

Bio 6077

Exercice, chapitre 9

I. Composantes et coordonnées principales

Mars 1998, 2002, 2004, 2005, 2006, 2008, 2011, 2012, 2018, 2019

L'objectif de cet exercice est de comparer les résultats d'une analyse en composantes principales à ceux d'une analyse en coordonnées principales.

Les fonctions R à utiliser sont:

- Analyse en composantes principales: votre propre fonction d'ACP¹.
- Mesure de distance euclidienne: fonction *dist()* de {stats}.
- Analyse des coordonnées principales: fonction *pcoa()* de {ape} ou *cmdscale()* de {stats}.

1. Choisissez un petit tableau de données contenant environ 10 objets et 5 variables quantitatives ou binaires. Décrivez brièvement ce tableau: nature et nombre d'objets et de variables, type mathématique des variables (quantitatives ou binaires). Attention: pouvez-vous traiter directement ces données? Faut-il d'abord centrer et réduire les variables?

2. Calculez une **analyse en composantes principales** basée sur la matrice des covariances; les données auront été centrées et réduites au no 1 si nécessaire. Faites rapport des covariances, des valeurs propres et de la position des objets (matrice **F**); conservez la matrice **F** dans un nouvel objet R en vue de l'étape 4. Tracez l'ordination des objets dans le plan des deux premières composantes principales. Quelle est la somme de toutes les valeurs propres que vous avez trouvées?

3. Utilisant le tableau de données sélectionné au no 1 (données brutes ou centrées-réduites) et utilisé pour l'ACP au no 2, calculez la distance euclidienne entre les objets (lignes) de ce tableau.

4. Utilisant *toutes les colonnes* de la matrice **F**, calculez maintenant la distance euclidienne entre les objets (lignes) de cette matrice.

5. Réalisez une **analyse en coordonnées principales** de la matrice de distances calculée au point 3. Faites rapport des valeurs propres et de la position des objets sur toutes les coordonnées principales. Tracez l'ordination des objets dans le plan des deux premières coordonnées principales. Quelle est la somme de toutes les valeurs propres que vous avez trouvées?

6. Réalisez une **analyse en coordonnées principales** de la matrice de distances calculée au point 4. Faites rapport des valeurs propres et de la position des objets sur toutes les coordonnées principales. Somme des valeurs propres?

¹ Avez-vous validé votre fonction R en comparant ses résultats à ceux de la fonction *rda()* de {vegan} ou de la fonction *PCA.newr()* du livre *Numerical ecology with R* (2018) ?

7. Comparez les valeurs propres trouvées aux points **2**, **5** et **6**. Sont-elles identiques? Si elles diffèrent, pourquoi en est-il ainsi?
8. Comparez les ordinations trouvées aux points **2**, **5** et **6**. Certains axes sont-ils inversés? Si oui, qu'est-ce que cela signifie?

II. Analyse des correspondances

- 9.1. Choisissez un petit tableau de données contenant environ 10 objets et 3-5 variables. Les données doivent être des *fréquences*, par exemple des abondances d'espèces ou des fréquences d'allèles. Décrivez brièvement ce tableau: nature et nombre d'objets et de variables.
- 9.2. Calculer une analyse des correspondances (AFC). Choisissez la méthode de cadrage à utiliser (cadrage = 1 ou 2); donnez les raisons de votre choix. Vous pourrez utiliser l'une ou l'autre des fonctions R suivantes:
- Fonction `cca()` de `{vegan}`. Un exemple est présenté à la page 9 du document "Travaux_pratiques_en_R.pdf".
 - Fonction `CA.newr.R` dans la section "Fonctions en langage R", paragraphe "Ordinations simples", sur la page <http://numericecology.com>. Aussi disponible dans le dossier *NEwR2-Functions* du livre *Numerical ecology with R* (2018).
- 9.3. Présentez les résultats obtenus: les valeurs propres, la position des objets et la position des espèces dans le diagramme de double projection (*biplot*) des axes I et II.

10. Réalisez l'AFC décrite à la diapo #36 du diaporama sur l'analyse des correspondances. Vous devrez d'abord générer les données nécessaires (coenocline) à l'aide des commandes de la diapo #35. Pour représenter uniquement les espèces dans l'ordination, vous devrez utiliser l'argument `display="sp"` :

```
plot(cca.res, display="sp")
```

Commentez brièvement l'ordination des espèces que vous avez obtenue.