RM-76-60

THE MULTIPLE INDICATOR - MULTIPLE CAUSE MODEL WITH SEVERAL LATENT VARIABLES*

Johannes Ledolter**

Bovas Abraham***

July 1976

*Revised paper originally presented at the 3rd World Congress of the Econometric Society, August 21-26, Toronto, Canada.

**Currently at IIASA, Laxenburg, Austria.

***Department of Mathematics, Dalhousie University, N.S., Canada

Preface

Latent variables, though not observed, are considerably useful in explaining relationships among observable variables and are frequently used in econometrics and psychometrics. This paper discusses the general multiple indicator - multiple cause model with several latent variables.

The Multiple Indicator - Multiple Cause Model With Several Latent Variables

Abstract

A model in which one observes multiple indicators and multiple causes of several latent variables is considered. The parameters of this model are estimated by maximum likelihood and restricted rank regression approaches. Also a likelihood ratio test statistic for testing the validity of the restrictions in the above model is derived.

1. Introduction

Latent variables, though not observed, are useful in explaining relationships among observable variables. Jöreskog and Goldberger (1975) utilize maximum likelihood and various other procedures for estimation of a model in which one observes multiple indicators and multiple causes of a single latent variable. In this paper we extend their analysis to cover the case of several latent variables.

In its most general form, our model is specified as follows: The structural equations are

$$\begin{cases} \underset{\sim}{y} = B\underset{\sim}{y}^* + \Gamma\underset{\sim}{z} + \underset{\sim}{u} \\ \underset{\sim}{y}^* = A'\underset{\sim}{x} + \underset{\sim}{\varepsilon} \end{cases} \tag{1.1}$$

where

$\underset{\sim}{y} = (y_1,\ldots,y_m)'$		observable endogenous indicators,
$\underset{\sim}{x} = (x_1,\ldots,x_k)'$		observable exogenous causes of latent variables,
$\underset{\sim}{z} = (z_1,\ldots,z_s)'$		observable additional exogenous variables directly affecting the indicators,

$$\underset{\sim}{y}^* = (y_1^*,\dots,y_r^*)' \qquad \text{latent variables,}$$

and

$$\underset{\sim}{u} = (u_1,\dots,u_m)' \qquad \underset{\sim}{\varepsilon} = (\varepsilon_1,\dots,\varepsilon_r)'$$

are the disturbances (error variables). The coefficient matrices are:

$$\underset{[m\times r]}{B} = \{\beta_{ij}\} \quad , \quad \underset{[m\times s]}{\Gamma} = \{\gamma_{ij}\} \quad , \quad \underset{[k\times r]}{A} = \{\alpha_{ij}\} \quad .$$

We make the following assumptions about the disturbances:

$$E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \Delta \quad , \quad E(\underset{\sim}{u}\underset{\sim}{u}') = \Theta^2 \text{ diagonal} \quad , \quad E(\underset{\sim}{\varepsilon}\underset{\sim}{u}') = 0 \quad .$$

The diagonal elements of Θ^2 are displayed in the vector

$$\underset{\sim}{\theta} = (\theta_1,\dots,\theta_m)' \quad .$$

Hence the reduced form of the model is given by

$$\begin{aligned} \underset{\sim}{y} &= B(A'\underset{\sim}{x} + \underset{\sim}{\varepsilon}) + \Gamma\underset{\sim}{z} + \underset{\sim}{u} \\ &= \Pi'\underset{\sim}{x} + \Gamma\underset{\sim}{z} + \underset{\sim}{v} \quad , \end{aligned} \tag{1.2}$$

where

$$\Pi' = BA' \quad \text{and} \quad \underset{\sim}{v} = B\underset{\sim}{\varepsilon} + \underset{\sim}{u} \quad .$$

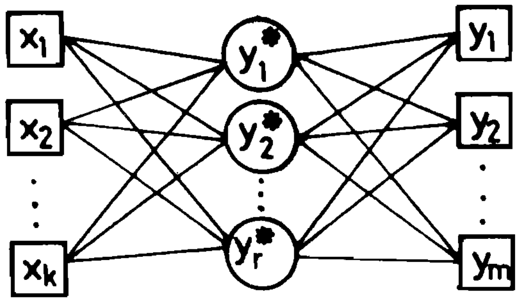
Then

$$E(\underset{\sim}{v}\underset{\sim}{v}') = B\Delta B' + \Theta^2 = \Omega \quad .$$

Several special cases of this general model have already been discussed in the literature: Zellner (1970) considers the generalized and modified least square estimation of a model with one latent variable ($r=1$), two observable endogenous indicators ($m=2$) and no observable exogenous variables directly affecting the indicators ($\Gamma=0$). Furthermore he assumes that the exogenous causes are not subject to stochastic errors. With $m > 2$, Goldberger (1972) and Jöreskog and Goldberger (1975) discuss the maximum likelihood (ML) estimation of a model where $\Gamma = 0$, $r = 1$ and the exogenous causes are subject to stochastic errors. Hauser (1972) extends the analysis to cover the situation where additional causal variables directly affect the indicators ($\Gamma \neq 0$). With Θ^2 not restricted to be diagonal, Hauser discusses the case of one latent variable ($r=1$). He shows that the observable additional exogenous variables directly affecting the indicators can be swept out by replacing $\underset{\sim}{y}$ and $\underset{\sim}{x}$ by the residuals from the regression of $\underset{\sim}{y}$ on $\underset{\sim}{z}$ and $\underset{\sim}{x}$ on $\underset{\sim}{z}$ respectively. Robinson (1974) covers the case of several latent variables ($r > 1$) with Θ^2 not restricted to be diagonal.

Since the reduced form of the general model remains unchanged when B is postmultiplied by a nonsingular $r \times r$ matrix H and A' is premultiplied by H^{-1}, we have an indeterminancy of the structural parameters. To remove this indeterminancy we adopt the normalization $\Delta = I$. Furthermore, following Hauser (1972) and Robinson (1974), we see that if Γ is unrestricted, the additional causal variables which directly affect the indicators can be reduced by sweeping out. Hence for unrestricted Γ there is no loss of generality in dropping $\underset{\sim}{z}$ from the general model.

A graphical interpretation of the model is given in the figure below. Furthermore we illustrate the model by way of two conceptual examples:



Example 1: Relationship between social status and social participation (Jöreskog and Goldberger (1975)).
$y_1, y_2, \ldots, y_m$ (church attendance, membership, friends seen,...) are viewed as indicators of a latent variable y* (social participation) which is linearly determined by the observable exogenous causes $x_1, x_2, \ldots, x_k$ measuring social status (income, occupation, education,...).

Example 2: Relationship between income distribution and hunger. In this example hunger is viewed as latent variable and $y_1, y_2, \ldots, y_m$ are the indicators of this latent variable, such as occurrence of malnutrition diseases, protein and calorie intake below standard levels. The latent variable y* is linearly determined by the observable exogenous causes $x_1, x_2, \ldots, x_k$ reflecting the income distribution of a country (percentage of people below average income, percentage of farms below average farm size,...).

2. Specification of the model

Thus, our reduced specification is

$$\begin{cases} \underset{\sim}{y} = B\underset{\sim}{y}^* + \underset{\sim}{u} \\ \underset{\sim}{y}^* = A'\underset{\sim}{x} + \underset{\sim}{\varepsilon} \end{cases} \tag{2.1}$$

where

$$\underset{[m\times r]}{B} = (\underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2, \ldots, \underset{\sim}{\beta}_r) \quad , \quad \underset{\sim}{\beta}_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{mj})' \quad ,$$

$$\underset{[k\times r]}{A} = (\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r) \quad , \quad \underset{\sim}{\alpha}_j = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{kj})' \quad ,$$

with

$$E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = I \quad , \quad E(\underset{\sim}{\varepsilon}\underset{\sim}{u}') = O \quad , \quad E(\underset{\sim}{u}\underset{\sim}{u}') = \Theta^2 \text{ (diagonal)} \quad .$$

Hence the reduced form is

$$\underset{\sim}{y} = \Pi'\underset{\sim}{x} + \underset{\sim}{v} \quad , \tag{2.2}$$

where

$$\Pi' = BA' \quad , \quad \underset{\sim}{v} = B\underset{\sim}{\varepsilon} + \underset{\sim}{u} \quad , \quad E(\underset{\sim}{v}\underset{\sim}{v}') = BB' + \Theta^2 = \Omega \ .$$

We notice two kinds of restrictions on the parameters of the reduced form:

(i) The matrix Π has rank r; the km elements of Π are expressed in terms of $r(k+m)$ elements of A and B. This is the type of restriction one encounters in the reduced form of conventional simultaneous equation models.

(ii) The m(m+1)/2 distinct elements of Ω are expressed in terms of the m(r+1) elements of B and Θ^2. This is the type of restriction which arises in conventional factor analysis models. It is also to be noted that the same matrix B appears both in Π and Ω.

We observe that there is an infinity of choices for B: The reduced form $\Pi' = BA'$ and $\Omega = BB' + \Theta^2$ will remain unchanged if we replace $\underset{\sim}{y}^*$ by $M\underset{\sim}{y}^*$, B by BM', and A by AM', where M is any orthogonal matrix. In the terminology of factor analysis this corresponds to a rotation of factors. Hence, following Lawley and Maxwell (1971), without loss of generality we choose B such that $B'\Theta^{-2}B = G^{-2}$, say, is diagonal. Then

$$\begin{aligned} \Omega^{-1} &= \Theta^{-2} - \Theta^{-2}B(I+G^{-2})^{-1}B'\Theta^{-2} \quad , \\ \Omega^{-1}B &= \Theta^{-2}B(I+G^{-2})^{-1} \quad , \qquad (2.3) \\ B'\Omega^{-1}B &= G^{-2}(I+G^{-2})^{-1} \quad . \end{aligned}$$

We consider two alternative specifications concerning the stochastic nature of $\underset{\sim}{x}$. In case 1, $\underset{\sim}{x}$ is taken as fixed and $\underset{\sim}{y}$ has a multivariate normal distribution, whereas in case 2, $(\underset{\sim}{x},\underset{\sim}{y})$ are jointly multivariate normal. In both cases successive observations are assumed to be independent.

3. Maximum Likelihood Estimation

Now consider a sample of T joint observations $\underset{\sim}{x}(t)$, $\underset{\sim}{y}(t)$ generated for t = 1,2,...,T by

$$\underset{\sim}{y}(t) = \Pi'\underset{\sim}{x}(t) + \underset{\sim}{v}(t) \quad , \quad \underset{\sim}{v}(t) \text{ are NID}(0,\Omega) \quad , \quad \underset{\sim}{x}(t) \text{ fixed.}$$

Here $\Pi' = BA'$, $\Omega = BB' + \Theta^2$. The log likelihood L_1 of the sample can be written as

$$L_1 = -\frac{1}{2}T[\log|\Omega| + \text{tr}(\Omega^{-1}W)] \tag{3.1}$$

where

$$W = (Y - X\Pi)'(Y-X\Pi)$$

$$X = T^{-\frac{1}{2}}(\underset{\sim}{x}(1),\ldots,\underset{\sim}{x}(T))'$$

$$Y = T^{-\frac{1}{2}}(\underset{\sim}{y}(1),\ldots,\underset{\sim}{y}(T))' \quad .$$

We define the usual multivariate regression statistics:

$$P = (X'X)^{-1}X'Y \quad , \quad Q = Y'XP \quad , \quad S = (Y - XP)'(Y - XP) \quad ,$$
$$R = Y'Y = S + Q \quad . \tag{3.2}$$

To maximize the likelihood, it suffices to minimize

$$F = (\frac{-2}{T})\ L_1 = \log|\Omega| + \text{tr}(\Omega^{-1}W) \quad . \tag{3.3}$$

The general formula for derivatives of a function of the form of F, as given in Jöreskog and Goldberger (1975), is

$$\frac{\partial F}{\partial \mu_i} = \text{tr}[\Omega^{-1}(\Omega - W)\Omega^{-1}\ \frac{\partial \Omega}{\partial \mu_i}] + \text{tr}[\Omega^{-1}\ \frac{\partial W}{\partial \mu_i}] \quad ,$$

where μ_i denotes any of the elements in A, B and $\underset{\sim}{\theta}$.

We now define $\underset{\sim}{d}_i$ and $\underset{\sim}{e}_i$ respectively to be $k \times 1$ and $m \times 1$ vectors with 1 at the i^{th} position and zeros at all other positions. Then

$$\frac{\partial\Omega}{\partial\theta_\ell} = 2\theta_\ell \underset{\sim}{e}_\ell \underset{\sim}{e}_\ell' \quad ; \qquad \frac{\partial W}{\partial\theta_\ell} = 0 \qquad \ell = 1,2,\ldots,m$$

$$\frac{\partial\Omega}{\partial\alpha_{ij}} = 0 \quad ; \quad \frac{\partial W}{\partial\alpha_{ij}} = -\underset{\sim}{\beta}_j \underset{\sim}{d}_i' X'Y - Y'X \underset{\sim}{d}_i \underset{\sim}{\beta}_j' + 2\underset{\sim}{d}_i' X'X\underset{\sim}{\alpha}_j \underset{\sim}{\beta}_j \underset{\sim}{\beta}_j'$$

$$+ \sum_{s\neq j}^{r} \{\underset{\sim}{\beta}_s \underset{\sim}{\alpha}_s' X'X\underset{\sim}{d}_i \underset{\sim}{\beta}_j' + \underset{\sim}{\beta}_j \underset{\sim}{d}_i' X'X\underset{\sim}{\alpha}_s \underset{\sim}{\beta}_s'\}$$

$$i = 1,2,\ldots,k \quad ; \quad j = 1,2,\ldots,r$$

$$\frac{\partial\Omega}{\partial\beta_{\ell j}} = \underset{\sim}{e}_\ell \underset{\sim}{\beta}_j' + \underset{\sim}{\beta}_j \underset{\sim}{e}_\ell'$$

$$\frac{\partial W}{\partial\beta_{\ell j}} = -\underset{\sim}{e}_\ell \underset{\sim}{\alpha}_j' X'Y - Y'X\underset{\sim}{\alpha}_j \underset{\sim}{e}_\ell' + \underset{\sim}{\alpha}_j' X'X\underset{\sim}{\alpha}_j (\underset{\sim}{\beta}_j \underset{\sim}{e}_\ell' + \underset{\sim}{e}_\ell \underset{\sim}{\beta}_j')$$

$$+ \sum_{s\neq j}^{r} \{\underset{\sim}{e}_\ell \underset{\sim}{\alpha}_j' X'X\underset{\sim}{\alpha}_s \underset{\sim}{\beta}_s' + \underset{\sim}{\beta}_s \underset{\sim}{\alpha}_s' X'X\underset{\sim}{\alpha}_j \underset{\sim}{e}_\ell'\}$$

$$\ell = 1,2,\ldots,m \quad ; \quad j = 1,2,\ldots,r \quad . \tag{3.4}$$

Hence we have

$$\frac{\partial F}{\partial\alpha_{ij}} = tr(\Omega^{-1}\frac{\partial W}{\partial\alpha_{ij}}) = 2\{-\underset{\sim}{\beta}_j'\Omega^{-1}Y'X + \underset{\sim}{\beta}_j'\Omega^{-1}\underset{\sim}{\beta}_j \underset{\sim}{\alpha}_j' X'X + \sum_{s\neq j}^{r} \underset{\sim}{\beta}_j'\Omega^{-1}\underset{\sim}{\beta}_s \underset{\sim}{\alpha}_s' X'X\}\underset{\sim}{d}_i$$

$$i = 1,2,\ldots,k \quad ; \quad j = 1,2,\ldots,r \quad .$$

Setting the derivatives $\partial F/\partial\alpha_{ij}$ equal to zero, we obtain the maximum likelihood (ML) estimate of A in terms of the ML estimates of B and $\underset{\sim}{\theta}$. The solutions we obtain are implicit and thus must be iterative.

The ML estimate of A is given by

$$\hat{A} = P\hat{\Omega}^{-1}\hat{B}(\hat{B}'\hat{\Omega}^{-1}\hat{B})^{-1} \quad . \tag{3.5}$$

Here and in the following carets denote ML estimates so that $\hat{\Omega} = \hat{\Theta}^2 + \hat{B}\hat{B}'$. We recall that B was chosen such that $B'\Theta^{-2}B = G^{-2} = \text{diag}\ (1/g_1^2,\ldots,1/g_r^2)$, say. Then

$$B'\Omega^{-1}B = G^{-2}(I + G^{-2})^{-1} = \text{diag}\left(\frac{1}{1+g_1^2},\ldots,\frac{1}{1+g_r^2}\right) \quad . \tag{3.6}$$

Hence using (3.6) in (3.5) we obtain

$$\hat{\underset{\sim}{\alpha}}_j = (1 + \hat{g}_j^2)P\hat{\Omega}^{-1}\hat{\underset{\sim}{\beta}}_j \quad . \qquad (j = 1,\ldots,r) \tag{3.7}$$

Now

$$\frac{\partial F}{\partial \beta_{\ell j}} = \text{tr}\left[\Omega^{-1}(\Omega - W)\Omega^{-1}\frac{\partial\Omega}{\partial\beta_{\ell j}}\right] + \text{tr}\left[\Omega^{-1}\frac{\partial W}{\partial\beta_{\ell j}}\right]$$

$$= 2\left[\underset{\sim}{\beta}_j'(\Omega^{-1}-\Omega^{-1}W\Omega^{-1}) - \underset{\sim}{\alpha}_j'X'Y\Omega^{-1} + \underset{\sim}{\alpha}_j'X'X\underset{\sim}{\alpha}_j\underset{\sim}{\beta}_j'\Omega^{-1} + \sum_{s\neq j}^{r}\underset{\sim}{\alpha}_j'X'X\underset{\sim}{\alpha}_s\underset{\sim}{\beta}_s'\Omega^{-1}\right]\underset{\sim}{e}_\ell$$

$$(\ell = 1,2,\ldots,m \quad ; \quad j = 1,2,\ldots,r)$$

We set the derivatives $\partial F/\partial\beta_{\ell j} = 0$ to obtain the ML estimates of B. The resulting equation turns out to be

$$\hat{\Omega}^{-1}[\hat{B}(I+\hat{C}) - (Y'X\hat{A} + \hat{W}\hat{\Omega}^{-1}\hat{B})] = 0 \quad , \tag{3.8}$$

where C is the ML estimate of

$$C = A'X'XA = \{\underset{\sim}{\alpha}_i'X'X\underset{\sim}{\alpha}_j\} = \{c_{ij}\} \quad . \tag{3.9}$$

It can be easily verified that $\hat{W}\hat{\Omega}^{-1}\hat{B} = S\hat{\Omega}^{-1}\hat{B}$ and using this and (3.5) in equation (3.8) we obtain

$$S\hat{\Omega}^{-1}\hat{B} + Q\hat{\Omega}^{-1}\hat{B}(\hat{B}'\hat{\Omega}^{-1}\hat{B}) = \hat{B}(I+\hat{C}) \quad . \qquad (3.10)$$

The above equation can be rewritten as

$$[\{S+(1+\hat{g}_1^2)Q\}\hat{\Omega}^{-1}\hat{\beta}_1,\ldots,\{S+(1+\hat{g}_r^2)Q\}\hat{\Omega}^{-1}\hat{\beta}_r] = \hat{B}(I+\hat{C}) \quad . \quad (3.11$$

For the ML estimate of θ^2, the derivatives $\partial F/\partial\theta_\ell$, $(\ell = 1,2,\ldots,m)$ are needed and these are given by

$$\frac{\partial F}{\partial\theta_\ell} = 2\theta_\ell e_\ell'\Omega^{-1}(\Omega - W)\Omega^{-1}e_\ell \quad , \qquad (\ell = 1,2,\ldots,m) \quad . \quad (3.12)$$

Also making use of (3.5) and (3.10) we obtain

$$\hat{\Omega}^{-1}(\hat{\Omega}-\hat{W})\hat{\Omega}^{-1} = \hat{\Theta}^{-2}[\hat{\Theta}^2 - R + \hat{B}(I+\hat{C})\hat{B}']\hat{\Theta}^{-2} \quad . \qquad (3.13)$$

Finally the ML estimate of θ^2 is obtained by setting $\partial F/\partial\theta_\ell = 0$ and using (3.13) in the resulting equation yields

$$\hat{\theta}_\ell^2 = [R]_{\ell\ell} - [\hat{B}(I+\hat{C})\hat{B}']_{\ell\ell} \quad , \qquad (\ell = 1,2,\ldots,m) \quad (3.14)$$

where $[M]_{\ell\ell}$ denotes the ℓ^{th} diagonal element of the matrix M.

Thus we have the implicit solutions for the ML estimates of A, B and θ given by (3.5), (3.11) and (3.14).

If we further assume that C = A'X'XA is a diagonal matrix, then the ML estimates of B and θ^2 are given respectively by

$$[S+(1+\hat{g}_j^2)Q]\hat{\Omega}^{-1}\hat{\beta}_j = (1+\hat{c}_{jj})\hat{\beta}_j \qquad (j = 1,\ldots,r) \quad (3.15)$$

$$\hat{\theta}_\ell^2 = [R]_{\ell\ell} - \sum_{s=1}^{r}(1+\hat{c}_{ss})\hat{\beta}_{\ell s}^2 \qquad (\ell = 1,\ldots,m) \quad (3.16)$$

Under the above assumption the ML estimate of $\underset{\sim}{\beta}_j$ is a characteristic vector of the matrix on the left of (3.15), normalized such that $B'\Theta^{-2}B$ is diagonal. It can also be shown that this characteristic vector is in fact the one corresponding to the largest root. The assumption that C is diagonal can be interpreted as requiring the latent variables to be uncorrelated.

When $\underset{\sim}{x}$ is random, say normal with mean zero and dispersion matrix Φ, the log likelihood of the sample is the sum of two parts. This is because the joint distribution of $\underset{\sim}{y}$ and $\underset{\sim}{x}$ is the product of the conditional distribution of $\underset{\sim}{y}$ given $\underset{\sim}{x}$ and the marginal distribution of $\underset{\sim}{x}$. Since the joint distribution is multivariate normal both the marginal and conditional distributions are multivariate normal. Hence the log likelihood is given by $L = L_1 + L_2$ where L_1 is given in (3.1) and $L_2 = -\frac{1}{2}T[\log|\Phi| + tr(X'X\Phi^{-1})]$. Then the ML estimate of Φ is $\hat{\Phi} = X'X$, and the ML estimates of the remaining parameters are unaffected.

Our results in (3.7), (3.15) and (3.16) are analogous to those obtained by Jöreskog and Goldberger (1975) for the single latent variable case (r = 1).

4. Testing for the Validity of Restrictions

We now derive an explicit expression for the likelihood-ratio test of the model.

First we evaluate the function in (3.3) at the ML estimates $\hat{A}$, $\hat{B}$ and $\hat{\Theta}^2$ to obtain $\hat{F}$, the minimum of F. We have

$$\hat{F} = \log|\hat{\Omega}| + \operatorname{tr}(\hat{\Omega}^{-1}\hat{W}) \quad . \tag{4.1}$$

Now

$$|\hat{\Omega}| = |\hat{\Theta}^2||I + \hat{B}'\hat{\Theta}^{-2}\hat{B}| = |\hat{\Theta}^2||I + \hat{G}^{-2}| \quad ,$$

so that

$$\log|\hat{\Omega}| = \sum_{i=1}^{m} \log \hat{\theta}_i^2 + \sum_{j=1}^{r} \log\left(\frac{1+\hat{g}_j^2}{\hat{g}_j^2}\right) \quad . \tag{4.2}$$

From (3.1) and (3.2) it can be seen that

$$\operatorname{tr}(\hat{\Omega}^{-1}\hat{W}) = \operatorname{tr}(\hat{\Omega}^{-1}R) - \operatorname{tr}(\hat{A}'X'Y\hat{\Omega}^{-1}\hat{B}) - \operatorname{tr}(\hat{B}'\hat{\Omega}^{-1}Y'X\hat{A}) + \operatorname{tr}(\hat{B}'\hat{\Omega}^{-1}\hat{B}\hat{A}'X'X\hat{A}) \quad . \tag{4.3}$$

We can also show that

$$\operatorname{tr}(\hat{A}'X'Y\hat{\Omega}^{-1}\hat{B}) = \operatorname{tr}(\hat{B}'\hat{\Omega}^{-1}Y'X\hat{A}) = \sum_{j=1}^{r} (\hat{\underset{\sim}{\alpha}}_j'X'Y\hat{\Omega}^{-1}\hat{\underset{\sim}{\beta}}_j) \quad , \tag{4.4}$$

and using (2.3)

$$\operatorname{tr}(\hat{B}'\hat{\Omega}^{-1}\hat{B}\hat{A}'X'X\hat{A}) = \sum_{j=1}^{r} (1/(1+\hat{g}_j^2))\hat{\underset{\sim}{\alpha}}_j'X'X\hat{\underset{\sim}{\alpha}}_j \tag{4.5}$$

From (3.7) it follows that

$$\hat{c}_{jj} = \hat{\underset{\sim}{\alpha}}_j'X'X\hat{\underset{\sim}{\alpha}}_j = (1+\hat{g}_j^2)\hat{\underset{\sim}{\alpha}}_jX'Y\hat{\Omega}^{-1}\hat{\underset{\sim}{\beta}}_j \quad , \quad j = 1,2,\ldots,r \quad . \tag{4.6}$$

Hence from (4.3)-(4.6) we obtain

$$\operatorname{tr}(\hat{\Omega}^{-1}\hat{W}) = \operatorname{tr}(\hat{\Omega}^{-1}R) - \sum_{j=1}^{r} (\hat{c}_{jj}/(1+\hat{g}_j^2)) \quad . \tag{4.7}$$

Now using (2.3), (3.2) and (3.6) we obtain

$$\operatorname{tr}(\hat{\Omega}^{-1}R) = \operatorname{tr}(\hat{\Theta}^{-2}R) - \sum_{j=1}^{r} (\hat{g}_j^2/(1+\hat{g}_j^2))\underset{\sim}{\beta}_j^{*\prime}(S^* + Q^*)\underset{\sim}{\beta}_j^{*} \quad , \tag{4.8}$$

where

$$S^* = \hat{\Theta}^{-1}S\hat{\Theta}^{-1} \quad , \quad Q^* = \hat{\Theta}^{-1}Q\hat{\Theta}^{-1} \quad \text{and} \quad B^* = \hat{\Theta}^{-1}\hat{B} \quad .$$

From (3.7) it can be shown that $\hat{\underset{\sim}{\alpha}}_j = \hat{g}_j^2 P\hat{\Theta}^{-2}\hat{\underset{\sim}{\beta}}_j$ $(j = 1,2,\ldots,r)$ and hence that

$$\underset{\sim}{\beta}_j^{*\prime} Q^* \underset{\sim}{\beta}_j^* = \hat{c}_{jj}/\hat{g}_j^2 \quad , \qquad j = 1,2,\ldots,r \quad . \tag{4.9}$$

Using (2.3) in equation (3.10) and after tedious algebraic manipulations we obtain

$$\sum_{j=1}^{r} (\hat{g}_j^2/(1+\hat{g}_j^2))\underset{\sim}{\beta}_j^{*\prime} S^* \underset{\sim}{\beta}_j^* = \sum_{j=1}^{r} (1/\hat{g}_j^2) \quad . \tag{4.10}$$

Therefore application of (4.8), (4.9) and (4.10) in (4.7) yields

$$\mathrm{tr}(\hat{\Omega}^{-1}\hat{W}) = \mathrm{tr}(\hat{\Theta}^{-2}R) - \sum_{j=1}^{r} ((1+\hat{c}_{jj})/\hat{g}_j^2) \quad . \tag{4.11}$$

It can, however, be shown from (3.16) that

$$\mathrm{tr}(\hat{\Theta}^{-2}R) = m + \sum_{j=1}^{r} (1+\hat{c}_{jj})/\hat{g}_j^2 \quad , \tag{4.12}$$

and hence (4.11) simplifies to

$$\mathrm{tr}(\hat{\Omega}^{-1}\hat{W}) = m \quad . \tag{4.13}$$

Therefore,

$$\hat{F} = m + \sum_{i=1}^{m} \log \hat{\theta}_i^2 + \sum_{j=1}^{r} \log\left(\frac{1+\hat{g}_j^2}{\hat{g}_j^2}\right) \quad . \tag{4.14}$$

Under the alternative hypothesis, Π and Ω are unconstrained, and are estimated by P and S respectively. The minimum value of

F is

$$F_0 = \log|S| + m \quad . \tag{4.15}$$

Hence the likelihood ratio test statistic is given by

$$\chi^2 = -2 \log \text{(likelihood ratio)} = T(\hat{F} - F_0)$$

$$= T\left[\sum_{i=1}^{m} \log \hat{\theta}_i^2 + \sum_{j=1}^{r} \log\left(\frac{1+\hat{g}_j^2}{\hat{g}_j^2}\right) - \log|S|\right] \quad . \tag{4.16}$$

Under the null hypothesis that the restrictions are valid, this statistic is asymptotically distributed as a Chi-square with the degrees of freedom equal to the number of restrictions, namely

$$\frac{m[m - (2r + 1)]}{2} + k(m - r) \quad .$$

It is to be noted that the above test is valid only when

(i) $m > r$

(ii) $\frac{m(m + 2k - 1)}{2(m + k)} > r$.

The condition (i) implies that the number of observable endogenous indicators is larger than the number of latent variables.

5. Restricted Rank Regression Approach

In this section we ignore restrictions on Ω and make use of restrictions on Π only. This is the model analyzed by Robinson (1974). The system is not identified but identification can be achieved by making the normalization $C = A'X'XA = I$. In other words we are getting estimates of $A^* = AC^{-\frac{1}{2}}$ and $B^* = BC^{\frac{1}{2}}$ where

$C^{\frac{1}{2}} = P_1'\Lambda^{\frac{1}{2}}P_1$; the orthogonal matrix P_1 and the diagonal matrix Λ of eigenvalues of $A'X'XA$ are such that $C = P_1'\Lambda P_1$.

The estimates of A^* and B^* can be derived by a "limited information maximum likelihood" analysis of P, the unrestricted estimator of Π. This is achieved by minimizing $F = \log|\Omega| + \mathrm{tr}(\Omega^{-1}W)$ subject to $\Pi = A^*B^{*\prime}$ with the normalization $A^{*\prime}X'XA^* = I$. In this case, the minimum distance principle (which minimizes $\mathrm{tr}(S^{-1}W)$) produces the same coefficient estimates as the maximum likelihood principle; see Goldberger (1970), Robinson (1974). These estimates, denoted by bars, are given by the equations:

$$\bar{B}^*(\bar{B}^{*\prime}S^{-1}\bar{B}^*) = QS^{-1}\bar{B}^* \quad , \quad \bar{A}^* = PS^{-1}\bar{B}^*(\bar{B}^{*\prime}S^{-1}\bar{B}^*)^{-1} \quad . \quad (5.1)$$

Using the normalization

$$\bar{B}^{*\prime}S^{-1}\bar{B}^* = \mathrm{diag}\left(\frac{1}{1+\bar{g}_1^2}, \ldots, \frac{1}{1+\bar{g}_r^2}\right) \quad , \quad (5.2)$$

we get from (5.1) that

$$QS^{-1}\bar{\beta}_j^* = \left(\frac{1}{1+\bar{g}_j^2}\right)\bar{\beta}_j^* \quad , \quad (j = 1,2,\ldots,r) \quad (5.3)$$

and

$$\bar{\alpha}_j^* = (1+\bar{g}_j^2)PS^{-1}\bar{\beta}_j^* \quad , \quad (j = 1,2,\ldots,r) \quad . \quad (5.4)$$

Here we see that the $\bar{\beta}_j^*$ are the characteristic vectors of QS^{-1}, and in fact it is easy to show that they correspond to the r largest characteristic roots of this matrix. The estimates obtained here are, apart from normalization, the same as those obtained by Robinson (1974). If we reintroduce the constraints on Θ^2, the estimates of Θ^2 can be obtained by undertaking a common factor analysis of $\bar{S} = (Y-X\bar{\Pi})'(Y-X\bar{\Pi})$.

6. Concluding Remarks

In this paper we studied a model with r latent variables, combining restrictions which occur in econometrics and psychometrics. We developed estimates using ML and restricted rank regression approaches. The solutions obtained were implicit and one needs to have some iterative scheme for the implementation of these solutions. Fortunately, the present model fits into Jöreskog's (1970) covariance structure model, for which the ML algorithm is already programmed.

Jöreskog (1970) develops a general covariance structure model for a multivariate normal vector $\underset{\sim}{z}$ with

$$E(\underset{\sim}{z}\underset{\sim}{z}') = D(\Lambda\Phi\Lambda' + \psi^2)D' + \Gamma^2 \quad .$$

Elements of the parameter matrices D, Λ, Φ (symmetric) and ψ, Γ (diagonal) may be fixed, constrained, or free. Taking $\underset{\sim}{z} = (\underset{\sim}{x}', \underset{\sim}{y}')'$ we have in the random case

$$E(\underset{\sim}{z}\underset{\sim}{z}') = \begin{pmatrix} \Phi & \Phi AB' \\ BA'\Phi & B(I+C)B'+\Theta^2 \end{pmatrix}$$

We choose

$$\underset{[(k+m)\times(k+r)]}{D} = \begin{pmatrix} I_{k\times k} & O_{k\times r} \\ O_{m\times k} & B_{m\times r} \end{pmatrix}, \qquad \underset{[(k+r)\times k]}{\Lambda} = \begin{pmatrix} I_{k\times k} \\ A'_{r\times k} \end{pmatrix},$$

$$\underset{[(k+r)\times(k+r)]}{\psi} = \begin{pmatrix} O_{k\times k} & O_{k\times r} \\ O_{r\times k} & I_{r\times r} \end{pmatrix}, \qquad \underset{[(m+k)\times(m+k)]}{\Gamma} = \begin{pmatrix} O_{k\times k} & O_{k\times m} \\ O_{m\times k} & \Theta_{m\times m} \end{pmatrix}.$$

Then the covariance structure of $\underset{\sim}{z}$ is specified in terms of Jöreskog's model.

Acknowledgement

We wish to thank A.S. Goldberger for many helpful comments.

References

Anderson, T.W. (1951), Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions, *Annals of Mathematical Statistics*, 22, 3, 327-351.

Goldberger, A.S. (1970), Criteria and Constraints in Multivariate Regression, *Social Systems Research Institute Workshop Paper 7026*, University of Wisconsin, Madison, Wisconsin.

Goldberger, A.S. (1972), Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables, *International Economic Review*, 13, 1, 1-15.

Hauser, R.M. and Goldberger, A.S. (1971), The Treatment of Unobservable Variables in Path Analysis, Chapter 4 in H.L. Costner, editor, *Sociological Methodology 1971*, San Francisco: Jossey-Bass.

Hauser, R.M. (1972), Disaggregating a Social-Psychological Model of Educational Attainment, *Social Science Research*, 1, 2, 159-188.

Jöreskog, K.G. (1970), A General Method for Analysis of Covariance Structures, *Biometrika*, 57, 231-251.

Jöreskog, K.G. and Goldberger, A.S. (1975), Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable, *Journal of the American Statistical Association*, 70, 351, 631-639.

Lawley, D.N. and Maxwell, E.E. (1971), *Factor Analysis as a Statistical Method*, New York: American Elsevier.

Robinson, P.M. (1974), Identification, Estimation and Large-Sample Theory for Regressions Containing Unobservable Variables, *International Economic Review*, 15, 3, 680-692.

Zellner, A. (1970), Estimation of Regression Relationships Containing Unobservable Variables, *International Economic Review*, 11, 441-454.