Producing weight matrices and fit functions

As one of its options, PRELIS produces the *asymptotic covariance matrix* of estimated covariances and correlations. We explain here what this matrix is and how it can be used to produce *weight matrices* for certain fit functions in LISREL.

A general family of fit functions for analysis-of-covariance structures may be written (see, for example, Browne, 1984)

$$F(\boldsymbol{\theta}) = (\mathbf{s}-\boldsymbol{\sigma})'\mathbf{W}^{-1}(\mathbf{s}-\boldsymbol{\sigma}) = \sum_{g=1}^{k}\sum_{h=1}^{g}\sum_{i=1}^{k}\sum_{j=1}^{i} w^{gh,ij}(s_{gh}-\sigma_{gh})(s_{ij}-\sigma_{ij}) \tag{1}$$

where

$$\mathbf{s}' = (s_{11}, s_{21}, s_{22}, s_{31}, \ldots, s_{kk})$$

is the vector of the elements in the lower half, including the diagonal, of the covariance matrix **S** of order $k \times k$ used to fit the model to the data;

$$\boldsymbol{\sigma}' = (\sigma_{11}, \sigma_{21}, \sigma_{22}, \sigma_{31}, \ldots, \sigma_{kk})$$

is a vector of corresponding elements of $\Sigma(\boldsymbol{\theta})$ reproduced from the model parameters $\boldsymbol{\theta}$; and $w^{gh,ij}$ is a typical element of a positive-definite matrix $\mathbf{W}^{-1}$ of order $p \times p$, where $p = k(k+1)/2$. In most cases, the elements of $\mathbf{W}^{-1}$ are obtained by inverting a matrix $\mathbf{W}$ whose typical element is denoted $w_{gh,ij}$. The usual way of choosing $\mathbf{W}$ in weighted least squares is to let $w_{gh,ij}$ be a consistent estimate of the asymptotic covariance between s_{gh} and s_{ij} but, in principle, any positive-definite matrix $\mathbf{W}$ may be used. To estimate the model parameters $\boldsymbol{\theta}$, the fit function is minimized with respect to $\boldsymbol{\theta}$.

Under very general assumptions, if the model holds in the population and if the sample variances and covariances in $\mathbf{S}$ converge in probability to the corresponding elements in the population covariance matrix $\boldsymbol{\Sigma}$ as the sample size increases, any such fit function will give a consistent estimator of $\boldsymbol{\theta}$. In practice, numerical results obtained by one fit function are often close enough to the results that would be obtained by another fit function, to allow the same substantive interpretation.

Further assumptions must be made, however, if one needs an asymptotically correct chi-square test of goodness of fit and asymptotically correct standard errors of parameter estimates.

"Classical" theory for covariance structures (see, for example, Browne, 1974 or Jöreskog, 1981) assumes that the asymptotic variances and covariances of the elements of **S** are of the form

$$ACov(s_{gh}, s_{ij}) = (1/N)(\sigma_{gi}\sigma_{hj} + \sigma_{gj}\sigma_{hi}) \tag{2}$$

where N is the total sample size. This holds, in particular, if the observed variables have a multivariate normal distribution, or if **S** has a Wishart distribution. The GLS and ML methods available in LISREL and their chi-square values and standard errors are based on these assumptions. The GLS method corresponds to using a matrix **W** in (1) whose general element is

$$w_{gh,ij} = (1/N)(s_{gi}s_{hj} + s_{gj}s_{hi}) \tag{3}$$

The fit function for ML is not of the form (1) but may be shown to be equivalent to using a **W** of the form (3), with **s** replaced with an estimate of $\boldsymbol{\sigma}$ that is updated in each iteration.

In fundamental work by Browne (1982, 1984), this classical theory for covariance structures has been generalized to any multivariate distribution for continuous variables satisfying very mild assumptions. This approach uses a **W** matrix with typical element

$$w_{gh,ij} = m_{ghij} - s_{gh}s_{ij} \tag{4}$$

where

$$m_{ghij} = (1/N)\sum_{\alpha=1}^{N}(z_{\alpha g} - \bar{z}_g)(z_{\alpha h} - \bar{z}_h)(z_{\alpha i} - \bar{z}_i)(z_{\alpha j} - \bar{z}_j)$$

are the fourth-order central moments. Using such a **W** in (1) gives what Browne calls "asymptotically distribution free best GLS estimators" for which correct asymptotic chi-squares and standard errors may be obtained. As shown by Browne, this **W** matrix also may be used to compute correct asymptotic chi-squares and standard errors for estimates that have been obtained by the classical ML and GLS methods. When **W** is defined by (4), we call the fit function WLS (weighted least squares) to distinguish it from GLS where **W** is defined by (3). WLS and GLS are different forms of weighted least squares; WLS is asymptotically distribution free, while GLS is based on normal theory.

While WLS is attractive in theory, it presents several difficulties in practical applications. First, the matrix **W** is of order $p \times p$ and has $p(p+1)/2$ distinct elements. This increases rapidly with k, demanding large amounts of computer memory when k is large. Second, to estimate moments of fourth order with reasonable precision requires very large samples. Third, when there are missing observations in the data, different moments involved in (4) may be based on different numbers of cases unless listwise deletion is used. When pairwise deletion is used, it is not clear how to deal with this problem.

Finally, Browne's (1984) development is a theory for sample covariance matrices for continuous variables. In practice, however, correlation matrices are often analyzed; that is, covariance matrices scaled by stochastic standard deviations. The elements of such a correlation matrix do not have asymptotic variances and covariances of the form (2), even if **S** has a Wishart distribution. In PRELIS, an estimate of the asymptotic covariance matrix of the estimated correlations can also be obtained under the same general assumptions of non-normality. This approach can be used when some or all of the variables are ordinal or censored, after the raw scores are replaced by normal scores. PRELIS can also compute estimates of the asymptotic variances and covariances of estimated polychoric and polyserial correlations. This approach is similar to that of Muthén (1984), but the PRELIS estimates are much simpler and faster to compute.

A correlation matrix estimated in PRELIS with the KM or PM option has $q = k(k-1)/2$ estimated correlations and, as a consequence, the asymptotic covariance matrix of these correlation is of order $q \times q$. To obtain the weight matrix to be used in LISREL, this covariance matrix must be inverted. The inversion is not performed by PRELIS but is part of LISREL. The asymptotic covariance matrix of estimated coefficients obtained by PRELIS may be saved in a file that can be read directly by LISREL.

To sum up: whenever possible in PRELIS, an estimate of the asymptotic covariance matrix of the elements of the estimated moment matrix is provided. Currently, such asymptotic covariance matrices are available for sample covariance, moment, and augmented moment matrices and matrices of product-moment (Pearson), polychoric, and/or polyserial correlations. Asymptotic covariance matrices are not yet available for OM, RM, and TM matrices.

Computation of asymptotic covariance matrices of estimated coefficients is very time-consuming and demands large amounts of memory when the number of carriable is large. An alternative approach, which may be used even when the number of variables is large, is to compute only the asymptotic variances of the estimated coefficients. Let w_{gh} be an estimate of the asymptotic variance of s_{gh} . These estimates may be used with a fit function of the form:

$$F(\boldsymbol{\theta}) = \sum_{g=1}^{k} \sum_{h=1}^{g} (1/w_{gh})(s_{gh} - \sigma_{gh})^2 \quad (5)$$

This corresponds to using a diagonal weight matrix $\mathbf{W}^{-1}$ in (1). In LISREL, this is called DWLS (diagonally weighted least squares). This does not lead to asymptotically efficient estimates of model parameters but is offered as a compromise between unweighted least squares (ULS) and fully weighted least squares (WLS). The DWLS method can also be used when correlation matrices (KM or PM) are analyzed.

References

Browne, M.W. (1974). Generalised Least Squares Estimators in the analysis of covariance structures. *South African Statistical Journal*, **8**, 1-24.

Browne, M.W. (1982). *Covariance Structures*. In D.M. Hawkins (Ed.), *Topics in Applied Multivariate Analysis*, pp. 72-141. Cambridge: Cambridge University Press.

Browne, M.W. (1984). Asymptotically Distribution-free Methods in the Analysis of Covariance Structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62-83.

Jöreskog, K.G. (1981). Analysis of covariance structures. *Scandinavian Journal of Statistics*, **8**, 65-92.