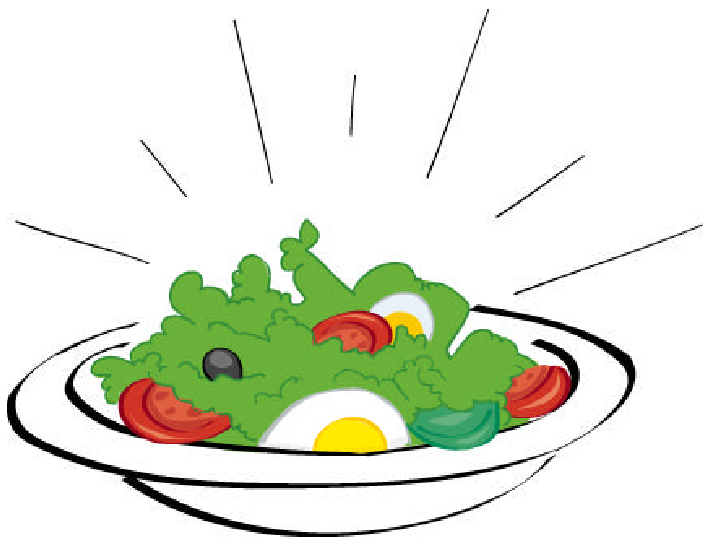
Analysis of sensory profiling data : from discrimination to description

M. Hanafi E. M. Qannari Ph. Courcoux

G. Organ

General context

- Sensory profiling data : **fixed vocabulary**



Pre-treatment of the data

Centre each data set in order to remove the judge effect.

Allow for an isotropic scaling factor in order to put the assessors on the same footing.

Usual methods of analysis

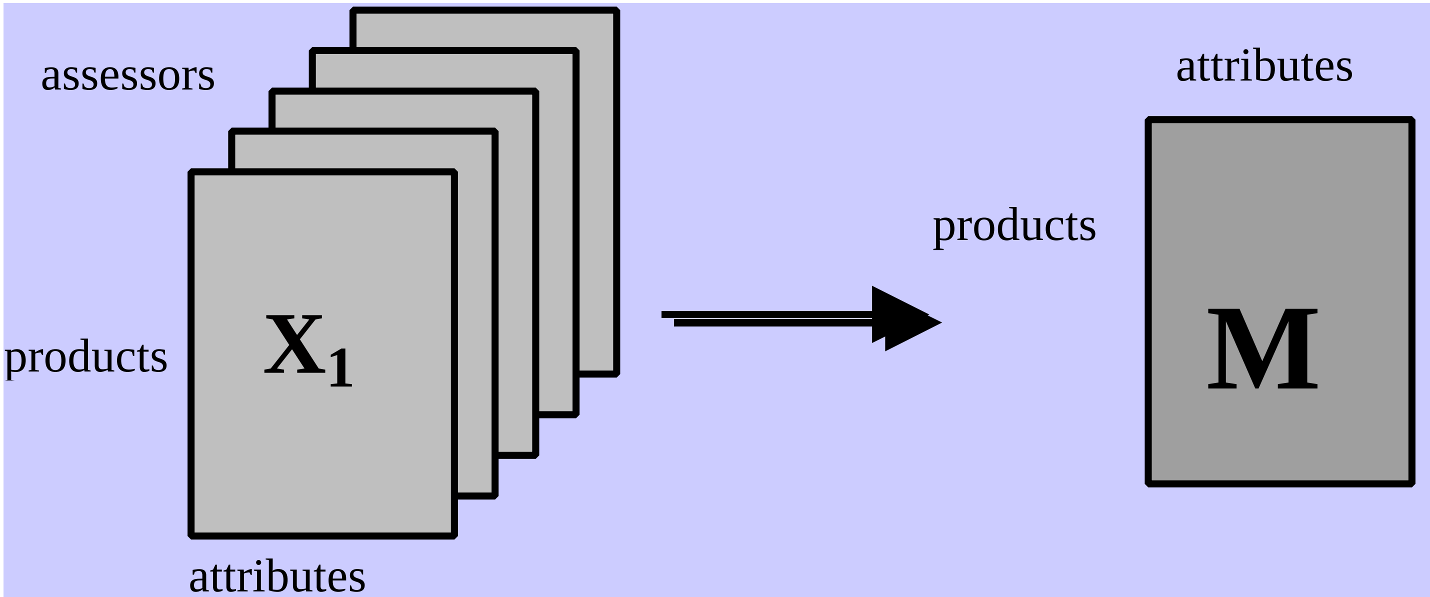
Discriminant analysis

PCA on the avarege data set

what do I know?

PCA on the merged data set

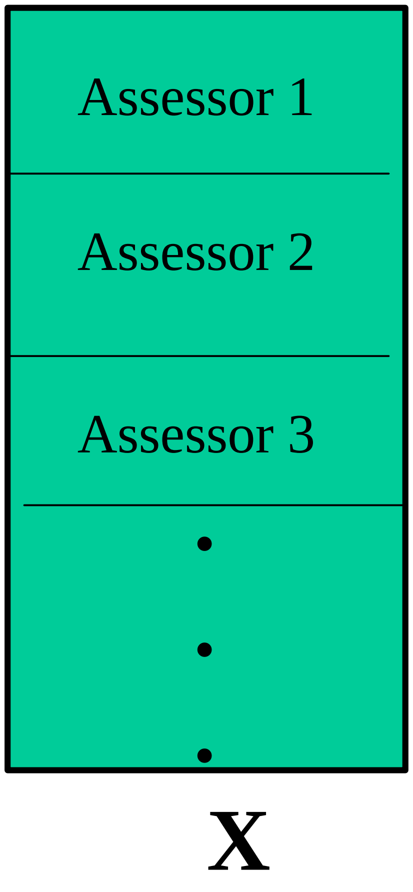
PCA on the average data set

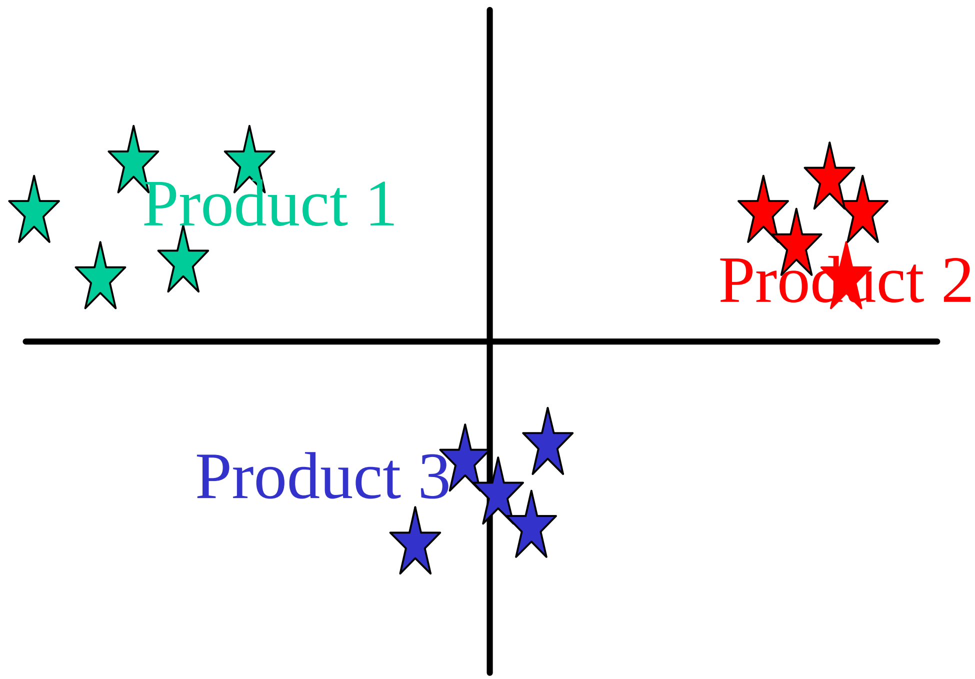


From a technical point of view

- The first principal component is a linear combination of the attributes with maximum variance : **best separation of the (average) products**.
- The loadings are obtained from the first eigenvector of
    - *$B=M^TM$*
  - ***Between groups variance-covariance matrix***

PCA on the whole data set

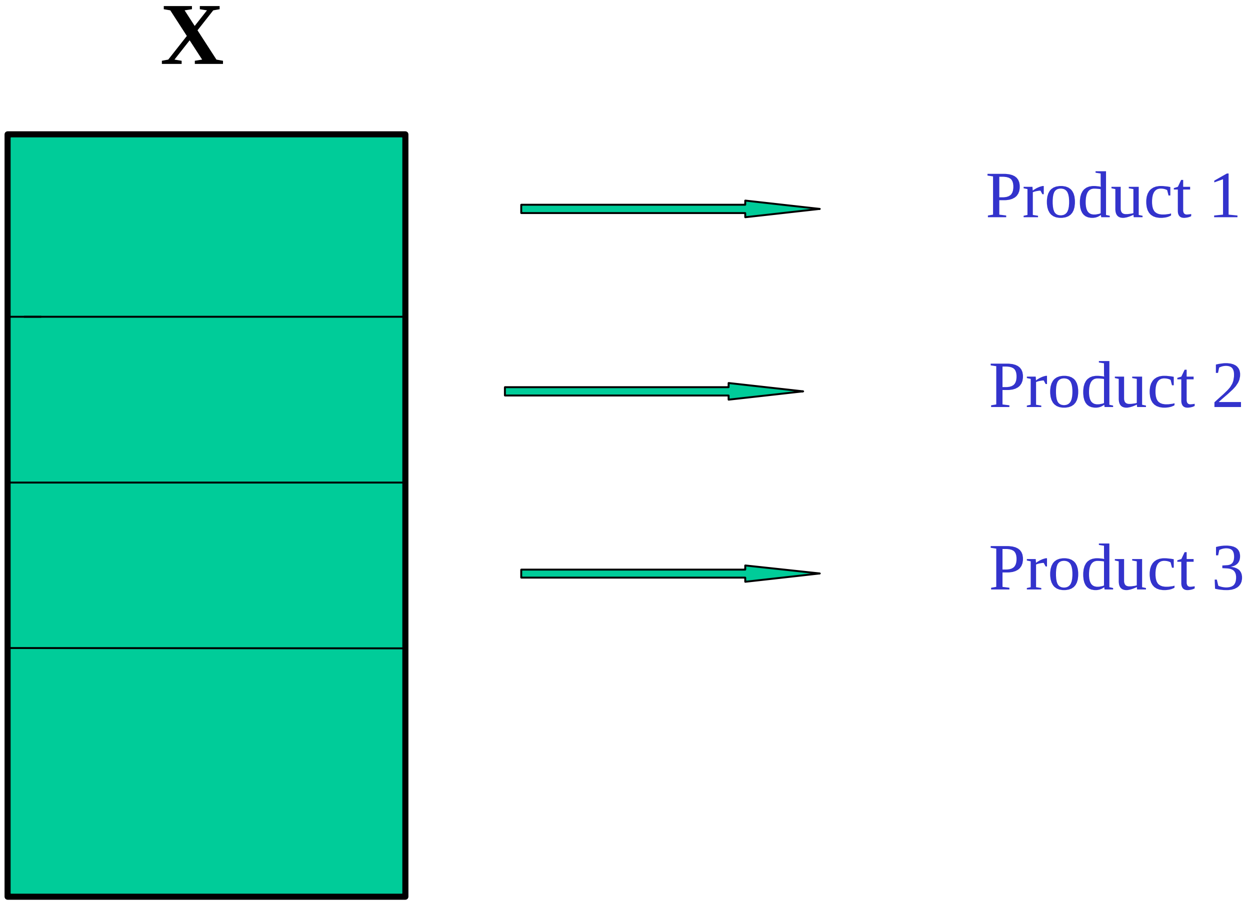




From a technical point of view

- The first principal component is a linear combination of the attributes with maximum variance : **we try to recover what the assessors are telling us.**
- The loadings are obtained from the first eigenvector of : $T=X^TX$
    - ***Total variance-covariance matrix***

Discriminant analysis



From a technical point of view

- The first discriminant function is a linear combination of the attributes **which operates the best discrimination among products.**
- The vector of loadings a_1 maximises the discrimination index (between groups to total variance ratio): $\boldsymbol{a_1^T B a_1 / a_1^T T a_1}$
- a_1 is the first eigenvector of $\boldsymbol{T^{-1}B}$

Discriminant analysis **:Which characteristics differentiate among products?**

The discrimination may be reflected by minor attributes!

The analysis may be fragile, unstable!

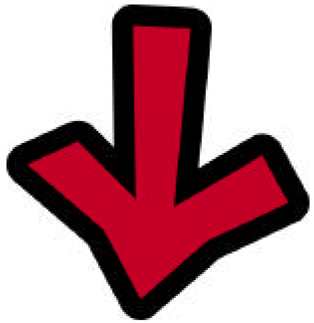
PCA on X : what are the assessors telling us?

- We do not fully take account of the structure of the data in terms of groups (products);

- Some marginal variations might blur the differences among groups.

Is PCA on the average data set a good compromise?

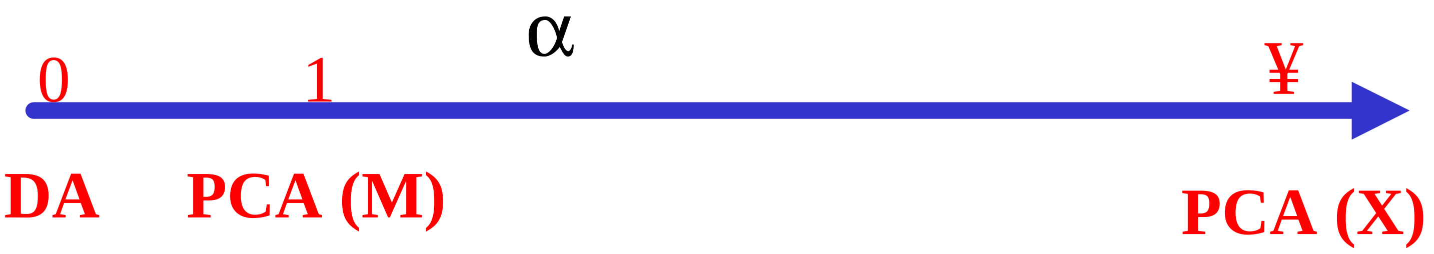
Is PCA on the average data set a good compromise?



Probably, yes!

But a better compromise may exist!

A continuum approach

- All three methods are, as a matter of fact, part of a **continuum approach.**
- 

- $\alpha=0$ → Disciminant analysis
- $\alpha=1$ → PCA (M)
- $\alpha \nearrow \infty$ → PCA (X)

Maximisation criterion

- Given α, a positive scalar,

- Find a (first) latent variable t=Xa such as to maximise :

$$f_a(\mathbf{a}) = \frac{\mathbf{a}^T\mathbf{B}\mathbf{a}}{\left(\mathbf{a}^T\mathbf{T}\mathbf{a}\right)^{1-a}}$$

Solution

- Iterative algorithm;
- **Step 0** : initialisation , choose a such that $a^{T}a=1$
- **Step 1** : compute $A = \dfrac{B}{a^{T}Ba} - (1-a)\dfrac{T}{a^{T}Ta}$
- **Step 2** :compute b : first eigenvector of matrix A
- **Step 3** : update $a \leftarrow$ b and **goto step** 1
- Other latent variables may be computed considering orthogonality constraints.

First property

- Consider $v_T(\alpha) = a^T T a$: the total variance explained by latent variable t=Xa.

- Then :

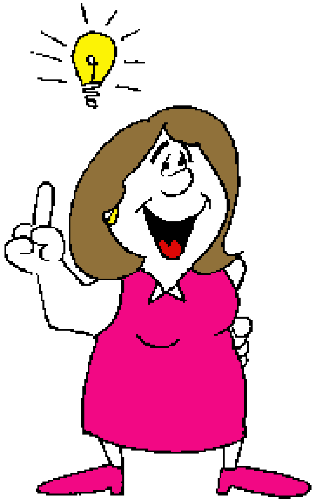
    - $v_T(\alpha)$ increases with α

Second property

- Consider $v_B(a) = a^T Ba$: the between groups variance explained by latent variable t=Xa.

- Then :

  - $v_B(\alpha)$ increases with α



Third property

- Consider the discrimination index (between groups to total variance ratio) associated with t:

- $I(a) = a^T Ba / a^T Ta$

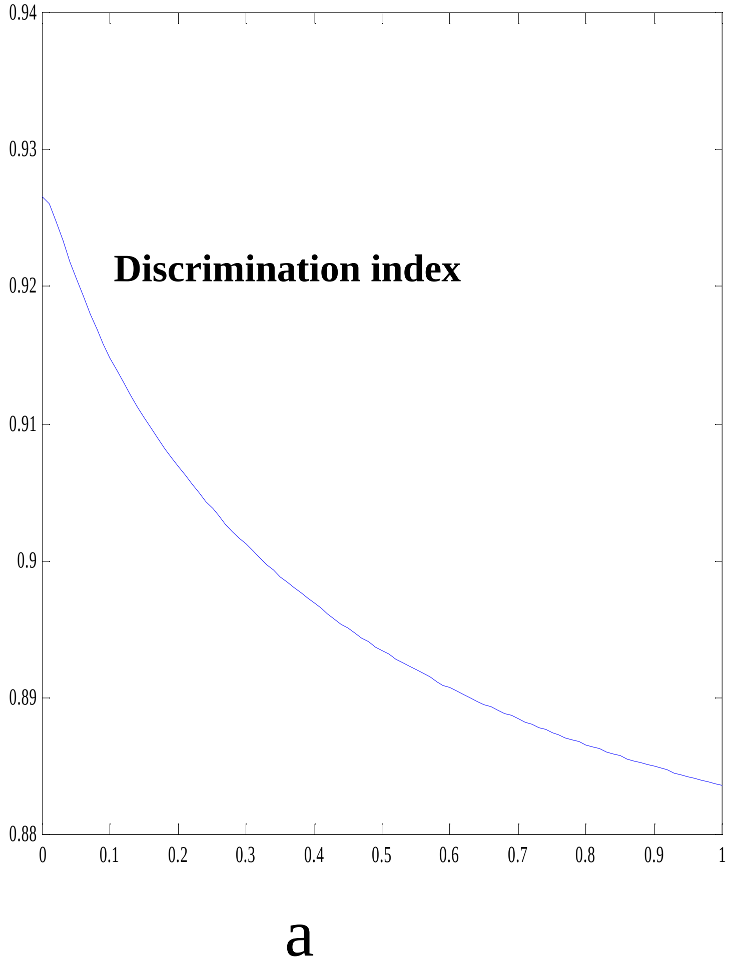
- *Then*

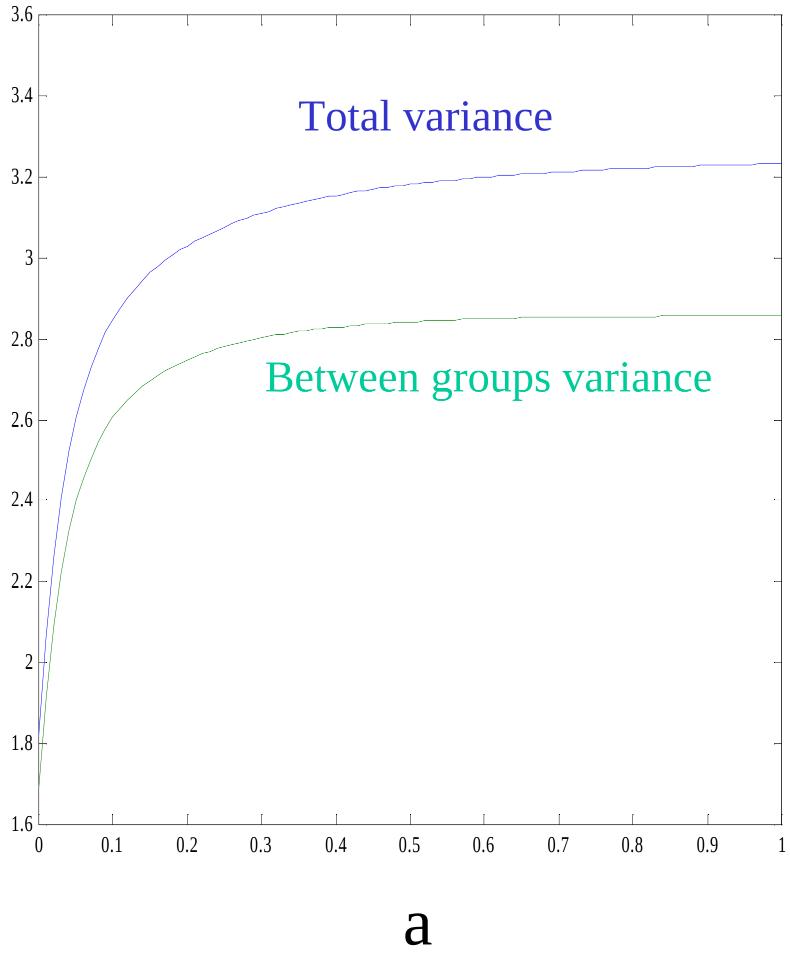
- I(α) decreases with α

A trade off!

- When α increases we lose in terms of discrimination but we gain in terms of total variance explained and between groups variance explained.
- What if there is an α that realises a good compromise?
- **The lost in terms of discrimination is more than compensated for by the gain in total variance explained**

See how it works!





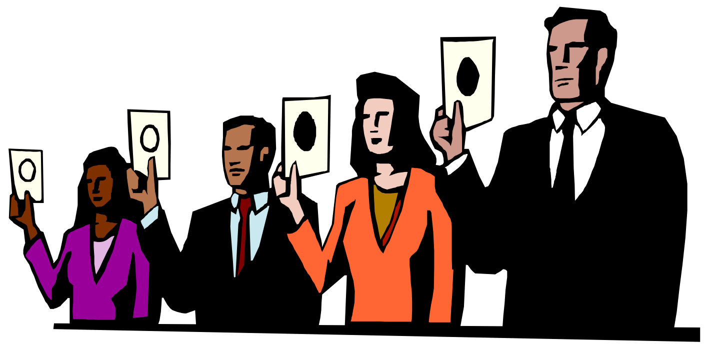
Cider Data

INTE intensity of odour

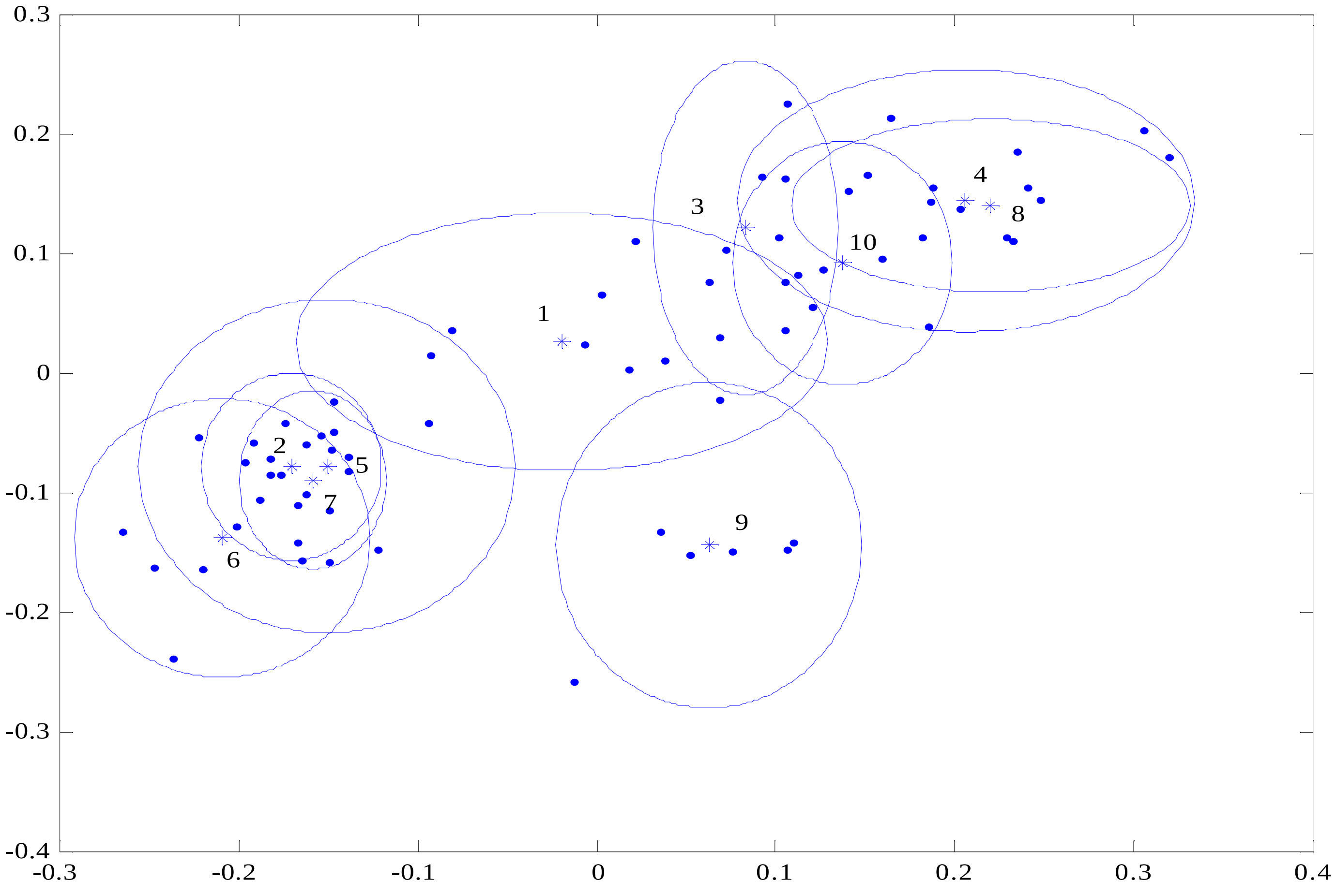
- **SUCR** sweetness
- **ACID** acid
- **Bitterness**
- **ASTR** astringency
- **Strength**
- **PIQU** spicy
- **ALCO** alcohol
- **PERF** perfume
- **FRUI** fruity

10 varieties of cider

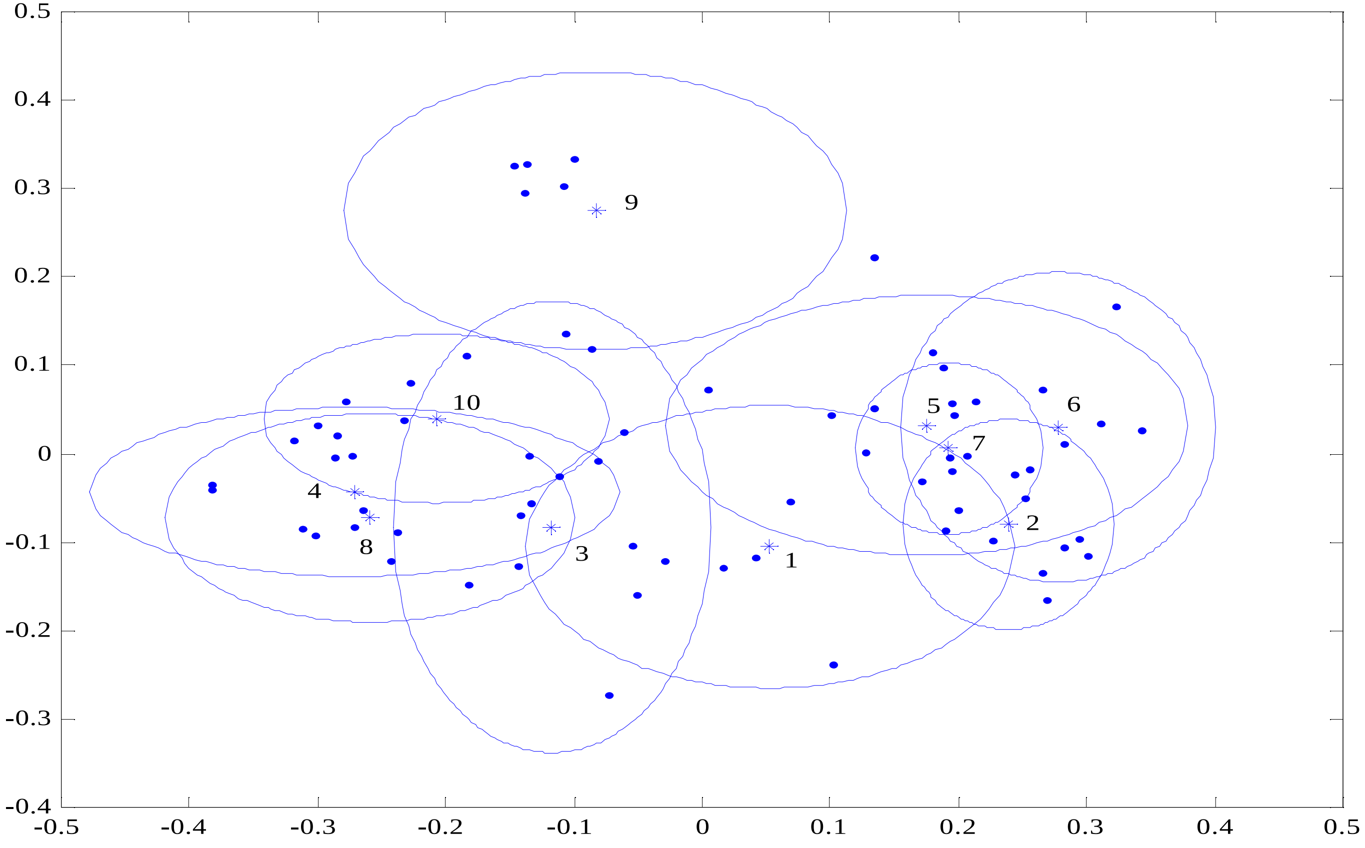
7 Assessors



a=0 → DA

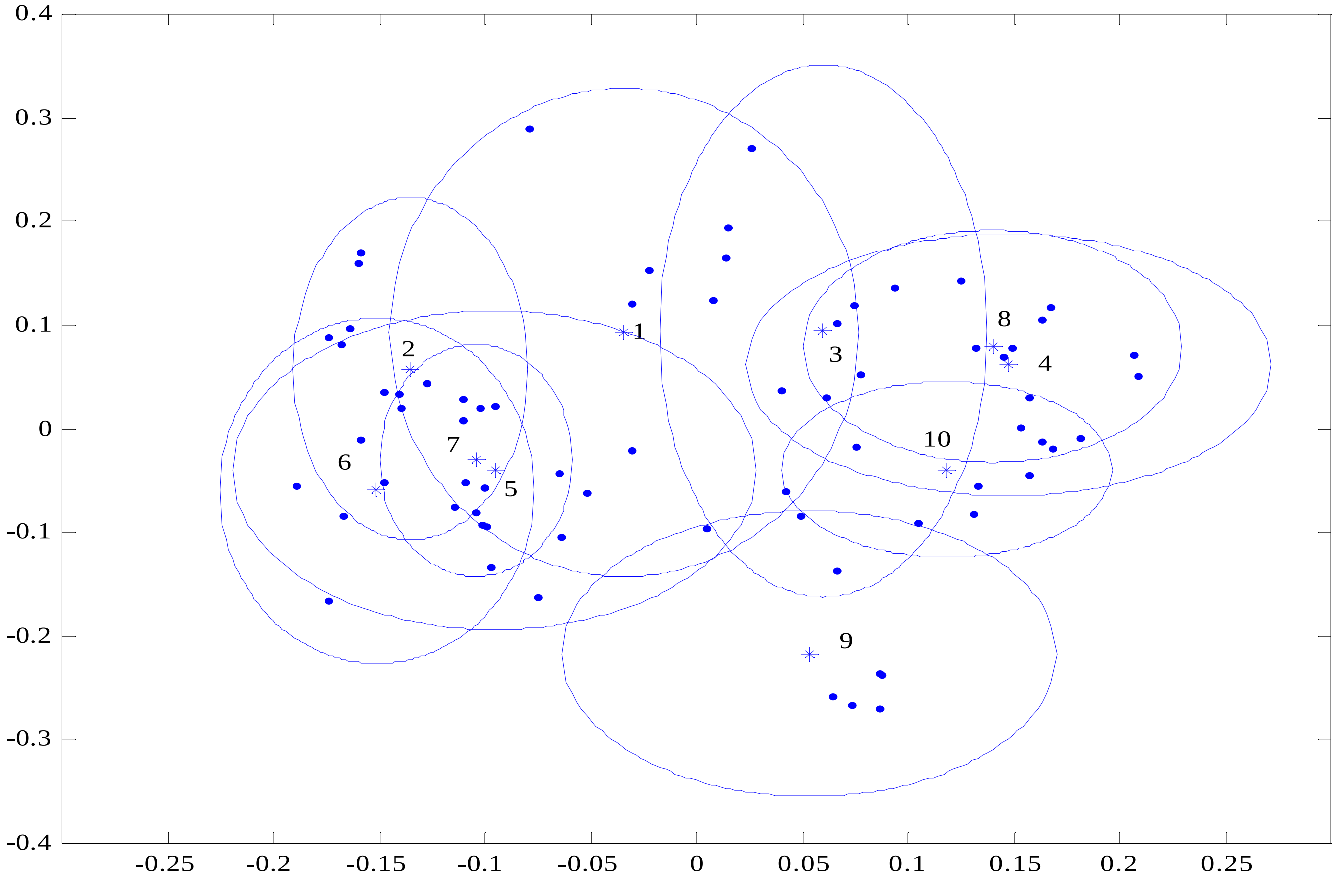


a=1 → PCA (M)



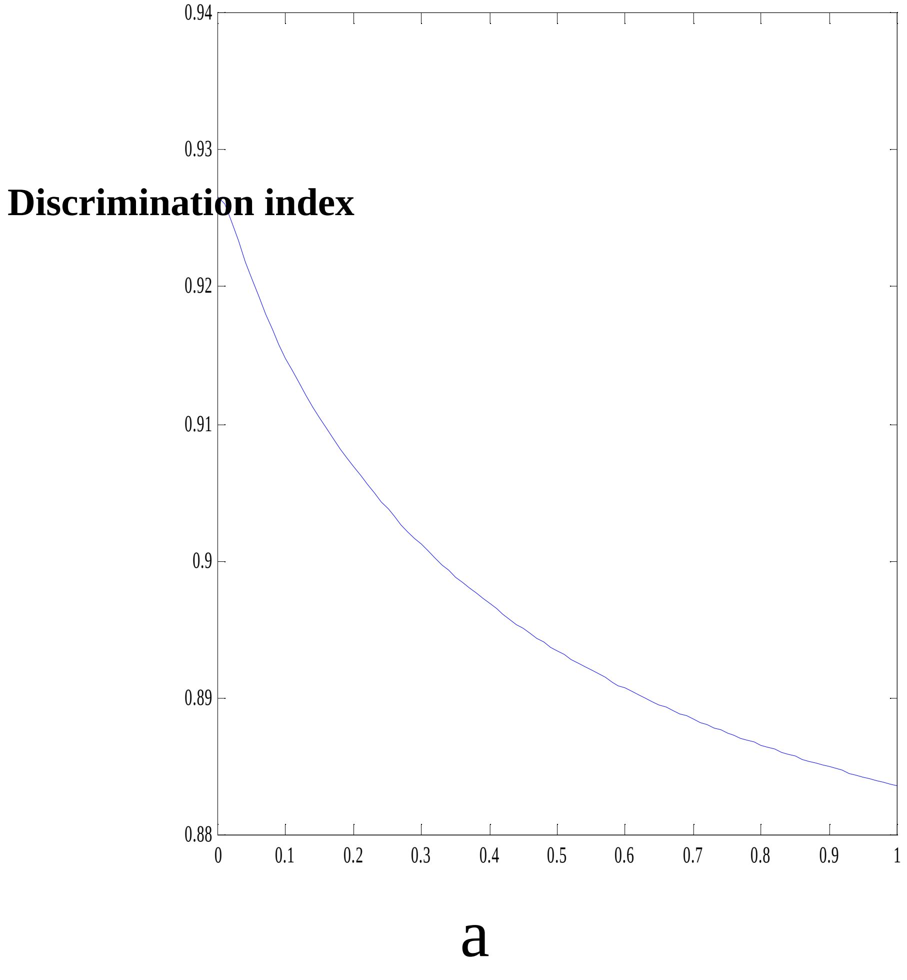
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
7
8
9
10

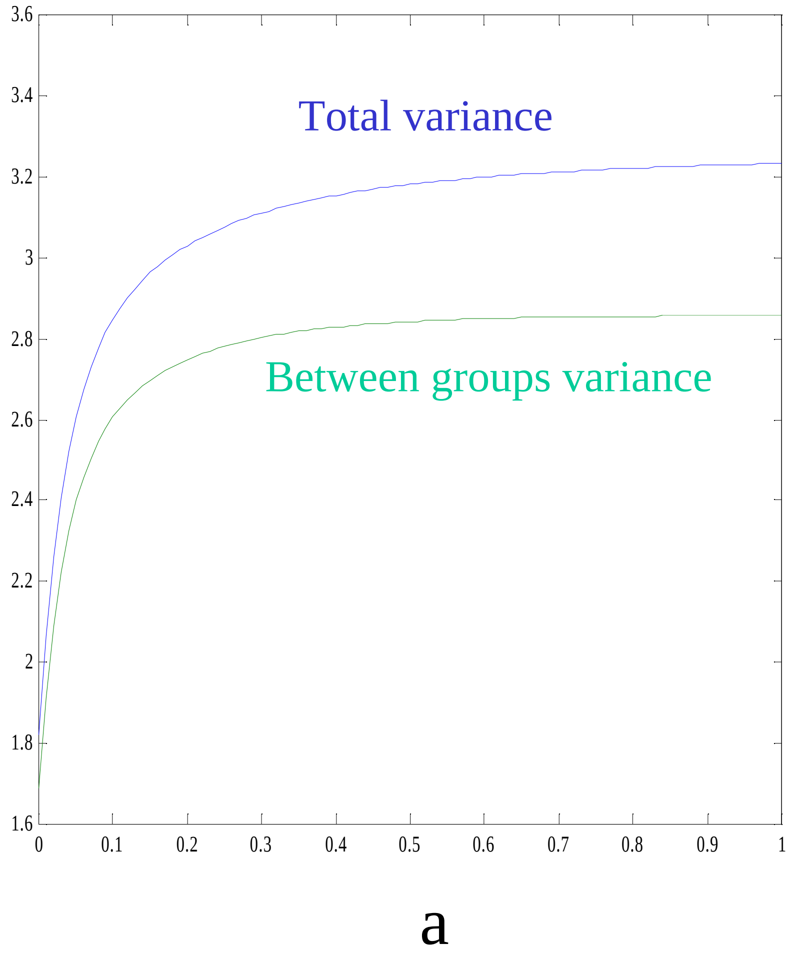
a ↗ ¥ ⟶ PCA (X)



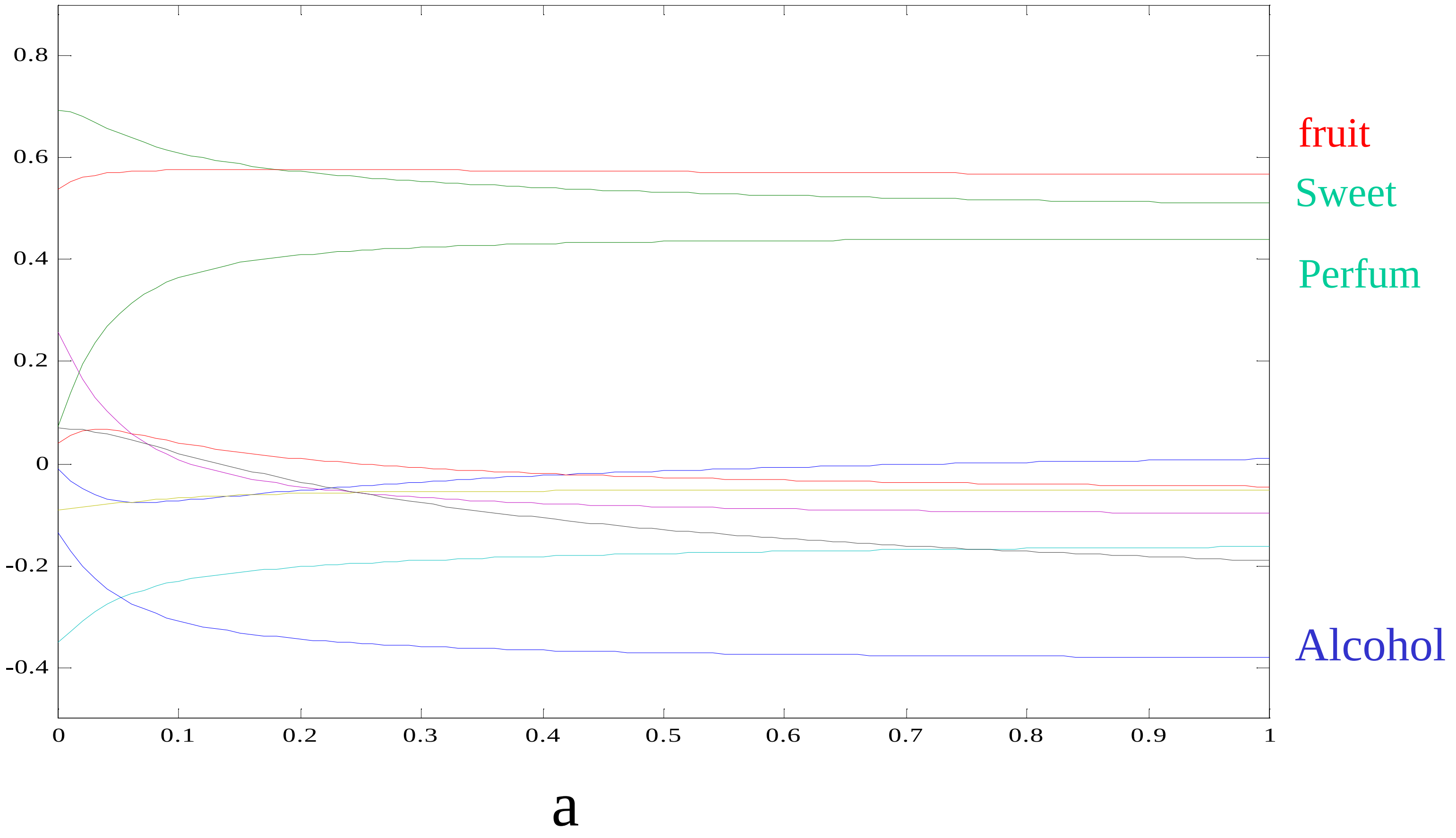
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.25
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
0.2
0.25
1
2
3
4
5
6
7
8
9
10

Which compromise to choose?

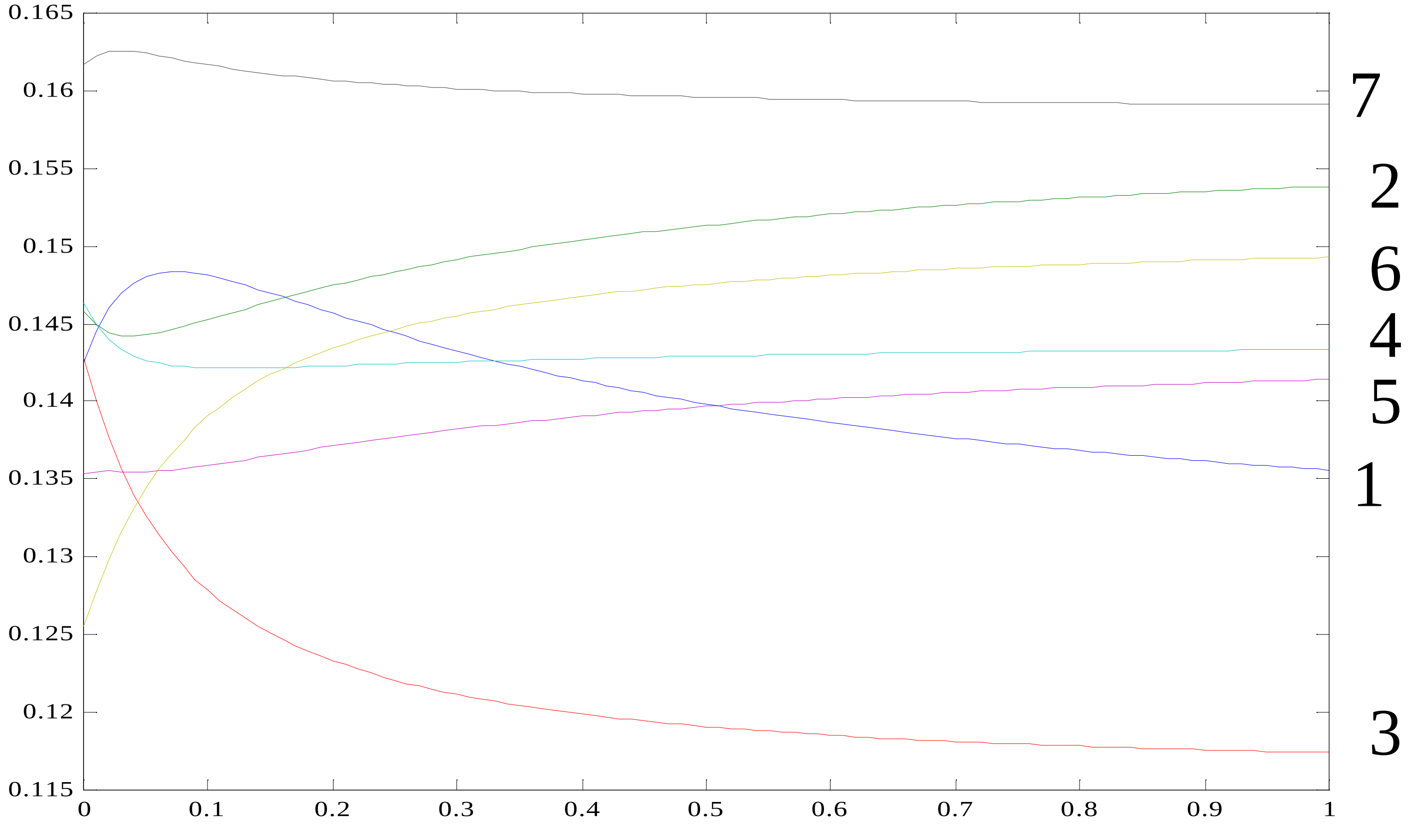




Loadings associated with the first latent variable

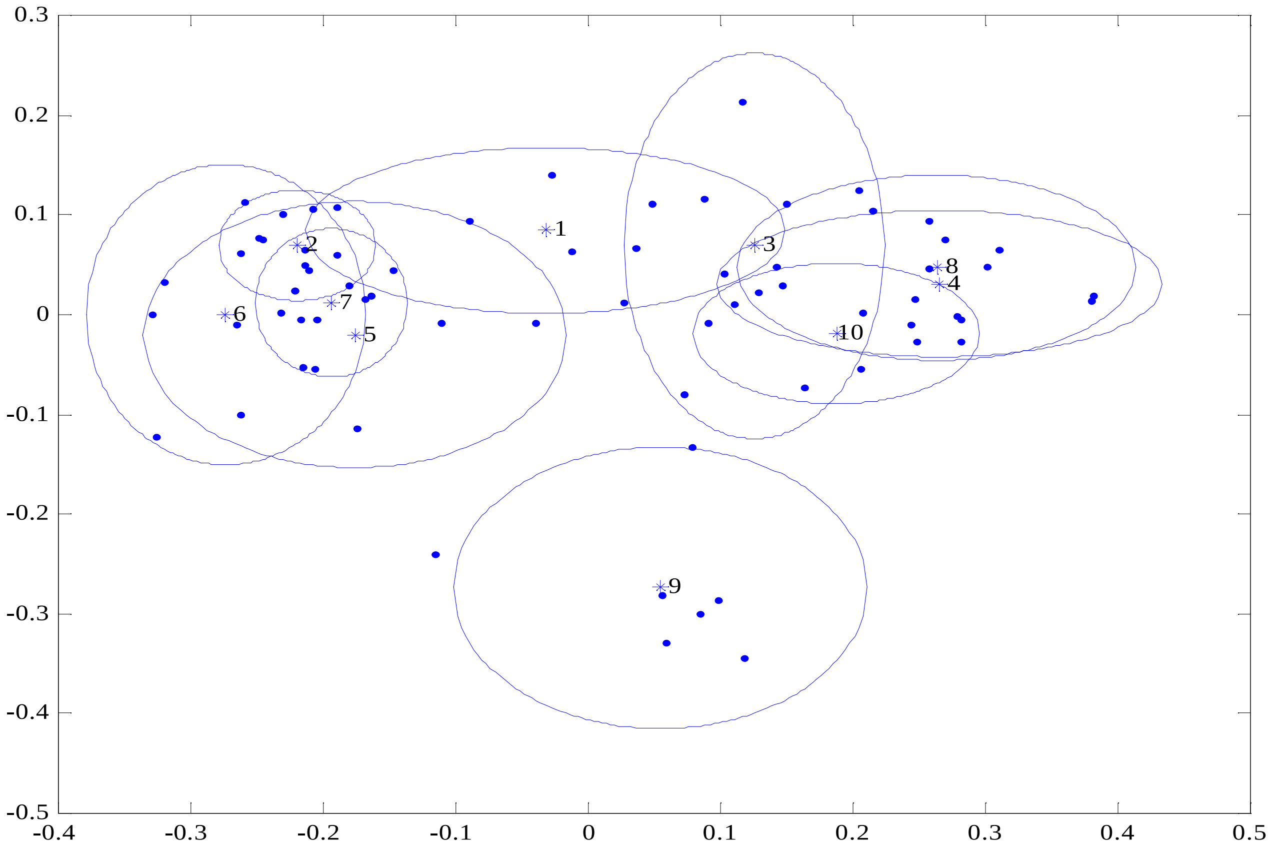


Assessors' contributions to the determination of the first latent component



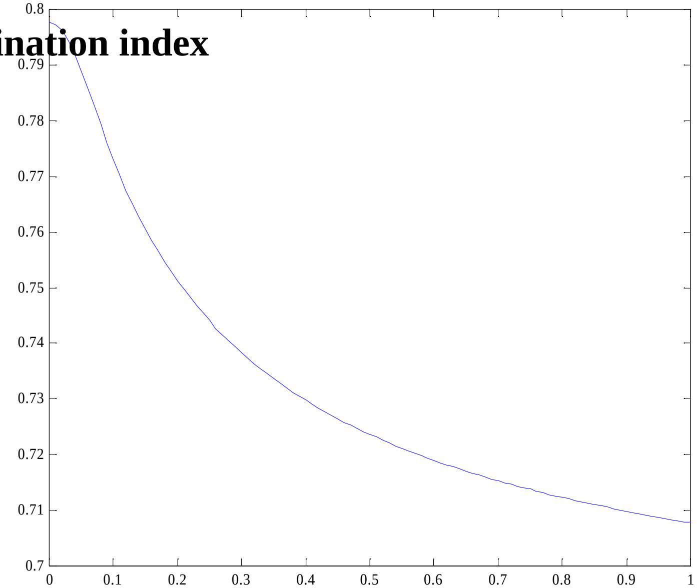
Alpha=0.15

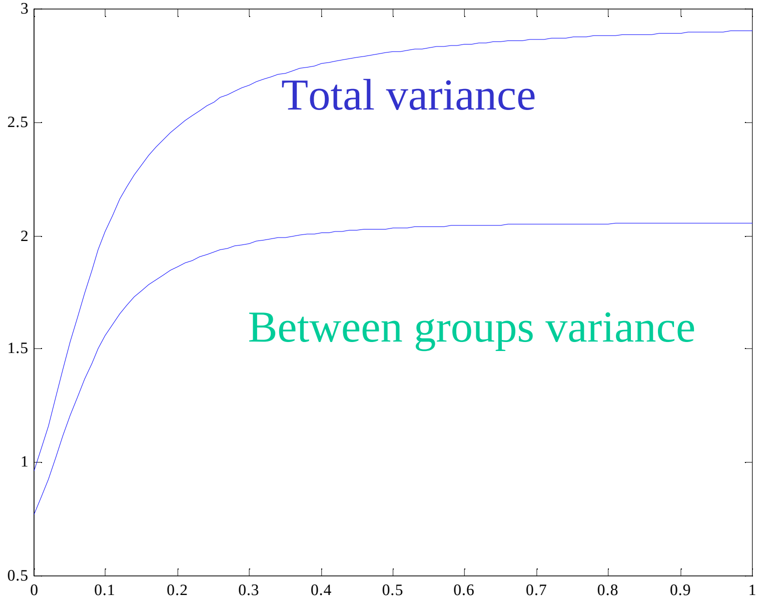
Between= 2.697
Total=2.963
Discrimination index=0.910



Noise was introduced in the data

Discrimination index

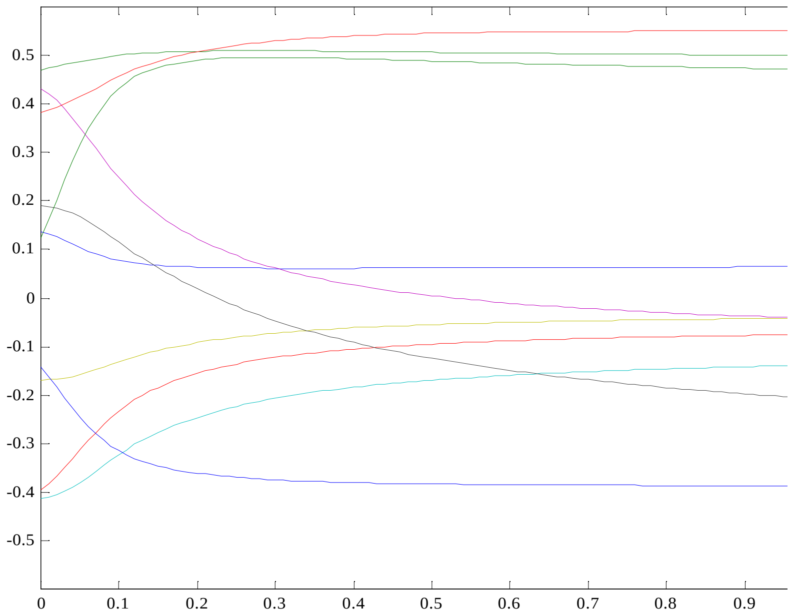


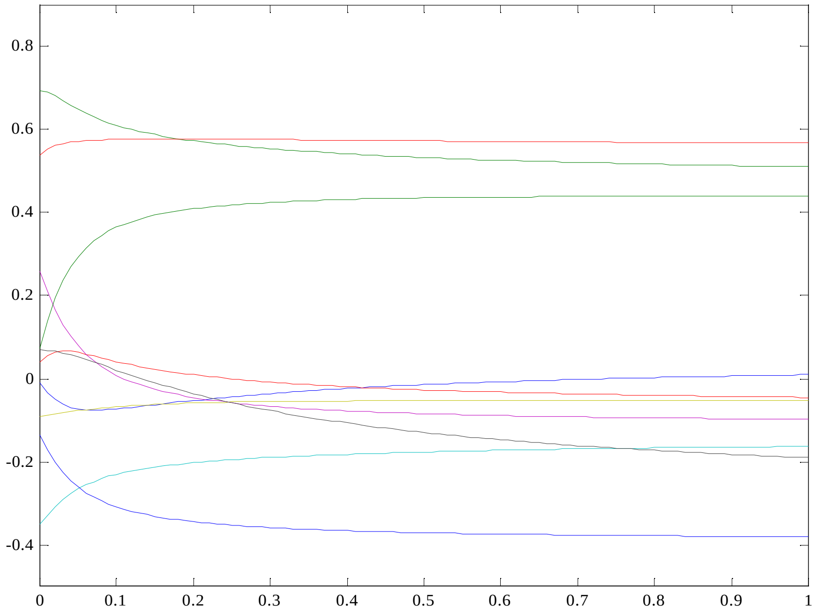


Loadings associated with the first latent variable

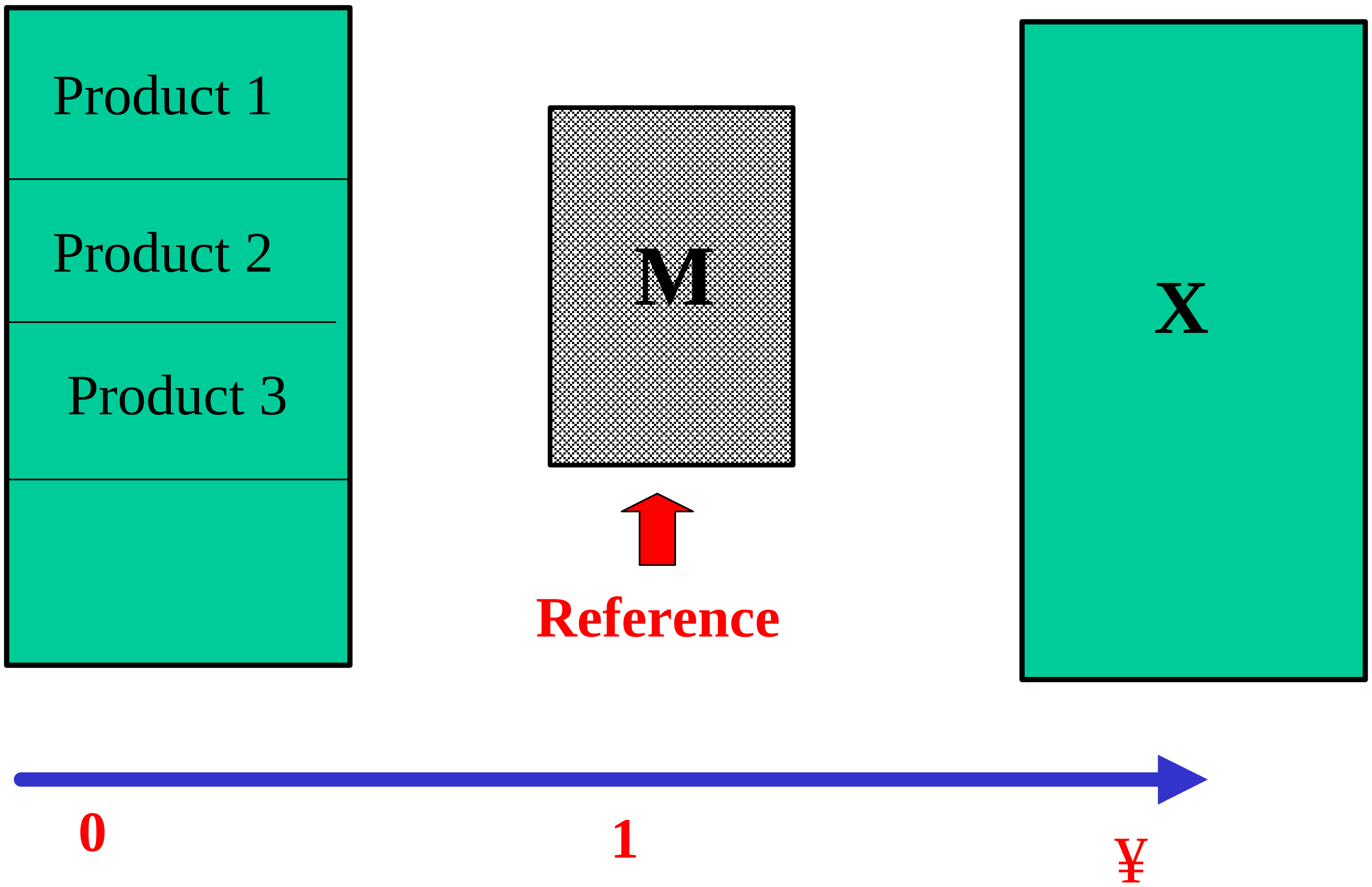
- Noisy

original

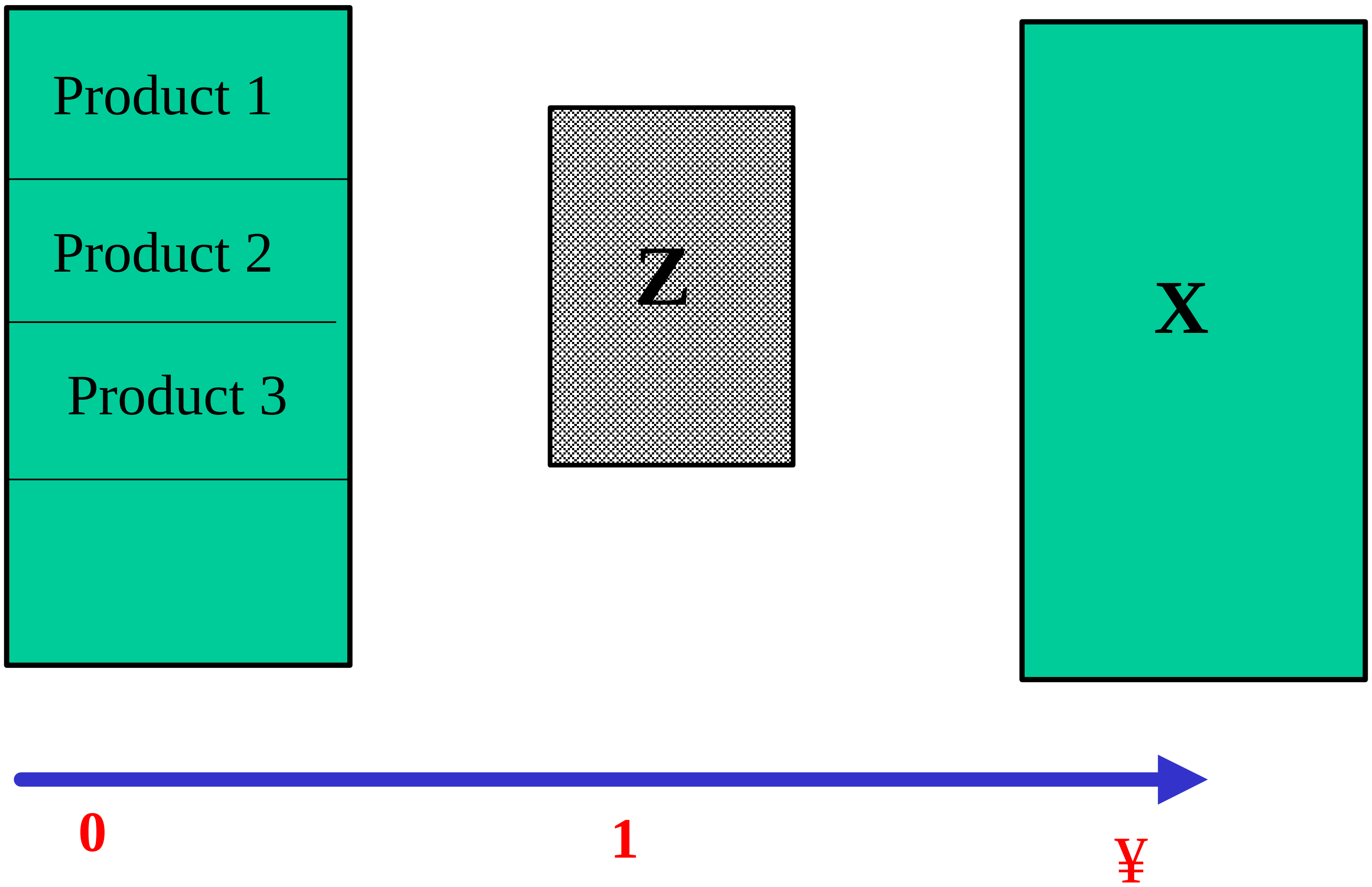




Underlying principle



Extension



Conclusion

- The strategy of analysis provides a **wide range of methods**. This make it possible to **safely** investigate the structure of the data.

  That is, it provides a stable model which realizes a good compromise between **recovering the total variance** (what the assessors are telling us) and **discriminating the products.**

I must go now!
my wife is
waiting for me!

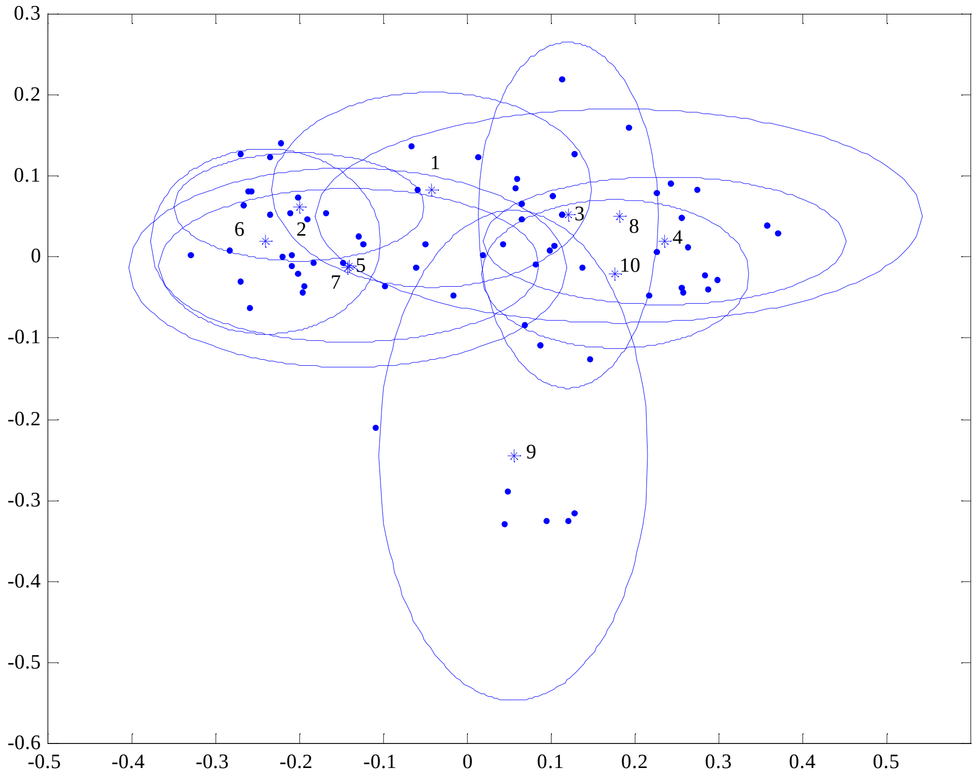
What's wrong with assessor 3

Individual (separate) Principal components analysis.

% of variation explained by the first PC.

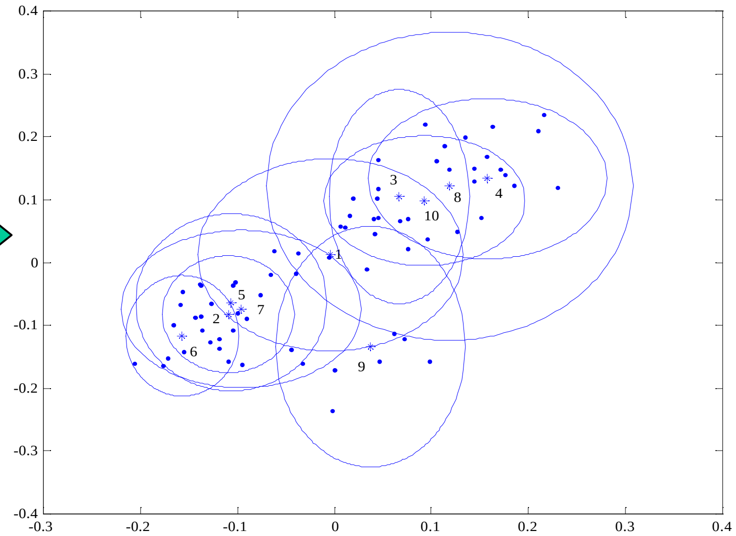
Assessor :	1	2	**3**	4	5	6	7
% var. Expl.	0.478	0.570	**0.395**	0.545	0.508	0.562	0.528

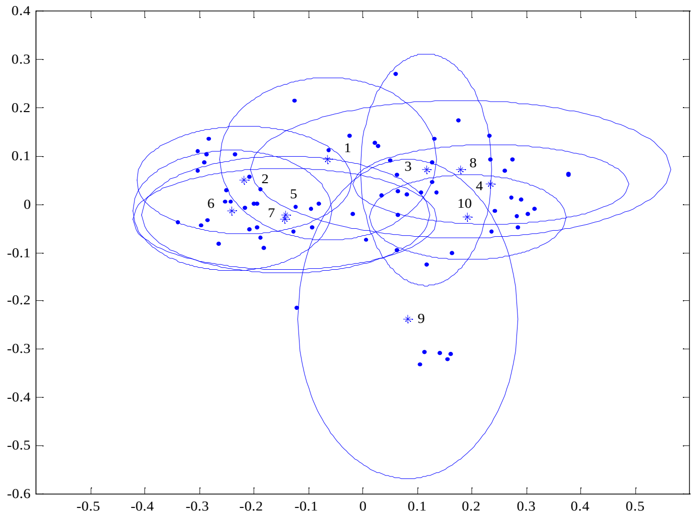
Alpha=0.25



0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
1
2
3
4
5
6
7
8
9
10

a=0 : Discrimination





a=1 : PCA (M)

a↗¥ : PCA (X)