

# TP: Positionnement multidimensionnel

## Résumé

Exemples élémentaires de *positionnement multidimensionnel* ou ACP d'un tableau de distances avec SAS et R.

## 1 Introduction

Cette séance a pour but d'expérimenter les techniques s'appliquant à un tableau de distances ou dissimilarités : ACP d'un tableau de distance ou MDS sur un exemple simple comparant des distances entre villes. Dans ce cas, le MDS permet une représentation des variables/gènes en fonction de similarités basés sur leurs corrélations. Le MDS sera revu en association avec la classification. Le [scénario](#) sur l'analyse élémentaire de données transcriptomiques fournit un exemple plus complet d'utilisation du MDS.

## 2 MDS avec SAS

Les données se présentent sous la forme du triangle inférieur d'une matrice symétrique, par construction, et contenant les distances kilométriques de villes françaises prises 2 à 2 (Source : carte IGN).

Charger le fichier `mdsville2.dat` puis lecture de la matrice triangulaire inférieure en précisant qu'il s'agit d'une matrice de distance :

```
data sasuser.mdsville (type=distance) ;
infile "mdsville2.dat" missover lrecl=200;
input ville $ 4. (d1-d47) (4.); run;
```

Calcul de l'ACP du tableau des distances et représentation graphique.

```
proc mds data=sasuser.mdsville dim=2 out=resul
level=abs;
var d1-d47;
object ville; run;
%gafcx(ident=ville);
```

Commenter la représentation par rapport à la carte "géographique".

## 3 MDS avec R

### 3.1 Lecture des données contenant les distances des villes

La lecture d'une matrice triangulaire inférieure en tant que matrice de distance pose quelques difficultés dans R dans la gestion du type des objets.

Lire le fichier `mdsville.dat`.

```
mdsville=read.table("mdsville.dat", fill=TRUE)
# extraction des noms des villes
villes=as.character(mdsville[2:48,1])
# extraction des valeurs des distances
m=mdsville[2:48,2:48]
# transformation en une matrice alphanumérique
m=as.matrix(m)
# retour au numérique avec données manquantes
m=as.numeric(m)
# reformattage en une matrice
m=matrix(m,47,47)
# adjonction des noms des villes en ligne et colonne
dimnames(m)[[1]]=villes
dimnames(m)[[2]]=villes
# transformation en un objet de type distance
d=as.dist(m,diag=TRUE)
d
```

### 3.2 MDS des villes avec R

En revanche, le MDS est immédiat. On remarque que le choix de la dimension, ici  $k = 2$  est à traiter comme en ACP. Une visite du manuel permet de voir comment calculer toutes les valeurs singulières afin d'aider ce choix par leur représentation.

```
mds = cmdscale(d, k=2)
plot(mds, type="n", xlab="cp1", ylab="cp2")
```

text (mds, villes)