A3 Index to mathematical notation and symbols

Matrices and vectors

A = matrix containing elements $a _ {ij} = - \frac{1}{2} d _ {ij} ^ 2 $
B = matrix of variables (*N* × *s*) that are linear combinations of normalised **X** variables having maximum correlation with CAP axes
C = matrix of CAP axes (*N* × *s*), standardised by the square root of their respective eigenvalues
D = matrix containing elements $d _ {ij}$ corresponding to distances or dissimilarities
G = Gower's centred matrix, consisting of elements $g _ {ij} = a_ {ij} - \bar{a} _ {i \bullet} - \bar{a} _ {\bullet j} + \bar{a} _ {\bullet \bullet}$
H = 'hat' matrix = **X[X′X]$^ {-1}$X′**, used as a projection matrix for regression models
I = identity matrix, with 1's along the diagonal and 0's elsewhere
Q = matrix of PCO axes, standardised by the square root of their respective eigenvalues
Q^0 = matrix of PCO axes, orthonormalised to SSCP = **I** ('sphericised') **U** = matrix whose columns contain the left singular vectors from a singular value decomposition (SVD) of a matrix (e.g., **X** = **UWV**′); if **X** is (*N* × *q*) and *q* < *N*, then **U** is (*N* × *q*)
V = matrix whose columns contain the right singular vectors from a singular value decomposition (SVD) of a matrix (e.g., **X** = **UWV**′); if **X** is (*N* × *q*) and *q* < *N*, then **V** is (*q* × *N*)
W = diagonal matrix of eigenvalues from a singular value decomposition (SVD) of a matrix (e.g., **X** = **UWV**′); if **X** is (*N* × *q*) and *q* < *N*, then **W** is (*q* × *q*)
X = matrix of predictor variables (*N* × *q*) (often a set of environmental variables)
X^0 = matrix of **X** variables, orthonormalised to SSCP = **I** ('sphericised')
Y = matrix of response variables (*N* × *p*) (often a set of species variables)
Y^0 = matrix of **Y** variables, orthonormalised to SSCP = **I** ('sphericised')
$\hat{ {\bf Y}}$ = **HY** = matrix of fitted values (*N* × *p*)
y$_ {ij} $ = vector of *p* response variables for the *j*th observation in the *i*th group
$\bar{ {\bf y}}$ = the centroid vector of *p* response variables for group *i*
Z = matrix of dbRDA canonical axes (*N* × *s*)

Letters

a, *b*, *c*, etc… = number of levels of factor A, B, C, etc… in an ANOVA experimental design
AIC = multivariate analogue to Akaike's information criterion
AIC$_c$ = multivariate analogue to the small-sample-size corrected version of *AIC*
B$_l$ = the *l*th variable in the space of normalised **X** variables that has maximum correlation with the *l*th coordinate axis (*C*$_l$) from a CAP analysis

BIC = multivariate analogue to Schwarz's 'Bayesian information criterion'
C$ _l $ = the *l*th coordinate axis scores from a CAP analysis
d$ _ {ij} $ = distance or dissimilarity between sample *i* and sample *j* *df* = degrees of freedom
F = pseudo-*F* statistic for testing hypotheses in PERMANOVA or DISTLM
i = index used for samples (i.e., *i* = 1, ..., *N*) or index used for groups (*i* = 1, ..., *a*)
j = second index used for samples (i.e., *j* = 1, ..., *N*) or index used for replicates within a group (*j* = 1,..., *n*)
k = index used for variables (i.e., *k* = 1, ..., *p* or else *k* = 1, ..., *q*)
l = index used for canonical axes or eigenvalues for either dbRDA or CAP (i.e., *l* = 1, ..., *s*) or either the abbreviation for 'log-likelihood' or the 'length' of a vector (depending on context).
m = number of PCO axes chosen as a subset for analysis by CAP *MC* = Monte Carlo
MS = mean square
N = total number of samples
n = number of samples (replicates) within a group or cell in an experimental design
P = *P*-value associated with the test of a null hypothesis *p* = number of multivariate response variables in matrix **Y** *q* = total number of predictor variables in matrix **X**
r = Pearson correlation coefficient
R = the ANOSIM *R* statistic (see Clarke 1993)
R^2 = proportion of explained variation from a model
s = number of canonical eigenvalues and associated canonical axes obtained from either a dbRDA *or* a CAP analysis *SS* = sum of squares
SSCP = sum of squares and cross products
SVD = singular value decomposition
t = pseudo-*t* statistic = $\sqrt{}$pseudo- *F*
tr = 'trace' of a matrix = the sum of the diagonal elements
X$ _ k $ = the *k*th predictor variable
Y$ _ k $ = the *k*th response variable
z$ _ {ij} $ = distance to group centroid for the *j*th replicate within the *i*th group.

Greek symbols and matrices

$ \alpha$ = significance level chosen for a test (usually α = 0.05).
$ \delta _ l ^ 2$ = the *l*th eigenvalue from a CAP analysis, a squared canonical correlation
$ \Delta$ = diagonal matrix containing the square roots of the eigenvalues from a CAP analysis (a capital delta)
$ \gamma _ l ^ 2$ = the *l*th eigenvalue from a dbRDA analysis, a portion of the explained (regression) sum of squares from a dbRDA model.
$ \Gamma$ = diagonal matrix containing the square roots of the eigenvalues from a dbRDA analysis (a capital gamma)
$ \lambda _ i $ = the *i*th eigenvalue from a PCO analysis
$ \Lambda $ = diagonal matrix of eigenvalues from a PCO analysis (a capital lambda)
$ \nu$ = number of parameters in a particular model during model selection
$ \rho $= Spearman rank correlation (rho)
$ \sum $ = sum over the relevant index

Revision #25
Created 18 August 2022 10:29:55 by Arden
Updated 18 August 2022 11:40:37 by Arden