

POSITIVE DEFINITENESS

The data matrix that you submit for analysis to an SEM computer tool should have the property that it is **positive definite** (PD), which is required for most estimation methods. A matrix that lacks this characteristic is **nonpositive definite** (NPD), and attempts to analyze such a data matrix will probably fail. A PD data matrix has the properties summarized next (Wothke, 1993):

1. The matrix is **nonsingular**, or invertible. In most kinds of multivariate analyses (SEM included), the computer needs to derive the inverse of the data matrix as part of linear algebra operations. A matrix that is not invertible is **singular**.

2. All eigenvalues of PD matrices are positive (> 0). An **eigenvalue** is the variance of an **eigenvector**, which is a linear combination of the observed variables where all the weights are not zero. An eigenvalue is the unstandardized proportion of variance explained by the corresponding eigenvector, and the variance of that composite (its eigenvalue) cannot logically be less than zero. The total number of pairs of eigenvalues and eigenvectors for a data matrix equals the number of observed variables. For example, if a covariance matrix is based on 10 variables, then there are a total of 10 eigenvalue-eigenvector pairs.

3. A related property is that the **determinant** of a PD matrix is greater than zero. If the determinant is zero, then the matrix is singular. A determinant equals the serial product (the first times the second times the third, and so on) of the eigenvalues, so if a determinant is negative, then some odd number of the eigenvalues (1 or 3 or 5, etc.) must be negative. A negative determinant indicates an NPD matrix.

4. In a PD data matrix, none of the correlations or covariances are **out of bounds**. An out-of-bounds matrix element is one that would be mathematically impossible to derive if all entries were calculated using data from the same cases. This property is explained next.

¹By the same token, it would also be problematic to convert raw scores to z scores and then submit for analysis the data file of standardized scores.

The value of the Pearson correlation between two variables X and Y is limited by the correlations between these variables and a third variable W . Specifically, the value of r_{XY} must fall within the following range:

$$(r_{XW} \times r_{YW}) \pm \sqrt{(1 - r_{XW}^2)(1 - r_{YW}^2)} \quad (3.1)$$

For example, if $r_{XW} = .60$ and $r_{YW} = .40$, then the value of r_{XY} must be within the range $.24 \pm .73$ (i.e., $-.49$ – $.97$). Any other value for r_{XY} would be out of bounds. Another way to view Equation 3.1 is that it specifies a **triangle inequality** for values of correlations among three variables measured in the same sample.²

In a PD data matrix, the maximum absolute value of cov_{XY} , the covariance between X and Y , must respect the upper limit defined next:

$$\max |cov_{XY}| \leq \sqrt{s_X^2 s_Y^2} \quad (3.2)$$

where s_X^2 and s_Y^2 are, respectively, the sample variances of X and Y . In words, the maximum absolute value for the covariance between any two variables is less than or equal to the square root of the product of their variances. Otherwise, the value of cov_{XY} is out of bounds. For example, given

$$cov_{XY} = 13.00, \quad s_X^2 = 12.00, \quad \text{and} \quad s_Y^2 = 10.00$$

then the covariance between X and Y would be out of bounds because

$$13.00 > (12.00 \times 10.00)^{1/2} = 10.95$$

which violates Equation 3.2. The value of r_{XY} for this example is also out of bounds because it equals 1.19. An exercise will ask you to verify this fact.

An NPD data matrix has at least one eigenvalue ≤ 0 . Many computer programs for multivariate statistical analyses, including those for SEM, print eigenvalues in the output, so this sign of trouble is apparent. An eigenvalue of zero indicates that the matrix is singular. A negative eigenvalue could indicate a few different problems. One is the presence of an out-of-bounds entry in the data matrix (i.e., Equations 3.1–3.2 do not hold). Another is perfect collinearity either between a pair of variables (e.g., $r_{XY} = 1.00$) or between a variable and at least two others (e.g., $R_{Y.XW} = 1.00$). It can also happen that near-perfect collinearity (e.g., $r_{XY} = .95$) manifested as positive but near-zero eigenvalues can cause matrix inversion operations to fail. It is easy to spot bivariate collinearity by inspecting the correlation matrix. A way to detect multivariate collinearity among three or more variables is described later in this chapter. See Topic Box 3.1 for more information about causes of nonpositive definiteness in the data matrix and possible solutions.

²In a geometric triangle, the length of a given side must be less than the sum of the lengths of the other two sides but greater than the difference between the lengths of the two sides.

TOPIC BOX 3.1

Causes of Nonpositive Definiteness and Solutions

Many points summarized here are from Wothke (1993).^{*} Some causes of non-positive definite (NPD) data matrices are listed next. Most can be detected through careful data screening:

1. Extreme bivariate or multivariate collinearity among the observed variables.
2. The presence of outliers, especially those that force values of correlations to be extremely high.
3. Pairwise deletion of cases with missing data.
4. Making a typing mistake when transcribing a data matrix from one source, such as a table in a journal article, to another, such as a command file for computer analysis, can also result in an NPD matrix. For example, if the value of a covariance in the original matrix is 15.00, then mistakenly typing 150.00 in the transcribed matrix could generate an NPD covariance matrix with elements that violate Equation 3.2. It is so easy to make a typing mistake during manual entry of a data matrix that errors are almost guaranteed, especially when the number of variables exceeds 10 or so. Follow this simple but effective advice from Wilkinson and the Task Force on Statistical Inference (1999) whenever you transcribe a data matrix: *look at the data*, that is, carefully compare, entry by entry, the original data matrix with your transcribed matrix before you attempt to analyze it with the computer.
5. Plain old sampling error can generate NPD data matrices, especially if the number of cases is relatively small or the sample is unrepresentative. The former condition can be addressed by increasing the sample size; the unrepresentativeness may be the result of using a sampling method that selects atypical cases.
6. Sometimes matrices of *estimated* Pearson correlations, such as polyserial or polychoric correlations derived for noncontinuous observed variables (Chapter 2), are NPD. This may be especially true if polyserial or polychoric correlations are estimated in a pairwise manner instead of simultaneously estimating the whole correlation matrix. Pairwise calculation of non-Pearson correlations is an older method that required less computer memory, but this goal is less relevant given today's personal computers with relatively large memory capacities. Modern computer tools, such as the PRELIS program of LISREL, can simultaneously estimate the whole correlation matrix.

^{*}See also E. Rigdon's webpage on nonpositive definite matrices in SEM at www2.gsu.edu/~mkteer/npdmatri.html

Here is a tip about diagnosing whether a data matrix is positive definite before submitting it for analysis to an SEM computer program: Copy the full matrix (with redundant entries above and below the diagonal) into a text (ASCII) file, such as Microsoft Windows Notepad. Next, point your Internet browser to a free, online matrix calculator and then copy the data matrix into the proper window on the calculating webpage.* Finally, select options on the webpage to derive the determinant and eigenvalues of the data matrix. Look for outcomes that indicate nonpositive definiteness, such as near-zero, zero, or negative eigenvalues.

Some SEM computer programs, such as LISREL, offer options for making a **ridge adjustment** to an NPD data matrix. The ridge technique iteratively multiplies the diagonal entries of the matrix by a constant > 1.0 until negative eigenvalues disappear (the matrix becomes positive definite). For covariance matrices, ridge adjustments increase the values of the variances until they are large enough to exceed any out-of-bounds covariance entry in the off-diagonal part of the matrix (Equation 3.2 will be satisfied). This technique "fixes up" a data matrix so that necessary algebraic operations can be performed (Wothke, 1993). However, the resulting parameter estimates, standard errors, and model fit statistics will be biased after applying a ridge correction. For this reason, I do not recommend that you use a ridge technique to analyze an NPD data matrix unless you are very familiar with linear algebra (i.e., you know what you are doing and why). Instead, you should try to solve the problem of nonpositive definiteness through data screening or increasing the sample size.

There are other contexts where you may encounter NPD matrices in SEM, but these generally concern (1) matrices of parameter estimates for your model or (2) matrices of covariances or correlations predicted from your model that could be compared with those observed in your sample. A problem in the analysis is indicated if any of these matrices is NPD. We will deal with these contexts in later chapters.

*www.bluebit.gr/matrix-calculator/