



Ordination en espace réduit

1.0. Introduction

Pierre Legendre

Département de sciences biologiques

Université de Montréal

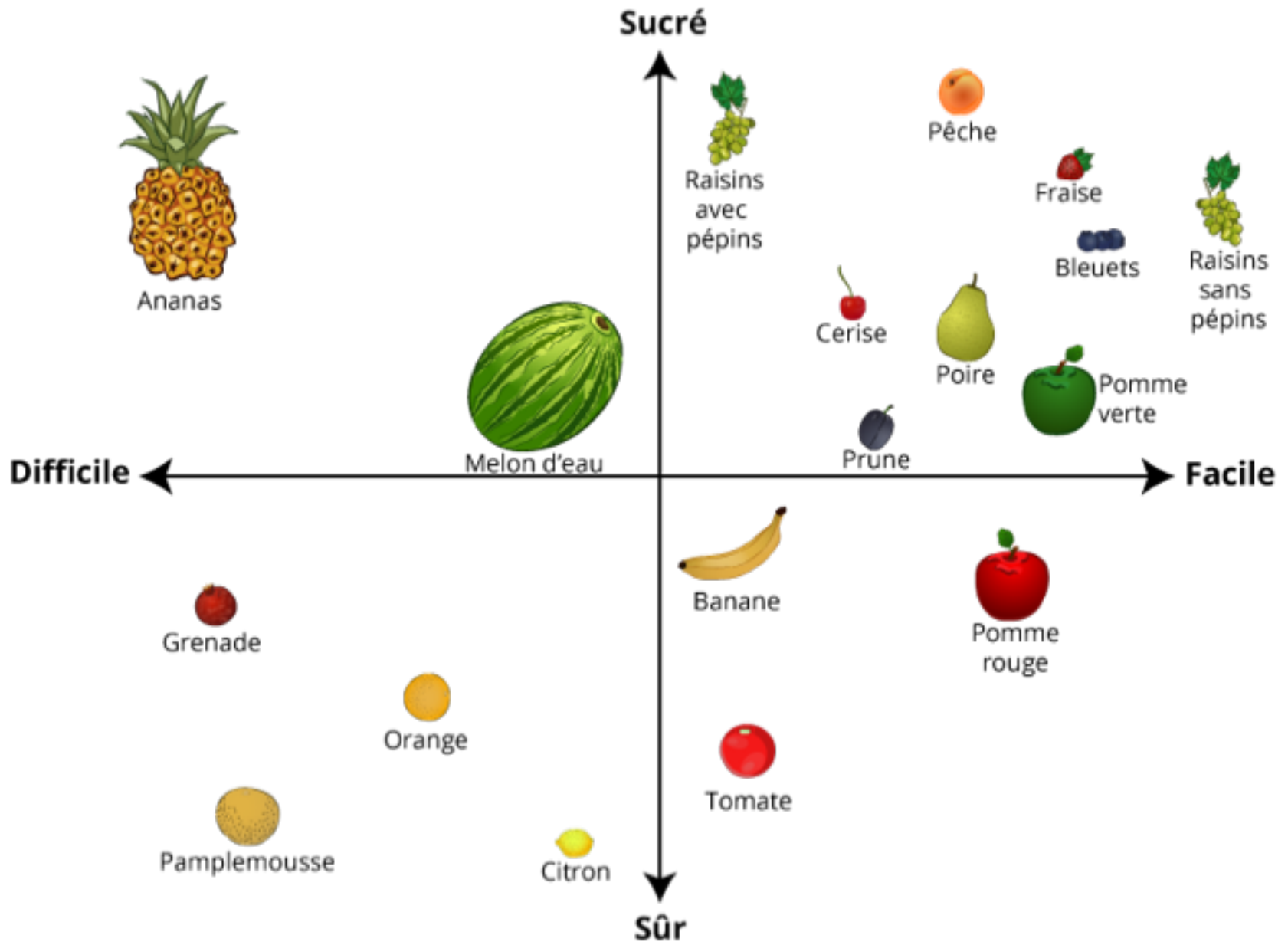
<http://www.NumericalEcology.com/>



Objectif de l'ordination

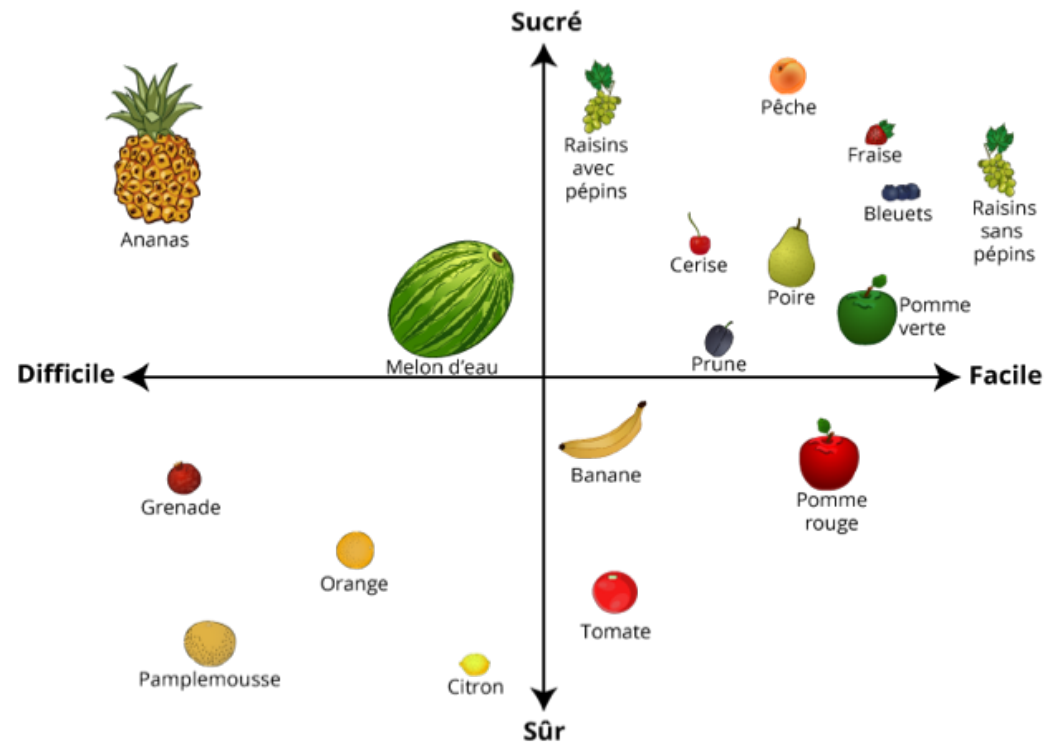
L'objectif principal de l'ordination est de représenter les objets (les sites en écologie) dans un *diagramme de dispersion* multivariable.

Voici un exemple de diagramme de dispersion en 2 dimensions:



Ordination en espace réduit

- Les axes permettent de disperser les fruits en fonction de critères identifiables: facile à peler ou non, goût sucré ou sûr.
- Les fruits semblables sont proches les uns des autres; leurs distances sont petites.
- Les fruits très différents sont loin les uns des autres; leurs distances sont grandes.
- Le graphique est une bonne représentation [ou *modèle*] **des distances** entre les fruits, considérant les variables formant les axes.



Ordination (du latin *ordinatio*, l'action de mettre en ordre) est l'arrangement d'éléments dans un certain ordre.

Le terme ordination, largement utilisé en statistique multivariable, vient de l'écologie où il réfère à la représentation d'objets (sites, stations d'échantillonnage, relevés) sous la forme de points le long d'axes de référence.

David W. Goodall

➤ En 1954, David Goodall, spécialiste de l'écologie végétale, fut le premier à utiliser l'analyse factorielle en écologie des communautés. Goodall proposa le terme "ordination" pour désigner ce type d'analyse, un terme qui est maintenant largement utilisé dans les publications en écologie et dans d'autres domaines.

David Goodall est né à Londres le 4 avril 1914. Cette photo le montre à son bureau à l'université Edith Cowan en Australie de l'Ouest en juin 2009. Il a eu 103 ans en 2017. David Goodall est le chercheur actif le plus âgé d'Australie.

Web page: https://en.wikipedia.org/wiki/David_W._Goodall



Faites tourner le code



'Spiders_28x12_spe.txt' est l'un des fichiers de données de ce cours

```
spiders = read.table(file.choose())
# Transformation de Hellinger des abondances d'araignées
spiders.hel = decostand(spiders,"hellinger")

# ACP avec la fonction prcomp() de {stats}
pca.spiders = prcomp(spiders.hel)
spiders.sites = summary(pca.spiders)$x[,1:3]

library(rgl)
rgl.open()
rgl.points(spiders.sites,color="green",size=6)

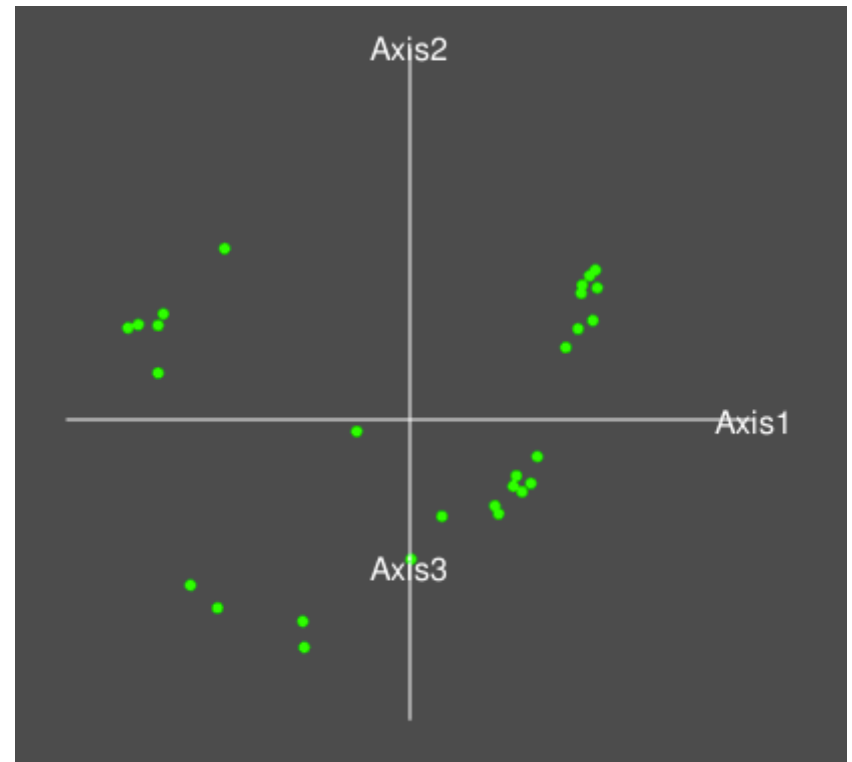
rgl.lines(x=c(-1,1),y=c(0,0),z=c(0,0))
rgl.lines(y=c(-1,1),x=c(0,0),z=c(0,0))
rgl.lines(z=c(-1,1),x=c(0,0),y=c(0,0))

rgl.texts(1,0,0,"Axis1")
rgl.texts(0,1,0,"Axis2")
rgl.texts(0,0,1,"Axis3")
```



Faites tourner le graphique 3-D en manipulant les axes

- Y a-t-il une position où les points ont le maximum de dispersion ?
- Quel est le terme statistique pour “dispersion” ?



```
# Quelle est la variance des points le long des 3 axes?  
apply(spiders.sites,2,var)
```

```
# Quelles sont les premières valeurs propres de l'ACP ?  
summary(pca.spiders)$sdev[1:3]^2
```


Les données écologiques sont multivariées, ce qui pose des difficultés pour leur analyse. Chaque variable représente une dimension d'une ordination qui aura de nombreuses dimensions.

En écologie, on observe habituellement de nombreuses variables à chaque site. – Par exemple, on peut y observer des centaines d'espèces ainsi que de nombreuses variables environnementales.

Dans la plupart des cas, les écologistes veulent décrire les tendances principales de la variation des objets [sites] à travers tous les descripteurs, pas seulement quelques-uns.

Dans le chapitre sur l'ordination, je décrirai les méthodes les plus importantes qui nous permettent d'identifier l'espace qui résume au mieux les données, habituellement en 2 ou 3 dimensions.



Quatre méthodes d'ordination

Méthode	Dissimilarité préservée	Types de variables
Analyse en composantes principales (ACP, <i>PCA</i>)	Distance euclidienne	Variables quantitatives en relations linéaires, dimensionnellement homogènes ou centrées-réduites
Analyse factorielle des correspondances (AFC, <i>CA</i>)	Distance de khi-carré	Données de type fréquence ou présence-absence, dimensionnellement homogènes, non négatives
Analyse en coordonnées principales (ACoP, <i>PCoA</i>) ; positionnement multidimensionnel classique (ou métrique)	Toute mesure de dissimilarité	Variables quantitatives, semi-quantitatives, qualitatives ou mixtes
Positionnement multidimensionnel non-métrique (nMDS)	Toute mesure de dissimilarité	Variables quantitatives, semi-quantitatives, qualitatives ou mixtes



Fin de la section