysis: A new eigenvalue technique for multivariate direct gradient analysis. *Ecology, 67*, 1167–1179.

Timm, N., & Carlson, J. (1976). Part and bipartial canonical correlation analysis. *Psychometrika, 41*, 159–176.

van Buuren, S., & de Leeuw, J. (1992). Equality constraints in multiple correspondence analysis. *Multivariate Behavioral Research, 27*, 567–583.

Yanai, H. (1986). Some generalizations of correspondence analysis in terms of projection operators. In E. Diday, Y. Escoufier, L. Lebart, J. P. Pagès, Y. Schektman, & R. Thomassone (Eds.), *Data analysis and informatics IV* (pp. 193–207). Amsterdam: North Holland.

Yanai, H. (1998). Generalized canonical correlation analysis with linear constraints. In C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H.-H. Bock, & Y. Baba (Eds.), *Data science, classification, and related methods* (pp. 539–546). Tokyo: Springer-Verlag.

Yanai, H., & Maeda, T. (2000). Partial multiple correspondence analysis. *Proceedings of the International Conference on Measurement and Multivariate Analysis, 28*, 110–113.

Manuscript received 27 MAR 2000
Final version received 12 FEB 2001