

# **Confirmatory Factor Analysis and “Exploratory-Confirmatory” Factor Analysis**

## ***Maximum Likelihood Factor Analysis***

Maximum likelihood factor analysis can be viewed as a special case of structural equation modeling. In structural equation modeling, we model the population covariance matrix  $\Sigma$  as a *matrix function* of a vector of free parameters (numbers that are free to vary)  $\theta$ .

Our model is thus

$$\Sigma = \mathbf{M}(\theta) \quad (1)$$

***Example.***

Consider the common factor model. What is  $\mathbf{M}(\boldsymbol{\theta})$ ?

The factor model states that  $\boldsymbol{\Sigma} = \mathbf{F}\mathbf{F}' + \mathbf{U}^2$ . Suppose we have 4 variables and one factor. Then

$$\mathbf{f} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \quad (2)$$

and

$$\mathbf{U}^2 = \begin{bmatrix} \theta_5 & 0 & 0 & 0 \\ 0 & \theta_6 & 0 & 0 \\ 0 & 0 & \theta_7 & 0 \\ 0 & 0 & 0 & \theta_8 \end{bmatrix} \quad (3)$$

What is  $\mathbf{M}[\boldsymbol{\theta}]$ ? (C.P.)

So we end up with a matrix equation. We can vectorize it, and say something like

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} \\ \sigma_{21} \\ \vdots \end{bmatrix} = \begin{bmatrix} \theta_1^2 + \theta_5^2 \\ \theta_2 \theta_1 \\ \vdots \end{bmatrix} \quad (4)$$

Note that we have a set of simultaneous nonlinear equations.

## ***Fitting the Model***

Almost inevitably a sample covariance matrix  $\mathbf{S}$  will not be equal to  $\mathbf{M}(\boldsymbol{\theta})$  for *any*  $\boldsymbol{\theta}$ , *even if we were so lucky to have the model fit perfectly in the population.*

What we can try to do is “come as close as we can” by *minimizing a discrepancy function.*

## ***Discrepancy Functions***

A discrepancy function is a scalar function that is measure of discrepancy between a covariance matrix and the model used to reproduce it.

Discrepancy functions to be reasonable must possess the following 3 basic properties.

$$F(\mathbf{S}, \mathbf{M}(\boldsymbol{\theta})) \geq 0$$

$$F(\mathbf{S}, \mathbf{M}(\boldsymbol{\theta})) = 0 \text{ if and only if } \mathbf{S} = \mathbf{M}(\boldsymbol{\theta})$$

$$F(\mathbf{S}, \mathbf{M}(\boldsymbol{\theta})) \text{ is continuous in } \mathbf{S} \text{ and } \mathbf{M}(\boldsymbol{\theta})$$

***Examples.***

**The “ordinary least squares discrepancy function”**

$$F_{OLS}(\mathbf{S}, \mathbf{M}(\boldsymbol{\theta})) = \frac{1}{2} \text{Tr}(\mathbf{S} - \mathbf{M}(\boldsymbol{\theta}))^2 \quad (5)$$

## The “maximum likelihood” discrepancy function

$$F_{ML} = \ln |\mathbf{M}(\boldsymbol{\theta})| - \ln |\mathbf{S}| + \text{Tr}(\mathbf{S}\mathbf{M}(\boldsymbol{\theta})^{-1}) - p \quad (6)$$

## ***The Chi Square Fit Statistic***

Under an assumption that  $\mathbf{S}$  has a Wishart distribution (which it will if the population is multivariate normal),

$$(N - 1) \text{Min}_{\theta}(F_{ML}) \quad (7)$$

has an asymptotic distribution that is  $\chi^2$  with  $p(p + 1)/2 - t$  degrees of freedom, where  $t$  is the number of effective free parameters in the model. This can be used to evaluate model fit.

## ***Chi-Square Difference Test***

A nested sequence of models can be tested using a chi square *difference test*. The difference between the two chi square statistics is a chi square, with degrees of freedom equal to the difference in degrees of freedom for the two individual chi-squares.

## ***Noncentral Chi Square Approximation***

Usually a model is not, strictly speaking, true. Consider the *population ML discrepancy function* for model  $\mathbf{M}_i$ . This is the value of the discrepancy function that we would obtain if we actually knew  $\Sigma$ , and computed

$$F_{M_i}^* = F(\Sigma, \mathbf{M}_i(\theta)) \quad (8)$$

Steiger, Shapiro, and Browne (1985) showed that under a set of assumptions called “population drift,” the chi square statistic  $(N - 1)F_{ML}$  has an asymptotic noncentral chi square distribution.

More importantly, they showed that this distribution has a noncentrality parameter equal to

$$\lambda = (N - 1)F_{M_i}^* \quad (9)$$

It is possible to get a confidence interval on  $\lambda$ . One can then divide its endpoints by  $N - 1$  to get a confidence interval on  $F_{M_i}^*$ . This can be used to assess “how bad” fit of a model is.

## ***Parameter Estimates and Standard Errors***

The value of  $\theta$  that minimizes the discrepancy function yields “maximum likelihood” estimators. As a byproduct of maximum likelihood estimation, one obtains the  $\hat{\theta}_i$ . These estimators are consistent, asymptotically normal, with standard errors that can be estimated in several ways.

Typically, a structural equation modeling program will output the  $\hat{\theta}_i$  and their estimated standard errors, then compute a “*t*-statistic” that is simply the ratio of the estimate to its standard error. You can use these statistics to gauge which parameters are “statistically significant.”

(C.P. What about the problem of multiple testing?)

## ***Confirmatory Factor Analysis***

This was developed in order to be able to test whether a factor pattern “really” has simple structure in a particular form.

For example, we could test whether our athletics data have, after rotation, a perfect simple structure of the form

$$\begin{bmatrix} X & 0 & 0 \\ X & 0 & 0 \\ X & 0 & 0 \\ 0 & X & 0 \\ 0 & X & 0 \\ 0 & X & 0 \\ 0 & 0 & X \\ 0 & 0 & X \\ 0 & 0 & X \end{bmatrix}$$

All the  $X$ 's are generally free parameters.

Whenever you put  $X$ 's the program computes  $\mathbf{M}(\boldsymbol{\theta})$ , then solves for the maximum likelihood estimates, and prints out the chi square statistic, confidence intervals on the population discrepancy function, and a variety of “model fit indices,” along with standard errors and  $t$  statistics for the parameters.

Let's try an example.