

Closeness Coefficients between Euclidean-Embeddable Homologous Configurations

Moonseong Heo*

Division of Biostatistics, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Belfer 1303E, Bronx, NY 10461, USA

Abstract: Measurement of closeness between homologous configurations is often of interest. For configurations that can be embedded onto the Euclidean space, we attempted to develop closeness coefficients between corresponding Euclidean coordinate matrices. A suitable closeness coefficient was required to satisfy the following five properties: 1) It must range between 0 and 1; 2) It must be invariant over translation, rotation and dilation of coordinate matrices, namely, TRD-invariance; 3) It must be one between equivalent coordinate matrices; 4) It must be zero between coordinate matrices whose corresponding configurations are orthogonal; and 5) It must be symmetric between any pair of coordinate matrices. We showed that the following two closeness coefficients derived based on different approaches were equivalent and both satisfied the five required properties: 1) a goodness of fit coefficient GF based on minimum distance fitting of coordinate matrices by translation, rotation and dilation; and 2) the Gower-Lingoes-Schönenman coefficient R_{GLS} based on the maximum of correlations of coordinate matrices over rotation. In addition, the Escoufier's RV coefficient was also shown to satisfy all the five properties. Finally, R_{GLS} , or equivalently GF, and RV were all shown to be a function of centered forms or singular values of coordinate matrices.

Key Words: Configuration, Coordinate matrix, Orthogonality of configurations, TRD-invariance, Closeness.

1. INTRODUCTION

Comparison between pairs of configurations can be made in terms of goodness of fit based on Procrustes methods [1, 2] or matrix correlations [3]. For instance, it may be of interest to measure how closely distinctive patterns of psychiatric disorders are fitted by a lower-rank approximation of a sample data matrix [4, 5] or to measure association between annual precipitations and geographical locations [6]. Recently, such interest extends to the area of bioinformatics; for instance, comparison of high-dimensional genomic data [7] and protein data [8]. Nevertheless, to our knowledge, a rigorous attempt has not been made to develop a closeness coefficient to measure such association between homologous configurations. In particular, properties required for a closeness coefficient have not been discussed. To this end, we introduced five properties that must be satisfied by a closeness coefficient. We also introduced a definition of orthogonality between homologous configurations that is necessary for satisfying a property.

The objective of this paper is to develop a closeness coefficient between homologous configurations. We considered only homologous configuration that are embeddable onto the Euclidean space. To this end, we first defined Euclidean-

embeddable configurations and then their corresponding Euclidean coordinate matrices in section 2. In section 3, we discussed unique representations of those coordinate matrices and introduced the concept of invariance under translation, rotation, and dilation (TRD-invariance). In section 4, we introduced a notion of orthogonality between configurations and between their corresponding uncorrelated coordinate matrices. In section 5, five properties that a closeness coefficient must satisfy are introduced. In section 6, we developed a closeness coefficient through application of a Procrustes fitting method to approximate coordinate matrices to each other as closely as possible by means of translation, rotation and dilation. In section 7, we examined maximization of correlations between coordinate matrices over rotation. In section 8, we examined correlations between coordinate matrices that are uniquely represented through centered forms or singular values of coordinate matrices. Discussion follows in section 9.

2. CONFIGURATION AND COORDINATE MATRIX

A configuration C is a set of ordered elements $\omega_1, \dots, \omega_n$ corresponding to a collection of given distances δ_{ie} , $i, e=1, \dots, n$ between all pairs of the objects. The inter-object distances δ should satisfy that: (1) $\delta_{ie} \geq 0$ for all i, e with equality if $i = e$; (2) $\delta_{ie} = \delta_{ei}$ for all i and e ; (3) $\delta_{ie} \leq \delta_{ij} + \delta_{je}$ for all i, j , and e . Therefore, $C \equiv C(\Omega, \delta)$ is a function of both a set $\Omega = \{\omega_1, \dots, \omega_n\}$ of the ordered objects and a set $\delta = \{\delta_{ie}, i, e=1, \dots, n\}$ of the inter-object distances.

*Address correspondence to this author at the Division of Biostatistics, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Belfer 1303E, Bronx, NY 10461, USA; Tel: +1 718 430 8838; Fax: +1 718 430 8780; E-mail: mheo@aecom.yu.edu

Two configurations for the same objects, $C_1 = C(\Omega_1, \delta_1)$ and $C_2 = C(\Omega_2, \delta_2)$, are *homologous* if there exists a one-to-one association between the ordered elements of Ω_1 and Ω_2 . For example, both Ω_1 and Ω_2 represent geographical coordinates of certain locations, while δ_1 in C_1 and δ_2 in C_2 represent sets of geographical distances and differences in annual precipitations, respectively, between the locations [6]. In this paper, we only consider such homologous configurations and therefore consider a single set Ω as given. This reduces the configurations to a function of distances $\underline{\delta}$ only and enables to write $C_{\underline{\delta}}$ for $C(\Omega, \underline{\delta})$.

To develop a closeness coefficient between homologous configurations, we embed configurations onto a space defined by *coordinates* $\underline{\mathbf{x}}_i$ associated with each element ω_i , $i = 1, \dots, n$. The n -rowed matrix $\mathbf{X} = ((x_{ij}))$ with the i -th row vector $\underline{\mathbf{x}}_i^T = (x_{i1}, x_{i2}, \dots, x_{ij})$ for some J is a list of the coordinates of all the elements of Ω , where $((x_{ij}))$ denote a matrix whose ij -th element is x_{ij} . The relation between the coordinates \mathbf{X} of Ω and the configuration $C_{\underline{\delta}}$ depends on a norm which equates the distances between coordinates $\underline{\mathbf{x}}_i$ and $\underline{\mathbf{x}}_e$ with the distances δ_{ie} between elements ω_i and ω_e . To this end, we consider only the Euclidean norm $\| \cdot \|$ in this paper. A configuration $C_{\underline{\delta}}$ is said to be *Euclidean embeddable* if and only if there exists \mathbf{X} such that $\underline{d}(\mathbf{X}) = \underline{\delta}$, where $\underline{d}(\mathbf{X}) = \{ \|\underline{\mathbf{x}}_i - \underline{\mathbf{x}}_e\|, i, e=1, \dots, n \}$ is the set of Euclidean distances measured by the Euclidean norm $\| \cdot \|$ between all pairs of the coordinates of the elements in Ω . Any such \mathbf{X} is called a (*Euclidean*) *coordinate matrix* of $C_{\underline{\delta}}$.

For a given $C_{\underline{\delta}}$ however, its associated coordinate matrices are not unique but there exist a *collection of coordinate matrices* denoted by $M_{\underline{\delta}} = \{ \mathbf{X} \mid \underline{d}(\mathbf{X}) = \underline{\delta} \}$, where $\{e \mid c\}$ represents a set of elements e that satisfy condition c . Furthermore, for a given set $\underline{\delta}$ of distances, there exists a *class of configurations* denoted by $\underline{C}_{\underline{\delta}} = \{ C_{\underline{\gamma}} \mid \underline{\delta} = \underline{\gamma} \}$ in which normalized distances of every configuration are identical to the given normalized distances $\underline{\underline{\delta}} \equiv \underline{\delta} / \|\underline{\delta}\|$. By this definition, $\underline{\underline{\delta}} = \underline{\underline{\gamma}}$ if and only if there exists a non-zero scalar β such that $\underline{\delta} = \beta \underline{\gamma}$. The normalization was considered because a closeness coefficient must be invariant over dilation and thus it is unnecessary to distinguish configurations whose distances are proportional to one another.

In that context, configurations $C_{\underline{\delta}_1}$ and $C_{\underline{\delta}_2}$ are said to be *equivalent* if they belong to the same class, i.e., if $\underline{\underline{\delta}}_1 = \underline{\underline{\delta}}_2$. Correspondingly, a *class of collections of coordinate matrices* of a configuration class $\underline{C}_{\underline{\delta}}$, denoted by $\underline{M}_{\underline{\delta}} = \{ M_{\underline{\gamma}} \mid \underline{\underline{\delta}} = \underline{\underline{\gamma}} \} = \{ \mathbf{X} \mid \underline{d}(\mathbf{X}) = \underline{\underline{\delta}} \}$, in which normalized inter-row distances $\underline{d}(\mathbf{X}) \equiv \underline{d}(\mathbf{X}) / \|\underline{d}(\mathbf{X})\|$ of any \mathbf{X} are identical to the given normalized inter-object distances $\underline{\underline{\delta}}$. Subsequently, coordinate matrices \mathbf{X} and \mathbf{Y} are said to be *equivalent* if they belong to the same class, i.e., if $\underline{d}(\mathbf{X}) = \underline{d}(\mathbf{Y})$.

3. UNIQUE REPRESENTATIONS OF CONFIGURATIONS AND TRD-INVARIANCE

For a configuration $C_{\underline{\delta}}$, the n -by- n matrix with (i, e) -th element δ_{ie}^2 is the *squared distances matrix* denoted by $\Delta_{\underline{\delta}}^2 = ((\delta_{ie}^2))$. This matrix is uniquely related to the configuration, since $\underline{\delta}$ uniquely defines $C_{\underline{\delta}}$. Similarly, the elements of $\underline{d}(\mathbf{X})$ can be arrayed in the squared distances matrix $\mathbf{D}^2(\mathbf{X}) = ((\|\underline{\mathbf{x}}_i - \underline{\mathbf{x}}_e\|^2))$. This matrix can be expressed as a function of the coordinate matrix \mathbf{X} through $\mathbf{X}\mathbf{X}^T$, which is referred to as its *form* [2], in the following way [9]:

$$\mathbf{D}^2(\mathbf{X}) = \text{Diag}(\mathbf{X}\mathbf{X}^T)\mathbf{1}\mathbf{1}^T + \mathbf{1}\mathbf{1}^T\text{Diag}(\mathbf{X}\mathbf{X}^T) - 2\mathbf{X}\mathbf{X}^T, \quad (1)$$

where $\text{Diag}(\mathbf{X}\mathbf{X}^T)$ is the diagonal matrix with the diagonal elements $\|\underline{\mathbf{x}}_i\|^2$, $i = 1, \dots, n$, and $\mathbf{1}$ is a column vector with all n components being 1.

For a given configuration, the form $\mathbf{X}\mathbf{X}^T$ of its associated coordinate matrices \mathbf{X} in $M_{\underline{\delta}}$ is not unique. It becomes unique only if the coordinate matrices are *centered*. As follows from the equation (1), if $\mathbf{D}^2(\mathbf{X}) = \Delta_{\underline{\delta}}^2 = \mathbf{D}^2(\mathbf{Y})$, then we have

$$\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H} = -0.5\mathbf{H}\mathbf{D}^2(\mathbf{X})\mathbf{H} = -0.5\mathbf{H}\Delta_{\underline{\delta}}^2\mathbf{H} = -0.5\mathbf{H}\mathbf{D}^2(\mathbf{Y})\mathbf{H} = \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}, \quad (2)$$

where $\mathbf{H} = \mathbf{I} - \mathbf{1}\mathbf{1}^T/n$ is an idempotent column-centering operator such that $\mathbf{H}\mathbf{H} = \mathbf{H}$, and \mathbf{I} is the n -by- n identity matrix.

In other words, $\Delta_{\underline{\delta}}^2$ is a one-to-one function not of $\mathbf{X}\mathbf{X}^T$ but of the *centered form* $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}$. Therefore, the centered form is unique for a given squared distances matrix and also for a configuration. Furthermore, the centered form $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}$ is invariant over translation and rotation of \mathbf{X} , i.e., *TR-invariant*, since $\mathbf{H}[(\mathbf{X} + \mathbf{1}\underline{\alpha}^T)\mathbf{R}][(\mathbf{X} + \mathbf{1}\underline{\alpha}^T)\mathbf{R}]^T\mathbf{H} = \mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}$ for any vector $\underline{\alpha}$ and any *orthonormal matrix* \mathbf{R} such that $\mathbf{R}\mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{I}$. This *TR-invariant* property is summarized in the following Proposition 1.

Proposition 1

For a given Euclidean embeddable configuration $C_{\underline{\delta}}$, its corresponding Euclidean coordinate matrices \mathbf{X} and \mathbf{Y} , both with full affine ranks, belong to the same collection $M_{\underline{\delta}}$ if and only if there exist a coordinate vector $\underline{\alpha}$ and an orthonormal matrix \mathbf{R} such that $\mathbf{X} = (\mathbf{Y} + \mathbf{1}\underline{\alpha}^T)\mathbf{R}$.

Proof

If the latter holds, then it is clear that $\mathbf{D}^2(\mathbf{X}) = \mathbf{D}^2(\mathbf{Y})$. It follows that both \mathbf{X} and \mathbf{Y} belong to the same $M_{\underline{\delta}}$ by definition. If the former holds, then any pair of coordinate matrices \mathbf{X} and \mathbf{Y} of the same $M_{\underline{\delta}}$ must satisfy $\Delta_{\underline{\delta}}^2 = \mathbf{D}^2(\mathbf{X}) = \mathbf{D}^2(\mathbf{Y})$. It follows from equation (1) that $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H} = \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}$. Therefore, the latter holds by the following claim.

Claim

$\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H} = \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}$, with \mathbf{X} and \mathbf{Y} of full affine rank, if and only if there exist a coordinate vector $\underline{\alpha}$ and an orthonormal matrix \mathbf{R} such that $\mathbf{X} = (\mathbf{Y} + \mathbf{1}\underline{\alpha}^T)\mathbf{R}$.

Proof

It is enough to show that the necessity part holds. Since $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H} = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^T = \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}$, there exists an orthonormal \mathbf{R}^* such that $\mathbf{H}\mathbf{X} = \mathbf{H}\mathbf{Y}\mathbf{R}^*$. This implies that $\nu((\mathbf{X} - \mathbf{Y}\mathbf{R}^*)^T)$ is orthogonal to $\nu(\mathbf{H}^T)$, where $\nu(\cdot)$ denotes the column space. It follows that $\mathbf{X} - \mathbf{Y}\mathbf{R} = \mathbf{1}\boldsymbol{\gamma}^T$ for some vector $\boldsymbol{\gamma}$. Hence, $\mathbf{X} = (\mathbf{Y} + \mathbf{1}\boldsymbol{\alpha}^T)\mathbf{R}$ for some $\boldsymbol{\alpha} = \boldsymbol{\gamma}\mathbf{R}^T$ and orthonormal matrix \mathbf{R} . QED.

In addition, by classical scaling, *principal coordinates* [10] of a configuration can be obtained through the singular value decomposition (SVD) [11, 12] of $-0.5\mathbf{H}\Delta_{\delta}^2\mathbf{H}$. The principal coordinates are arrayed in a principal coordinate matrix $\mathbf{U}\mathbf{\Lambda}^{1/2}$, where $-0.5\mathbf{H}\Delta_{\delta}^2\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ [13]. (We denote the SVD equation of a matrix \mathbf{X} in general by $\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, where $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{V}\mathbf{V}^T = \mathbf{I}$, and $\mathbf{\Lambda}$ is a diagonal matrix of the *singular values* of \mathbf{X} .) If there are no multiples of same singular values, the principal coordinate matrix $\mathbf{U}\mathbf{\Lambda}^{1/2}$ is unique for a configuration C_{δ} .

For a configuration class, the squared distances matrix and the centered form are unique up to a dilation factor. However, the normalized squared distances matrices of \overline{C}_{δ} and of \mathbf{X} , $\underline{\Delta}_{\delta}^2 = \Delta_{\delta}^2/\mathbf{1}^T\Delta_{\delta}^2\mathbf{1}$ and $\underline{\mathbf{D}}^2(\mathbf{X}) = \mathbf{D}^2(\mathbf{X})/\mathbf{1}^T\mathbf{D}^2(\mathbf{X})\mathbf{1}$ respectively, are unique and related uniquely to the centered form of $\mathbf{X}/\sqrt{\mathbf{D}^2(\mathbf{X})/\mathbf{1}^T\mathbf{D}^2(\mathbf{X})\mathbf{1}}$. Therefore, for a configuration class the coordinates allow $\beta(\mathbf{X} + \mathbf{1}\boldsymbol{\alpha}^T)\mathbf{R}$ transformation to be invariant over translation, rotation and dilation of any coordinate matrix in \overline{M}_{δ} , i.e., *TRD-invariant*, as summarized in the following proposition, which is a natural extension of Proposition 1. Due to the equations in (2), both the normalized squared distances matrix and the normalized centered form are also TRD-invariant.

Proposition 2

For a given Euclidean embeddable configuration C_{δ} , its corresponding Euclidean coordinate matrices \mathbf{X} and \mathbf{Y} , with the same full affine ranks, belong to the same class of collections \overline{M}_{δ} if and only if there exist a coordinate vector $\boldsymbol{\alpha}$, a non-zero scalar β , and an orthonormal matrix \mathbf{R} such that $\mathbf{X} = \beta(\mathbf{Y} + \mathbf{1}\boldsymbol{\alpha}^T)\mathbf{R}$; also, if and only if their centered forms satisfy $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H} = \beta\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}$ for some non-zero β .

Proof

Since both $\underline{\mathbf{D}}^2(\mathbf{X})$ and $\underline{\mathbf{D}}^2(\mathbf{Y})$ are invariant over choice of the non-zero dilation factor β , the proof of proposition 1 can naturally be applied after normalization of \mathbf{X} and \mathbf{Y} . QED.

4. ORTHOGONALITY OF CONFIGURATION CLASSES

In univariate (or single dimensional) case, as the opposite of perfect correlation, there is zero correlation between two variates, i.e., lack of association. Analogously, in comparison of high-dimensional configurations, the opposite of perfect closeness/similarity may be said to be *orthogonality* between configurations C_{δ_1} and C_{δ_2} if univariate correlation is zero for every direction in one configuration with every direction in the other. Therefore, as summarized in the following Proposition 3, the orthogonality holds if $\mathbf{X}^T\mathbf{Y} = \mathbf{0}$ for any $\mathbf{X} \in \overline{M}_{\delta_1}$ and any $\mathbf{Y} \in \overline{M}_{\delta_2}$. The condition $\mathbf{X}^T\mathbf{Y} = \mathbf{0}$ implies that $\nu(\mathbf{X})$ is orthogonal to $\nu(\mathbf{Y})$, i.e., $\langle \mathbf{X}, \mathbf{Y} \rangle = 0$, where $\mathbf{0} = ((0_{ij}))$ with $0_{ij} = 0$ for all i and j , and $\langle \mathbf{X}, \mathbf{Y} \rangle = \sum_i \sum_j x_{ij}y_{ji}$ denotes the inner product of matrices $\mathbf{X} = ((x_{ij}))$ and $\mathbf{Y} = ((y_{ij}))$.

Proposition 3

$\mathbf{X}\mathbf{a}$ and $\mathbf{Y}\mathbf{b}$ are orthogonal, i.e., $\langle \mathbf{X}\mathbf{a}, \mathbf{Y}\mathbf{b} \rangle = 0$ for every pair of vectors \mathbf{a} and \mathbf{b} , if and only if $\mathbf{X}^T\mathbf{Y} = \mathbf{0}$, where $\mathbf{0}$ denotes a matrix with all elements zeros.

Proof

If the former holds, then we have $\max_{\mathbf{a}, \mathbf{b}} |\mathbf{a}^T \mathbf{X}^T \mathbf{Y} \mathbf{b}| = 0$. It follows that the first singular value of $\mathbf{X}^T\mathbf{Y}$ is zero. The converse is obvious. QED.

The orthogonality between configuration classes \overline{C}_{δ_1} and \overline{C}_{δ_2} however, requires that $\mathbf{X}^T\mathbf{H}\mathbf{Y} = \mathbf{0}$ for any $\mathbf{X} \in \overline{M}_{\delta_1}$ and for any $\mathbf{Y} \in \overline{M}_{\delta_2}$. Specifically, the inner product of \mathbf{X} and \mathbf{Y} are required to be translation invariant. This condition can be satisfied if $\mathbf{H}\mathbf{X}\mathbf{a}$ and $\mathbf{H}\mathbf{Y}\mathbf{b}$ are orthogonal for every pair of \mathbf{a} and \mathbf{b} . Conversely, if so, then the associated configuration classes are orthogonal. Then, $\mathbf{X}\mathbf{a}$ and $\mathbf{Y}\mathbf{b}$ are said to be *uncorrelated* as summarized in the following Corollary to Proposition 3. It should, however, be noted that even if $\mathbf{X}^T\mathbf{H}\mathbf{Y} = \mathbf{Z}^T\mathbf{H}\mathbf{Y} = \mathbf{0}$ for $\mathbf{Z} \in \overline{M}_{\delta_1}$, it does not necessarily follow that $\overline{C}_{\delta_1} = \overline{C}_{\delta_2}$.

Corollary 1

$\mathbf{X}\mathbf{a}$ and $\mathbf{Y}\mathbf{b}$ are uncorrelated, i.e., $\langle \mathbf{H}\mathbf{X}\mathbf{a}, \mathbf{H}\mathbf{Y}\mathbf{b} \rangle = 0$ for every pair of vectors \mathbf{a} and \mathbf{b} if and only if $\mathbf{X}^T\mathbf{H}\mathbf{Y} = \mathbf{0}$.

5. FIVE PROPERTIES REQUIRED FOR A CLOSENESS COEFFICIENT

The following properties from (C-1) to (C-5) are required a closeness coefficient, say ρ_C , between two configurations: (C-1): $0 \leq \rho_C(C_{\delta_1}, C_{\delta_2}) \leq 1$, for all C_{δ_1} and C_{δ_2} ; (C-2): $\rho_C(C_{\delta_1}, C_{\delta_3}) = \rho_C(C_{\delta_2}, C_{\delta_3})$, for any C_{δ_3} if $\underline{\delta}_1 = \underline{\delta}_2$;

(C-3): $\rho_C(C_{\delta_1}, C_{\delta_2}) = 1$ if and only if $\underline{\delta}_1 = \underline{\delta}_2$; (C-4): $\rho_C(C_{\delta_1}, C_{\delta_2}) = 0$ if and only if \overline{C}_{δ_1} and \overline{C}_{δ_2} are orthogonal; and (C-5): $\rho_C(C_{\delta_1}, C_{\delta_2}) = \rho_C(C_{\delta_2}, C_{\delta_1})$, for any pair of C_{δ_1} and C_{δ_2} . Specifically, the coefficient is required to be: between 0 and 1 for any pair of configurations (C-1); TRD-invariant, i.e., the same for all pairs of members from classes to which they belong (C-2); zero between orthogonal configurations from orthogonal classes (C-3); one between configurations from the same class (C-4); and symmetric between any pair of configurations (C-5).

Embedding configurations onto coordinate matrices, we may quantify the closeness coefficient between configurations through a coefficient, say ρ_M , between coordinate matrices. In particular, this coefficient ρ_M is required to be invariant over choice of $\mathbf{X} \in \overline{M}_{\delta_1}$ and any $\mathbf{Y} \in \overline{M}_{\delta_2}$, i.e., TRD-invariant (M-2). Again, ρ_M must satisfy the following five properties corresponding to (C-1) to (C-5):

- (M-1): $0 \leq \rho_M(\mathbf{X}, \mathbf{Y}) \leq 1$, for all \mathbf{X} and \mathbf{Y} .
- (M-2): $\rho_M(\mathbf{X}, \mathbf{Y}) = \rho_M(\mathbf{X}, \mathbf{Z})$ for any \mathbf{Z} if $\underline{d}(\mathbf{X}) = \underline{d}(\mathbf{Y})$.
- (M-3): $\rho_M(\mathbf{X}, \mathbf{Y}) = 1$ if and only if $\underline{d}(\mathbf{X}) = \underline{d}(\mathbf{Y})$.
- (M-4): $\rho_M(\mathbf{X}, \mathbf{Y}) = 0$ if and only if $\mathbf{X}\underline{\mathbf{a}}$ and $\mathbf{Y}\underline{\mathbf{b}}$ are uncorrelated for all $\underline{\mathbf{a}}$ and $\underline{\mathbf{b}}$.
- (M-5): $\rho_M(\mathbf{X}, \mathbf{Y}) = \rho_M(\mathbf{Y}, \mathbf{X})$, for any pair of \mathbf{X} and \mathbf{Y} .

6. TRD-INVARIANT GOODNESS OF FIT OF PROCRUSTES METHOD

The closeness of two configuration classes \overline{C}_{δ_1} and \overline{C}_{δ_2} can be defined by means of the distances between their associated matrices in \overline{M}_{δ_1} and \overline{M}_{δ_2} , respectively. However, the distances should be the same over choice of pairs of \mathbf{X} in \overline{M}_{δ_1} and \mathbf{Y} in \overline{M}_{δ_2} , i.e., the distances should be TRD-invariant. Specifically, the distances between $\beta_x(\mathbf{X} + \mathbf{1}\alpha_x^T)\mathbf{R}_x$ and $\beta_y(\mathbf{Y} + \mathbf{1}\alpha_y^T)\mathbf{R}_y$ must be identical over non-zero scalars β_x and β_y , vectors α_x and α_y , and orthogonal matrices \mathbf{R}_x and \mathbf{R}_y . Such a TRD-invariant distance can be achieved by the following infimum, that is,

$$\min_{\{\mathbf{X} \in \overline{M}_{\delta_1}\}} \min_{\{\mathbf{Y} \in \overline{M}_{\delta_2}\}} \|\mathbf{X} - \mathbf{Y}\|^2 = \inf_{\{\beta_x \neq 0, \beta_y \neq 0\}} \min_{\{\alpha_x, \alpha_y, \mathbf{R}_x, \mathbf{R}_y\}} \|\beta_x(\mathbf{X} + \mathbf{1}\alpha_x^T)\mathbf{R}_x - \beta_y(\mathbf{Y} + \mathbf{1}\alpha_y^T)\mathbf{R}_y\|^2.$$

This infimum is in fact the same as

$\inf_{\{\beta \neq 0\}} \min_{\{\alpha, \mathbf{R}\}} \|\mathbf{X} - \mathbf{1}\alpha^T - \beta\mathbf{Y}\mathbf{R}\|^2$, which is again the same as $\inf_{\beta \neq 0} \min_{\mathbf{R}} \|\mathbf{H}\mathbf{X} - \beta\mathbf{H}\mathbf{Y}\mathbf{R}\|^2$. The orthonormal matrix that minimizes the distance can be obtained by the Procrustes method [1,14-16].

It is, however, necessary to normalize this infimum distance so that (\mathbf{X}, \mathbf{Y}) closeness can be compared to (\mathbf{X}, \mathbf{Z}) closeness for any \mathbf{Z} . This can be achieved by normalization of the maximum of the above infima over all matrices with the same order. It follows that a goodness of fit coefficient, denoted by GF, of the Procrustes method can be constructed as:

$$GF(\mathbf{X}, \mathbf{Y}) = 1 - \frac{\inf_{\beta \neq 0} \min_{\mathbf{R}} \|\mathbf{H}\mathbf{X} - \beta\mathbf{H}\mathbf{Y}\mathbf{R}\|^2}{\max_{\mathbf{w}} \inf_{\beta_w \neq 0} \min_{\mathbf{R}_w} \|\mathbf{H}\mathbf{X} - \beta_w \mathbf{H}\mathbf{W}\mathbf{R}_w\|^2}.$$

It is clear by this definition that $0 \leq GF(\mathbf{X}, \mathbf{Y}) \leq 1$, which satisfies property (M-1). It will be shown in what follows that GF satisfies all the properties required for a closeness coefficient. To this end, we start with the following proposition.

Proposition 4

For any given matrix \mathbf{X} ,

$$\max_{\mathbf{w}} \left(\inf_{\beta_w \neq 0} \min_{\mathbf{R}_w} \|\mathbf{X} - \beta_w \mathbf{W}\mathbf{R}_w\|^2 \right) = \|\mathbf{X}\|^2.$$

Proof

Since $\min_{\mathbf{R}} \|\mathbf{X} - \mathbf{Y}\mathbf{R}\|^2 = \|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2 - \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\}$ for orthogonal matrix \mathbf{R} [15], $\inf_{\beta} \min_{\mathbf{R}} \|\mathbf{X} - \beta\mathbf{Y}\mathbf{R}\|^2 = \|\mathbf{X}\|^2 + \inf_{\beta} \left\{ \beta^2 \|\mathbf{Y}\|^2 - 2\beta \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} \right\}$. Therefore, the infimum can be achieved when $\beta = \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} / \|\mathbf{Y}\|^2$. Furthermore, we have $\inf_{\beta_w} \min_{\mathbf{R}_w} \|\mathbf{X} - \beta_w \mathbf{W}\mathbf{R}_w\|^2 = \|\mathbf{X}\|^2 - \text{trace}\{(\mathbf{X}^T\mathbf{W}\mathbf{W}^T\mathbf{X})^{1/2}\} / \|\mathbf{W}\|^2$ for any \mathbf{W} . It is because $\inf_{\beta} \min_{\mathbf{R}} \|\mathbf{X} - \beta\mathbf{Y}\mathbf{R}\|^2 = \|\mathbf{X}\|^2 - \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} / \|\mathbf{Y}\|^2$ for orthogonal matrix \mathbf{R} and non-zero β . Therefore we may define the second term on the right hand side as 0 if $\mathbf{W} = \mathbf{0}$, so that $\inf_{\beta} \min_{\mathbf{R}} \|\mathbf{X} - \beta\mathbf{0}\mathbf{R}\|^2 = \inf_{\beta} \min_{\mathbf{R}} \|\mathbf{X}\|^2 = \|\mathbf{X}\|^2$. It follows that $\max_{\mathbf{w}} \inf_{\beta_w \neq 0} \min_{\mathbf{R}_w} \|\mathbf{H}\mathbf{X} - \beta_w \mathbf{H}\mathbf{W}\mathbf{R}_w\|^2 \leq \|\mathbf{X}\|^2$. On the other hand, we have: $\max_{\mathbf{w}} \inf_{\beta_w \neq 0} \min_{\mathbf{R}_w} \|\mathbf{H}\mathbf{X} - \beta_w \mathbf{H}\mathbf{W}\mathbf{R}_w\|^2 \geq \inf_{\beta_0} \min_{\mathbf{R}_0} \|\mathbf{X} - \beta_0 \mathbf{0}\mathbf{R}_0\|^2 = \|\mathbf{X}\|^2$. QED.

Due to this Proposition, the GF coefficient can be reduced to

$$\text{GF}(\mathbf{X}, \mathbf{Y}) = 1 - \inf_{\beta} \min_{\mathbf{R}} \|\mathbf{HX} - \beta\mathbf{HYR}\|^2 / \|\mathbf{HX}\|^2.$$

Furthermore, since $\inf_{\beta} \min_{\mathbf{R}} \|\mathbf{X} - \beta\mathbf{YR}\|^2 = \|\mathbf{X}\|^2 - \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} / \|\mathbf{Y}\|^2$ for orthogonal matrix \mathbf{R} and non-zero β , the $\text{GF}(\mathbf{X}, \mathbf{Y})$ can be rewritten as

$$\text{GF}(\mathbf{X}, \mathbf{Y}) = \left[\text{trace}\{(\mathbf{X}^T\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}\mathbf{X})^{1/2}\} \right]^2 / \left\{ \|\mathbf{HX}\|^2 \|\mathbf{HY}\|^2 \right\}, \quad (3)$$

where the function *trace* is the sum of diagonal elements of a square matrix. This equation shows that the GF coefficient satisfies property (M-2) because it is free of β and \mathbf{R} , and thus TRD-invariant. For property (M-3), we have the following Proposition.

Proposition 5

$\text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} = \|\mathbf{HX}\| \|\mathbf{HY}\|$ if and only if $\mathbf{HX} = \beta\mathbf{HYR}$ for some non-zero scalar β and orthogonal matrix \mathbf{R} .

Proof

It is clear that the latter is sufficient for the former. The former implies that $\inf_{\beta} \min_{\mathbf{R}} \|\mathbf{HX} - \beta\mathbf{HYR}\|^2 = 0$ because $\left[\text{trace}\{(\mathbf{X}^T\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}\mathbf{X})^{1/2}\} \right]^2 / \left\{ \|\mathbf{HX}\|^2 \|\mathbf{HY}\|^2 \right\} = 1 - \inf_{\beta} \min_{\mathbf{R}} \|\mathbf{HX} - \beta\mathbf{HYR}\|^2 / \|\mathbf{HX}\|^2$. Therefore, we have $\mathbf{HX} = \beta\mathbf{HYR}$ for some non-zero scalar β and orthogonal matrix \mathbf{R} . QED.

It follows that $\text{GF}(\mathbf{X}, \mathbf{Y}) = 1$ if and only if \mathbf{X} and \mathbf{Y} belong to the same \overline{M}_{δ} , i.e., $\overline{M}_{\delta_1} = \overline{M}_{\delta_2} = \overline{M}_{\delta}$. For property (M-4), we have the following Proposition.

Proposition 6

$\text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} = 0$ if and only if $\mathbf{X}^T\mathbf{Y} = \mathbf{0}$.

Proof

It can be seen in the process of proof of Proposition 5 above that the former implies that all the singular values of $\mathbf{X}^T\mathbf{Y}$ are zeros. Hence the latter follows. The converse is clear. QED.

It follows that $\text{trace}\{(\mathbf{X}^T\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}\mathbf{X})^{1/2}\} = 0$ if and only if $v(\mathbf{HX})$ is orthogonal to $v(\mathbf{HY})$, i.e., $\mathbf{X}\mathbf{a}$ and $\mathbf{Y}\mathbf{b}$ are uncorrelated for any pair of \mathbf{a} and \mathbf{b} . Finally, for property (M-5), we have the following Proposition that shows that the GF coefficient is symmetric, i.e., $\text{GF}(\mathbf{X}, \mathbf{Y}) = \text{GF}(\mathbf{Y}, \mathbf{X})$ for any pair of \mathbf{X} and \mathbf{Y} .

Proposition 7

$\text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} = \text{trace}\{(\mathbf{Y}^T\mathbf{X}\mathbf{X}^T\mathbf{Y})^{1/2}\}$.

Proof

Let $\mathbf{X}^T\mathbf{Y} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$. Then $(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$. It follows that $\text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\} = \sum_i \lambda_i$ where λ_i 's are the singular values. Similarly, $(\mathbf{Y}^T\mathbf{X}\mathbf{X}^T\mathbf{Y})^{1/2} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$. It follows that $\text{trace}\{(\mathbf{Y}^T\mathbf{X}\mathbf{X}^T\mathbf{Y})^{1/2}\} = \sum_i \lambda_i$. QED.

7. TRD-INVARIANT MATRIX CORRELATION COEFFICIENTS

Closeness of homologous configurations C_{δ_1} and C_{δ_2} can also be quantified by means of a correlation between two coordinate matrices $\mathbf{X} \in \overline{M}_{\delta_1}$ and $\mathbf{Y} \in \overline{M}_{\delta_2}$ as follows [3]:

$$\text{corr}(\mathbf{X}, \mathbf{Y}) = \langle \mathbf{HX}, \mathbf{HY} \rangle / \{ \|\mathbf{HX}\| \|\mathbf{HY}\| \} = \text{trace}(\mathbf{X}^T\mathbf{HY}) / \{ \|\mathbf{HX}\| \|\mathbf{HY}\| \}.$$

Although the correlation $\text{corr}(\mathbf{X}, \mathbf{Y})$ clearly satisfies required properties (M-3), (M-4) and (M-5), it can be negative and is not invariant over orthogonal \mathbf{R} because $\text{trace}(\mathbf{X}^T\mathbf{HY}) \neq \text{trace}(\mathbf{X}^T\mathbf{HYR})$ for every orthogonal matrix \mathbf{R} .

We, therefore, examined whether the maximum correlation

$$\text{CORR}(\mathbf{X}, \mathbf{Y}) = \max_{\beta, \mathbf{R}} \text{corr}(\mathbf{X}, \beta\mathbf{YR})$$

between any pair of members (one from each set of \overline{M}_{δ_1} and \overline{M}_{δ_2}) satisfies the five required properties including the TRD-invariance. First, it is obvious that $\text{CORR}(\mathbf{X}, \mathbf{Y})$ ranges between 0 and 1, satisfying (M-1). Second, due to the following Proposition 8, $\text{CORR}(\mathbf{X}, \mathbf{Y})$ is TRD-invariant, thus satisfies (M-2), and is in fact the Gower-Lingoes-Schönenman coefficient [14, 17], denoted here by R_{GLS} . That is,

$$R_{GLS}(\mathbf{X}, \mathbf{Y}) = \text{trace}\{(\mathbf{X}^T\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}\mathbf{X})^{1/2}\} / \{ \|\mathbf{HX}\| \|\mathbf{HY}\| \}. \quad (4)$$

Proposition 8

The orthogonal matrix \mathbf{R} which minimizes $\|\mathbf{X} - \beta\mathbf{YR}\|^2$, for some non-zero β , also maximizes $\langle \mathbf{X}, \mathbf{YR} \rangle$ and $\max_{\mathbf{R}} \langle \mathbf{X}, \mathbf{YR} \rangle = \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\}$.

Proof

For the first part, observe that $\min_{\mathbf{R}} \|\mathbf{X} - \beta\mathbf{YR}\|^2 = \|\mathbf{X}\|^2 + \beta^2 \|\mathbf{Y}\|^2 - 2\beta \max_{\mathbf{R}} \langle \mathbf{X}, \mathbf{YR} \rangle$. The second part is clear since $\text{trace}(\mathbf{X}^T\mathbf{YR}) \leq \text{trace}\{(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\}$ [15], with equality if $\mathbf{X} = \mathbf{YR}$ for some orthogonal \mathbf{R} . QED.

It follows that

$$\text{CORR}(\mathbf{X}, \mathbf{Y}) = \max_{\beta, \mathbf{R}} \text{corr}(\mathbf{X}, \beta\mathbf{YR}) =$$

$$\max_{\beta, \mathbf{R}} \left[\langle \mathbf{HX}, \beta\mathbf{HYR} \rangle / \{ \|\mathbf{HX}\| \|\beta\mathbf{HYR}\| \} \right]$$

$$= \max_{\mathbf{R}} \left[\frac{\langle \mathbf{H}\mathbf{X}, \mathbf{H}\mathbf{Y}\mathbf{R} \rangle}{\{\|\mathbf{H}\mathbf{X}\| \|\mathbf{H}\mathbf{Y}\|\}} \right] = \text{trace}\{(\mathbf{X}^T \mathbf{H}\mathbf{Y}\mathbf{Y}^T \mathbf{H}\mathbf{X})^{1/2}\} / \{\|\mathbf{H}\mathbf{X}\| \|\mathbf{H}\mathbf{Y}\|\} = R_{GLS}(\mathbf{X}, \mathbf{Y}).$$

Subsequently, this equation shows that based on (3) and (4),

$$R_{GLS}(\mathbf{X}, \mathbf{Y}) = \sqrt{\text{GF}(\mathbf{X}, \mathbf{Y})}. \tag{5}$$

This equation implies that R_{GLS} is TRD-invariant, and proves that the maximization of inner product over rotation is equivalent to minimization of distances between coordinate matrices by translation, rotation and dilation.

In fact, the R_{GLS} coefficient (or GF) depends on $\mathbf{X} \in \overline{M}_{\delta_1}$ and $\mathbf{Y} \in \overline{M}_{\delta_2}$ through their centered forms $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}$ and $\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}$. This relation is due to the fact that $\text{trace}\{(\mathbf{X}^T \mathbf{H}\mathbf{Y}\mathbf{Y}^T \mathbf{H}\mathbf{X})^{1/2}\} = \text{trace}\{(\mathbf{Y}^T \mathbf{H}\mathbf{X}\mathbf{X}^T \mathbf{H}\mathbf{Y})^{1/2}\}$ based on Proposition 7, and both $\|\mathbf{H}\mathbf{X}\|$ and $\|\mathbf{H}\mathbf{Y}\|$ are a function of the centered forms. Moreover, the R_{GLS} coefficient can be expressed as a function of singular values of $\mathbf{H}\mathbf{X}$ and $\mathbf{H}\mathbf{Y}$ for $\mathbf{X} \in \overline{M}_{\delta_1}$ and $\mathbf{Y} \in \overline{M}_{\delta_2}$, respectively. To this end, let us denote the singular value decompositions of $\mathbf{H}\mathbf{X}$, $\mathbf{H}\mathbf{Y}$, and $\mathbf{X}^T \mathbf{H}\mathbf{Y}$ by

$$\mathbf{H}\mathbf{X} = \mathbf{U}_x \mathbf{\Lambda}_x \mathbf{V}_x^T, \mathbf{H}\mathbf{Y} = \mathbf{U}_y \mathbf{\Lambda}_y \mathbf{V}_y^T, \text{ and } \mathbf{X}^T \mathbf{H}\mathbf{Y} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T. \tag{6}$$

Although $\text{corr}(\mathbf{U}_x \mathbf{\Lambda}_x, \mathbf{U}_y \mathbf{\Lambda}_y)$ does not satisfy the invariance property, again its maximum over orthonormal matrix \mathbf{R} does since $\max_{\mathbf{R}} \text{corr}(\mathbf{U}_x \mathbf{\Lambda}_x, \mathbf{U}_y \mathbf{\Lambda}_y \mathbf{R}) = \max_{\mathbf{R}} \text{corr}(\mathbf{X}, \mathbf{Y}\mathbf{R}) = R_{GLS}(\mathbf{X}, \mathbf{Y})$. It follows that

$$R_{GLS}(\mathbf{X}, \mathbf{Y}) = \text{trace}(\mathbf{\Lambda}) / \sqrt{\text{trace}(\mathbf{\Lambda}_x^2) \text{trace}(\mathbf{\Lambda}_y^2)}. \tag{7}$$

8. COEFFICIENTS BASES ON CENTERED FORMS

A correlation between centered forms $\text{corr}(\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}, \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H})$ is the same by definition as the RV coefficient of Escoufier [18,19]:

$$\text{corr}(\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}, \mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H}) = \text{trace}(\mathbf{H}\mathbf{X}\mathbf{X}^T \mathbf{H}\mathbf{Y}\mathbf{Y}^T \mathbf{H}) / \{\|\mathbf{H}\mathbf{X}\mathbf{X}^T \mathbf{H}\| \|\mathbf{H}\mathbf{Y}\mathbf{Y}^T \mathbf{H}\|\} = \text{RV}(\mathbf{X}, \mathbf{Y}).$$

It is a normalized inner product between two centered forms, and can be interpreted as a pooled correlation coefficient of inner products of row vectors between $\mathbf{H}\mathbf{X}$ and $\mathbf{H}\mathbf{Y}$. The RV coefficient also satisfies all the required properties. In order to show this, it would be enough to hand the dilation factor by minimizing over β since every configuration class has a centered form that is unique up to a dilation factor (section 2). To this end, based on Proposition 8, the RV can be rewritten as a minimum distance between two centered forms as follows:

$$\text{RV}(\mathbf{X}, \mathbf{Y}) = 1 - \sqrt{\inf_{\beta} \|\mathbf{H}\mathbf{X}\mathbf{X}^T \mathbf{H} - \beta \mathbf{H}\mathbf{Y}\mathbf{Y}^T \mathbf{H}\|^2} / \|\mathbf{H}\mathbf{X}\mathbf{X}^T \mathbf{H}\|}.$$

Furthermore, the RV coefficient can be re-expressed as $\text{RV}(\mathbf{X}, \mathbf{Y}) = \text{corr}(\mathbf{U}_x \mathbf{\Lambda}_x^2 \mathbf{U}_x^T, \mathbf{U}_y \mathbf{\Lambda}_y^2 \mathbf{U}_y^T)$, where $\mathbf{H}\mathbf{X}\mathbf{X}^T\mathbf{H}$

$\stackrel{\text{svd}}{=} \mathbf{U}_x \mathbf{\Lambda}_x^2 \mathbf{U}_x^T$ and $\mathbf{H}\mathbf{Y}\mathbf{Y}^T\mathbf{H} \stackrel{\text{svd}}{=} \mathbf{U}_y \mathbf{\Lambda}_y^2 \mathbf{U}_y^T$ as in (6). It follows that $\text{RV}(\mathbf{X}, \mathbf{Y})$ can be reduced to

$$\text{RV}(\mathbf{X}, \mathbf{Y}) = \text{trace}(\mathbf{\Lambda}^2) / \sqrt{\text{trace}(\mathbf{\Lambda}_x^4) \text{trace}(\mathbf{\Lambda}_y^4)},$$

a function of singular values only of the singular value decompositions like the R_{GLS} coefficient as shown in equation (7). Finally, based on the equation (1), it can be shown that

$$\text{RV}(\mathbf{X}, \mathbf{Y}) = \text{corr}(\mathbf{H} \Delta_{\delta_1}^2 \mathbf{H}, \mathbf{H} \Delta_{\delta_2}^2 \mathbf{H}),$$

which is a correlation between the double centered squared distances matrices. This implies that the RV coefficient is ready to be obtained without obtaining coordinate matrices when squared distances matrices are given.

9. DISCUSSION

We have shown that development of a closeness coefficient that satisfies the five properties was possible by minimization of distances between coordinate matrices, maximization of correlation between them, or correlations between unique representations of configurations. The notion of minimization of distances between matrices over the two sets is shown to be equivalent to that of maximization of correlation by Proposition 8 that resulted in equation (5) between GF and R_{GLS} . A univariate analogy is that the goodness of fit R^2 obtained from a simple linear regression is the square of the Pearson correlation between the dependent and independent variables. Furthermore, both R_{GLS} and RV are shown to be a function of unique representations of configurations through centered forms or singular values of the coordinate matrices. As a result, those existing matrix correlation coefficients are now proven to satisfy the five properties, including the “zero” property under the newly introduced orthogonality of configurations in section 5.

Comparison between R_{GLS} and RV is discussed in [4] and ranges of RV in terms of R_{GLS} , or vice versa, are suggested in [20]. The permutational distribution of the RV coefficient under permutation of homologous objects is derived in [21]. The distribution can be used to test significance of RV between homologous configurations. As far as lower-rank approximations are concerned, Heo and Gabriel [5] discussed behaviors of R_{GLS} with varying dimensions of coordinate matrices, and Heo [22] discussed distributions of R_{GLS} under a null situation.

Another approach of developing a closeness coefficient could be to deal directly with the distance sets $\underline{\delta}_1$ and $\underline{\delta}_2$. This approach is attractive because the distances are unique up to a dilation factor for every configuration class, even if not Euclidean. In this context, Mantel’s cosine coefficient [23], $\text{cos}(\underline{\delta}_1, \underline{\delta}_2) = \langle \underline{\delta}_1, \underline{\delta}_2 \rangle / \{\|\underline{\delta}_1\| \|\underline{\delta}_2\|\}$, can serve for that purpose. This coefficient is simple, intuitive, and practical because the distances are given in practice more often than not. In addition, this coefficient is clearly unique between configuration classes because it is invariant over dilation. Therefore, $\text{cos}(\underline{\delta}_1, \underline{\delta}_2)$ is TRD-invariant. Nevertheless, $\text{cos}(\underline{\delta}_1, \underline{\delta}_2)$ is not

necessarily zero when two configuration classes are orthogonal. In fact, this coefficient is zero if and only if either of the two configuration classes is a point, i.e., of zero dimension.

In summary, R_{GLS} , GF, and RV can be used as a closeness coefficient of homologous configurations, and each depends on coordinate matrices through their centered forms or singular values.

ACKNOWLEDGEMENT

This paper is written in memory of the late Professor Ruben K. Gabriel, who guided the author's doctoral dissertation, the result of which is a part of this paper.

REFERENCES

- [1] J.M.F. Ten Berge, and D.L. Knol, "Orthogonal rotations to maximal agreement for two or more matrices of different column orders", *Psychometrika*, vol. 49, pp. 49-55, 1984.
- [2] C. Goodall, "Procrustes methods in the statistical analysis of shape (with discussion)", *J. R. Stat. Soc. B*, vol. 53, pp. 283-339, 1991.
- [3] J.O. Ramsay, J.M.F. Ten Berge, and G.P.H. Styan, "Matrix Correlation", *Psychometrika*, vol. 49, pp. 403-423, 1984.
- [4] M. Heo, *On the Fit of Sample Graphical Displays to Patterns in Populations*. Unpublished Ph.D. Dissertation, The University of Rochester, NY, Rochester, 1996.
- [5] M. Heo, and K.R. Gabriel, "The fit of graphical displays of data to patterns of expectations", *Comput. Stat. Data Anal.*, vol. 36, pp. 47-67, 2001.
- [6] K.R. Gabriel, "Relating multivariate data to geographical location", in *Data Analysis, Expert Knowledge and Decisions*, W. Gaul, M. Schader, Eds. Berlin: Springer, 1988, pp. 341-354.
- [7] A.K. Smilde, H.A.L. Kiers, S. Bijlsma, C.M. Rubingh, and M.J. van Erk, "Matrix correlations for high-dimensional data: the modified RV-coefficient", *Bioinformatics*, vol. 25, pp.410-405, 2009.
- [8] O. Bastien, P. Ortet, S. Roy, and E. Maréchal, "A configuration space of homologous proteins conserving mutual information and allowing a phylogeny inference based on pair-wise Z-score probabilities", *BMC Bioinformatics*, vol. 6, p. 49, 2005.
- [9] W.S. Torgerson, *Theory and Methods of Scaling*. New York: Wiley, 1958.
- [10] W.J. Krzanowski, *Principles of Multivariate Analysis*, Oxford: Clarendon Press, 1988.
- [11] C. Eckart, and G. Young, "The approximation of one matrix by another of lower rank", *Psychometrika*, vol. 1, p. 211-218, 1936.
- [12] A.S. Householder, and G. Young, "Matrix approximation and latent roots", *Am. Math. Mon.*, vol. 45, pp. 165-171, 1938.
- [13] J.C. Gower, "Some distance properties of latent root and vector methods used in multivariate analysis", *Biometrika*, vol. 53, pp. 325-338, 1966.
- [14] J.C. Gower, "Statistical methods of comparing different multivariate analyses of the same data", in *Mathematics in the Archeological and Historical Sciences*, F.R. Hodson, D.G. Kendall, P. Tautu Eds. Edinburgh: University Press, 1971, pp.138-149.
- [15] R. Sibson, "Studies in the robustness of multidimensional scaling: Procrustes statistics", *J. R. Statist. Soc. B*, vol. 41, pp. 217-219, 1978.
- [16] K.R. Gabriel, and M. Heo, *Rotational fitting and correlation of coordinates and configuration*, Technical Report 96/10, Department of Biostatistics, University of Rochester, 1996.
- [17] J.C. Lingoes, and P.H. Schönemann, "Alternative measures for fit for the Schönemann-Carroll matrix fitting algorithm", *Psychometrika*, vol. 39, pp. 423-427, 1974.
- [18] Y. Escoufier, "Le traitement des variables vectorielles", *Biometrics*, vol. 29, pp. 751-769, 1973.
- [19] P. Robert, and Y. Escoufier, "A unifying tool for linear multivariate statistical methods: The RV-coefficient", *Appl. Stat.*, vol. 25, pp. 257-265, 1976.
- [20] A. Lazraq, R. Cléroux, and H.A.L. Kiers, "Mesures de liaison vectorielle et generalization de l'analyse canonique", *Rev Statistique Appliquée*, vol. 39, pp. 23-35, 1992.
- [21] M. Heo, and K.R. Gabriel, "A permutation test of association between configurations by means of the RV Coefficients", *Commun. Stat. Simul. Comput.*, vol. 27, pp. 843-856, 1998.
- [22] M. Heo, "Exploration of distributions of ratios of a partial sum of sample eigenvalues when all population eigenvalues are the same", *J. Mod. Appl. Stat. Meth.*, vol. 1, pp. 420-427, 2002.
- [23] N. Mantel, "The detection of disease clustering and a generalized regression approach", *Cancer Res.*, vol. 27, pp. 209-220, 1967.

Received: January 28, 2009

Revised: February 24, 2009

Accepted: March 12, 2009

© Moonseong Heo; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.