

# 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 ( $N \times s$ ) that are linear combinations of normalised **X** variables having maximum correlation with CAP axes

**C** = matrix of CAP axes ( $N \times s$ ), 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  $g_{ij} = a_{ij} - \bar{a}_{i.} - \bar{a}_{.j} + \bar{a}_{..}$

**H** = 'hat' matrix =  $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ , 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

$\mathbf{Q}^0$  = 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.,  $\mathbf{X} = \mathbf{U}\mathbf{W}\mathbf{V}'$ ); if **X** is ( $N \times q$ ) and  $q < N$ , then **U** is ( $N \times q$ )

**V** = matrix whose columns contain the right singular vectors from a singular value decomposition (SVD) of a matrix (e.g.,  $\mathbf{X} = \mathbf{U}\mathbf{W}\mathbf{V}'$ ); if **X** is ( $N \times q$ ) and  $q < N$ , then **V** is ( $q \times N$ )

**W** = diagonal matrix of eigenvalues from a singular value decomposition (SVD) of a matrix (e.g.,  $\mathbf{X} = \mathbf{U}\mathbf{W}\mathbf{V}'$ ); if **X** is ( $N \times q$ ) and  $q < N$ , then **W** is ( $q \times q$ )

**X** = matrix of predictor variables ( $N \times q$ ) (often a set of environmental variables)

$\mathbf{X}^0$  = matrix of **X** variables, orthonormalised to SSCP = **I** ('sphericised')

**Y** = matrix of response variables ( $N \times p$ ) (often a set of species variables)

$\mathbf{Y}^0$  = matrix of **Y** variables, orthonormalised to SSCP = **I** ('sphericised')

$\hat{\mathbf{Y}} = \mathbf{H}\mathbf{Y}$  = matrix of fitted values ( $N \times p$ )

$\mathbf{y}_{ij}$  = vector of  $p$  response variables for the  $j$ th observation in the  $i$ th group

$\bar{\mathbf{y}} =$  the centroid vector of  $p$  response variables for group  $i$

**Z** = matrix of dbRDA canonical axes ( $N \times s$ )

## 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 'An 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.,  $i = 1, \dots, N$ ) or index used for groups ( $i = 1, \dots, a$ )  
 $j$  = second index used for samples (i.e.,  $j = 1, \dots, N$ ) **or** index used for replicates within a group ( $j = 1, \dots, n$ )  
 $k$  = index used for variables (i.e.,  $k = 1, \dots, p$  or else  $k = 1, \dots, q$ )  
 $l\ \$$  = index used for canonical axes or eigenvalues for either dbRDA **or** CAP (i.e.,  $l\ \$ = 1, \dots, s$ ) **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 =  $\sqrt{\text{pseudo-} F}$   
 $tr$  = '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  $\alpha = 0.05$ ).  
 $\Delta_l^2$  = 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 #45

Created 18 August 2022 10:29:55 by Arden

Updated 11 December 2024 03:43:26 by Abby Miller