The Basic Space Model

Let x_{ij} be the ith individual's (i=1, ..., n) reported position on the jth issue (j = 1, ..., m) and let $\mathbf{X}_0$ be the n by m matrix of observed data where the "0" subscript indicates that elements are missing from the matrix -- not all individuals report their positions on all issues. Let ψ_{ik} be the ith individual's position on the kth (k = 1, ..., s) basic dimension. The model estimated is:

$$\mathbf{X}_0 = [\mathbf{\Psi W'} + \mathbf{J}_n\underline{\mathbf{c}}']_0 + \mathbf{E}_0 \qquad (1A)$$

where $\mathbf{\Psi}$ is the n by s matrix of coordinates of the individuals on the basic dimensions, $\mathbf{W}$ is an m by s matrix of weights, $\underline{\mathbf{c}}$ is a vector of constants of length m, $\mathbf{J}_n$ is an n length vector of ones, and $\mathbf{E}_0$ is a n by m matrix of error terms. $\mathbf{W}$ and $\underline{\mathbf{c}}$ map the individuals from the basic space onto the issue dimensions.

Equation (1A) can be written as the product of partitioned matrices

$$\mathbf{X_0} = \left[\mathbf{\Psi} \mid \mathbf{J_n}\right]\left[\frac{\mathbf{W'}}{\underline{\mathbf{c}}'}\right]_{\mathbf{0}} + \mathbf{E_0} \qquad (1B)$$

where $\left[\mathbf{\Psi} \mid \mathbf{J_n}\right]$ is a n by s+1 matrix and $\left[W|\underline{c}\right]$ is a m by s+1 matrix. If n > m and there is no error or missing data, then the rank of $\mathbf{X}$ is s and the rank of $\mathbf{X} - \mathbf{J_n}\underline{\mathbf{c}}'$ is less than or equal to s.

No Missing Data

To solve (1) when there is no missing data, set $\underline{\mathbf{c}}$ equal to the column means of $\mathbf{X}$; that is

$$\mathbf{c_j} = \frac{\sum_{i=1}^{n} \mathbf{x_{ij}}}{\mathbf{n}} = \bar{\mathbf{x}}_{\mathbf{j}}$$

and perform a singular value decomposition of $\mathbf{X} - \mathbf{J_n}\underline{\mathbf{c}}'$:

$$\mathbf{X} - \mathbf{J_n}\underline{\mathbf{c}}' = \mathbf{U}\mathbf{\Lambda}\mathbf{V}' = \mathbf{\Psi}\mathbf{W}'$$

where $\mathbf{U}$ is an n by m matrix, $\mathbf{\Lambda}$ is a m by m matrix, and $\mathbf{V}$ is a m by m matrix.

A simple solution for $\mathbf{\Psi}$ and $\mathbf{W}$ is

$$\begin{aligned} \mathbf{\Psi} &= \mathbf{U}\,\Lambda^{\frac{1}{2}} \\ \mathbf{W} &= \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}} \end{aligned} \qquad (2)$$

where the diagonal elements of $\Lambda^{\frac{1}{2}}$ are the square roots of $\mathbf{\Lambda}$. Let $\mathbf{I_m}$ be the m by m identity matrix. Equation (2) implies that $\mathbf{\Psi}'\mathbf{\Psi} = \mathbf{W}'\mathbf{W}$. That is:

$$\mathbf{\Psi}'\mathbf{\Psi} = \Lambda^{\frac{1}{2}}\mathbf{U}'\mathbf{U}\,\Lambda^{\frac{1}{2}} = \Lambda^{\frac{1}{2}}\mathbf{I_m}\,\Lambda^{\frac{1}{2}} = \mathbf{\Lambda}$$

and

$$\mathbf{W}'\mathbf{W} = \Lambda^{\frac{1}{2}}\mathbf{V}'\mathbf{V}\,\Lambda^{\frac{1}{2}} = \Lambda^{\frac{1}{2}}\mathbf{I_m}\,\Lambda^{\frac{1}{2}} = \mathbf{\Lambda}$$

In addition, by construction, $\mathbf{J_n}'[\mathbf{X} - \mathbf{J_n}\underline{\mathbf{c}}'] = \underline{\mathbf{0}}'$, so that $\mathbf{J_n}'\mathbf{U} = \mathbf{J_n}'\mathbf{\Psi} = \underline{\mathbf{0}}'$, where $\underline{\mathbf{0}}$ is a m length vector of zeros.

When an s < m is preferred, the Eckart-Young Theorem may be used in (2) to arrive at solutions for $\mathbf{\Psi}$ and $\mathbf{W}$. That is, the s + 1 to m singular values are set equal to zero so that $\mathbf{\Psi}$ and $\mathbf{W}$ from (2) are n by s and m by s matrices respectively.

Missing Data

Because of the presence of missing data, SVD and the Eckart-Young Theorem cannot be used directly. Instead, I work with the loss function

$$\xi = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\{[\sum_{k=1}^{s}\psi_{ik}w_{jk}] + c_j - x_{ij}\}^2 \qquad (3)$$

The notation m_i means that the total of the summation over j may vary from s + 1 to m depending on how many entries there are in the i[th] row of $\mathbf{X}_0$. That is, each individual must report at least s + 1 issue positions in order to be identified. Furthermore, the number of missing entries in the columns of $\mathbf{X}_0$ must also be restricted. In most practical applications n will be much larger than m. Consequently, I will adopt the convention that there must be at least 2m entries in each column of $\mathbf{X}_0$.

In line with the discussion above, the following two restrictions are applied to the loss function:

$$\mathbf{\Psi}'\mathbf{\Psi} = \mathbf{W}'\mathbf{W} \quad \text{and} \quad \mathbf{J_n}'\mathbf{\Psi} = \underline{\mathbf{0}}'$$

These restrictions produce the Lagrangean multiplier problem

$$\mu = \xi + 2\gamma'[\,\mathbf{\Psi}'\mathbf{J_n}] + \mathrm{tr}[\mathbf{\Phi}(\mathbf{\Psi}'\mathbf{\Psi} - \mathbf{W}'\mathbf{W})] \tag{4}$$

where $\underline{\gamma}$ is an s length vector of Lagrangean multipliers and $\mathbf{\Phi}$ is a symmetric s by s matrix of Langrangean multipliers.

Given that the Lagrangean multipliers are all zero, the partial derivatives of $\mathbf{\Psi}$, $\mathbf{W}$, and $\underline{\mathbf{c}}$ from equations (3) and (4) are identical. In particular:

$$\frac{\partial \mu}{\partial \psi_{ik}} = 2\sum_{j=1}^{m_i}\left[\left(\sum_{\ell=1}^{s} w_{j\ell}\psi_{i\ell}\right) + c_j - x_{ij}\right] w_{jk} \tag{5A}$$

$$\frac{\partial \mu}{\partial w_{jk}} = 2\sum_{i=1}^{n_j}\left[\left(\sum_{\ell=1}^{s} w_{j\ell}\psi_{i\ell}\right) + c_j - x_{ij}\right] \psi_{ik} \tag{5B}$$

$$\frac{\partial \mu}{\partial c_j} = 2\sum_{i=1}^{n_j}\left[\left(\sum_{\ell=1}^{s} w_{j\ell}\psi_{i\ell}\right) + c_j - x_{ij}\right] \tag{5C}$$

where n_j means that the total of the summation over i may vary from 2m to n depending upon how many entries there are in the ith column of $\mathbf{X}_0$.

Setting (5A) to zero and collecting the s partial derivatives of the ith row of $\mathbf{\Psi}$ into a vector and dividing by 2 produces

$$[\mathbf{W^{*\prime}W^*}]\underline{\boldsymbol{\psi}}_{\mathbf{i}} - \mathbf{W^{*\prime}}[\underline{\mathbf{x}}_{oi} - \underline{\mathbf{c}}_{o}] = \underline{\mathbf{0}}$$

where $\mathbf{W^*}$ is an m_i by s matrix with the appropriate rows corresponding to missing entries in $\mathbf{X}_o$ removed, $\underline{\boldsymbol{\psi}}_{\mathbf{i}}$ is the ith row of $\mathbf{\Psi}$, $\underline{\mathbf{x}}_{oi}$ is the ith row of $\mathbf{X}_0$ and is of length m_i , $\underline{\mathbf{c}}_o$ is the m_i length vector of constants corresponding to the elements of $\underline{\mathbf{x}}_{oi}$, and $\underline{\mathbf{0}}$ is an s length vector of zeroes.

If $\mathbf{W^{*\prime}W^*}$ is nonsingular, then

$$\hat{\underline{\boldsymbol{\psi}}}_{\mathbf{i}} = (\mathbf{W^{*\prime}W^*})^{-1}\mathbf{W^{*\prime}}[\underline{\mathbf{x}}_{oi} - \underline{\mathbf{c}}_{o}] \tag{6}$$

and the rows of $\mathbf{\Psi}$ can be estimated through ordinary least squares.

The s partial derivatives of the jth row of $\mathbf{W}$ from equation (5B) and the partial derivative for $\underline{\mathbf{c}}_j$ from (5C) can be collected into the vector

$$\left[\mathbf{\Psi}_j^{*\prime}\mathbf{\Psi}_j^{*}\right]\begin{bmatrix}\underline{\mathbf{w}}_j \\ \hline \mathbf{c}_j\end{bmatrix} - \mathbf{\Psi}_j^{*\prime}\underline{\mathbf{x}}_{\mathbf{oj}} = \underline{\mathbf{0}}$$

where $\mathbf{\Psi_j^* = [\Psi_o|J_o]}$ is an n_j by s + 1 matrix (the matrix $\mathbf{\Psi}$ with the appropriate rows corresponding to missing data removed and then bordered by ones), $\underline{\mathbf{w}}_j$ is the s length vector of the jth row elements of $\mathbf{W}$, c_j is the jth element of $\underline{\mathbf{c}}$, $\underline{\mathbf{x}}_{oj}$ is the jth column of $\mathbf{X}_o$ and is of length n_j, and $\underline{\mathbf{0}}$ is an s+1 length vector of zeroes.

If $\mathbf{\Psi_j^{*\prime}\Psi_j^*}$ is nonsingular, then

$$\begin{bmatrix}\hat{\underline{\mathbf{w}}}_j \\ \hline \hat{\mathbf{c}}_j\end{bmatrix} = (\mathbf{\Psi}_j^{*\prime}\mathbf{\Psi}_j^{*})^{-1}\mathbf{\Psi}_j^{*\prime}\underline{\mathbf{x}}_{\mathbf{oj}} \tag{7}$$

and the rows of $\mathbf{W}$ and the elements of $\underline{\mathbf{c}}$ can be estimated through ordinary least squares.

The easiest way to estimate $\mathbf{W}$ and $\boldsymbol{\Psi}$ is to select some suitable starting estimate of either matrix and then iterate between (6) and (7) until convergence is achieved. The constraints on $\mathbf{W}$ and $\boldsymbol{\Psi}$ can be met at any stage of the iteration by simply setting the column means of $\hat{\boldsymbol{\Psi}}$ equal to zero, forming the matrix product $\hat{\boldsymbol{\Psi}}\hat{\mathbf{W}}'$, and performing the singular value decomposition:

$$\hat{\boldsymbol{\Psi}}\hat{\mathbf{W}}' = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}'$$

where $\boldsymbol{\Lambda}$ is an s by s diagonal matrix containing the s singular values in descending order, and $\mathbf{U}$ and $\mathbf{V}$ are n by s and s by s matrices respectively such that $\mathbf{U}'\mathbf{U} = \mathbf{V}'\mathbf{V} = \mathbf{I}_s$. Setting $\hat{\boldsymbol{\Psi}} = \mathbf{U}\,\Lambda^{\frac{1}{2}}$ and $\hat{\mathbf{W}} = \mathbf{V}\Lambda^{\frac{1}{2}}$ as in (2) satisfies the constraints.

A simple way to proceed with the estimation is to exploit the orthogonality of $\boldsymbol{\Psi}$ and estimate one column of $\boldsymbol{\Psi}$ and $\mathbf{W}$ at a time. This is motivated by the fact that if the n_j are close to n, $\boldsymbol{\Psi}_j^{*\prime}\boldsymbol{\Psi}_j^*$ in (7) will be very close to a diagonal matrix.

Table 1

Summary of the Estimation Procedure

1) Obtain starting estimates of $\hat{\underline{c}}$, denoted by $\hat{\underline{c}}^{(1)}$, using the column means of $\mathbf{X_0}$. Obtain starting estimates of $\hat{\underline{w}}_1$, denoted by $\hat{\underline{w}}_1^{(1)}$, by finding the vector of plus and minus ones that maximizes the number of positive elements in the covariance matrix $[\mathbf{X_o} - \mathbf{J_n}\hat{\underline{c}}']'[\mathbf{X_o} - \mathbf{J_n}\hat{\underline{c}}']$ (see Appendix B).

2) Use $\hat{\underline{c}}^{(1)}$ and $\hat{\underline{w}}_1^{(1)}$ in equation (8) to obtain a starting estimate of $\hat{\underline{\psi}}_1$, denoted by $\hat{\underline{\psi}}_1^{(1)}$, and set the mean of $\hat{\underline{\psi}}_1^{(1)}$ equal to zero.

3) Use $\hat{\underline{\psi}}_1^{(1)}$ in equation (7) to obtain a second estimate of $\hat{\underline{c}}$ and $\hat{\underline{w}}_1$ -- $\hat{\underline{c}}^{(2)}$ and $\hat{\underline{w}}_1^{(2)}$ respectively.

4) Use $\hat{\underline{c}}^{(2)}$ and $\hat{\underline{w}}_1^{(2)}$ in equation (6) to obtain a second estimate of $\hat{\underline{\psi}}_1$, $\hat{\underline{\psi}}_1^{(2)}$. Set the mean of $\hat{\underline{\psi}}_1^{(2)}$ equal to zero and set the sum of squares of $\hat{\underline{\psi}}_1^{(2)}$ equal to the sum of squares of $\hat{\underline{\psi}}_1^{(1)}$; that is $\sum_{i=1}^{n} \hat{\psi}_{i1}^{(2)^2} = \sum_{i=1}^{n} \hat{\psi}_{i1}^{(1)^2}$.

5) Repeat steps (3) and (4) until convergence.

6) Compute $\mathbf{E_{01} = X_0 - \underline{\hat{\psi}}_1 \underline{\hat{w}}'_1 - J_n \underline{\hat{c}}'}$.

7) Obtain starting estimates of $\hat{\underline{w}}_2$, $\hat{\underline{w}}_2^{(1)}$, by finding the vector of plus and minus ones that maximizes the number of positive elements in the covariance matrix $E_{01}^{'} E_{01}$.

8) Use $\hat{\underline{w}}_2^{(1)}$ in equation (10) to obtain starting estimates of $\hat{\underline{\psi}}_2$, $\hat{\underline{\psi}}_2^{(1)}$.

9) Use $\underline{\hat{\psi}}_2^{(1)}$ in equation (11) to obtain $\underline{\hat{w}}_2^{(2)}$.

10) Use $\underline{\hat{w}}_2^{(2)}$ in equation (12) to obtain $\underline{\hat{\psi}}_2^{(2)}$. Set the mean of $\underline{\hat{\psi}}_2^{(2)}$ equal to zero and set the sum of squares of $\underline{\hat{\psi}}_2^{(2)}$ equal to the sum of squares of $\underline{\hat{\psi}}_{\mathbf{2}}^{\mathbf{(1)}}$ as in step (4) above.

11) Repeat steps (9) and (10) until convergence.

12) Compute $\mathbf{E_{02} = X_0 - \underline{\hat{\psi}}_1 \underline{\hat{w}}'_1 - J_n \underline{\hat{c}}' - \underline{\hat{\psi}}_2 \underline{\hat{w}}'_2 = E_{01} - \underline{\hat{\psi}}_2 \underline{\hat{w}}'_2}$.

13) Repeat steps (7) - (12) to estimate remaining dimensions; that is: $\underline{\hat{w}}_3$ and $\underline{\hat{\psi}}_3$, $\underline{\hat{w}}_4$ and $\underline{\hat{\psi}}_4$, ... , and $\underline{\hat{w}}_s$ and $\underline{\hat{\psi}}_s$.

14) Use the full n by s matrix $\hat{\mathbf{\Psi}}$ in equation (7) to obtain the full m by s matrix $\hat{W}$ and the m length vector of constants $\underline{\hat{c}}$.

15) Use $\hat{W}$ and $\underline{\hat{c}}$ in equation (6) to obtain a new estimate of $\hat{\mathbf{\Psi}}$.

16) Repeat steps (14) and (15) until convergence.