Journal of the Korean Statistical Society (2002), **31**: *2, pp 163–175*

Minimax Eccentricity Estimation for Multiple Set Factor Analysis

Hyuncheol Kang[1] and Keeyoung Kim[2]

ABSTRACT

An extended version of the minimax eccentricity factor estimation for multiple set case is proposed. In addition, two more simple methods for multiple set factor analysis exploiting the concept of generalized canonical correlation analysis is suggested. Finally, a certain connection between the generalized canonical correlation analysis and the multiple set factor analysis is derived which helps us clarify the relationship.

Keywords. Multiple set factor analysis, maximum eccentricity, generalized canonical correlation, sets of variables, internal independence.
AMS 2000 subject classifications. Primary 62H25; Secondary 62H20.

1. Introduction

Method of interbattery factor analysis, due to Tucker (1958), is to determine factors common to two sets of variables, providing unweighted least squares estimates of interbattery factor loadings. This is a factor analysis model in that a single set of unobservable factor variables accounts for all correlation coefficients between two batteries of tests. Generalizations for more than two sets have been investigated by several authors (Joreskog, 1970, 1971; McDonald, 1970; Browne, 1979, 1980). All these extensions appear to be similar to the canonical correlation analysis conceptually where there are several sets of variables.

Consider an observable p-vector $\mathbf{x}' = (\mathbf{x}_1' \cdots \mathbf{x}_g' \cdots \mathbf{x}_m')$ constituting partitioned m (≥ 2) sets of random variables, where $\mathbf{x}_g$ $(g = 1, 2, \cdots, m)$ is the g-th subvector of $\mathbf{x}$ and $p = \sum_{g=1}^{m} p_g$. We shall let inter-set common factors be denoted by $\mathbf{f}$ $(d \times 1)$, set-specific factors by $\mathbf{y}_g$ $(r_g \times 1)$, $r_g \leq p_g$, and uniqueness by

Received January 2001; accepted March 2002.
[1]Department of Informational Statistics, Hoseo University, Asan 336-795, Korea
[2]Department of Statistics, Korea University, Seoul 136-701, Korea

$\mathbf{e}_g$ ($p_g \times 1$) for the g-th set, respectively. The multiple set factor analysis (MSFA) model for the g-th set is then stated as, by setting $\mathrm{E}(\mathbf{x}) = \mathbf{0}$ without loss of generality,

$$\mathbf{x}_g = \mathbf{\Lambda}_g \mathbf{f} + \mathbf{\Gamma}_g \mathbf{y}_g + \mathbf{e}_g, \; g = 1, \cdots, m, \tag{1.1}$$

where $\mathbf{\Lambda}_g$ ($p_g \times d$) and $\mathbf{\Gamma}_g$ ($p_g \times r_g$) represent the inter-set factor loading matrix and the set-specific factor loading matrix, respectively. By denoting the factor loading matrix $\mathbf{\Lambda}' = (\mathbf{\Lambda}'_1 \cdots \mathbf{\Lambda}'_g \cdots \mathbf{\Lambda}'_m)$, the model can be rewritten as

$$\mathbf{x} = \mathbf{\Lambda} \mathbf{f} + \mathbf{u}, \tag{1.2}$$

where $\mathbf{u}_g = \mathbf{\Gamma}_g \mathbf{y}_g + \mathbf{e}_g$ ($g = 1, \cdots, m$) and $\mathbf{u}' = (\mathbf{u}'_1 \cdots \mathbf{u}'_g \cdots \mathbf{u}'_m)$.

Based on the assumptions for orthogonality that $\mathrm{var}(\mathbf{f}) = \mathbf{I}$ and $\mathrm{cov}(\mathbf{f}, \mathbf{u}') = 0$, the covariance matrix $\mathbf{\Sigma}$ of the entire set becomes

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Lambda}' + \mathbf{\Psi} \tag{1.3}$$

where $\mathrm{cov}(\mathbf{u})$ takes a block diagonal form

$$\mathbf{\Psi} = \mathrm{Diag}(\mathbf{\Psi}_1 \mathbf{\Psi}_2 \cdots \mathbf{\Psi}_m) = \begin{pmatrix} \mathbf{\Psi}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Psi}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Psi}_m \end{pmatrix} \tag{1.4}$$

where $\mathbf{\Psi}_g = \mathrm{cov}(\mathbf{u}_g)$ ($g = 1, 2, \cdots, m$).

It is worth noting that in the standard factor analysis model the block diagonal residual covariance matrix $\mathbf{\Psi}$ in (1.4) is replaced by a diagonal matrix. Also, as a special case for $m = 2$, above expressions (1.1)–(1.4) provide the same results as in the interbattery factor analysis model.

Brown (1980) has shown that the maximum likelihood estimates of $\mathbf{\Lambda}$ are, under the assumption mentioned above, the solution of the following equations

$$\tilde{\mathbf{\Lambda}}_g = \sum_{h \neq g}^{m} \mathbf{S}_{gh} \tilde{\mathbf{\Psi}}_h^{-1} \tilde{\mathbf{\Lambda}}_h \left\{ \sum_{h \neq g}^{m} \tilde{\mathbf{\Lambda}}'_h \tilde{\mathbf{\Psi}}_h^{-1} \tilde{\mathbf{\Lambda}}_h \right\}^{-1}, \; g = 1, \cdots, m, \tag{1.5}$$

where

$$\tilde{\mathbf{\Psi}}_h = \mathbf{S}_{hh} - \tilde{\mathbf{\Lambda}}_h \tilde{\mathbf{\Lambda}}'_h, \; h = 1, \cdots, m, \tag{1.6}$$

and $\mathbf{S}_{gh}$ ($p_g \times p_h$) is the sample covariance matrix of $\mathbf{x}_g$ and $\mathbf{x}_h$.

On the other hand, the maximum-determinant method is to maximize the determinant of the partial correlation matrix $\mathbf{R}_p$ of $\mathbf{x}$ given common factors $\mathbf{f}$,

$$\mathbf{R}_p = \mathbf{\Psi}^{-1/2}(\mathbf{S} - \mathbf{\Lambda\Lambda}')\mathbf{\Psi}^{-1/2}. \tag{1.7}$$

It is proved that the maximum likelihood and maximum-determinant estimators are equivalent as in the standard single set factor analysis (Browne, 1979; Kang, 1998).

In this paper, we propose a generalized version of the minimax eccentricity estimation for multiple set case. In addition, two more simple methods for multiple set factor analysis exploiting the concept of generalized canonical correlation analysis is suggested. Also, a certain connection between the generalized canonical correlation analysis and the multiple set factor analysis is derived which helps us clarify the relationship.

2. Minimax Eccentricity Estimation

In one-set factor analysis, common factors could be constructed in such a way that a single measure of dependence in a set of partial correlation coefficients given common factors is optimized. One of the measures is, for instance, $\det(\mathbf{R}_p)$ for maximum-determinant solution. Schuenemeyer and Bargmann (1978) proposed the minimax eccentricity estimation method based on the union-intersection principle. This approach provides a solution minimizing the maximum eccentricity of the correlation ellipsoid where the maximum eccentricity regarded as a dependency measure turns out to be the maximum obtainable canonical correlation coefficient between any two conceivable subvectors of a variable set.

In this section, we make an attempt to extend Schuenemeyer and Bargmann's (1978) method to the case of several sets of variables as a generalized minimax eccentricity estimation. To derive the solution, we need the following lemma.

Lemma 2.1. *Let l_k be the k-th largest eigenvalue of the partial correlation matrix $\mathbf{R}_p$ and $\mathbf{e}_k$ the corresponding unit-normed eigenvector partitioned as $\mathbf{e}'_k = (\mathbf{e}'_{k1}\mathbf{e}'_{k2} \cdots \mathbf{e}'_{km})$. Then the partial derivative of l_k with respect to $\mathbf{\Lambda}_g$ is given by*

$$\frac{\partial l_k}{\partial \mathbf{\Lambda}_g} = -2\mathbf{v}_{kg}\left(\mathbf{v}'_k\mathbf{\Lambda} - l_k\mathbf{v}'_{kg}\mathbf{\Lambda}_g\right), \; g = 1, \cdots, m, \tag{2.1}$$

where $\mathbf{v}'_k = \mathbf{e}'_k\mathbf{\Psi}^{-1/2} = (\mathbf{v}'_{k1}\mathbf{v}'_{k2} \cdots \mathbf{v}'_{km})$ and $\mathbf{v}_{kg} = \mathbf{\Psi}_g^{-1/2}\mathbf{e}_{kg}$.

Proof. Since $\mathbf{R}_p\mathbf{e}_k = l_k\mathbf{e}_k$, we have

$$(\mathbf{S} - \mathbf{\Lambda\Lambda}')\mathbf{v}_k = l_k\mathbf{\Psi}\mathbf{v}_k. \tag{2.2}$$

The partial derivative of (2.2) with respect to $\mathbf{\Lambda} = (\lambda_{ij})$, $i = 1, 2, \cdots, p$, $j = 1, 2, \cdots, d$, is given by

$$\frac{\partial(\mathbf{S} - \mathbf{\Lambda\Lambda}')}{\partial\lambda_{ij}}\mathbf{v}_k + (\mathbf{S} - \mathbf{\Lambda\Lambda}')\frac{\partial\mathbf{v}_k}{\partial\lambda_{ij}} = \frac{\partial l_k}{\partial\lambda_{ij}}\mathbf{\Psi}\mathbf{v}_k + l_k\frac{\partial\mathbf{\Psi}}{\partial\lambda_{ij}}\mathbf{v}_k + l_k\mathbf{\Psi}\frac{\partial\mathbf{v}_k}{\partial\lambda_{ij}}. \tag{2.3}$$

Pre-multiplying $\mathbf{v}_k'$ to both side of (2.3) gives us

$$\frac{\partial l_k}{\partial\lambda_{ij}} = \mathbf{v}_k'\frac{\partial(\mathbf{S} - \mathbf{\Lambda\Lambda}')}{\partial\lambda_{ij}}\mathbf{v}_k - l_k\mathbf{v}_k'\frac{\partial\mathbf{\Psi}}{\partial\lambda_{ij}}\mathbf{v}_k, \tag{2.4}$$

since $\mathbf{v}_k'\mathbf{\Psi}\mathbf{v}_k = 1$. If we let $\boldsymbol{\lambda}_j$ denote the j-th column of $\mathbf{\Lambda}$, then

$$\frac{\partial(\mathbf{S} - \mathbf{\Lambda\Lambda}')}{\partial\lambda_{ij}} = -(\mathbf{\Lambda}_{(j)} + \mathbf{\Lambda}_{(j)}'), \tag{2.5}$$

where $\mathbf{\Lambda}_{(j)} = (\mathbf{0}\cdots\boldsymbol{\lambda}_j\cdots\mathbf{0})$. And Supposing that the i-th variable is contained in the g-th set, we have

$$\frac{\partial\mathbf{\Psi}}{\partial\lambda_{ij}} = -(\mathbf{\Lambda}_{(j)g} + \mathbf{\Lambda}_{(j)g}'), \tag{2.6}$$

where

$$\mathbf{\Lambda}_{(j)g} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \boldsymbol{\lambda}_{jg} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \tag{2.7}$$

and $\boldsymbol{\lambda}_{jg}$ stands for the j-th column of $\mathbf{\Lambda}_g$. When we substitute the results of (2.5) and (2.6) into the right-hand side of (2.4), we obtain

$$\frac{\partial l_k}{\partial\lambda_{ij}} = -2v_{ki}\left(\mathbf{v}_k'\boldsymbol{\lambda}_j - l_k\mathbf{v}_{kg}'\boldsymbol{\lambda}_{jg}\right), \tag{2.8}$$

where v_{ki} is the i-th element of $\mathbf{v}_k$. The proof is completed by expressing the result of (2.8) to all cell (i, j). □

For the multiple set minimax eccentricity factor solution, the maximum eccentricity η is expressed as $\eta = (l_1 - l_p)/(l_1 + l_p)$. In order to achieve the minimization of η, we may exploit Lemma 2.1 to derive the gradient of the maximum eccentricity, which leads us to Theorem 2.1.

Theorem 2.1. *Minimax eccentricity estimates* $(\hat{\mathbf{v}},\ \hat{\mathbf{\Lambda}})$ *for* $\mathbf{v}$ *and* $\mathbf{\Lambda}$ *satisfy the following system of equations,*

$$\left(l_p \hat{\mathbf{v}}_{1g} \hat{\mathbf{v}}_1' - l_1 \hat{\mathbf{v}}_{pg} \hat{\mathbf{v}}_p'\right) \hat{\mathbf{\Lambda}} = l_1 l_p \left(\hat{\mathbf{v}}_{1g} \hat{\mathbf{v}}_{1g}' - \hat{\mathbf{v}}_{pg} \hat{\mathbf{v}}_{pg}'\right) \hat{\mathbf{\Lambda}}_g,\ g = 1, \cdots, m. \tag{2.9}$$

Proof. Since

$$\frac{\partial \eta}{\partial \mathbf{\Lambda}_g} = \frac{2}{(l_1 + l_p)^2} \left(l_p \frac{\partial l_1}{\partial \mathbf{\Lambda}_g} - l_1 \frac{\partial l_p}{\partial \mathbf{\Lambda}_g} \right), \tag{2.10}$$

by applying Lemma 2.1 to the right-hand side of (2.10), we obtain the following gradient of the maximum eccentricity η

$$\frac{\partial \eta}{\partial \mathbf{\Lambda}_g} = -\frac{4}{(l_1 + l_p)^2} \left[\left(l_p \mathbf{v}_{1g} \mathbf{v}_1' - l_1 \mathbf{v}_{pg} \mathbf{v}_p' \right) \mathbf{\Lambda} - l_1 l_p \left(\mathbf{v}_{1g} \mathbf{v}_{1g}' - \mathbf{v}_{pg} \mathbf{v}_{pg}' \right) \mathbf{\Lambda}_g \right]. \tag{2.11}$$

Hence the final results for $\hat{\mathbf{v}}$ and $\hat{\mathbf{\Lambda}}$ are setting (2.11) equal to zero. □

We propose two more approaches for multiple set factor analysis in view of its relationship with generalized canonical correlation analysis (GCCA). It is well-known that the extreme eigenvalues of a certain matrix play an important role in measuring the degree of the association among sets of variables in GCCA (Kettenring, 1971; Kang, 1998). Among others MAXVAR and MINVAR are thought as typical criteria used in GCCA. In this connection, simply minimizing or maximizing the extreme eigenvalues of $\mathbf{R}_p$ could be useful approaches for the multiple set factor analysis (MSFA) depending on the intrinsic nature of those eigenvalues. Two additional methods are MIN-MAXVAR and MAX-MINVAR. The former is to minimize l_1 of $\mathbf{R}_p$ while the latter to maximize l_p of $\mathbf{R}_p$. It is easy to see, using Lemma 2.1, that the corresponding solutions to these two methods turn out to satisfy the following relationships

$$\mathbf{v}_{kg} \mathbf{v}_k' \mathbf{\Lambda} = l_k \mathbf{v}_{kg} \mathbf{v}_{kg}' \mathbf{\Lambda}_g,\ g = 1, \cdots, m, \tag{2.12}$$

where $k = 1$ for MIN-MAXVAR and $k = p$ for MAX-MINVAR.

We have presented several equations associated with the methodologies to solve for solutions on the multiple set factor analysis. Because all those equations appear to be of the implicit form, it might be hard to compare the characteristics of each derived solution directly. However, the property of the solutions may be viewed by considering the meaning of criteria used for developing the corresponding equations. All the methods should be related to the way of achieving the overall independence after partialing out the common factors. Maximum likelihood (equivalently maximum-determinant) method refers maximizing the associated volume of the parallel-pipe (generalized variance) for which $\mathbf{R}_p$ forms to attaining the partial independence upon common factors. Meanwhile, minimax eccentricity method is to attain the partial independence by minimizing the degree of the most elongated ellipsoid for which $\mathbf{R}_p$ forms. And the characteristics of both MIN-MAXVAR and MAX-MINVAR are similar to the ones discussed in Kettenring (1971) and Kang (1998). The relationship of the methods would be further looking by the results on the data given in Section 4.

As a special case for two sets ($m = 2$), the following theorem shows that all the factor solutions provided by the methods mentioned above are equivalent.

Theorem 2.2. *For two sets of variables, maximum likelihood, minimax eccentricity, MIN-MAXVAR, and MAX-MINVAR methods are all equivalent.*

Proof. Suppose we have two sets of variables of size s and t, respectively. Then the partial correlation matrix takes the following form

$$\mathbf{R}_p = \begin{pmatrix} \mathbf{I}_s & \mathbf{A} \\ \mathbf{A}' & \mathbf{I}_t \end{pmatrix}, \tag{2.13}$$

where $\mathbf{A} = \mathbf{\Psi}_1^{-1/2}(\mathbf{S}_{12} - \mathbf{\Lambda}_1\mathbf{\Lambda}_2')\mathbf{\Psi}_2^{-1/2}$ and $\mathbf{I}_s$ stands for the identity matrix of size s ($\geq t$). Let $l_1 \geq l_2 \geq \cdots \geq l_{s+t}$ be the eigenvalues of $\mathbf{R}_p$ and $\mathbf{e}_i' = (\mathbf{e}_{i1}' \ \mathbf{e}_{i2}')$ the unit-normed eigenvector corresponding to l_i. That is, l_i and $\mathbf{e}_i$ satisfy following equations

$$\begin{cases} \mathbf{e}_{i1}' + \mathbf{e}_{i2}'\mathbf{A}' = l_i\mathbf{e}_{i1}', \\ \mathbf{e}_{i1}'\mathbf{A} + \mathbf{e}_{i2}' = l_i\mathbf{e}_{i2}'. \end{cases} \tag{2.14}$$

By setting the determinant equal to zero,

$$|\mathbf{R}_p - l\mathbf{I}| = \begin{vmatrix} (1-l)\mathbf{I}_s & \mathbf{A} \\ \mathbf{A}' & (1-l)\mathbf{I}_t \end{vmatrix} = (1-l)^{s-t}\left|(1-l)^2\mathbf{I}_t - \mathbf{A}'\mathbf{A}\right|, \tag{2.15}$$

the eigenvalues of $\mathbf{R}_p$ becomes

$$\begin{cases} l_i = 1 + l^{(i)} & , i = 1, 2, \cdots, t, \\ l_{t+1} = \cdots = l_s = 1, & \\ l_i = 1 - l^{(s+t-i+1)} & , i = s+1, \cdots, s+t, \end{cases} \tag{2.16}$$

where $l^{(i)}$ is the i-th singular value of $\mathbf{A}$. Therefore, since $(l_1, \cdots, l_{s+t}) = (l_1, \cdots, l_t, 1, \cdots, 1, 2 - l_t, \cdots, 2 - l_1)$, minimax eccentricity, MIN-MAXVAR and MAX-MINVAR methods are to optimize the same criterion l_1, causing the relevant solutions being equivalent. On the other hand, maximum likelihood estimates satisfy from the equation (1.5) the following equation

$$\begin{cases} (\mathbf{S}_{12} - \tilde{\mathbf{\Lambda}}_1\tilde{\mathbf{\Lambda}}_2')\tilde{\mathbf{\Psi}}_2^{-1}\tilde{\mathbf{\Lambda}}_2 = \mathbf{0}, \\ (\mathbf{S}_{21} - \tilde{\mathbf{\Lambda}}_2\tilde{\mathbf{\Lambda}}_1')\tilde{\mathbf{\Psi}}_1^{-1}\tilde{\mathbf{\Lambda}}_1 = \mathbf{0}. \end{cases} \tag{2.17}$$

Let $\tilde{\mathbf{e}}_i$ and $\tilde{l}_i$ be the i-th eigenvector and eigenvalue of the matrix $\mathbf{R}_p$ with $\mathbf{A}$ replaced by $\tilde{\mathbf{A}} = \tilde{\mathbf{\Psi}}_1^{-1/2}(\mathbf{S}_{12} - \tilde{\mathbf{\Lambda}}_1\tilde{\mathbf{\Lambda}}_2')\tilde{\mathbf{\Psi}}_2^{-1/2}$. Now pre-multiplying $\tilde{\mathbf{e}}_{11}\tilde{\mathbf{\Psi}}_1^{-1/2}$ to the first equation and $\tilde{\mathbf{e}}_{12}\tilde{\mathbf{\Psi}}_2^{-1/2}$ to the second equation in (2.17) yields, since $\tilde{\mathbf{e}}_{11}'\tilde{\mathbf{A}} = (\tilde{l}_1 - 1)\tilde{\mathbf{e}}_{12}'$ and $\tilde{\mathbf{e}}_{12}'\tilde{\mathbf{A}}' = (\tilde{l}_1 - 1)\tilde{\mathbf{e}}_{11}'$ from (2.14),

$$\begin{cases} (\tilde{l}_1 - 1)\tilde{\mathbf{e}}_{12}'\tilde{\mathbf{\Psi}}_2^{-1/2}\tilde{\mathbf{\Lambda}}_2 = \mathbf{0}, \\ (\tilde{l}_1 - 1)\tilde{\mathbf{e}}_{11}'\tilde{\mathbf{\Psi}}_1^{-1/2}\tilde{\mathbf{\Lambda}}_1 = \mathbf{0}. \end{cases} \tag{2.18}$$

Thus, if $\tilde{l}_1 \neq 1$, since $\tilde{\mathbf{v}}_{11}'\tilde{\mathbf{\Lambda}}_1 = \tilde{\mathbf{v}}_{12}'\tilde{\mathbf{\Lambda}}_2 = \mathbf{0}$ from (2.18), maximum likelihood estimates satisfy the equation (2.9). Otherwise $\tilde{l}_i = 1$ for all i, thus all the solutions are equivalent. Therefore, the proof is completed. □

To find the solution of minimax eccentricity, MIN-MAXVAR and MAX-MINVAR, we use the (dual) quasi-Newton optimization techniques. It is worth noting that when all sets contain only one variable *i.e.*, $p_1 = \cdots = p_m = 1$, all the expressions given in this chapter provide the same results as in the standard single set factor analysis.

3. An Interpretation of the Relationship between MSFA and GCCA

For the case of two sets ($m = 2$), Browne (1979) have shown that maximum likelihood solution of the first set in factor analysis can be obtained by an eigenanalysis of the matrix

$$\mathbf{S}_{11}^{-1/2}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1/2}. \tag{3.1}$$

As a matter of fact, the eigenvalues and associated eigenvectors of (3.1) are directly related to the two-set canonical correlation analysis. In terms of factor model, the solution for the first set is given by

$$\hat{\mathbf{\Lambda}}_1 = \mathbf{S}_{11}^{1/2}\mathbf{U}_1\mathbf{D}_r^{1/2}, \tag{3.2}$$

where $\mathbf{D}_r$ is a diagonal matrix with the k largest eigenvalues of the matrix (3.1) and $\mathbf{U}_1$ $(p_1 \times k)$ consists of the associated eigenvectors of unit norm. The solution for the second set is obtained simply by exchanging the corresponding subscript from 1 to 2. That is, for the second set, the formula for solution is given by

$$\hat{\mathbf{\Lambda}}_2 = \mathbf{S}_{22}^{1/2}\mathbf{U}_2\mathbf{D}_r^{1/2}, \tag{3.3}$$

where $\mathbf{U}_2$ $(p_2 \times k)$ consists of the unit normed eigenvectors associated the k largest eigenvalues of the matrix

$$\mathbf{S}_{22}^{-1/2}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1/2}. \tag{3.4}$$

Browne (1979) further discussed the relationship between interbattery factor analysis and canonical correlation analysis from the expression (3.3) and (3.4) for the case of two sets. It can be seen that the matrix

$$\mathbf{S}_{gg}^{-1/2}\mathbf{U}_g = \mathbf{S}_{gg}^{-1}\hat{\mathbf{\Lambda}}_g\mathbf{D}_r^{-1/2}, \; g = 1, 2, \tag{3.5}$$

actually corresponds to the canonical correlation loading matrix for each set g and the k-th element of $\mathbf{D}_r$ is the k-th canonical correlation coefficient. This idea was extended to more general case by linking up the concept of generalized canonical coefficients (Browne, 1980). However, it will be explained that this type of relationship could be more specific when introducing the concept of common factor scores appeared in the multiple set factor analysis model (1.2). It would be helpful to interpret the relationship between GCCA and MSFA.

Suppose we are interested in obtaining estimated values of the factor scores for the multiple set factor model with d common factors. Let us consider the following model for estimating the matrix of factor scores $\mathbf{F}$ $(n \times d)$ by using only the g-th set of variables,

$$\mathbf{F} = \mathbf{X}_g\mathbf{B}_g + \boldsymbol{\delta}_g, \tag{3.6}$$

where $\mathbf{X}_g$ is a $(n \times p_g)$ data matrix of the g-th set, $\mathbf{B}_g$ is a $(p_g \times d)$ matrix of factor scoring coefficients and $\boldsymbol{\delta}_g$ the additive error vector. When $\mathbf{F}$ is known, $\mathbf{B}_g$ could be estimated by the least squares method such as

$$\hat{\mathbf{B}}_g = (\mathbf{X}_g'\mathbf{X}_g)^{-1}\mathbf{X}_g'\mathbf{F}. \tag{3.7}$$

However, in the situation that $\mathbf{F}$ is unknown, we may consider replacing $\mathbf{X}_g'\mathbf{F}$ in (3.7) by the usual factor loading matrix $\hat{\mathbf{\Lambda}}_g$ to estimate $\mathbf{B}_g$. Namely,

$$\hat{\mathbf{B}}_g = \mathbf{S}_{gg}^{-1}\hat{\mathbf{\Lambda}}_g. \tag{3.8}$$

Denoting $\hat{\mathbf{b}}_{sg}$ and $\hat{\lambda}_{sg}$ by the s-th column of $\hat{\mathbf{B}}_g$ and $\hat{\Lambda}_g$ respectively, we have $\hat{\mathbf{b}}'_{sg}\mathbf{S}_{gg}\hat{\mathbf{b}}_{sg} = \hat{\lambda}'_{sg}\mathbf{S}_{gg}^{-1}\hat{\lambda}_{sg}$. And the standardized version of the estimated regression weight $\hat{\mathbf{B}}_g^*$ of $\hat{\mathbf{B}}_g$ can be written as

$$\hat{\mathbf{B}}_g^* = \mathbf{S}_{gg}^{-1}\hat{\Lambda}_g\mathbf{D}_g^{-1/2}, \tag{3.9}$$

where $\mathbf{D}_g$ is a diagonal matrix with $\hat{\lambda}'_{sg}\mathbf{S}_{gg}^{-1}\hat{\lambda}_{sg}$ as its s-th element. That is, $\hat{\mathbf{B}}_g^*$ is rescaled so as to satisfy $\hat{\mathbf{f}}'_s\hat{\mathbf{f}}_s = \hat{\mathbf{b}}^{*\prime}_{sg}\mathbf{S}_{gg}\hat{\mathbf{b}}^*_{gs} = 1$ $(s = 1, \cdots, k)$, where $\hat{\mathbf{f}}_s$ stands for the s-th column of the estimated factor score matrix $\hat{\mathbf{F}}$.

For the case of two sets of variables considered in the MSFA model, it is interesting to see that the standardized regression weights

$$\hat{\mathbf{B}}_g^* = \mathbf{S}_{gg}^{-1}\hat{\Lambda}_g\mathbf{D}_r^{-1/2},\ g = 1, 2, \tag{3.10}$$

becomes the canonical loading matrices as mentioned in (3.5), since $\hat{\Lambda}'_g\mathbf{S}_{gg}^{-1}\hat{\Lambda}_g = \mathbf{D}_r$ $(g = 1, 2)$. This relationship implies that multiple set factor loading matrices $\hat{\Lambda}_g$ are obtained by rescaling the canonical loading matrices $\mathbf{S}_{gg}^{-1/2}\mathbf{U}_g$ $(g = 1, 2)$ or vice-versa, when there are two sets of variables. There is, however, a substantial difference in terms of the interpretation between canonical loadings and multiple set factor loadings in the following sense. In canonical correlation analysis, the matrices of canonical scores $\mathbf{Z}_g$ are calculated by the formula $\mathbf{Z}_g = \mathbf{X}_g\hat{\mathbf{B}}_g$ $(g = 1, 2)$. We note that $\hat{\mathbf{B}}_g$ are the regression weights corresponding to the observable values $\mathbf{X}_g$. Thus $\mathbf{Z}_g$ is a set of scores which are calculable. Furthermore $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are correlated each other and not necessarily identical. On the other hand, $\hat{\Lambda}_g$ $(g = 1, 2)$ can be regarded as the regression weights of either one of $\mathbf{X}_1$ or $\mathbf{X}_2$ on the common factor $\mathbf{F}$. But here $\mathbf{F}$ is presumably unobservable in multiple set factor analysis.

As we have seen earlier, the maximum likelihood or minimax eccentricity solutions can be expressed explicitly for two sets of variables, causing no iteration process in calculation. And the direct relationship between the usual canonical correlation analysis and two-set factor analysis enables us to derive the solutions to one analysis from the solutions of another. But for more than two sets of variables, the relationship in GCCA and MSFA is to be an implicit form, being complicated to visualize it directly, and the desired solutions require certain iteration process. Also we may not enjoy the exchangeability as in the two-set case. However, if we follow the arguments appeared in (3.6)–(3.9), a similar type of multiset generalization could be proceeded so as to regard $\hat{\mathbf{B}}_g^*$ in (3.9) as canonical loadings.

4. Numerical Illustration

We provide an illustrative example using Horst's (1961) "ability data" which has been used by McDonald (1970), Kettenring (1971) and Browne (1980). The data consist of scores on three sets of tests and in each set there are three tests designed to measure verbal, numerical, and spatial ability ($m = 3$, $p_1 = p_2 = p_3 = 3$). To obtain the solutions to three estimation methods (minimax eccentricity, MIN-MAXVAR and MAX-MINVAR) suggested in Section 2, from the correlation matrix of ability data, we have used NLPQN subroutine in SAS/IML (SAS Institute Inc., 1996) which exploits (dual) quasi-Newton optimization techniques. But for the maximum likelihood estimation, Gauss-Seidel algorithm is applied to (1.5). Three factors are extracted for each of four estimation methods.

TABLE 4.1 *Factor loadings and eigenvalues of partial correlation matrix*

	MIN-MAXECC			*MAX-LIKELI*		
	1	2	3	1	2	3
	.626	−.438	−.104	.626	−.441	−.090
Λ_1	.602	.315	−.455	.597	.312	−.477
	.667	.313	.439	.672	.311	.436
	.565	−.649	−.000	.561	−.651	−.000
Λ_2	.592	.477	−.318	.584	.469	−.319
	.561	.283	.347	.564	.293	.336
	.610	−.565	−.006	.608	−.559	−.013
Λ_3	.698	.206	−.110	.698	.206	−.117
	.629	.235	.244	.632	.237	.228
l_1		1.0883			1.0843	
l_p		0.9298			0.9140	

	MAX-MINVAR			*MIN-MAXVAR*		
	1	2	3	1	2	3
	.627	−.432	−.092	.628	−.442	−.109
Λ_1	.589	.304	−.441	.604	.317	−.449
	.661	.308	.431	.670	.315	.434
	.553	−.641	−.003	.567	−.651	−.000
Λ_2	.590	.481	−.324	.593	.476	−.325
	.557	.281	.343	.565	.286	.351
	.619	−.564	−.008	.606	−.562	−.009
Λ_3	.697	.207	−.102	.701	.207	−.112
	.637	.227	.239	.628	.240	.250
l_1		1.0550			1.1104	
l_p		0.8900			0.9440	

TABLE 4.2 *Residual matrix for minimax eccentricity solution*

Set	I			II			III		
I	.405	−.038	.036	−.003	−.012	−.007	−.004	.010	.013
	−.038	.330	.097	.001	.001	−.007	−.003	−.009	.013
	.036	.097	.263	.006	.001	−.003	−.002	−.012	.008
II	−.003	.001	.006	.258	.065	.013	−.003	−.007	−.012
	−.012	.001	.001	.065	.320	−.061	.008	−.007	−.014
	−.007	−.007	−.003	.013	−.061	.483	−.002	.024	−.009
III	−.004	−.003	−.002	−.003	.008	−.002	.308	−.019	−.005
	.010	−.009	−.012	−.007	−.007	.024	−.019	.456	−.033
	.013	.013	.008	−.012	−.014	−.009	−.005	−.033	.488

TABLE 4.3 *Regression weights for minimax eccentricity solution*

$\hat{\mathbf{B}}_1^*$			$\hat{\mathbf{B}}_2^*$			$\hat{\mathbf{B}}_3^*$		
.498	−.885	−.157	.552	−.832	−.055	.456	−.985	−.057
.379	.474	−.882	.517	.535	−.818	.497	.402	−.757
.485	.503	.927	.422	.287	.869	.395	.391	1.055

All solutions obtained satisfy the absolute function convergence criterion of value 10^{-10}. We have chosen maximum likelihood solution derived from the standard factor analysis as an initial parameter matrix. The number of iterations required for convergence are 30, 60, 41 and 54 for minimax eccentricity (MIN-MAXECC), maximum likelihood (MAX-LIKELI), MIN-MAXVAR and MAX-MINVAR, respectively. The resulted factor loadings of four methods are presented in Table 4.1. It is interesting to observe that most of the factor loadings in minimax eccentricity lie between those of the remaining three methods. And the same is for the eigenvalues as well. This appears to be frequently observed phenomenon in MSFA as well as GCCA. With this experience we may say that minimax eccentricity solutions tend to take intermediate position among most of the possible methods in general. Maximum likelihood method provides similar results to those of minimax eccentricity among other methods.

The residual matrix, $\mathbf{R} - \hat{\Lambda}\hat{\Lambda}'$ after extraction the common factors using minimax eccentricity method is given in Table 4.2. It is observed that most of the off-block diagonal elements in the residual matrix of Table 4.2 are well within a bound around zero, showing that the corresponding fit looks good (the residual matrices of the other three methods show similar pattern with that of minimax eccentricity method, although we do not provide those here).

The matrix of standardized regression weights for factor scoring coefficient, $\hat{\mathbf{B}}_g^*$ in (3.9), are presented in Table 4.3 when we consider minimax eccentricity. The first stage between set correlation matrix, $\hat{\mathbf{b}}_{1i}^{*\prime}\mathbf{R}_{ij}\hat{\mathbf{b}}_{1j}^*$ $(i, j = 1, 2, 3)$, is calculated as follows

$$\begin{pmatrix} 1 & .736 & .754 \\ .736 & 1 & .743 \\ .754 & .743 & 1 \end{pmatrix},$$

where $\hat{\mathbf{b}}_{1g}^*$ is the first column of $\hat{\mathbf{B}}_g^*$. These elements could be regarded as the generalized version of canonical correlation coefficients.

REFERENCES

Browne, M. W. (1979). "The maximum likelihood solution in inter-battery factor analysis", *The British Journal of Mathematical and Statistical Psychology*, **32**, 75–86.

Browne, M. W. (1980). "Factor analysis of multiple batteries by maximum likelihood", *The British Journal of Mathematical and Statistical Psychology*, **33**, 184–199.

Horst, P. (1961). "Relations among *m* sets of measures", *Psychometrika*, **26**, 129–149.

Joreskog, K. (1970). "A general method for analysis of covariance structures", *Biometrika*, **57**, 239–251.

Joreskog, K. (1971). "Simultaneous factor analysis in several populations", *Psychometrika*, **36**, 409–426.

Kang, H. (1998). *A Study of Multivariate Structural Relationships for Sets of Variables*, Ph.D. Dissertation, Korea University, Seoul.

Kettenring, J. R. (1971). "Canonical analysis of several sets of variables", *Biometrika*, **58**, 433–451.

McDonald, R. P. (1970). "Three common factor models for groups of variables", *Psychometrika*, **35**, 111–128.

SAS Institute Inc. (1996). *SAS/IML Software : Changes and Enhancements through Release 6.11*, Cary, NC.

Schuenemeyer, J. H. and Bargmann R. E. (1978). "Maximum eccentricity as a union-intersection test statistic in multivariate analysis", *Journal of Multivariate Analysis*, **8**, 268–273.

Tuker, L. R. (1958). "An inter-battery method of factor analysis", *Psychometrika*. **23**. 111–136.