Journal of Classification 34:35-48 (2017)
DOI: 10.1007/s00357-017-9224-z

On the Correspondence Between Procrustes Analysis and Bidimensional Regression

Justin L. Kern

University of Illinois at Urbana-Champaign

Abstract: Procrustes analysis is defined as the problem of fitting a matrix of data to a target matrix as closely as possible (Gower and Dijksterhuis, 2004). The problem can take many forms, but the most common form, orthogonal Procrustes analysis, has as allowable transformations, a translation, a scaling, an orthogonal rotation, and a reflection. Procrustes analysis and other rotation methods have a long history in quantitative psychology, as well as in other fields, such as biology (Siegel and Benson, 1982) and shape analysis (Kendall, 1984). In the field of quantitative geography, the use of bidimensional regression (Tobler, 1965) has recently become popular. Tobler (1994) defines bidimensional regression as "an extension of ordinary regression to the case in which both the independent and dependent variables are two-dimensional." In this paper, it is established that orthogonal Procrustes analysis (without reflection) and Euclidean bidimensional regression are the same. As such, both areas of development can borrow from the other, allowing for a richer landscape of possibilities.

Keywords: Procrustes analysis; Bidimensional regression.

1. Introduction

Procrustes analysis has a long tradition in the field of psychometrics and quantitative psychology. The term "Procrustes analysis" comes from the Greek myth of Procrustes (Hurley and Cattell, 1962), wherein Procrustes invited strangers into his inn to sleep in his all-fitting bed. According to the tale, the bed was "all-fitting" because Procrustes fit his guests to the bed; he stretched his short guests with racks and chopped off the limbs of his tall

Corresponding Author's Address: J.L. Kern, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820, USA, e-mail: kern4@illinois.edu.

Published online: 20 March 2017

guests so that, no matter their size, they fit exactly. Similarly, the goal of a Procrustes analysis is to make one matrix of data fit a target matrix as closely as possible, usually using the least squares criterion.

Many different forms of Procrustes problems exist. In orthogonal Procrustes problems (Gruvaeus, 1970; Schönemann, 1966; Schönemann and Carroll, 1970), the goal is to transform one matrix to a target matrix as closely as possible, with the only permissible transformations being a translation, a scaling, an orthogonal rotation, and a reflection. Other constraints on the form of the transformation may be applied as well (Gower and Dijksterhuis, 2004). There have also been attempts to find orthogonal solutions that are robust to outliers (Groenen, Giaquinto, and Kiers, 2003; Siegel and Benson, 1982; Verboon and Heiser, 1992; Rohlf and Slice, 1990); this is accomplished by either underweighting points with large residuals, or by substituting a median criterion for the usual sums-of-squares criterion. If the constraints on the fitted matrix are lessened so that it no longer must be orthogonal, then this is an oblique Procrustes problem (Browne, 1967; Korth and Tucker, 1976; ten Berge and Nevels, 1977). Much has been written about how to best do this in different circumstances, and much of that has been rooted in the factor analytic tradition. Other techniques in the Procrustes literature include fitting more than two matrices to each other (generalized Procrustes analysis; Gower, 1975; Peay, 1988; ten Berge, 1977; ten Berge and Bekker, 1993), using alternative loss functions (Trendafilov, 2003; Trendafilov and Watson, 2004), differential weighting of columns (Koschat and Swayne, 1991; Mooijaart and Commandeur, 1990), and finding best fitting matrices to targets in a lower dimensionality (projection Procrustes problems; Cliff, 1966; Constantine and Gower, 1978, 1982; ten Berge, 1979; ten Berge and Knol, 1984). For this paper, the focus will be on two-dimensional orthogonal Procrustes analysis.

One field that has used Procrustean methods to great success is in the field of quantitative biology. The field's interest in Procrustes analysis is borne from the problem of examining shape differences in anatomical features within and between species. A wide array of shape analysis techniques were developed in quantitative biology for this purpose, under the umbrella of morphometrics (Bookstein, 1991). Siegel and Benson (1982) developed a resistant (via the median approach) extension of the Procrustes method, and illustrated the approach comparing the lateral profiles of human skulls (*Australopithecus*, *Homo erectus*, and *Homo sapiens*) with a chimpanzee skull. In this approach, the vectors of the residuals after fitting to the chimpanzee skull superimposed onto the skull show the way the skulls differ. Following this same approach, Rohlf and Slice (1990) developed a series of resistant extensions of the generalized Procrustes method, and illustrated the approaches with a set of mosquito wings. Other biological examples include those by Mardia and Dryden (1989b), Goodall (1991), and Smith, Crespi, and Bookstein (1997).

Another way to view Procrustes analysis (in two-dimensional space) is as a method to quantify the "Procrustes distance" between two shapes. Here, shape is defined as the remaining information about a configuration of points after accounting for information about position, orientation, and scaling (Bookstein, 1991). This information on shape can be characterized by a set of equivalence classes on a sphere (with respect to rotations) called "Kendall's shape space." Kendall (1981, 1984) showed that there is a shortest length from one shape to another on this sphere in shape space, and that distance is preserved when returned to the original Euclidean space of the configurations, and, thus, can easily be used as a metric of the closeness between two shapes. As such, the sum of the squared residuals between two configurations (after translating, rotating, and scaling one configuration to the other) can be seen as the distance between the two configurations (Goodall, 1991). A statistical theory following Kendall's definition of closeness between shapes was developed by Mardia and Dryden (1989a,b) and Goodall (1991).

Bidimensional regression is a shape analysis technique that was developed by Tobler (1965, 1966, 1978, 1994). In essence, the idea is to estimate (much like with the Procrustes distance) the degree of similarity between two different point location patterns. We may recognize this as simply a product-moment correlation and, thus, the least-squares regression. However, we must take into account the two-dimensional nature of the coordinate system. Indeed, Tobler (1994) defines bidimensional regression as "an extension of ordinary regression to the case in which both the independent and dependent variables are two-dimensional." Through this, a spatial correlation, or a distortion index (Waterman and Gordon, 1984), can then be computed to assess similarity.

Bidimensional regression was originally developed within the quantitative geography literature as a tool for mapping. In particular, it has been used numerous times as a way to assess the accuracy of cognitive maps (Carbon, 2013; Friedman and Kohler, 2003; Giraudo, Gayraud, and Habib, 1997; Ishikawa, 2013; Kitchin, 1996; Ohuchi, Iwaya, Suzuki, and Munekata, 2006; Schinazi et al., 2013). In this paradigm, cognitive maps can be obtained either via direct sketch-mapping or by asking for distances between places and obtaining a map through a multidimensional scaling solution (Waller and Haun, 2003); the real map is then regressed on the cognitive map. Other applications include the relationship between spatial patterns and travel times (Ahmed and Miller, 2007; Dusek, 2012), decision-making in virtual environments (Bakdash, Linikenauger, and Proffitt, 2008), and assessing facial similarities (Kare, Samuel, and Marx, 2010; Shi, Samal, and Marx, 2005). It has also been used to assess the similarity between historical geographic maps (Tobler, 1966; Symington, Charlton, and Brunsdon, 2002).

In recent years, many extensions to bidimensional regression have been developed. These include differential weighting of dimensions (Schmid, Marx, and Samal, 2011), using Gibbs sampling to estimate a Bayesian bidimensional regression and a hierarchical bidimensional regression model (Lee, 2012), an extension to three-dimensional configurations called tridimensional regression (Schmid, Marx, and Samal, 2012), and nonlinear bidimensional regression using kernel smoothing (Symington, Charlton, and Brunsdon, 2002). Other work includes developing statistical inference techniques for the method (Nakaya, 1997), and the development of a statistical package (Carbon and Leder, 2005).

As should be clear, the goals of Procrustes analysis and bidimensional regression are very similar to each (that is, to determine the closest configuration to another by simple transformations and assessing their similarity). This paper seeks to show that in a special case—two-dimensional orthogonal Procrustes analysis and Euclidean bidimensional regression—these two techniques are exactly equivalent.

2. Procrustes Analysis

In this section, a description of the solutions to Procrustes analysis is provided. Though these are well-known (e.g., Schönemann, 1966; Schönemann and Carroll, 1970; Gower and Dijksterhuis, 2004), it is instructive for the reader for these to be reproduced here.

2.1 Translation

Procrustes analysis is usually formulated as a least-squares problem where we are trying to find a $p \times p$ transformation matrix $\mathbf{T}$ that minimizes the squared distance from entries in one $n \times p$ matrix $\mathbf{X}$ to entries in a second $n \times p$ matrix $\mathbf{Y}$. Usually, we are mainly interested in the shape of the configurations $\mathbf{X}$ and $\mathbf{Y}$, and in this case, the matrices can be translated to the origin by instead considering the matrices $\mathbf{X} - \mathbf{1a}_1'$ and $\mathbf{Y} - \mathbf{1a}_2'$, where $\mathbf{1}$ is an appropriately sized vector of ones. Furthermore, the size of the shape can be considered arbitrary, so if we are additionally only interested in similarity of shape, then a scaling constant s may be incorporated. Thus, the Procrustes problem can be determined as finding the values $\mathbf{a}_1$, $\mathbf{a}_2$, s, and $\mathbf{T}$ such that

$$||s(\mathbf{X} + \mathbf{1a}_1')\mathbf{T} - (\mathbf{Y} + \mathbf{1a}_2')|| \tag{1}$$

is minimized.

In Equation 1, it is easy to see that the translation parameters are not unique:

$$\begin{aligned}\|s(\mathbf{X}+\mathbf{1a}_1')\mathbf{T}-(\mathbf{Y}+\mathbf{1a}_2')\| &= \|s\mathbf{XT}+s\mathbf{1a}_1'\mathbf{T}-\mathbf{Y}-\mathbf{1a}_2'\| \\ &= \|s\mathbf{XT}-\mathbf{Y}+\mathbf{1}(s\mathbf{a}_1'\mathbf{T}-\mathbf{a}_2')\| \\ &= \|s\mathbf{XT}-\mathbf{Y}+\mathbf{1a}'\|. \end{aligned} \tag{2}$$

Since $\|\mathbf{A}+\mathbf{1a}'\|$, where $\mathbf{A}$ is an arbitrary $n \times p$ matrix, is minimized when $\mathbf{a}' = -\frac{\mathbf{1}'\mathbf{A}}{n}$ (the column means of $\mathbf{A}$), then the translation term in Equation 2 is given by

$$\hat{\mathbf{a}}' = \frac{\mathbf{1}'(\mathbf{Y}-s\mathbf{XT})}{n}. \tag{3}$$

Importantly, we can find that s and $\mathbf{T}$ can be optimized without considering $\mathbf{a}'$ by simply centering $\mathbf{X}$ and $\mathbf{Y}$ and using the centered matrices in the Procrustes analysis. This can be seen in the following:

$$\mathbf{A}+\mathbf{1a}' = \mathbf{A}-\frac{\mathbf{11}'\mathbf{A}}{n} = \left(\mathbf{I}-\frac{\mathbf{11}'}{n}\right)\mathbf{A} = \mathbf{CA} = \mathbf{A}^c,$$

where $\mathbf{C} = \mathbf{I}-\frac{\mathbf{11}'}{n}$ is the centering matrix, and $\mathbf{A}^c \equiv \mathbf{CA}$ is the matrix $\mathbf{A}$ centered at the origin. Then,

$$\mathbf{CA} = \mathbf{C}(s\mathbf{XT}-\mathbf{Y}) = s\mathbf{CXT}-\mathbf{CY} = s\mathbf{X}^c\mathbf{T}-\mathbf{Y}^c.$$

Thus,

$$\|s\mathbf{XT}-\mathbf{Y}+\mathbf{1a}'\| = \|s\mathbf{X}^c\mathbf{T}-\mathbf{Y}^c\|. \tag{4}$$

2.2 Scaling

In the case of a general transformation matrix $\mathbf{T}$, then the scaling parameter can be absorbed into $\mathbf{T}$. If, on the other hand, we insist that $\mathbf{T}$ be constrained to have a structure in some way, then a least-squares estimate of s can be found by minimizing Equation 4 with respect to s.

$$\begin{aligned} F = \|s\mathbf{X}^c\mathbf{T}-\mathbf{Y}^c\| &= \mathrm{tr}[(s\mathbf{X}^c\mathbf{T}-\mathbf{Y}^c)'(s\mathbf{X}^c\mathbf{T}-\mathbf{Y}^c)] \\ &= \mathrm{tr}(s^2\mathbf{T}'\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T}-2s\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}+\mathbf{Y}^{c\prime}\mathbf{Y}^c) \\ &= s^2\mathrm{tr}(\mathbf{T}'\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T})-2s\,\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T})+\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{Y}^c). \end{aligned} \tag{5}$$

Then

$$\frac{\partial F}{\partial s} = 2s\,\mathrm{tr}(\mathbf{T}'\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T})-2\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}),$$

and

$$\hat{s} = \frac{\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T})}{\mathrm{tr}(\mathbf{T}'\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T})}. \tag{6}$$

2.3 Rotation

Finally, we only need to find $\mathbf{T}$. For this, we are interested in the case where $\mathbf{T}$ is constrained to be orthogonal; that is, when $\mathbf{T}'\mathbf{T} = \mathbf{T}\mathbf{T}' = \mathbf{I}$. When $\mathbf{T}$ is orthogonal, then $\det(\mathbf{T}) = \pm 1$, with a determinant of -1 corresponding to a reflection over an axis, and a determinant of 1 corresponding to no reflection. Since we are going to show that bidimensional regression is equivalent to orthogonal Procrustes analysis without a reflection, we will insist on the additional constraint that $\det(\mathbf{T}) = 1$. In this case, $\mathbf{T}$ is a matrix representation of a simple rotation. As such, in two-dimensions, where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2]$ are $n \times 2$ matrices, it will take the structure

$$\mathbf{T} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{7}$$

where θ defines the degree of rotation, so $\mathbf{T}$ is a function of θ.

From Equation 5, we find that if $\mathbf{T}$ is orthogonal, then

$$\begin{aligned} F &= s^2 \mathrm{tr}(\mathbf{T}'\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T}) - 2s\,\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}) + \mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{Y}^c) \\ &= s^2 \mathrm{tr}(\mathbf{X}^{c\prime}\mathbf{X}^c\mathbf{T}\mathbf{T}') - 2s\,\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}) + \mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{Y}^c) \\ &= s^2 \mathrm{tr}(\mathbf{X}^{c\prime}\mathbf{X}^c) - 2s\,\mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}) + \mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{Y}^c). \end{aligned}$$

Thus, to find the optimal θ, we need only maximize $F_2 = \mathrm{tr}(\mathbf{Y}^{c\prime}\mathbf{X}^c\mathbf{T}) = \mathrm{tr}(\mathbf{Z}\mathbf{T})$ with respect to θ, where $\mathbf{Z} = \mathbf{Y}^{c\prime}\mathbf{X}^c$. By finding the derivative of $\mathrm{tr}(\mathbf{Z}\mathbf{T})$ with respect to θ and setting that equal to zero, we find that

$$\hat{\theta} = \tan^{-1}\left(\frac{z_{21} - z_{12}}{z_{11} + z_{22}}\right) = \tan^{-1}\left(\frac{u}{v}\right), \tag{8}$$

where $u = z_{21} - z_{12}$, $v = z_{11} + z_{22}$, and $z_{jk} = \sum_{i=1}^{n} x_{ji} y_{ki} - n\bar{x}_j \bar{y}_k$.

Finally, using Equations 3, 6, and 8, and some algebra, we can find that the full solution of the two-dimensional orthogonal Procrustes problem without a reflection is

$$\hat{\theta} = \tan^{-1}\left(\frac{u}{v}\right) \tag{9}$$

$$\hat{s} = \frac{\sqrt{v^2 + u^2}}{\sum_{i=1}^{n}(x_{1i} - \bar{x}_1)^2 + \sum_{i=1}^{n}(x_{2i} - \bar{x}_2)^2} \tag{10}$$

$$\hat{a}_1 = \bar{y}_1 + \frac{\hat{s}u\bar{x}_2 - \hat{s}v\bar{x}_1}{\sqrt{u^2 + v^2}} \tag{11}$$

$$\hat{a}_2 = \bar{y}_2 - \frac{\hat{s}u\bar{x}_1 + \hat{s}v\bar{x}_2}{\sqrt{u^2 + v^2}}. \tag{12}$$

3. Euclidean Bidimensional Regression

Euclidean bidimensional regression is described as an extension of the usual univariate regression to a case where the independent and dependent variables are two-dimensional. However, the two-dimensional nature of the independent and dependent variables is modeled using complex numbers, rather than simply a two-dimensional vector. In this case, the coefficients are also complex numbers. Thus, the regression is simply

$$\mathbf{y} = c\mathbf{1} + \beta\mathbf{x} \tag{13}$$

where $\mathbf{y}$ and $\mathbf{x}$ are $n \times 1$ vectors where the jth entry is $y_{1j} + iy_{2j}$ and $x_{1j} + ix_{2j}$, respectively, $c = c_1 + ic_2$, and $\beta = \beta_1 + i\beta_2$. Equation 13 can be expanded as

$$\begin{aligned}\mathbf{y} &= c\mathbf{1} + \beta\mathbf{x} \\ &= c_1\mathbf{1} + ic_2\mathbf{1} + \beta_1\mathbf{x}_1 + i\beta_1\mathbf{x}_2 - \beta_2\mathbf{x}_2 + i\beta_2\mathbf{x}_1 \\ &= (c_1\mathbf{1} + \beta_1\mathbf{x}_1 - \beta_2\mathbf{x}_2) + i(c_2\mathbf{1} + \beta_1\mathbf{x}_2 + \beta_2\mathbf{x}_1). \end{aligned} \tag{14}$$

Equation 14 can then be broken into two equations representing the real and imaginary parts of $\mathbf{y} = \mathbf{y}_1 + i\mathbf{y}_2$ as:

$$\begin{aligned}\mathbf{y}_1 &= c_1\mathbf{1} + \beta_1\mathbf{x}_1 - \beta_2\mathbf{x}_2 \\ \mathbf{y}_2 &= c_2\mathbf{1} + \beta_1\mathbf{x}_2 + \beta_2\mathbf{x}_1,\end{aligned}$$

and then finally recombined into a single matrix equation as

$$\mathbf{y}_s = \begin{pmatrix}\mathbf{y}_1 \\ \mathbf{y}_2\end{pmatrix} = \begin{pmatrix}\mathbf{1} & \mathbf{0} & \mathbf{x}_1 & -\mathbf{x}_2 \\ \mathbf{0} & \mathbf{1} & \mathbf{x}_2 & \mathbf{x}_1\end{pmatrix}\begin{pmatrix}c_1 \\ c_2 \\ \beta_1 \\ \beta_2\end{pmatrix} = \mathbf{X}_d\mathbf{b} \tag{15}$$

To find least-squares estimates of the parameters, we can then simply use the standard regression estimator $\hat{\mathbf{b}} = (\mathbf{X}_d'\mathbf{X}_d)^{-1}\mathbf{X}_d'\mathbf{y}_s$. It is relatively simple, though algebraically time-consuming, to find these estimates in scalar form. They are as follows:

$$\hat{c}_1 = \bar{y}_1 - \hat{\beta}_1\bar{x}_1 + \hat{\beta}_2\bar{x}_2 \tag{16}$$

$$\hat{c}_2 = \bar{y}_2 - \hat{\beta}_1\bar{x}_2 - \hat{\beta}_2\bar{x}_1 \tag{17}$$

$$\hat{\beta}_1 = \frac{\bar{x}_1\bar{y}_1 + \bar{x}_2\bar{y}_2 - \frac{1}{n}(\sum_i x_{1i}y_{1i} + x_{2i}y_{2i})}{\bar{x}_1^2 + \bar{x}_2^2 - \frac{1}{n}\sum_i(x_{1i}^2 + x_{2i}^2)} \tag{18}$$

$$\hat{\beta}_2 = \frac{\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1 - \frac{1}{n}(\sum_i x_{1i}y_{2i} - x_{2i}y_{1i})}{\bar{x}_1^2 + \bar{x}_2^2 - \frac{1}{n}\sum_i(x_{1i}^2 + x_{2i}^2)}. \tag{19}$$

There is one thing to note here. In Equation 13, the complex regression parameter β can also be written in polar coordinate notation as

$$\beta = \beta_1 + i\beta_2 = \sqrt{\beta_1^2 + \beta_2^2} e^{i \arg(\beta)} = se^{i\theta},$$

where

$$\arg(\beta) = \begin{cases} \tan^{-1}\left(\frac{\beta_2}{\beta_1}\right), & \text{if } \beta_1 > 0 \\ \tan^{-1}\left(\frac{\beta_2}{\beta_1}\right) + \pi, & \text{if } \beta_1 < 0 \text{ and } \beta_2 \geq 0 \\ \tan^{-1}\left(\frac{\beta_2}{\beta_1}\right) - \pi, & \text{if } \beta_1 < 0 \text{ and } \beta_2 < 0 \\ \frac{\pi}{2}, & \text{if } \beta_1 = 0 \text{ and } \beta_2 > 0 \\ -\frac{\pi}{2}, & \text{if } \beta_1 = 0 \text{ and } \beta_2 < 0 \end{cases}$$

s is the distance from the origin and θ is the counter-clockwise angle from the real axis. Similarly, the jth element of $\mathbf{x}$ can be written as

$$x_j = x_{1j} + ix_{2j} = \sqrt{x_{1j}^2 + x_{2j}^2} e^{i \arg(x_j)} = r_j e^{i\phi_j}.$$

Combining them together, Equation 13 can be written in scalar form as

$$y_j = c + (sr_j)e^{i(\theta + \phi_j)}, \; j = 1, \ldots, n. \tag{20}$$

As such, the regression parameter can be thought of as stretching or shrinking $\mathbf{x}$ by a factor of s and rotating $\mathbf{x}$ counter-clockwise by an angle of θ. Thus, since the goal of the orthogonal Procrustes analysis without a reflection is the same (finding an optimal scaling, rotation, and translation) and both use a least-squares estimation approach, one would expect that the estimates obtained should be one-to-one transformations of each other.

A more general form of bidimensional regression analogous to polynomial regression also exists; it is simply called curvilinear bidimensional regression. In scalar form, the model is expressed as

$$y_j = c + \sum_{k=1}^{K} \beta_k x_j^k, \; j = 1, \ldots, n, \tag{21}$$

where $y_j = y_{1j} + iy_{2j}$, $x_j = x_{1j} + ix_{2j}$, $c = c_1 + ic_2$, and $\beta_k = \beta_{1k} + i\beta_{2k}$. In like manner as Equation 20, the curvilinear bidimensional regression model may be rewritten using polar coordinate notation as

$$y_j = c + \sum_{k=1}^{K} (s_k r_j^k) e^{i(\theta_k + k\phi_j)},$$

with a similar interpretation of the scaling and rotation parameters.

4. Linking Orthogonal Procrustes Analysis and Euclidean Bidimensional Regression

It can be easily established through some algebra that the parameters of the two-dimensional orthogonal Procrustes problem without a reflection are one-to-one functions of the parameters of the Euclidean bidimensional regression. To transform from the Procrustes parameters to the bidimensional regression parameters, we find that for $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$,

$$c_1 = a_1,\ c_2 = a_2,\ \beta_1 = \frac{s}{\sqrt{1+\tan^2\theta}},\ \text{and}\ \beta_2 = \frac{s\tan\theta}{\sqrt{1+\tan^2\theta}}, \tag{22}$$

for $-\pi < \theta < -\frac{\pi}{2}$ or $\frac{\pi}{2} < \theta \leq \pi$,

$$c_1 = a_1,\ c_2 = a_2,\ \beta_1 = -\frac{s}{\sqrt{1+\tan^2\theta}},\ \text{and}\ \beta_2 = -\frac{s\tan\theta}{\sqrt{1+\tan^2\theta}}, \tag{23}$$

and for $\theta = \pm\frac{\pi}{2}$,

$$c_1 = a_1,\ c_2 = a_2,\ \beta_1 = 0,\ \text{and}\ \beta_2 = \pm s. \tag{24}$$

Conversely, to transform from the bidimensional regression parameters to the Procrustes parameters, we have

$$a_1 = c_1,\ a_2 = c_2,\ s = \sqrt{\beta_1^2 + \beta_2^2},\ \theta = \arg(\beta). \tag{25}$$

5. Discussion

By establishing that the orthogonal Procrustes analysis problem without reflection and Euclidean bidimensional regression are the same, it's easy to see that their respective extensions can be seen as generalizations of the other as well. For instance, orthogonal Procrustes analysis with a forced reflection can be done in bidimensional regression by simply substituting $\overline{\mathbf{x}} = \mathbf{x}_1 - i\mathbf{x}_2$ (the complex conjugate of $\mathbf{x}$) in for $\mathbf{x}$ in Equation 13. As such, a standard orthogonal Procrustes analysis is equivalent to doing two Euclidean bidimensional regressions—one with $\mathbf{x}$ and one with $\overline{\mathbf{x}}$—and choosing the set of parameters with the highest corresponding R^2.

It is also possible to consider the curvilinear bidimensional regression model given in Equation 21 as a generalization of both bidimensional regression and Procrustes analysis. This can be seen by expanding Equation 21, breaking it up into two equations representing the real and imaginary parts of the model, and then recombining the model into a single matrix equation as

$$\mathbf{y}_s = \begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix} = \begin{pmatrix}\mathbf{1} & \mathbf{0} & \mathrm{Re}(\mathbf{x}) & -\mathrm{Im}(\mathbf{x}) & \cdots & \mathrm{Re}(\mathbf{x}^k) & -\mathrm{Im}(\mathbf{x}^k)\\ \mathbf{0} & \mathbf{1} & \mathrm{Im}(\mathbf{x}) & \mathrm{Re}(\mathbf{x}) & \cdots & \mathrm{Im}(\mathbf{x}^K) & \mathrm{Re}(\mathbf{x}^K)\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ \boldsymbol{\beta}_s\end{pmatrix},$$

where $\boldsymbol{\beta}_s = (\beta_{11} \quad \beta_{21} \quad \cdots \quad \beta_{1K} \quad \beta_{2K})'$. The corresponding Procrustes problem is

$$\|s_1\mathbf{X}_1\mathbf{T}_1 + s_2\mathbf{X}_2\mathbf{T}_2 + \cdots + s_K\mathbf{X}_K\mathbf{T}_K + \mathbf{1a}' - \mathbf{Y}\| = \|\mathbf{XST} + \mathbf{1a}' - \mathbf{Y}\|,$$

where $\mathbf{X}_k = \left[\mathrm{Re}(\mathbf{x}^k) \;\; \mathrm{Im}(\mathbf{x}^k)\right]'$,

$$\mathbf{T}_k = \begin{pmatrix} \cos\theta_k & \sin\theta_k \\ -\sin\theta_k & \cos\theta_k \end{pmatrix},$$

s_k and θ_k are one-to-one transformations of β_{1k} and β_{2k} following the transformations given in Equations 22–25, and

$$\begin{aligned} \mathbf{X} &= (\,\mathbf{X}_1 \mid \mathbf{X}_2 \mid \cdots \mid \mathbf{X}_K\,), \\ \mathbf{S} &= K^{1/2}\mathrm{diag}(s_1, \ldots, s_2) \otimes \mathbf{I}_2, \\ \mathbf{T} &= K^{-1/2}\left(\,\mathbf{T}_1' \mid \mathbf{T}_2' \mid \cdots \mid \mathbf{T}_K'\,\right)'. \end{aligned}$$

The rescaling constant $K^{-1/2}$ in $\mathbf{T}$ is present so that the columns of $\mathbf{T}$ remain orthonormal due to the following:

$$\left(K^{1/2}\mathbf{T}\right)'\left(K^{1/2}\mathbf{T}\right) = \mathbf{T}_1'\mathbf{T}_1 + \cdots + \mathbf{T}_K'\mathbf{T}_K = \underbrace{\mathbf{I} + \cdots + \mathbf{I}}_{K} = K\mathbf{I}.$$

Thus, we find that curvilinear bidimensional regression can be seen as a constrained projection Procrustes problem with an anisotropic scaling. This identity then suggests a a simple solution to this problem that might otherwise require an alternating least squares algorithm and Lagrange multipliers.

Finally, as bidimensional regression comes from the regression tradition, it is possible that new techniques in Procrustes analysis can borrow from this literature. While these are not pursued here, some possibilities include incorporating correlated errors, such as in time series analysis, and allowing for random effects, such as in mixed models and hierarchical models. Other nonlinear transformations other than polynomial transformations may also be possible. These possibilities may spur on more interest in how Procrustes methods can be used in unique analysis designs in the future. For instance, as described here, it would be possible to directly incorporate covariates into the design.

Suppose that there are two groups of subjects, Groups 1 and 2, and we are interested in the effect that some experimental condition has on the cognitive map $\mathbf{y}$ of some real map $\mathbf{x}$. In this case, it would be best to have the subjects draw a sketch map of their conception of the locations of places of interest on the map, where the locations of two places of interest have already been provided. The effect of the experiment can be modeled by simply including a covariate for group membership $\mathbf{w}$, coded as $w_j = 1+0i$ if participant j was in Group 2, and $w_j = 0 + 0i$ otherwise. An interaction effect could also be included. In a bidimensional regression, this would amount to

$$\hat{\mathbf{y}} = c\mathbf{1} + \beta_1\mathbf{x} + \beta_2\mathbf{w} + \beta_3\mathbf{xw}.$$

For those in Group 1, this would reduce to

$$\hat{\mathbf{y}} = c\mathbf{1} + \beta_1\mathbf{x},$$

and would be

$$\hat{\mathbf{y}} = (c + \beta_2)\mathbf{1} + (\beta_1 + \beta_3)\mathbf{x}$$

for those in Group 2. Thus, β_2 is the effect of the experimental group on the translation, whereas β_3 is the effect of the experimental group on the rotation and scaling of the real map to the cognitive map. From the interpretation given above of these parameters as transformations of rotation and scaling parameters, we can now find the effects. Let us take $\beta_k = \beta_{1k} + i\beta_{2k}$ and $c = c_1 + ic_2$. Then, Group 1 has translation parameters $a_{11} = c_1$ and $a_{21} = c_2$, a scaling of $s_1 = \sqrt{\beta_{11}^2 + \beta_{12}^2}$, and a rotation of $\theta_1 = \arg(\beta_1)$, whereas Group 2 has translation parameters $a_{12} = c_1 + \beta_{12}$ and $a_{22} = c_2 + \beta_{22}$, a scaling of $s_2 = \sqrt{(\beta_{11} + \beta_{31})^2 + (\beta_{21} + \beta_{32})^2}$, and a rotation of $\theta_2 = \arg(\beta_1 + \beta_3)$. In this way, we can parse the nonlinear effect that the experimental group had into a more standard "linear" interpretation. This, of course, is just one description of an analysis design that can be handled with the recognition that bidimensional regression and Procrustes analysis are related; many others are possible.

References

AHMED, N., and MILLER, H.J. (2007), "Time-Space Transformations of Geographic Space for Exploring, Analyzing and Visualizing Transportation Systems", *Journal of Transport Geography, 15*, 2–17.

BAKDASH, J.Z., LINKENAUGER, S.A., and PROFFITT, D. (2008), "Comparing Decision-Making and Control for Learning a Virtual Environment: Backseat Drivers Learn Where They Are Going", *Proceedings of the Human Factors and Ergonomics Society, 52*, 2117–2121.

BOOKSTEIN, F.L. (1991), *Morphometric Tools for Landmark Data: Geometry and Biology*, New York, NY: Cambridge University Press.

BROWNE, M.W. (1967), "On Oblique Procrustes Rotation", *Psychometrika, 32*, 125–132.

CARBON, C.-C. (2013), BiDimRegression: Bidimensional Regression Modeling Using R, *Journal of Statistical Software, 52*, 1–11.

CARBON, C.-C., and LEDER, H. (2005), "The Wall Inside the Brain: Overestimation of Distances Crossing the Former Iron Curtain", *Psychonomic Bulletin and Review, 12*, 746–750.

CLIFF, N. (1966), "Orthogonal Rotation to Congruence", *Psychometrika, 31*, 33–42.

CONSTANTINE, A.C., and GOWER, J.C. (1978), " Some Properties and Applications of Simple Orthogonal Matrices", *IMA Journal of Applied Mathematics, 21*, 445–454.

CONSTANTINE, A.C., and GOWER, J.C. (1982), "Models for the Analysis of Inter-Regional Migration", *Environmental Planning A, 14*, 477–497.

DUSEK, T. (2012), "Bidimensional Regression in Spatial Analysis", *Regional Statistics (English Edition), 52*, 61–73.

FRIEDMAN, A., and KOHLER, B. (2003), "Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and Other Two-Dimensional Data Sets", *Psychological Methods, 8*, 468–491.

GIRAUDO, M.-D., GAYRAUD, D., and HABIB, M. (1997), "Visuospatial Ability of Parkinsonians and Elderly Adults in Location Memory Tasks", *Brain and Cognition, 34*, 259–273.

GOODALL, C. (1991), "Procrustes Methods in the Statistical Analysis of Shape", *Journal of the Royal Stastistical Society. Series B (Methodological), 53*, 285–339.

GOWER, J.C. (1975), "Generalized Procrustes Analysis", *Psychometrika, 40*, 33–51.

GOWER, J.C., and DIJKSTERHUIS, G.B. (2004), *Procrustes Problems*, New York, NY: Oxford University Press.

GROENEN, P.J.F., GIAQUINTO, P., and KIERS, H.A.L. (2003), "An Improved Majorization Algorithm for Robust Procrustes Analysis", in *New Developments in Classification and Data Analysis,* eds. M. Vichi, P. Monari, S. Mignani, and A. Montanari, Heidelberg, Germany: Springer, pp. 151–158.

GRUVAEUS, G.T. (1970), "A General Approach to Procrustes Pattern Rotation", *Psychometrika, 35*, 493–505.

HURLEY, J.R., and CATTELL, R.B. (1962), "The Procrustes Program: Producing Direct Rotation to Test a Hypothesized Factor Structure", *Behavioral Science, 7*, 258–262.

ISHIKAWA, T. (2013), "Retention of Memory for Large-Scale Spaces", *Memory, 21(7)*, 807–817.

KARE, S., SAMAL, A., and MARX, D. (2010), "Using Bidimensional Regression to Assess Face Similarity", *Machine Vision and Applications, 21*, 261–274.

KENDALL, D.G. (1981), "The Statistics of Shape", in *Interpreting Multivariate Data*, ed. V. Barnett, New York, NY: Wiley, pp. 75–80.

KENDALL, D.G. (1984), "Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces", *Bulletin of the London Mathematical Society, 16*, 81–121.

KITCHIN, R.M. (1996), "Methodological Convergence in Cognitive Mapping Research: Investigating Configurational Knowledge", *Journal of Environmental Psychology, 16*, 163–185.

KORTH, B., and TUCKER, L.R. (1976), "Procrustes Matching by Congruence Coefficients", *Psychometrika, 41*, 531–535.

KOSCHAT, M.A., and SWAYNE, D.F. (1991), "A Weighted Procrustes Criterion", *Psychometrika, 56*, 229–239.

LEE, H. (2012), "Bayesian Bidimensional Regression and Its Extension", Master's thesis, Washington University, St. Louis, MO.

MARDIA, K.V., and DRYDEN, I.L. (1989a), "Shape Distributions for Landmark Data", *Advances in Applied Probability, 21*, 742–755.

MARDIA, K.V., and DRYDEN, I.L. (1989b), "The Statistical Analysis of Data", *Biometrika, 76*, 271–281.

MOOIJAART, A., and COMMANDEUR, J.J.F. (1990), "A General Solution of the Weighted Orthonormal Procrustes Problem", *Psychometrika, 55*, 657–663.

NAKAYA, T. (1997), "Statistical Inferences in Bidimensional Regression Models", *Geographical Analysis, 29*, 169–186.

OHUCHI, M., IWAYA, Y., SUZUKI, Y., and MUNEKATA, T. (2006), "Cognitive-Map Formation of Blind Persons in a Virtual Sound Environment", *Proceedings of the 12th International Conference on Auditory Display*, pp. 1–7.

PEAY, E.R. (1988), "Multidimensional Rotation and Scaling of Configurations to Optimal Agreement", *Psychometrika, 53*, 199–208.

ROHLF, F.J., and SLICE, D. (1990), "Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks", *Systematic Zoology, 39*, 40–59.

SCHINAZI, V.R., NADI, D., NEWCOMBE, N.S., SHIPLEY, T.F., and EPSTEIN, R.A. (2013), "Hippocampal Size Predicts Rapid Learning of a Cognitive Map in Humans", *Hippocampus, 23*, 515–528.

SCHMID, K.K., MARX, D.B., and SAMAL, A. (2011), "Weighted Bidimensional Regression", *Geographical Analysis, 43*, 1–13.

SCHMID, K.K., MARX, D.B., and SAMAL, A. (2012), "Tridimensional Regression for Comparing and Mapping 3D Anatomical Structures", *Anatomy Research International*, 1–9. f

SCHÖNEMANN, P.H. (1966), "A Generalized Solution of the Orthogonal Procrustes Problem", *Psychometrika, 31*, 1–10.

SCHÖNEMANN, P.H., and CARROLL, R.M. (1970), "Fitting One Matrix to Another Under Choice of a Central Dilation and a Rigid Motion", *Psychometrika, 35*, 245–256.

SHI, J., SAMAL, A., and MARX, D. (2005), "Face Recognition Using Landmark-Based Bidimensional Regression", *Proceedings of the Fifth IEEE International Conference on Data Mining*, pp. 1–4.

SIEGEL, A.F., and BENSON, R.H. (1982), "A Robust Comparison of Biological Shapes", *Biometrics, 38*, 341–350.

SMITH, D.R., CRESPI, B.J., and BOOKSTEIN, F.L. (1997), "Fluctuating Asymmetry in the Honey Bee, *Apis Mellifera*: Effects of Ploidy and Hybridization", *Journal of Evolutionary Biology, 10*, 551–574.

SYMINGTON, A., CHARLTON, M. E., and BRUNSDON, C.F. (2002), "Using Bidimensional Regression to Explore Map Lineage", *Computers, Environment and Urban Systems, 26*, 201–218.

TEN BERGE, J.M.F. (1977), "Orthogonal Procrustes Rotation for Two or More Matrices", *Psychometrika, 42*, 267–276.

TEN BERGE, J.M.F. (1979), "On the Equivalence of Two Oblique Congruence Rotation Methods, and Orthogonal Approximations", *Psychometrika, 44*, 359–364.

TEN BERGE, J.M.F., and BEKKER, P.A. (1993), "The Isotropic Scaling Problem in Generalized Procrustes Analysis", *Computational Statistics and Data Analysis, 16*, 201–204.

TEN BERGE, J.M.F., and KNOL, D.L. (1984), "Orthogonal Rotations to Maximal Agreement for Two or More Matrices of Different Column Orders", *Psychometrika, 49*, 49–55.

TEN BERGE, J.M.F., and NEVELS, K. (1977), "A General Solution to Mosier's Oblique Procrustes Problem", *Psychometrika, 42*, 593–600.

TOBLER, W.R. (1965), "Computation of the Correspondence of Geographical Patterns", *Papers of the Regional Science Association, 15*, 131–139.

TOBLER, W.R. (1966), "Medieval Distortions: The Projections of Ancient Maps", *Annals of the Association of American Geographers, 56*, 351–360.

TOBLER, W.R. (1978), "Comparing Figures by Regression", *Computer Graphics, 12*, 193–195.

TOBLER, W.R. (1994), "Bidimensional Regression", *Geographical Analysis, 26*, 187–212.

TRENDAFILOV, N.T. (2003), "On the $\ell 1$ Procrustes Problem", *Future Generation Computer Systems, 19*, 1177–1186.

TRENDAFILOV, N.T., and WATSON, G.A. (2004), "The $\ell 1$ Oblique Procrustes Problem", *Statistics and Computing, 14*, 39–51.

VERBOON, P., and HEISER, W.J. (1992), "Resistant Orthogonal Procrustes Analysis", *Journal of Classification, 9*, 237–256.

WALLER, D., and HAUN, D.B.M. (2003), "Scaling Techniques for Modeling Directional Knowledge", *Behavior Research Methods, Instruments, and Computers, 35*, 285–293.

WATERMAN, S., and GORDON, D. (1984), "A Quantitative-Comparative Approach to Analysis of Distortion in Mental Maps", *Professional Geographer, 36*, 326–337.