# Variance and covariance components

Covariance structure analysis may be used to study differences in test performance when the tests have been constructed by assigning items or subtests according to objective features of content or format to subclasses of a factorial or hierarchical classification.

Bock (1960) suggested that the scores of $N$ subjects on a set of tests classified in a $2^q$ factorial design may be viewed as data from an $N \times 2^q$ experimental design, where the subjects represent a random mode of classification and the tests represent $n$ fixed modes of classification. Bock pointed out that conventional mixed-model analysis of variance gives useful information about the psychometric properties of the tests. In particular, the presence of non-zero variance components for the random mode of classification and for the interaction of the random and fixed modes of classification provides information about the number of dimensions in which the tests are able to discriminate among subjects. The relative size of these components measures the power of the tests to discriminate among subjects along the respective dimensions.

In this example (Browne (1970)), the Rod and Frame (RF) test is used as a measure of field dependence. A subject is seated in a darkened room on a chair which may be titled to the left or to the right. In front of the subject is a luminous rod located in a luminous square frame. The chair, frame, and rod are tilted to pre-specified positions. By operating push buttons connected to an electric motor, the subject should move the rod to the vertical position. The score on the trial is the angle of the rod from the vertical, as positive and negative values are possible. Each subject undergoes 12 trials. The last two columns of the design matrix $\mathbf{A}$ below give initial positions of the frame and chair for each trial.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

A value of +1 denotes that the position of the frame or chair was at 28 degrees from the vertical, a value of -1 indicates that the angle was – 28 degrees and a value of 0 denotes that the initial position was vertical.

The covariance matrix between trials of the RF test obtained from a sample of 107 eighteen-year-old males is given in the table below.

**Table: Inter-trial covariance matrix for the rod and frame test**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 51.6 | | | | | | | | | | | |
| -27.7 | 72.1 | | | | | | | | | | |
| 38.9 | -41.1 | 69.9 | | | | | | | | | |
| -36.4 | 40.7 | -39.1 | 75.8 | | | | | | | | |
| 13.8 | -5.2 | 17.9 | 1.9 | 84.8 | | | | | | | |
| -13.6 | 10.9 | 9.5 | 17.8 | -37.4 | 91.1 | | | | | | |
| 21.5 | -9.4 | 8.5 | -13.1 | 59.7 | -54.4 | 79.9 | | | | | |
| -12.8 | -17.2 | -3.1 | 22.0 | -43.3 | 52.7 | -49.9 | 87.2 | | | | |
| 11.0 | -8.9 | 19.2 | -11.2 | -12.6 | 21.9 | -10.6 | 17.5 | 27.6 | | | |
| -4.5 | 10.2 | -7.6 | 12.7 | 20.4 | -11.5 | 16.5 | -14.8 | -8.8 | 19.9 | | |
| 9.2 | -0.3 | 18.9 | -13.6 | -3.9 | 19.0 | -8.3 | 13.1 | 17.7 | -2.8 | 27.3 | |
| -3.7 | -4.5 | -4.5 | 12.8 | 19.9 | -8.8 | 15.5 | -8.6 | -5.4 | 13.3 | -1.0 | 16.0 |

We want to estimate the variance components associated with general bias, frame effect, chair effect, and error.

Let *a*, *b*, and *c* be uncorrelated random components associated with general bias, frame effect, and chair effect, respectively, and let *e* denote an error component uncorrelated with *a*, *b*, and *c* and uncorrelated over trials. Let
‘

$$\mathbf{u}'_{\upsilon} = [a_{\upsilon} \quad b_{\upsilon} \quad c_{\upsilon}]$$

be the values of *a*, *b*, and *c* for subject $\upsilon$. Then the scores on the twelve trials for subject $\upsilon$ is

$$\mathbf{x}_\upsilon = \mathbf{A}\mathbf{u}_\upsilon + \mathbf{e}_\upsilon$$

with covariance matrix

$$\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Phi}\mathbf{A}' + \sigma_e^2\mathbf{I},$$

where

$$\boldsymbol{\Phi} = diag\left(\sigma_a^2, \sigma_b^2, \sigma_c^2\right).$$

The equation shows that all the 78 variances and covariances in $\boldsymbol{\Sigma}$ are *linear* functions of the four parameters $\sigma_a^2$, $\sigma_b^2$, $\sigma_c^2$, and $\sigma_e^2$. To see this explicitly, consider the covariance matrix generated by trials 1, 2 5, and 9:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & & & \\ \sigma_a^2-\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & & \\ \sigma_a^2-\sigma_b^2+\sigma_c^2 & \sigma_a^2+\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & \\ \sigma_a^2+\sigma_b^2 & \sigma_a^2-\sigma_b^2 & \sigma_a^2-\sigma_b^2 & \sigma_a^2+\sigma_b^2+\sigma_e^2 \end{bmatrix}.$$

This is an example of a *linear covariance structure*. If this structure holds, the four parameters can be solved in terms of the elements of $\boldsymbol{\Sigma}$. For example, $\sigma_a^2 = \frac{1}{2}\left(\sigma_{41}+\sigma_{42}\right)$, $\sigma_b^2 = \frac{1}{2}\left(\sigma_{41}+\sigma_{42}-\sigma_{21}-\sigma_{31}\right)$, $\sigma_c^2 = \frac{1}{2}\left(\sigma_{31}-\sigma_{42}\right)$, etc. There are many ways in which the four parameters can be solved in terms of the $\sigma$ 's. If the ten equations are consistent, however, all solutions are identical. In this case the parameters are *overidentified*.

Consider the covariance structure $\boldsymbol{\Sigma}$, generated by the first four rows of $\mathbf{A}$:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & & & \\ \sigma_a^2-\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & & \\ \sigma_a^2+\sigma_b^2+\sigma_c^2 & \sigma_a^2-\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 & \\ \sigma_a^2-\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2 & \sigma_a^2-\sigma_b^2-\sigma_c^2 & \sigma_a^2+\sigma_b^2+\sigma_c^2+\sigma_e^2 \end{bmatrix}.$$

In this case we can solve for $\sigma_a^2 = \frac{1}{2}(\sigma_{21} + \sigma_{31})$, say, but it is impossible to solve for $\sigma_b^2$ and $\sigma_c^2$ separately. This is an example of a *non-identified model* in which some parameters are identified and others are not. The reason for this is that the matrix $\mathbf{A}$ has rank 2 and not rank 3 as in the previous case.

Estimation of the variance components according to the model above gives

$$\sigma_a^2 = 3.52 \quad \sigma_b^2 = 14.23 \quad \sigma_c^2 = 27.45 \quad \sigma_e^2 = 22.56.$$

However, examination of the fit of the model to the data reveals that the fit is very poor: $\chi^2$ = 464.3 with 74 degrees of freedom. We shall therefore seek an alternative model that better accounts for the data. This is obtained by structuring the error component **e**.

There are six distinct experimental conditions among the twelve trials, each one repeated twice. Let $\tau_i$ ($i$ = 1, 2, …, 6) be random components associated with the experimental conditions. Then

$$x_{i\alpha} = \tau_i + e_{i\alpha},$$

where $\alpha$ = 1, 2 indexes the two replications. This simply means that one should allow the error variances to be different for different experimental conditions but still equal within replications of the same condition. An analysis according to this model gives $\chi^2$ = 311.2 with 69 degrees of freedom. the reduction in $\chi^2$ clearly indicates that the error variances depend on the experimental condition.

The command file for this analysis (**EX63.LIS** in the **LISREL Examples** folder) is:

```
The Rod and Frame Test
DA  NI = 12  NO = 107
CM  FI = EX63.COV
MO  NX = 12  NK = 3  LX = FI  PH = DI
MA  LX
1   1   1
1  -1  -1
1   1   1
1  -1  -1
1  -1   1
1   1  -1
1  -1   1
1   1  -1
1   1   0
1  -1   0
1   1   0
1  -1   0
```

```
EQ  TD(1)  TD(3)
EQ  TD(2)  TD(4)
EQ  TD(5)  TD(7)
EQ  TD(6)  TD(8)
EQ  TD(9)  TD(11)
EQ  TD(10) TD(12)
OU SE TV
```

The ML estimates of the variance components are now, with standard errors below the estimates:

```
PHI
         Note: This matrix is diagonal.

               KSI 1      KSI 2      KSI 3
            --------   --------   --------
               4.075     11.273     26.190
             (0.739)    (1.717)    (4.115)
               5.516      6.566      6.365


         THETA-DELTA

               VAR 1      VAR 2      VAR 3      VAR 4      VAR 5      VAR 6
            --------   --------   --------   --------   --------   --------
              22.034     37.632     22.034     37.632     28.573     40.503
             (2.680)    (4.177)    (2.680)    (4.177)    (3.275)    (4.609)
               8.221      9.010      8.221      9.010      8.724      8.788


         THETA-DELTA

               VAR 7      VAR 8      VAR 9     VAR 10     VAR 11     VAR 12
            --------   --------   --------   --------   --------   --------
              28.573     40.503     11.633      5.072     11.633      5.072
             (3.275)    (4.609)    (1.398)    (0.660)    (1.398)    (0.660)
               8.724      8.788      8.321      7.680      8.321      7.680
```

The results indicate that most of the variance in the trials is associated with the chair effect. The variance due to the frame effect is less than half of this and the variance due to general bias is still smaller. the error variances are generally quite large, except for the two experimental conditions in which the chair is already vertical.

Since the fit of the model is still not satisfactory one could allow the variance components to be correlated. However, an analysis with $\mathbf{\Phi}$ free reveals than none of the covariances in $\mathbf{\Phi}$ is significant.