

Factor Analysis of Local Formalism

François Bavaud¹ and Christelle Cocco²

¹ University of Lausanne, `francois.bavaud@unil.ch`

² University of Lausanne, `christelle.cocco@unil.ch`

Abstract. Local formalism deals with weighted unoriented networks, specified by an exchange matrix, determining the selection probabilities of pairs of vertices. It permits to define local inertia and local autocorrelation relatively to arbitrary networks. In particular, free partitioned exchanges amount in defining a categorical variable (hard membership), together with canonical spectral scores, identical to Fisher’s discriminant functions. One demonstrates how to extend the construction of the latter to any unoriented network, and how to assess the similarity between canonical and original configurations, as illustrated on four datasets.

1 Introduction

Introducing a *neighborhood relation* between pairs of observations permits to construct a *local formalism*, and a local variance in particular. Comparing the latter to the ordinary variance defines the Durbin-Watson, Moran or Geary measures of autocorrelation. Local formalisms are central to spatial statistics ever since the fifties (Moran 1950; Geary 1954); they have also been considered by a few authors in the data analytic community (e.g. Lebart 1969; Le Foll 1982; Meot et al. 1993; Thioulouse et al. 1995).

Section 2 exposes a quite general yet tractable local formalism, based upon *two primitives only*, namely an *exchange matrix* E between observations, determining the selection probabilities of pairs of observations, defining an un-oriented weighted network together with the observation weights, and a *dissimilarity matrix* D between observations, chosen as squared Euclidean. This formalism, closely related to reversible Markov chain theory and spectral clustering, defines in particular a *local inertia* and a *relative inertia*, generalising the local variance and Moran’s I .

Categorical variables, that is groups of observations, emerge as a particular instance of the local formalism, namely under *free partitioned exchange matrices* (section 3.1). This circumstance enables to generalise *for any network* the computation of *canonical scores* (Fisher discriminant functions), maximising the between-groups dispersion (section 3.2). Canonical scores and factor

scores (the latter maximising the global inertia) are compared by means of two presumably original *weighted configuration similarity* indices (section 3.3).

In the last part (section 4), the theory is illustrated on four datasets. The higher the relative inertia, the more similar are the canonical and factor configurations. Also, the Markovian origin of the formalism allows to consider higher-order exchange matrices, and to construct iterated canonical configurations, in contrast to ordinary discriminant analysis, idempotent in nature.

2 Definitions and notations

2.1 Exchange matrix and local variance

In spatial or temporal contexts, neighborhood relations between n observations can be expressed by means of an $n \times n$ *exchange matrix* (Berger and Snell 1957; Bavaud 2008), describing an *unoriented weighted network*:

$$E = (e_{ij}) \quad e_{ij} \geq 0 \quad e_{ij} = e_{ji} \quad f_i := \sum_j e_{ij} > 0 \quad \sum_{ij} e_{ij} = 1 .$$

Here e_{ij} can be interpreted as the probability to select the pair of individuals (i, j) , and f_i as the probability to select individual i , that is the *weight* of the observation. The ordinary (weighted) variance reads as

$$\text{var}(x) = \sum_i f_i (x_i - \bar{x})^2 = \frac{1}{2} \sum_{ij} f_i f_j (x_i - x_j)^2 \quad \text{where} \quad \bar{x} := \sum_i f_i x_i .$$

The *local variance* (Lebart 1969; see also e.g. Bavaud 2013) is defined as

$$\text{var}_{\text{loc}}(x) = \frac{1}{2} \sum_{ij} e_{ij} (x_i - x_j)^2 = \text{var}(x) - \text{cov}(x, Wx) \quad w_{ij} := \frac{e_{ij}}{f_i} \quad (1)$$

where $W = (w_{ij})$ is the transition matrix of a reversible Markov chain. The canonical measure of autocorrelation - *Moran's I* - is defined as

$$I(x) := \frac{\text{var}(x) - \text{var}_{\text{loc}}(x)}{\text{var}(x)} \quad \text{with} \quad -1 \leq I(x) \leq 1 .$$

Iterated exchange matrices are given by $E^{(r)} := \Pi W^r$, where W is given in (1) and $\Pi = \text{diag}(f)$. As $E^{(0)} = \Pi$ (frozen network) and, at least for regular chains, $E^{(\infty)} = f f'$ (complete network), one gets $\delta^{(r=0)} = 1$ and $\delta^{(r=\infty)} = 0$.

2.2 Eigen-decomposition

The *standardised exchange matrix* $E^s = (e_{ij}^s)$ and its spectral decomposition

$$e_{ij}^s := \frac{e_{ij} - f_i f_j}{\sqrt{f_i f_j}} \quad E^s = T \Gamma T' \quad T = (t_{i\alpha}) \quad \Gamma = \text{diag}(\gamma_\alpha) \quad (2)$$

enjoy numerous properties of interest. First, E^s can be iterated: $(E^s)^2$ points to the neighbors of the neighbors, and $(E^s)^r$ generates neighbors of order r . The dependence among nodes of the network can be measured by the chi-square $\chi^2(E) = \text{Tr}((E^s)^2)$, whose minimum 0 is attained with the *complete network* as described by the *free exchange* $e_{ij} = f_i f_j$, and its maximum $n - 1$ for the *frozen network* $e_{ij} = f_i \delta_{ij}$, where δ_{ij} is Kronecker delta.

The standardised exchange matrix is related to the so-called *normalised Laplacian* \mathcal{L} (e.g. Chung 1997) as $E^s = I - \mathcal{L} - \sqrt{f} \sqrt{f}'$, with essentially the same eigenstructure. In particular, the eigenvalues $\gamma_\alpha \in [-1, 1]$ of E^s are those of W with the exception of the trivial Perron-Frobenius unit value, transformed to $\gamma_0 = 0$, with corresponding eigenvector $t_{i0} = \sqrt{f_i}$. The largest eigenvalue obeys $\gamma_1 \leq 1$, with equality iff the network is *reducible*, that is composed of two or more disconnected components. The smallest eigenvalue obeys $\gamma_{n-1} \geq -1$, with equality iff the network is bipartite. The network is *diffusive* if E is positive semi-definite (p.s.d.), that is if $\gamma_\alpha \geq 0$, and *off-diagonal* if $e_{ii} = 0$. Note that off-diagonal networks cannot be diffusive. Proofs of the above elements, exposed in Bavaud (2010 and 2013) in the present context, can be found in the standard literature on Markov chains.

2.3 Distances, kernels, MDS, covariances and factor scores

Squared Euclidean dissimilarities between observations, as obtained from the $n \times p$ (possibly pre-transformed) data matrix $X = (x_{ik})$, are

$$D_{ij} = \sum_{k=1}^p (x_{ik} - x_{jk})^2 = \|x_i - x_j\|^2 .$$

Scalar products B , and weighted scalar products or *kernels* K are defined as

$$B = -\frac{1}{2} H D H' = X^c (X^c)' \quad K = \Pi^{\frac{1}{2}} B \Pi^{\frac{1}{2}}$$

where $H := I - \mathbf{1} f'$ is the centering matrix, centering the data as $X^c = H X$, and $\Pi = \text{diag}(f)$ is the diagonal matrix of object weights. By construction, K possesses a trivial eigenvalue $\lambda_0 = 0$ corresponding to the eigenvector \sqrt{f} , and non-negative non-trivial eigenvalues $\lambda_\beta \geq 0$ for $\beta \geq 1$, decreasingly ordered.

Weighted MDS consists in extracting, from K , a set of $\min(n - 1, p)$ uncorrelated coordinates $\tilde{X} = (\tilde{x}_{i\beta})$ reproducing the distances as $D_{ij} = \|\tilde{x}_i - \tilde{x}_j\|^2$ and expressing a maximum amount of variance in the first non-trivial spectral dimensions. The solution is

$$\tilde{x}_{i\beta} = \frac{\sqrt{\lambda_\beta}}{\sqrt{f_i}} u_{i\beta} \quad \text{where } K = U \Lambda U', \quad U = (u_{i\beta}) \text{ and } \Lambda = \text{diag}(\lambda_\beta). \quad (3)$$

Weighted covariances are given by $S = (X^c)' \Pi X^c$. By the singular value decomposition, the non-zero eigenvalues of S are identical to those of K , that is $S = V \Lambda V'$, where the *loadings* V serve at obtaining the *factor scores* of principal component analysis as $F = (F_{i\beta}) := X^c V$.

2.4 Local inertia, relative inertia and local covariance

Moran's I (section 2.1) can be generalised to the multivariate setting by defining the *inertia* Δ , the *local inertia* Δ_{loc} and the *relative inertia* $\delta(E, D)$ as

$$\Delta := \frac{1}{2} \sum_{ij} f_i f_j D_{ij} = \text{Tr}(K) = \text{Tr}(S) \quad \Delta_{\text{loc}} := \frac{1}{2} \sum_{ij} e_{ij} D_{ij} \quad \delta := \frac{\Delta - \Delta_{\text{loc}}}{\Delta}$$

Here $\Delta_{\text{loc}} = \text{Tr}(S_{\text{loc}})$, where $S_{\text{loc}} := (X^c)'(\Pi - E)X^c$ is the *local covariance*.

Under the null hypothesis of absence of autocorrelation, the expected value of δ is $\delta_0 = -1/(n-1)$ for off-diagonal networks (e.g. Cliff and Ord 1973), and $\delta_0 = (m-1)/(n-1)$ for m isolated complete sub-networks described by the free partitioned exchange matrices of section 3.1 (e.g. Bavaud 2013). Alternative expressions read (cf. (2) and (3)) :

$$\delta = \frac{\text{Tr}(E^s K)}{\text{Tr}(K)} = \frac{\sum_{\alpha, \beta \geq 1} \gamma_\alpha \lambda_\beta C_{\alpha\beta}}{\sum_{\beta \geq 1} \lambda_\beta} \quad \text{where} \quad C_{\alpha\beta} := \left(\sum_i t_{i\alpha} u_{i\beta} \right)^2 \leq 1 .$$

In particular, $-1 \leq \gamma_{n-1} \leq \delta \leq \gamma_1 \leq 1$, where the maximum of $\delta(E, D)$ for E fixed is attained for $D_{ij} = C(t_{i1}/\sqrt{f_i} - t_{j1}/\sqrt{f_j})^2$ with $C > 0$ (Bavaud 2010).

3 Local formalism as a generalised discriminant analysis

3.1 Free partitioned exchanges and within-groups covariance

An important special case consists of *partitioned complete weighted networks* made of m components, described by *free partitioned exchange matrices*

$$e_{ij} := f_i f_j \sum_{g=1}^m \frac{I(i \in g) I(j \in g)}{\rho_g} \quad \rho_g := \sum_{i \in g} f_i . \quad (4)$$

where $I(i \in g)$ is the 0/1 indicator function of the event ‘‘observation i belongs to group g ’’. In this setup, $\Delta_{\text{loc}} = \Delta_W$, the *within-groups inertia* associated to the hard partitioning of the n observations into m groups (e.g. Le Foll 1982, Meot et al. 1993, Lebart 2005). Similarly, S_{loc} is the *within-groups covariance* S_W . It is thus tempting to define, in the general case, a *between-groups covariance* and a *between-groups inertia* as

$$S_B := S - S_{\text{loc}} = (X^c)' E X^c \quad \Delta_B := \Delta - \Delta_{\text{loc}}$$

with the caveat that, in contrast to S or S_{loc} , S_B is not p.s.d., unless E is diffusive. This is indeed the case for reducible exchanges (4), obeying $(E^s)^2 = E^s$, thus making $\gamma_\alpha = 1$ or $\gamma_\alpha = 0$.

Hence the relative autocorrelation reads as $\delta = \Delta_B/\Delta$, behaving as a kind of F -ratio for any general neighborhood structure E - reducible or not, diffusive or not. Also, *hard partitioning* and *discriminant analysis* appear as particularly cases of the local formalism.

3.2 Factor analysis and canonical scores for unoriented networks

Fisher *canonical scores* consist of mutually orthogonal *linear combinations* of scores maximizing the relative between-groups dispersion Δ_B/Δ (or, equivalently, maximising δ under the linear requirement; compare with section 2.4 for the unconstrained maximising configuration). They can be shown to be $F^{\text{can}} := X^c U^{\text{can}}$, where U^{can} contains eigenvectors of $S^{-1} S_B u_\sigma^{\text{can}} = \lambda_\sigma^{\text{can}} u_\sigma^{\text{can}}$, yet to be normalised (e.g. Mardia et al. 1979; Flury 1997; Saporta 2006). Canonical scores are related to the normalised eigenvectors of the symmetric matrix $A := S^{-\frac{1}{2}} S_B S^{-\frac{1}{2}} = \dot{U} A^{\text{can}} \dot{U}'$ as $U^{\text{can}} = S^{-\frac{1}{2}} \dot{U} \Xi$, where the freely adjustable matrix $\Xi = \text{diag}(\xi)$ is diagonal and fixes the normalisation: canonical scores are centered, with covariance

$$S^{\text{can}} = (F^{\text{can}})' \Pi F^{\text{can}} = \Xi^2 = (\delta_{\sigma\tau} \xi_\sigma^2) . \quad (5)$$

Canonical dissimilarities turn out to be of the form

$$D_{ij}^{\text{can}} = \sum_{\sigma} (F_{i\sigma}^{\text{can}} - F_{j\sigma}^{\text{can}})^2 = \sum_{kl} m_{kl} (x_{ik} - x_{jk})(x_{il} - x_{jl}) \quad (6)$$

where $M = (m_{kl}) = S^{-\frac{1}{2}} \dot{U} \Xi^2 \dot{U}' S^{-\frac{1}{2}}$. Equivalently, canonical scores F^{can} are given by MDS coordinates resulting of the spectral decomposition of the kernel $K^{\text{can}} = \Pi^{\frac{1}{2}} X^c M (X^c)' \Pi^{\frac{1}{2}}$.

Canonical scores and canonical dissimilarities are, ultimately, fully specified by requiring a normalisation relation of the form $\xi_\sigma^2 = h(\lambda_\sigma^{\text{can}})$, making the canonical and local canonical covariances diagonal:

$$S^{\text{can}} = h(A^{\text{can}}) \quad S_{\text{loc}}^{\text{can}} := (F^{\text{can}})' (\Pi - E) F^{\text{can}} = h(A^{\text{can}}) (I - A^{\text{can}}) .$$

Then

$$M = S^{-\frac{1}{2}} \dot{U} h(A^{\text{can}}) \dot{U}' S^{-\frac{1}{2}} = S^{-\frac{1}{2}} h(A) S^{-\frac{1}{2}} = S^{-\frac{1}{2}} h(S^{-\frac{1}{2}} S_B S^{-\frac{1}{2}}) S^{-\frac{1}{2}} .$$

The choice $h(\lambda) \equiv 1$ yields $M = S^{-1}$, that is canonical dissimilarities (6) identical to Mahalanobis distances. The choice $h(\lambda) = 1/(1 - \lambda)$, *which we shall adopt in this paper*, turns out to produce the *local metric* $M = S_{\text{loc}}^{-1}$. Note that $M = S^{-1} S_B S^{-1}$ for $h(\lambda) = \lambda$, a legitimate choice provided E is diffusive (section 3.1), insuring that M is p.s.d. and D^{can} in (6) is Euclidean.

3.3 Similarity between two weighted configurations

For future use, let us propose two similarity indices between two weighted configurations (f_A, D_A) and (f_B, D_B) on n observations, with associated kernels K_A and K_B , namely

$$\begin{aligned} \text{configuration similarity 1} \quad \text{CS1}_{AB} &:= \frac{\text{Tr}(K_A K_B)}{\sqrt{\text{Tr}(K_A^2) \text{Tr}(K_B^2)}} \\ \text{configuration similarity 2} \quad \text{CS2}_{AB} &:= \frac{\text{Tr}(K_A^{\frac{1}{2}} K_B^{\frac{1}{2}})}{\sqrt{\text{Tr}(K_A) \text{Tr}(K_B)}} = \frac{\text{Tr}(K_A^{\frac{1}{2}} K_B^{\frac{1}{2}})}{\sqrt{\Delta_A \Delta_B}} . \end{aligned}$$

The similarity indices constitute a simple alternative to Procrustean analysis, as well as to the *RV-coefficient* (Robert and Escoufier 1976) or to the *distance correlation coefficient* (Székely et al. 2007). By construction, $0 \leq CS1_{AB} \leq CS2_{AB} \leq 1$. Also, $\tilde{D}_{AB} := 1 - CS1_{AB}$ and $\hat{D}_{AB} := 1 - CS2_{AB}$ constitute squared Euclidean distances between weighted configurations.

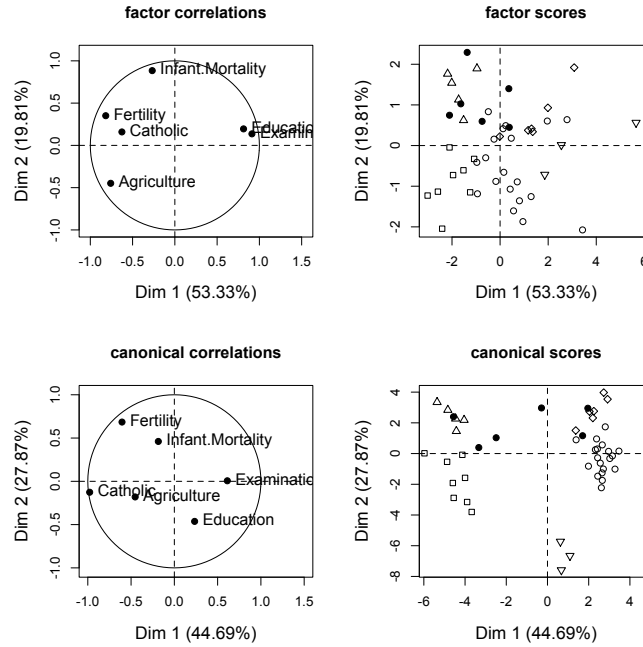


Fig. 1. Factor (top) and canonical (bottom) correlations circle and spectral scores (section 4.1). Plotting symbols refer to cantons: $\triangle = FR$, $\nabla = GE$, $\circ = VD$, $\square = VS$, $\diamond = NE$, $\bullet = JU$. Groups are fairly homogeneous ($\delta = 0.63$, to be compared to $\delta_0 = (6-1)/(47-1) = 0.11$), resulting in high similarities ($CS1 = 0.72$ and $CS2 = 0.89$). Canonical configuration further increases the groups homogeneity, as it must.

4 Illustrations and cases studies

Let us illustrate the theory on four datasets, each endowed with a local structure (E, D) . After computing δ , the factor scores F and canonical scores F^{can} , together with their correlations with the original features X (correlations circle), are determined and plotted. Also, the similarity coefficients $CS1(X, F^{can})=CS1(F, F^{can})$ and $CS2(X, F^{can})=CS2(F, F^{can})$ are computed.

4.1 Swiss 19th century socio-economic data

$p = 6$ socio-economic standardised variables are measured on $n = 47$ districts, partitioned into $m = 6$ cantons (source: `swiss{datasets}` in R). Here E is the free partitioned exchange matrix (section 3.1), and the setup amounts to ordinary discriminant analysis (Figure 1).

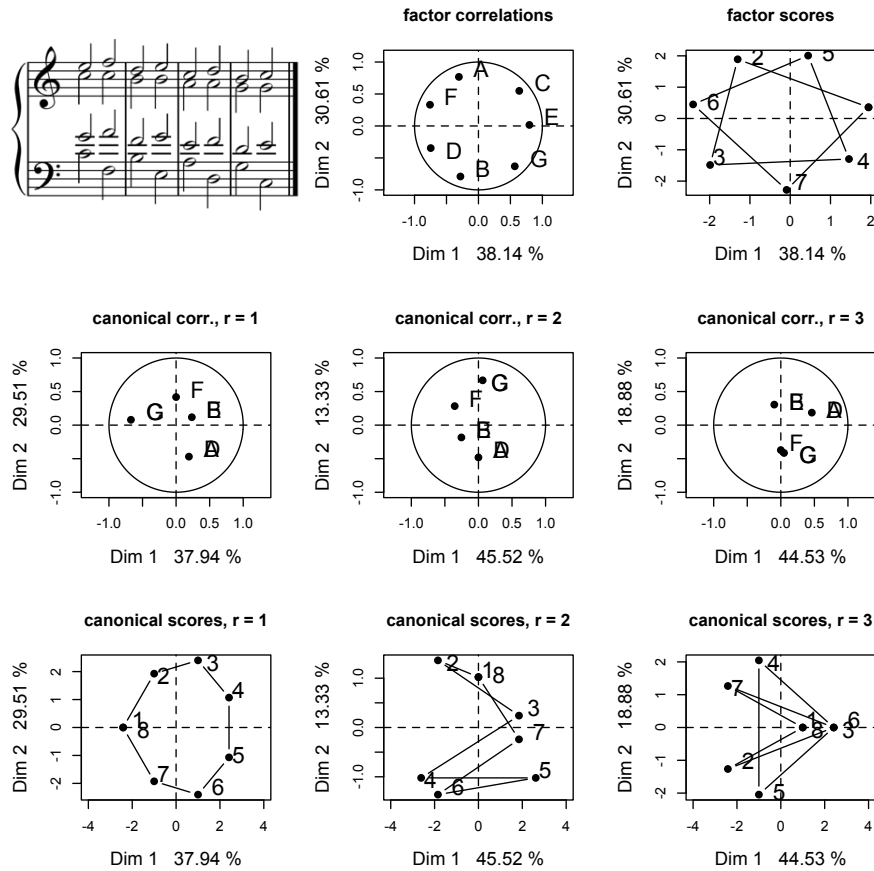


Fig. 2. Factor and canonical analysis of the musical piece, with periodic neighbourhoods of order $r = 1, 2, 3$. For $r = 1$: $\delta = 0.03$, $CS1 = 0.61$, $CS2 = 0.87$. For $r = 2$: $\delta = -0.51$, $CS1 = 0.40$, $CS2 = 0.78$. For $r = 3$: $\delta = -0.12$, $CS1 = 0.54$, $CS2 = 0.84$.

4.2 Correspondence analysis of musical scores

A simple musical illustration is provided by the *circle of fifths progression*, split into $n = 8$ time intervals of one half-note each (Figure 2, top left). First, one constructs the 8×7 contingency table $\Psi = (\psi_{ik})$ counting the relative duration

of note k ($k = C, D, E, F, G, A, B$) in each interval i , with $\sum_k \psi_{ik} = 1$, on which *chi-square dissimilarities* D_{ij} between the rows of Ψ (that is between time intervals) are computed. Secondly, one considers *periodic neighbourhoods of order* $r \geq 1$, in which each column and each row of $E^{[r]}$ have two non-zero entries $1/(2n)$ at distance r (modulo n) from the main diagonal. Note that $f_i = 1/n$ and $E^{[2r]} \neq (E^{[r]})^2$, that is periodic neighbourhoods of order r do not form an iterated family (see section 2.1).

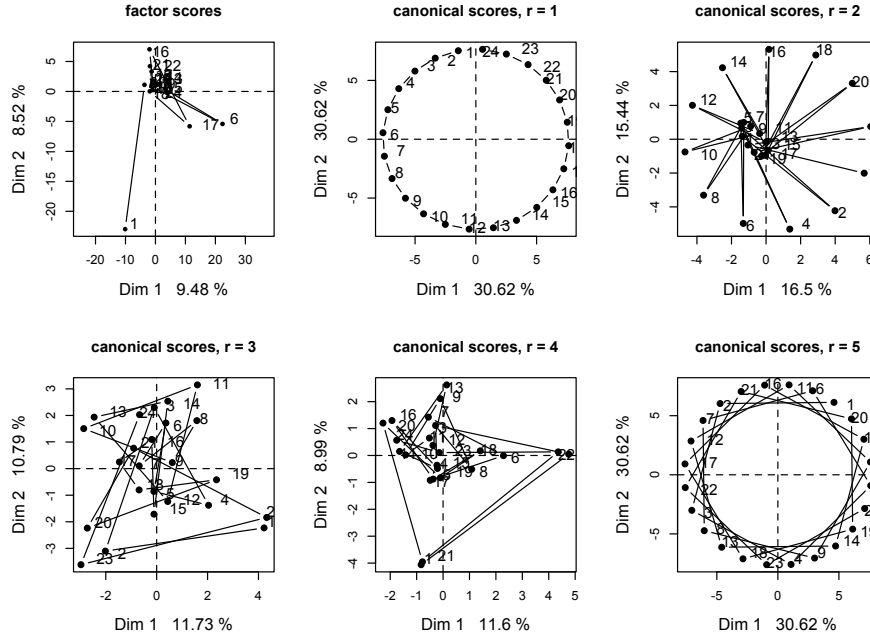


Fig. 3. Factor and canonical scores for the textual example of section 4.2. Auto-correlation is absent ($\delta^{(r)} \cong \delta_0 = -1/(n - 1)$), yet similarities between factorial and canonical configurations remain fairly high, despite appearances: CS1 $\sim .45$ and CS2 $\sim .75$ for $r = 1, 5$; CS1 = .61 and CS2 = .84 for $r = 2$; CS1 $\sim .745$ and CS2 $\sim .88$ for $r = 3, 4$.

4.3 Local CA for text-document matrix

The play *Sganarelle ou le Cocu imaginaire* (Molière 1660) contains $n = 24$ short scenes (the observations) including $p = 339$ verbs. Analysing the scenes-verbs matrix yields fairly constant chi-square distances D^x between scenes, in the range $[9.9, 25.7]$, with a non-significant $\chi^2[\text{df} = 7774] = 195.9$, impeding the emergence of a textual structure by low-dimensional CA compression.

Scene profiles x_i are chosen as standardised coordinates resulting from uniform MDS on D^x . As in the previous section, exchange matrices are periodic

neighbourhoods of order r , resulting in uniform weights $f_i = 1/n$. Autocorrelation is absent, with a δ close, for various r , to its expected value under independence $\delta_0 = -1/23 = -0.043$. Canonical scores, minimizing the relative local dispersion, exhibit unexpected patterns (Figure 3), yet to be elucidated.

4.4 Distances between world cities

$n = 313$ world cities over 10^6 inhabitants, with latitudes θ_i , longitudes α_i and relative weights f_i (proportional to the population size) have been extracted from the R file `world.cities{maps}`.

Geodesic or *arc-length* dissimilarities $D_{ij} = \arccos^2(\kappa_{ij})$, where $\kappa_{ij} = \sin \theta_i \sin \theta_j + \cos \theta_i \cos \theta_j \cos(\alpha_i - \alpha_j)$, are squared Euclidean. They may serve at constructing “gravity-like” exchange matrices of the form $e_{ij} = C b_i b_j \exp(-\beta D_{ij})$ for $i \neq j$, where b is the *eigenvector centrality*, that is the dominant eigenvector of $\exp(-\beta D_{ij})$, and e_{ii} is adjusted so that $\sum_j e_{ij} = f_i$ (see Bavaud 2013 for details). Figure 4 shows the results obtained with $\beta = 1$.

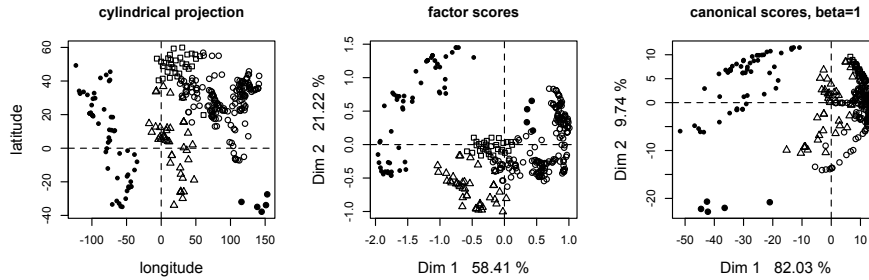


Fig. 4. World cities scatterplots (section 4.4), where plotting symbols refer to the continents : \triangle = Africa, \bullet = Oceania, \circ = Asia, \square = Europa, \blacklozenge = America. The central figure results from weighted MDS. The depicted factor scores, whose two first dimensions account for 79.63% of the inertia, are highly similar to the canonical scores (91.77% of the inertia): $\delta = 0.98$, $CS1 = 0.73$ and $CS2 = 0.96$.

5 Conclusion

Exchange matrices specify weighted unoriented networks. This contribution underlines their role as *soft categorical variables* associated to the neighborhood relation between observations, permitting to generalise linear discriminant analysis and hard partitioning to any local setup.

In general, the higher the relative autocorrelation is, the more similar are the factor and canonical configurations accordingly to $CS1$ and $CS2$ - but not the other way round (see Figure 3).

Further studies could develop the concepts of *cross-autocorrelation* between two distinct data sets, and of *partial autocorrelation* beyond the time-series framework. *Local clustering*, favouring the grouping of observations with small dissimilarities *and* strong exchanges, is to be further investigated.

References

- BAVAUD, F. (2008): Local concentrations. *Papers in Regional Science* 87, 357–370.
- BAVAUD, F. (2010): Euclidean Distances, Soft and Spectral Clustering on Weighted Graphs. In: *Proceedings of the ECML PKDD10*. Lecture Notes in Computer Science 6321, 103–118
- BAVAUD, F. (2013): Testing Spatial Autocorrelation in Weighted Networks: The Modes Permutation Test. *Journal of Geographical Systems* 15, 233–247.
- BERGER, J. and SNELL, J.L. (1957): On the concept of equal exchange. *Behavioral Science* 2, 111–118.
- CHUNG, F.R.K. (1997): *Spectral graph theory*. CBMS Regional Conference Series in Mathematics 92. American Mathematical Society, Washington.
- CLIFF, A.D. and ORD, J.K., (1973): *Spatial autocorrelation*. Pion, London.
- FLURY, B. (1997): *A First Course in Multivariate Statistics*. Springer, New York.
- GEARY, R. (1954): The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician* 5, 115–145.
- LEBART, L. (1969): Analyse Statistique de la Contiguïté. *Publications de l'Institut de Statistique des Universités de Paris, XVIII*, 81–112.
- LEBART, L. (2005): Contiguity Analysis and Classification. In : W. Gaul, O. Opitz and M. Schader M. (Eds.): *Data Analysis*. Springer, Berlin, 233–244.
- LE FOLL, Y. (1982): Pondération des distances en analyse factorielle. *Statistique et Analyse des Données*, 7, 13–31.
- MARDIA, K. V., KENT, J.T. and BIBBY, J.M. (1979): *Multivariate analysis*. Academic Press, London.
- MEOT A., CHESSEL D. and SABATIER R. (1993): Opérateurs de voisinage et analyse des données spatio-temporelles. In : D. Lebreton and B. Asselain (Eds.): *Biométrie et Environnement*. Masson, Paris, 45–71.
- MORAN, P.A.P. (1950): Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23.
- ROBERT, P. and ESCOUFIER, Y. (1976): A Unifying Tool for Linear Multivariate Statistical Methods: The *RV*-Coefficient. *Applied Statistics* 25, 257–265.
- SAPORTA, G. (2006): *Probabilités, Analyse des Données et Statistique*. Technip, Paris.
- SZÉKELY, G. J., RIZZO, M. L. and BAKIROV, N. K. (2007): Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics* 35, 2769–2794.
- THIOULOUSE, J., CHESSEL, D. and CHAMPELY, S. (1995): Multivariate analysis of spatial patterns - a unified approach to local and global structures. *Environmental and Ecological Statistics* 2, 1–14.