



*Tutorials and worked examples for simulation, curve fitting, statistical analysis, and plotting.*

<https://simfit.org.uk>

<https://simfit.silverfrost.com>

Procrustes analysis is useful when there are two matrices  $X$  and  $Y$  with the same dimensions, and it wished to see how closely the  $X$  matrix can be made to fit the target matrix  $Y$  using only distance preserving transformations, like translation and rotation. For instance,  $X$  could be a matrix of loadings, and the target matrix  $Y$  could be a reference matrix of loadings from another data set.

From the main SIMFIT menu choose [Statistics], [Multivariate], then [Procrustes] and analyze the following two matrices

$$X = \begin{bmatrix} 0.63 & 0.58 \\ 1.36 & 0.39 \\ 1.01 & 1.76 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0.00 & 0.00 \\ 1.00 & 0.00 \\ 0.00 & 2.00 \end{bmatrix}$$

contained in the default test files g03bcf.tf1 with  $X$ -data to be rotated, and g03bcf.tf2 containing the target matrix  $Y$ , to obtain the following results.

**Results from Procrustes analysis**

$X$ -data for rotation: g03bcf.tf1

$Y$ -data for target: g03bcf.tf2

Number of rows: 3

Number of columns: 2

Type: To origin then  $Y$ -centroid

Scaling: Least squares scaling

$\alpha = 1.5563$

Residual sum of squares = 0.019098

**Residuals from Procrustes rotation**

0.09644

0.08455

0.05145

**Rotation matrix from Procrustes rotation**

0.9673            0.2536

-0.2536           0.9673

**$Y$ -hat matrix from Procrustes rotation**

-0.0934           0.0239

1.0805            0.0259

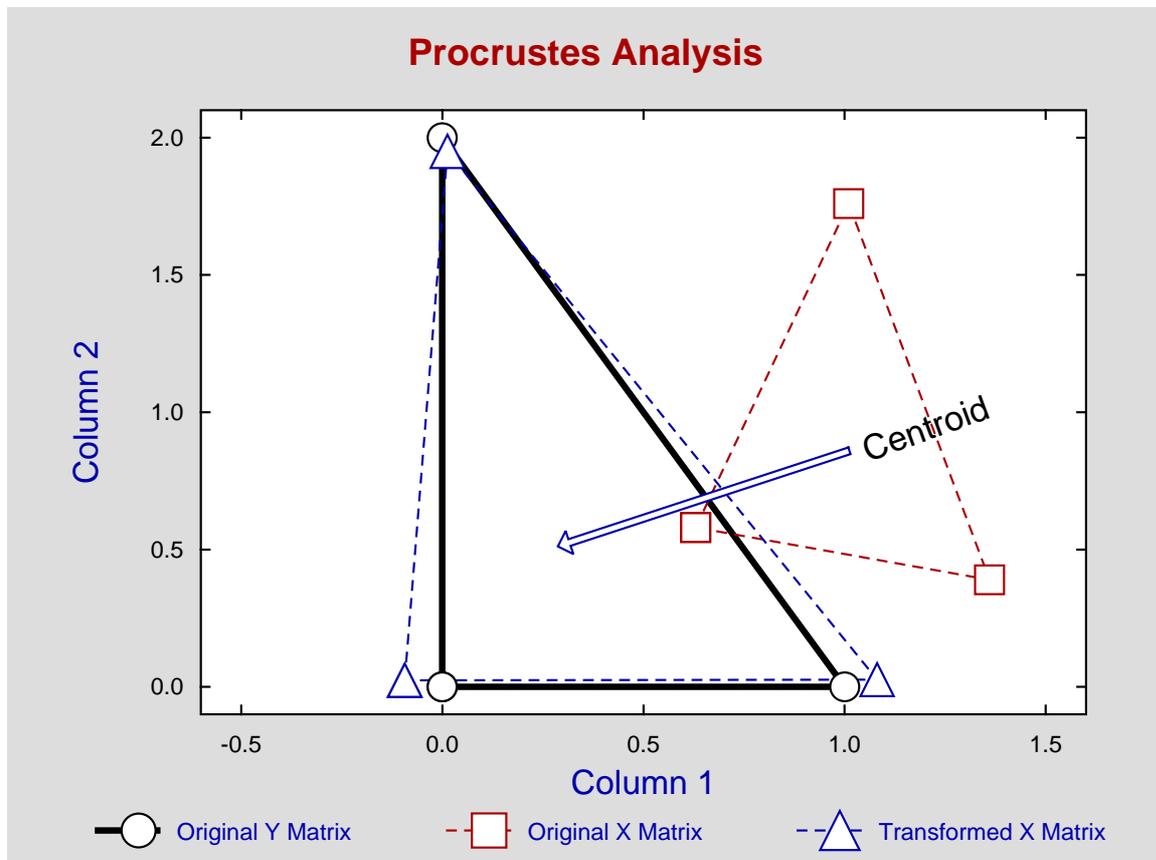
0.0130            1.9502

There are numerous options for performing this calculation including the following alternative types

- No translation or normalization
- Translation to the origin
- Translation to the origin then to the  $Y$  centroid after rotation
- Unit normalization
- Translation and normalization (i.e. standardization)

with or without least squares scaling after the rotation.

The following diagram shows the result from transforming the original  $X$  matrix which is initially distant from the target  $Y$  matrix into the transformed  $X$  matrix  $\hat{Y}$  which has been brought almost into coincidence with the target  $Y$  matrix by movement of the centroid, rotating and scaling.



Also, as well as displaying the residuals, the sum of squares, the rotation and best fit matrices, options are provided to plot arbitrary rows or columns of these matrices.

## Theory

First the centroids of  $X$  and  $Y$  are translated to the origin to give  $X_c$  and  $Y_c$ . Then the matrix of rotations  $R$  that minimize the sum of squared residuals is found from the singular value decomposition as

$$X_c^T Y_c = U D V^T$$

$$R = U V^T,$$

and after rotation a dilation factor  $\alpha$  can be estimated by least squares, if required, to give the estimate

$$\hat{Y}_c = \alpha X_c R.$$

Additional options from the SIMFIT Procrustes interface include normalizing both matrices to have unit sums of squares, normalizing the  $X$  matrix to have the same sum of squares as the  $Y$  matrix, and translating to the original  $Y$  centroid after rotation. Note that these Procrustes options can often be done interactively in SIMFIT whenever loadings are calculated.