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 RGLS 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, RGLS, 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 \( \omega_1, ..., \omega_n \) corresponding to a collection of given distances \( \delta_{ie} \), \( i, e=1, ..., n \) between all pairs of the objects. The inter-object distances \( \hat{\delta} \) should satisfy that: (1) \( \hat{\delta}_{ie} \geq 0 \) for all \( i, e \) with equality if \( i = e \); (2) \( \hat{\delta}_{ie} = \hat{\delta}_{ei} \) for all \( i \) and \( e \); (3) \( \hat{\delta}_{ij} \leq \hat{\delta}_{ie} + \hat{\delta}_{ej} \) for all \( i, j, \) and \( e \). Therefore, \( C = C(\Omega, \hat{\delta}) \) is a function of both a set \( \Omega = \{ \omega_1, ..., \omega_n \} \) of the ordered objects and a set \( \hat{\delta} = \{ \hat{\delta}_{ie}, i,e=1,....,n \} \) of the inter-object distances.
Two configurations for the same objects, \( C_1 = C(\Omega_1, \delta) \) and \( C_2 = C(\Omega_2, \delta) \), are homologous if there exists a one-to-one association between the ordered elements of \( \Omega_1 \) and \( \Omega_2 \). For example, both \( \Omega_1 \) and \( \Omega_2 \) represent geographical coordinates of certain locations, while \( \delta \) in \( C_1 \) and \( \delta \) 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 \( \Omega \) as given. This reduces the configurations to a function of distances \( \delta \) only and enables to write \( C_\delta \) for \( C(\Omega, \delta) \).

To develop a closeness coefficient between homologous configurations, we embed configurations onto a space defined by coordinates \( \mathbf{x} \) associated with each element \( \omega_i \), \( i = 1, ..., n \). The \( n \)-rowed matrix \( X = ((x_{ij})) \) with the \( i \)-th row vector \( \mathbf{x}_i^T = (x_{i1}, x_{i2}, ..., x_{ij}) \) for some \( J \) is a list of the coordinates of all the elements of \( \Omega \), where \( ((x_{ij})) \) denote a matrix whose \( j \)-th element is \( x_{ij} \). The relation between the coordinates \( \mathbf{X} \) of \( \Omega \) and the configuration \( C_\delta \) depends on a norm which equates the distances between coordinates \( \mathbf{x}_1 \) and \( \mathbf{x}_2 \) with the distances \( \delta_1 \) between elements \( \omega_1 \) and \( \omega_2 \). To this end, we consider only the Euclidean norm \( \| \| \) in this paper. A configuration \( C_\delta \) is said to be Euclidean embeddable if and only if there exists \( \mathbf{X} \) such that \( d(X) = \delta \) where \( d(X) = (\|x_i - x_j\|, \omega = 1, ... , n) \) is the set of Euclidean distances measured by the Euclidean norm \( \| \| \) between all pairs of the coordinates of the elements in \( \Omega \). Any such \( \mathbf{X} \) is called a (Euclidean) coordinate matrix of \( C_\delta \).

For a given \( C_\delta \), its associated coordinate matrices are not unique but there exist a collection of coordinate matrices denoted by \( M_\delta = \{ \mathbf{X} | d(X) = \delta \} \), where \( \{ e \} \subseteq \mathbb{R} \) represents a set of elements \( e \) that satisfy condition \( c \). Furthermore, for a given set \( \delta \) of distances, there exists a class of configurations denoted by \( C_\delta = \{ C_\gamma | \gamma = \gamma \} \) in which normalized distances of every configuration are identical to the given normalized distances \( \delta = \gamma / \| \gamma \| \). By this definition, \( \delta = \gamma \) if and only if there exists a non-zero scalar \( \beta \) such that \( \delta = \beta \gamma \). 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 \( C_\delta \) and \( C_\gamma \) are said to be equivalent if they belong to the same class, i.e., if \( \delta = \gamma \).

Correspondingly, a class of collections of coordinate matrices of a configuration class \( C_\delta \), denoted by \( M_\delta = \{ M_j | \delta = \gamma \} \) = \{ \mathbf{X} | d(X) = \delta \}, in which normalized inter-row distances \( d(X) = d(X) / \|d(X)\| \) of any \( \mathbf{X} \) are identical to the given normalized inter-object distances \( \delta \). Subsequently, coordinate matrices \( \mathbf{X} \) and \( \mathbf{Y} \) are said to be equivalent if they belong to the same class, i.e., if \( d(X) = d(Y) \).

### 3. UNIQUE REPRESENTATIONS OF CONFIGURATIONS AND TRD-ININVANCE

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

\[
D^2(X) = \text{Diag}(\mathbf{X}^T)11^T + 11^T \text{Diag}(\mathbf{X}^T)11^T - 2\mathbf{X}^T \mathbf{X},
\]

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

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

\[
\mathbf{HXX}^T = -0.5\mathbf{HDD}^2(\mathbf{X})\mathbf{H} = -0.5\mathbf{HH}^2\mathbf{Y} = -0.5\mathbf{HDD}^2(\mathbf{Y})\mathbf{H} = \mathbf{HYY}^T \mathbf{H},
\]

where \( \mathbf{H} = \mathbf{I} - \mathbf{11}^T/n \) is an idempotent column-centering operator such that \( \mathbf{HH} = \mathbf{H} \) and \( \mathbf{I} \) is the \( n \)-by-\( n \) identity matrix. In other words, \( \Delta_\varphi^2 \) is a one-to-one function not of \( \mathbf{X}^T \) but of the centered form \( \mathbf{HXX}^T \mathbf{H} \). Therefore, the centered form is unique for a given squared distances matrix and also for a configuration. Furthermore, the centered form \( \mathbf{HXX}^T \mathbf{H} \) is invariant over translation and rotation of \( \mathbf{X} \), i.e., TR invariant, since \( \mathbf{H}[(\mathbf{X} + \mathbf{1}q^T)\mathbf{R}][(\mathbf{X} + \mathbf{1}q^T)\mathbf{R}^T] = \mathbf{HXX}^T \mathbf{H} \) for any vector \( \mathbf{q} \) and any orthonormal matrix \( \mathbf{R} \) such that \( \mathbf{RR}^T = \mathbf{R}^T \mathbf{R} = \mathbf{I} \). This TR-invariant property is summarized in the following Proposition 1.

**Proposition 1**

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

**Proof**

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

**Claim**

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

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

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

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 \( C \) and \( \frac{X}{\sqrt{D^T(X)/I^T D^T(X)}} \) respectively, are unique and related uniquely to the centered form of \( X \). Therefore, for a configuration class the coordinates allow \( \beta X + 1 \alpha^T \) transformation to be invariant over translation, rotation and dilation of any coordinate matrix in \( M \), 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 \( C \), its corresponding Euclidean coordinate matrices \( X \) and \( Y \), with the same full affine ranks, belong to the same class of collections \( M \) if and only if there exist a coordinate vector \( \alpha \), a non-zero scalar \( \beta \), and an orthonormal matrix \( R \) such that \( X = \beta (Y + 1 \alpha^T) R \); also, if and only if their centered forms satisfy \( HX^TH = \beta HY^TH \) for some non-zero \( \beta \).

Proof

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

4. ORTHOGONALITY OF CONFIGURATION CLASSES

In univariate (or single dimensional) case, as the opposite of perfect correlation, there is zero correlation between two variables, 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 \( C_{\delta} \) and \( C_{\delta} \) 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 \( X^T Y = 0 \) for any \( X \in M_\delta \) and any \( Y \in M_\delta \). The condition \( X^T Y = 0 \) implies that \( v(X) \) is orthogonal to \( v(Y) \), i.e., \( \langle X, Y \rangle = 0 \), hence \( 0 = \sum_i \sum_j x_i y_j \) denotes the inner product of matrices \( X = (x_i) \) and \( Y = ((y_j)) \).

Proposition 3

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

Proof

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

The orthogonality between configuration classes \( \overline{C_\delta} \) and \( \overline{C_\delta} \) however, requires that \( X^T Y = 0 \) for any \( X \in \overline{M_\delta} \) and any \( Y \in \overline{M_\delta} \). Specifically, the inner product of \( X \) and \( Y \) are required to be translation invariant. This condition can be satisfied if \( HX_a \) and \( HY_b \) are orthogonal for every pair of \( a \) and \( b \). Conversely, if so, then the associated configuration classes are orthogonal. Then, \( X_a \) and \( Y_b \) are said to be uncorrelated as summarized in the following Corollary to Proposition 3. It should, however, be noted that even if \( X^T Y = Z^T \) \( Y = 0 \) for \( Z \in \overline{M_\delta} \), it does not necessarily follow that \( \overline{C_\delta} = \overline{C_\delta} \).

Corollary 1

\( X_a \) and \( Y_b \) are uncorrelated, i.e., \( \langle HX_a, HY_b \rangle = 0 \) for every pair of vectors \( a \) and \( b \) if and only if \( X^T Y = 0 \).

5. FIVE PROPERTIES REQUIRED FOR A CLOSENESS COEFFICIENT

The following properties from (C-1) to (C-5) are required a closeness coefficient \( \rho \), between two configurations:

\( C \): \( 0 \leq \rho(C, C_{\delta}) \leq 1 \), for all \( C \) and \( C_{\delta} \); \( C \): \( C_{\delta} = C_{\delta} \), for any \( C_{\delta} \) if \( \delta = \delta \);
(C-3): \( \rho_c(C_{\delta}, C_{\delta}) = 1 \) if and only if \( \delta_\ast = \delta_\ast \); (C-4): 
\( \rho_c(C_{\delta}, C_{\delta}) = 0 \) if and only if \( \overline{C_{\delta}} \) and \( \overline{C_{\delta}} \) are orthogonal; and (C-5): 
\( \rho_c(C_{\delta}, C_{\delta}) = \rho_c(C_{\delta}, C_{\delta}) \) for any pair of 
\( C_{\delta} \) and \( C_{\delta} \). 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 \( \rho_M \), between coordinate matrices. In particular, this coefficient \( \rho_M \) is required to be 
TRD-invariant over choice of \( X \in \overline{M_{\delta}} \) and any \( Y \in \overline{M_{\delta}} \), i.e., 
TRD-invariant (M-2). Again, \( \rho_M \) must satisfy the following 
five properties corresponding to (C-1) to (C-5):

(M-1): \( 0 \leq \rho_M(X, Y) \leq 1 \), for all \( X \) and \( Y \).

(M-2): \( \rho_M(X, Y) = \rho_M(X, Z) \) for any \( Z \) if \( d(X) = d(Y) \).

(M-3): \( \rho_M(X, Y) = 1 \) if and only if \( d(X) = d(Y) \).

(M-4): \( \rho_M(X, Y) = 0 \) if and only if \( X_a \) and \( Y_b \) are uncorrelated for all \( a \) and \( b \).

(M-5): \( \rho_M(X, Y) = \rho_M(Y, X) \), for any pair of \( X \) and \( Y \).

6. TRD-INARIANT GOODNESS OF FIT OF PROCRUSTES METHOD

The closeness of two configuration classes \( \overline{C_{\delta}} \) and \( \overline{C_{\delta}} \) can be defined by means of the distances between their associated matrices in \( \overline{M_{\delta}} \) and \( \overline{M_{\delta}} \), respectively. However, 
the distances should be the same choice of pairs of \( X \) in 
\( \overline{M_{\delta}} \) and \( Y \) in \( \overline{M_{\delta}} \), i.e., the distances should be TRD- 
invariant. Specifically, the distances between \( \beta_0(X + 1 \alpha_x^T R_x) \) and \( \beta_0(Y + 1 \alpha_y^T R_y) \) must be identical over non-zero 
scalars \( \beta_0 \) and \( \beta_0 \), vectors \( \alpha_x \) and \( \alpha_y \), and orthogonal matrices \( R_x \) and \( R_y \). Such a TRD-invariant distance can be achieved by the following infimum, that is,

\[ \min_{X \in \overline{M_{\delta}}} \min_{Y \in \overline{M_{\delta}}} \| X - Y \|^2 = \inf_{\beta_0 \neq 0, \beta_0 \neq 0} \| \beta_0(X + 1 \alpha_x^T R_x) - \beta_0(Y + 1 \alpha_y^T R_y) \|^2. \]

This infimum is in fact the same as

\[ \inf_{\beta_0 \neq 0} \min_{R_x, R_y} \| X - \beta_0 Y \|^2, \]

which is again the same as

\[ \inf_{\beta_0 \neq 0} \min_{R} \| HX - \beta Y \|^2. \]

This infimum is in fact the same as

\[ \inf_{\beta_0 \neq 0} \| HX - \beta Y \|^2. \]

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 \( (X, Y) \) closeness can be compared to \( (X, Z) \) 
closeness for any \( Z \). 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:

\[ \text{GF}(X, Y) = 1 - \frac{\inf_{\beta_0 \neq 0} \| HX - \beta Y \|^2}{\max_{\beta_0 \neq 0} \| HX - \beta Y \|^2}. \]

It is clear by this definition that \( 0 \leq \text{GF}(X, Y) \leq 1 \), 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 \( X \),

\[ \max_{w} \left\{ \inf_{\beta_0 \neq 0} \| X - \beta_0 Y \|^2 \right\} = \| X \|^2. \]

Proof

Since \( \min_{R} \| X - YR \|^2 = \| X \|^2 + \| Y \|^2 - \text{trace}((X^T Y Y^T X)^{1/2}) \) for orthogonal matrix \( R \) [15],

\[ \inf_{\beta_0 \neq 0} \| X - \beta_0 Y \|^2 = \| X \|^2 + \| Y \|^2 - \text{trace}((X^T Y Y^T X)^{1/2}) \]

Therefore, the infimum can be achieved when \( \beta_0 \) is selected such that

\[ (X^T Y Y^T X)^{1/2} = \| Y \|^2. \]

Furthermore, we have

\[ \inf_{\beta_0 \neq 0} \| X - \beta_0 Y \|^2 = \| X \|^2 - \text{trace}((X^T Y Y^T X)^{1/2}) \]

for any \( Y \). It is because

\[ \inf_{\beta_0 \neq 0} \| X - \beta_0 Y \|^2 = \inf_{\beta_0 \neq 0} \| X \|^2 - \text{trace}((X^T Y Y^T X)^{1/2}) \]

for orthogonal matrix \( R \) and non-zero \( \beta_0 \). Therefore we may define the second term on the right hand side as \( 0 \) if \( W = 0 \), 
so that \( \inf_{\beta_0 \neq 0} \| X - \beta_0 R \|^2 = \inf_{\beta_0 \neq 0} \| X \|^2 = \| X \|^2. \) It follows

that max \( \inf_{\beta_0 \neq 0} \| HX - \beta_0 Y \|^2 \leq \| X \|^2. \) On the other

hand, we have: \( \max \inf_{\beta_0 \neq 0} \| HX - \beta_0 Y \|^2 \geq \inf_{\beta_0 \neq 0} \| X - \beta_0 Y \|^2 = \| X \|^2. \) QED.
Due to this Proposition, the GF coefficient can be reduced to
\[
\text{GF}(X, Y) = 1 - \inf_{\beta} \min_{R} \frac{\|HX - \beta HYR\|^2}{\|HX\|^2}.
\]

Furthermore, since \( \inf_{\beta} \min_{R} \|X - \beta YR\|^2 = \|X\|^2 - \text{trace}(X^T Y Y^T X)^{1/2}/\|Y\|^2 \) for orthogonal matrix \( R \) and non-zero \( \beta \), the GF coefficient can be rewritten as
\[
\text{GF}(X, Y) = \left[ \text{trace}\left( \left( X^T H Y Y^T X \right)^{1/2} \right) \right]^2/\|HX\|^2 \|HY\|^2,
\]
where the function \( \text{trace} \) 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 \( \beta \) and \( R \), and thus TRD-invariant. For property (M-3), we have the following Proposition.

**Proposition 5**

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

**Proof**

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

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

**Proposition 6**

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

**Proof**

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

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

**Proposition 7**

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

**Proof**

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

**7. TRD-INVariant MATRIX CORRelaTION COEFFICIENTS**

Closeness of homologous configurations \( C_\delta \) and \( C_\delta \) can also be quantified by means of a correlation between two coordinate matrices \( X \in \bar{M}_\delta \) and \( Y \in \bar{M}_\delta \) as follows [3]:
\[
\text{corr}(X, Y) = \frac{(HX, HY)/(\|HX\| \|HY\|)}{\|HX\| \|HY\|} = \frac{\text{trace}(X^T HY)}{\|HX\| \|HY\|}.
\]

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

We, therefore, examined whether the maximum correlation
\[
\text{CORR}(X, Y) = \max_{\beta, R} \text{corr}(X, \beta Y R)
\]

between any pair of members (one from each set of \( \bar{M}_\delta \) and \( \bar{M}_\delta \) ) satisfies the five required properties including the TRD-invariance. First, it is obvious that \( \text{CORR}(X, Y) \) ranges between 0 and 1, satisfying (M-1). Second, due to the following Proposition 8, \( \text{CORR}(X, Y) \) 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,
\[
\text{R}_{GLS}(X, Y) = \frac{\text{trace}(X^T H Y Y^T X)^{1/2}}{\|HX\| \|HY\|}.
\]

**Proposition 8**

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

**Proof**

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

It follows that
\[
\text{CORR}(X, Y) = \max_{\beta, R} \text{corr}(X, \beta Y R) = \max_{\beta, R} \left\{ \frac{\langle HX, \beta HYR \rangle}{\|HX\| \|\beta HYR\|} \right\}.
\]
Closeness Coefficients

\[
\begin{align*}
= \max_{R} \left[ \frac{\langle HX, HYR \rangle}{\|HX\| \|HY\|} \right] = \\
\text{trace} \left( \left( X^T HY^T HX \right)^{1/2} / \|HX\| \|HY\| \right) = R_{GLS}(X,Y).
\end{align*}
\]

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

\[
R_{GLS}(X,Y) = \sqrt{\text{GF}(X,Y)} .
\]

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 \( X \in M_\delta \) and \( Y \in M_\delta \) through their centered forms \( HX^T H \) and \( HY^T \). This relation is due to the fact that \( \text{trace} \left( \left( X^T HY^T HX \right)^{1/2} \right) = \text{trace} \left( \left( Y^T HX^T HY \right)^{1/2} \right) \) based on Proposition 7, and both \( \|HX\| \) and \( \|HY\| \) are a function of the centered forms. Moreover, the \( R_{GLS} \) coefficient can be expressed as a function of singular values of \( HX \) and \( HY \) for \( X \in M_\delta \) and \( Y \in M_\delta \), respectively. To this end, let us denote the singular value decompositions of \( HX \), \( HY \), and \( X^T HY \) by

\[
HX = U_s A_s V_s^T , \quad HY = U_s A_s V_s^T , \quad \text{and } X^T HY = UAV^T .
\]

Although \( \text{corr}(U_s A_s, U_s A_s) \) does not satisfy the invariance property, again its maximum over orthonormal matrix \( R \) does since \[ \max_{R} \text{corr}(U_s A_s, U_s A_s R) = \max_{R} \text{corr}(X Y R) = R_{GLS}(X,Y) . \]

It follows that

\[
R_{GLS}(X,Y) = \frac{\text{trace}(\Lambda^2)}{\sqrt{\text{trace}(\Lambda^2) \text{trace}(\Lambda^2)}} .
\]

8. COEFFICIENTS BASES ON CENTERED FORMS

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

\[
\text{corr}(HXX^T H, HY^T Y) = \text{trace}(HXX^T H Y Y^T Y) / \left( \|HXX^T H\| \|HY^T Y\| \right) = \text{RV}(X,Y).
\]

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 \( HX \) and \( HY \). 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 \( \beta \) 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:

\[
\text{RV}(X,Y) = 1 - \inf_{\beta} \frac{\|HXX^T H - \beta HY^T Y\|}{\|HXX^T H\|} .
\]

Furthermore, the RV coefficient can be re-expressed as

\[
\text{RV}(X,Y) = \text{corr}(U_s A_s^T U_s, U_s A_s^T U_s), \quad \text{where } HXX^T H
\]

s \( = U_s A_s^2 U_s^T \) and \( HY^T Y \) s \( = U_s A_s^2 U_s^T \) as in (6). It follows that \( \text{RV}(X,Y) \) can be reduced to

\[
\text{RV}(X,Y) = \frac{\text{trace}(\Lambda^2)}{\sqrt{\text{trace}(\Lambda^2) \text{trace}(\Lambda^2)}} ,
\]

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

\[
\text{RV}(X,Y) = \text{corr}(H \Delta^2 \Delta^2, H \Delta^2 \Delta^2).
\]

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 [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 \( \delta_1 \) and \( \delta_2 \). 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],

\[
\cos(\delta_1, \delta_2) = \langle \delta_1, \delta_2 \rangle / \left( \|\delta_1\| \|\delta_2\| \right) ,
\]

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, \( \cos(\delta_1, \delta_2) \) is TRD-invariant. Nevertheless, \( \cos(\delta_1, \delta_2) \) 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, $R_{GFS}$, 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