Constraints

LISREL does not directly impose constraints on the covariance matrices $\mathbf{\Phi}$, $\mathbf{\Psi}$, $\mathbf{\Theta}_{\varepsilon}$ and $\mathbf{\Theta}_{\delta}$ so as to make these positive definite. If an estimate of any one of these matrices is not positive definite, *when they are supposed to be so*, this is an indication that the model is wrong.

LISREL checks the positive definiteness of $\mathbf{\Phi}$, $\mathbf{\Psi}$, $\mathbf{\Theta}_{\varepsilon}$ and $\mathbf{\Theta}_{\delta}$ as part of the admissibility test (see AD keyword on the OU command). If after AD iterations (default value = 10) one of these matrices is not positive definite, the iterations will stop and the current "solution" will be printed. If one believes strongly that there is an admissible solution for the model and the data analyzed, one should increase the value of AD and rerun the problem with different and better starting values. Our experience suggests, however, that when the "solution" is non-admissible after 10 iterations, and the program is allowed to continue to iterate, it will either converge to a non-admissible solution or not converge at all. As already stated, this is usually the fault of the model rather than the program. However, this problem has also occurred in Mote Carlo studies for occasional odd samples, where it cannot be blamed on the model. The example below illustrates this case.

Suppose we believe in the model and cannot accept a negative estimate of an error variance. If the LISREL estimates are admissible, there is no problem. However, if one or more of the estimates of error variances is negative one must reparameterize the model in a way to prevent this.

Suppose, for example, we are estimating a LISREL Submodel 1 with $\mathbf{\Theta}_{\delta}$ diagonal. The model can be reparameterized as

$$\mathbf{\Lambda}_x^* = \begin{bmatrix} \mathbf{\Lambda}_x & \mathbf{D}_{\delta} \end{bmatrix} \quad \mathbf{\Phi}^* = \begin{bmatrix} \mathbf{\Phi} & \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \mathbf{\Theta}_{\delta}^* = \mathbf{0},$$

where $\mathbf{D}_{\delta}$ is a diagonal matrix such that $\mathbf{\Theta}_{\delta} = \mathbf{D}_{\delta}^2$. It is easily verified that

$$\mathbf{\Lambda}_x^*\mathbf{\Phi}^*\mathbf{\Lambda}_x^{*'} + \mathbf{\Theta}_{\delta}^* = \mathbf{\Lambda}_x\mathbf{\Phi}\ \mathbf{\Lambda}_x^{'} + \mathbf{D}_{\delta}^2 = \mathbf{\Lambda}_x\mathbf{\Phi}\ \mathbf{\Lambda}_x^{'} + \mathbf{\Theta}_{\delta}.$$

Estimates of error variances are obtained as $\hat{\mathbf{D}}_\delta^2$. The elements d_i of $\mathbf{D}_\delta$ may come out positive or negative (or zero), but the error variances are estimated as d_i^2 , which cannot be negative.

It should be noted that the reparameterized model has NK > NX so that the IV and TSLS procedures for generating starting values will not work. Starting values must be provided by the user.

In a Monte Carlo study (Hägglund (1982)) random samples were generated from the following population parameters:

$$\mathbf{\Lambda}_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.889 & 0 \\ 0 & 0.857 \\ 0 & 0.714 \\ 0.333 & 0 \end{bmatrix}$$

$$\mathbf{\Phi} = \begin{bmatrix} 0.810 & \\ 0.378 & 0.490 \end{bmatrix}$$

$$\mathbf{\Theta}_\delta = diag\begin{bmatrix} 0.19 \\ 0.51 \\ 0.36 \\ 0.64 \\ 0.75 \\ 0.91 \end{bmatrix}.$$

The population covariance matrix

$$\mathbf{\Sigma} = \mathbf{\Lambda}_x \mathbf{\Phi} \mathbf{\Lambda}_x' + \mathbf{\Theta}_\delta$$

formed from these parameter matrices is

$$\mathbf{\Sigma} = \begin{bmatrix} 1.000 & & & & & \\ 0.378 & 1.000 & & & & \\ 0.720 & 0.336 & 1.000 & & & \\ 0.324 & 0.420 & 0.288 & 1.000 & & \\ 0.270 & 0.350 & 0.240 & 0.300 & 1.000 & \\ 0.270 & 0.126 & 0.240 & 0.108 & 0.090 & 1.000 \end{bmatrix}.$$

LISREL can be used to compute $\boldsymbol{\Sigma}$ from $\boldsymbol{\Lambda}_x$, $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}_\delta$. The following LISREL command file (**EX81A.LIS**) will do the job:

```
Ex8.1a: Computing Population Sigma
DA NI=6 NO=10
CM
1 0 1 2*0 1 3*0 1 4*0 1 5*0 1
MO NX=6 NK=2 PH=FI TD=FI
MA LX
1 0 0 1 .889 0 0 .857 0 .714 .333 0
MA PH
.810 .378 .490
MA TD
.19 .51 .36 .64 .75 .91
OU SI=EX81.SIG
```

The trick is to specify all parameter matrices to be fixed and enter the parameter matrices by MA commands. There are no parameters to estimate, so LISREL will just compute $\boldsymbol{\Sigma}$ and save it in the file **EX81.SIG** as requested in the OU command. However, LISREL always need some data to analyze. The matrix to be "analyzed" may be *any* positive-definite matrix. Here we use an identity matrix as this is particularly easy to enter.

Hägglund (1982) was concerned with unrestricted factor analysis. In this context, this corresponds to the model where all elements in rows 3 – 6 in $\boldsymbol{\Lambda}_x$ are unknown parameters. To see that LISREL works correctly, when the population $\boldsymbol{\Sigma}$ is analyzed, we run the following command file (**EX81B.LIS**):

```
Ex8.1b: Analyzing Population Sigma
DA NI=6 NO=200
CM FI=EX81.SIG
MO NX=6 NK=2
FR LX(5)-LX(12)
VA 1 LX(1) LX(4)
OU SE
```

The output file reveals that LISREL correctly recovers all the parameters, including the zeros in $\boldsymbol{\Lambda}_x$. The output file also shows that all residuals are zero.

The sample size specified on the DA command is of course arbitrary and irrelevant. For reasons which will be obvious in a moment, we chose NO = 200. The output file suggests that the *true* standard error of $\theta_{11}^{(\delta)}$ is 0.166. This suggests that if we were to take repeated random samples of size 200 from the population, the estimated $\hat{\theta}_{11}^{(\delta)}$ in these samples should fall in the interval $0.19 \pm (1.96 \times 0.166) = 0.19 \pm 0.33$ in 95% of the cases. This interval is from -0.14 to 0.52. Thus, some of the estimates are likely to be negative. In fact, if the normality approximation holds, we should expect 13 percent negative estimates. It should therefore not come as a surprise that negative estimates of error variances (Heywood cases) occur in a Monte Carlo study designed like this.

Gösta Hägglund generated many random sample covariance matrices based on $\boldsymbol{\Sigma}$. One of these is given in the file **EX81.COV**.

Although this matrix does not "look strange", it is an extremely odd sample. No matter what one does, one almost always end up with a non-admissible solution. We shall illustrate how LISREL behaves in a case like this.

Suppose we proceed in the usual way to estimate the model with free λ 's in rows 3 – 6. The command file is (**EX81C.LIS**):

```
Ex8.1c: Analyzing Gosta's Bad Sample
DA NI=6 NO=200
CM FI=EX81.COV
MO NX=6 NK=2
FR LX(5)-LX(12)
VA 1 LX(1) LX(4)
OU
```

Since AD is default on the OU command, LISREL stops after 10 iterations with a large negative estimate of TD(1). There is no admissible solution for this data and model. This can be verified by starting at different initial values and allowing the program to iterate further. For example, if one sets AD = OFF and IT = 250, say, LISREL iterates "forever" producing larger and larger negative estimates of TD(1). This can also be verified by fixing TD(1) at zero and plotting the fit function against it, as shown in the figure below.

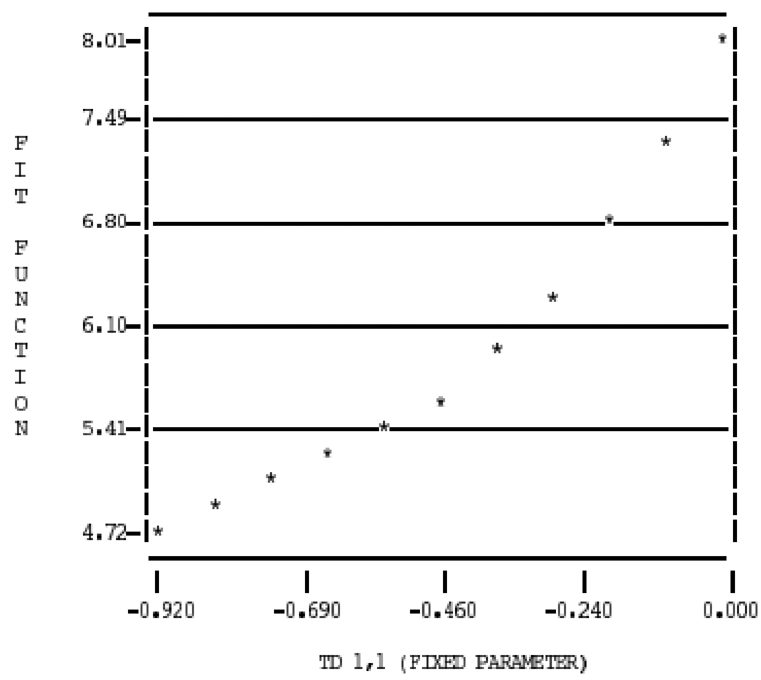

Figure: Plot of Fit Function against TD(1,1)

Next, let us constrain the error variances to be non-negative using the reparameterized model

$$\mathbf{\Lambda}_x^* = \begin{bmatrix} \mathbf{\Lambda}_x & \mathbf{D}_\delta \end{bmatrix} \quad \mathbf{\Phi}^* = \begin{bmatrix} \mathbf{\Phi} & \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \mathbf{\Theta}_\delta^* = \mathbf{0}.$$

This can be done as follows (**EX81D.LIS**):

```
Ex8.1d: Analyzing Gosta's Bad Sample by Reparameterized Model
DA NI=6 NO=200
CM FI=EX81.COV
MO NX=6 NK=8 LX=FI PH=FI TD=ZE
FR LX 1 3 LX 2 4 LX 3 5 LX 4 6 LX 5 7 LX 6 8
FR LX 3 1 LX 3 2 LX 4 1 LX 4 2 LX 5 1 LX 5 2 LX 6 1 LX 6 2
FR PH 1 1 PH 2 1 PH 2 2
VA 1 LX 1 1 LX 2 2 PH 3 3 PH 4 4 PH 5 5 PH 6 6 PH 7 7 PH 8 8
ST .5 ALL; ST 0 PH 2 1
OU NS AD=OFF
```

Note that starting values for $\mathbf{D}_\delta$ must be given. Otherwise, the matrix $\mathbf{\Sigma}$ computed at initial values is singular which means that the ML function cannot be computed. Also, as this is a very problematic case, we have to allow for more iterations than does the default value of IT (in this case 51 iterations). The resulting solution shows that all the error variances are positive except for variable 1, which is zero. There is no admissible solution for this data.

It should be emphasized that constraining error variances to be non-negative does not really solve the problem. Zero estimates of error variances are as unacceptable as are negative estimates. The root of the problem is that the model is empirically overparameterized. Instead of estimating *all* the λ 's in row 3 – 6 of $\mathbf{\Lambda}_x$, one should ask the question which of these eight λ 's are statistically (and substantively) zero and which are statistically (and substantively) non-zero? One can answer this question by starting with the model in which all eight λ 's are fixed at zero, and free one λ at a time using the information provided by the modification index and estimated change. In the non-zero λ 's should be positive, one should not free an λ for which the estimated change is negative even if its modification index is large. When the four λ 's which are zero in the population are fixed at zero (see **EX81E.LIS** below), the following solution is obtained.

```
Ex8.1e: Analyzing Gosta's Bad Sample
DA NI=6 NO=200
CM FI=EX81.COV
MO NX=6 NK=2
VA 1 LX(1) LX(4)
FR LX 3 1 LX 4 2 LX 5 2 LX 6 1
OU


LISREL Estimates (Maximum Likelihood)

         LAMBDA-X

               KSI 1      KSI 2
            --------   --------
    VAR 1      1.000       - -
```

```
VAR 2       - -       1.000

VAR 3     0.820       - -
        (0.076)
         10.832

VAR 4       - -       0.824
                    (0.124)
                      6.647

VAR 5       - -       0.788
                    (0.127)
                      6.205

VAR 6     0.142       - -
        (0.069)
          2.055


     PHI

          KSI 1      KSI 2
        --------   --------
KSI 1     1.143
        (0.154)
          7.434

KSI 2     0.569      0.546
        (0.092)    (0.117)
          6.200      4.675


     THETA-DELTA

          VAR 1      VAR 2      VAR 3      VAR 4      VAR 5      VAR 6
        --------   --------   --------   --------   --------   --------
          0.134      0.579      0.380      0.581      0.723      0.989
        (0.088)    (0.088)    (0.070)    (0.075)    (0.087)    (0.099)
          1.530      6.599      5.462      7.727      8.354      9.978
```

This may be compared with the population parameters.

Constraining covariance matrices to be non-negative definite:

Consider LISREL Submodel 1 with $\mathbf{\Phi}$ free and suppose we want to constrain $\mathbf{\Phi}$ to be non-negative definite. We can then specify $\mathbf{\Phi}$ as

$$\boldsymbol{\Phi} = \mathbf{T}_{\phi}\mathbf{T}_{\phi}^{'},$$

where $\mathbf{T}_{\phi}$ is a lower-triangular matrix. This can be reparameterized as a Submodel 3A with

$$\boldsymbol{\Lambda}_x = \boldsymbol{\Lambda}_y, \mathbf{B} = \mathbf{0}, \boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{\phi}, \boldsymbol{\Phi} = \mathbf{I}, \boldsymbol{\Theta}_{\varepsilon} = \boldsymbol{\Theta}_{\delta}.$$

The covariance matrix $\hat{\boldsymbol{\Phi}} = \hat{\mathbf{T}}_{\phi}\hat{\mathbf{T}}_{\phi}^{'}$ is obtained as the covariance matrix of $\boldsymbol{\eta}$ in the output.

To constrain $\boldsymbol{\Psi}$ to be non-negative definite in the full model, we specify this as

$$\begin{aligned} \boldsymbol{\eta} &= \mathbf{B}\boldsymbol{\Phi} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \\ &= \mathbf{B}\boldsymbol{\eta} + \begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{T}_{\psi} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta}^{*} \end{bmatrix}, \end{aligned}$$

where $\mathbf{T}_{\psi}$ is a lower triangular matrix such that $\boldsymbol{\Psi} = \mathbf{T}_{\psi}\mathbf{T}_{\psi}^{'}$ and $Cov\left(\boldsymbol{\zeta}^{*}\right) = \mathbf{I}$.