Title

---

procrustes — Procrustes transformation

Syntax

procrustes (*varlist$_y$*) (*varlist$_x$*) [*if*] [*in*] [*weight*] [, *options*]

options	Description
Model	
transform(<u>or</u>thogonal)	orthogonal rotation and reflection transformation; the default
transform(<u>ob</u>lique)	oblique rotation transformation
transform(<u>un</u>restricted)	unrestricted transformation
<u>nocon</u>stant	suppress the constant
<u>norho</u>	suppress the dilation factor ρ (set $\rho = 1$)
force	allow overlap and duplicates in *varlist$_y$* and *varlist$_x$* (advanced)
Reporting	
<u>nofit</u>	suppress table of fit statistics by target variable

bootstrap, by, jackknife, and statsby are allowed; see **[U] 11.1.10 Prefix commands**.
Weights are not allowed with the bootstrap prefix; see **[R] bootstrap**.
aweights are not allowed with the jackknife prefix; see **[R] jackknife**.
aweights and fweights are allowed; see **[U] 11.1.6 weight**.
See **[U] 20 Estimation and postestimation commands** for more capabilities of estimation commands.

Menu

Statistics > Multivariate analysis > Procrustes transformations

Description

procrustes performs the Procrustean analysis, one of the standard methods of multidimensional scaling. For given "target" variables *varlist$_y$* and "source" variables *varlist$_x$*, the goal is to transform the source **X** to be as close as possible to the target **Y**. The permitted transformations are any combination of dilation (uniform scaling), rotation and reflection (that is, orthogonal or oblique transformations), and translation. Closeness is measured by the residual sum of squares. procrustes deals with complete cases only.

procrustes assumes equal weights or scaling for the dimensions. It would be inappropriate, for example, to have the first variable measured in grams ranging from 5,000 to 8,000, the second variable measured in dollars ranging from 3 to 12, and the third variable measured in meters ranging from 100 to 280. In such cases, you would want to operate on standardized variables.

Options

Model

transform(*transform*) specifies the transformation method. The following transformation methods
   are allowed:

   orthogonal specifies that the linear transformation matrix $\mathbf{A}$ should be orthogonal, $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}$. This is the default.

   oblique specifies that the linear transformation matrix $\mathbf{A}$ should be oblique, $\text{diag}(\mathbf{A}\mathbf{A}') = \mathbf{1}$.

   unrestricted applies no restrictions to $\mathbf{A}$, making the procrustes transformation equivalent to
   multivariate regression with uncorrelated errors; see [MV] **mvreg**.

noconstant specifies that the translation component $\mathbf{c}$ is fixed at $\mathbf{0}$ (the 0 vector).

norho specifies that the dilation (scaling) constant ρ is fixed at 1. This option is not relevant with
   transform(unrestricted); here ρ is always fixed at 1.

force, an advanced option, allows overlap and duplicates in the target variables *varlist*$_y$ and source
   variables *varlist*$_x$.

Reporting

nofit suppresses the table of fit statistics per target variable. This option may be specified during
   estimation and upon replay.

Remarks and examples

stata.com

Remarks are presented under the following headings:

> Introduction to Procrustes methods
> Orthogonal Procrustes analysis
> Is an orthogonal Procrustes analysis symmetric?
> Other transformations

Introduction to Procrustes methods

The name *Procrustes analysis* was applied to optimal matching of configurations by Hurley and
Cattell (1962) and refers to Greek mythology. The following account follows Cox and Cox (2001,
123). Travelers from Eleusis to Athens were kindly invited by Damastes to spend the night at his place.
Damastes, however, practiced a queer kind of hospitality. If guests would not fit the bed, Damastes
would either stretch them to make them fit, or chop off extremities if they were too long. Therefore,
he was given the nickname Procrustes—ancient Greek for "the stretcher". Theseus, a warrior, finally
gave Procrustes some of his own medicine.

Procrustes methods have been applied in many areas. Gower and Dijksterhuis (2004) mention
applications in psychometrics (for example, the matching of factor loading matrices), image analysis,
market research, molecular biology, biometric identification, and shape analysis.

Formally, procrustes solves the minimization problem

$$\text{Minimize} \, | \, \mathbf{Y} - (\mathbf{1}\mathbf{c}' + \rho \, \mathbf{X} \, \mathbf{A}) \, |$$

where $\mathbf{c}$ is a row vector representing the translation, ρ is the scalar "dilation factor", $\mathbf{A}$ is the rotation
and reflection matrix (orthogonal, oblique, or unrestricted), and $|.|$ denotes the L2 norm.

Some of the early work on Procrustes analysis was done by Mosier (1939), Green (1952), Hurley
and Cattell (1962), and Browne (1967); see Gower and Dijksterhuis (2004).

Orthogonal Procrustes analysis

▷ Example 1

   We illustrate `procrustes` with John Speed's historical 1610 map of the Worcestershire region in England, engraved and printed by Jodocus Hondius in Amsterdam in 1611–1612. Used with permission of Peen (2007).
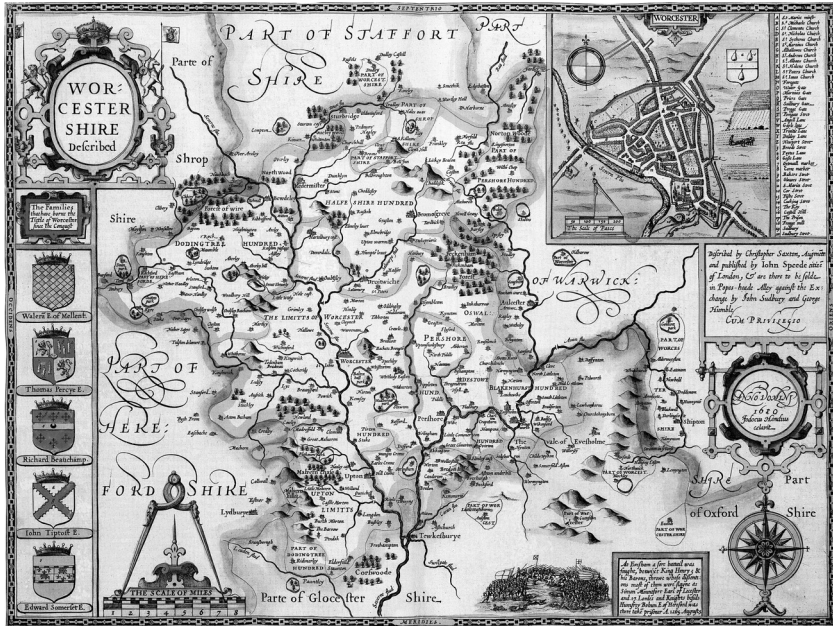
   We analyze the accuracy of this map. Cox and Cox (2001) present data on the locations of 20 towns and villages on this old map, as well as the locations on a modern map from the Landranger Series of Ordnance Survey Maps. The locations were measured relative to the lower-left corner of the maps. We list this small dataset, specifying the `noobs` option to prevent wrapping and `sep(0)` to suppress internal horizontal lines.

```
. use http://www.stata-press.com/data/r13/speed_survey
(Data on Speed's Worcestershire map (1610))

. list name lname speed_x speed_y survey_x survey_y, sep(0) noobs
```

name	lname	speed_x	speed_y	survey_x	survey_y
Alve	Alvechurch	192	211	1027	725
Arro	Arrow	217	155	1083	565
Astl	Astley	88	180	787	677
Beck	Beckford	193	66	976	358
Beng	Bengeworth	220	99	1045	435
Crad	Cradley	79	93	736	471
Droi	Droitwich	136	171	893	633
Ecki	Eckington	169	81	922	414
Eves	Evesham	211	105	1037	437
Hall	Hallow	113	142	828	579
Hanb	Hanbury	162	180	944	637
Inkb	Inkberrow	188	156	1016	573
Kemp	Kempsey	128	108	848	490
Kidd	Kidderminster	104	220	826	762
Mart	Martley	78	145	756	598
Stud	Studley	212	185	1074	632
Tewk	Tewkesbury	163	40	891	324
UpSn	UpperSnodsbury	163	138	943	544
Upto	Upton	138	71	852	403
Worc	Worcester	125	132	850	545

You will probably conclude immediately that the scales of the two maps differ and that the coordinates are expressed with respect to different reference points; the lower-left corners of the maps correspond to different physical locations. Another distinction will not be so obvious—at least not by looking at these numbers: the orientations of the maps may well differ. We display as scatterplots Speed's data (speed_x, speed_y) and the modern survey data (survey_x, survey_y).

```
. scatter speed_y speed_x, mlabel(name)
> ytitle("") xtitle("") yscale(off) xscale(off) ylabel(,nogrid)
> title(Historic map of 20 towns and villages in Worcestershire)
> subtitle((Speed 1610))
```


Historic map of 20 towns and villages in Worcestershire
(Speed 1610)
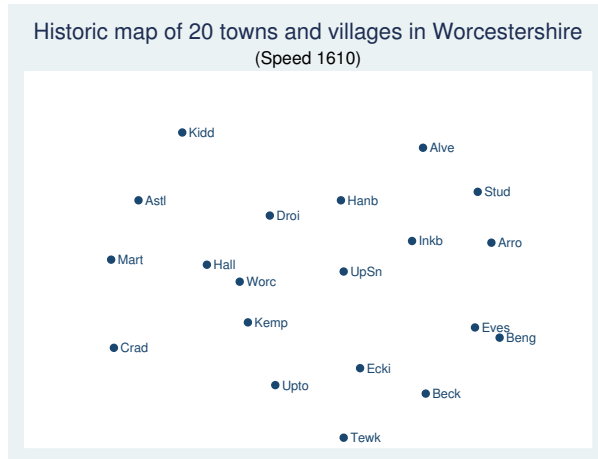
```
. scatter survey_y survey_x, mlabel(name)
> ytitle("") xtitle("") yscale(off) xscale(off) ylabel(,nogrid)
> title(Modern map of 20 towns and villages in Worcestershire)
> subtitle((Landranger series of Ordnance Survey Maps))
```


Modern map of 20 towns and villages in Worcestershire
(Landranger series of Ordnance Survey Maps)
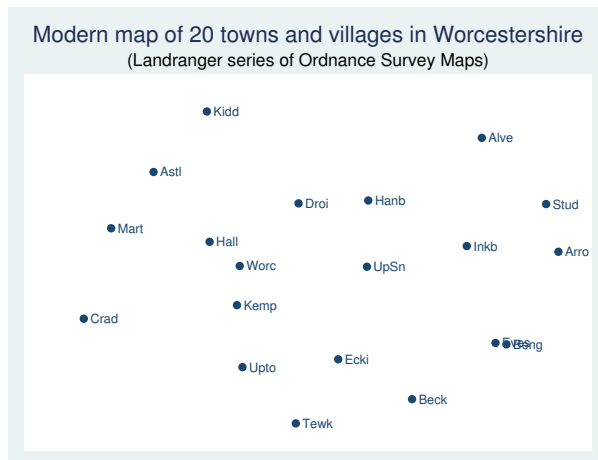
To gauge the accuracy of the historic map, we must account for differences in scale, location, and orientation. Because the area depicted on the map is relatively small, we think that it is justified to ignore the curvature of the earth and approximate distances along the globe with Euclidean distances, and apply a Procrustes analysis,

$$\text{survey\_map} = \text{transformation}(\text{speed\_map}) + \text{residual}$$

choosing the transformation (from among the allowed transformations) to minimize the residual in terms of the residual sum of squares. The transformation should allow for, in mathematical terms,

translation, uniform scaling, and two-dimensional orthogonal rotation. The uniform scaling factor is often described as the dilation factor, a positive scalar. The transformation from source to target configuration can be written as

$$(\text{survey\_x} \quad \text{survey\_y}) = (c_x \quad c_y) + \rho (\text{speed\_x} \quad \text{speed\_y}) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + (\text{res\_x} \quad \text{res\_y})$$

or simply as

$$\text{survey\_map} = \text{translation} + \text{dilation} \times \text{speed\_map} \times \text{rotation} + \text{residual}$$

The matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

should satisfy the constraint that it represents an orthogonal rotation—it should maintain the lengths of vectors and the angles between vectors. We estimate the translation $(c_x \quad c_y)$, dilation factor ρ, and the rotation matrix $\mathbf{A}$ with the procrustes command.

```
. procrustes (survey_x survey_y) (speed_x speed_y)
Procrustes analysis (orthogonal)          Number of observations =         20
                                          Model df (df_m)        =          4
                                          Residual df (df_r)     =         36
                                          SS(target)             =     495070
                                          RSS(target)            =   1973.384
                                          RMSE = root(RSS/df_r)  =   7.403797
                                          Procrustes = RSS/SS    =     0.0040
```

Translation c

	survey_x	survey_y
_cons	503.8667	293.9878

Rotation & reflection matrix A (orthogonal)

	survey_x	survey_y
speed_x	.9841521	-.1773266
speed_y	.1773266	.9841521

Dilation factor

```
        rho =    2.3556
```

Fit statistics by target variable

Statistics	survey_x	survey_y
SS	216310.2	278759.8
RSS	1081.36	892.0242
RMSE	7.750841	7.039666
Procrustes	.0049991	.0032
Corr_y_yhat	.9976669	.9985076

We can read the elements of the transformation from the output: the translation from the Speed map onto the survey map is (504, 294). The scale of the survey and Speed maps differ by a factor of 2.36. The orientations of the maps also differ somewhat; if the maps had been oriented the same, we would have expected the rotation to be an identity matrix. Note that $.984^2 + .177^2 = 1$, subject to rounding error—indeed the rotation is "norm preserving". A counterclockwise rotation in a plane over θ can be written as

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

See appendix B in Gower and Dijksterhuis (2004). Here $\cos(\theta) = 0.984$, so the difference in orientation of the two maps is $\theta = 10.2$ degrees.

The other output produced by procrustes may be more familiar. procrustes estimated four parameters: the angle of rotation, two translation parameters, and the dilation factor ρ. SS(target) is the centered sum of squares of the survey data, and is meaningful mostly in relation to the residual sum of squares RSS(target). The Procrustes statistic, defined as RSS/SS, measures the size of the residuals relative to the variation in the target variables; it is equivalent to $1 - R^2$ in a regression analysis context. The smaller the Procrustes statistic, the closer the correspondence of Speed's map to the survey map. The number in this case, 0.004, is small indeed. Another way of looking at fit is via the square root of the mean squared residual error, RMSE, a measure for the average size of residuals.

The last output table describes how well the transformed Speed coordinates match the survey coordinates, separately for the horizontal (x) and the vertical (y) coordinates. In this case, we do not see disturbing differences between the coordinates. By definition, the overall Procrustes statistic and the overall RMSE are averages of the coordinate statistics. Because Procrustes analysis treats (weights) both coordinates the same and independently, analogous to the sphericity assumption in multivariate regression or MANOVA analysis, the comparable statistics for the different coordinates is reassuring.

This example is continued in [MV] **procrustes postestimation**, demonstrating how to generate fitted values and residual sum of squares with predict, how to produce a graph showing the target overlaid with the transformed source values with procoverlay, and how to produce various summaries and comparisons with estat.

◁

A Procrustes analysis fits the transformation subject to the constraint that **A** is orthogonal; for other constraints, see below. In two dimensions, there are actually two types of orthogonal matrices: rotations and reflections. Think of left and right hands. A rotation transforms a left hand into a left hand, never into a right hand; rotation preserves orientation. A reflection changes a left hand into a right hand; reflections invert orientation. In algebraic terms, an orthogonal matrix **A** satisfies $\det(\mathbf{A}) = \pm 1$. **A** is a rotation if $\det(\mathbf{A}) = 1$, and **A** is a reflection if $\det(\mathbf{A}) = -1$. In more than two dimensions, the classification of orthogonal transformations is more complicated.

▷ Example 2

In example 1, we treated the location, dilation, and orientation as estimable aspects of the transformation. It is possible to omit the location and dilation aspects—though, admittedly, from a casual inspection as well as the substantial understanding of the data, these aspects are crucial. For instance, we may omit the dilation factor—that is, assume $\rho = 1$—with the norho option.

```
. procrustes (survey_x survey_y) (speed_x speed_y), norho
```

Procrustes analysis (orthogonal)

Number of observations	=	20
Model df (df_m)	=	3
Residual df (df_r)	=	37
SS(target)	=	495070
RSS(target)	=	165278.1
RMSE = root(RSS/df_r)	=	66.83544
Procrustes = RSS/SS	=	0.3338

Translation c

	survey_x	survey_y
_cons	741.4458	435.6215

Rotation & reflection matrix A (orthogonal)

	survey_x	survey_y
speed_x	.9841521	-.1773266
speed_y	.1773266	.9841521

Dilation factor

           rho =   1.0000

Fit statistics by target variable

Statistics	survey_x	survey_y
SS	216310.2	278759.8
RSS	70385.78	94892.36
RMSE	61.68174	71.61925
Procrustes	.3253928	.340409
Corr_y_yhat	.9976669	.9985076

As expected, the optimal transformation without dilation implies a much weaker relation between the Speed and Survey maps; the Procrustes statistic has increased from 0.0040 to 0.3338. We conclude that we cannot adequately describe the correspondence between the maps if we ignore differences in scale.

◁

Is an orthogonal Procrustes analysis symmetric?

In examples 1 and 2, we transformed the Speed map to optimally match the modern Survey map. We could also have reversed the procedure, that is, transform the Survey map to match the Speed map.

▷ Example 3

Here we change the order of the Speed and Survey map in our call to `procrustes` from example 1.

```
. procrustes (speed_x speed_y) (survey_x survey_y)
```

Procrustes analysis (orthogonal)

Number of observations	=	20
Model df (df_m)	=	4
Residual df (df_r)	=	36
SS(target)	=	88862.75
RSS(target)	=	354.2132
RMSE = root(RSS/df_r)	=	3.136759
Procrustes = RSS/SS	=	0.0040

Translation c

	speed_x	speed_y
_cons	-187.0142	-159.5801

Rotation & reflection matrix A (orthogonal)

	speed_x	speed_y
survey_x	.9841521	.1773266
survey_y	-.1773266	.9841521

Dilation factor

rho =   0.4228

Fit statistics by target variable

Statistics	speed_x	speed_y
SS	41544.95	47317.8
RSS	218.3815	135.8317
RMSE	3.483146	2.747036
Procrustes	.0052565	.0028706
Corr_y_yhat	.9975074	.9986641

The implied transformations are similar but not identical. For instance, the product of estimated scale factors is $2.3556 \times 0.4228 = 0.9959$, which is close to 1 but not identical to 1—this is not due to roundoff error. Why do the results differ? Think about the analogy with regression analysis. The regression of Y on X and the regression of X on Y generally imply different relationships between the variables. In geometric terms, one minimizes the sum of squares of the "vertical" distances between the data point and the regression line, whereas the other minimizes the "horizontal" distances. The implied regression lines are the same if the variance in X and Y are the same. Even if this does not hold, the proportion of explained variance R^2 in both regressions is the same. In Procrustes analysis, an analogous relationship holds between the analyses "Speed = transformed(Survey) + E" and "Survey = transformed(Speed) + E". Both analyses yield the same Procrustes statistic. The implied analyses are equivalent (that is, the implied transformation in one analysis is the mathematical inverse of the transformation in the other analysis) only if the Speed and Survey data are scaled so that the trace of the associated covariance matrices is the same.

◁

Other transformations

A Procrustes analysis can also be applied with other classes of transformations. Browne (1967) analyzed Procrustes analyses with oblique rotations. Cramer (1974) and ten Berge and Nevels (1977) identified and solved some problems in Browne's solution (but still ignore the problem that the derived oblique rotations are not necessarily orientation preserving). `procrustes` supports oblique transformations. `procrustes` also allows dilation; see *Methods and formulas*.

▷ Example 4

Even though the orthogonal Procrustes analysis of example 1 demonstrated a similarity between the two configurations assuming an orthogonal transformation, we now investigate what happens with an oblique transformation.

```
. procrustes (survey_x survey_y) (speed_x speed_y), trans(oblique)
```

Procrustes analysis (oblique)

Number of observations =	20
Model df (df_m) =	5
Residual df (df_r) =	35
SS(target) =	495070
RSS(target) =	1967.854
RMSE = root(RSS/df_r) =	7.498294
Procrustes = RSS/SS =	0.0040

Translation c

	survey_x	survey_y
_cons	503.0093	292.4346

Rotation & reflection matrix A (oblique)

	survey_x	survey_y
speed_x	.9835969	-.1737553
speed_y	.1803803	.9847889

Dilation factor

           rho =   2.3562

Fit statistics by target variable

Statistics	survey_x	survey_y
SS	216310.2	278759.8
RSS	1080.677	887.1769
RMSE	7.858307	7.1201
Procrustes	.004996	.0031826
Corr_y_yhat	.9976685	.9985163

We see that the optimal oblique transformation is almost orthogonal; the columns of the oblique rotation and reflection matrix are almost perpendicular. The dilation factor and translation vector hardly differ from the orthogonal case shown in example 1. Finally, we see that the residual sum of squares decreased little, namely, from 1,973.4 to 1,967.9.

◁

Procrustes analysis can be interpreted as multivariate regression $\mathbf{Y} = \mathbf{c} + \mathbf{x}\mathbf{B} + \mathbf{e}$ in which some nonlinear restriction is applied to the coefficients $\mathbf{B}$. Procrustes analysis assumes $\mathbf{B} = \rho\mathbf{A}$ with $\mathbf{A}$

assumed to be orthogonal or **A** assumed to be oblique. The intercepts of the multivariate regression are, of course, the translation of the Procrustean transform. In contrast to multivariate regression, it is assumed that the distribution of the residuals **e** is spherical; that is, all that is assumed is that $\mathrm{var}(\mathbf{e}) = \sigma^2 \mathbf{I}$. This assumption affects standard errors, not the estimated coefficients. Multivariate regression serves as a useful baseline to gauge the extent to which the Procrustean analysis is appropriate. `procrustes` supports the `transform(unrestricted)` option and displays the fitted model in a format comparable to Procrustes analysis.

▷ Example 5

We demonstrate with Speed's map data.

```
. procrustes (survey_x survey_y) (speed_x speed_y), trans(unrestricted)
```

Procrustes analysis (unrestricted)	Number of observations	=	20
	Model df (df_m)	=	6
	Residual df (df_r)	=	34
	SS(target)	=	495070
	RSS(target)	=	1833.435
	RMSE = root(RSS/df_r)	=	7.343334
	Procrustes = RSS/SS	=	0.0037

Translation c

	survey_x	survey_y
_cons	510.8028	288.243

Rotation & reflection matrix A (unrestricted)

	survey_x	survey_y
speed_x	2.27584	-.4129564
speed_y	.4147244	2.355725

Fit statistics by target variable

Statistics	survey_x	survey_y
SS	216310.2	278759.8
RSS	1007.14	826.2953
RMSE	7.696981	6.971772
Procrustes	.004656	.0029642
Corr_y_yhat	.9976693	.9985168

Because we already saw that there is almost no room to improve on the orthogonal Procrustes transform with this particular dataset, dropping the restrictions on the coefficients hardly improves the fit. For instance, the residual sum of squares further decreases from 1,967.9 in the oblique case to 1,833.4 in the unrestricted case, with only a small reduction in the value of the Procrustes statistic.

◁

Stored results

procrustes stores the following in e():

Scalars
e(N)	number of observations
e(rho)	dilation factor
e(P)	Procrustes statistic
e(ss)	total sum of squares, summed over all y variables
e(rss)	residual sum of squares, summed over all y variables
e(rmse)	root mean squared error
e(urmse)	root mean squared error (unadjusted for # of estimated parameters)
e(df_m)	model degrees of freedom
e(df_r)	residual degrees of freedom
e(ny)	number of y variables (target variables)

Macros
e(cmd)	procrustes
e(cmdline)	command as typed
e(ylist)	y variables (target variables)
e(xlist)	x variables (source variables)
e(transform)	orthogonal, oblique, or unrestricted
e(uniqueA)	1 if rotation is unique, 0 otherwise
e(wtype)	weight type
e(wexp)	weight expression
e(properties)	nob noV
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins

Matrices
e(c)	translation vector
e(A)	orthogonal transformation matrix
e(ystats)	matrix containing fit statistics

Functions
e(sample)	marks estimation sample

Methods and formulas

Methods and formulas are presented under the following headings:

> Introduction
> Orthogonal transformations
> Oblique transformations
> Unrestricted transformations
> Reported statistics

Introduction

A Procrustes analysis is accomplished by solving a matrix minimization problem

$$\text{Minimize} \, | \, \mathbf{Y} - (\mathbf{1}\mathbf{c}' + \rho \, \mathbf{X} \, \mathbf{A}) \, |$$

with respect to $\mathbf{A}$, $\mathbf{c}$, and ρ. $\mathbf{A}$ is a matrix representing a linear transformation, $\rho > 0$ is a scalar called the "dilation factor", $\mathbf{c}$ is a translation (row-) vector, and $|.|$ is the Frobenius (or L2) norm. Three classes of transformations are available in procrustes: orthogonal, oblique, and unrestricted. The orthogonal class consists of all orthonormal matrices $\mathbf{A}$, that is, all square matrices that satisfy $\mathbf{A}'\mathbf{A} = \mathbf{I}$, representing orthogonal norm-preserving rotations and reflections. The oblique class comprises all normal matrices $\mathbf{A}$, characterized by $\text{diag}(\mathbf{A}'\mathbf{A}) = \mathbf{1}$. Oblique transformations preserve the length of vectors but not the angles between vectors—orthogonal vectors will generally not remain orthogonal under oblique transformations. Finally, the unrestricted class consists of all conformable regular matrices $\mathbf{A}$.

Define $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{X}}$ as the centered $\mathbf{Y}$ and $\mathbf{X}$, respectively, if a constant $\mathbf{c}$ is included in the analysis and as the uncentered $\mathbf{Y}$ and $\mathbf{X}$ otherwise.

The derivation of the optimal $\mathbf{A}$ obviously differs for the three classes of transformations.

Orthogonal transformations

The solution for the orthonormal case can be expressed in terms of the singular value decomposition of $\widetilde{\mathbf{Y}}'\widetilde{\mathbf{X}}$,

$$\widetilde{\mathbf{Y}}'\widetilde{\mathbf{X}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}'$$

where $\mathbf{U}'\mathbf{U} = \mathbf{V}'\mathbf{V} = \mathbf{I}$. Then

$$\widehat{\mathbf{A}} = \mathbf{V}\mathbf{U}'$$

$\widehat{\mathbf{A}}$ is the same whether or not scaling is required, that is, whether ρ is a free parameter or a fixed parameter. When ρ is a free parameter, the optimal ρ is

$$\widehat{\rho} = \frac{\text{trace}(\widehat{\mathbf{A}}\widetilde{\mathbf{Y}}'\widetilde{\mathbf{X}})}{\text{trace}(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})}$$

See ten Berge (1977) for a modern and elementary derivation; see Mardia, Kent, and Bibby (1979) for a derivation using matrix differential calculus.

Oblique transformations

Improving on earlier studies by Browne (1967) and Cramer (1974), ten Berge and Nevels (1977) provide a full algorithm to compute the optimal oblique rotation without dilation, that is, with uniform scaling $\rho = 1$. In contrast to the orthogonal case, the optimal oblique rotation $\widehat{\mathbf{A}}$ depends on ρ. To the best of our knowledge, this case has not been treated in the literature (J. M. F. ten Berge, 2004, pers. comm.). However, an "alternating least squares" extension of the ten Berge and Nevels (1977) algorithm deals with this case.

For each iteration, step (a) follows ten Berge and Nevels (1977) for calculating $\widetilde{\mathbf{Y}}$ and $\widehat{\rho}\widetilde{\mathbf{X}}$. In step (b) of an iteration, ρ is optimized, keeping $\widehat{\mathbf{A}}$ fixed, with solution

$$\widehat{\rho} = \frac{\text{trace}(\widehat{\mathbf{A}}\widetilde{\mathbf{Y}}'\widetilde{\mathbf{X}})}{\text{trace}(\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}')}$$

Iteration continues while the relative decrease in the residual sum of squares is large enough. This algorithm is ensured to yield a local optimum of the residual sum of squares as the RSS decreases both when updating the rotation $\mathbf{A}$ and when updating the dilation factor ρ. Beware that the algorithm is not guaranteed to find the global minimum.

Unrestricted transformations

In the unrestricted solution, the dilation factor ρ is fixed at 1. The computation of the Procrustes transformation is obviously equivalent to the least-squares solution of multivariate regression

$$\widehat{\mathbf{A}} = (\widetilde{\mathbf{X}}'\widetilde{\mathbf{X}})^{-1}\widetilde{\mathbf{X}}'\widetilde{\mathbf{Y}}$$

Given $\widehat{\mathbf{A}}$ and $\widehat{\rho}$, the optimal translation $\widehat{\mathbf{c}}$ can be written as

$$\widehat{\mathbf{c}} = \mathbf{Y}'\mathbf{1} - \widehat{\rho}\widehat{\mathbf{A}}\mathbf{X}$$

If the constant is suppressed, $\mathbf{c}$ is simply set to $\mathbf{0}$.

Reported statistics

`procrustes` computes and displays the following statistics for each target variable separately and globally by adding the appropriate sums of squares over all target variables. The predicted values $\widehat{\mathbf{Y}}$ for $\mathbf{Y}$ are defined as

$$\widehat{\mathbf{Y}} = \mathbf{1}\widehat{\mathbf{c}}' + \widehat{\rho}\mathbf{X}\widehat{\mathbf{A}}$$

The Procrustes statistic, P, is a scaled version of the squared distance of $\mathbf{Y}$:

$$P = \text{RSS}/\text{SS}$$

where

$$\text{RSS} = \text{trace}((\mathbf{Y} - \widehat{\mathbf{Y}})(\mathbf{Y} - \widehat{\mathbf{Y}})')$$

$$\text{SS} = \text{trace}(\widetilde{\mathbf{Y}}'\widetilde{\mathbf{Y}})$$

Note that $0 \leq P \leq 1$, and a small value of P means that $\mathbf{Y}$ is close to the transformed value of $\mathbf{X}$, that is, the $\mathbf{X}$ and $\mathbf{Y}$ configurations are similar. In the literature, this statistic is often denoted by R^2. It is easy to confuse this with the R^2 statistic in a regression context, which is actually $1 - P$.

A measure for the size of the residuals is the root mean squared error,

$$\text{RMSE} = \sqrt{\text{RSS}/\text{df}_{\text{r}}}$$

Here df_{r} are $Nn_y - \text{df}_{\text{m}}$, with $\text{df}_{\text{m}} = n_y n_x + n_y + 1 - k$, and with N the number of observations, n_y and n_x the number of target variables and source variables, respectively, and k, the number of restrictions, defined as

$$
\begin{array}{ll}
\text{orthogonal:} & k = n_x(n_x - 1)/2 \\
\text{oblique:} & k = n_y \\
\text{unrestricted:} & k = 1
\end{array}
$$

`procrustes` computes RMSE(j) for target variable y_j as

$$\text{RMSE}(j) = \sqrt{\text{RSS}(j)/(\text{df}_{\text{r}}/n_y)}$$

Finally, `procrustes` computes the Pearson correlation between y_j and $\widehat{y}_j$. For the unrestricted transformation, this is just the square root of the explained variance $1 - P(j)$, where $P(j) = \text{RSS}(j)/\text{SS}$. For the orthogonal and oblique transformation, this relationship does not hold.

References

Browne, M. W. 1967. On oblique Procrustes rotation. *Psychometrika* 32: 125–132.

Cox, T. F., and M. A. A. Cox. 2001. *Multidimensional Scaling.* 2nd ed. Boca Raton, FL: Chapman & Hall/CRC.

Cramer, E. M. 1974. On Browne's solution for oblique Procrustes rotation. *Psychometrika* 39: 159–163.

Gower, J. C., and G. B. Dijksterhuis. 2004. *Procrustes Problems.* Oxford: Oxford University Press.

Green, B. F. 1952. The orthogonal approximation of an oblique structure in factor analysis. *Psychometrika* 17: 429–440.

Hurley, J. R., and R. B. Cattell. 1962. The Procrustes program: Producing direct rotation to test a hypothesized factor structure. *Behavioral Science* 7: 258–262.

Mardia, K. V., J. T. Kent, and J. M. Bibby. 1979. *Multivariate Analysis.* London: Academic Press.

Mosier, C. I. 1939. Determining a simple structure when loadings for certain tests are known. *Psychometrika* 4: 149–162.

Peen, C. 2007. Old towns of England. http://www.oldtowns.co.uk.

ten Berge, J. M. F. 1977. Orthogonal Procrustes rotation for two or more matrices. *Psychometrika* 42: 267–276.

ten Berge, J. M. F., and K. Nevels. 1977. A general solution to Mosier's oblique Procrustes problem. *Psychometrika* 42: 593–600.

Also see