Université de Technologie de Compiègne
SY09 : Analyse des données et Apprentissage automatique

SY09 : TP 03 : Représentation Euclidienne des Données

Julien Jerphanion

Printemps 2018

Table des matières

1 Représentation euclidienne **2**
1.1 Représentation des données . . . 2
1.1.1 Identification des groupes d'élèves en fonction des résultats scolaires . . . 6
1.2 Projection et qualité de réprésentation . . . 8
1.2.1 Premier changement de base . . . 8
1.2.2 Second changement de base . . . 12
1.2.3 Troisième changement de base . . . 16
1.3 Choix d'une représentation . . . 20

2 Question Théorique **21**
2.1 Relation entre inertie et trace de la matrice de covariance empirique . . . 21
2.2 Invariance de la trace par permutation circulaire . . . 22
2.3 Préservation du centrage par changement de bases . . . 23
2.4 Préservation de l'inertie du nuage de points par changement de bases orthonormées . . . 23

1 Représentation euclidienne

1.1 Représentation des données

Ici, nous allons nous intéresser à un jeu de données contenant les notes de 9 individus ($n = 9$) pour 5 matières ($p = 5$):

— `math`: mathématiques;
— `scie`: sciences « naturelles »
— `fran`: français
— `lati`: latin
— `d.m`: « arts »

On veut trouver une *base de representation* permettant de visualiser les données de la *meilleure façon possible*.

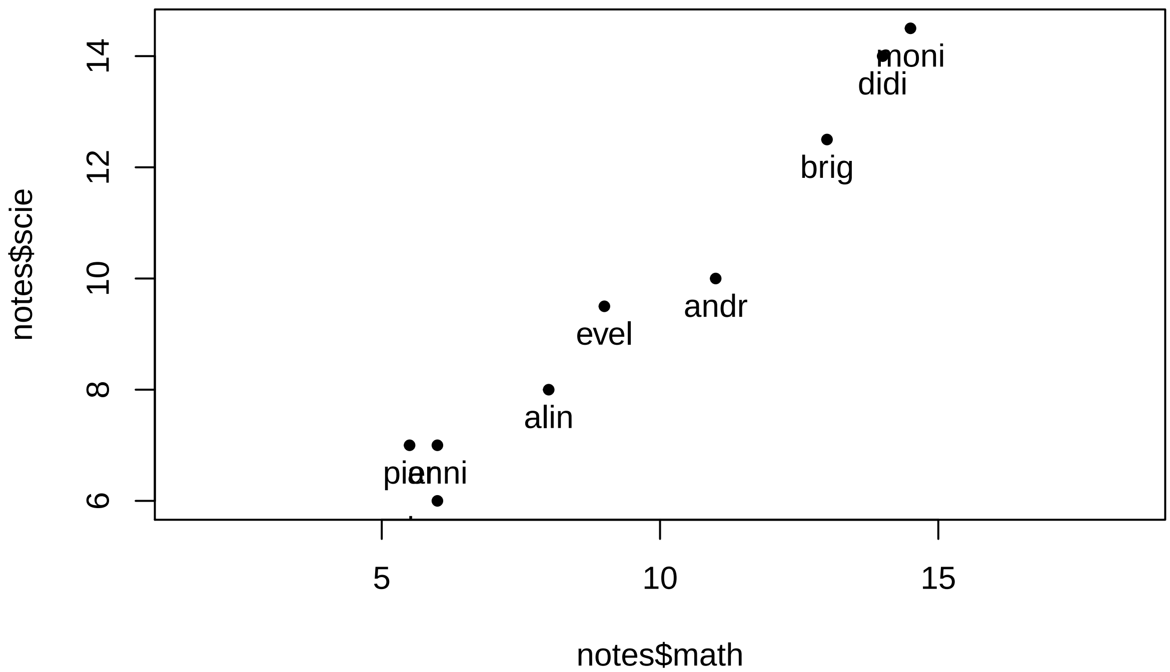
On peut, tout d'abord, charger les données ainsi:

```
notes = read.table("donnees/notes.txt", header = TRUE)
notes
```

```
##      math scie fran lati d.m
## jean  6.0  6.0  5.0  5.5   8
## alin  8.0  8.0  8.0  8.0   9
## anni  6.0  7.0 11.0  9.5  11
## moni 14.5 14.5 15.5 15.0   8
## didi 14.0 14.0 12.0 12.5  10
## andr 11.0 10.0  5.5  7.0  13
## pier  5.5  7.0 14.0 11.5  10
## brig 13.0 12.5  8.5  9.5  12
## evel  9.0  9.5 12.5 12.0  18
```

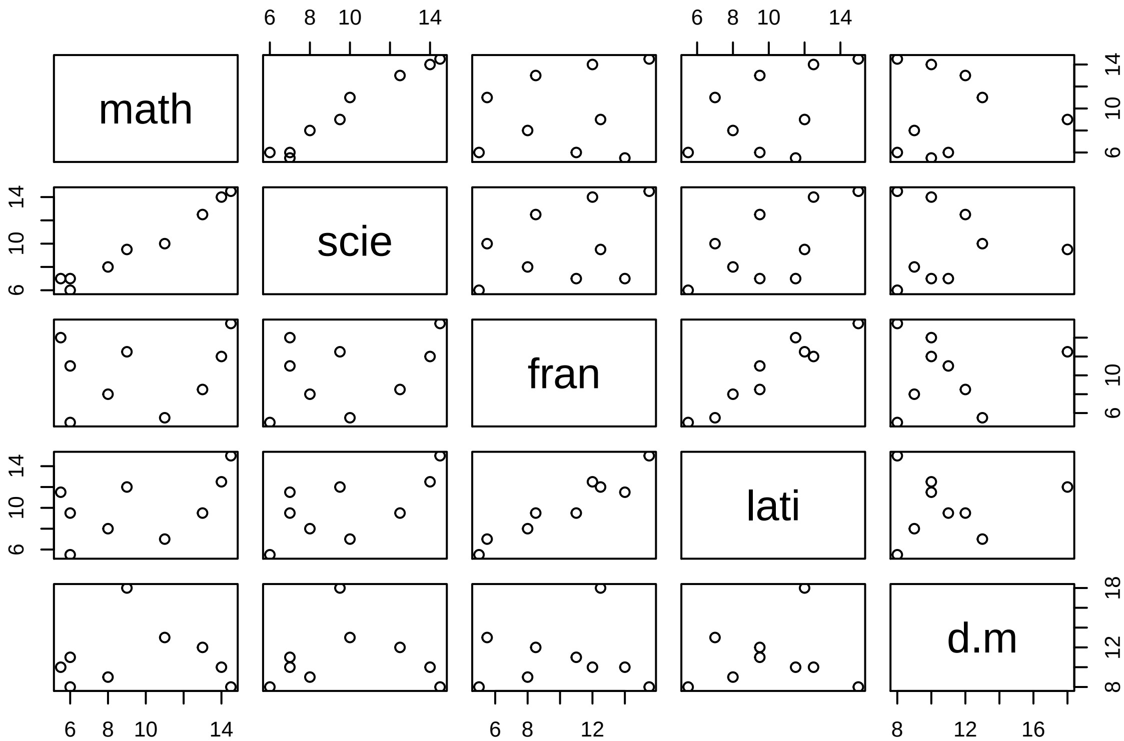
Pour se donner un premier aperçu, on peut plotter deux séries de notes ainsi – avec ici les sciences en fonction des maths:

```
plot(notes$math, notes$scie, pch=20, asp=1)
text(notes[,c("math","scie")], row.names(notes), pos=1)
```



On peut, de manière plus générale, générer un ensemble de graphes de dispersion pour avoir une meilleure intuition:

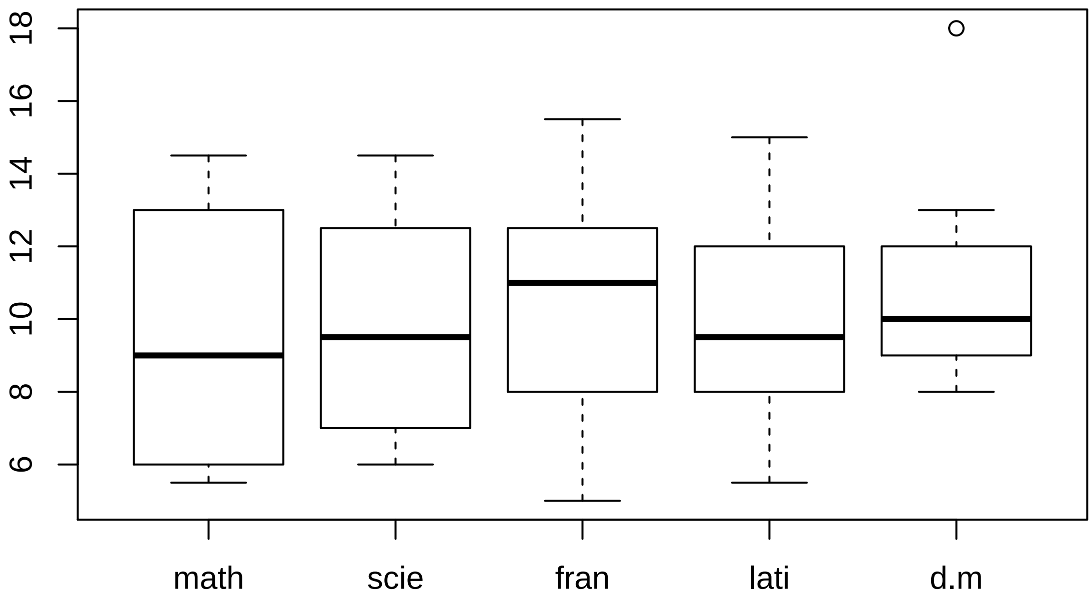
```
plot(notes)
```



Ici, on peut voir que `math` et `scie` semblent être corrélés. Cela est aussi vrai pour `fran` et `lati`. `d.m` ne semble pas particulièrement être corrélé avec le reste.

On peut utiliser des boîtes à moustaches pour avoir rapidement un aperçu univarié de chaque variable.

```
boxplot(notes)
```



Après avoir tracé ces quelques diagrammes et représentations graphiques, on peut s'intéresser à la corrélation des variables:

```
corMat = cor(notes)
corMat
```

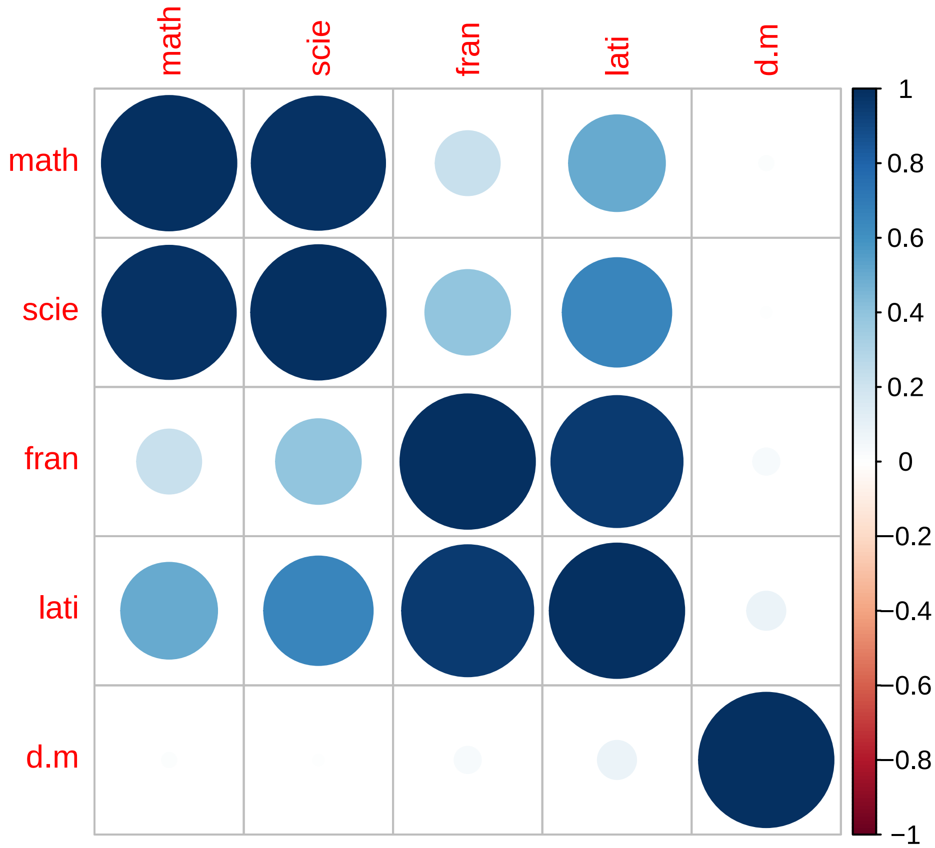
```
##             math        scie       fran       lati         d.m
## math 1.00000000 0.982535729 0.22673193 0.50814398 0.011183835
## scie 0.98253573 1.000000000 0.39669324 0.65153051 0.006309933
## fran 0.22673193 0.396693238 1.00000000 0.95120575 0.038035989
## lati 0.50814398 0.651530514 0.95120575 1.00000000 0.080500057
## d.m  0.01118384 0.006309933 0.03803599 0.08050006 1.000000000
```

Ce résultat sous forme numérique ne laisse que difficilement percevoir les tendance. Pour pallier à ce problème, on peut représenter cette matrice sous la forme d'une *heatmap* pour avoir un visuel plus probant que des données chiffrées. On utilise pour ce faire, la fameuse bibliothèque `corrplot`.

```
library(corrplot)
```

```
## corrplot 0.84 loaded
```

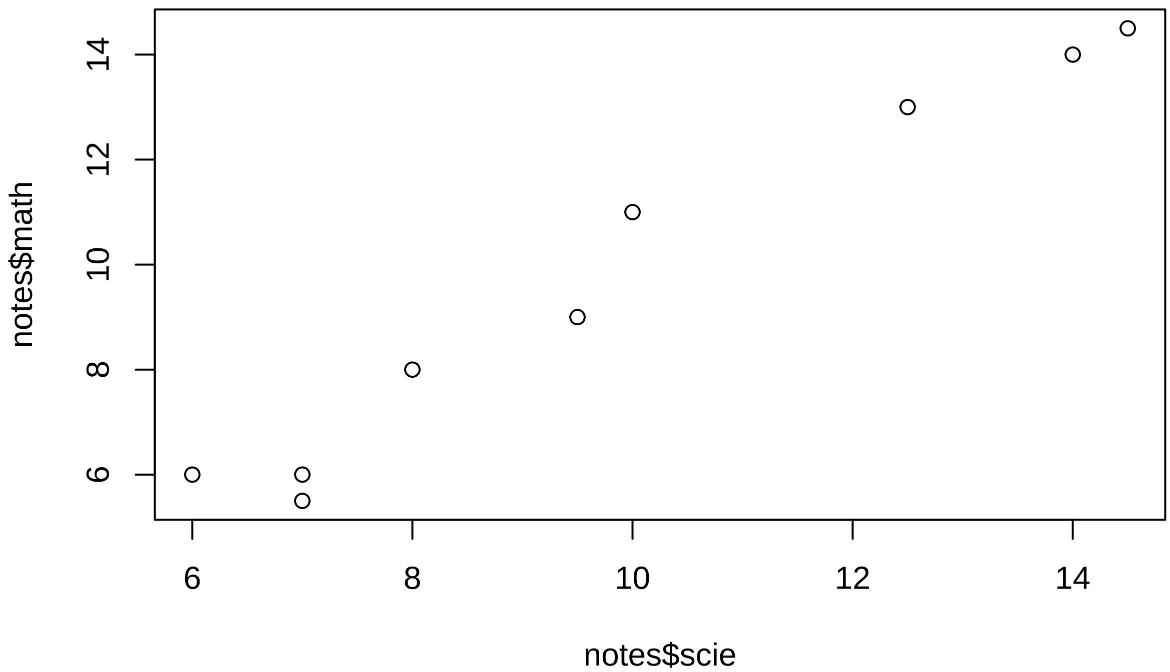
```
corrplot(cor(notes))
```



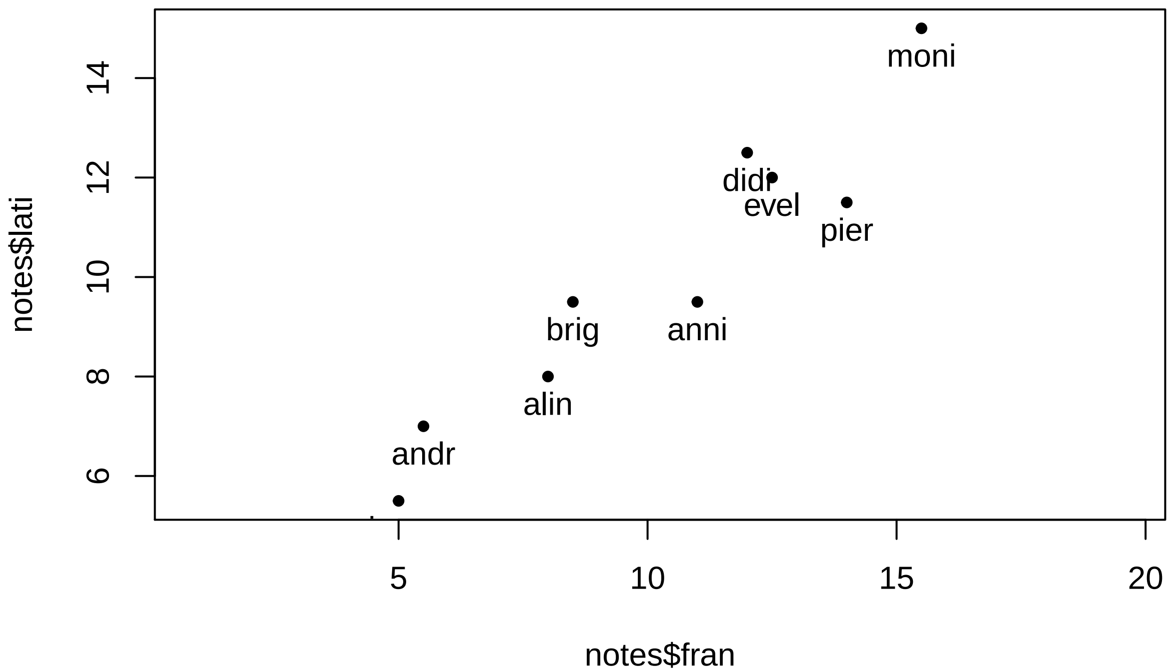
On voit facilement que les matières scientiques sont fortement corrélées ; il en est de même pour les matières littéraires. En revanche les matières artistiques, ne semblent pas être corrélés avec d'autres matières.

1.1.1 Identification des groupes d'élèves en fonction des résultats scolaires

```
plot(notes$math~notes$scie)
```




```
plot(notes$fran, notes$lati, pch=20, asp=1)
text(notes[,c("fran","lati")], row.names(notes), pos=1)
```



1.2 Projection et qualité de réprésentation

1.2.1 Premier changement de base

Soit la matrice A_1 suivante:

$$A_1 = \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & -1/2 & 0 & 0 \\ 0 & 1/2 & 10 & 1/2 & 0 \\ 0 & 1/2 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Les colonnes de A_1 sont orthogonales donc linéairement indépendantes ; puisque A_1 est carrée, elle définit donc une base.

On peut la représenter dans R ainsi:

```
A1 = matrix(c(1, 0, 1, 0, 0,
              1, 0,-1, 0, 0,
              0, 1, 0, 1, 0,
              0, 1, 0,-1, 0,
              0, 0, 0, 0, 2)/2,
            ncol = 5,
            byrow = TRUE)
A1
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]  0.5  0.0  0.5  0.0    0
## [2,]  0.5  0.0 -0.5  0.0    0
## [3,]  0.0  0.5  0.0  0.5    0
## [4,]  0.0  0.5  0.0 -0.5    0
## [5,]  0.0  0.0  0.0  0.0    1
```

Dès lors, on peut vérifier que A_1 définit bien une base en montrant que son déterminant est non nul.

```
det(A1)
```

```
## [1] -0.25
```

On peut alors chercher à représenterles `notes` dans la nouvelle base A_1. Ce *changement de base* peut se faire avec une multiplication de matrices.

Pour cela on crée la matrice X des notes:

```
X = data.matrix(notes, rownames.force = NA)
X
```

```
##      math scie fran lati d.m
## jean  6.0  6.0  5.0  5.5   8
## alin  8.0  8.0  8.0  8.0   9
## anni  6.0  7.0 11.0  9.5  11
## moni 14.5 14.5 15.5 15.0   8
## didi 14.0 14.0 12.0 12.5  10
## andr 11.0 10.0  5.5  7.0  13
## pier  5.5  7.0 14.0 11.5  10
## brig 13.0 12.5  8.5  9.5  12
## evel  9.0  9.5 12.5 12.0  18
```

Dans la suite, on aura besoin de l'inverse de A_1, que l'on calcule ici:

```
AInv = solve(A1)
AInv
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    1    0    0    0
## [2,]    0    0    1    1    0
## [3,]    1   -1    0    0    0
## [4,]    0    0    1   -1    0
## [5,]    0    0    0    0    1
```

Et on peut maintenant passer de la base canonique sous-jacente:

```
XA = X %*% t(AInv)
XA
```

```
##      [,1] [,2] [,3] [,4] [,5]
## jean 12.0 10.5  0.0 -0.5    8
## alin 16.0 16.0  0.0  0.0    9
## anni 13.0 20.5 -1.0  1.5   11
## moni 29.0 30.5  0.0  0.5    8
## didi 28.0 24.5  0.0 -0.5   10
## andr 21.0 12.5  1.0 -1.5   13
## pier 12.5 25.5 -1.5  2.5   10
## brig 25.5 18.0  0.5 -1.0   12
## evel 18.5 24.5 -0.5  0.5   18
```

On peut se ramener à la base initiale en utilisant A_1 ainsi

```r
XInit = XA %*% t(A1)
XInit == X
```

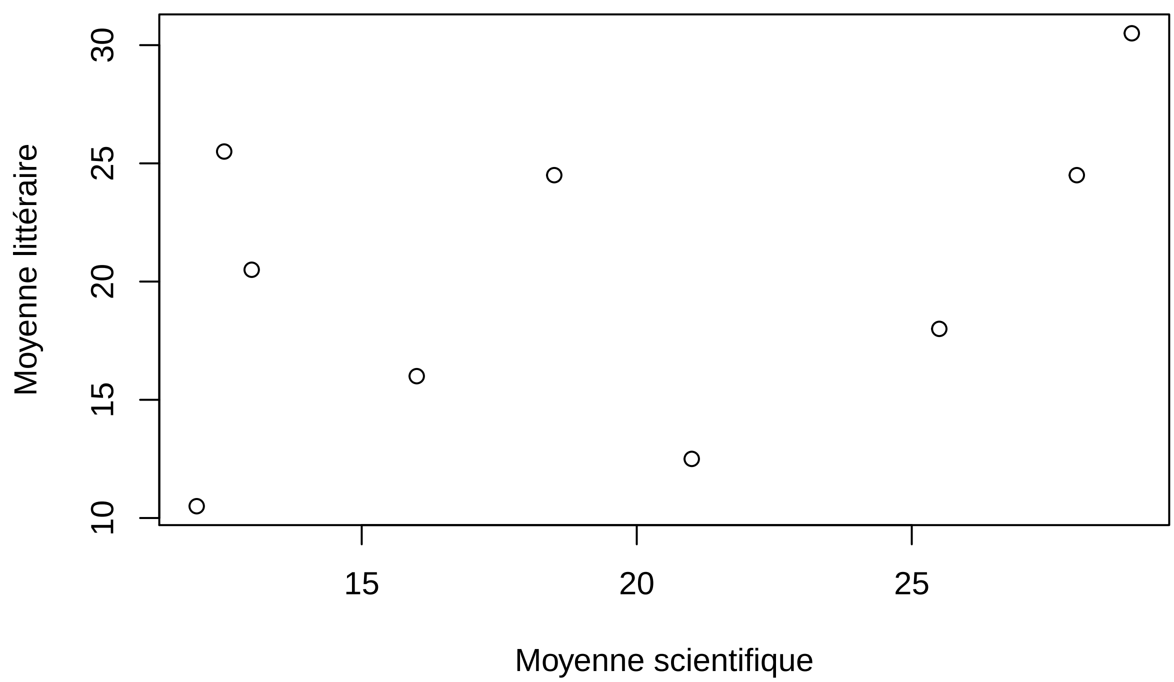
```
##      [,1] [,2] [,3] [,4] [,5]
## jean TRUE TRUE TRUE TRUE TRUE
## alin TRUE TRUE TRUE TRUE TRUE
## anni TRUE TRUE TRUE TRUE TRUE
## moni TRUE TRUE TRUE TRUE TRUE
## didi TRUE TRUE TRUE TRUE TRUE
## andr TRUE TRUE TRUE TRUE TRUE
## pier TRUE TRUE TRUE TRUE TRUE
## brig TRUE TRUE TRUE TRUE TRUE
## evel TRUE TRUE TRUE TRUE TRUE
```

Les données que l'on obtient dans la base définie par A_1 représentent, pour chaque individu:

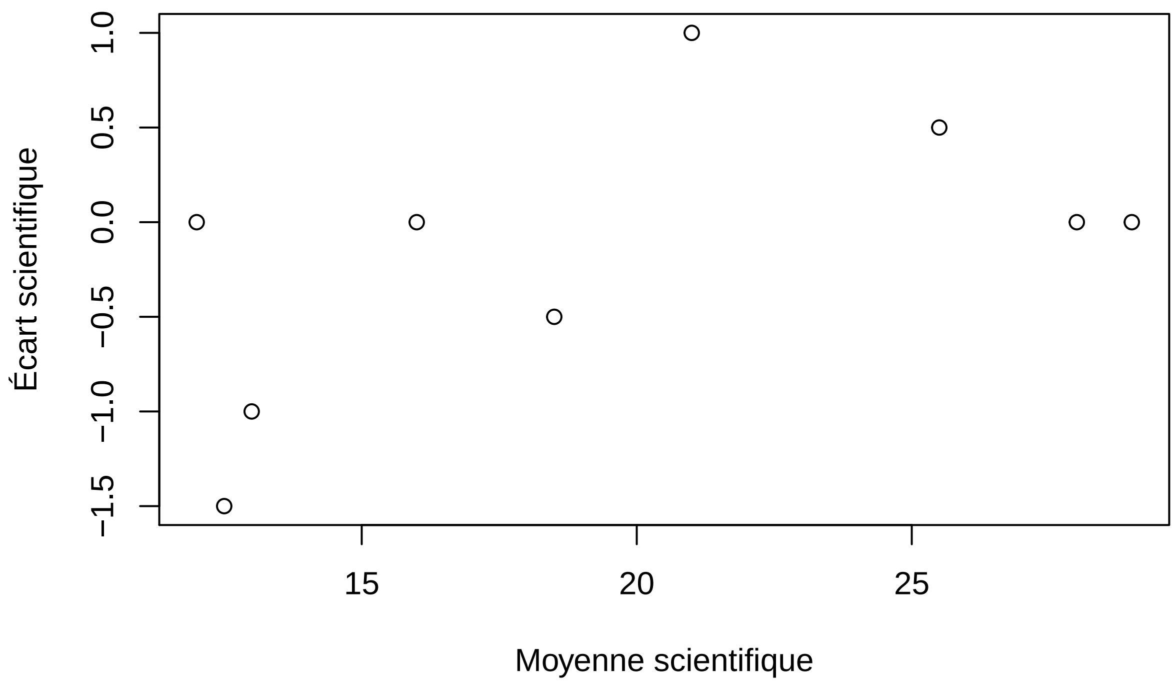
- la moyenne de ses notes des matières scientifiques
- la moyenne de ses notes des matières littéraires
- l'écart normalisé entre ses moyennes de `math` et de `sci`
- l'écart normalisé entre ses moyennes de `fran` et de `lati`
- sa moyenne d'arts

On peut représenter les composantes (X^1, X^2), ainsi que les composantes (X^1, X^3):

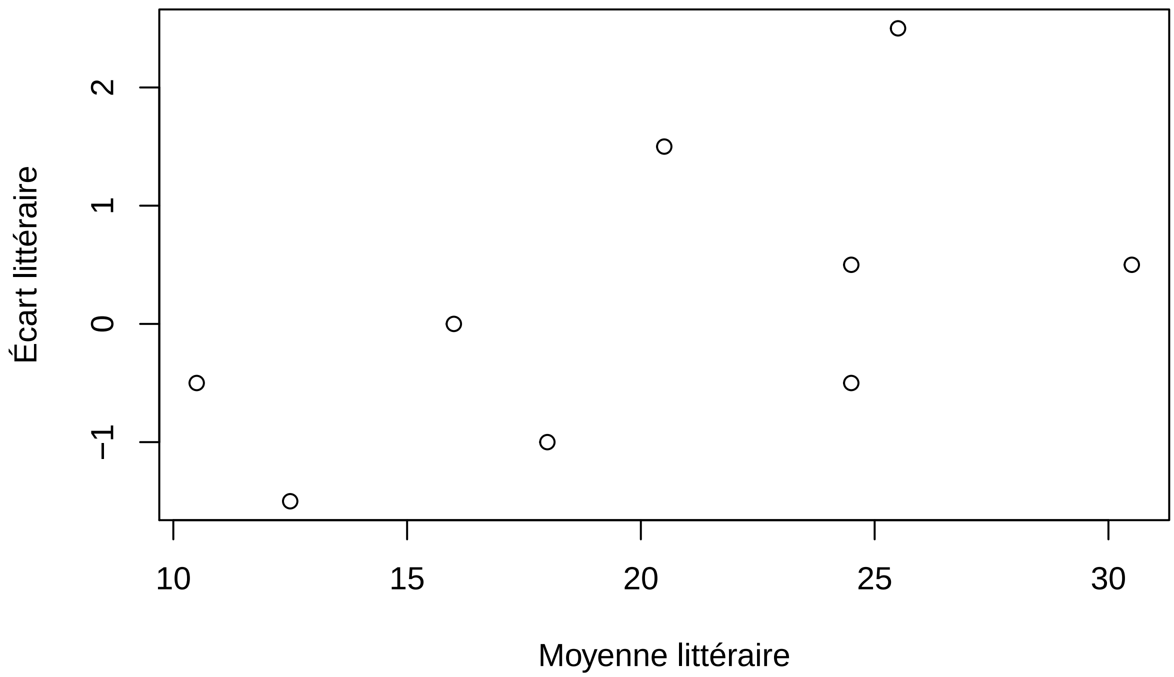
```r
plot(XA[,1],XA[,2],xlab = "Moyenne scientifique", ylab = "Moyenne littéraire")
```




```r
plot(XA[,1],XA[,3],xlab = "Moyenne scientifique", ylab = "Écart scientifique")
```




```r
plot(XA[,2],XA[,4],xlab = "Moyenne littéraire", ylab = "Écart littéraire")
```



1.2.2 Second changement de base

On considère maintenant la matrice B_1 suivante:

$$B_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Elle définit aussi une base car elle est orthogonale ; voyez-plutôt:

```
B1 = matrix(c(1, 0, 1, 0, 0,
              1, 0,-1, 0, 0,
              0, 1, 0, 1, 0,
              0, 1, 0,-1, 0,
              0, 0, 0, 0, 2/sqrt(2))*sqrt(2)/2,
            ncol = 5,
            byrow = TRUE)
B1
```

```
##           [,1]      [,2]       [,3]       [,4] [,5]
## [1,] 0.7071068 0.0000000  0.7071068  0.0000000    0
## [2,] 0.7071068 0.0000000 -0.7071068  0.0000000    0
## [3,] 0.0000000 0.7071068  0.0000000  0.7071068    0
```

```
## [4,] 0.0000000 0.7071068  0.0000000 -0.7071068    0
## [5,] 0.0000000 0.0000000  0.0000000  0.0000000    1
```

```
round(t(B1) %*% B1,3)
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
```

On peut changer de base comme tout à l'heure:

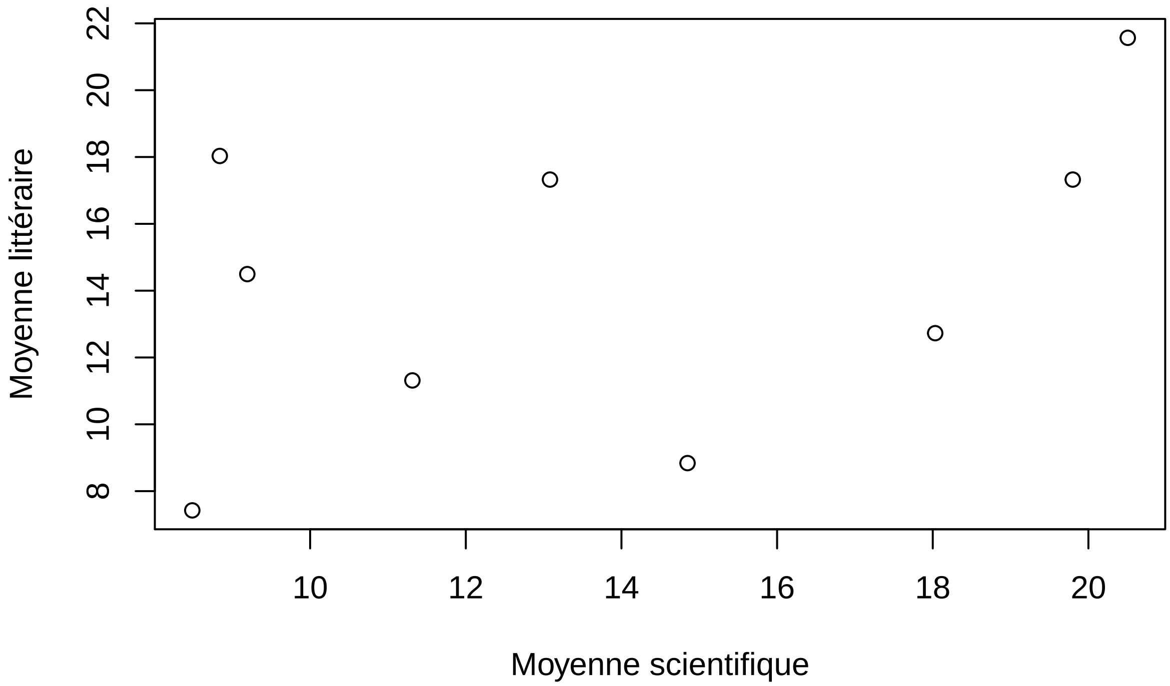
```
B1Inv = t(B1)
B1Inv
```

```
##           [,1]       [,2]      [,3]       [,4] [,5]
## [1,] 0.7071068  0.7071068 0.0000000  0.0000000    0
## [2,] 0.0000000  0.0000000 0.7071068  0.7071068    0
## [3,] 0.7071068 -0.7071068 0.0000000  0.0000000    0
## [4,] 0.0000000  0.0000000 0.7071068 -0.7071068    0
## [5,] 0.0000000  0.0000000 0.0000000  0.0000000    1
```

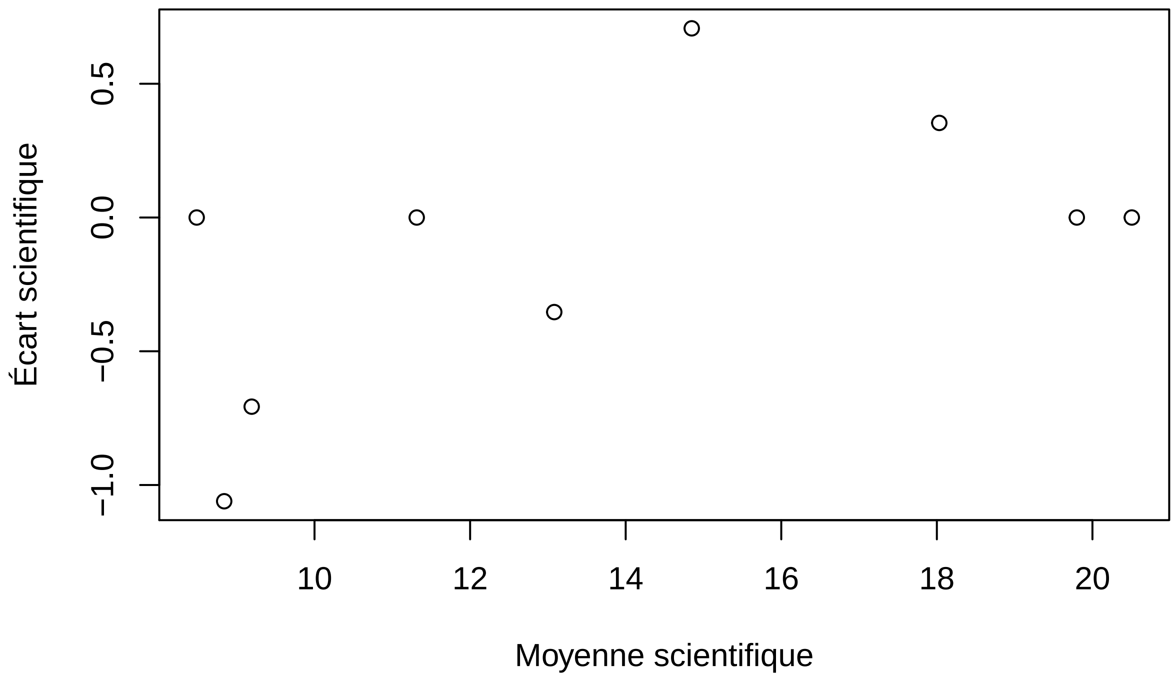
```
XB1 = X %*% B1
XB1
```

```
##           [,1]      [,2]          [,3]       [,4] [,5]
## jean  8.485281  7.424621  2.220446e-16 -0.3535534    8
## alin 11.313708 11.313708  0.000000e+00  0.0000000    9
## anni  9.192388 14.495689 -7.071068e-01  1.0606602   11
## moni 20.506097 21.566757  3.885781e-16  0.3535534    8
## didi 19.798990 17.324116 -6.661338e-16 -0.3535534   10
## andr 14.849242  8.838835  7.071068e-01 -1.0606602   13
## pier  8.838835 18.031223 -1.060660e+00  1.7677670   10
## brig 18.031223 12.727922  3.535534e-01 -0.7071068   12
## evel 13.081475 17.324116 -3.535534e-01  0.3535534   18
```

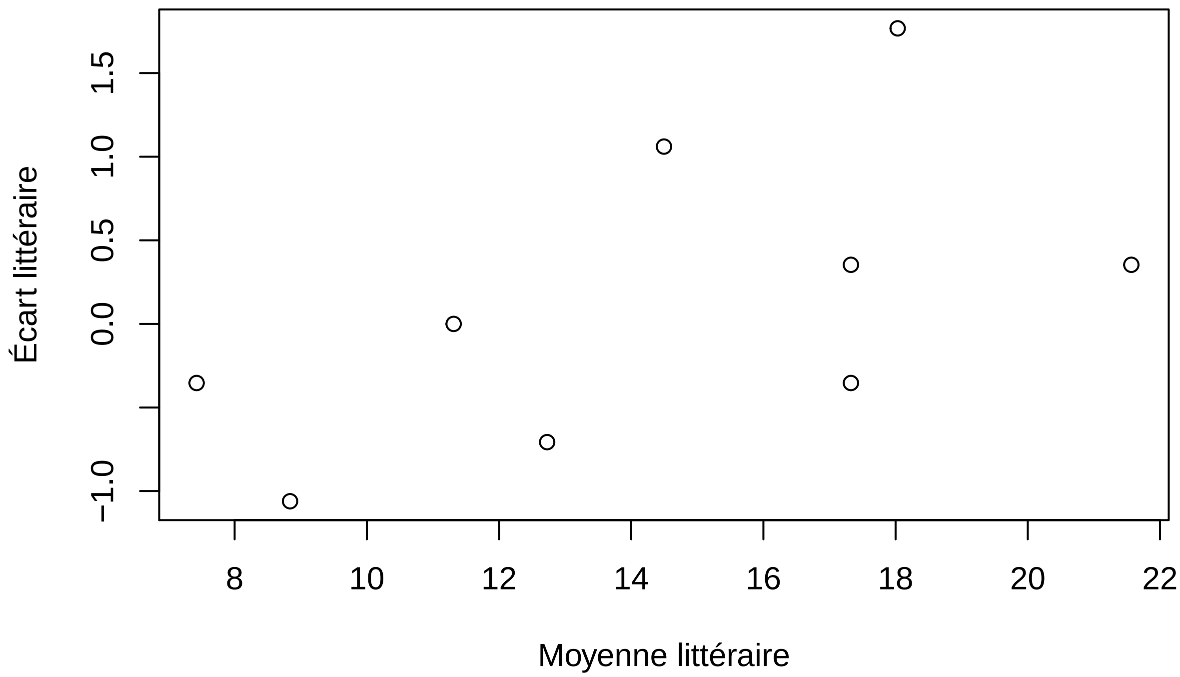
```
plot(XB1[,1],XB1[,2],xlab = "Moyenne scientifique", ylab = "Moyenne littéraire")
```




```
plot(XB1[,1],XB1[,3],xlab = "Moyenne scientifique", ylab = "Écart scientifique")
```




```r
plot(XB1[,2],XB1[,4],xlab = "Moyenne littéraire", ylab = "Écart littéraire")
```



1.2.3 Troisième changement de base

$$B_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

```
B2 = matrix(c(1, 0, 1, 0, 0,
              0, 1, 0, 1, 0,
              1, 0,-1, 0, 0,
              0, 1, 0,-1, 0,
              0, 0, 0, 0, 2/sqrt(2))*sqrt(2)/2,
            ncol = 5,
            byrow = TRUE)
B2
```

```
##           [,1]      [,2]       [,3]       [,4] [,5]
## [1,] 0.7071068 0.0000000  0.7071068  0.0000000    0
## [2,] 0.0000000 0.7071068  0.0000000  0.7071068    0
## [3,] 0.7071068 0.0000000 -0.7071068  0.0000000    0
## [4,] 0.0000000 0.7071068  0.0000000 -0.7071068    0
## [5,] 0.0000000 0.0000000  0.0000000  0.0000000    1
```

Cette matrice est orthogonale:

```r
round(t(B2) %*% B2,3)
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
```

Elle définit une base à laquelle on peut se reporter:

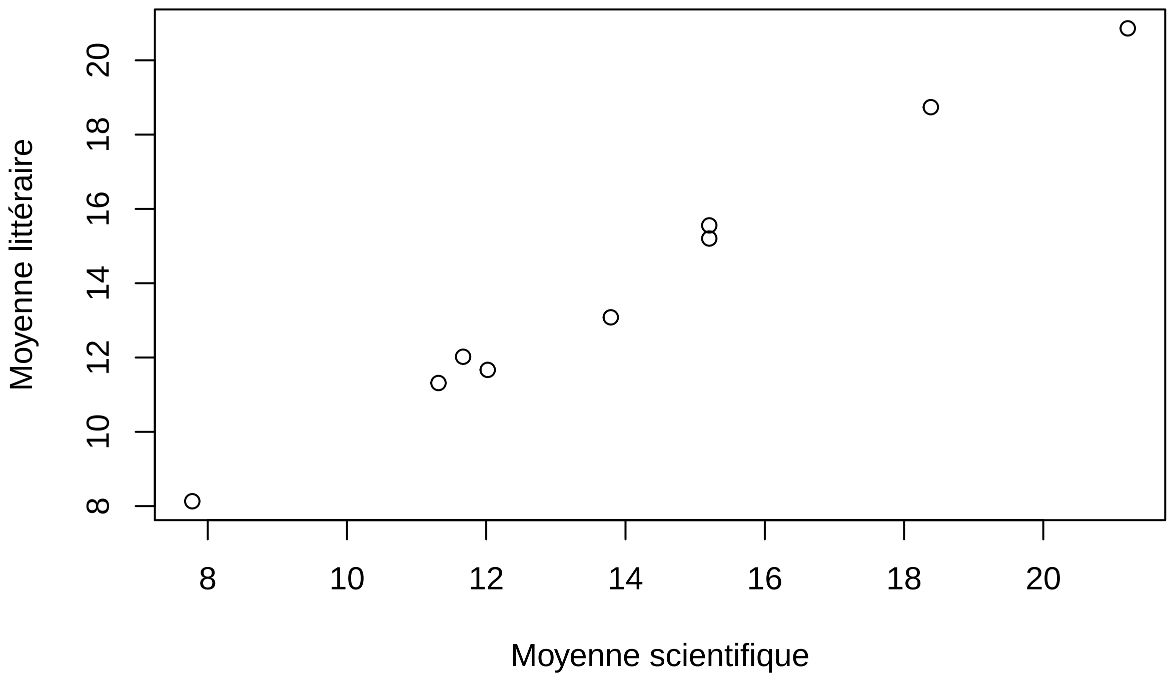
```r
B2Inv = t(B2)
B2Inv
```

```
##           [,1]      [,2]       [,3]       [,4] [,5]
## [1,] 0.7071068 0.0000000  0.7071068  0.0000000    0
## [2,] 0.0000000 0.7071068  0.0000000  0.7071068    0
## [3,] 0.7071068 0.0000000 -0.7071068  0.0000000    0
## [4,] 0.0000000 0.7071068  0.0000000 -0.7071068    0
## [5,] 0.0000000 0.0000000  0.0000000  0.0000000    1
```

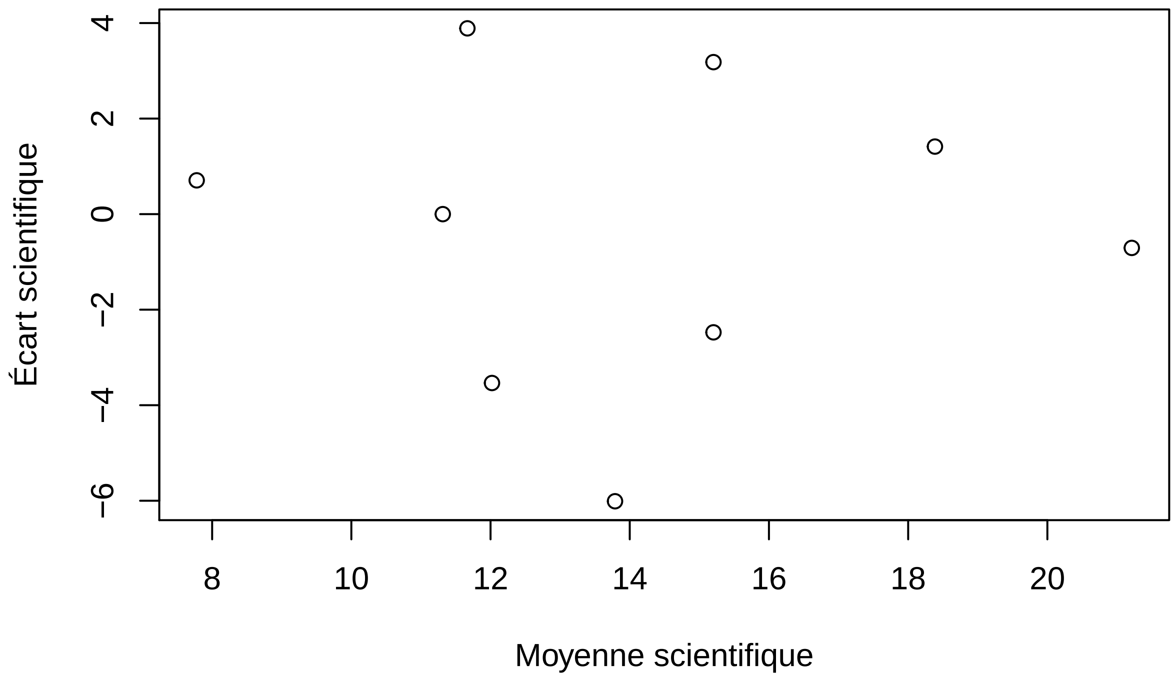
```r
XB2 = X %*% B2
XB2
```

```
##           [,1]      [,2]       [,3]       [,4] [,5]
## jean  7.778175  8.131728  0.7071068  0.3535534    8
## alin 11.313708 11.313708  0.0000000  0.0000000    9
## anni 12.020815 11.667262 -3.5355339 -1.7677670   11
## moni 21.213203 20.859650 -0.7071068 -0.3535534    8
## didi 18.384776 18.738330  1.4142136  1.0606602   10
## andr 11.667262 12.020815  3.8890873  2.1213203   13
## pier 13.788582 13.081475 -6.0104076 -3.1819805   10
## brig 15.202796 15.556349  3.1819805  2.1213203   12
## evel 15.202796 15.202796 -2.4748737 -1.7677670   18
```

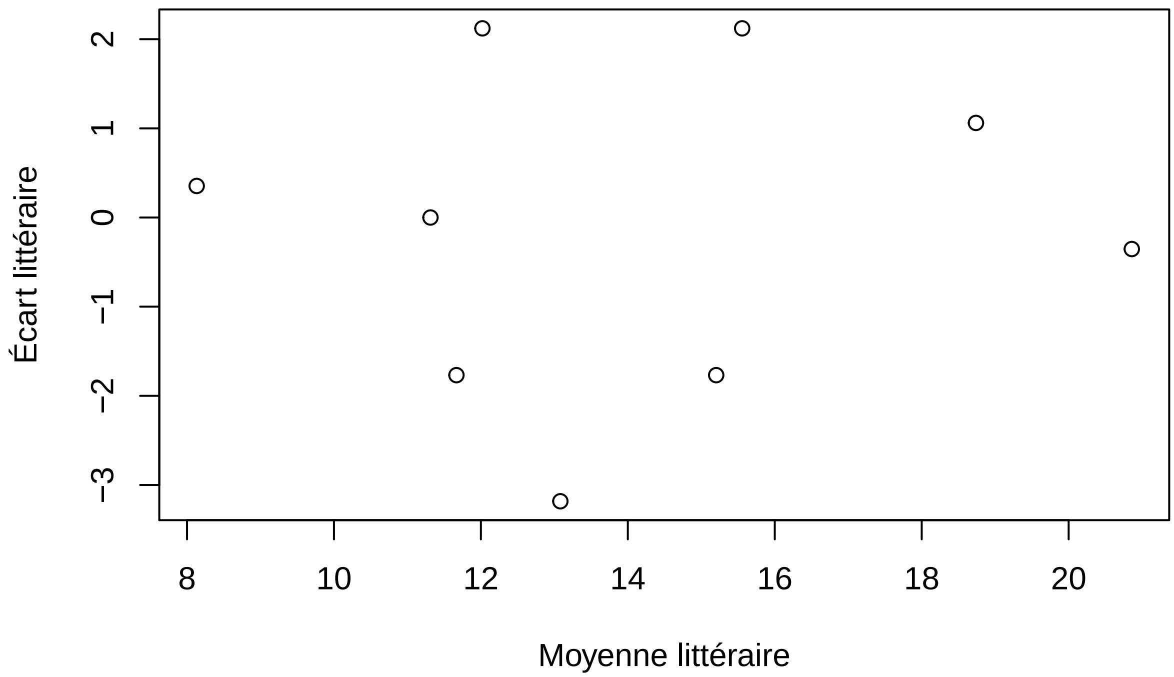
```r
plot(XB2[,1], XB2[,2], xlab = "Moyenne scientifique", ylab = "Moyenne littéraire")
```




```r
plot(XB2[,1], XB2[,3], xlab = "Moyenne scientifique", ylab = "Écart scientifique")
```




```
plot(XB2[,2], XB2[,4], xlab = "Moyenne littéraire", ylab = "Écart littéraire")
```



1.3 Choix d'une représentation

Quantité d'inertie du nuage de points expliquée par chacun des axes, si l'on considère la base canonique de $\mathbb{R}^5$? Quelle est la quantité d'inertie totale du nuage de points ?

```
G = apply(X = X,FUN = mean, MARGIN = 2)
G
```

```
##      math      scie      fran      lati       d.m
##  9.666667  9.833333 10.222222 10.055556 11.000000
```

```
XCentre = X - G
XCentre
```

```
##           math       scie      fran      lati        d.m
## jean -3.666667 -5.0000000 -5.055556 -4.722222 -1.8333333
## alin -1.833333 -1.6666667 -3.000000 -2.055556 -1.2222222
## anni -4.222222 -2.8333333  1.333333 -1.500000  0.9444444
## moni  4.444444  4.2777778  5.666667  5.333333 -3.0000000
## didi  3.000000  3.9444444  1.777778  2.666667  0.3333333
## andr  1.333333 -1.0000000 -4.555556 -3.222222  3.1666667
## pier -4.333333 -2.6666667  3.000000  1.444444 -0.2222222
## brig  2.777778  2.6666667 -1.166667 -1.500000  1.9444444
## evel -1.055556 -0.7222222  2.666667  2.333333  7.0000000
```

```
XNorms = apply(X = XCentre, FUN = function(x) sum(x^2), MARGIN = 1)
XNorms
```

```
##      jean      alin      anni      moni      didi      andr      pier
##  89.66358  20.85802  30.77469 107.60802  34.94136  43.94136  37.02469
##      brig      evel
##  22.21914  63.19136
```

```
I_g = sum(XNorms) / (9 - 1)
I_g
```

```
## [1] 56.27778
```

On peut aussi refaire cela avec la matrix de covariance:

```
covMat = cov.wt(X)
sum(diag(covMat$cov)) # method = "ML" pour changer l'estimateur
```

```
## [1] 55.09722
```

```
inertia = function(X) {
  Xs = scale(X,center = TRUE, scale = FALSE)
  Xsum = apply(X = Xs, FUN = function(x) sum(x^2), MARGIN = 1)

  inert = mean(Xsum)
  inert
}
```

L'inertie est la même pour des bases orthonormées :

```
inertia(X)
```

```
## [1] 48.97531
```

```
inertia(XB1)
```

```
## [1] 48.97531
```

```
inertia(XB2)
```

```
## [1] 48.97531
```

2 Question Théorique

On considère un nuage de points de coordonnées X exprimées dans un espace euclidien. On suppose que les individus ont tous le même poids.

2.1 Relation entre inertie et trace de la matrice de covariance empirique

Question : *Montrer que l'inertie I_X d'un nuage de points de coordonnées X est égale à la trace de sa matrice de covariance empirique (non corrigée) Σ_X.*

On a:

$$(\Sigma_X)_{i,j} = X^\top X$$

$$(\Sigma_X)_{i,j} = \begin{cases} Cov(X_i, X_j) & \text{si } i \neq j \\ Var(X_i) & \text{sinon.} \end{cases}$$

Ainsi:

$$trace(\Sigma_X) = \sum_{j=1}^{n} (\Sigma_X)_{j,j} = \sum_{j=1}^{n} Var(X_j) = \frac{1}{n} \sum_{j=1}^{n} \ \|X_j\|_2^2 = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \ |x_{ij} - \bar{x}_j|_2^2$$

C'est à dire par inversion des sommes :

$$trace(\Sigma_X) = \frac{1}{n} \sum_{i=1}^{n} \ \|x_i - \bar{x}\|_2^2$$

Or puisque $\bar{x} = g$ avec g le centre de gravité de $\mathcal{N}_X$ on a:

$$trace(\Sigma_X) = \frac{1}{n} \sum_{i=0}^{n} \ d^2(x_i, g)$$

C'est à dire par la définition de l'inertie de $\mathcal{N}_X$:

$$trace(\Sigma_X) = I_X$$

2.2 Invariance de la trace par permutation circulaire

Question : *Montrer l'invariance de la trace par permutation circulaire, i.e. pour trois matrices A, B* et C : $trace(ABC) = trace(CAB)$.

Soit A, B et C trois matrices de $\mathbb{R}^{n\times n}$ avec $n \in \mathbb{N}^*$.

Soit i et j deux entiers naturels inférieurs à n, on a :

$$(ABC)_{i,j} = \underline{(AB)_i} C_j$$

Soit k un autre entier naturel inférieur à n, on a:

$$\underline{(AB)_i}_k = \underline{A}_i B_k = \sum_{l=1}^{n} A_{i,l} B_{l,k}$$

Ainsi:

$$(ABC)_{i,j} = \begin{bmatrix} \sum_l A_{i,l}B_{l,1} & \sum_l A_{i,l}B_{l,2} & \cdots & \sum_l A_{i,l}B_{l,n} \end{bmatrix} \begin{bmatrix} C_{1,j} \\ C_{2,j} \\ \vdots \\ C_{n,j} \end{bmatrix}$$

$$(ABC)_{i,j} = \sum_{k=1}^{n} C_{k,j} \sum_{l=1}^{n} [A_{i,l} B_{l,k}]$$

$$(ABC)_{i,j} = \sum_{k=1}^{n} \sum_{l=1}^{n} B_{l,k} A_{i,l} C_{k,j}$$

On a donc, de manière identique en identifiant les matrices:

$$(CAB)_{i,j} = \sum_{k=1}^{n} \sum_{l=1}^{n} A_{l,k} C_{i,l} B_{k,j}$$

De plus:

$$trace(ABC) = \sum_{i=1}^{n} (CAB)_{i,i} = \sum_{i=1}^{n} \sum_{l=1}^{n} \sum_{k=1}^{n} A_{i,l} C_{k,j} B_{l,k}$$

En effectuant un renommage des indices ($i \to l, l \to k, k \to i$) on a:

$$trace(ABC) = \sum_{i=1}^{n} (CAB)_{i,i} = \sum_{l=1}^{n} \sum_{k=1}^{n} \sum_{i=1}^{n} A_{l,k} C_{i,l} B_{k,j}$$

C'est à dire:

$$trace(ABC) = trace(CAB)$$

2.3 Préservation du centrage par changement de bases

Question : *Montrer que le centrage est préservé par la projection, c'est-à-dire que la projection $X_c B$ d'un nuage de points X_c centré sur une base B est centrée.*

Montrons que $X_c B$ est centré c'est à dire :

$$\forall j \in [1, n], \ \sum_{i=1}^{m} (X_c B)_{i,j} = 0$$

On a :

$$\sum_{i=0}^{m} \underline{(X_c B)}_i = \sum_{i=0}^{m} \left[\sum_{l} (X_c)_{i,l} B_{l,1} \ \sum_{l} (X_c)_{i,l} B_{l,2} \cdots \ \sum_{l} (X_c)_{i,l} B_{l,n} \right]$$

$$\sum_{i=0}^{m} \underline{(X_c B)}_i = \left[\sum_{i=0}^{m} \sum_{l} (X_c)_{i,l} B_{l,1} \ \sum_{i=0}^{m} \sum_{l} (X_c)_{i,l} B_{l,2} \cdots \ \sum_{i=0}^{m} \sum_{l} (X_c)_{i,l} B_{l,n} \right]$$

Or X_c est centré, on a donc:

$$\forall j \in [1, n], \ \sum_{i=1}^{m} (X_c)_{i,j} = 0$$

Ainsi:

$$\sum_{i=0}^{m} \underline{(X_c B)}_i = \left[\sum_{i=0}^{m} 0 \ B_{l,1} \ \sum_{i=0}^{m} 0 \ B_{l,2} \cdots \ \sum_{i=0}^{m} 0 \ B_{l,n} \right] = \mathbf{0}^{\top}$$

C'est à dire $X_c B$ est centré.

2.4 Préservation de l'inertie du nuage de points par changement de bases orthonormées

Question : *Montrer que l'inertie totale des points représentés dans la base B est constante, quelle que soit la base orthonormée B considérée.*

Soit B_1 et B_2 deux bases orthonormées. Soit x et y deux vecteur exprimés dans B_1. On peut représenter x et y dans B_2 comme $B_2^{\top} x$ et $B_2^{\top} y$.

On peut facilement montrer que le changements de bases préserve les distances ; on a:

$d(B_2^\top x, B_2^\top y) = (B_2^\top x)^\top B_2^\top y = x^\top B_2 B_2^\top y = x^\top I_n y = x^\top y = d(x, y)$

Soit g le centre de gravité de exprimé dans la base B_1 ; $B_2^\top g$ est sa représentation dans B_2. On a ainsi:

$$\sum_{i=1}^{n} d^2(x_i, g) = \sum_{i=1}^{n} d^2(B_2^\top x_i, B_2^\top g)$$

C'est à dire que l'inertie est constante quelque soit les bases orthonormées considérées.