

An exact algebraic solution for the recursive path model with manifest variables

PETER H. SCHÖNEMANN
Purdue University, Lafayette, Indiana

It is shown that the classical recursive path "model" $By = Ax + Uz$, where $\text{Cov}(x,z) = 0$ and B is lower triangular, has a very simple algebraic solution that is essentially given by the chol-eski factor of the partial variance-covariance matrix, $\text{Var}(y|x)$. This solution always exists, and it is unique provided only that $\text{Var}(x)$ and $\text{Var}(y|x)$ are nonsingular. It may be useful for exploratory work and also as a starting configuration to speed up convergence of iterative estimation schemes.

NOTATION

Let $y' = (y_1, \dots, y_p)$ a vector of p (≥ 1) observed "dependent" (or "endogenous") real-valued random variables with $p \times p$ variance-covariance matrix $\text{Var}(y)$, $x' = (x_1, \dots, x_m)$ a vector of m (≥ 1) observed "independent" (or "exogenous") real-valued random variables with nonsingular $m \times m$ variance-covariance matrix $\text{Var}(x)$. The covariance matrix between the dependent and independent variables will be denoted $\text{Cov}(y,x)$.

It is customary to call the structural relations

$$By = Ax + Uz \quad (1.1)$$

a "recursive path analysis model" (see, e.g., Li, 1975, and Wright, 1918, 1921) if

$$\text{Var}(z) = I_p, \quad (1.2a)$$

$$\text{Cov}(x,z) = 0, \quad (1.2b)$$

$$B \text{ is lower triangular with } \text{diag}(B) = \text{diag}(I), \quad (1.2c)$$

$$U = \text{diagonal, positive definite.} \quad (1.2d)$$

The lower triangular matrix B will be called the "recursion matrix" of the structural relations (1.1). It carries the observed dependent variables y_i into a set of images y_i^* in such a way that the n th image depends only on y_n and possibly the preceding $n-1$ y_i , thus implying an order relation among the dependent variables. In view of (1.2), B always has an inverse of the same form, that is, also lower triangular with 1s in the diagonal. The $p \times m$ matrix A will be called the "regression pattern" of the structural relations (1.1), since its rows contain the regression weights for predicting $y^* =$

By from the independent variables in x (with heteroscedastic errors, $\text{Var}(e) = U^2$). Finally, the p -variate random vector $z' = (z_1, \dots, z_p)$ can be interpreted as a vector of p standardized "residuals" that are uncorrelated with the m independent variables x_j .

It should be noted that these residual random variables, although "unobserved," are perfectly well defined in the population by the structural relations (1.1) and the stochastic conditions (1.2) as linear combinations of the observed y_i and x_j . This is no longer the case in the presently popular "LISREL model" (e.g., Jöreskog, 1973), which is a generalization of the relations (1.1) and (1.2) and allows for "latent" variables, in analogy to the classical factor model. In this case, there is an excess of postulated (linearly) independent latent variables over observed variables. Hence, the latent variables ("causes") of the general LISREL model are indeterminate, for the same reasons that they are indeterminate in the classical factor model (see, e.g., Guttman, 1955, Schönemann & Steiger, 1978, Schönemann & Wang, 1972, and Wilson, 1928).

ALGEBRAIC SOLUTION

Since B is lower triangular with $\text{diag}(B) = \text{diag}(I)$, it has a unique inverse B^{-1} . On premultiplying (1.1) with B^{-1} , one finds for the covariance matrix of y with x , in view of (1.2b),

$$\text{Cov}(y,x) \text{Var}^{-1}(x) = B^{-1}A, \quad (2.1)$$

and for the variance-covariance matrix of the dependent variables

$$\text{Var}(y) = B^{-1}(A \text{Var}(x)A' + U^2)B^{-1'}. \quad (2.2)$$

On substituting (2.2) into (2.1), one finds that

$$\frac{\text{Var}(y) - B^{-1}A \text{Var}(x)A' B^{-1'}}{\text{Var}(y) - \text{Cov}(y,x) \text{Var}^{-1}(x) \text{Cov}(x,y)} = \quad (2.3)$$

The author's mailing address is: Department of Psychological Sciences, Purdue University, Lafayette, IN 47907.

is the partial variance-covariance matrix, $\text{Var}(y|x)$. Hence, (2.2) is equivalent to

$$HH' = (B^{-1}U)(UB^{-1'}) = \text{Var}(y|x), \tag{2.4}$$

where

$$H = B^{-1}U \tag{2.5}$$

is a choleski ("triangular") factor of $\text{Var}(y|x)$. It is well-known that such factorizations exist for all gramian matrices and that they are unique up to column reflections. Since $\text{rank}(\text{Var}(y|x)) = \text{rank}(H)$, a necessary and sufficient condition that U can be chosen positive definite is that $\text{Var}(y|x)$ be nonsingular. Since it is also gramian by definition, the reflections can then always be chosen so that H and, hence, also $U = \text{diag}(H)$ are positive definite. One thus arrives at

$$U = \text{diag}(H), \text{ (positive definite),} \tag{2.6a}$$

$$B = \text{diag}(H) H^{-1}, \tag{2.6b}$$

$$A = \text{diag}(H) H^{-1} \text{Cov}(y,x) \text{Var}^{-1}(x), \tag{2.6c}$$

where H is the positive definite choleski factor of the partial covariance matrix, $\text{Var}(y|x)$, and $\text{diag}(X)$ denotes the diagonal of a square matrix X .

It is easy to see that the map from $[\text{Var}(y), \text{Var}(x), \text{Cov}(y,x)]$ to $[\text{Var}(x), A, B, U]$ is bijective (1:1 onto), since

$$\text{Var}(x) = \text{Var}(x), \tag{2.7a}$$

$$\text{Var}(y) = B^{-1} (A \text{Var}(x) A' + U^2) B^{-1'}, \tag{2.7b}$$

$$\text{Cov}(y,x) = B^{-1} A \text{Var}(x). \tag{2.7c}$$

One thus arrives at Theorem 1: A necessary and sufficient condition that the structural relations (2.1) and (2.2) have a solution is that $\text{Var}(x)$ and $\text{Var}(y|x)$ be of full rank, in which case the solution is unique and given by (2.6).

The relations (1.1) and (1.2), therefore, do not define a "model" in the conventional sense of a formalized falsifiable theory, but rather a canonical decomposition of the observed (p+m)-variate random vector $(x'; y')$ akin to a principal components decomposition or a canonical correlation decomposition. On reading the earlier literature on path analysis, one gathers the impression that it has not always been realized that the classical recursive path "model" is actually a tautology (see also Numerical Illustrations section below).

As was shown, such a decomposition can always be found and therefore is empirically not falsifiable without imposing additional constraints. It is interesting to note that its existence and uniqueness do not depend on A having full column rank.

NUMERICAL ILLUSTRATIONS

Although the computations are fairly straightforward, a brief numerical illustration will be given. The data are taken from Li (1975, p. 324), who, in turn, borrowed them from Jencks, Smith, Acland, Baine, Cohen, Gintis, Heyns, and Michaelson (1972). In the analysis presented by Li, the first two variables are treated as independent variables:

- x_1 : father's education (FED)
- x_2 : father's occupation (FOC).

And the next three variables are treated as dependent variables:

- y_1 : early child IQ (IQ1)
- y_2 : child's education (CED)
- y_3 : child's adult IQ (IQ2)

Thus, in this case, the recursive path paradigm is used to explain the adult child's IQ as a function of the preceding four variables, two of which are treated as exogenous. The correlations matrix for the five variables are given in Table 1.

The computations can be summarized in six steps:

- (1) $G = \text{Cov}(y,x) \text{Var}^{-1}(x)$
- (2) $\text{Var}(y|x) = \text{Var}(y) - \text{Cov}(y,x) \text{Var}^{-1}(x) \text{Cov}(y,x)$
- (3) $\text{Var}(y|x) = HH'$ (choleski decomposition)
- (4) $U = \text{diag}(H)$
- (5) $B = UH^{-1}$
- (6) $A = BG$.

Table 1
Exact Solution of Recursive Path "Model" for Jencks's et al. (1972) Data

Variables:	1: father's education (FED)
	2: father's occupation (FOC)
	3: early child's IQ (IQ1)
	4: child's education (CED)
	5: adult child's IQ (IQ2)

Correlations (from Li, 1975, p.324):					
	1	2	3	4	5
1	1.00	.509	.300	.382	.305
2		1.00	.300	.420	.314
3			1.00	.550	.830
4				1.00	.630
5					1.00
	symmetric				

Algebraic solution obtained with

A. Variables 1 and 2 independent		B. Variable 1 independent	
Step 1:	$G = \text{Cov}(y x) \text{Var}^{-1}(x)$:		
		1	2
3	.199	.199	
4	.227	.304	
5	.196	.214	
			1
			2
			3
			4
			5
			.509
			.300
			.382
			.305

Step 2: $HH' = \text{Var}(y|x)$:

	3	4	5		2	3	4	5
3	.881	-.391	.707		2	.741	.147	.226
4		.785	.465		3		.910	.435
5		sym.	.873		4			.854
					5	sym.		.907

Step 3: choleski factor H:

	3	4	5		2	3	4	5
3	.938	0	0		2	.861	0	0
4		.416	.782		3		.938	0
5		.707	.465	.873	4			.782
					5	.184	.753	.194

Step 4: $U = \text{diag}(H)$

$U = \text{diag}(.938 \ .782 \ .873)$ $U = \text{diag}(.861 \ .938 \ .782 \ .518)$

Step 5: $B = UH^{-1}$:

	3	4	5		2	3	4	5
3	1.00	0	0		2	1.00	0	0
4		.443	1.00		3	-.199	1.00	0
5		-.693	-.248	1.00	4	-.216	-.443	1.00
					5	-.001	-.693	-.248

Step 6: $A = BG$:

	1	2		1
3	.199	.199		2
4	.139	.216		3
5	.002	.001		4
				5

Check: $B \text{Var}(y) B' - A \text{Var}(x) A' = U^2$:

	3	4	5		2	3	4	5
3	.881	0	0		2	.741	0	0
4		0	.612		3	0	.881	0
5		0	0	.268	4	0	0	.612
					5	0	0	0

The matrix G at Step 1 is simply the (sample estimate of the) matrix of regression weights of y on x, and the matrix at Step 2 is the estimate of the partial covariance matrix, $\text{Var}(y|x)$. Choleski factorization of $\text{Var}(y|x)$ then gives the gram factor H at Step 3 and its diagonal the diagonal matrix U at Step 4. Inversion of H gives the recursion matrix $B = UH^{-1}$ for the dependent variables at Step 5, which is used to obtain the regression pattern $A = BG$ at Step 6. Since none of these computations involve iterations of any kind, this algebraic solution is computationally very efficient and can be obtained with the most elementary equipment, for example, a microcomputer.

These computations are illustrated, step by step, in Table 1 for the Jencks et al. (1972) data, both for $m = 1$ (one independent, four dependent variables) and $m = 2$ (two independent, three dependent variables). The residual matrices, $B \text{Var}(y) B' - A \text{Var}(x) A'$, are exactly equal to U^2 in both instances.

The numerical results obtained for $m = 2$ coincide exactly with those given by Li (1975, path diagram, p. 325), who, however, was apparently unaware that the structural relations (1.1) and (1.2) are tautological:

In the analysis above, we have used 8 of the 10 correlations. The two correlations not used are $r_{Q_1} = 0.305$ and $r_{Q_2} = 0.314$ [these are the correlations of the two independent variables, FED and FOC, with the ultimate dependent variable, IQ2 - PHS]. On the basis of Fig. 325 [the path diagram], two additional equations may be written down for these two correlations. These are called "excess" equations and the system is said to be over-determined or over-identified" (Li, 1975, p. 325).

The author then used his path coefficients to compute the two correlations supplied by the "excess equations" to obtain an "estimate" of .304 for r_{Q_1} . On the basis of the small numerical discrepancy, he concluded:

According to the diagram, the expected correlation between X_1 and Q is .304, while the observed correlation is $r_{Q_1} = 0.305$, in spite of the sampling and rounding off of errors. A similar situation exists for r_{Q_2} . We conclude that the path coefficients indicated in Fig. 325 are consistent with the data" (emphasis added) (Li, 1975, p. 327).

This example illustrates once again that counting equations and unknowns may lead to false inferences about the degree of identifiability of systems of structural relations (see also Guttman's, 1958, p. 306, critique of the once widespread trust in "Lederman's number" by factor analysis). In the above example, the system is "just identified" (in the sense that it always has exactly one and only one solution), because the "excess equations" are redundant. The small numerical discrepancies Li (1975) observed are solely due to rounding error, not to sampling error.

The fact that the defining structural relations (1.1) and (1.2) of the recursive path "model" with manifest variables are tautological in the sense that no data can be found that permit rejection of these postulated relations does not mean, of course, that they cannot be rendered empirically falsifiable by imposing additional constraints.

REFERENCES

GUTTMAN, L. (1955). The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. *British Journal of Statistical Psychology*, 8, 65-81.
 GUTTMAN, L. (1958). To what extent can communalities reduce rank? *Psychometrika*, 23, 297-308.
 JENCKS, C., SMITH, M., ACLAND, H., BAINE, M. J., COHEN, D., GINTIS, H., HEYNS, B., & MICHAELSON, S. (1972). *Inequality, a reassessment of family and schooling in America*. New York: Basic Books.
 JÖRESKOG, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85-112). New York: Seminar Press.
 LI, C. C. (1975). *Path analysis—A primer*. Pacific Grove: Boxwood.
 SCHÖNEMANN, P. H., & STEIGER, J. H. (1978). On the validity of indeterminate factor scores. *Bulletin of the Psychonomic Society*, 12, 287-290.

SCHÖNEMANN, P. H., & WANG, M. M. (1972). Some new results on factor determinacy. *Psychometrika*, **37**, 61-91.

WILSON, E. B. (1928). Review of *the abilities of man, their nature and measurement*, by C. Spearman. *Science*, **67**, 244-248.

WRIGHT, S. (1918). On the nature of size factors. *Genetics*, **3**, 367-374.

WRIGHT, S. (1921). Correlation and causation. *Journal of Agricultural Research*, **20**, 557-585.

(Manuscript received for publication April 12, 1984.)