Application of hypercorrelated matrices in ecological research

Branko Karadžić, Snežana Jarić, Pavle Pavlović, Saša Marinković, Miroslava Mitrović

Institute for Biological Research, 'Siniša Stanković' of Belgrade University, Despota Stefana 142, 11000, Belgrade, Serbia

Email address:

branko@ibiss.bg.ac.rs (B. Karadžić), nena2000@ibiss.bg.ac.rs (S. Jarić), ppavle@ibiss.bg.ac.rs (P. Pavlović),
grifon@ibiss.bg.ac.rs (S. Marinković), mmit@ibiss.bg.ac.rs (M. Mitrović)

Abstract: Ecological data matrices often require some form of pre-processing so that any undesirable effects (*e.g.* the variable size effect) may be removed from multivariate analyses. This paper describes hypercorrelation, a simple data transformation that improves ordination methods significantly. Hypercorrelated matrices efficiently eliminate the 'arch' (or Guttman) effect, a spurious polynomial relation between ordination axes. These matrices reduce the sensitivity of correspondence analysis to outliers. Canonical analyses (canonical correspondence analysis and redundancy analysis) of hypercorrelated matrices are resistant to undesirable effects of missing data. Finally, the hypercorrelation extends applicability of "linear ordination method" (principal components analysis and redundancy analysis) to sparse (high beta diversity) matrices.

Keywords: Arch Effect, Beta Diversity, (Canonical) Correspondence Analysis, Hypercorrelation, Missing Data, Outliers, Principal Components Analysis, Redundancy Analysis

1. Introduction

Correspondence analysis (CA) and principal components analysis (PCA) with their canonical forms are the most frequently used ordination methods in ecology [1-6].

A spurious polynomial relation between ordination axes (the arch effect or Guttman effect) is a well-known drawback of CA and PCA. Compared to CA, PCA is more sensitive to the arch effect, especially in the case when beta diversity (species turnover) along spatial or environmental gradients is high. Therefore, the principal components analysis and its canonical variant (redundancy analysis) are inappropriate for analyses of long environmental gradients.

Sensitivity to 'outliers' is another fault of CA [1-6]. Application of CA to matrices with sparse vectors often produces uninterpretable results. In such cases, CA highlights the importance of outliers (sparse vectors), obscuring the remaining data variability.

Application of canonical correspondence analysis (CCA) and redundancy analysis (RDA) to matrices with missing data may produce quite distorted and ecologically uninterpretable results.

In this article, we propose a simple solution for these problems. The solution is based on hypercorrelated matrices. Comparative tests with simulated data revealed that hypercorrelated matrices significantly improve the performance of (canonical) correspondence analysis, PCA and RDA.

2. Decorrelated and Hypercorrelated Matrices

Suppose that $X_{(nxm)}$ is a matrix that describes the distribution of *n* species in *m* sites. The matrix specifies the position of *m* points in *n*-dimensional Euclidean space. The axes of referent space are mutually orthogonal, but not necessarily independent. We may assess the statistical dependence between two variables using either squared or absolute value of Pearson correlation coefficient. Both quantities may vary from 0 (if two variables are statistically independent) to 1 (if two variables are linearly dependent and perfectly correlated).

We may either eliminate or increase linear dependence between rows of **X**. *Decorrelation* and *hypercorrelation*, two opposite processes that reduce and increase statistical dependence between variables, may be performed using

simple transformations of data matrices.

The Mahalanobis transformation

$$\mathbf{M}=(\mathbf{XX}^T)^{-0.5} \tag{1}$$

decorrelates variables that constitute the basis of n-dimensional Euclidean space. After this transformation, the correlation between each pair of axes (each pair of species) is reduced to zero.
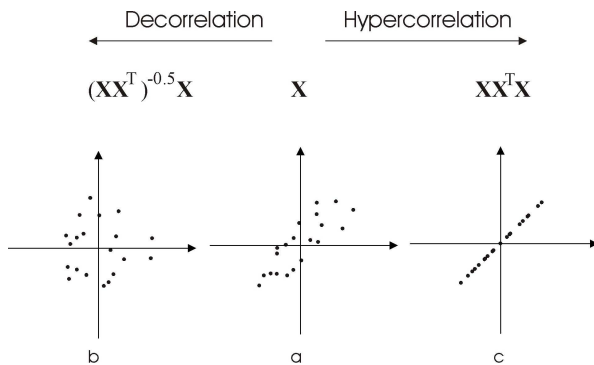
The hypercorrelated matrix $\mathbf{X}'=\mathbf{XX}^T\mathbf{X}$ increases statistical (linear) dependence between variables. It is easy to prove this regularity. The factorization of X by the singular value decomposition (SVD) gives:

$$\mathbf{X}=\mathbf{P}\Sigma\mathbf{Q}^T \tag{2}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are orthogonal matrices and $\Sigma=diag(\sigma)$ is the matrix of singular values. Therefore,

$$\mathbf{XX}^T\mathbf{X}=(\,\mathbf{P}\Sigma\mathbf{Q}^T)\,\mathbf{Q}\Sigma\mathbf{P}^T(\,\mathbf{P}\Sigma\mathbf{Q}^T)=\mathbf{P}\Sigma^3\mathbf{Q}^T \tag{3}$$

because $\mathbf{P}^T\mathbf{P}=\mathbf{I}_n$ and $\mathbf{Q}^T\mathbf{Q}=\mathbf{I}_m$. Since $\sigma_1>\sigma_2>...>\sigma_r$, it is obvious that $\sigma_1^3>>\sigma_2^3>>...>>\sigma_r^3$, $r=\min(m,n)$. This relationship indicates that hypercorrelated matrices elongate multidimensional ellipsoids. If $\sigma_1>>\sigma_2$ (e.g. if $\sigma_1/\sigma_2>4$), then hypercorrelation transforms multidimensional ellipsoids into multidimensional lines (Fig.1). In that case, all rows of the hypercorrelated matrix $\mathbf{X}'$ are linearly dependent and perfectly correlated.

Fig 1. *Mathematical operations that alter the linear dependence between rows (columns) of data matrices. a) Graphical equivalent of a data matrix X. b) The Mahalanobis transformation decorrelates variables. c) The hypercorrelated matrix increases statistical (linear) dependence between variables.*

2.1. Ecological Interpretation and Properties of Hypercorrelated Matrices

The ecological meaning of $\mathbf{X}'=\mathbf{XX}^T\mathbf{X}$ is hidden. However, the hypercorrelated matrix may be represented using an alternative equation:

$$\mathbf{X}'=\mathbf{XS} \tag{4}$$

where $\mathbf{S}=\mathbf{X}^T\mathbf{X}$ is the mxm-dimensional similarity matrix containing information on floristic (or faunistic) similarity between pairs of sites. Above equation clearly indicates that the hypercorrelated matrix is specifically transformed species-by-sites matrix. The most important effects of this

transformation involve a selective reduction of data variability and the elimination of zeroes from data matrices.

The selective reduction of data variability is an unique and very useful property of hypercorrelated matrices. The hypercorrelation does not affect the main data variability that is associated with first principal axis. The variability reduction gradually increases with decreasing importance of principal axes. This regularity may be proved easily. The matrix $\mathbf{X}=\mathbf{Q}\Sigma\mathbf{P}^T$ of rank r may be represented as the sum of r matrices:

$$X = \sum_{i=1}^{r} \sigma_i p_i q_i^T \tag{5}$$

where for each $i=1,...,r$ the singular value σ_i, and corresponding singular vectors $\mathbf{p}_i$ and $\mathbf{q}_i^T$ satisfy the condition $\mathbf{p}_i\mathbf{X}=\sigma_i\mathbf{q}_i^T$ [7]. Similarly, the matrix $\mathbf{X}=\mathbf{Q}\Sigma^3\mathbf{P}^T$ may be represented as the sum of r matrices:

$$X' = \sigma_1^2\left(\sum_{i=1}^{r} a_i\sigma_i p_i q_i^T\right) \tag{6}$$

where $a_i=\sigma_i^2/\sigma_1^2$. The matrices $\mathbf{X}$ and $\mathbf{X}'$ are closely related. They differ with respect to the scaling factor σ_1^2 and the weighting coefficients a_i. The multidimensional configuration of rows (columns) of a data matrix remains unaltered if we multiply the matrix by a scalar. Therefore, the scaling factor σ_1^2 has no impact on data variability. On the other hand, the weighting coefficients a_i have profound effects on variability patterns. We may assume that all weighting coefficients in $\mathbf{X}$ are the same ($a_1=a_2=...=a_r=1$). Starting from $a_1=1$, the weighting coefficients of the hypercorrelated matrix subsequently decrease, gradually reducing total data variability. The selective reduction of data variability is very important because it may contribute to a reduction of the 'arch effect'.

Since

$$x'_{i,j} = \sum_{h=1}^{m} x_{i,h} s_{h,j} \tag{7}$$

it is obvious that $x'_{i,j} \geq x_{i,j}$. This relation indicates that the *hypercorrelation may eliminate zero values from sparse matrices*. Suppose that a species is unrecorded in a site ($x_{i,j} = 0$). The corresponding $x'_{i,j}$ will be greater than zero if the species is present in another site, and if these two sites share at least one species. Since hypercorrelated matrices eliminate zeros from sparse vectors, they may reduce the sensitivity of CA to rare categories. Moreover, the elimination of zeroes may reduce the sensitivity of "linear ordination methods" to the effects of long gradients.

2.2. Singular Vectors of Hypercorrelated Matrices

Matrices $\mathbf{X}$ and $\mathbf{X}'$ share the same set of singular vectors. Therefore, principal component analyses of $\mathbf{X}$ and $\mathbf{X}'$ produce trivial results that differ only with respect to singular values. We may avoid the trivial solution if we

perform a data transformation of these matrices. Suppose, for example, that we performed centering of $\mathbf{X}$ and $\mathbf{X}$'. Resulting matrices $\mathbf{X}_c$ and $\mathbf{X}_c$' have different sets of singular vectors since $\mathbf{X}_c' \neq \mathbf{X}_c\mathbf{X}_c^T\mathbf{X}_c$. Any transformation applied to the matrices $\mathbf{X}$ and $\mathbf{X}$' assures that singular vectors of transformed matrices are different.
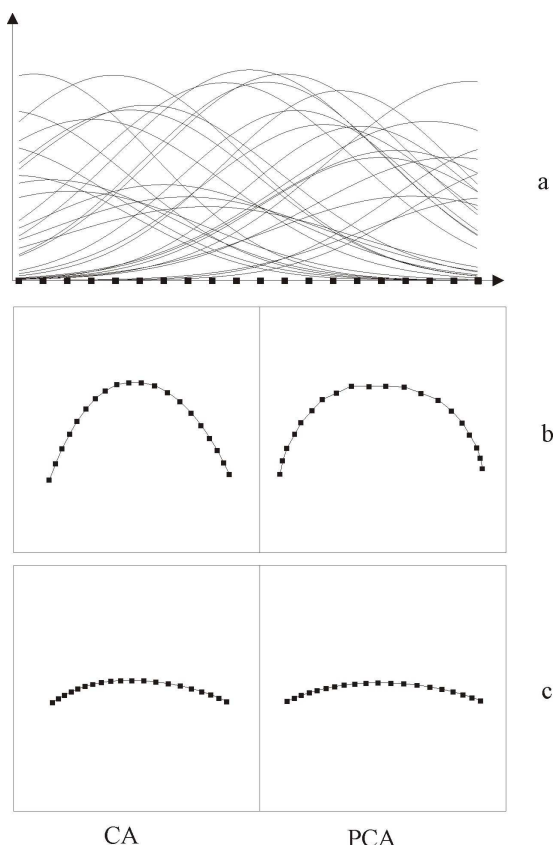
Fig 2. *Using a coenocline model (a), we created a data matrix that describes the distribution of 30 species in 20 sites. Sites (squares) are linearly distributed along a single environmental gradient. b) Correspondence analysis (CA) and principal components analysis (PCA) represented the sites as an arch-shaped, rather than a linear sequence of points. c) Applying CA and PCA to the hypercorrelated matrix, we obtained almost ideal results.*

3. Methods

We evaluated the effects of hypercorrelated matrices on different ordination methods using the simulated data, derived from explicit models of community variation along environmental gradients. The use of simulated data in comparative tests has apparent advantage since the ordination performance may be assessed by comparing ordination configurations with the configuration of samples in the simulated environmental space [8].

Artificial species-by-sites matrices may be created using a wide spectrum of ecological models [9-11]. Using either univariate Gaussian curves or bivariate Gaussian surfaces to simulate distribution (response) of species along one or two independent environmental gradients, we created a set of data matrices. Sites were positioned at regular intervals along the environmental gradients. Modal coordinates of

species (species optima) were randomly distributed on the gradients. The β diversity along simulated gradients (expressed in Half Change-HC units, *sensu* Gauch & Whittaker [12]) varied from 3 to 6 HC.

We analysed the effects of hypercorrelation on CA, CCA, PCA and RDA using the 'FLORA' package [13], specifically the latest version of the package [14]

Initial processing of data matrices (*e. g.* data centering and/or data standardization) has essential impact on PCA and RDA results [15-19] Most authors [2, 17, 19] recommend use of PCA (and RDA) after the Hellinger transformation. Following these recommendations, we performed PCA and RDA using a stepwise preparation of data matrices. The first step involved the Hellinger transformation of either simulated matrix X or its hypercorrelated equivalent X'. The second step involved centering of transformed matrices. We also performed other standardization options (*e. g.* standardizations by site totals, Euclidean norm and general norm) instead of the Hellinger transformation. Different options produced essentially similar results. Therefore we presented only results obtained by PCA and RDA of matrices that are preprocessed using the Hellinger transformation.

4. Results

4.1. Sensitivity to the Arch Effect

Using a coenocline model (Fig 2a), we created a data matrix that describes the distribution of 30 species in 20 sites. Despite a low beta diversity (3 HC), CA and PCA produced distorted results. These methods represented the sites as an arch-shaped, rather than a linear sequence of points (Fig. 2b). Applying both methods to the hypercorrelated matrix produced almost ideal results.

Compared to CA, the "linear ordination methods" (PCA, principal coordinate analysis and their canonical variants) are more sensitive to the arch effect, especially in the case when the species turnover along spatial or environmental gradients (directional beta diversity *sensu* Anderson *et al*. [20]) is high. To confirm this regularity we used coenoplanes with relatively high species turnover. Using a coenoplane model, we generated a data matrix with 77 species and 60 sites. The sites were distributed in a regular 10x6 grid along two independent gradients with high beta diversities (5 x 3 HC). Applying CA and PCA to the coenoplane, we obtained different results. CA generated square instead of rectangular pattern of sites (Fig 3a). Nevertheless, CA accurately recovered the rank order of samples along simulated gradients. PCA, however, severely distorted the planar pattern of sites (Fig 3a). The sites that are located at opposite ends of environmental gradients have no species in common. Therefore, these sites should be maximally separated in ordination plane. However, PCA locates these samples in a near proximity. Interpretability of such results may be questioned. Legendre & Galagher [18] emphasized that the Hellinger transformation removes the horseshoe effect of

PCA. However, our results clearly indicate that the Hellinger transformation is not sufficient for removing undesirable curvilinear distortion. Applying CA and PCA to the hypercorrelated matrix, we obtained more acceptable results (Fig 3b). We repeated experiments using the coenoplanes with higher beta diversities (*e. g.* 6x6 HC coenoplanes), and found out that PCA of hypercorrelated matrices is resistant to the arch distortion. It may be concluded that hypercorrelation extends applicability of PCA to matrices with long gradients.
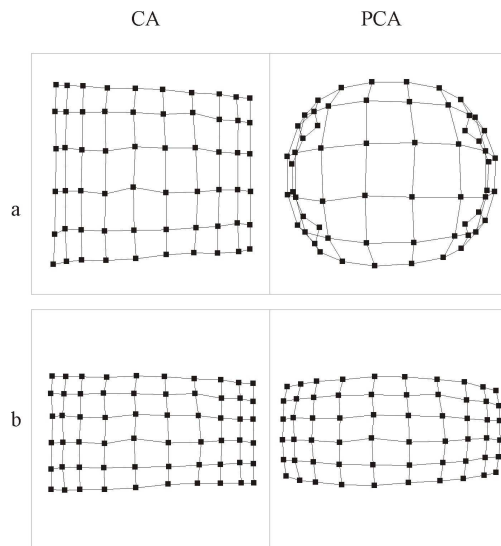
Fig 3. *CA and PCA of a 5 x 3 half-change coenoplane. a) CA generated a square instead of rectangular pattern of sites. Nevertheless, CA accurately recovered the rank order of samples along simulated gradients. Due to the arch effect, PCA severely distorted the planar pattern of sites. b) Both CA and PCA of the hypercorrelated matrix accurately recovered expected (known) configuration of sites along simulated gradients.*

4.2. The Rarity Problem

The application of CA to matrices with sparse vectors often produces uninterpretable results. In such cases, CA highlights sparse vectors, obscuring the remaining data variability. To simulate outliers, we generated an artificial 130x81 data matrix. We generated the main part of the matrix (the 120x80 submatrix), assuming that samples are positioned in a regular 10x8 grid along two independent gradients. We added ten rare species and one atypical site to the main submatrix. Each of the rare species occurs in two sites (the atypical site and another, randomly selected site). Applying CA to the extended matrix, we obtained quite unacceptable results. CA overemphasized the importance of outliers. A cluster of overlapping points (Fig. 4) represents all the other sites. Such a result is confusing, since a large proportion of data variability is hidden.

Downweighting of rare categories may reduce the undesirable effects of outliers [21, 22]. If F*max* is the frequency of the most common species, then species rarer than F*max*/5 are downweighted in proportion to their frequency. We used downweighting of either sites, or species, or both sites and species in order to improve CA. However,

each of the downweighting options failed to eliminate the undesirable effect of outliers (Fig. 4).

Applying CA to the hypercorrelated matrix, we obtained more acceptable result. CA of hypercorrelated matrix also emphasized the outlier, but contrary to ordinary CA, it revealed the latent coenoplane structure (Fig 4d).
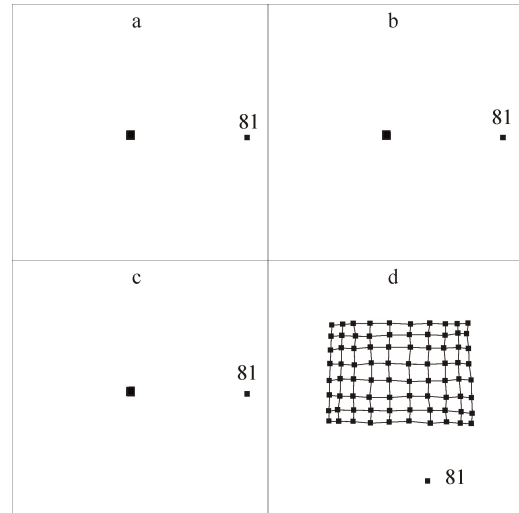
Fig 4. *The effects of outliers on CA. The data matrix describes the distribution of 130 species in 81 sites. Atypical site #81 has 10 rare species. a) Amplifying the importance of the atypical site, CA obscured the main variability pattern of the coenoplane. Down-weighting of either species (b) or both species and sites (c) is unable to improve CA results. d) Correspondence analysis of the hypercorrelated matrix successfully recovered latent structure of the coenoplane.*

4.3. Canonical Analyses of Hypercorrelated Matrices

Redundancy analysis [23] and canonical correspondence analysis [24] are explanatory methods that may be used for testing the null hypothesis of species-environment independence. Since these methods operate using two data matrices (the 'response' data matrix $X_{(n \times m)}$, which describes the distribution of n species in m sites, and the 'explanatory' matrix $Z_{(q \times m)}$, which contains values of q environmental variables in m sites), we may hypercorrelate either the 'response' matrix, or the 'explanatory' matrix, or both. In order to evaluate these options, we performed a simple test.

We used the coenocline with 30 species and 20 sites (Fig 2a) as a response matrix. The 'explanatory' matrix consisted of 20 variables (one moderately noisy and 19 random variables). The moderately noisy variable was generated using a stepwise procedure. Firstly, we defined a noiseless variable using the site positions along the gradient. The noiseless variable was perfectly correlated with community samples. Then, we added a small amount of stochastic variability to each element of the noiseless variable. Despite this randomization procedure, the resulting variable was highly correlated with sampling sites (r1=0.96). Another set of 19 variables consisted of random numbers.

The statistical significance of the effect of each variable was tested by the Monte Carlo permutation test, using 1000 randomized runs for each analysis. We revealed that all CCA options accurately detected the statistically significant

relationship of species with the moderately noisy variable.

Irrespective of the CCA variant we used, statistical tests should have confirmed that species and random variables are unrelated. The opposite result is referred to as a type I error. Hypercorrelation of either explanatory or both explanatory and response data matrices significantly increased the level of type I errors. More precisely, in both cases, we always obtained erroneous results (Fig. 5). Applying RDA to the same data set, we obtained essentially the same result.

Hypercorrelation of the explanatory matrix increases collinearity among environmental variables. Increased collinearity among variables may spuriously increase statistical importance of a random variable. In such a situation, the statistical tests may mislead to wrong conclusion about species-environment relationship. To avoid this problem, we recommend only hypercorrelating the response matrix.
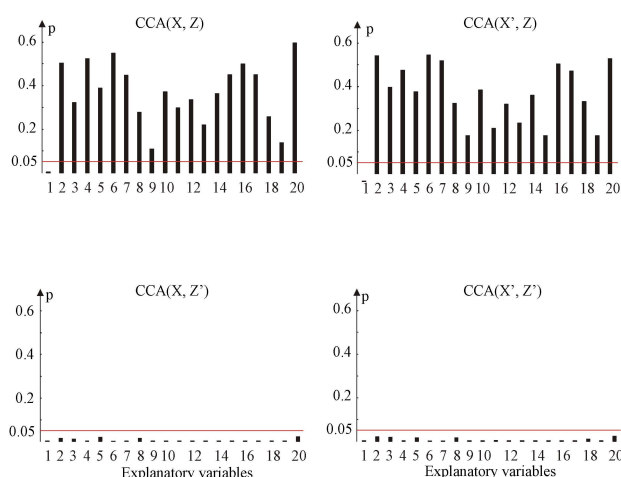
Fig 5. Sensitivity of different CCA variants to type I error. We used the coenocline with 30 species and 20 sites (Fig. 2a) as a response matrix. Species are unrelated to a set of 19 explanatory variables (variables 2-20), since each variable consisted of random numbers. For each random variable the probability of the null hypothesis on species-environment independence should be greater than a significance limit of 5%. Both CCA of untransformed matrices [CCA(X,Z)] and CCA of hypercorrelated response matrix [CCA(X',Z)] accurately detected the statistical significance of the species-environment relationship. CCA of hypercorrelated explanatory matrix [CCA(X,Z')] and CCA of hypercorrelated response and explanatory matrices [CCA(X',Z')] produced unacceptable amount of type I errors.

4.4. The Missing Data Problem

Both, CCA and RDA produce two kinds of site scores [2, 25, 26]. Palmer [25] denoted dual site scores of CCA as linear combination scores (LC) and weighted averages scores (WA). LC site scores are linear combinations of environmental variables. WA site scores are weighted averages of species scores. The term WA score is inappropriate for RDA, since RDA has nothing in common with the weighted averaging procedures. The more suitable term for dual scores of RDA are linear combination (LC) scores and weighted summation (WS) scores.

The choice of site scores is a controversial topic [25 - 27]. The most important consequence of the choice is how the

ordination reacts to noise [26]. LC scores can be highly sensitive to moderate noise in the environmental data. WA scores in CCA and WS scores in RDA are insensitive to environmental noise. However, as we illustrate in the following example, even if we use WA scores in CCA (or WS scores in RDA), canonical analyses may produce unacceptable results if some of environmental data are missing.

Using the Gaussian surface model, we generated a 'response' matrix with 88 species and 70 sites. Sites were distributed in a regular 10x7 grid along two independent gradients. The 'explanatory' matrix consisted of two moderately noisy environmental variables. We deleted records of environmental variables in 3 out of 70 sites in order to simulate missing data. When performing either CCA or RDA to such a modified matrix, we obtained unacceptable results. Both WA scores of CCA and WS scores of RDA severely distorted regular pattern of sites (Fig. 6). CCA and RDA of the hypercorrelated response matrix successfully recovered the expected (known) configuration of sites along the gradients. This indicates that hypercorrelated matrices reduce the sensitivity of canonical analyses to missing data.
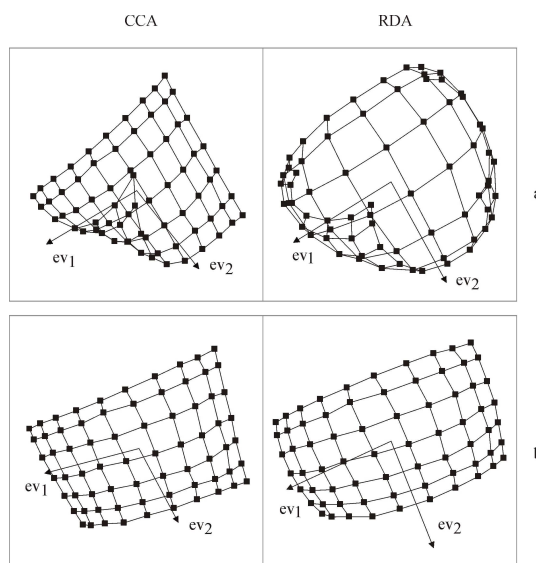
Fig 6. The sensitivity of CCA to missing data. The 'explanatory' matrix consisted of two moderately noisy environmental variables (ev_1 and ev_2). We deleted records of enivironmental variables in 3 out of 70 sites, in order to simulate missing data. a) Both CCA and RDA severely inverted position of sites along the environmental gradients. b) Hypercorrelation of the response matrix improved CCA and RDA since both methods accurately revealed the expected position of sites.

5. Conclusion

Correspondence analysis (CA) and principal components analysis (PCA) with their detrended and/or canonical forms are the most popular ordination methods in ecology. Due to well-known problems that arise because of outliers, long gradients or missing data, these methods may produce ecologically non-interpretable results.

Hypercorrelation is a powerful solution of these

problems.

By eliminating zeros from sparse vectors, the hypercorrelated matrices reduce the sensitivity of CA to rare categories. Compared to ordinary CA, the CA of hypercorrelated matrices is more resistant to outliers.

Moreover, due to reduction of zeroes from sparse matrices, the hypercorrelation extends applicability of linear ordination methods significantly. Both PCA and RDA of hypercorrelated matrices produce ecologically meaningful results, irrespective on beta diversity along simulated gradients.

Canonical analyses are sensitive to missing data. A small proportion of missing data may severely distort CCA and RDA results. CCA and RDA of hypercorrelated matrices are resistant to undesirable effects of missing data.

Acknowledgements

References

[1] Jongman, R.H., ter Braak, C.J.F. & van Tongeren, O.F.R. (*eds*.). *Data analysis in community and landscape ecology.* Pudoc, Wageningen. 1987.

[2] Legendre, P. & Legendre, L. *Numerical Ecology.* Elsevier, Amsterdam. 1998.

[3] Karadžić, B. & Marinković, S. *Quantitative ecology.* IBISS, Belgrade. (in Serbian). 2009.

[4] Greenacre, M.J. *Biplots in Practice.* Fundación BBVA, Barcelona. 2010.

[5] Gauch, H.G. *Multivariate analysis in community ecology.* Cambridge University Press, Cambridge. 1982.

[6] Greenacre, M.J. *Theory and Applications of Correspondence Analysis.* Academic Press, London. 1984.

[7] Gabriel, K.R. The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58, 453-467. 1971.

[8] Minchin, P.R. An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio*, 69, 89-107. 1987.

[9] Austin, M.P. Continuum concept, ordination methods and niche theory. *Annual Review of Ecology and Systematics*, 16, 39-61. 1985.

[10] Minchin, P.R. Simulation of multidimensional community patterns: towards a comprehensive model. *Vegetatio,* 71, 145-156. 1987.

[11] Karadžić, B., Marinković, S. & Kataranovski, D. Use of the β-function to estimate the skewness of species responses. *Journal of Vegetation Science,* 14, 799-805. 2003.

[12] Gauch, H.G. & Whittaker, R.H. Simulation of community patterns. *Vegetatio*, 33, 13-16. 1976.

[13] Karadžić, B., Šašo-Jovanović, V., Jovanović, Z., & Popović, R. 'FLORA' a database and software for floristic and vegetation analyses. *Progress in Botanical Research* (eds I. Tsekos, & M. Moustakas). Kluwer Academic Publishers, Dodrecht, pp. 69-72. 1998.

[14] Karadžić, B. FLORA: A Software Package for Statistical Analysis of Ecological Data. *Water Research and Management,*3, 45-54. 2013.

[15] Noy-Meir, I., Walker, D. & W.T. Williams. Data transformations in ecological ordination. II. On the meaning of data standardization. *J. Ecol.* 63, 779-800. 1975.

[16] Karadžić, B. & Popović, R. The generalized standardization procedure in ecological ordination: Tests with "metric" ordination methods. *Journal of Vegetation Science*, 5, 259-262. 1994.

[17] Rao, C.R. A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Qüestiió (Quaderns d'Estadística i Investigació Operativ*a), 19, 23-63. 1995.

[18] Karadžić, B., Šašo-Jovanović, V., Jovanović, Z. & Karadžić, D. On detrending in correspondence Analysis and Principal Components Analysis. *Ecoscience*, 6, 110-116. 1999.

[19] Legendre, P. & Gallagher, E.D. Ecologically meaningful transformations for ordination of species data. *Oecologia* 129:271–280. 2001.

[20] Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N. J.B., Stegen, J.C. & Swenson, N.G. Navigating the multiple meanings of b diversity: a roadmap for the practicing ecologist. *Ecology Letters*, 14, 19–28. 2010.

[21] McCune, B. & Mefford, M.J. *Multivariate analysis on the PC-ORD system*. Version 2.0. MjM Software, Gleneden Beach, Oregon, USA. 1995.

[22] Hill, M.O. & Šmilauer, P. *TWINSPAN for Windows version 2.3.* Centre for Ecology and Hydrology & University of South Bohemia, Huntingdon & Ceske Budejovice. 2005.

[23] Rao, C.R. The Use and Interpretation of Principal Component Analysis in Applied Research. *Sankhyā: The Indian Journal of Statistics*, Series A, 26, 329-358. 1964.

[24] ter Braak, C.J.F. Canonical Correspondence Analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, 67, 1167-1179. 1986.

[25] Palmer, M.W. Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology* 74, 2215-2230. 1993.

[26] McCune, B. Influence of noisy environmental data on canonical correspondence analysis. *Ecology,* 78, 2617-2623. 1997.

[27] Oksanen, J., Blanchet, F.G, Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. *Vegan: Community ecology package*. R package Version 1.17-3. 2010.