

I. Positionnement Multidimensionnel: Multidimensional Scaling

► Téléchargez les données

Les données se trouvent à l'adresse:

<http://www-stat.stanford.edu/~susan/summer09/eckman.txt>.

Il faut lire les données dans R: Pour cette session vous aurez besoin des fonctions de la librairie `ade4`, et éventuellement `ellipse`.

```
>eck=read.table("eckman.txt",header=TRUE)
>eck
```

`eck` est la matrice des confusion de couleurs les étiquettes sont les longueurs d'ondes.

Exercice 1 Trouver la fonction `myImagePlot` sur l'internet et le sourcer avec R vous devez avoir une fonction qui s'appelle: `myImagePlot`.

```
>myImagePlot(eck)
```

Que se passe-t-il quand vous faites:

```
>myImagePlot(1-eck)
```

Quelle est la meilleure mesure de dissimilarité: `eck` ou `1-eck` ou `1- exp(-eck)`?

► Fonction de base multidimensionnelle R : `cmdscale`

Exercice 2 Quel est la sortie par défaut de `cmdscale`?

Changez ces sorties par défaut, par exemple demander la sortie des 3 valeurs propres.

```
>eck.cmds=cmdscale(1-eck,k=3,eig=TRUE)
```

Faites un graphique du premier plan.

► Download the package `ade4` from the cran website. Load the library `ade4`. What happens when you type:

Exercice 3

```
>eck.pco=dudi.pco(as.dist(1-eck))
```

You can fix this by typing:

```
>queck=quasieuclid(as.dist(1-eck))
>eck.pco=dudi.pco(queck)
>scatter(eck.pco)
```

Look at the function `scatter` in two ways

```
>scatter
>scatter.pco
```

II. Principal Component Analysis

► Student Scores Data

Exercise 4

```
> scor88=read.table('scor88.txt',header=T)
> dim(scor88)
```

Calculer la moyenne et la mediane pour les 5 colonnes:

```
>apply(scor88,2,median)
>apply(scor88,2,mean)
>apply(scor88,2,sd)
```

► Regarder les correlations

Exercise 5 Vous pouvez utiliser `pairs` ou `plotcorr` dans le paquet `ellipse`.

► PCA avec princomp

Exercise 6 Faites un simple acp avec la fonction `princomp`, regarder qu'est-ce qui sort dans la liste de sortie de la fonction:

```
> scor.pc=princomp(scor88)
> biplot(scor.pc)
> scor.pc
> names(scor.pc)
```

Comment faire un eboulis des valeurs propres?

```
> barplot(scor.pc$sdev^2)
> barchart(scor.pc$sdev^2)
> sum(scor.pc$sdev^2)
```

► ACP avec ade4

Exercise 7

```
> scor.pca=dudi.pca(scor88)
> scatter(scor.pca)
> sum(scor.pca$eig)
> sum(diag(cor88))
```

Comment expliquer les deux derniers nombres?

Faites le cercle des correlations.

```
> s.corcircle(scor.pca$co)
```

Faites le graphiques des points observations.

```
> s.label(scov.pca$li,box=F)
```

Exercise 8 Comparer avec les sorties de positionnement multidimensionnel.

```
> plot(cmdscale(dist(scov88)),type='n')  
> text(cmdscale(dist(scov88)),labels=1:88)
```

III. Analyse des Correspondances

► Entrer les données du tableau de contingence

Exercise 9 Suivez les instructions ci-dessous:

```
> eyes = matrix( c(68, 119, 26, 7, 20, 84, 17, 94,  
15, 54, 14, 10, 5, 29, 14, 16 ),byrow=T,ncol=4)  
> eyesc=c('Black','Brunette','Red','Blonde')  
> eyesn=c('brown','blue','hazel','green')  
> dimnames(eyes)=list(eyesn,eyesc)  
> chisq.test(eyes)  
> outchisq=chisq.test(eyes)  
> names(outchisq)
```

Que vous dit la sortie?

► Faites une analyse multidimensionnelle du tableau

Exercise 10 Utilisez les fonctions `dudi.coa`, `scatter`, `s.label` pour faire une analyse des correspondances et faire les graphiques des sorties.

Faites un e-bouillis des valeurs propres. Mettez la sortie dans un fichier pdf.

► Decomposing Information

Exercise 11 Calculer:

```
> sum(eyes)*sum(eyes.coa$eig)
```

Qu'est-ce que ce nombre?