Glossary

Affine transformation (Also called "uniform"). Transformation (or mapping) that leaves parallel lines parallel. The possible affine transformations include those that do not alter shape (scaling, translation, rotation) and those that do (shear and contraction/dilation). See also **Explicit uniform terms, Implicit uniform terms** (Chapter 6).

Allometry Shape change correlated with size change, sometimes more narrowly defined as a change in the size of a part according to the power law $Y = bX^k$, where Y is the size of the part, X is either the size of another part or overall body size, and k and b are constants. There are three distinct types of allometry: (1) ontogenetic, an ontogenetic change in shape correlated with an ontogenetic increase in size; (2) static, variation in shape correlated with variation size among individuals at a common developmental stage; and (3) evolutionary, an evolutionary change in shape correlated with evolutionary changes in size (Chapters 10, 13).

***Alpha* (α)** (1) The acceptable Type I error rate, typically 5%; (2) a factor multiplying partial warps before computing principal components of them; if $\alpha = 0$, principal components of partial warps are conventional principal components; when $\alpha \neq 0$, the partial warps are differentially weighted. Either those with lower bending energy are weighted more highly ($\alpha > 0$) or those with greater bending energy are weighted more highly ($\alpha < 0$). Typically, values of $+1$ or -1 are used. See also **Relative warps**.

ANCOVA Analysis of covariance. A method for testing the hypothesis that samples do not differ in their means when the effects of a covariate are taken into account. See also **ANOVA, MANOVA** and **MANCOVA** (Chapters 9, 10).

Anisotropic Not isotropic, having a preferred direction. In general, anisotropy is a measure of the degree to which variation in some parameter is a function of its direction relative to some axis. In geometric morphometrics, anisotropy usually refers to a measure of an affine transformation – either the ratio between principal strains, or a ratio of variances along principal axes. See also **Isotropic** (Chapter 3).

ANOVA Analysis of variance. A method for testing the hypothesis that samples do not differ in their means. ANOVA differs from MANOVA in that the means are unidimensional scalars. See also **ANCOVA, MANOVA** and **MANCOVA** (Chapter 9).

Baseline A line joining two landmarks, used in some superimposition methods to register shapes by assigning fixed values to one or more coordinates of those landmarks. See also **Baseline registration, Bookstein coordinates, Sliding baseline registration** (Chapters 3, 5).

Geometric Morphometrics for Biologists
ISBN 0–12–77846–08

Baseline registration A method of superimposing landmark configurations by assigning two landmarks fixed values (the two landmarks are the endpoints of the baseline). The most common method of baseline registration is the two-point registration developed by Bookstein, in which the ends of the baseline are fixed at (0, 0) and (1, 0), yielding Bookstein coordinates. Other methods of baseline registration fix the endpoints at different values (see Dryden and Mardia, 1998) or only fix one coordinate of each baseline point (see **Sliding baseline registration**) (Chapters 3, 5).

Basis A set of linearly independent vectors that span the entire vector space, also the smallest necessary set of vectors that span the space. The basis can serve as a coordinate system for the space because every vector in that space is a unique linear combination of the basis vectors. However, the basis itself is not unique; any vector space has infinitely many bases that differ by a rotation. An orthonormal basis is a set of mutually orthogonal axes, all of unit length. Partial warps and principal components are two common orthonormal bases used in shape analysis. See also **Eigenvectors** (Chapters 6, 7).

Bending energy (1) A measure of the amount of non-uniform shape difference based on the thin-plate spline metaphor. In this metaphor, bending energy is the amount of energy required to bend an ideal, infinite and infinitely thin steel plate by a given amplitude between chosen points. Applying this concept to the deformation of a two-dimensional configuration of landmarks involves modeling the displacements of landmarks in the X, Y plane as if they were displacements above or below the plane ($\pm Z$). (2) Eigenvalues of the bending-energy matrix, representing the amount of bending energy per unit deformation along a single principal warp (eigenvector of the bending-energy matrix). This concept of bending energy is useful because it provides a measure of spatial scale; it takes more energy to bend the plate by a given amount between closely spaced landmarks than between more distantly spaced landmarks. Thus, principal warps with large eigenvalues represent more localized components of deformation than principal warps with smaller eigenvalues. The total bending energy (definition 1) of an observed deformation is a sum of multiples of the eigenvalues, and accounts for the non-uniform deformation of the reference shape into the target shape. See also **Thin-plate spline, Principal warps, Partial warps** (Chapter 6).

Bending-energy matrix The matrix used to compute principal warps and their bending energies (eigenvectors and eigenvalues, respectively). This matrix is a function of the distances between landmarks in the reference shape. See also **Principal warps**, **Partial warps** (Chapter 6).

Biorthogonal directions Principal axes of a deformation; the term was used in Bookstein et al., 1985; more recently, workers refer to principal axes (Chapter 3).

Black Book Marcus, L. F., Bello, E. and Garcia-Valdcasas, A. (eds) (1993). *Contributions to Morphometrics*. Madrid, Monografias del Museo Nacional de Ciencias Naturales 8. (See also **Blue Book**, **Orange Book**, **Red Book** and **White Book**.)

Blue Book: Rohlf, F. J. and Bookstein, F. L. (eds) (1990). *Proceedings of the Michigan Morphometrics Workshop*. University of Michigan Museum of Zoology, Special Publication No. 2 (See also **Black Book**, **Orange Book**, **Red Book** and **White Book**.)

Bonferroni correction, Bonferroni adjustment An adjustment of the α-value to protect against inflating Type I error rate when testing multiple *a posteriori* hypotheses. The adjustment is done by dividing the acceptable Type I error rate (α) by the number of tests. That quotient is the adjusted α-value for each of the *a posteriori* hypotheses. For example, if the desired Type I error rate is 5%, and there are 10 *a posteriori* hypotheses to test, $0.05/10 = 0.005$ is the α-value for each of those 10 tests. A less conservative approach uses a sequential Bonferroni adjustment in which the desired α-value is divided by the number of remaining tests. Thus, the adjusted α for the first test would be 0.05/10; for the second it would be 0.05/9; for the third it would be 0.05/8, etc. To apply this sequential adjustment, hypotheses are ordered from lowest to highest p-value; the null hypothesis is rejected for each in turn until reaching one that cannot be rejected (the analysis stops at that point).

Bookstein coordinates (**BC**) The shape variables produced by the two-point registration, in which the configuration is translated to fix one end of the baseline at $(0, 0)$, and then rescaled and rigidly rotated to fix the other end of the baseline at $(1, 0)$. See also **Baseline registration** (Chapter 3).

Bookstein two-point registration (**BTR**) See **Two-point registration, Bookstein coordinates**.

Bootstrap test A statistical test based on random resampling (with replacement) of the data. Usually, the method is used to simulate the null model that one wishes to test. For example, if using a bootstrap test of the difference between means, the null hypothesis of no difference is simulated. Bootstrap tests are used when the data are expected to violate distributional assumptions of conventional analytic statistical tests. Rather than assuming that the data meet the distributional assumptions, bootstrapping produces an empirical distribution that can be used either for hypothesis testing or for generating confidence intervals. See also **Jackknife test, Permutation test** (Chapter 8).

Canonical variates analysis (**CVA**) A method for finding the axes along which groups are best discriminated. These axes (canonical variates) maximize the between-group variance relative to the within-group variance. Scores for individuals along these axes can be used to assign specimens (including unknowns) to the groups, and can be plotted to depict the distribution of specimens along the axes. CVA is an ordination rather than statistical method. See also **Ordination methods, Principal components analysis** (Chapter 7).

Cartesian coordinates Coordinates that specify the location of a point as displacements along fixed, mutually perpendicular axes. The axes intersect at the origin, or zero point, of all axes. Two Cartesian coordinates are needed to specify positions in a plane (flat surface); three are required to specify positions in a three-dimensional space. These coordinates are called “Cartesian” after the philosopher Descartes, a pioneer in the field of analytic geometry.

Centered A matrix is centered when its centroid is at the origin of a Cartesian coordinate system; i.e. at $(0, 0)$ of a two-dimensional system or at $(0, 0, 0)$ of a three-dimensional system (Chapter 6).

Centroid See **Centroid position**.

Centroid position The position of the averaged coordinates of a configuration of landmarks. The centroid position has the same number of coordinates as the landmarks. The X-component of the centroid position is the average of the X-coordinates of all landmarks of an individual configuration. Similarly, the Y-component is the average of the Y-coordinates of all landmarks of an individual configuration. It is common to place the centroid position at $(0, 0)$, because this often simplifies other computations (Chapter 6).

Centroid size (CS) A measure of geometric scale, calculated as the square root of the summed squared distances of each landmark from the centroid of the landmark configuration. This is the size measure used in geometric morphometrics. It is favored because centroid size is uncorrelated with shape in the absence of allometry, and also because centroid size is used in the definition of the Procrustes distance (Chapters 3, 4, 5).

Coefficient A number multiplying a function. For example, in the equation $Y = mX$, m is the coefficient for the slope, which is the function that relates X and Y.

Column vector A vector whose entries are arranged in a column. Contrast to a **Row vector**.

Complex numbers A number consisting of both a real and an imaginary part. An imaginary number is a real number multiplied by i, where i is $\sqrt{-1}$. A complex number is written as $Z = X + iY$, where X and Y are real numbers. In that notation, X is said to be the real part of Z and Y is the imaginary part. A complex number is often used to represent a vector in two dimensions. The mathematics of two-dimensional vectors and complex numbers are similar, so it is sometimes useful to perform calculations or derivations in complex number form.

Configuration see **Landmark configuration**.

Configuration matrix A matrix representing the configuration of K landmarks, each of which has M dimensions. A configuration matrix is a $K \times M$ matrix in which each row represents a landmark and each column represents one Cartesian coordinate of that landmark; $M = 2$ for landmarks of two-dimensional configurations (planar shapes), and $M = 3$ for landmarks of three-dimensional configurations. Two configuration matrices can differ in location, size and orientation, as well as shape (Chapter 4).

Configuration space The set of all possible configuration matrices describing all possible configurations of K landmarks with M coordinates (all with the same values of K and M). Because there are $K \times M$ elements in the configuration matrices, there are $K \times M$ dimensions in the configuration space. In statistical analyses, the configuration space accounts for $K \times M$ degrees of freedom because that is the number of independent pieces of information (e.g. landmark coordinates) needed to specify a particular configuration (Chapter 4).

Consensus configuration The mean (average) configuration of landmarks in a sample of configurations. Usually, this is calculated after superimposing coordinates. See also **Generalized Procrustes superimposition, Reference form** (Chapters 4, 5).

Contraction A mathematical mapping that "shrinks" a configuration along one axis. A contraction along the X-axis would map the point (X, Y) to the point (AX, Y), where A is less than one. A contraction along the Y-axis would map (X, Y) to (X, AY). **Expansion** or **dilation** is the opposite of contraction ($A > 1$).

Coordinates The set of values that specify the location of a point along a set of axes (see **Cartesian coordinates**).

Correlation A measure of the association between two or more variables. In morphometrics, correlation is most often measured using Pearson's product-moment correlation, which is the covariance divided by the product of the variances:

$$\mathrm{R}_{XY} = \frac{\sum (X - X_{\mathrm{mean}})(Y - Y_{\mathrm{mean}})}{\sqrt{\sum (X - X_{\mathrm{mean}})^2 \sum (Y - Y_{\mathrm{mean}})^2}}$$

where the sums are taken over all specimens. When variables are highly correlated we can predict one from the other (e.g. Y from X), and the more highly correlated they are, the better our predictions will be. Uncorrelated variables are considered independent. See also **Covariance.**

Covariance Like correlation, a measure of the association between variables. The sample estimate of the covariance between X and Y is:

$$\mathrm{S}_{XY} = \left(\frac{1}{N-1}\right) \sum (X - X_{\mathrm{mean}})(Y - Y_{\mathrm{mean}})$$

where the summation is over all N specimens.

Curved space A metric space in which the distance measure is not linear. The ordinary rules of Euclidean geometry do not apply in such spaces. The consequences of the curvature depend upon the distance between points; we can treat the surface of the earth as flat as long as the maps cover only small areas, but in long-distance navigation the curvature must be taken into account. Shape space is curved, so the rules of Euclidean geometry do not apply, which is why shapes are mapped onto a Euclidean space tangent to shape space.

D A generalized statistical distance between means of two groups (**X1** and **X2**) relative to the variance within the groups:

$$\mathrm{D} = \sqrt{(\mathbf{X1} - \mathbf{X2})^{\mathrm{T}} \mathbf{S}_{\mathrm{p}}^{-1} (\mathbf{X1} - \mathbf{X2})}$$

where $(\ )^{\mathrm{T}}$ refers to the transpose of the enclosed matrix, and $\mathbf{S}_{\mathrm{p}}^{-1}$ is the inverse of the pooled variance–covariance matrix. This distance takes into account the correlations among variables when computing the distance between means. The generalized distance is used in Hotelling's T^2-test. Also known as the **Mahalanobis' distance**.

D^2 The squared generalized distance, **D**. See **D**.

Deformation A smooth, continuous mapping or transformation; in morphometrics, it is usually the transformation of one shape into another. The deformation refers not only to the change in positions of landmarks, but also to the interpolated changes in locations of unanalyzed points between landmarks (Chapter 6).

Degrees of freedom In general, the number of independent pieces of information. In statistical analyses, the total degrees of freedom are approximately the product of the number of variables and the number of individuals (the total may be partitioned into separate components for some tests). If every measurement on every individual were completely independent, the degrees of freedom would be the product of the number of variables and the number of individuals, but if one statistic is known (or estimated), the number of degrees of freedom that remain to estimate a second statistic will be reduced. For example, the estimate of the mean height of N individuals in a sample will have $N \times 1 = N$ degrees of freedom, because all N measurements are needed and there is only one measured variable. In contrast, the estimate of the variance in height will have $N - 1$ degrees of freedom because only $N - 1$ deviations from mean height are independent (the deviation of the Nth individual can be calculated from the mean and the other $N - 1$ observed heights). In geometric morphometrics, when configurations of landmarks are superimposed, degrees of freedom are lost for a different reason; namely, information that is not relevant to comparison of shapes (location, scale and rotation) is removed from the coordinates.

Dilation Opposite of **Contraction**.

Discriminant function The linear combination of variables optimally discriminating between two groups. It is produced by discriminant function analysis. Scores on the discriminant function can be used to identify members of the groups (Chapter 7).

Discriminant function analysis A two-group **canonical variates analysis**. See **Canonical variates analysis** (Chapter 7).

Disparity, morphological disparity (*MD*) Phenotypic variety, usually morphological. Several metrics can be used to measure disparity, but the one most commonly used in studies of continuous variables is:

$$MD = \frac{\sum_{j=1}^{N} D_j^2}{(N-1)}$$

where D_j is the distance of species j from the overall centroid (i.e. the grand mean calculated over N groups, e.g. species) (Chapter 12).

Distance A function measuring the separation between points. Within any space there are multiple possible distances. For this reason, it is necessary to specify the type of distance used. See also **D, D^2, Euclidean distance, Generalized distance, Geodesic distance, Great circle distance, Partial Procrustes distance, Full Procrustes distance, Mahalanobis' distance** (Chapter 4).

Dot product (Also called inner product.) Given two vectors $\mathbf{A} = \{A_1, A_2, A_3 \ldots A_N\}$, $\mathbf{B} = \{B_1, B_2, B_3 \ldots B_N\}$, the dot product of **A** and **B** is:

$$\mathbf{A} \cdot \mathbf{B} = A_1B_1 + A_2B_2 + A_3B_3 + \ldots . + A_NB_N$$

and

$$\mathbf{A} \cdot \mathbf{B} = |A||B| \cos(\theta)$$

where $|A|$ is the magnitude of **A**, $|B|$ is the magnitude of **B**, and θ is the angle between **A** and **B**. If the magnitude of **A** is 1, then $\mathbf{A} \cdot \mathbf{B} = |B| \cos(\theta)$, which is the component of **B** along the direction specified by **A**. The dot product is used to calculate scores on coordinate axes, by projecting the data onto those axes (this is how partial warp scores and scores on principal components are calculated). It is also used to find the vector correlation, R_V, between two vectors (that correlation is the cosine of the angle between vectors).

Edge registration See **Baseline registration.**

Eigenvalues See **eigenvectors.**

Eigenvectors Eigenvectors are the non-zero vectors, **A**, satisfying the eigenvector equation:

$$(\mathbf{X} - \lambda\mathbf{I})\mathbf{A} = 0$$

The values of λ that satisfy this equation are eigenvalue*s* of **X**. Eigenvectors are orthogonal to one another, and provide the smallest necessary set of axes for a vector space (i.e. they provide a basis for that space). The eigenvectors of a variance–covariance matrix are called principal components; the eigenvalue corresponding to each axis gives the variance associated with it. The eigenvectors of the bending-energy matrix are the principal warps; the eigenvalue corresponding to each axis gives the bending energy associated with it. See also **Basis** (Chapters 4, 6, 7).

Element of a matrix A number in a matrix, typically referenced by the symbol designating the matrix with subscripts indicating its row and column; for example, $X_{4,5}$ refers to the element on the fourth row and fifth column of the matrix **X**.

Euclidean distance The square root of the summed squared distances along all orthogonal axes. A Euclidean distance does not change when the axes of the space are rotated (in contrast to a Manhattan distance, which is simply the sum of the distances). See also **D, D^2, Distance, Generalized distance, Geodesic distance, Great circle distance, Procrustes distance, Full Procrustes distance, Partial Procrustes distance** (Chapter 4).

Euclidean space A coordinate space in which the metric is a Euclidean distance.

Explicit uniform term, explicit uniform component A uniform component describes affine or uniform deformations. Some of these do not alter shape (i.e. rotation, translation and rescaling) whereas others do (i.e. shear and dilation). Accordingly, we divide affine deformations into two sets: (1) implicit uniform terms, which do not alter shape and are used in superimposing forms but are not explicitly recorded; and (2) explicit uniform terms, which do alter shape and therefore are typically reported as components of the deformation. All uniform terms must be known to model a deformation correctly (Chapter 4).

Fiber In geometric morphometrics, the set of all points in pre-shape space representing all possible rigid rotations of a landmark configuration that has been centered and scaled to unit centroid size; in other words, the set of pre-shapes that have the same shape. Fibers are collapsed to a point in shape space (Chapter 4).

Form Size-plus-shape of an object; form includes all the geometric information not removed by rotation and translation. Form is also called **Size-and-shape**.

Full Procrustes distance (D_F) The distance between two landmark configurations in the linear space tangent to Kendall's shape space (i.e. the tangent space) when centroid size of one is allowed to vary to minimize the distance between the shapes rather than fixed to unit size. See also **Partial Procrustes distance** (Chapters 4, 5).

Full Procrustes superimposition A superimposition minimizing the full Procrustes distance. See also **Partial Procrustes distance** (Chapters 4, 5).

Generalized distance See **D**.

Generalized least squares superimposition A generalized superimposition method that uses a least squares fitting criterion, meaning that the parameters are estimated to minimize the sum of squared distances over all landmarks over all specimens. Usually, in geometric morphometrics, GLS refers specifically to a generalized least squares Procrustes superimposition – a different approach is used in generalized resistant-fit methods (Chapter 5).

Generalized least squares Procrustes superimposition (GLS) A generalized superimposition minimizing the partial Procrustes distance over all shapes in the sample, using a least squares fitting function. This is the method usually used in geometric morphometrics (Chapters 4, 5).

Generalized superimposition The superimposition of a set of specimens onto their mean. This involves an iterative approach because the mean cannot be calculated without superimposing specimens, which cannot be superimposed on the mean before the mean is calculated (an alternative approach is used in ordinary Procrustes analysis). See also **Consensus configuration** (Chapters 4, 5).

Geodesic distance The shortest distance between points in a space. On a flat planar surface, this is the length of the straight line joining the points – i.e. the Euclidean distance. On curved surfaces, this distance is the length of an arc.

Great circle The intersection of the surface of a sphere and a plane passing through its center. A great circle divides the surface of the sphere in half. On the surface of the sphere, the shortest distance between two points lies along the great circle that passes through those points. If the Earth were perfectly spherical, the equator and all lines of latitude would be great circles.

Great circle distance The arc length of the segment of the great circle connecting two points on the surface of a sphere; this is the geodesic distance between those points, the shortest distance between the points in the space of the surface of the sphere.

Homology (1) Similarity due to common evolutionary origin. In morphometrics, landmarks are considered homologous by virtue of the homology of the structures defining their locations. (2) Some morphometricians use the term for the correspondences between points that are imputed by a mathematical function, called a "homology function" (e.g. see Bookstein et al., 1985). Homology is the primary criterion for selecting landmarks (Chapter 2).

Hypersphere The generalization of a three-dimensional sphere to more than three dimensions. In three dimensions, points on the surface of a sphere of radius R that is centered at the origin satisfy the equation $X^2 + Y^2 + Z^2 = R^2$.

Implicit uniform terms See **Explicit uniform terms**.

Induced correlation A correlation induced by dividing two values by a third which is common to both. The induced correlation between the (rescaled) variable is not present in the original variables.

Inner product See **Dot product**.

Invariant A quantity is invariant under a mathematical operation or transformation when it is not changed by that operation. For example, centroid size is invariant under translation, centroid position is not.

Isometric In general, a transformation that leaves distances between points unaltered. In morphometrics, isometry usually means that shape is uncorrelated with size. In statistical tests of allometry, isometry is the null hypothesis (Chapters 10, 13).

Isotropic A property is said to be isotropic if it is uniform in all directions, i.e. if it does not differ as a function of direction. When an error is isotropic, it is equal in all directions, and there is no correlation among errors. Isotropic is the opposite of anisotropic.

Jackknife test An approach to statistical testing that involves resampling the original observations to generate an empirical distribution. Jackknifing is carried out by omitting one specimen at a time. See also **Bootstrap test, Permutation test** (Chapter 8).

Kendall's shape space The space in which the distance between landmark configurations is the Procrustes distance. This space is constructed by using operations that do not alter shape to minimize differences between all configurations of landmarks that have the same values of K (number of landmarks) and M (number of coordinates of a landmark). Kendall's shape space is the curved surface of a hypersphere, so conventional statistical analyses are conducted in a Euclidean tangent space (Chapter 4).

Landmark Biologically, landmarks are discrete, homologous anatomical loci; mathematically, landmarks are points of correspondence, matching within and between populations (Chapter 2).

Landmark configuration The positions (coordinates) of a set of landmarks representing a single object, containing information about size, shape, location and orientation. The number of landmarks is typically represented by K, and the dimensionality of the landmarks (number of coordinates) is typically represented by M. Therefore, if there are 16 landmarks, each with an X- and Y-coordinate, then $K = 16$ and $M = 2$ (Chapter 4).

Least squares A method of choosing parameters that minimizes the summed square differences over all individuals (and variables) (Chapter 10).

Linear A function $f(X)$ is linear if it depends only on the first power of X; e.g. $f(X) = 2(X)$ is linear, but $f(X) = 2(X)^2$ is not.

Linear combination A vector produced by multiplying and summing coefficients of one or more vectors. For example, given the vector $\mathbf{X}^{\mathbf{T}} = \{X_1, X_2 \ldots X_N\}$ and $\mathbf{A}^{\mathbf{T}} = \{A_1, A_2 \ldots A_N\}$, then $\mathbf{Y} = A_1X_1 + A_2X_2 + \cdots A_NX_N$ is a linear combination of the vectors. We can write this as $\mathbf{Y} = \mathbf{A}^{\mathbf{T}}\mathbf{X}$.

Linear transformation A transformation producing a set of new vectors that are linear combinations of the original variables. See **Linear combination**.

Linear vector space The set of all linear combinations of a set of vectors. The space spans all possible linear combinations of the basis vectors, as well as all sums or differences of any linear combination of those basis vectors. The two-dimensional Cartesian plane is the linear vector space formed by the linear combinations of two vectors of unit length, one along the X-axis, the other along the Y-axis.

Mahalanobis' distance (D) The squared distance between two means divided by the pooled sample variance–covariance matrices. This is a generalized statistical distance, adjusting for correlations among variables. See also **D**, **Generalized distance**.

MANCOVA Multivariate analysis of covariance. A method for testing the hypothesis that samples do not differ in their means when the effects of a covariate are taken into account. See also **ANOVA**, **ANCOVA** and **MANOVA** (Chapters 9, 10).

MANOVA Multivariate analysis of variance. A method for testing the hypothesis that samples do not differ in their means; MANOVA differs from ANOVA in that the means are multidimensional vectors. See also **ANOVA**, **ANCOVA** and **MANCOVA** (Chapter 9).

Map A mathematical function relating **X** to **Y** by stating the correspondence between elements in **X** and **Y**. Each element in **X** is placed in correspondence with one element in **Y**. Multiple elements in **X** may map to the same element in **Y** (landmark configurations differing only in rotation for example would all map to the same shape). A map is written as: $f : \mathbf{X} \rightarrow \mathbf{Y}$ where f is the map from the set **X** to the set **Y**.

Matrix A rectangular array of numbers (real or complex). The numbers in a matrix are referred to as elements of the matrix. The size of a matrix is always given as the number of rows followed by the number of columns; e.g. a 4×2 matrix has four rows and two columns.

Mean Also known as the average; an estimate of the center of the distribution calculated by summing all observations and dividing by the sample size.

Median An estimate of the center of a distribution calculated such that half the observed values are above and the other half are below.

Metric A non-negative real-valued function, $D(X, Y)$, of the points X and Y in a space such that:

1. The only time that the function is zero is when X and Y are the same point, i.e. $D(X, Y) = 0$, if and only if $X = Y$
2. If we measure from X to Y, we get the same distance as when we measure from Y to X, so $D(X, Y) = D(Y, X)$ for all X and Y
3. The triangle inequality holds true. The triangle inequality states the distance between any two points, X and Y, is less than or equal to the sum of distances from each to a third point, Z, so $D(X, Y) \leq D(X, Z) + D(Y, Z)$, for all X, Y and Z.

Multiple regression Regression of a single (univariate) dependent variable on more than one independent variable. See also **Multivariate regression**, **Regression**.

Multivariate analysis of variance See **MANOVA**.

Multivariate multiple regression Regression of several dependent variables on more than one independent variable. In morphometrics, this method is used to regress shape (the dependent variables) onto multiple independent variables. See also **Multiple regression**, **Multivariate regression**, **Regression**.

Multivariate regression Regression of several dependent variables onto one independent variable. In morphometrics, this method is used to regress shape onto a single independent variable, such as size. The coefficients obtained by multivariate regression are the same as those estimated by simple bivariate regression of each dependent variable on the independent variable. However, the statistical test of the null hypothesis differs. See also **Multiple regression**, **Multivariate multiple regression**, **Regression** (Chapters 10, 13).

Non-uniform Not **Uniform; Non-affine**. See **Non-uniform deformation**.

Non-uniform deformation The component of a deformation that is not uniform. In contrast to a uniform deformation, which leaves parallel lines parallel and has the same effect everywhere across a form, a non-uniform deformation turns squares into trapezoids or diamonds (shapes that do not have parallel sides) and has different effects over different regions of the form. Most deformations comprise both uniform and non-uniform parts. The non-uniform component can be further subdivided, see **Partial warps** (Chapter 6).

Normalize To set the magnitude to one. Normalizing a vector sets the length of the vector to one; this is done by dividing each component of the vector by the length of the vector, calculated by taking the square root of the summed squared coefficients.

Null hypothesis, or **null model** Usually, the hypothesis that the factor of interest has no effect beyond that expected by chance. For example, in an analysis of allometry, the null hypothesis being tested by regression of shape on size is that shape does not depend on size (i.e. isometry). Similarly, in a comparison of two means using Hotelling's T^2-test, the null hypothesis is that the two groups do not differ beyond what is expected by chance.

Orange Book Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data. Geometry and Biology*. Cambridge University Press. (See also **Black Book, Blue Book, Red Book** and **White Book**.)

Ordinary Procrustes analysis (OPA) An approach to superimposition in which one landmark configuration is fitted to another, differing from a **Generalized superimposition** in that it involves only two forms. This approach has rarely been used since iterative methods became available for generalized superimpositions. See also **Generalized superimposition, Consensus form** (Chapter 5).

Ordination Ordering specimens along one or more axes based on some criterion (e.g. from youngest to oldest, or shortest to tallest). Ordination methods include principal components analysis and canonical variates analysis; the scores on the axes provide a basis for ordering specimens (Chapter 7).

Orthogonal Perpendicular (at right angles to each other). Two vectors are orthogonal if the angle between them is 90°; when they are, their dot product is zero.

Orthonormal Perpendicular and of unit length; vectors are orthonormal if they are mutually orthogonal and of unit length.

Orthonormal basis See **Basis**.

Population The set of all possible individuals of a specific type, such as all members of a species, or all leaves on a particular kind of tree. See also **Sample** (Chapter 8).

Outline A curve around the perimeter of an object (or around a distinct part of it).

Partial least squares analysis A method of exploring patterns of covariance or correlation between two blocks of variables measured on the same set of specimens. A singular value decomposition is used to determine the pair of vectors (each a linear combination of variables within one of the blocks) that expresses the greatest proportion of the covariance between blocks. See also **Singular value decomposition**, **Singular warps** (Chapter 11).

Partial Procrustes distance (D_p) The distance between two landmark configurations in the linear tangent space to Kendall's shape space when both shapes are centered, fixed to unit centroid size, and rotated to minimize the sum of squared distances between their corresponding landmarks. See also **Full Procrustes distance**, **Procrustes distance** (Chapters 4, 5).

Partial Procrustes superimposition A superimposition that minimizes the partial Procrustes distance between shapes. See also **Full Procrustes distance**, **Procrustes distance** (Chapters 4, 5).

Partial warps The term partial warps sometimes refers solely to the components of the non-uniform deformation, which are computed as eigenvectors of the bending-energy matrix projected onto the X, Y-plane of the data (they are projections of principal warps), ordered from least to most bending energy. These eigenvectors provide an orthonormal basis for the non-uniform part of a deformation. Sometimes "partial warps" also includes the components of the uniform deformation, as the zeroth partial warp – in which case the scores on this component are included among the partial warp scores (Chapter 6).

Partial warp scores Coefficients indicating the position of an individual, relative to the reference, along partial warps. They are calculated by taking the dot product between the partial warps and the data for a specimen. When appropriate scores on the uniform component are also included among the partial warps scores, the sum of the squared scores equals the squared partial Procrustes distance of that specimen from the reference. This full set of scores can be used as shape variables in any conventional statistical analysis because they are based on the appropriate distance measure and have the same number of coordinates as degrees of freedom. See also **Non-uniform deformation**, **Partial warps, Principal warps, Uniform deformation** (Chapter 6).

Permutation test An approach to statistical testing that involves permuting (rather than randomly sampling) observed values. See also **Bootstrap test**, **Jackknife test**, **Monte Carlo simulations** (Chapter 8).

Pinocchio effect A large change concentrated at one landmark, with little or none at others; a highly localized change. In the presence of the Pinocchio effect, Procrustes superimpositions imply that the shape difference is distributed over all landmarks. Resistant-fit methods, such as RFTRA, were devised to avoid that implication (Chapter 5).

Position See **Centroid position**.

Pre-shape A centered landmark configuration, scaled to unit centroid size (Chapter 4).

Pre-shape space The set of all possible pre-shapes for a given number of landmarks with a given number of dimensions. This is the surface of a sphere of $KM - M - 1$ dimensions, where K is the number of landmarks and M is the number of dimensions of each landmark (Chapter 4).

Principal axes The set of orthogonal axes used in modeling the change of one shape into another as an affine transformation. This transformation can be parameterized by its effect on a circle or sphere (for two or three dimensional shapes, respectively). In two dimensions, an affine transformation takes a circle into an ellipse and the principal axes are the directions of the circle that undergo the greatest relative elongation or shortening mapped onto the major and minor axes of the ellipse. The ratio of the lengths of these axes is the anisotropy, a measure of the amount of affine shape change. Principal axes are invariant under a change in the coordinate system. See also **Principal strains** (Chapter 3).

Principal components analysis (PCA) A method for reducing the dimensionality of multivariate data, performed by extracting the eigenvectors of the variance–covariance matrix. These eigenvectors are called principal components. Their associated eigenvalues are the variance explained by each axis. Principal components provide an orthonormal basis. The position of a specimen along a principal component is represented as its principal component score, calculated by taking the dot product between that principal component and the data for that specimen (Chapter 7).

Principal strain In an affine deformation, the ratio of the length of a principal axis in the ellipse to the original diameter of the circle. See also **Principal axes** (Chapter 3).

Principal warp An eigenvector of the bending-energy matrix interpreted as a warped surface over the surface of the X, Y-plane of the landmark coordinates. Principal warps are ordered from least to most bending energy (smallest to largest eigenvalue), which corresponds to the least to most spatially localized deformation. Principal warps differ from partial warps in that partial warps are projections of principal warps onto the X, Y-plane of the data. See also **Bending energy, Bending-energy matrix, Orthonormal basis, Partial warp, Thin-plate spline** (Chapter 6).

Probability distribution A mathematical function that describes the probability of a measurement taking on either a particular value or a range of values, depending on whether the variable is discrete or continuous, respectively (Chapter 8).

Procrustes distance The distance between two landmark configurations in Kendall's shape space. It is approximately the square root of the summed squared distances between homologous landmarks when the configurations are in Procrustes superimposition. This distance is measured in the curved shape space (Chapter 4).

Procrustes methods A general term referring to the superimposition of matrices based on a least squares criterion. The term comes from the Greek mythological figure, Procrustes, who fitted visitors to a bed by stretching them or amputating overhanging parts (Chapter 5).

Procrustes residuals Coordinates of a landmark configuration obtained by a Procrustes superimposition. They are residuals in the sense that they indicate the deviation of each specimen from the mean (i.e. the consensus configuration) or other reference. See also **Consensus configuration**, **Procrustes superimposition**, **Reference** (Chapter 5).

Procrustes superimposition A superimposition of shapes that minimizes the Procrustes distances over the sample. The term is used whether the distance being minimized is the full or the partial Procrustes distance (Chapters 4, 5).

Red Book Bookstein, F. L., Chernoff, B., Elder, R. L. et al. (eds) (1985). *Morphometrics in Evolutionary Biology: The Geometry of Size and Shape Change, with Examples from Fishes*. Academy of Natural Sciences of Philadelphia, Special Publication No. 15. (See also **Black Book**, **Blue Book**, **Orange Book** and **White Book**.)

Reference, Reference form The shape to which all others are compared. It is the point of tangency between Kendall's shape space and the tangent space. Because the linear approximation to Kendall's shape space may be inaccurate when the point of tangency is far from the center of the distribution of specimens, the reference is usually chosen to minimize the distances between it and the other specimens – i.e. it is chosen to be the consensus shape (Chapters 4, 5).

Regression An analytic procedure for fitting a predictive model to data and assessing the validity of that model. One variable is expressed as a function of the other, e.g. $Y = mX + b$ expresses Y as a linear function of X. The predictor variable(s) are the independent variable(s), and those variables predicted by the model are the dependent variable(s). In the linear model above, X is the independent variable that predicts the dependent variable, Y. The term "regression" comes from Francis Galton (1889), who concluded that offspring tend towards (regress towards) the mean of the population. As stated by Galton in his law of universal regression, "each peculiarity in a man is shared by his kinsman, but *on the average*, in a less degree." Thus, the offspring of unusually tall fathers regress towards the mean height of the population (Chapters 10, 13).

Relative warps Principal components of partial warp scores, sometimes weighted to emphasize components of low or high bending energy (that weighting is done by setting the parameter α to a value other than 0). Originally, the term referred to an eigenanalysis of the variance–covariance matrix relative to the bending-energy matrix, hence a new term was coined for these components (Bookstein, 1991). Currently, the term usually refers to a conventional principal components analysis of partial warp scores. See also ***Alpha*** (**$\boldsymbol{\alpha}$**), **Bending energy**, **Partial warp scores**, **Principal components analysis** (Chapter 7).

Repeated median The median of medians, used in estimating the scaling factor and rotation angle by resistant-fit superimposition methods such as RFTRA. The repeated median is more robust to large deviations than the median or a least squares estimator. See also **Resistant-fit superimposition**, **RFTRA** (Chapter 5).

Resampling A method whereby a new data set is constructed by randomly selecting from the original data (either values recorded on specimens or residuals from a model).

Construction of a large series of resampled data sets can be used to simulate either the distribution of measured values or the distribution of a test statistic under the null model. Under some conditions, resampling can also be used to produce confidence intervals around the statistic. This approach permits hypothesis tests when the data are expected to deviate from the distributional assumptions of conventional analytic tests. Resampling may be done with replacement, meaning that each observation can appear more than once in a resampled data set; resampling without replacement means each observation appears only once in a set. See also **Bootstrap test**, **Jackknife test**, **Permutation test** (Chapter 8).

Rescale Multiply or divide by a scalar value; used in geometric morphometrics to change the centroid size of a configuration (Chapters 3, 4).

Residual Deviation of an observation from the expected value under a model. For example, a residual from a regression is the deviation between the observed and expected values of the dependent variable at a given value of the independent variable. The term is also used for the coordinates obtained by a Procrustes superimposition, the Procrustes residuals, which are deviations between individual specimens and the reference.

Resistant-fit superimposition A superimposition method that uses medians or repeated medians (rather than a least squares error criterion) to superimpose forms. The method is intended to be resistant to large localized shape differences, such as those produced by the Pinocchio effect. RFTRA is an example of this type of method. See also **Repeated medians, RFTRA** (Chapter 5).

RFTRA (Resistant fit theta-rho analysis) A resistant-fit superimposition method using the method of repeated medians to determine the scaling factor and rotational angle. See also **Resistant fit** and **Repeated median** (Chapter 5).

Rigid rotation A rotation of an entire vector or matrix by a single angle. Rigid rotations do not alter the size, shape or location of the object. Rotations are often represented by square matrices. The rotation matrix:

$$R = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}$$

rotates a $2 \times N$ matrix through an angle θ. When different vectors are multiplied by different angles, the rotation is oblique, not rigid.

Row vector A vector with coefficients in a row. Contrast to a **Column vector**.

Sample The collection of observed individuals representing members of a population. An individual observation is the smallest sampling unit in the study, which might be an individual organism or one of its parts, or a collection of organisms such as a species or a bacterial colony (Chapter 8).

Scalar A real or complex number.

Scale (1) Noun – size of an object (given some definition of size); (2) verb – to change the size of an object (equivalent to rescale).

Scaling factor A constant which is used to change the scale or size of a matrix or vector. This is done by multiplying or dividing the matrix or vector by the constant.

Score In morphometrics, a coefficient locating a specimen along a vector, calculated by projecting the specimen onto an axis. Usually, scores locate the position of a specimen relative to the axes of a coordinate system. They are calculated by taking the dot product between an axis of the coordinate system and the data of a specimen. The scores are linear combinations of the original variables. Partial warp scores locate the position of an individual specimen relative to the coordinate system provided by the partial warps. Similarly, principal component scores locate the position of an individual specimen relative to the coordinate system provided by the principal components. Scores can be calculated relative to any basis of a vector space because each basis provides a coordinate system for that space. See **Dot product**.

Semilandmark A point on a geometric feature (curve, edge or surface) defined in terms of its position on that feature (e.g. at 10% of the length of the curve from one end). Semilandmarks are used to incorporate information about curvature in a geometric shape analysis. Because semilandmarks are defined in terms of other features, they represent less information (fewer degrees of freedom) than landmarks (Chapter 15).

Shape Shape has a variety of inconsistent definitions. In geometric morphometrics, the definition of shape is Kendall's: all the geometric information remaining in an object (such as a landmark configuration) after differences in location, scale and rotational effects are removed (Chapters 1, 4, 5).

Shape coordinates Within geometric morphometrics, coordinates of landmarks after superimposition, whether by a two-point registration (which yields Bookstein shape coordinates), or Procrustes superimposition (which yields Procrustes residuals) (Chapters 3, 5).

Shape space Within geometric morphometrics, shape space refers to Kendall's shape space. The term is more general, however, as it can apply to any space defined by a particular mathematical definition of shape. There are shape spaces for outline measurements, for example. There are also shape spaces based on different definitions of size. The characteristics of these various shape spaces are not necessarily the same as those of Kendall's shape space (Chapter 4).

Shape variable A general term for any variable expressing the shape of an object, including ratios, angles, shape coordinates obtained by a superimposition method, or vectors of coefficients obtained from partial warp analysis, principal components analysis, regression, etc. Shape variables are invariant under translation, scaling and rotation.

Shear An affine or uniform deformation that leaves the Y-coordinate fixed while the X-coordinate is displaced along the X-axis by a multiple of Y. Under a shear, the point

(X, Y) maps to $(X + AY, Y)$, where A is the magnitude of the shear. Visually, this looks like altering a square by sliding the top side to the left or right, without altering its height or the lengths of the top and bottom (Chapter 5, 6).

Singular axes Orthonormal vectors produced by singular value decomposition. See **Singular value decomposition** (Chapter 11).

Singular value In a singular value decomposition, a quantity expressing a relationship between two singular axes; an element λ_i of the diagonal matrix **S**. In partial least squares analysis, each singular value represents the covariance explained by the corresponding pair of singular axes. See **Singular value decomposition** (Chapter 11).

Singular value decomposition (SVD) A mathematical technique for taking an $M \times N$ matrix **A** (where N is greater than or equal to M) and decomposing it into three matrices:

$$\mathbf{A} = \mathbf{USV}^{\mathrm{T}}$$

where **U** is an $M \times N$ matrix whose columns are orthonormal vectors, **S** is an $N \times N$ diagonal matrix with on-diagonal elements λ_i, and **V** is an $N \times N$ matrix whose columns are orthonormal vectors. The values λ_i are called the *singular values* of the decomposition, and the columns of **U** and **V** are called the *singular vectors* or *singular axes* corresponding to a given singular value. In partial least squares analysis, **A** is the matrix of covariances between the two blocks, the columns of **U** are linear combinations of the variables in one of the two data sets, the columns of **V** are linear combinations of the variables in the other data set, and each λ_i is the portion of the total covariance explained by the corresponding pair of singular axes (Chapter 11).

Singular warps Singular axes computed from shape data (partial warp scores or residuals of a Procrustes superimposition), so that the singular axes describe patterns of differences in shape. See **Singular value decomposition** (Chapter 11).

Size Any positive real valued function $g(\mathbf{X})$, where **X** is a configuration or set of points, such that $g(A\mathbf{X}) = Ag(\mathbf{X})$, where A is any positive, real scalar value. In other words, multiplying every element in **X** by A multiplies $g(\mathbf{X})$ by A. There are a wide variety of measures of size, including lengths measured between landmarks, sums or differences of interlandmark distances, square roots of area, etc. The size measure used in geometric morphometrics is centroid size. See also **Centroid size** (Chapters 3, 4).

Size-and-shape All the geometric information remaining in an object (such as a landmark configuration) after differences in location and rotational effects are removed. See **Form**.

Space A set of objects (or measurements thereof) that satisfies some definition. For example, a space might be defined as the set of all four-landmark configurations measured in two dimensions.

Statistic Any mathematical function based on an analysis of all measured individuals, e.g. the mean, standard deviation, variance, maximum, minimum, and range. The true value of the statistic in the population is called the parameter, which we are trying to estimate from our sample (Chapter 8).

Superimposition A method for matching two landmark configurations (or matrices) prior to further analysis. A number of different optimality criterion may be used. See also **Bookstein coordinates**, **Procrustes superimposition** (also **Full Procrustes superimposition** and **Partial Procrustes superimposition, RFTRA, Sliding baseline registration** (Chapters 3, 5).

Strain See **Principal strain**.

Tangent space The linear vector space tangent to a curved space. In geometric morphometrics, the Euclidean space tangent to Kendall's shape space. In the tangent space, distances between shapes are linear functions, which allows for analysis of shape variation by ordinary multivariate statistical methods. When the linear approximation to the curved surface is accurate (when all shapes in a study are close to the point of tangency), distances in the tangent space approximate distances in the curved space. The point of tangency between Kendall's shape space and the tangent space is the reference form. See also **Kendall's shape space, Reference form** (Chapter 4).

Target shape A shape being compared to the reference shape. See **Reference**.

Thin-plate spline An interpolation function used to predict the difference in shape between a reference and another shape over all points on the form, not just at landmarks. This interpolation function minimizes the bending energy of the deformation, which is equivalent to modeling that deformation as smoothly as possible given the observed landmarks (thus taking a parsimonious approach to interpolation). Thin-plate spline analysis produces scores for the non-uniform component of the deformation – scores for the uniform component are produced by a different analysis (Chapter 6).

Transformation See **Map**.

Two-point shape coordinates See **Bookstein coordinates**.

Type I, Type II error Type I error is invalidly rejecting a true null hypothesis. Type II error is failing to reject a false null hypothesis.

Type 1 landmark A landmark that can be defined in terms of local information, such as a landmark located at the junction of three bones or two bones and a muscle (i.e. anatomical features that meet at a point). There is no need to refer to any distant structures or maxima/minima of curvature. The typology of landmarks is based on Bookstein, 1991. See also **Type 2** and **Type 3 landmarks** (Chapter 2).

Type 2 landmark A landmark defined by a relatively local property, such as the maximum or minimum of curvature of a small bulge or at the endpoint of a structure. These are considered less useful than Type 1 landmarks because their evidence of homology is at least partly geometric rather than purely histological or osteological. See also **Type 1** and **Type 3 landmarks** (Chapter 2).

Type 3 landmark A landmark defined in terms of extremal points, such as the landmark on the rostrum *furthest away from* the foramen magnum. Such landmarks are regarded as deficient because they have one less degree of freedom than they have coordinates (the other degree of freedom is lost when specifying how to locate the landmark). Such landmarks can be used in geometric morphometric studies, but the loss of a degree of freedom must be taken into account when conducting statistical tests. See also **Type 1** and **Type 2 landmarks** (Chapter 2).

Uniform components The components describing the uniform deformation. For two-dimensional configurations, the uniform deformation is described by two components: compression/dilation and shear. The uniform deformation is sometimes considered the zeroth partial warp (Chapter 6).

Uniform deformation A deformation that is purely uniform (or affine), or the purely uniform component of a deformation. The uniform deformations include only the uniform transformations that alter shape (compression/dilation and shear). They do not include transformations that do not alter shape (translation, scaling and rotation). See also **Uniform shape component** (Chapters 5, 6).

Uniform component scores Scores locating a specimen, relative to the reference, along the uniform components. The summed squared scores on the uniform components and partial warps equal the Procrustes distance between each specimen and the reference. Taken together, the uniform and non-uniform scores fully describe the shape difference between the reference and that specimen (Chapter 6).

Vector A set of *P* coordinates that specify the location of a point in *P* dimensions.

Vector space A set of vectors, together with rules for adding and multiplying them (thereby obtaining all permissible linear combinations of them). Addition and scalar multiplication are required to meet eight rules:

1. $\mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}$
2. $\mathbf{X} + (\mathbf{Y} + \mathbf{Z}) = (\mathbf{X} + \mathbf{Y}) + \mathbf{Z}$
3. A unique zero vector exists such that $\mathbf{X} + \mathbf{0} = \mathbf{X}$, for all $\mathbf{X}$
4. For each $\mathbf{X}$ there exists a unique vector $-\mathbf{X}$ such that $\mathbf{X} + (-\mathbf{X}) = \mathbf{0}$
5. $1\mathbf{X} = \mathbf{X}$
6. $(C_1 C_2)\mathbf{X} = C_1(C_2\mathbf{X})$
7. $C(\mathbf{X} + \mathbf{Y}) = C\mathbf{X} + C\mathbf{Y}$
8. $(C_1 + C_2)\mathbf{X} = C_1\mathbf{X} + C_2\mathbf{X}$.

White Book Marcus, L. F., Corti, M., Loy, A. et al. (1996). *Advances in Morphometrics*. Plenum Press. (See also **Black Book**, **Blue Book**, **Orange Book** and **Red Book**.)