

Initiation à l'analyse en composantes principales

Jean R. Lobry & Anne B. Dufour

Une première approche très intuitive et interactive de l'ACP avec des données réelles, voire historiques, de températures cardinales de croissance de micro-organismes (T_{\min} , T_{opt} et T_{\max}). Ayez confiance, nous ne vous jetterons pas dans le grand bassin multidimensionnel de l'ACP sans bouées.

Table des matières

1	ACP intuitive à la main (pataugeoire)	3
1.1	Note liminaire	3
1.2	De la nature des données	3
1.3	De l'imposition d'une échelle commune	6
1.4	De la visualisation des données en 3D	7
1.5	De la forme idéalisée du nuage de points	8
1.6	Des symétries des solutions (important)	9
2	ACP centrée-réduite avec ade4 (petit bassin)	10
2.1	De la mise en œuvre	10
2.2	De la projection des individus	12
2.3	De la projection des variables	13
2.4	Des limitations intrinsèques	15
3	Dans le grand bassin des dimensions élevées (exercice)	15
4	Pour aller plus loin	17
	Références	17

code	organisme	T_{\min}	T_{opt}	T_{\max}
Ab1	<i>Achromobacter</i> sp.483	-10.0	26.2	31.5
Ab2	<i>Achromobacter</i> sp.5	-7.8	24.6	30.0
Ab3	<i>Achromobacter</i> sp.7	-10.5	25.3	32.4
Ac1	<i>Acinetobacter</i> sp.2.55	1.4	28.4	35.1
Ac2	<i>Acinetobacter</i> sp.4.41	0.1	30.2	37.3
Aer	<i>Aeromonas</i> sp.4.29	5.1	35.4	44.0
Alt	<i>Alteromonas</i> sp. CLD38	-5.2	25.5	33.4
Bco	<i>Bacillus coagulans</i>	12.1	53.8	62.9
Bme	<i>Bacillus megaterium</i>	19.8	50.3	59.2
Bs1	<i>Bacillus stearothermophilus</i> 238	30.8	57.2	65.5
Bs2	<i>Bacillus stearothermophilus</i>	33.6	64.8	71.8
Bsu	<i>Bacillus subtilis</i>	13.4	38.7	51.0
Can	<i>Candida</i> Y1	-9.7	21.8	26.8
CbA	<i>Clostridium botulinum</i> type A	11.0	39.3	45.8
CbB	<i>Clostridium botulinum</i> type B	13.0	37.9	46.1
Cel	<i>Cellulomonas</i> sp.	-2.9	25.4	29.0
Cj1	<i>Cytophaga johnsonae</i>	-11.3	24.9	32.0
Cj2	<i>Cytophaga johnsonae</i>	-7.5	27.8	33.1
Ec1	<i>Escherichia coli</i>	5.6	40.3	47.3
Ec2	<i>Escherichia coli</i> K-12	11.2	41.0	48.0
Ec3	<i>Escherichia coli</i>	4.9	41.3	47.5
Fla	<i>Flavobacterium</i> sp. 2.4	-2.8	29.3	35.5
Geo	<i>Geotrichoides</i> sp. Y9	-9.0	20.1	26.4
Gf1	<i>Gibberella fujikuroi</i>	2.3	30.8	39.7
Gf2	<i>Gibberella fujikuroi</i>	1.5	30.9	39.8
Mc1	<i>Micrococcus cryophilus</i> ATCC15174	-7.1	21.5	25.5
Mc2	<i>Micrococcus cryophilus</i> M19	0.9	30.9	37.9
Mc3	<i>Micrococcus cryophilus</i> T8	0.1	28.3	35.2
Mor	<i>Moraxella</i> sp. 4.16	-0.9	30.5	37.2
Myc	<i>Mycotorula</i> sp. Y15	-5.7	28.4	30.7
Pa1	<i>Pseudomonas aeruginosa</i>	5.0	38.1	46.2
Pa2	<i>Pseudomonas aeruginosa</i> NRRL B23	4.4	40.3	46.0
Pfl	<i>Pseudomonas fluorescens</i>	9.8	38.6	45.4
PI1	<i>Pseudomonas</i> sp. type I	-0.5	30.0	34.4
PI2	<i>Pseudomonas</i> sp. type I	-3.5	26.0	29.4
PII	<i>Pseudomonas</i> sp. type II	-2.9	27.2	34.4
Pmo	<i>Pseudomonas morgani</i> M68	2.1	36.7	42.6
Pp1	<i>Pseudomonas psychrophiles</i> p-200	-5.0	31.6	35.1
Pp2	<i>Pseudomonas psychrophiles</i> 1-3b	-10.1	30.8	34.0
Pp3	<i>Pseudomonas psychrophiles</i> 31-3c	-5.0	29.8	34.3
Psy	<i>Pseudomonas syringae</i>	-1.8	27.8	36.3
Sma	<i>Serratia marcescens</i>	-2.1	25.8	31.3
Taq	<i>Thermus aquaticus</i> TY	36.8	70.9	81.8
Vm1	<i>Vibrio marinus</i> ATCC15381	2.7	15.6	19.5
Vm2	<i>Vibrio marinus</i> ATCC15382	6.5	23.8	29.7
Vps	<i>Vibrio psychroerythrus</i> ATCC27364	3.8	12.4	18.5
Xpr	<i>Xanthomonas pruni</i>	1.1	30.8	36.8

TABLE 1 – Les températures (°C) de croissance cardinales de 47 micro-organismes d’après les tables 1 et 3 de [5].

1 ACP intuitive à la main (pataugeoire)

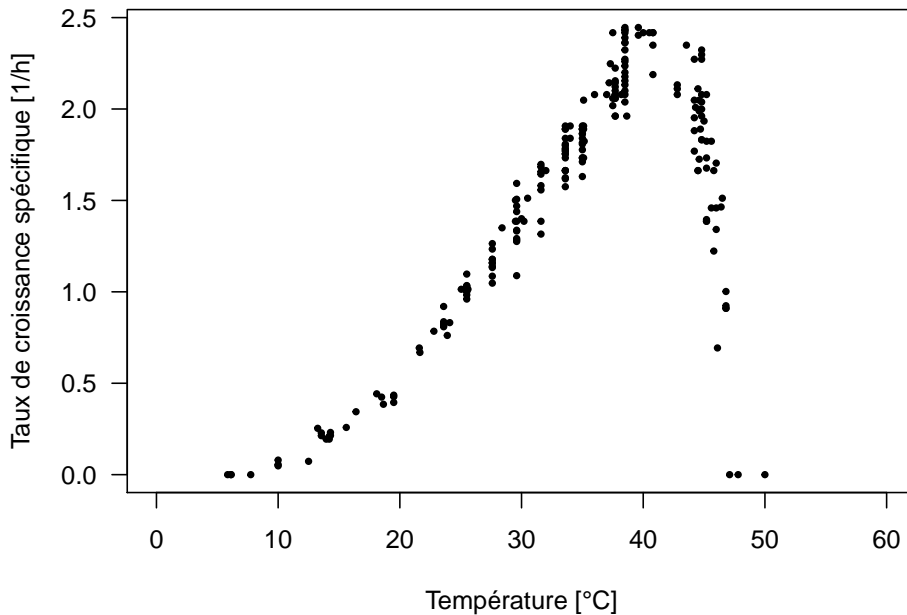
1.1 Note liminaire

LES méthodes d'analyses multivariées sont une branche à part entière de la statistique. Nous vous proposons ici une introduction, très sommaire, à l'une d'entre elles, connue sous le nom d'Analyse en Composantes Principales (ACP).

1.2 De la nature des données

LORSQUE l'on s'intéresse à l'influence de la température, T , sur le taux de croissance spécifique, μ , des micro-organismes, on obtient une réponse caractéristique illustrée ci-après avec les données publiées par BARBER en 1908 [2] dans le cas de *Escherichia coli*.

Influence de la température sur la croissance d'E. coli



ON peut résumer l'information en ajustant une courbe aux points, par exemple avec le modèle CTMI¹ [5] donné par les équations 1 et 2 ci-après.

$$\begin{cases} \mu(T) = 0 & \text{si } T < T_{\min} \\ \mu(T) = \mu_{\text{opt}}\phi(T) & \text{si } T_{\min} \leq T \leq T_{\max} \\ \mu(T) = 0 & \text{si } T > T_{\max} \end{cases} \quad (1)$$

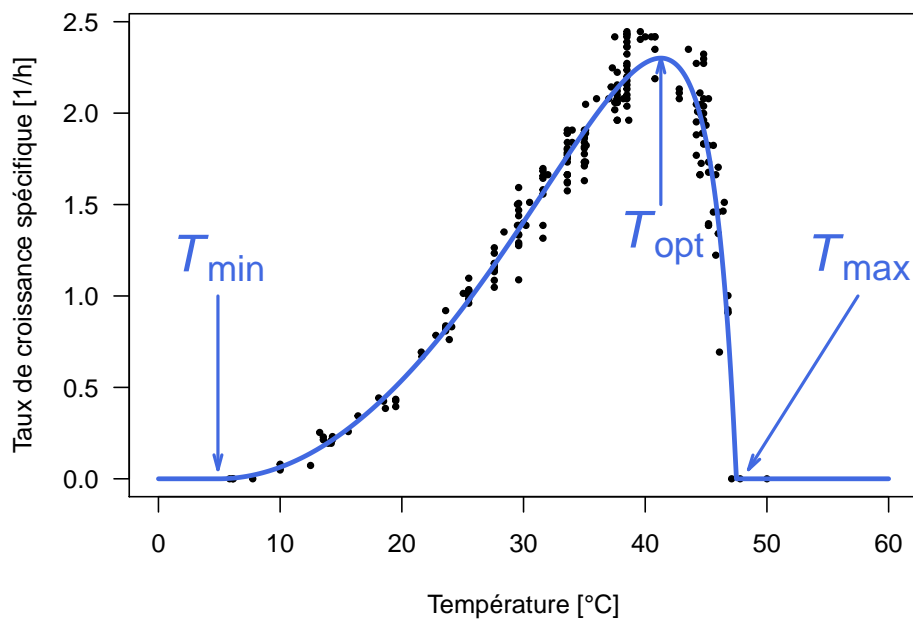
avec

1. Cardinal Temperature Model with Inflection.

$$\phi(T) = \frac{(T - T_{\max})(T - T_{\min})^2}{(T_{\text{opt}} - T_{\min})[(T_{\text{opt}} - T_{\min})(T - T_{\text{opt}}) - (T_{\text{opt}} - T_{\max})(T_{\text{opt}} + T_{\min} - 2T)]} \quad (2)$$

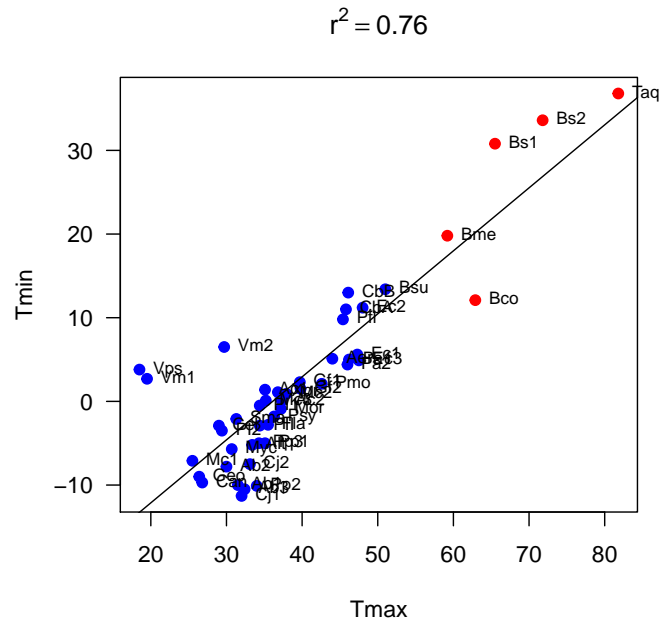
LES trois températures cardinales sont T_{\min} , la température en deçà de laquelle la croissance n'est plus possible, T_{opt} , la température pour laquelle la croissance est optimale et vaut μ_{opt} , et T_{\max} , la température en delà de laquelle la croissance n'est plus possible.

Influence de la température sur la croissance d'E. coli



LES données utilisées [5] sont les trois températures cardinales observées chez 47 micro-organismes (voir la table 1 page 2 pour le détail). Pour faciliter la lecture nous définissons un vecteur de couleurs pour distinguer en rouge les micro-organismes poussant à haute température ($T_{\text{opt}} \geq 50^\circ\text{C}$) et en bleu ceux poussant à basse température ($T_{\text{opt}} < 50^\circ\text{C}$). On commence par représenter T_{\min} en fonction de T_{\max} .

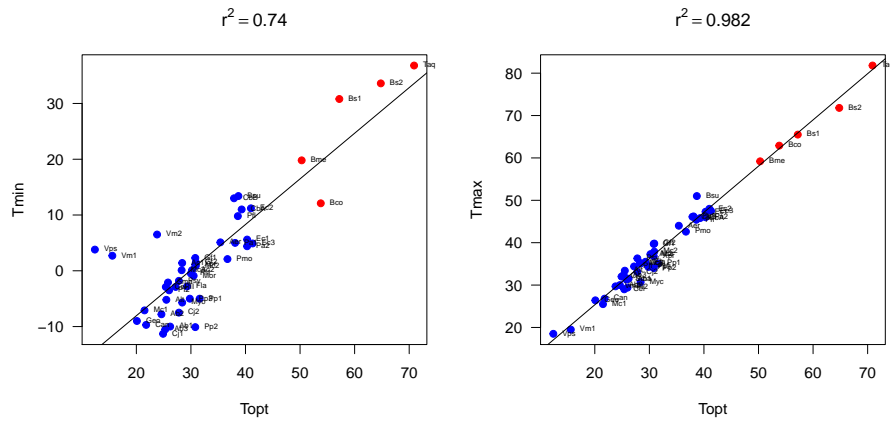
```
load(url("https://pbil.univ-lyon1.fr/R/donnees/tdr605/card.Rda"))
couleurs <- ifelse(card$Topt < 50, "blue", "red")
with(card, {
  plot(Tmax, Tmin, col = couleurs, pch = 19, las = 1)
  abline(lm(Tmin~Tmax))
  r2 <- signif(cor(Tmin, Tmax)^2, 3)
  title(main = bquote(r^2 == .(r2)))
  text(Tmax, Tmin, rownames(card), cex = 0.75, pos = 4, xpd = NA)
})
```



COMMENTEZ la relation entre ces deux températures cardinales. Quelles sont les espèces (voir la table 1 page 2 pour le nom des espèces) qui se démarquent par rapport à la tendance générale ?

Réponse :

COMME nous n'avons que $n = 3$ variables, le nombre de couples de variables possibles est faible, $\frac{n(n-1)}{2} = 3$, on peut se lancer dans une étude exhaustive de tous les couples de variables. Avec plus de variables, cette approche n'est plus envisageable. Par exemple avec 25 variables il faudrait examiner 300 graphiques. C'est dans ce cas de figure que l'ACP prendra utilement le relais. Donnez le code R permettant de reproduire les figures ci-après.



Réponse :

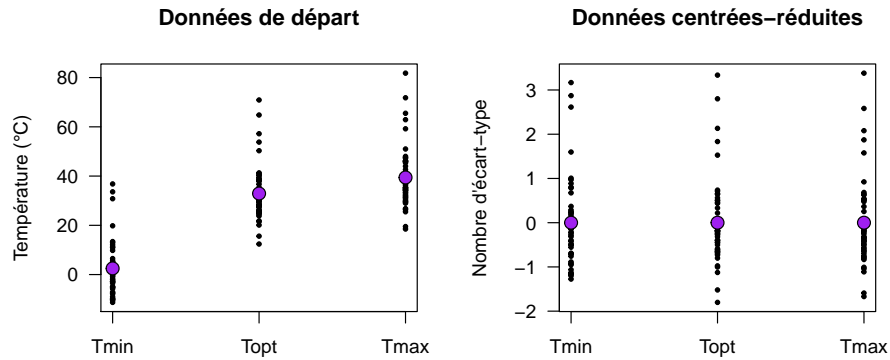
QUAND pour un ensemble de variables tous les couples de variables sont corrélés positivement on parle « *d'effet taille* ». C'est un terme qui vient à l'origine des études sur les variables morphométriques mais qui a maintenant un sens technique. Dans le cas des températures cardinales des micro-organismes on ne sait pas vraiment pourquoi il en est ainsi. C'est une observation qui a été confirmée en 2016 [4] avec un jeu de données bien plus conséquent de 464 points.

1.3 De l'imposition d'une échelle commune

IL n'y a en général pas de raison pour que toutes les variables étudiées soient exprimées comme ici dans une même unité (°C). C'est la raison pour laquelle tous les programmes d'ACP travaillent par défaut sur les variables centrées-réduites. On a alors des variables sans dimension dont la valeur dit à combien d'écart-type on se trouve au dessus ou au dessous de la moyenne.

```
library(ade4)
par(mfrow = c(1, 2))
card.cr <- as.data.frame(scalewt(card)) # centrage et réduction
stripchart(card, main = "Données de départ", las = 1,
            ylab = "Température (°C)", vertical = TRUE, pch = 19, cex = 0.5)
points(1:3, colMeans(card), pch = 21, bg = "purple", cex = 1.5)
```

```
stripchart(card.cr, main = "Données centrées-réduites", vertical = TRUE,
           pch = 19, cex = 0.5, ylab = "Nombre d'écart-type", las = 1)
points(1:3, colMeans(card.cr), pch = 21, bg = "purple", cex = 1.5)
```



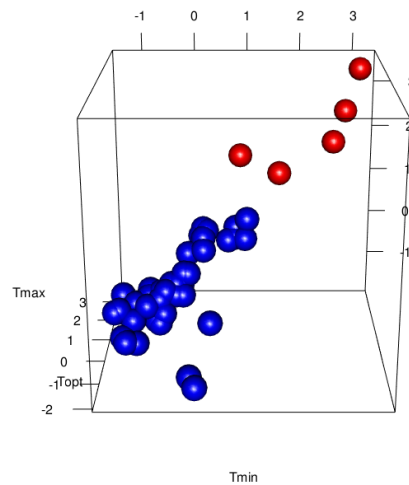
SOIT x une variable, sa moyenne \bar{x} et son écart-type s_x . Si on note x^* sa version centrée-réduite, quelle relation permet de calculer x^* ?

Réponse :

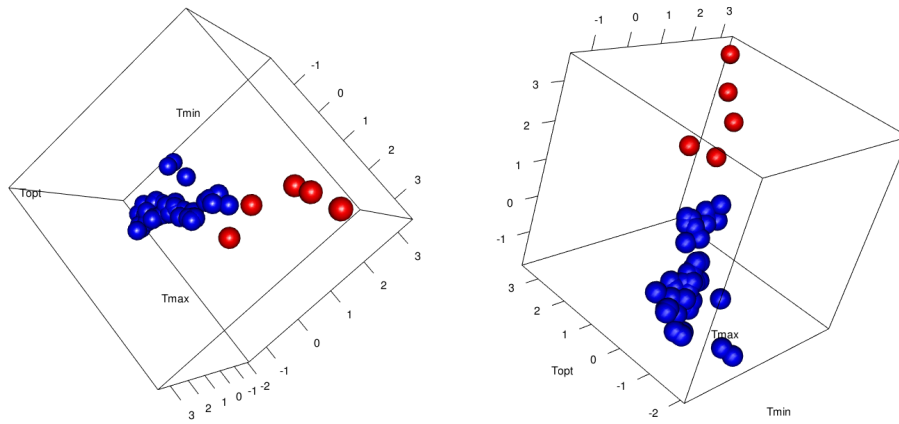
1.4 De la visualisation des données en 3D

CHACQUE individu est caractérisé par trois variables, soit par un point dans \mathbb{R}^3 . La fonction `plot3d()` de la bibliothèque `rgl` [1] vous permet d'explorer facilement un nuage de points en 3 dimensions.

```
library(rgl)
plot3d(card.cr, type = "s", col = couleurs)
```



FAITES tourner avec le curseur de la souris cette représentation pour en avoir différents points de vue :



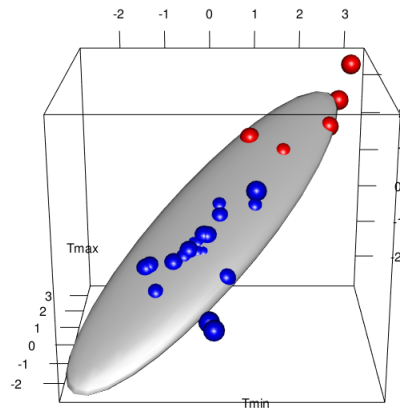
1.5 De la forme idéalisée du nuage de points

DANS l'exemple que nous avons choisi, la forme générale du nuage de points est celui d'une dragée (le terme technique est un ellipsoïde).



Des dragées

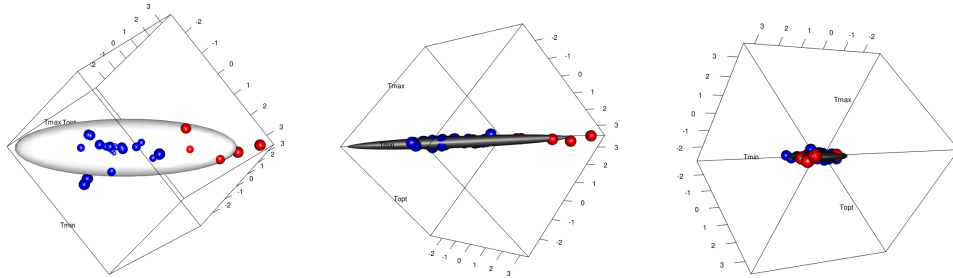
```
plot3d(card.cr, type = "s", col = couleurs)
plot3d(ellipse3d(cor(card.cr)), col = "grey", add = TRUE)
```



Une dragée est définie par ses trois axes :

- 1° le premier axe correspondant au plus grand diamètre de l'ellipsoïde, la longueur de la dragée ;
- 2° le deuxième axe correspondant au diamètre moyen de l'ellipsoïde, la largeur de la dragée ;
- 3° le troisième axe correspondant au plus petit diamètre de l'ellipsoïde, l'épaisseur de la dragée.

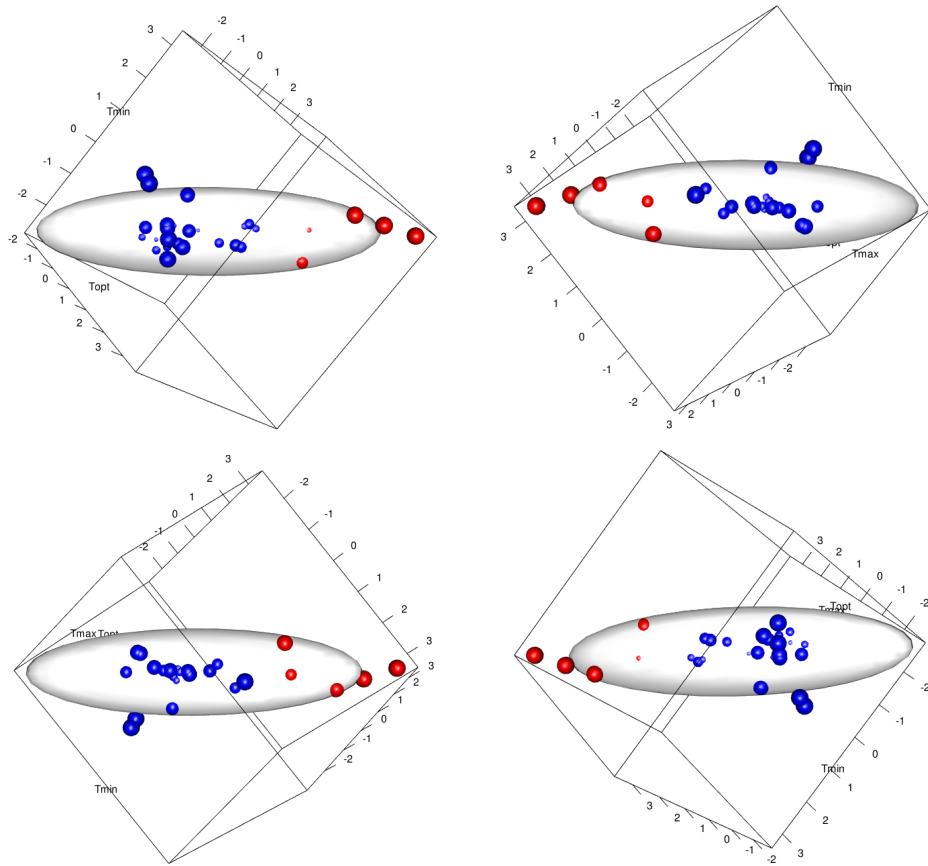
FAITES tourner le graphique précédent pour représenter le nuage de points dans le plan des axes (1, 2), puis (1, 3) puis (2, 3) :



SI vous aviez à choisir entre les trois plans que nous avons envisagés ci-dessus, vous n’hésiteriez pas à prendre le plan défini par les deux premiers axes de la dragée parce que c’est dans cette représentation que l’on a le moins de perte d’information par rapport au nuage de points dans \mathbb{R}^3 : c’est dans cette projection que les points sont les plus étalés dans le plan (on dit aussi que l’on a conservé le maximum possible de l’inertie totale du nuage de points). Ce faisant, vous avez en fait réalisé une ACP à la main.

1.6 Des symétries des solutions (important)

NOUS allons faire ici une remarque qui est valable pour toutes les méthodes d’analyses multivariées quand on cherche à interpréter un plan factoriel. Faites tourner le graphique précédent pour représenter le nuage de points dans le plan des axes (1,2) et montrez que l’on peut le faire de 4 façons parfaitement équivalentes :



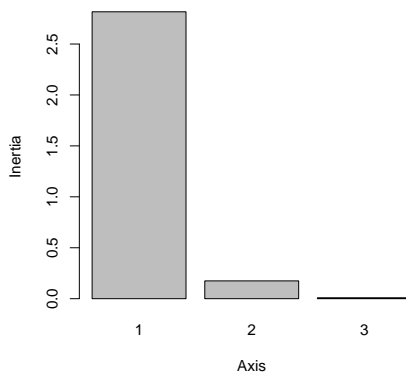
LES figures du haut sont symétriques par rapport à l'axe 1 de la dragée de celles du bas. Les figures de gauche sont symétriques par rapport à l'axe 2 de la dragée de celles de droite. Toutes ces représentations sont parfaitement équivalentes du point de vue de la minimisation de la perte d'information par rapport au nuage de points initial dans \mathbb{R}^3 . Ce qu'il faut retenir c'est que l'orientation des axes des analyses multivariées est arbitraire et donc qu'il ne faut pas chercher à donner un sens à cette orientation.

2 ACP centrée-réduite avec ade4 (petit bassin)

2.1 De la mise en œuvre

UTILISER la fonction `dudi.pca()` du paquet `ade4` [3] pour exécuter une ACP centrée réduite :

```
acp <- dudi.pca(card, center = TRUE, scale = TRUE)
```



Select the number of axes:

RÉPONDRE 2 à la question "Select the number of axes:". Dans la pratique, on ne conserve qu'un nombre réduit d'axes ; l'ACP est une technique de réduction dimensionnelle. Ici, nous en conservons 2 pour refaire par le calcul la projection $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ que nous avons fait manuellement précédemment. Le graphique produit s'appelle le « graphe des valeurs propres ». Le vecteur des valeurs correspondant est donné par `acp$eig`.

```
acp$eig
[1] 2.816425046 0.174647391 0.008927564
100*acp$eig/sum(acp$eig)
[1] 93.8808349 5.8215797 0.2975855
```

La fonction `summary()` permet de retrouver facilement toutes ces informations.

```
summary(acp)
Class: pca dudi
Call: dudi.pca(df = card, scannf = FALSE, nf = 2)
Total inertia: 3

Eigenvalues:
  Ax1      Ax2      Ax3
2.816425 0.174647 0.008928

Projected inertia (%):
  Ax1      Ax2      Ax3
93.8808  5.8216  0.2976

Cumulative projected inertia (%):
  Ax1  Ax1:2  Ax1:3
93.88  99.70 100.00
```

LES valeurs propres nous renseignent sur la fraction de l'inertie totale prise en compte par chaque axe. Elles sont rangées par convention dans l'ordre décroissant :

- 1° le premier axe correspondant au plus grand diamètre de l'ellipsoïde, la longueur de la dragée, représente 93.9 % de l'inertie totale ;
- 2° le deuxième axe correspondant au diamètre moyen de l'ellipsoïde, la largeur de la dragée, représente 5.8 % de l'inertie totale ;

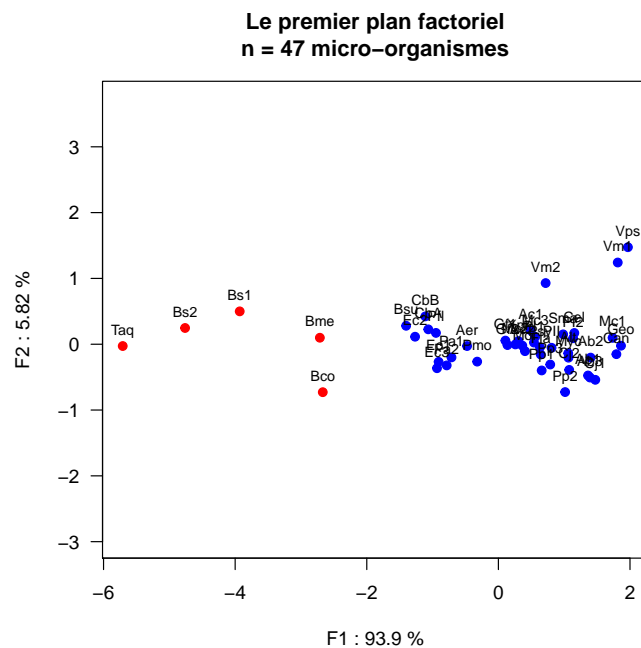
3° le troisième axe correspondant au plus petit diamètre de l'ellipsoïde, l'épaisseur de la dragée, représente 0.3 % de l'inertie totale.

CECI signifie que lorsque nous projetons le nuage de points initial dans \mathbb{R}^3 sur le plan défini par les deux premiers axes factoriels, nous n'avons perdu que très peu (0.3 %) d'information.

2.2 De la projection des individus

POUR reproduire par le calcul la projection manuelle de $\mathbb{R}^3 \rightarrow \mathbb{R}^2$, nous avons besoin de connaître les coordonnées des individus, ici nos micro-organismes, dans le repère défini par le premier et le deuxième axes. Elle sont données dans `acp$li` avec `acp$li[, 1]` pour le premier axe et `acp$li[, 2]` pour le deuxième. Nous pouvons produire ainsi le premier plan dit « factoriel » parce qu'en ACP, on parle indifféremment d'« axe » de « composante » ou de « facteur » pour désigner les axes des ellipsoïdes. On les note généralement F1, F2, F3, etc. et l'usage est de préciser le pourcentage d'inertie totale qui leur est associé.

```
x <- acp$li[, 1] ; y <- acp$li[, 2]
main <- paste("Le premier plan factoriel\nn =", nrow(card), "micro-organismes")
xlab <- paste("F1 :", signif(100*acp$eig[1]/sum(acp$eig), 3), "%")
ylab <- paste("F2 :", signif(100*acp$eig[2]/sum(acp$eig), 3), "%")
plot(x, y, pch = 19, col = couleurs, asp = 1, las = 1, main = main, xlab = xlab, ylab = ylab)
text(x, y, rownames(card), cex = 0.75, pos = 3)
```

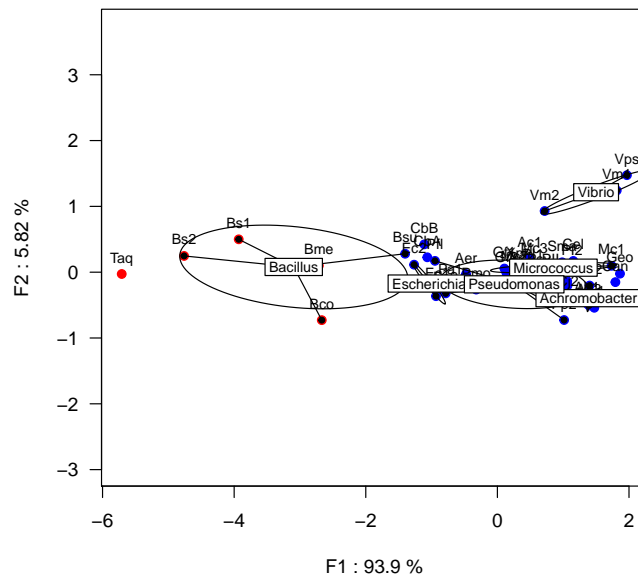


NOUS avons retrouvé ici par le calcul une des quatre solutions symétriques obtenues manuellement dans la section 1.6 page 9. Le premier facteur oppose les micro-organismes qui vivent à basse température de ceux qui vivent à haute

température. La quasi-totalité (93.9 %) de l'information de notre jeu de données dans \mathbb{R}^3 se résume donc à un simple gradient thermique dans \mathbb{R} . Quant au deuxième facteur, il met en évidence trois micro-organismes : **Vps** (*Vibrio psychroerythrus* ATCC27364), **Vm1** (*Vibrio marinus* ATCC15381) et **Vm2** (*Vibrio marinus* ATCC15382). Il est intéressant de constater que toutes ces souches appartiennent à un même genre. Inversement, on peut se demander où se situent sur le plan factoriel toutes les souches du genre *Vibrio*. Dans le graphique ci-après, on porte comme information complémentaire la localisation des genres représentés par au moins 3 souches.

```
load(url("https://pbil.univ-lyon1.fr/R/donnees/tdr605/cardGenre.Rda"))
x <- acp$li[, 1] ; y <- acp$li[, 2]
main <- paste("Le premier plan factoriel\nn =", nrow(card), "micro-organismes")
xlab <- paste("F1 :", signif(100*acp$eig[1]/sum(acp$eig), 3), "%")
ylab <- paste("F2 :", signif(100*acp$eig[2]/sum(acp$eig), 3), "%")
plot(x, y, pch = 19, col = couleurs, asp = 1, las = 1, main = main, xlab = xlab, ylab = ylab)
text(x, y, rownames(card), cex = 0.75, pos = 3)
selection <- names(table(cardGenre))[table(cardGenre) >= 3] # au moins trois souches
isel <- which(cardGenre %in% selection)
par(bty = "n")
s.class(acp$li[isel, ], cardGenre[isel], clabel = 0.75, add.plot = TRUE, axesell = FALSE)
```

**Le premier plan factoriel
n = 47 micro-organismes**



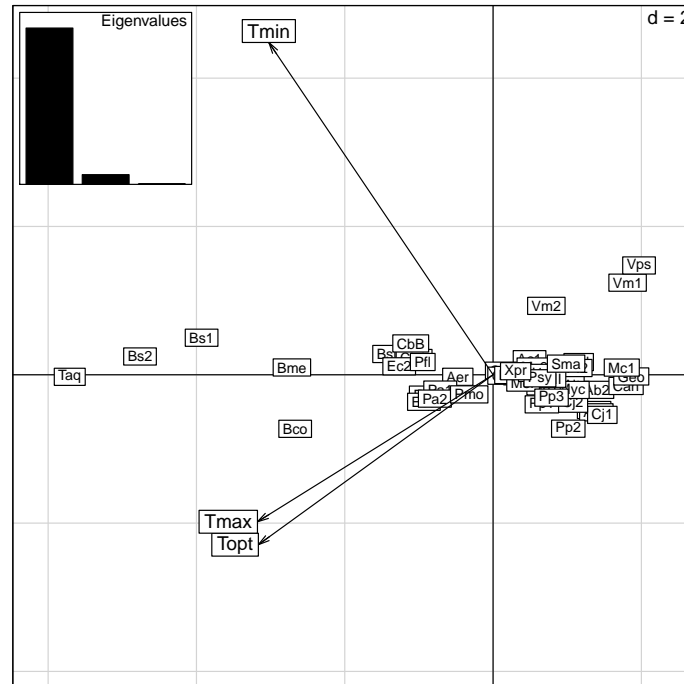
ON constate que toutes les souches du genre *Vibrio* sortent du lot sur le deuxième facteur. On ne sait pas à ce stade en quoi ces trois souches se singularisent, c'est là que les projections des variables vont aider à l'interprétation.

2.3 De la projection des variables

LES trois variables sont T_{\min} , T_{opt} et T_{\max} . À quoi peut bien correspondre la projection des variables? Quand vous avez fait l'ACP à la main, vous

avez modifié la position des points (les individus) mais aussi celle des axes (les variables). La projection des variables correspond à la projection de la base de départ (T_{\min} , T_{opt} , T_{\max}) sur le plan factoriel, on les représente avec des vecteurs pour rappeler leur nature d'axe. La fonction `scatter()` va faire cette opération pour vous.

`scatter(acp)`



DANS le cartouche en haut à gauche, on a un rappel du graphe des valeurs propres : la première valeur propre très dominante correspond au premier facteur sur l'axe horizontal et la deuxième valeur propre, beaucoup plus faible, correspond au deuxième facteur sur l'axe vertical. Interprétons maintenant les deux facteurs.

SUR le premier facteur les trois variables pointent toutes du même côté, c'est typique d'un « effet taille », quand les variables sont toutes corrélées positivement deux à deux. Les micro-organismes poussant à haute température sont « du côté flèche » des variables ce qui signifie qu'ils ont des valeurs plus élevées pour les trois températures cardinales que les autres micro-organismes : on confirme ici le gradient thermique que l'on avait mis en évidence précédemment. Notez l'angle très aigu formé par les vecteurs T_{opt} et T_{\max} , c'est la conséquence de la très forte corrélation entre ces deux variables ($r^2 = 0.982$). Dans le cas limite $r^2 = 1$ on aurait deux variables redondantes et les deux vecteurs seraient confondus.

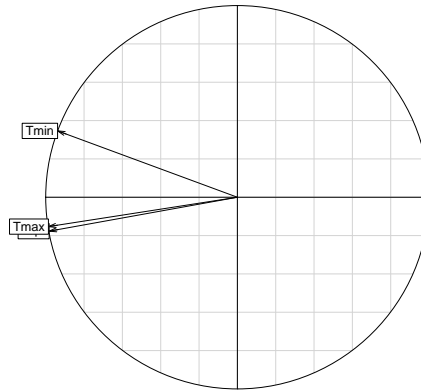
SUR le deuxième facteur, on voit que nos trois souches de *Vibrio* sont « du côté flèche » pour T_{\min} . Comme le deuxième facteur est orthogonal au premier, cela signifie que, une fois que l'on a pris en compte le gradient thermique global du jeu de données, nos trois souches se singularisent par des valeurs élevées pour

T_{\min} . La projection des variables nous aide à comprendre en quoi les individus d'intérêt diffèrent des autres.

2.4 Des limitations intrinsèques


SI deux individus ont des valeurs proches pour les variables étudiées, ils seront proches sur les plans factoriels. Si deux variables sont fortement corrélées, elles formeront un angle aigu sur les plans factoriels. Mais la réciproque n'est pas vraie, l'ACP est une méthode de projection, deux points peuvent être projetés à proximité sur un plan factoriel mais être distants dans une dimension orthogonale. L'ACP minimise globalement cette nuisance mais rien ne garanti qu'elle n'est plus présente. Pour les variables on dispose d'un outil très puissant, le cercle des corrélations :

```
s.corcircle(acp$co)
```

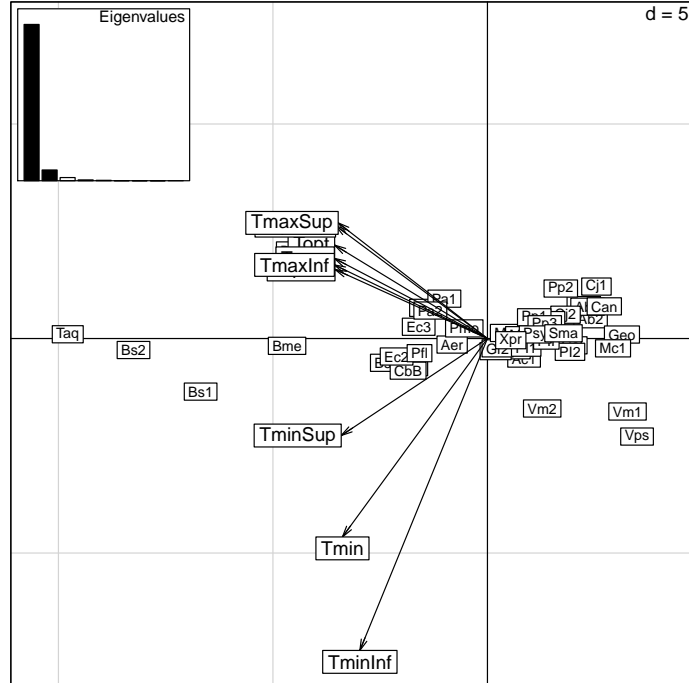


PLUS une variable est bien prise en compte dans un plan factoriel, plus la longueur du vecteur projeté correspondant sera proche de 1. On a comme repère ici le cercle de rayon unité. En l'espèce, nous sommes à « à bout touchant » pour reprendre une expression utilisée en balistique. Il n'y a pas de problème ici, la projection n'est pas trompeuse (ce qui est rassurant puisque nous sommes dans \mathbb{R}^3 avec deux variables quasi-redondantes).

3 Dans le grand bassin des dimensions élevées (exercice)

NOUS complétons le jeu de données précédent en ajoutant pour chaque température cardinale les bornes inférieures et supérieures de l'intervalle de confiance associé à l'estimation de sa valeur. Chaque micro-organisme est caractérisé maintenant par $3 \times 3 = 9$ variables, il faudrait faire 36 graphiques différents pour les confronter toutes deux-à-deux, ce qui serait fastidieux. C'est ici que l'ACP va révéler tout son potentiel en faisant une projection $\mathbb{R}^9 \rightarrow \mathbb{R}^2$, impossible à faire à la main. Importez les données puis donnez le code  permettant de produire le premier plan factoriel reproduit ci-après.

```
load(url("https://pbil.univ-lyon1.fr/R/donnees/tdr605/card9.Rda"))
```



Réponse :

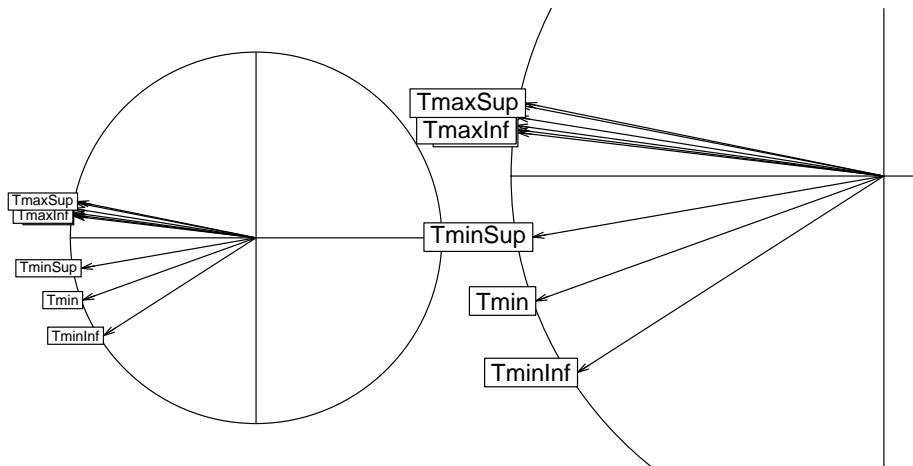
DONNEZ le pourcentage de l'inertie totale pris en compte par le premier facteur, par le second facteur et globalement par le premier plan factoriel.

Réponse :

SACHANT que nous partons d'un jeu de données dans \mathbb{R}^9 , le résumé de la projection dans \mathbb{R}^2 vous semble-t-il satisfaisant ? Comment interprétez-vous le premier et le second facteurs ?

Réponse :

DONNEZ le code **R** permettant de produire le cercle des corrélations reproduit à gauche ci-après. Les vecteurs pour T_{\min} sont plus dispersés que ceux pour T_{opt} et T_{\max} , comment l'interprétez-vous ?



Réponse :

TOUS les vecteurs pointent du même côté sur l'axe horizontal. Quel est le terme technique décrivant cette situation ? Que signifie-t-il ?

Réponse :

4 Pour aller plus loin

POUR aller plus loin nous vous conseillons de travailler avec les données de la fiche tdr602². Elles présentent l'avantage d'être faciles à interpréter, ce sont les notes de 109 étudiants pour 9 matières. De plus, elles montrent que l'ACP centrée-réduite, qui est souvent celle que l'on utilise par défaut, n'est pas forcément la méthode la plus adaptée.

Références

- [1] ADLER, D., AND MURDOCH, D. *rgl : 3D visualization device system (OpenGL)*, 2014. R package version 0.93.996.
- [2] BARBER, M. The rate of multiplication of *Bacillus coli* at different temperatures. *Journal of Infectious diseases* 5 (1908), 379–400.
- [3] CHESSEL, D., DUFOUR, A.-B., AND THIOULOUSE, J. The ade4 package-I-One-table methods. *R News* 4 (2004), 5–10.
- [4] GRIMAUD, G. *Modelling the temperature effect on phytoplankton : from acclimation to adaptation*. PhD thesis, Université Nice Sophia-Antipolis, France, 2016.

2. <http://pbil.univ-lyon1.fr/R/pdf/tdr602.pdf>

- [5] ROSSO, L., LOBRY, J., AND FLANDROIS, J.-P. An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. *Journal of Theoretical Biology* 162, 4 (1993), 447–463.