Multiple Factor Analysis

Julie Josse

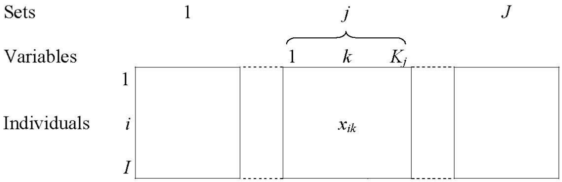
Applied Mathematics Department, Agrocampus Ouest

Stat 300
Stanford, July 2015

Multiple Factor Analysis

1. Data - Issues
2. Common Structure
3. Groups Study
4. Partial Analyses
5. Example

Multi-blocks data set



- Sensory analysis: products - sensorial, physico-chemical
- Survey: individuals - questionnaires themes (students health: addicted consumptions, psychological conditions, sleep, id)
- Economy: countries - economic indicators each year
- Biology: samples - Omics data (brain tumors: CGH, transcriptome; mouse: transcriptome, hepatic fatty acid measurements)

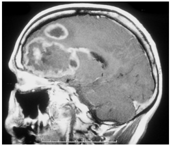
⇒ Generalized Canonical Correlation, Procrustes, Statis, etc.

⇒ MFA (Escofier & Pagès, 1998)

⇒ Continuous / categorical / contingency sets of variables

Example: gliomas brain tumors

Gliomas: Brain tumors, WHO classification

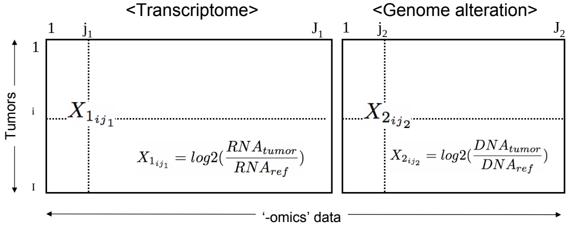


- astrocytoma (A) x5
- oligodendroglioma (O) x8
- oligo-astrocytoma (OA) x6
- glioblastoma (GBM) x24

43 tumor samples

(Bredel *et al.*,2005)

- Transcriptional modification (RNA), microarrays: 489 variables
- Damage to DNA (CGH array): 113 variables



Objectives

- Study the similarities between individuals with respect to all the variables
- Study the linear relationships between variables

$\Rightarrow$ taking into account the structure on the data (balance the influence of each group)

- Find the common structure with respect to all the groups - highlight the specificities of each group
- Compare the typologies obtained from each group of variables (separate analyses)

Principal component methods

The core of principal component methods is PCA on particular matrices

"Doing a data analysis, in good mathematics, is simply searching eigenvectors, all the science of it (the art) is just to find the right matrix to diagonalize"
Benzécri

MFA is a particular weighted PCA!

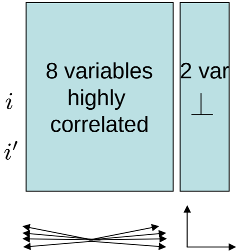
Balancing the groups of variables

MFA is a weighted PCA:

- compute the first eigenvalue λ_1^j of each group of variables
- perform a global PCA on the weighted data table:

$$\left[\frac{X_1}{\sqrt{\lambda_1^1}}; \frac{X_2}{\sqrt{\lambda_1^2}}; \ldots; \frac{X_J}{\sqrt{\lambda_1^J}} \right]$$

$\Rightarrow$ Same idea as in PCA when variables are standardized: variables are weighted to compute distances between individuals i and i'



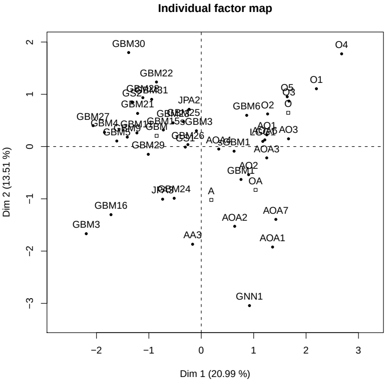
Balancing the groups of variables

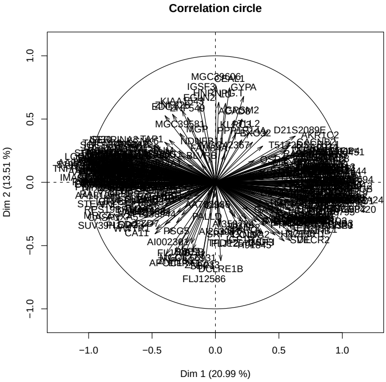
	Transcriptome	Genome
λ_1	162	12
λ_2	35	10
λ_3	21	5

This weighting allows that:

- Same weight for all the variables of one group: the structure of the group is preserved
- For each group the variance of the main dimension of variability (first eigenvalue) is equal to 1
- No group can generate by itself the first global dimension
- A multidimensional group will contribute to the construction of more dimensions than a one-dimensional group

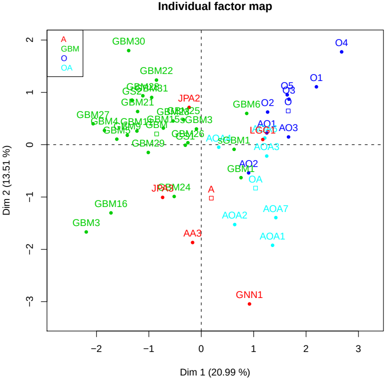
Individuals and variables representations

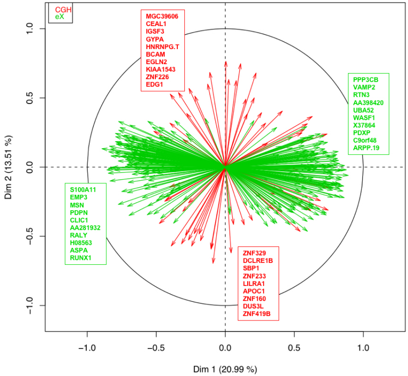




⇒ What can be added to interpret?

Individuals and variables representations





$\Rightarrow$ Do the means of the tumors coordinates per stage on each dimension significantly differ from each other?

Groups study

⇒ Synthetic comparison of the groups

⇒ Are the relative positions of individuals globally similar from one group to another? Are the partial clouds similar?

⇒ Do the groups bring the same information?

Principal component in MFA

MFA = weighted PCA $\Rightarrow$ first principal component of MFA maximizes

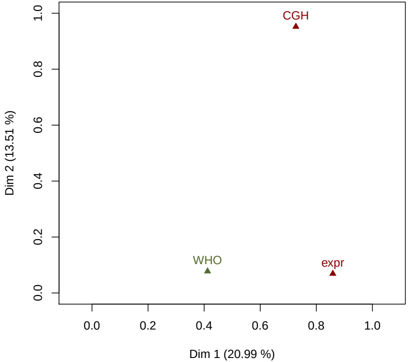
$$\sum_{j=1}^{J} \sum_{k \in K_j} cov^2 \left(\frac{x_{.k}}{\sqrt{\lambda_1^j}}, F_1 \right) = \sum_{j=1}^{J} \mathcal{L}_g(F_1, K_j)$$

$$\mathcal{L}_g(F_1, K_j) = < \frac{W_{K_j}}{\lambda_1}, F_1 F_1' > = \text{trace}(W_{K_j}' F_1 F_1')$$

$\Rightarrow$ F_1 the most related to the groups in the $\mathcal{L}_g$ sense

Representation of the groups

Group j has the coordinates $(\mathcal{L}_g(F_1, K_j), \mathcal{L}_g(F_2, K_j))$



- 2 groups are all the more close that they induce the same structure
- The 1st dimension is common to all the groups
- 2nd dimension mainly due to CGH

$$0 \leq \mathcal{L}_g(F_1, K_j) = \frac{1}{\lambda_1^j} \underbrace{\sum_{k \in K_j} cov^2(x_{.k}, F_1)}_{\leq \lambda_1^j} \leq 1$$

$\Rightarrow$ Could you predict the results of the PCA for each group?

The RV coefficient

$X_{j_{(I \times K_j)}}$ and $X_{m_{(I \times K_m)}}$ not directly comparable
$W_{j_{(I \times I)}} = X_j X_j'$ and $W_{m_{(I \times I)}} = X_m X_m'$ can be compared
Inner product matrices = relative position of the individuals

Covariance between two groups:

$$< W_j, W_m >= \sum_{k \in K_j} \sum_{l \in K_m} cov^2(x_{.k}, x_{.l})$$

Correlation between two groups (Escoufier, 1973):

$$RV(K_j, K_m) = \frac{< W_j, W_m >}{\|W_j\| \; \|W_m\|} \qquad 0 \leq RV \leq 1$$

$RV = 0$: variables of K_j are uncorrelated with variables of K_m
$RV = 1$: the two clouds of points are homothetic

$\Rightarrow$ Extension of the notion of correlation matrix

Similarity between two groups

Measure of similarity between groups K_j and K_m:

$$\mathcal{L}_g(K_j, K_m) = \sum_{k \in K_j} \sum_{l \in K_m} cov^2 \left(\frac{x_{.k}}{\lambda_1^k}, \frac{x_{.l}}{\lambda_1^l} \right)$$

Ramsay (1984): "Matrices may be similar or dissimilar in a many ways"

Canonical correlation (Hotteling, 1936), Mantel (1967), Procrustes (Gower, 1971), dCov (Szekely et al., 2007), kernel based HSIC (Gretton et al., 2005), etc...

Numeric indicators

```
> res.mfa$group$Lg
      CGH expr  WHO
CGH  2.51 0.60 0.46
expr 0.60 1.10 0.36
WHO  0.46 0.36 0.50
```

$$\mathcal{L}_g(K_j, K_j) = \frac{\sum_{k=1}^{K_j} (\lambda_k^j)^2}{(\lambda_1^j)^2} = 1 + \frac{\sum_{k=2}^{K_j} (\lambda_k^j)^2}{(\lambda_1^j)^2}$$

```
> res.mfa$group$RV
      CGH expr  WHO
CGH  1.00 0.36 0.41
expr 0.36 1.00 0.48
WHO  0.41 0.48 1.00
```

- CGH gives richer description ($\mathcal{L}_g$ greater)
- RV: a standardized $\mathcal{L}_g$
- CGH and expr are not linked (RV=0.36)

Contribution of each group to each component of the MFA

```
> res.mfa$group$contrib
     Dim.1 Dim.2 Dim.3
CGH   45.8  93.3  78.1
expr  54.2   6.7  21.9
```

- Similar contribution of the 2 groups to the first dimension
- Second dimension only due to CGH

Partial analyses

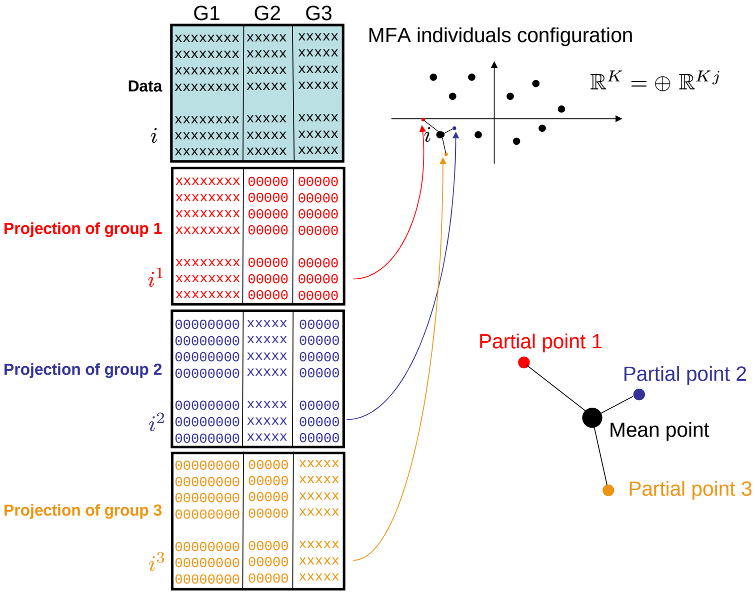
⇒ Comparison of the groups through the individuals

⇒ Comparison of the typologies provided by each group in a common space

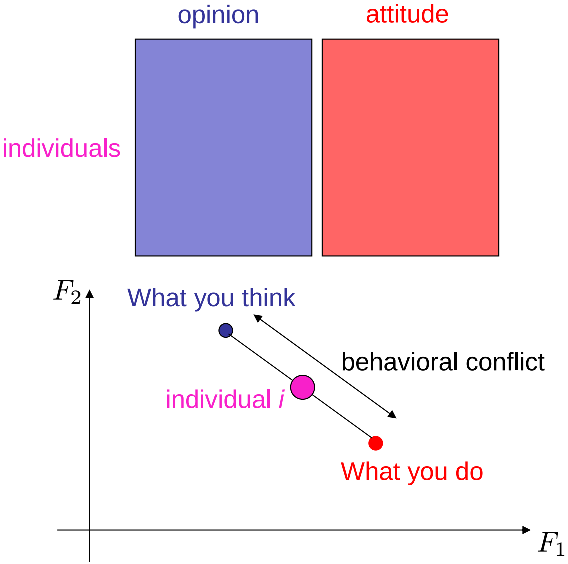
⇒ Are there individuals very particular with respect to one group?

⇒ Comparison of the separate PCA

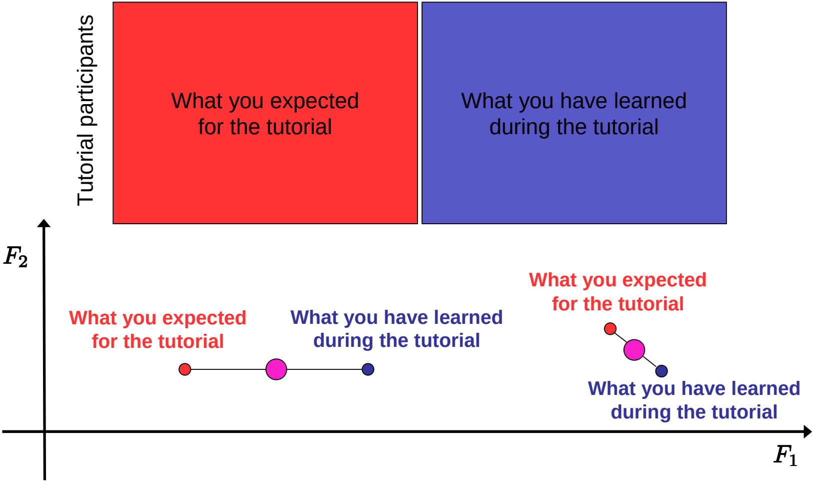
Projection of partial points



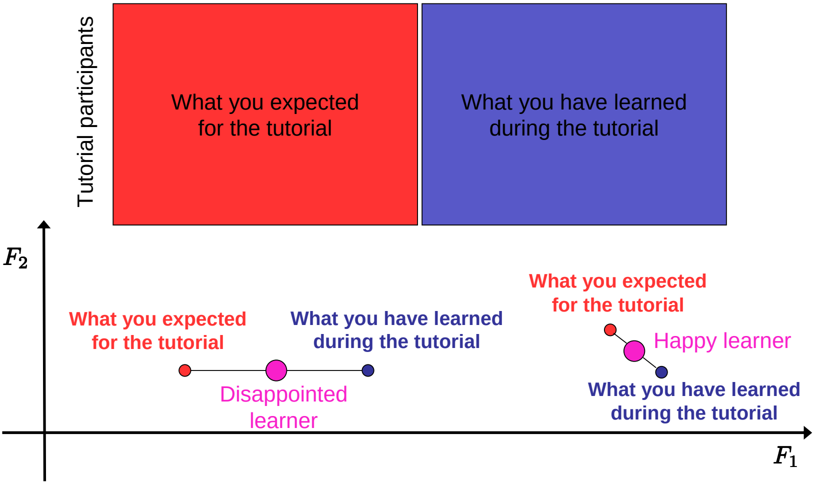
Partial points



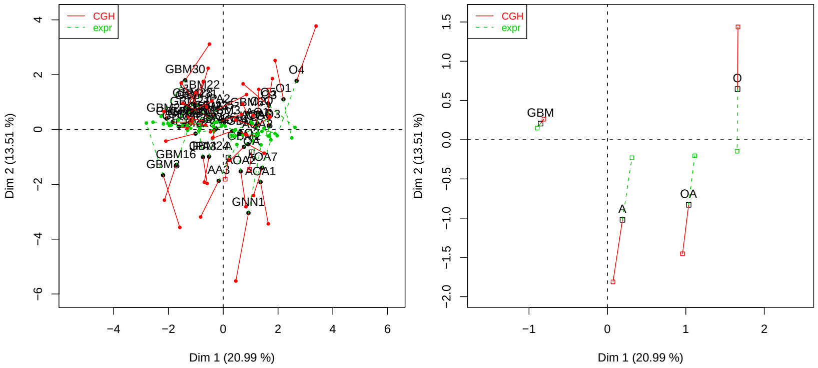
Partial points



Partial points

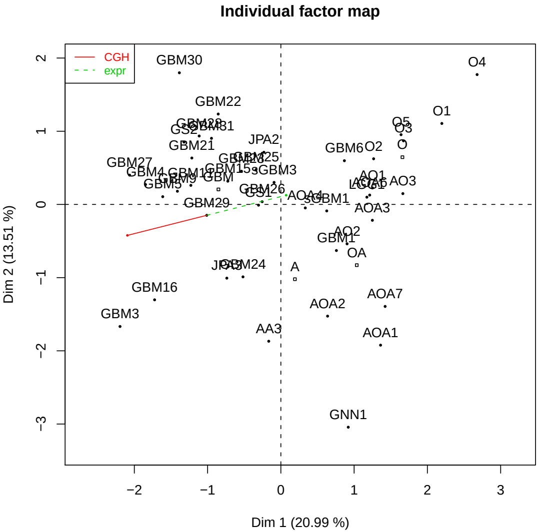


Representation of the partial points



- an individual is at the barycentre of its partial points
- an individual is all the more "homogeneous" that its superposed representations are close

Identify particular individuals



Numeric indicators

$$\sum_{i=1}^{I}\sum_{j=1}^{J}(F_{i^j q})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}(F_{iq})^2 + \sum_{i=1}^{I}\sum_{j=1}^{J}(F_{i^j q} - F_{iq})^2$$

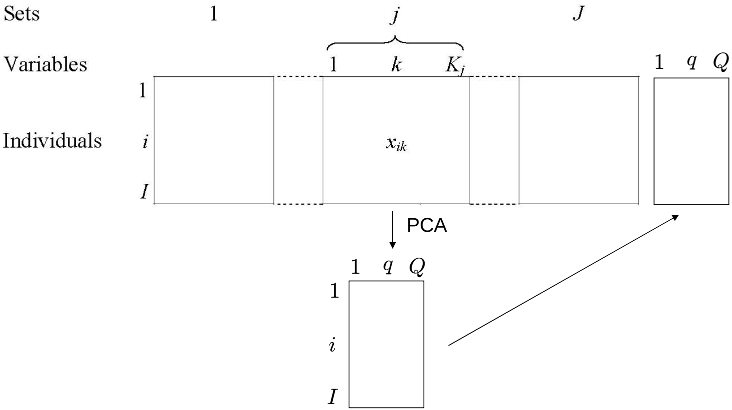
Total inertia = Between indiv. inertia + Within indiv. inertia

```
> res.mfa$inertia.ratio
Dim.1 Dim.2 Dim.3 Dim.4 Dim.5
 0.84  0.56  0.44  0.59  0.43
```

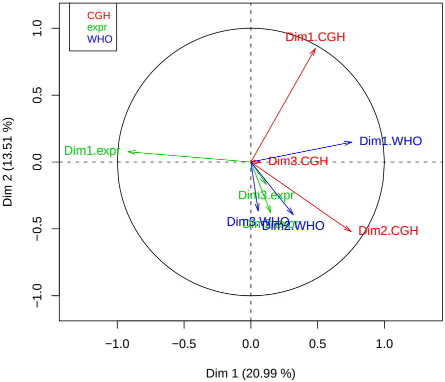
- For the first dimension, the coordinates of each partial points are close (0.84 close to 1)
- The within inertia can be decomposed by individuals `res.mfa$ind$within.inertia`

Representation of the partial components

Do the separate analyses give similar dimensions as MFA?



Representation of the partial components



- The first dimension of each group is well projected
- CGH has same dimensions as MFA

Use of biological knowledge

Genes can be grouped by gene ontology (GO) biological process

GO:0006928
cell motility

ANXA1
CALD1
EGFR
ENPP2
FN1
FPRL2
LSP1
MSN
PDPN
PLAUR
PRSS3
SAA2
SPINT2
TNFRSF12A
VEGF
WASF1
YARS

GO:0009966
regulation of signal transduction

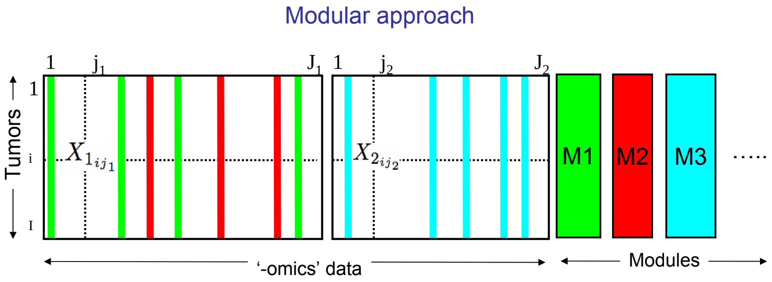
CASP1
EDG2
F2R
HCLS1
HMOX1
IGFBP3
IQSEC1
LYN
MALT1
TCF7L1
TNFAIP3
TRIO
VEGF
YWHAG
YWHAH

GO:0052276
chromosome organisation and biogenesis

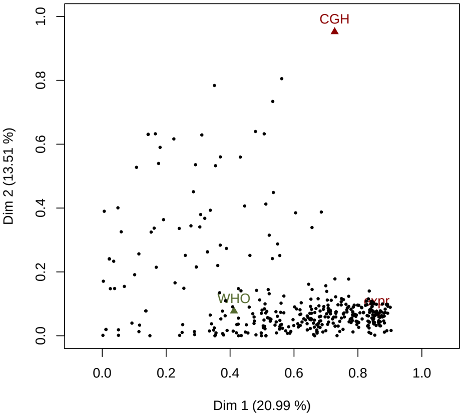
CBX6
NUSAP1
PCOLN3
PTTG1
SUV39H1
TCF7L1
TSPYL1

Use of biological knowledge

- Biological processes considered as supplementary groups of variables



Use of biological knowledge

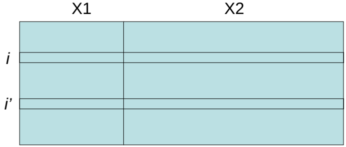


Many biological processes induce the same structure on the individuals than MFA

To go further

- Mixed data: MFA with 1 group = 1 variable
  continuous variables: PCA is recovered; categorical variables: MCA is recovered
  mixed: FAMD
- MFA used for methodological purposes:
  - comparison of coding (continuous or categorical)
  - comparison between preprocessing (standardized PCA and unstandardized PCA)
  - comparison of results from different analyses
- Hierarchical Multiple Factor Analysis:
  Takes into account a hierarchy on the variables: variables are grouped and subgrouped (like in questionnaires structured in topics and subtopics)

Clustering: MFA as a preprocessing



MFA balances the influence of the groups when computing distances between individuals

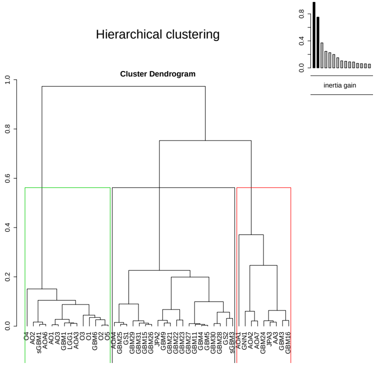
$$d^2(i,i') = \sum_{j=1}^{J} \frac{1}{\sqrt{\lambda_j}} \sum_{k=1}^{K_j} (x_{ik} - x_{i'k})^2$$

AHC or k-means onto the first principal components $(F_{.1}, ..., F_{.Q})$ obtained from MFA allows to

- take into account the groups structure in the clustering
- make the clustering more robust by deleting the last dimensions

Clustering

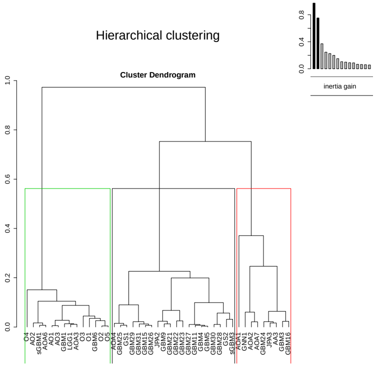
AHC onto the first 5 principal components from MFA



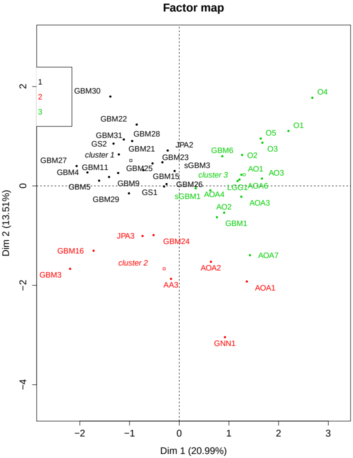
Individuals are sorted according to their coordinate on $F_{.1}$

Partition from the tree

An empirical number of clusters is suggested



Partition on the principal component map



Continuous vision (principal component) and discontinuous (clusters)

Cluster description by variables

$$\text{v.test} = \frac{\bar{x}_\ell - \bar{x}}{\sqrt{\frac{s^2}{I_\ell}\left(\frac{I-I_\ell}{I-1}\right)}}$$ H_0 : random sampling of I_ℓ values from I

with $\bar{x}_\ell$ the mean of variable x in cluster ℓ, $\bar{x}$ (s) the mean (standard deviation) of the variable x in the data set, I_ℓ the cardinal of cluster ℓ

```
$desc.var$quanti$‘1‘
             v.test Mean in category Overall mean sd in category Overall sd p.value
TMEM49        4.488           -0.430       -1.424          0.722      1.277   0.000
TNFRSF12A     4.433           -0.794       -1.838          0.789      1.357   0.000
LGALS3        4.369           -0.222       -1.216          0.861      1.312   0.000
S100A11       4.300           -0.737       -1.500          0.525      1.024   0.000
BGN           4.273            2.105        1.106          0.697      1.348   0.000
IFI30         4.264            0.987        0.026          0.979      1.300   0.000
....
....
C9orf48      -4.411           -0.686       -0.037          0.540      0.848   0.000
PSD3         -4.594           -1.684       -1.024          0.419      0.829   0.000
AA398420     -4.635            0.324        1.134          0.635      1.007   0.000
```

Cluster description by observations

- parangon: the closest observations to the centroid of the cluster

$$\min_{i \in \ell} d(x_{i.}, C_\ell) \text{ with } C_\ell \text{ the centroid of cluster } \ell$$

- specific observations: the furthest observations to the centroids of the other clusters (the observations sorted according to their distance from the highest to the smallest to the closest centroid)

$$\max_{i \in \ell} \min_{\ell' \neq \ell} d(x_{i.}, C_{\ell'})$$

```
desc.ind$para
cluster: 1
    GBM11     GBM28      GBM5     GBM25     GBM31
0.6649847 0.7001998 0.7973604 0.8869271 0.9674042
---------------------------------------------------------------
desc.ind$dist
cluster: 1
  GBM30      GS2    GBM21    GBM22    GBM27
3.227968 3.096048 3.031256 2.904327 2.778950
---------------------------------------------------------------
```

Cluster description

- by the principal components (observations coordinates): same description than for continuous variables

```
$desc.axes$quanti$‘1‘
       v.test Mean in category Overall mean sd in category Overall sd p.value
Dim.2  2.919             0.511            0           0.465      1.010   0.004
Dim.1 -4.458            -0.974            0           0.560      1.259   0.000
```

- by categorical variables: chi-square and hypergeometric test

```
$test.chi2
   p.value df
type 8.433474e-06  6
```

⇒ Active and supplementary elements are used
⇒ Only significant results are presented

Cluster description

```
$`1`
         Cla/Mod  Mod/Cla   Global       p.value    v.test
type=GBM      75 94.73684 55.81395 3.300966e-06  4.651145
type=OA        0  0.00000 13.95349 2.207775e-02 -2.289028
type=O         0  0.00000 18.60465 5.071916e-03 -2.802430

$`2`
       Cla/Mod Mod/Cla   Global   p.value   v.test
type=A      60    37.5 11.62791 0.0398214 2.055597

$`3`
         Cla/Mod Mod/Cla   Global       p.value    v.test
type=O     100.0   50.00 18.60465 8.875341e-05  3.919444
type=GBM    12.5   18.75 55.81395 2.319983e-04 -3.681354
```

Complementarity between hierarchical clustering and partitioning

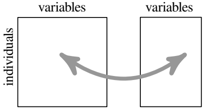
- Partitioning after AHC: the k-means algorithm is initialized from the centroids of the partition obtained from the tree
  - consolidate the partition
  - loss of the hierarchy

- AHC with many individuals: time-consuming
  $\Rightarrow$ partitioning before AHC
  - compute k-means with approximately 100 clusters
  - AHC on the weighted centroids obtained from the k-means
    $\Rightarrow$ top of the tree is approximately the same

Other methods: ade4

Two-table analysis

Available methods

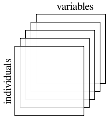


Function name	Analysis name
between	Between-class analysis
within	Within-class analysis
discrimin	Discriminant analysis
coinertia	Coinertia analysis
cca	Canonical correspondence analysis
pcaiv	PCA on Instrumental Variables
pcaivortho	Orthogonal PCAIV
procuste	Procustes analysis
niche	Niche (OMI) analysis

Other methods: ade4

Other functions

K-table



Function name	Analysis name
sepan	K-table separate analyses
pta	Partial triadic analysis
foucart	Foucart analysis
statis	STATIS analysis
mfa	Multiple factor analysis
mcoa	Multiple coinertia analysis
statico	2 K-table analysis

Other methods

Predict one block with others:

- Multi-block PLS regression
- Multi-block PCA on instrumental variables..

RV Tests

Is there any (linear) relationship between the 2 sets? $H_0 : \rho V = 0$

Asymptotic tests: distributions normal, elliptical - rank (Robert *et al*, 1985, Cléroux, 1995, Cléroux & Ducharme, 1989) $n\text{RV} \sim \sum \lambda_i Z_i^2$
$\Rightarrow$ sensitive to the departure from the distribution and to n

RV Tests

Is there any (linear) relationship between the 2 sets? $H_0 : \rho V = 0$

Asymptotic tests: distributions normal, elliptical - rank (Robert *et al*, 1985, Cléroux, 1995, Cléroux & Ducharme, 1989) $n\text{RV} \sim \sum \lambda_i Z_i^2$
$\Rightarrow$ sensitive to the departure from the distribution and to n

Permutation tests:
permute one matrix's rows - compute the RV for $n!$ permutations
p-value: proportion of the values greater than the observed one
$\Rightarrow$ computationally costly (“old fashion" argument?)

RV Tests

Is there any (linear) relationship between the 2 sets? $H_0 : \rho V = 0$

Asymptotic tests: distributions normal, elliptical - rank (Robert *et al*, 1985, Cléroux, 1995, Cléroux & Ducharme, 1989) $n\text{RV} \sim \sum \lambda_i Z_i^2$
$\Rightarrow$ sensitive to the departure from the distribution and to n

Permutation tests:
permute one matrix's rows - compute the RV for $n!$ permutations
p-value: proportion of the values greater than the observed one
$\Rightarrow$ computationally costly (“old fashion" argument?)

Approximation of the permutation distribution

- sampling from the permutations - package ade4 (RV.rtest)
- moment matching: Pearson family, Edgeworth expansion

Moments matching

The first three moments under H_0 (Kazi-Aoual *et al.*, 1995)

$$\mathbb{E}_{H_0}(RV) = \frac{\sqrt{\beta_x \times \beta_y}}{n-1} \qquad \beta_x = \frac{(tr(X'X))^2}{tr((X'X)^2)} = \frac{(\sum \lambda_i)^2}{\sum \lambda_i^2}.$$

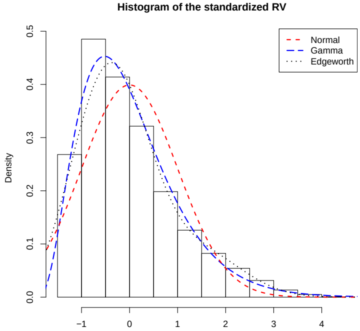
β_x a measure of complexity $1 \leq \beta_x \leq p$
RV large: n small and many orthogonal variables per group

$\Rightarrow$ Normal approximation:

$$\mathrm{RV}_{\mathrm{std}} = \frac{RV - \mathbb{E}_{H_0}(RV)}{\sqrt{\mathbb{V}_{H_0}(RV)}}$$

Moments matching

Problem: the exact distribution of the RV_{std} is often skewed



$\Rightarrow$ Pearson type III

f(x) (skewness= γ):

$$\frac{(2/\gamma)^{4/\gamma^2}}{\Gamma(4/\gamma^2)} \left(\frac{2+\gamma x}{\gamma}\right)^{(4-\gamma^2)/\gamma^2} e^{-2(2+x\gamma)/\gamma^2}$$

$\Rightarrow$ package FactoMineR (coeffRV) (Josse *et al.*, 2008)

Back to the wine example!

- 10 white wines from Val de Loire (5 Vouvray - 5 Sauvignon)
- 27 continuous variables: sensory descriptors

	O.fruity	O.passion	O.citrus	...	Sweetness	Acidity	Bitterness	Astringency	Aroma.intensity	Aroma.persistency	Visual.intensity	Odor.preferene	Overall.preference	Label
S Michaud	4.3	2.4	5.7	...	3.5	5.9	4.1	1.4	7.1	6.7	5.0	6.0	5.0	Sauvignon
S Renaudie	4.4	3.1	5.3	...	3.3	6.8	3.8	2.3	7.2	6.6	3.4	5.4	5.5	Sauvignon
S Trotignon	5.1	4.0	5.3	...	3.0	6.1	4.1	2.4	6.1	6.1	3.0	5.0	5.5	Sauvignon
S Buisse Domaine	4.3	2.4	3.6	...	3.9	5.6	2.5	3.0	4.9	5.1	4.1	5.3	4.6	Sauvignon
S Buisse Cristal	5.6	3.1	3.5	...	3.4	6.6	5.0	3.1	6.1	5.1	3.6	6.1	5.0	Sauvignon
V Aub Silex	3.9	0.7	3.3	...	7.9	4.4	3.0	2.4	5.9	5.6	4.0	5.0	5.5	Vouvray
V Aub Marigny	2.1	0.7	1.0	...	3.5	6.4	5.0	4.0	6.3	6.7	6.0	5.1	4.1	Vouvray
V Font Domaine	5.1	0.5	2.5	...	3.0	5.7	4.0	2.5	6.7	6.3	6.4	4.4	5.1	Vouvray
V Font Brûlés	5.1	0.8	3.8	...	3.9	5.4	4.0	3.1	7.0	6.1	7.4	4.4	6.4	Vouvray
V Font Coteaux	4.1	0.9	2.7	...	3.8	5.1	4.3	4.3	7.3	6.6	6.3	6.0	5.7	Vouvray

Back to the wine example!

- 3 panels (oenologists, naive consumers, our students!)
- 60 preference scores: taste evaluation 1 - 10

Continuous variables / Categorical

	Expert (27)	Consumer (15)	Student (15)	Preference (60)	Label (1)
wine 1					
wine 2					
...					
wine 10					

- How are the products described by the panels?
- Do the panels describe the products in a same way? Is there a specific description done by one panel?

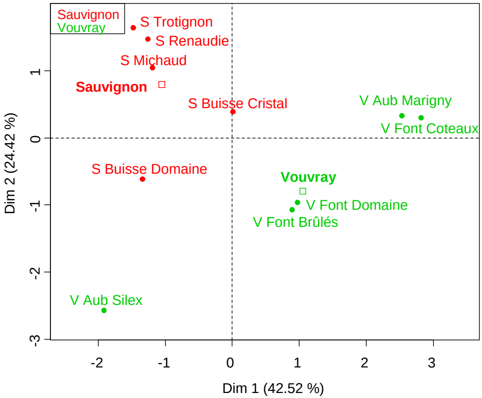
Practice with R

1. Define groups of active and supplementary variables
2. Scale or not the variables
3. Perform MFA
4. Choose the number of dimensions to interpret
5. Simultaneously interpret the individuals and variables graphs
6. Study the groups of variables
7. Study the partial representations
8. Use indicators to enrich the interpretation

Practice with R

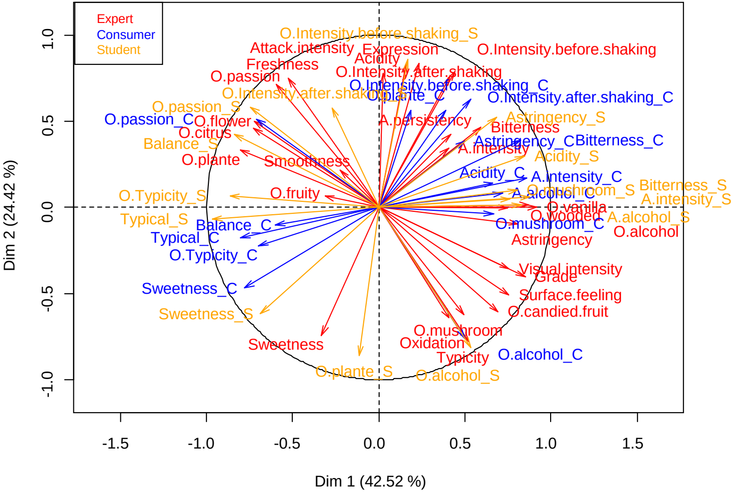
```
library(FactoMineR)
Expert <- read.table("http://factominer.free.fr/docs/Expert_wine.csv",
   header=TRUE, sep=";", row.names=1)
Consu <- read.table(".../Consumer_wine.csv",header=T,sep=";",row.names=1)
Stud <- read.table(".../Student_wine.csv",header=T,sep=";",row.names=1)
Pref <- read.table(".../Preference_wine.csv",header=T,sep=";",row.names=1)

palette(c("black","red","blue","orange","darkgreen","maroon","darkviolet"))
complet <- cbind.data.frame(Expert[,1:28],Consu[,2:16],Stud[,2:16],Pref)
res.mfa <- MFA(complet,group=c(1,27,15,15,60),type=c("n",rep("s",4)),
   num.group.sup=c(1,5),graph=FALSE,
   name.group=c("Label","Expert","Consumer","Student","Preference"))
plot(res.mfa,choix="group",palette=palette())
plot(res.mfa,choix="var",invisible="quanti.sup",hab="group",palette=palette())
plot(res.mfa,choix="ind",partial="all",habillage="group",palette=palette())
plot(res.mfa,choix="axes",habillage="group",palette=palette())
dimdesc(res.mfa)
write.infile(res.mfa,file="my_wine_results.csv") #to export a list
```

Representation of the individuals

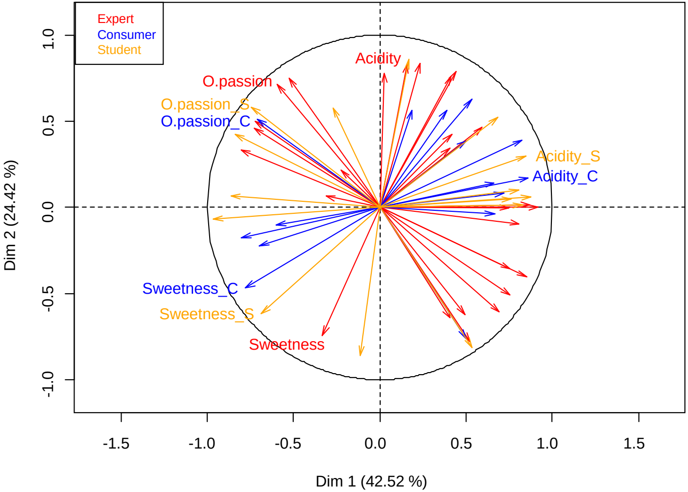


- The two labels are well separated
- Vouvray are sensorially more different
- Several groups of wines, ...

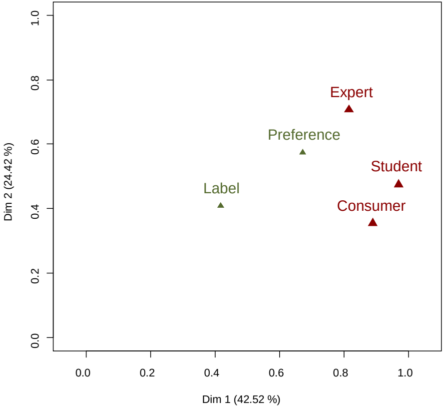
Representation of the active variables



Representation of the active variables

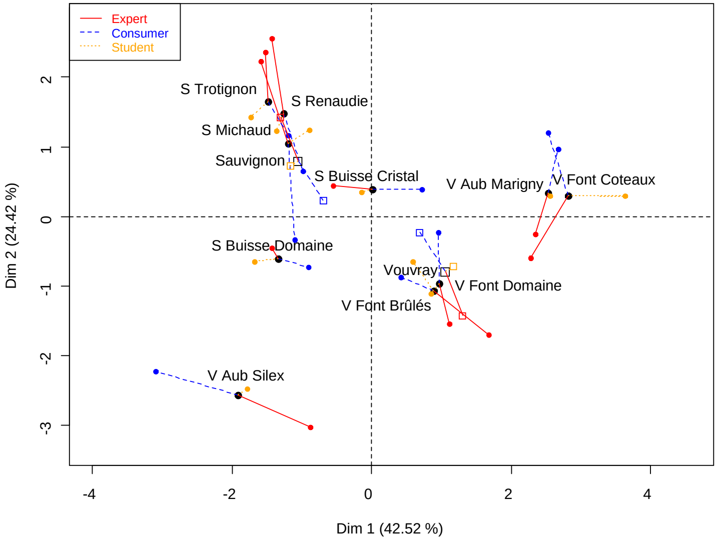


Representation of the groups

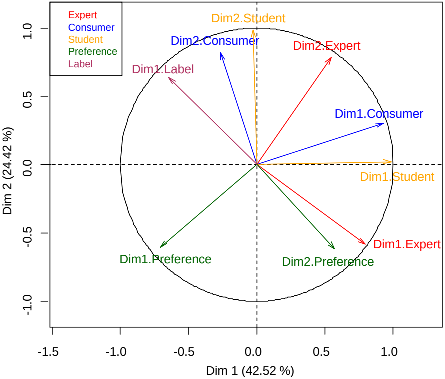


- 2 groups are all the more close that they induce the same structure
- The 1st dimension is common to all the panels
- 2nd dimension mainly due to the experts
- Preference linked to sensory description

Representation of the partial points

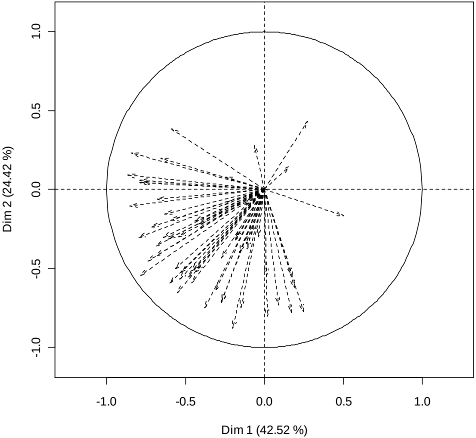


Representation of the partial dimensions



- The two first dimensions of each group are well projected
- Consumer has same dimensions as MFA

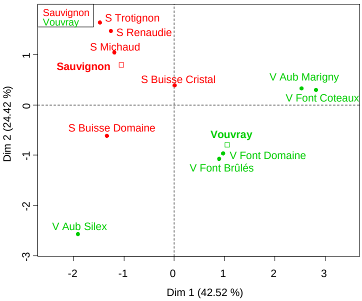
Representation of supplementary continuous variables

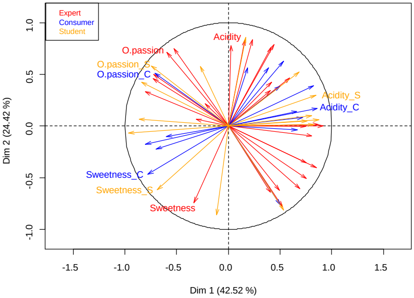


⇒ Preferences do not participated to the construction of the dimensions

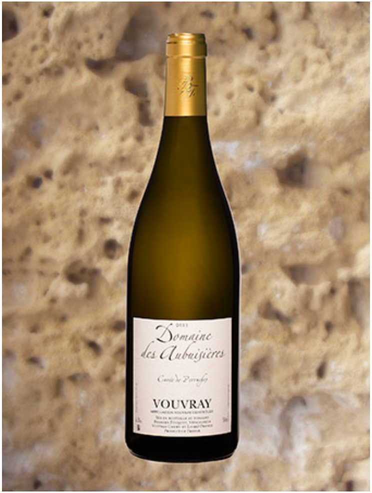
⇒ Preferences are linked to sensory description

Representation of supplementary continuous variables





Representation of supplementary continuous variables



⇒ Main information: the favourite is *Vouvray Aubuisières Silex*

Helps to interpret

- Contribution of each group of variables to each component of the MFA

```
> res.mfa$group$contrib
         Dim.1 Dim.2 Dim.3
Expert    30.5  46.0  33.7
Consumer  33.2  23.1  31.2
Student   36.3  30.9  35.1
```

- Similar contribution of the 3 groups to the first dimension
- Second dimension mainly due to the expert

- Correlation between the global cloud and each partial cloud

```
> res.mfa$group$correlation
         Dim.1 Dim.2 Dim.3
Expert    0.95  0.95  0.96
Consumer  0.95  0.83  0.87
Student   0.99  0.99  0.84
```

First components are highly linked to the 3 groups: the 3 clouds of points are nearly homothetic

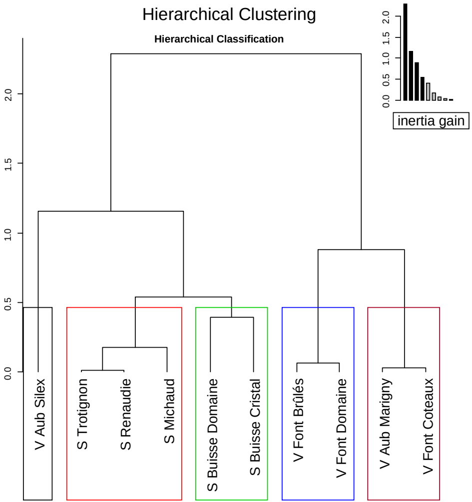
Similarity measures between groups

```
> res.mfa$group$Lg
           Expert Consumer Student Preference Label  MFA
Expert       1.45     0.94    1.17       1.01  0.89 1.33
Consumer     0.94     1.25    1.04       1.11  0.28 1.21
Student      1.17     1.04    1.29       1.03  0.62 1.31
Preference   1.01     1.11    1.03       1.47  0.37 1.18
Label        0.89     0.28    0.62       0.37  1.00 0.67
MFA          1.33     1.21    1.31       1.18  0.67 1.44

> res.mfa$group$RV
           Expert Consumer Student Preference Label  MFA
Expert       1.00     0.70    0.85       0.69  0.74 0.92
Consumer     0.70     1.00    0.82       0.82  0.25 0.90
Student      0.85     0.82    1.00       0.75  0.55 0.96
Preference   0.69     0.82    0.75       1.00  0.31 0.81
Label        0.74     0.25    0.55       0.31  1.00 0.56
MFA          0.92     0.90    0.96       0.81  0.56 1.00
```

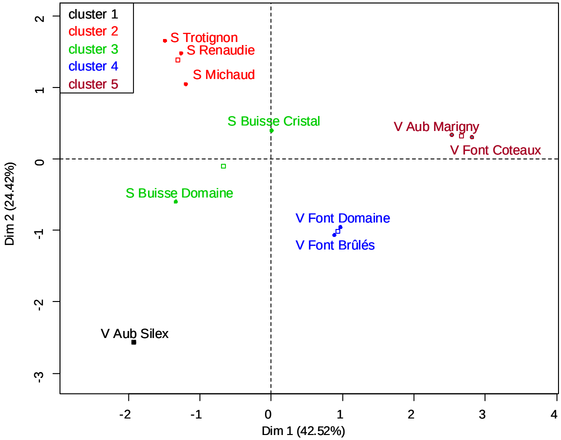
- Expert gives a richer description ($\mathcal{L}_g$ greater)
- Groups Student and Expert are linked (RV $=$ 0.85)
- Group Student is the closest to the overall (RV $=$ 0.96)

Partition from the tree

An empirical number of clusters is suggested $(\min_q \frac{W_q - W_{q+1}}{W_{q-1} - W_q})$



Partition on the principal component map



Continuous vision (principal component) and discontinuous (clusters)