MARRAKECH
جامعة القاضي عياض
UNIVERSITÉ CADI AYYAD

كلية العلوم
والتقنيات – مراكش
FACULTÉ DES SCIENCES ET
TECHNIQUE – MARRAKECH

AFD :
ANALYSE FACTORIELLE DISCRIMINANTE

Préparé par :

- **Labrassi Mohammed**
- **Laamari Salma**
- **Essarghi Hiba Allah**

Avec notre professeur:

- **Ouaarab Aziz**

Plan:

I. Introduction

II. Principe de l'analyse factorielle discriminante

III. Les axes discriminants

IV. étapes et démarche à suivre

V. Règle de classification

VI. Exemples d'application

I

INTRODUCTION

I - Introduction

I -1 - Définition

On distingue deux aspects en analyse discriminante :

1- Analyse discriminante descriptive :

- Il s'agit de chercher les combinaisons linéaires de variables qui permettent de séparer le mieux possible les K classes et donner une représentation graphique, qui rende compte au mieux de cette séparation.
- Ces combinaisons linéaires sont appelées fonctions linéaires discriminantes. Il s’agit donc d’une étape de discrimination des classes.

2- Analyse discriminante prédictive :

- Un nouvel individu se présente pour lequel on connait les valeurs des prédicteurs. Il s’agit alors de décider dans quelle classe il faut l’affecter.

I - Introduction

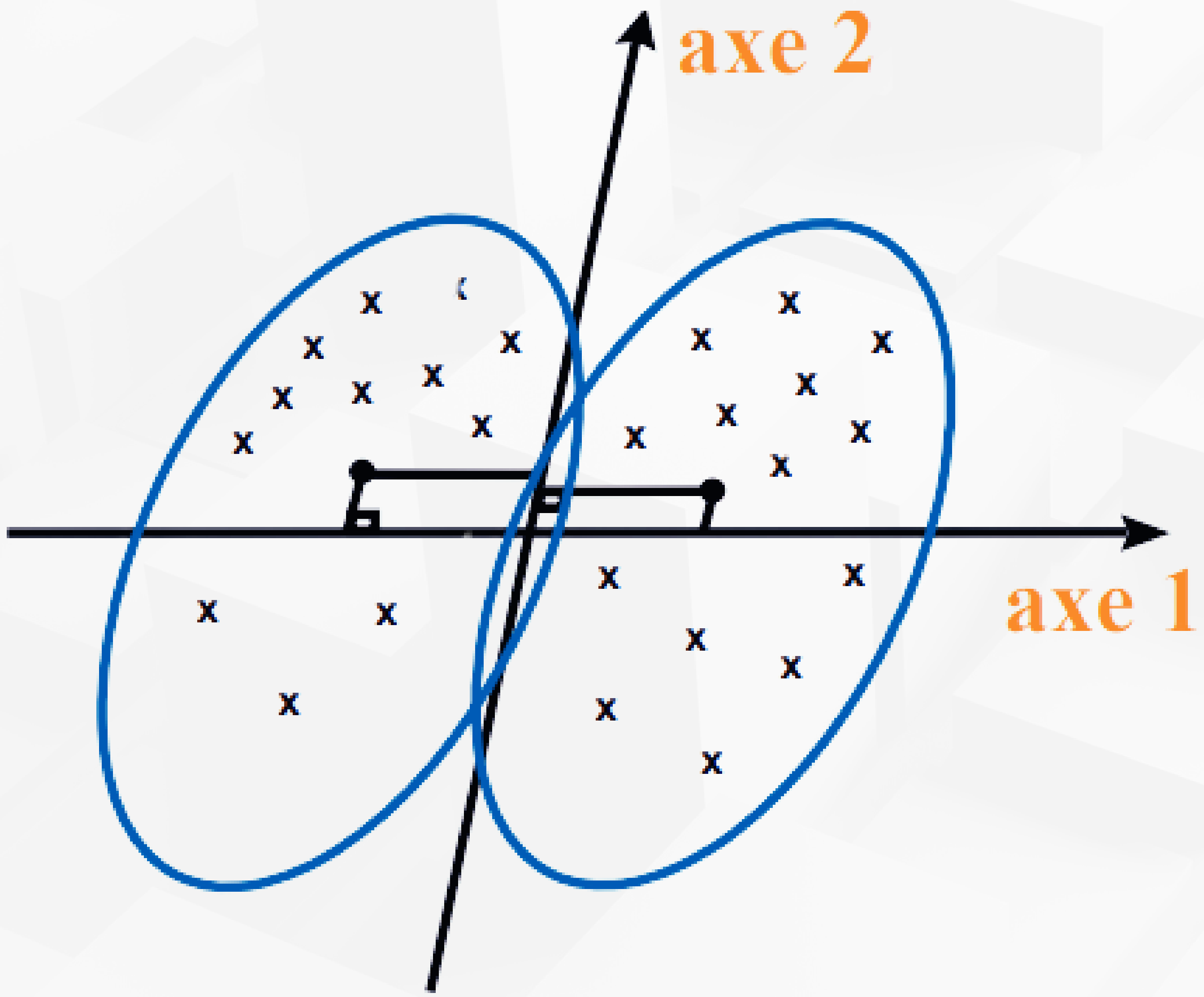
I -2 - Objectifs

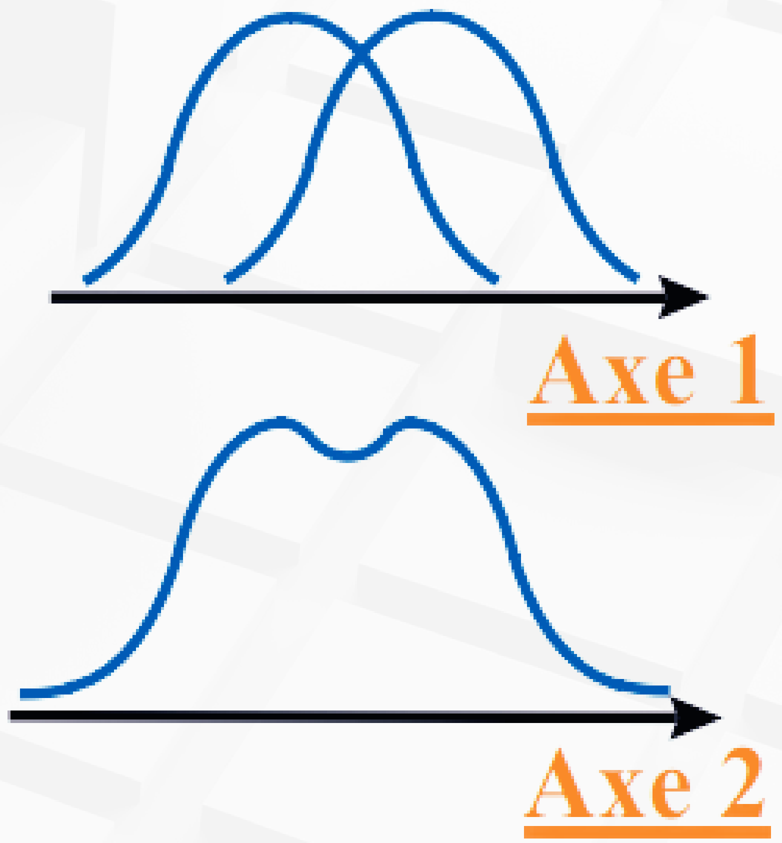
Parmi les objectifs principaux peuvent être assignés à l'analyse discriminate:

- Déterminer les variables explicatives les plus discriminantes vis à vis des classes detérminées .
- Déterminer à quel groupe appartient un individu à partir de ses caractéristique.
- Valider une classification ou à faire un choix entre plusieurs classifications pour savoir laquelle la plus pertinente.
- Vérifier sur un graphique à deux ou trois dimensions si les groupes auxquels appartiennent les observations sont bien distincts.
- Maximiser la séparabilité des données :

I - Introduction

I -2 - Objectifs



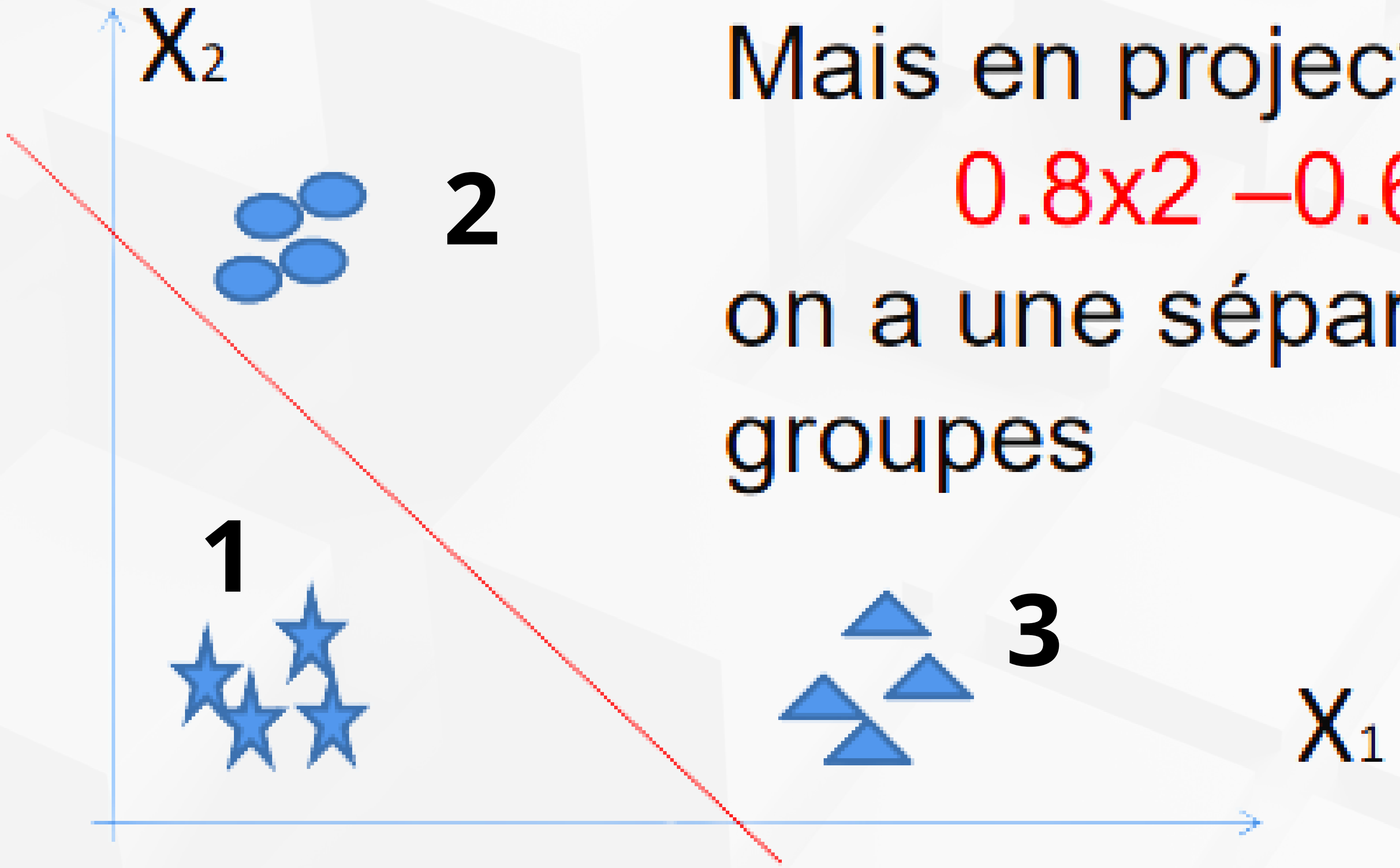


L'axe 1 possède un bon pouvoir discriminant
L'axe 2 ne permet pas de séparer en projection les 2 groupes.

I - Introduction

I -2 - Objectifs

- **3 groupes, 2 variables explicatives:**
  - Groupes 1 et 3 confondus sur l'axe 2
  - Groupes 1 et 2 confondus sur l'axe 1



Mais en projection sur l'axe

$0.8x2 - 0.6x1$

on a une séparation des 3 groupes

I - Introduction

I -3 - Conditions

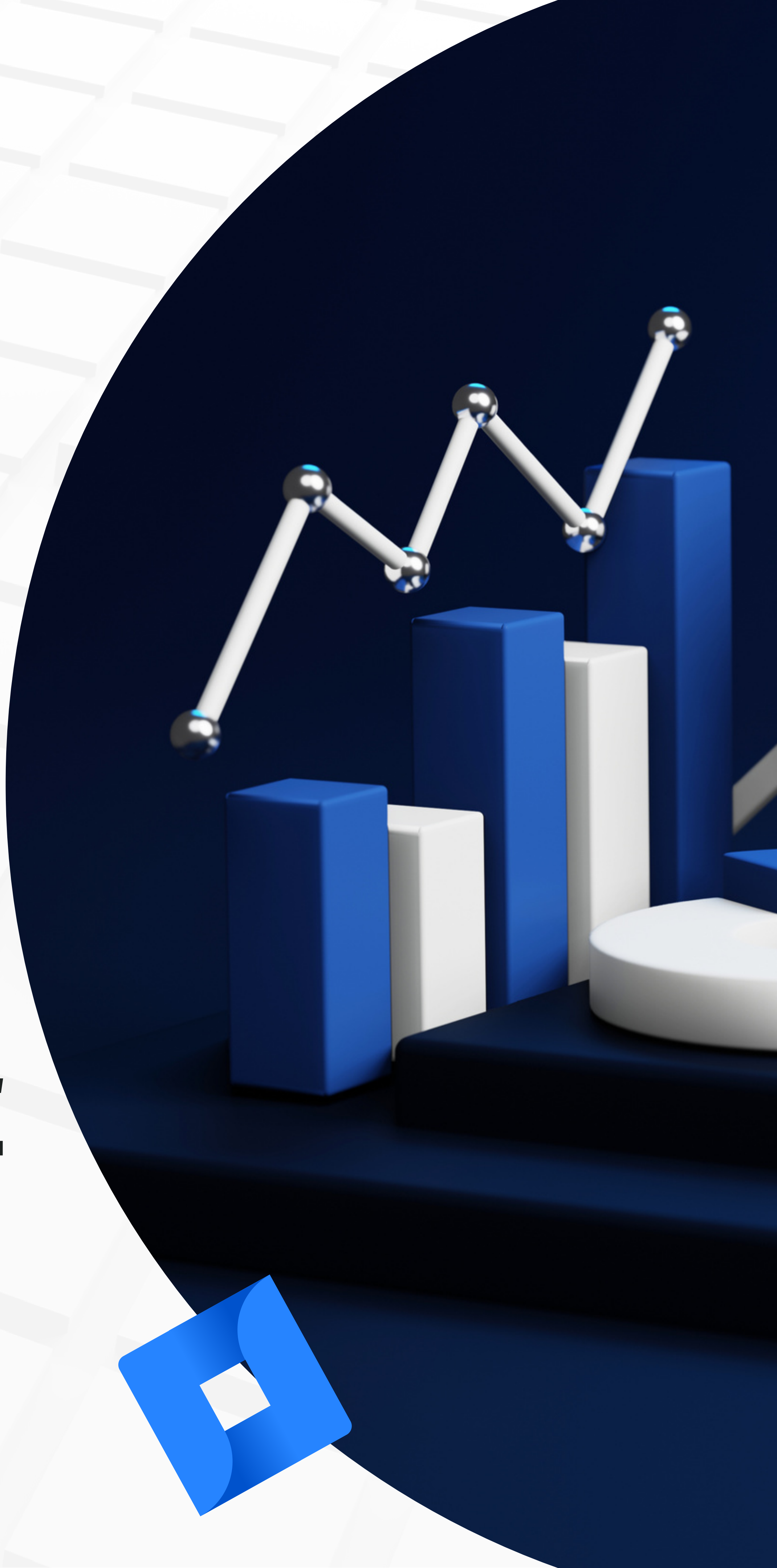
Deux conditions à remplir

Les variables explicatives doivent étre métrique

Elles ne doivent pas étre trop corrélées entre elles

II

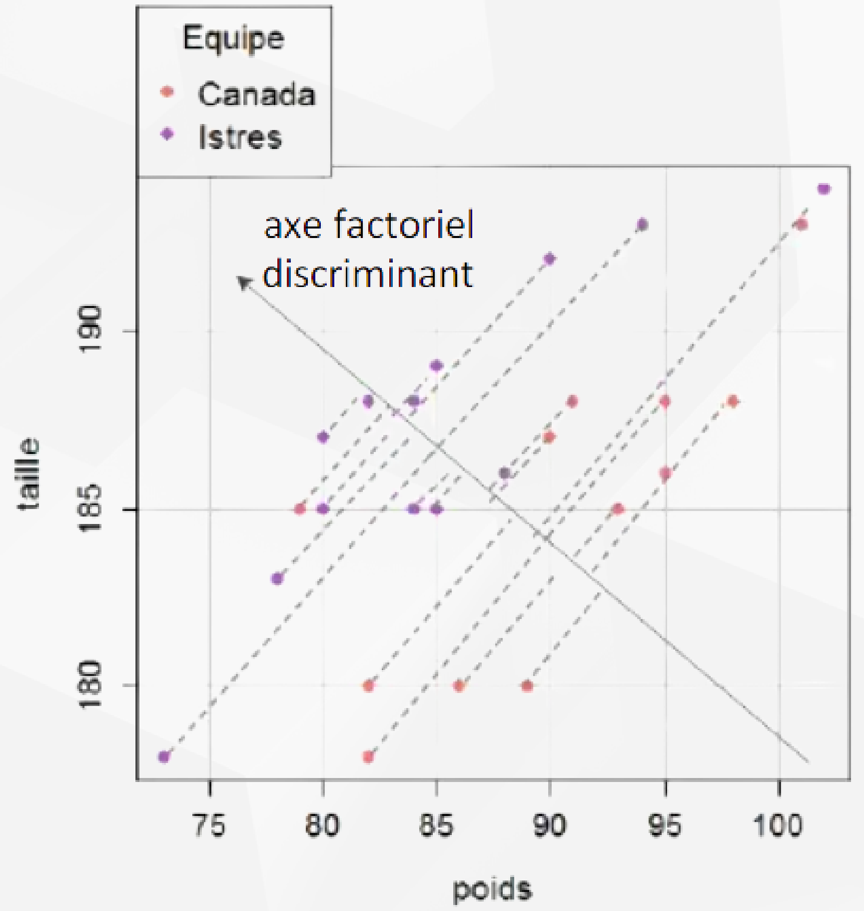
PRINCIPE DE L'ANALYSE FACTORIELLE DISCRIMINANTE

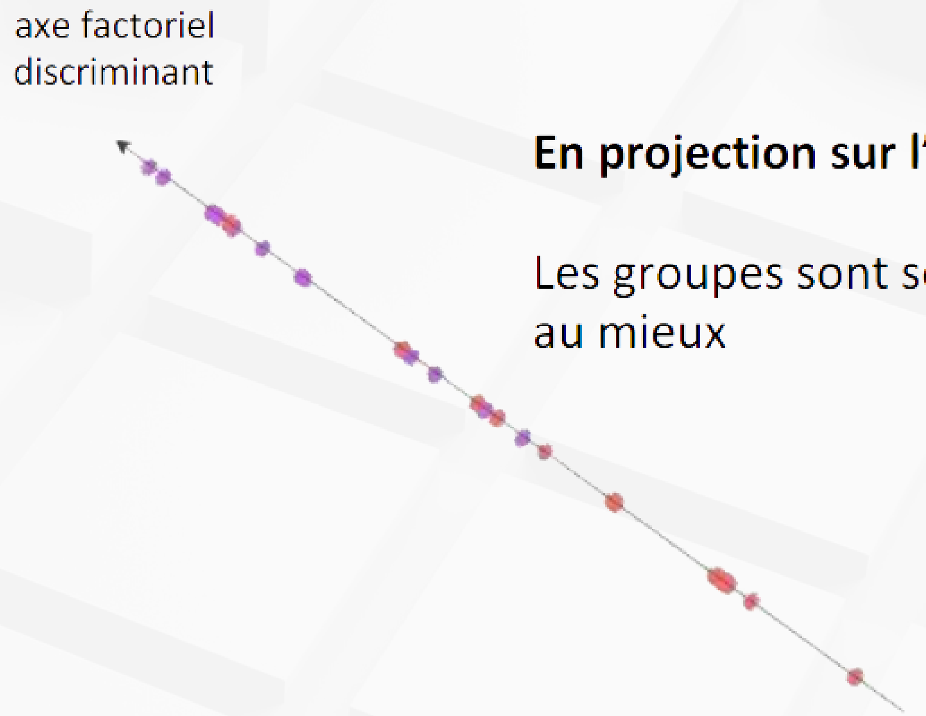


II - Principe de l’analyse factorielle discriminante

II -1 - Principe de l’AFD :

- Rechercher une ou des frontières entre les groupes en se servant de l’ensemble des variables explicatives
- Construire un ou des axes factoriels discriminants, à partir de combinaisons linéaires des variables initiales





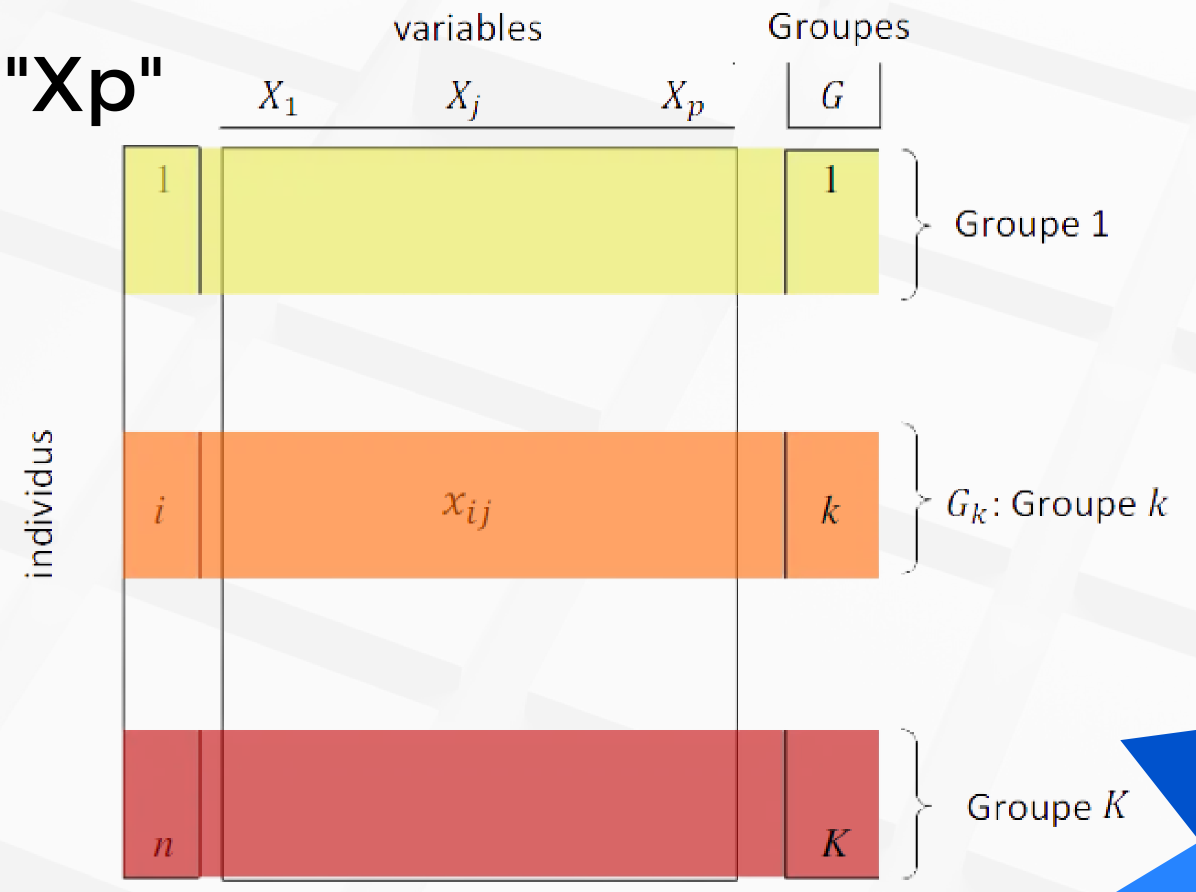
En projection sur l’axe :

Les groupes sont séparés au mieux

II - Principe de l'analyse factorielle discriminante

II -2 - Données et Notations :

- **Y variable à expliquer** à **k modalités** partage la population en k sous populations.
- On dispose d'un échantillon de n observations partagé en k groupes de taille n1 n2… nk.
- Les lignes sont étiquetées avec les individus "i" à "n".
- Les colonnes sont étiquetées avec les variables "X1" à "Xp"
- p variables explicatives numériques .



II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :

- Deux équipes de 13 joueurs chacune
- On connaît le poids et la taille de chaque joueur

	Equipe	taille	poids
1	Canada	178	82
2	Canada	100	02
3	Canada	180	86
4	Canada	180	89
5	Canada	185	79
6	Canada	185	84
7	Canada	185	93
0	Canada	166	95
9	Canada	187	90
10	Canada	188	91
11	Canada	100	95
12	Canada	188	98
13	Canada	193	101
14	Istres	169	85
15	Istres	185	84
16	Istres	186	88
17	Istres	192	90
18	Istres	193	94
19	Istres	194	102
20	Istres	185	80
21	Istres	185	85
22	Istres	188	84
23	Istres	183	78
24	Istres	187	80
25	Istres	178	73
26	Istres	188	82

II - Principe de l'analyse factorielle discriminante

II -2 - Données et Notations :

- gk centre de gravité du nuage de n points du chaque groupe k :

$$\mathbf{g}_k = \frac{1}{n_k} \sum_{i \in G_k} x_i.$$

- T matrice de variance-covariance globale de dimension p x p :

$$T = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mathbf{g})(x_i - \mathbf{g})'.$$

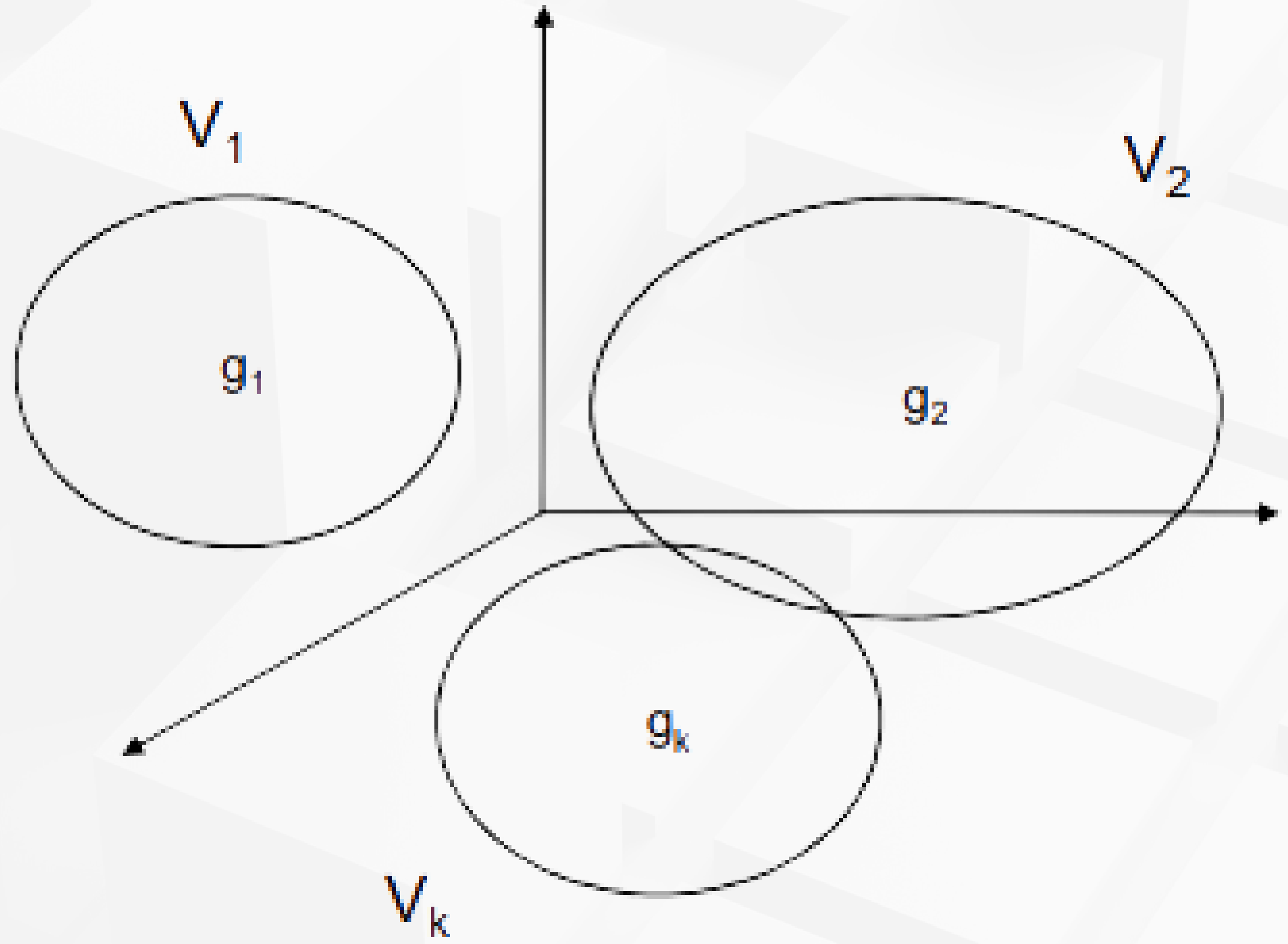
- Vk matrice de variance-covariance du groupe gk de dimension p x p :

$$\mathbf{V}_k = \frac{1}{n_k} \sum_{i \in G_k} (x_i - \mathbf{g}_k)(x_i - \mathbf{g}_k)'.$$

II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :

- exemple de gk centre de gravité et Vk matrice de variance du nuage des nk points de la classe k .



II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :

- B est variance-covariance inter-groupe de dimention p x p :

$$\mathbf{B} = \sum_{k=1}^{K} \frac{n_k}{n} (\mathbf{g}_k - \mathbf{g})(\mathbf{g}_k - \mathbf{g})'.$$

- W est variance-covariance intra-groupe de dimention p x p :

$$\mathbf{W} = \sum_{k=1}^{K} \mu_k V_k = \sum_{k=1}^{K} \frac{n_k}{n} V_k.$$

II - Principe de l'analyse factorielle discriminante

II -2 - Données et Notations :

- T est variance totale :

  variance totale = moyenne des variances marginales + variance des moyennes marginales.

$$\mathbf{T = W + B}$$

II - Principe de l'analyse factorielle discriminante

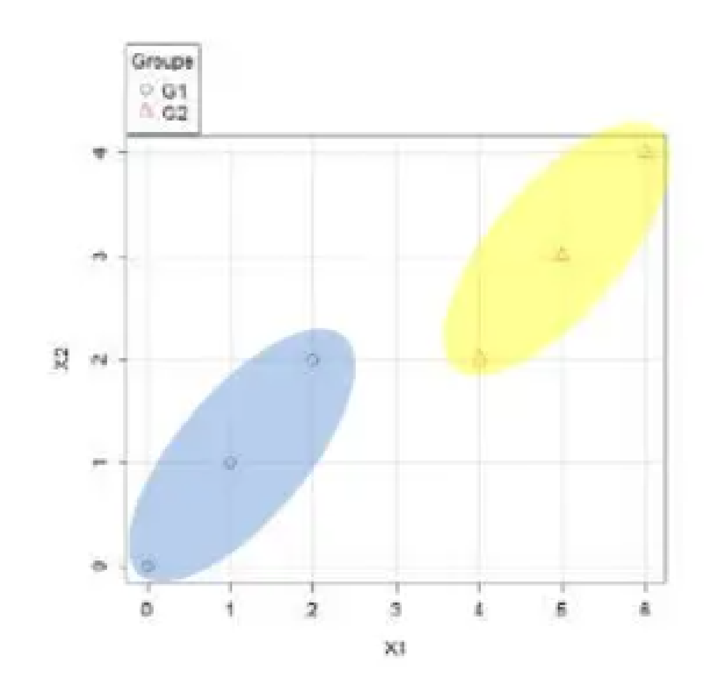
II -2 - Données et Notations :

- T est variance totale :

  Contexte

  - Deux groupes G1 et G2
  - Deux variables X1 et X2

X1	X2	Groupe
0	0	G1
1	1	G1
2	2	G1
4	2	G2
5	3	G2
6	4	G2



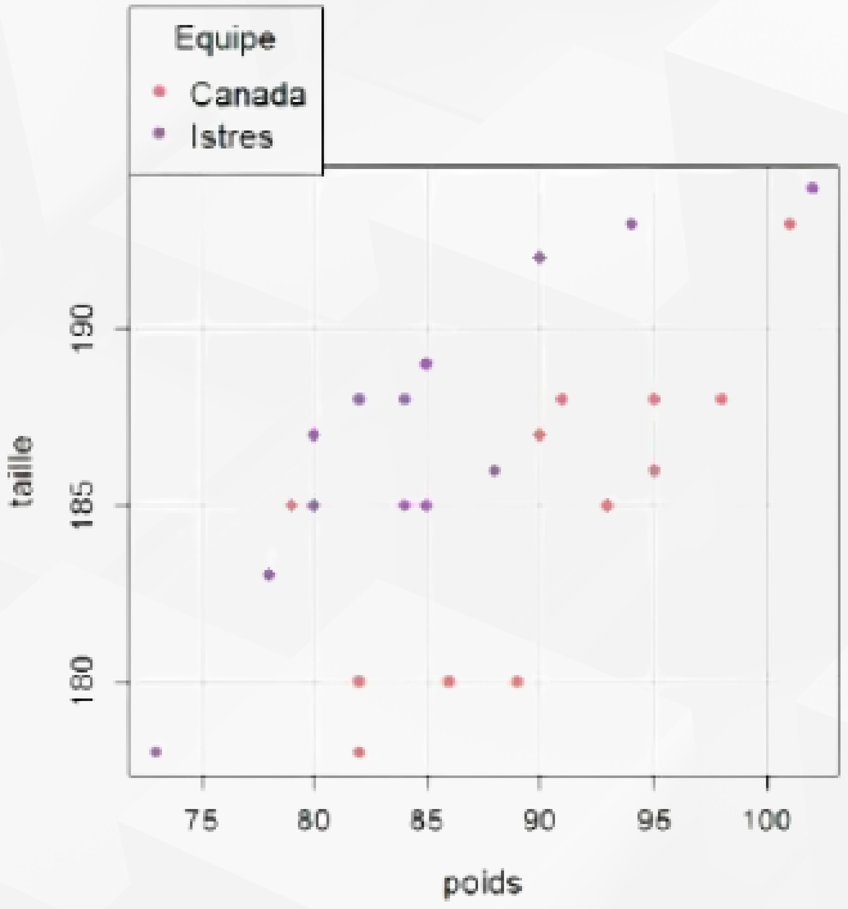
$$\begin{pmatrix} \mathrm{V}(X_1) & \mathrm{cov}(X_1,X_2) \\ \mathrm{cov}(X_1,X_2) & \mathrm{V}(X_2) \end{pmatrix} = \begin{pmatrix} \text{Variance des } \bar{x}_{1k} & \text{Covariance } (X_1,X_2) \text{ sur cg} \\ \text{Covariance } (X_1,X_2) \text{ sur cg} & \text{Variance des } \bar{x}_{2k} \end{pmatrix} + \begin{pmatrix} \text{Moyenne des } \mathrm{V}(X_1) & \text{moyenne des cov Intra} \\ \text{Moyenne des cov Intra} & \text{Moyenne des } \mathrm{V}(X_2) \end{pmatrix}$$

$$T \quad = \quad B \quad + \quad W$$

II - Principe de l'analyse factorielle discriminante

II -2 - Données et Notations :

Application à l'exemple « handball »



matrice de variance-covariance globale T

	taille	poids
taille	480,000	525,000
poids	525,000	1341,538

=

Matrice de Variance-Covariance Inter-Classe B

	taille	poids
taille	34,615	-69,231
poids	-69,231	138,462

+

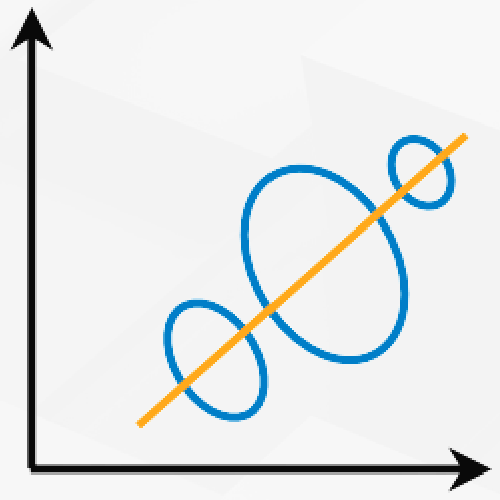
Matrice de Variance-Covariance Intra-Classe W

	taille	poids
taille	445,385	594,231
poids	594,231	1203,077

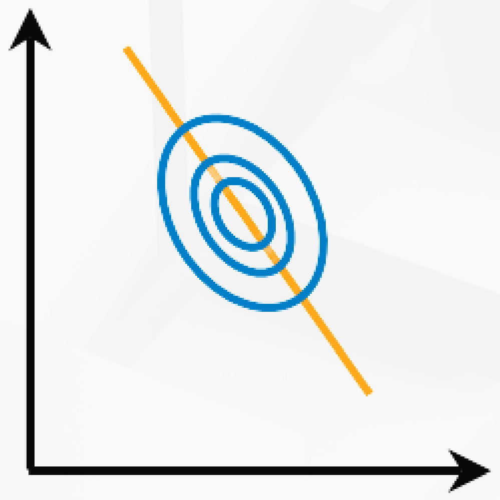
II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :

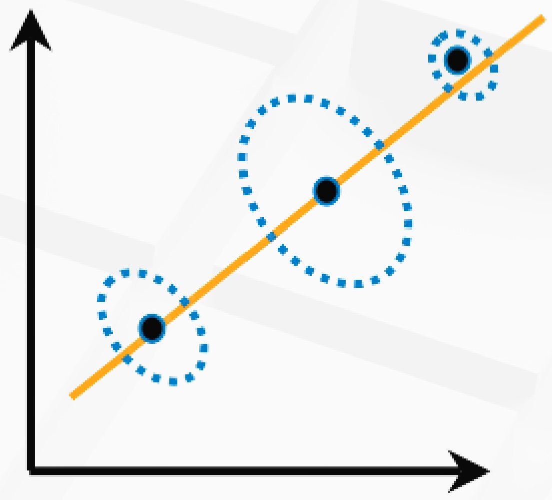
- En analyse discriminante, on considère trois types de matrices de variances-covariances et donc trois types de corrélations:



corrélation totale



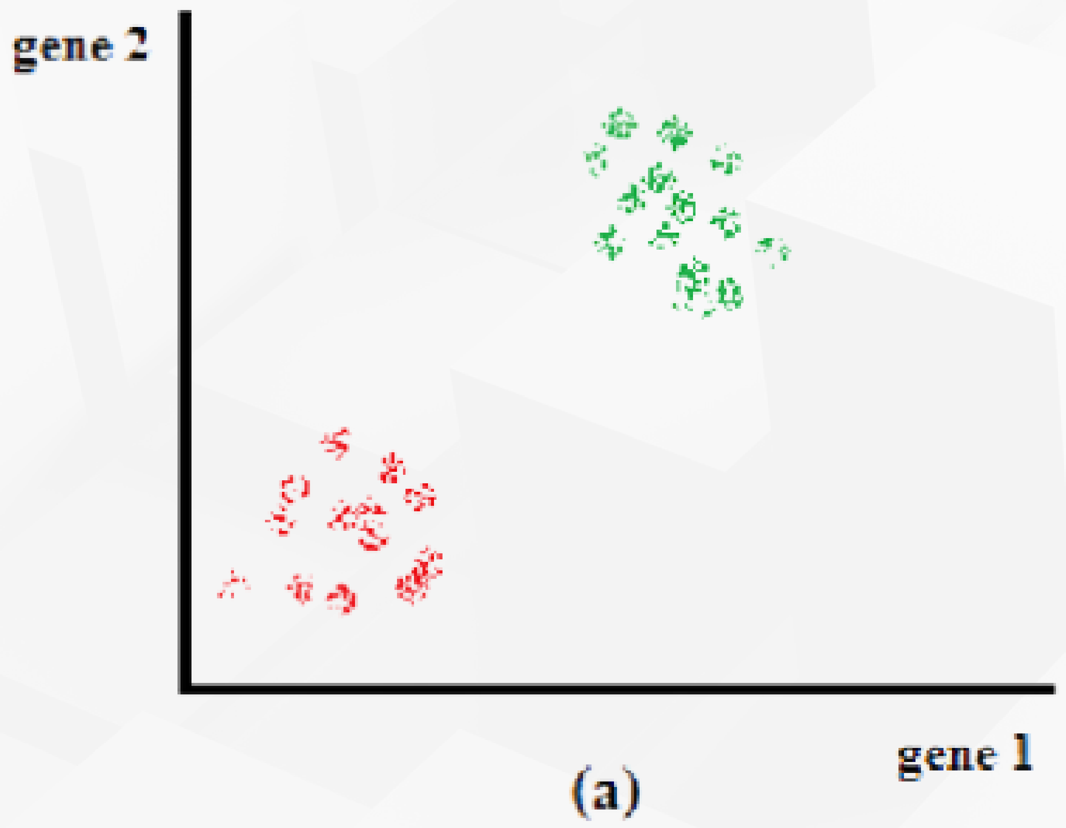
corrélation intra-classes



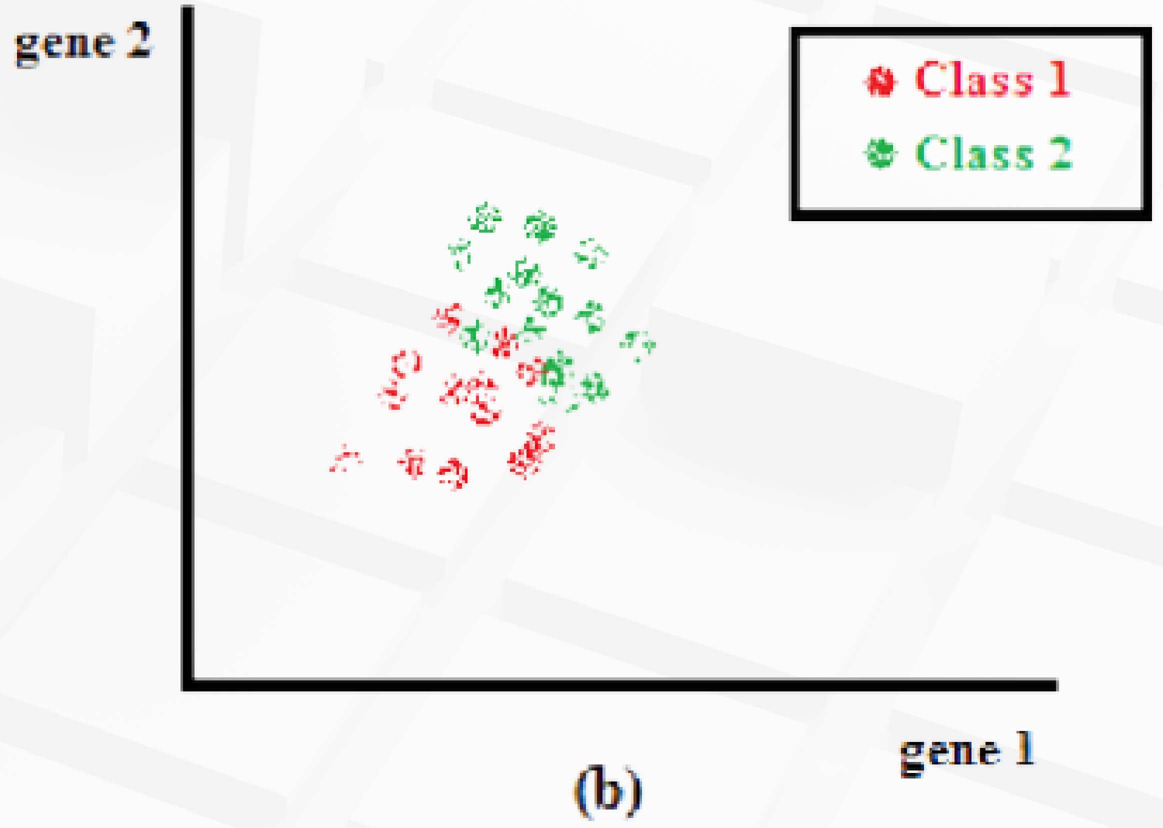
corrélation inter-classes

II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :



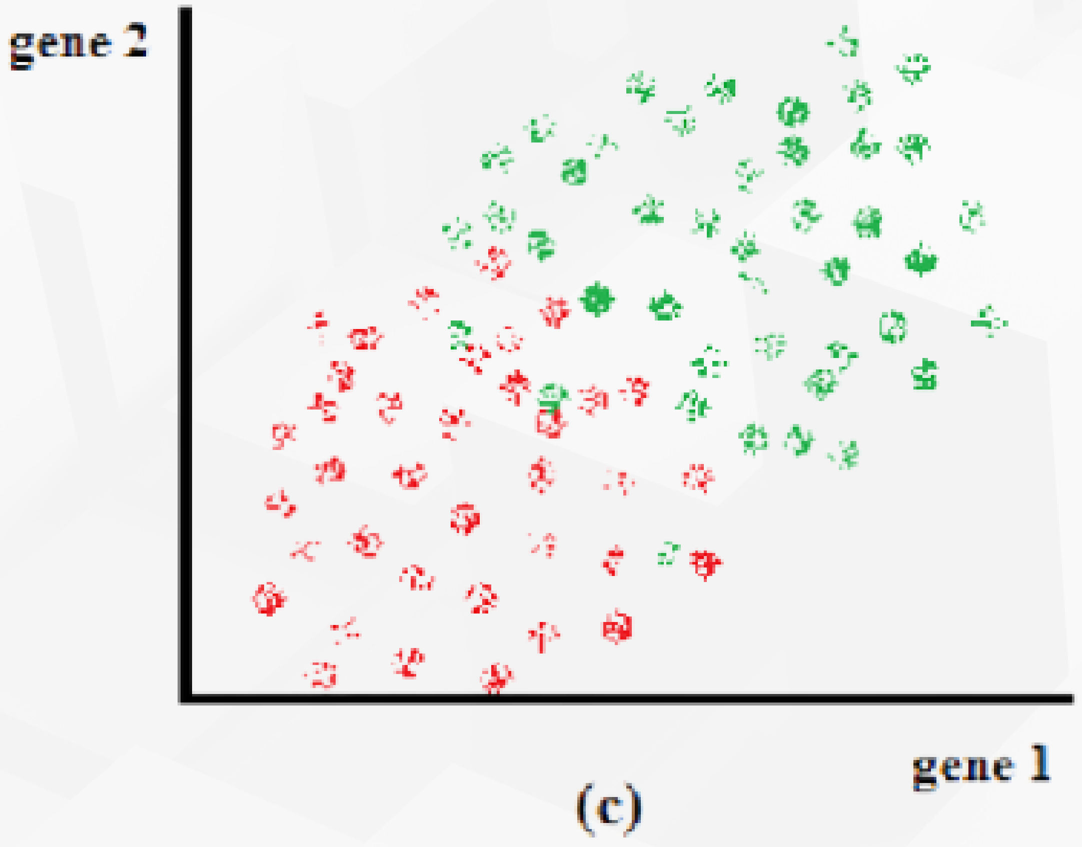
(a) "Distance moyenne élevée et faible variance intra-classe
la meilleure séparabilité."



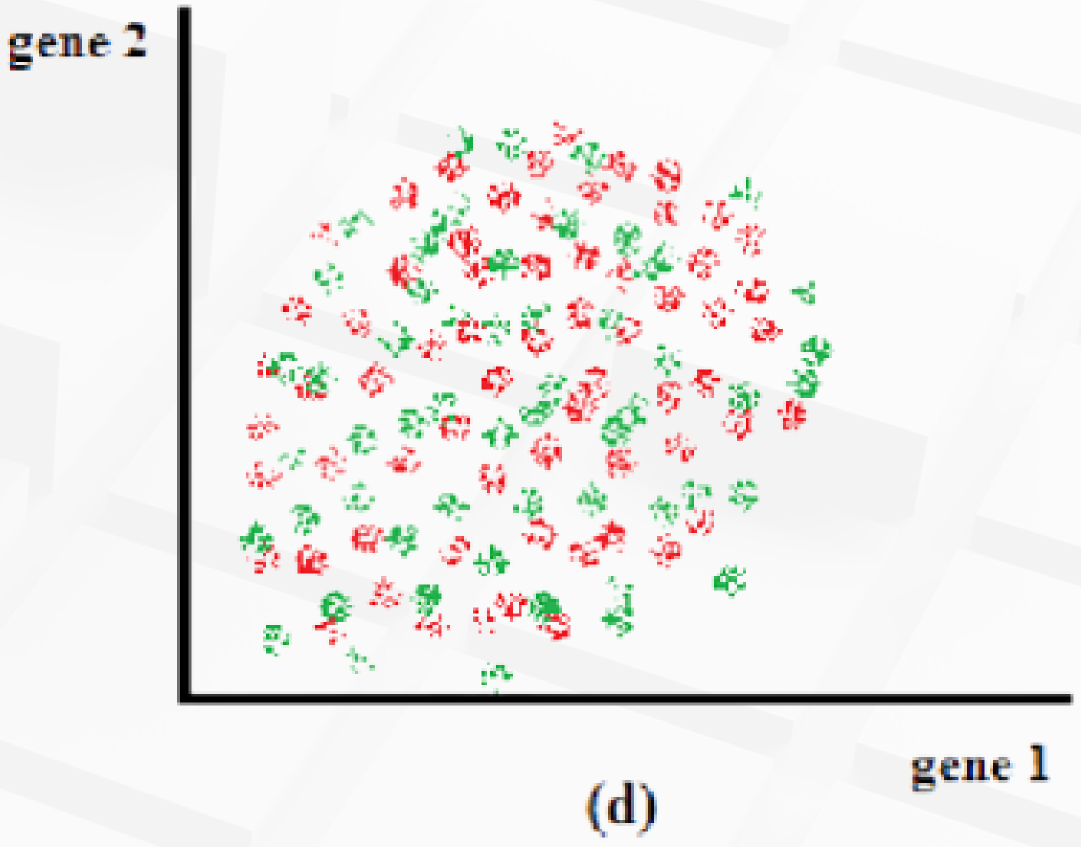
(b) "Distance moyenne faible et faible variance intra-classe."

II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :



(c) Distance moyenne élevée et variance intra-classe élevée..



(d) Faible distance moyenne et variance intra-classe élevée
mauvaise séparabilité

II - Principe de l’analyse factorielle discriminante

II -2 - Données et Notations :

- **Centrage des données** : En AFD comme en analyse en composantes principales (ACP), on suppose que g = 0, c’est à dire que les données sont centrées .
- En particulier, l’écriture des matrices de variance-covariance globale et inter-classe est simplifié :

$$\mathbf{V} = \sum_{i=1}^{n} \frac{1}{n} x_i x_i' \quad \text{et} \quad \mathbf{B} = \sum_{k=1}^{K} \frac{n_k}{n} \mathbf{g}_k \mathbf{g}_k',$$

- et on obtient l’ écriture matricielle

$$\mathbf{V} = \frac{1}{n} \mathbf{X}' \mathbf{X}.$$

III

LES AXES DISCRIMINANTS

III - Les axes discriminants

III -1 -Premier axe discriminant :

Un axe discriminant :

- Une **variable synthétique** capable de discriminer les groupes
- Une **combinaison linéaire** des variables initiales
- Cette variable notée D est définie ici comme un vecteur de Rn combinaison linéaire des vecteurs x1, . . . , xp

$$D = Xu = u_1 x^1 + u_2 x^2 + \ldots + u_p x^p$$

- u = (u1, . . . , up)‘∈ Rp est le vecteur des coefficients de cette combinaison linéaire.

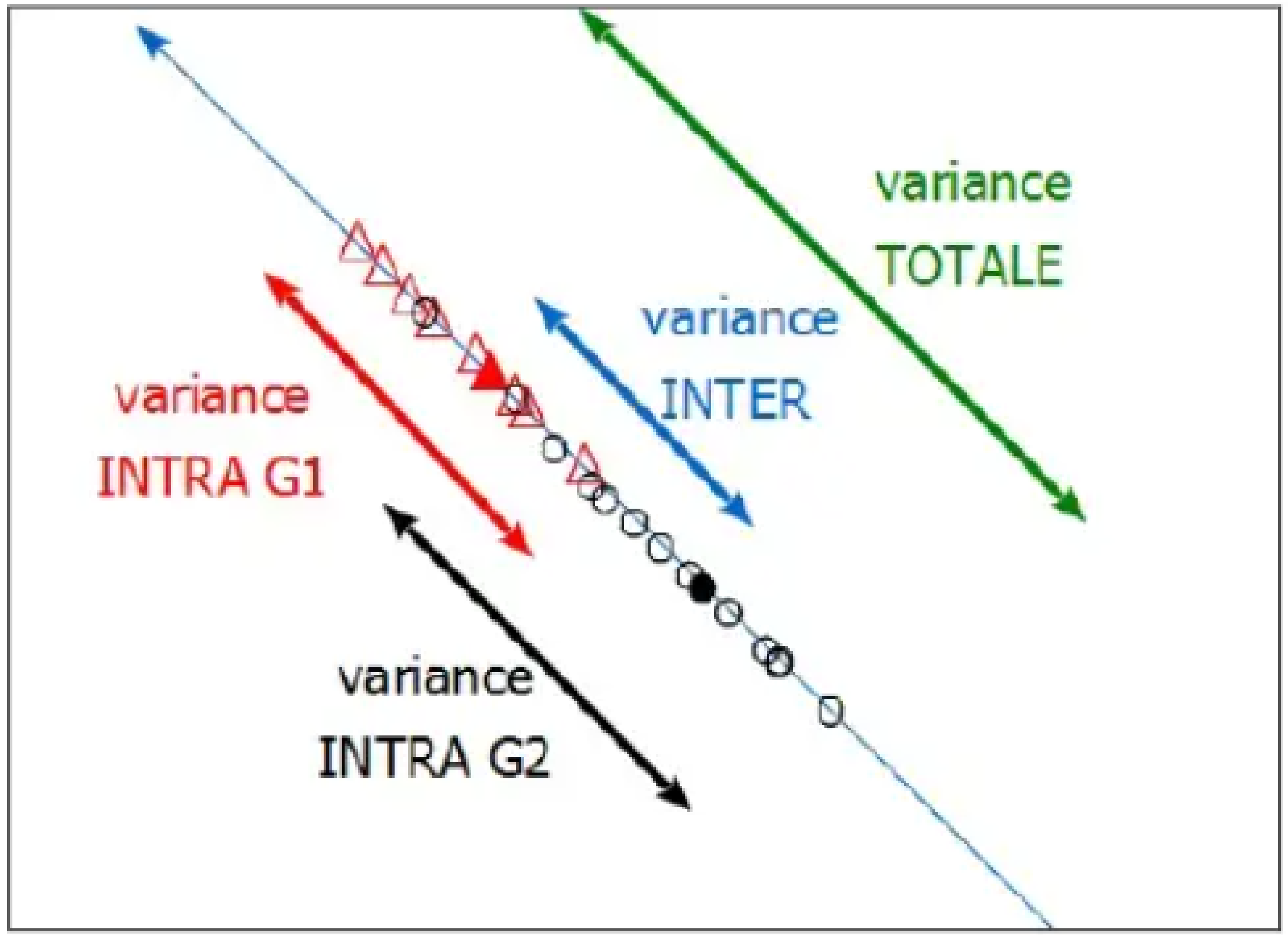
III - Les axes discriminants

III -1 -Premier axe discriminant :

On cherche u1 tel que sur cet axe:

En projection:
- Les groupes se **séparent au mieux**
- Les groupes sont **peu dispersés**

En terme de dispersion / variance:
- la variance INTER groupes est maximale
- la variance INTRA groupes est minimale



Théorème de Huyghens

$$\text{Var Totale} = \text{Var Intra} + \text{Var Inter}$$

III - Les axes discriminants

III -1 -Premier axe discriminant :

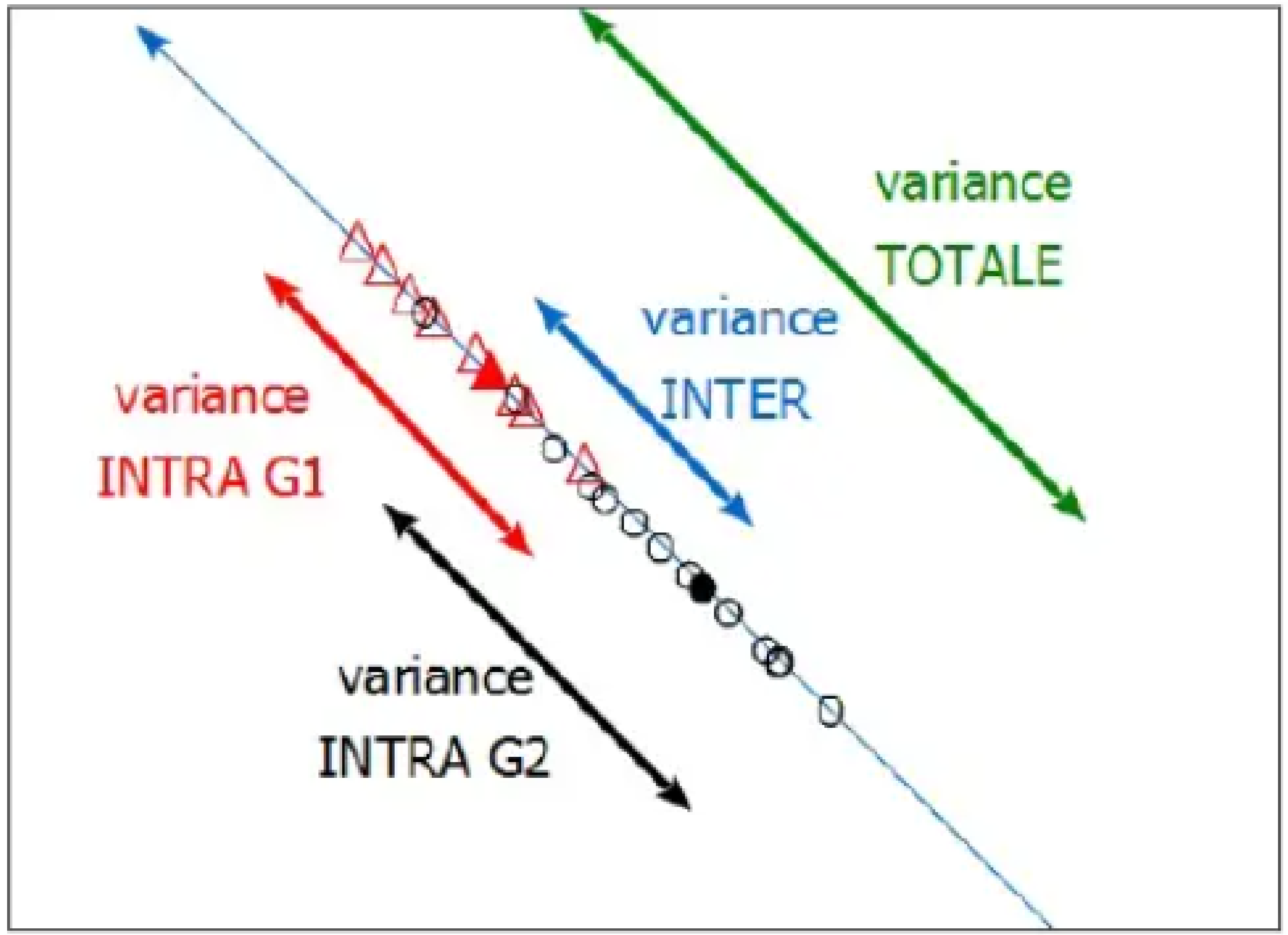
- Le problème d'analyse factorielle discriminante peut être formulé différemment. On cherche maximiser le rapport de corrélation .

- **Le rapport de corrélation:**

$$\lambda_1 = \frac{u'Bu}{u'Tu}$$ **MAXIMAL**

Ou

$$\eta^2 = \frac{\text{Var Inter}}{\text{Var Totale}} = \frac{\text{Var Inter}}{\text{Var Intra}}$$ **MAXIMAL**



III - Les axes discriminants

III -2 -Cas limites:

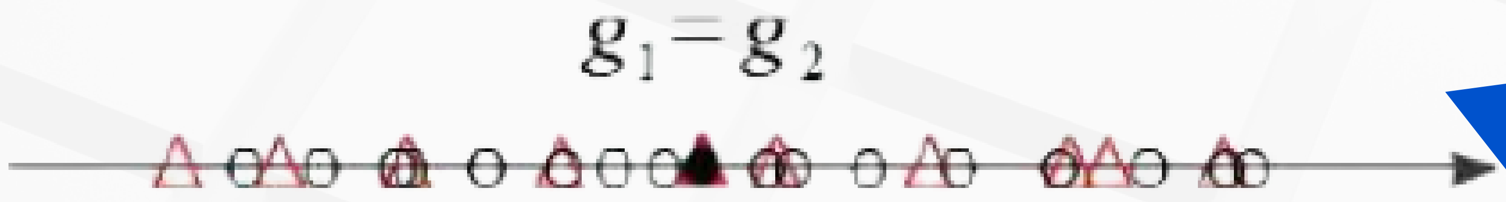
$\lambda_1 \in [0,1]$

$\lambda_1 = 1$

- discrimination parfaite
- les dispersions intra-groupe sont nulles en projection sur la droite discriminante

$\lambda_1 = 0$

- Les centres de gravité des nuages de points sont confondus (aucune discrimination n'est possible).
- le meilleur axe discriminant ne permet pas de séparer les m centres de gravité



III - Les axes discriminants

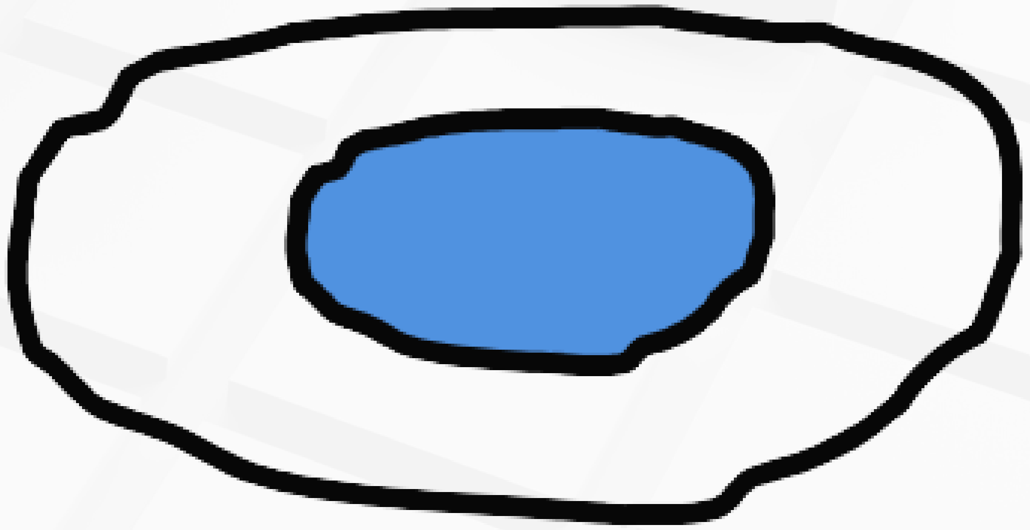
III -2 -Cas limites:

- $\lambda_1 = 1$ séparation parfaite,

pas de variance intra en projection

- 

$\lambda_1 = 0$ groupes concentriques,
pas de variance inter en projection
Pas de séparation linéaire possible
(distance au centre est fonction quadratique des variables)



III - Les axes discriminants

III -2 -Cas limites:

Une séparation parfaite, mais un $\eta^2 \neq 1$

III - Les axes discriminants

Calcul des axes discriminants :

- **Recherche du premier axe factoriel discriminant:**
  - Recherche d'une direction Δu engendrée par un vecteur u.
  - Sur cet axe, le rapport Inter/Intra doit être maximal (ou Inter/Total)

- **Détermination des variances en projection sur Δu**
  - la variance **TOTALE** : $u'Tu$
  - la variance **INTRA** : $u'Wu$
  - la variance **INTER** : $u'Bu$

Avec $u'T\,u = u'W\,u + u'Bu$

III - Les axes discriminants

Calcul des axes discriminants :

- **Problème de maximisation.** De façon équivalente:
  - Déterminer une direction u maximisant

$$\frac{u'Bu}{u'Tu} \quad (1) \qquad \text{ou} \qquad \frac{u'Bu}{u'Wu} \quad (2)$$

- **Solution des problèmes de maximisation (1) et (2)**
  - On choisit le vecteur u associé à la plus grande valeur propre λ de $T^{-1}B$
  - On choisit le vecteur v associé à la plus grande valeur propre μ de $W^{-1}B$

III - Les axes discriminants

Calcul des axes discriminants :

- **Relation entre les deux solutions:**
  - Les vecteurs propres sont identiques
  - Les valeurs propres sont reliées par $\lambda = \frac{\mu}{\mu+1}$ et $\mu = \frac{\lambda}{1-\lambda}$

- **Exemple :**

Valeurs propres :

	F1
Valeur propre	1,016
Discrimination (%)	100,000
% cumulé	100,000

Problème de maximisation (1)

$\lambda = 0,504$

Problème de maximisation (2)

$\mu = \frac{0,504}{1-0,504} = 1,016$

III - Les axes discriminants

III -2 -Kème axes discriminants:

- L'AFD du tableau X s'obtient en cherchant les vecteurs propres uk et les valeurs propres associées de $T^{-1}B$: le $\mathrm{k}^{\mathrm{eme}}$ axe discriminant est le vecteur propre associé à la valeur propre de rang k de cette matrice.

- Ils sont orthogonaux deux à deux

- Nombre d'axes discriminants = min (K-1 , p)

  Dans le cas le plus fréquent (n>p>K) : (K-1) **axes discriminants**

III - Les axes discriminants

III -3 - Illustration :

Illustration pour deux groupes

- L'axe discriminant passe par les 2 centres de gravité = Fonction linéaire discriminante de Fisher

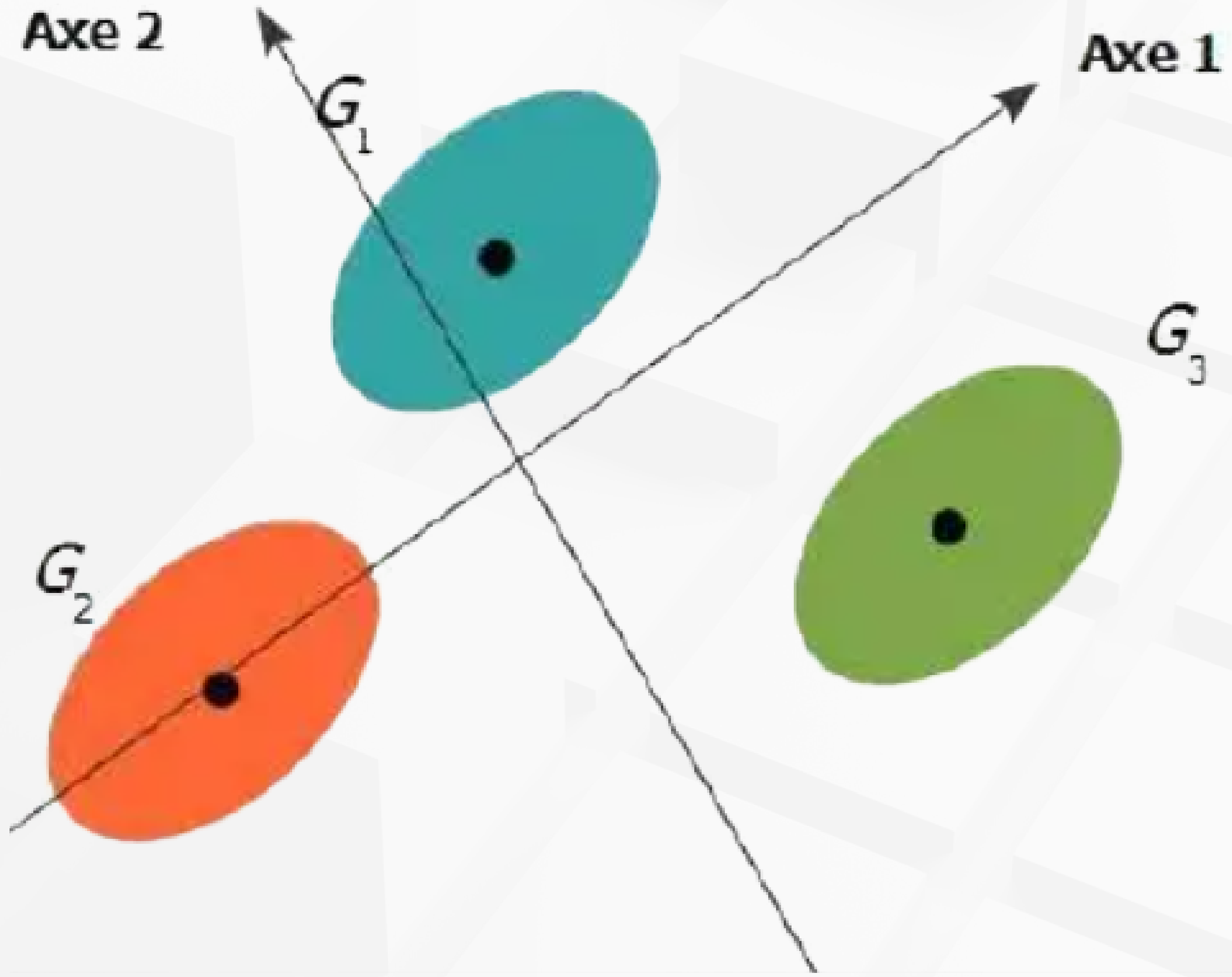
- Il n'y a qu'une seule variable discriminante puisque k-1=1.
- L'axe discriminant est alors nécessairement la droite reliant les deux centres de gravité g2 et g1 .

III - Les axes discriminants

III -3 - Illustration :

Illustration pour trois groupes

- L'interprétation des axes en termes des variables initiales est réalisée au moyen du cercle des corrélation



III - Les axes discriminants

-L'AFD : une ACP particulière

D'après les équations précédentes, l'analyse factorielle discriminante n'est autre que l'A.C.P. du nuage des $\boldsymbol{k}$ centres de gravité avec la métrique V^{-1} .

On en déduit que les variables discriminantes sont non corrélées deux à deux.

Dans le cas où il existe plusieurs axes discriminants $(k > 2)$ on peut utiliser les représentations graphiques usuelles de l'A.C.P. : cercle des corrélations...

IV

LES ÉTAPES ET DÉMARCHES À SUIVRE

IV - Les étapes et démarches à suivre

IV - Les étapes de l'analyse discriminante

Vérification de l'existence de différences entre les groupes

ETAPE 1

Vérification de validité d'étude

ETAPE 2

Vérification du pouvoir discriminant des axes

ETAPE 3

Jugement de la qualité de la représentation du modéle

ETAPE 4

IV - Les étapes et démarches à suivre

IV -1 -1er étape :

On vérifie s'il existe bien des différences entre les groupes grâce à trois indicateurs .

1 . la moyenne/ la variance :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x \qquad V = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Après le calcul des moyennes ou des variances on va les comparer :

- S’il y a une différence on dit qu'il y a une influence.
- S’il n’y a pas une différence, on dit qu'il n y a pas d'influence

IV - Les étapes et démarches à suivre

IV -1 -1er étape :

2. Test F :

C’est un terme générique désignant tout test statistique dans lequel la statistique de test sur la loi de Fisher sous l’hypothèse nulle.

Après la réalisation du test F, il y a deux situations possibles :

- Si F tend vers 0,000 on dit qu'il y a une influence.
- Si F >= 0,01 ou 0,05, on dit qu'il n y a pas d'influence

IV - Les étapes et démarches à suivre

IV -1 -1er étape :

3 . Le lambda de Wilkes :

C'est une statistique qui compare simultanément plusieurs moyennes en mesurant la part d'inerti intra classe dans l'inertie totale .

Après le calcule de lambda de Wilkes on le trouve :

- soit inférieur à 0.9, on dit qu'il y a une influence .
- soit il tend vers 1 et on dit qu'il n y a pas d'influence .

$$\boxed{\text{Wilks' Lambda} = \frac{\text{Somme des carrés intra-classes (X)}}{\text{Somme des carrés totale (X)}} = 1 - \eta^2}$$

IV - Les étapes et démarches à suivre

IV -1 -2eme étape :

On estime la validité d'une analyse discriminante à partir de trois indicateurs :

Le test M de Box: doit être le plus élevé possible.

- Nous utilisons le test de Box pour déterminer si deux ou plusieurs matrices de covariance sont égales.
- La significativité du test de F doit tendre vers 0. S'il est supérieur à 0,05, l'analyse n'est pas valide

La corrélation globale:

- Une corrélation canonique proche de 1 indique une forte corrélation entre les variables originales et l'axe discriminant correspondant. Cela signifie que cet axe discrimine bien entre les groupes définis par la variable dépendante.

IV - Les étapes et démarches à suivre

IV -1 -2eme étape :

Le Lambda de Wilkes : Plus la valeur du Lambda de Wilkes (deuxième colonne) est faible, plus le modèle est bon.
On observe également sa significativité : plus elle est tend vers 0, meilleur, plus le modèle est bon.

IV -1 -3eme étape :

On observe le pouvoir discriminant des axes grâce au tableau << Canonical Discriminant Fonction Coefficients.
On obtient une fonction discriminante sous la forme :

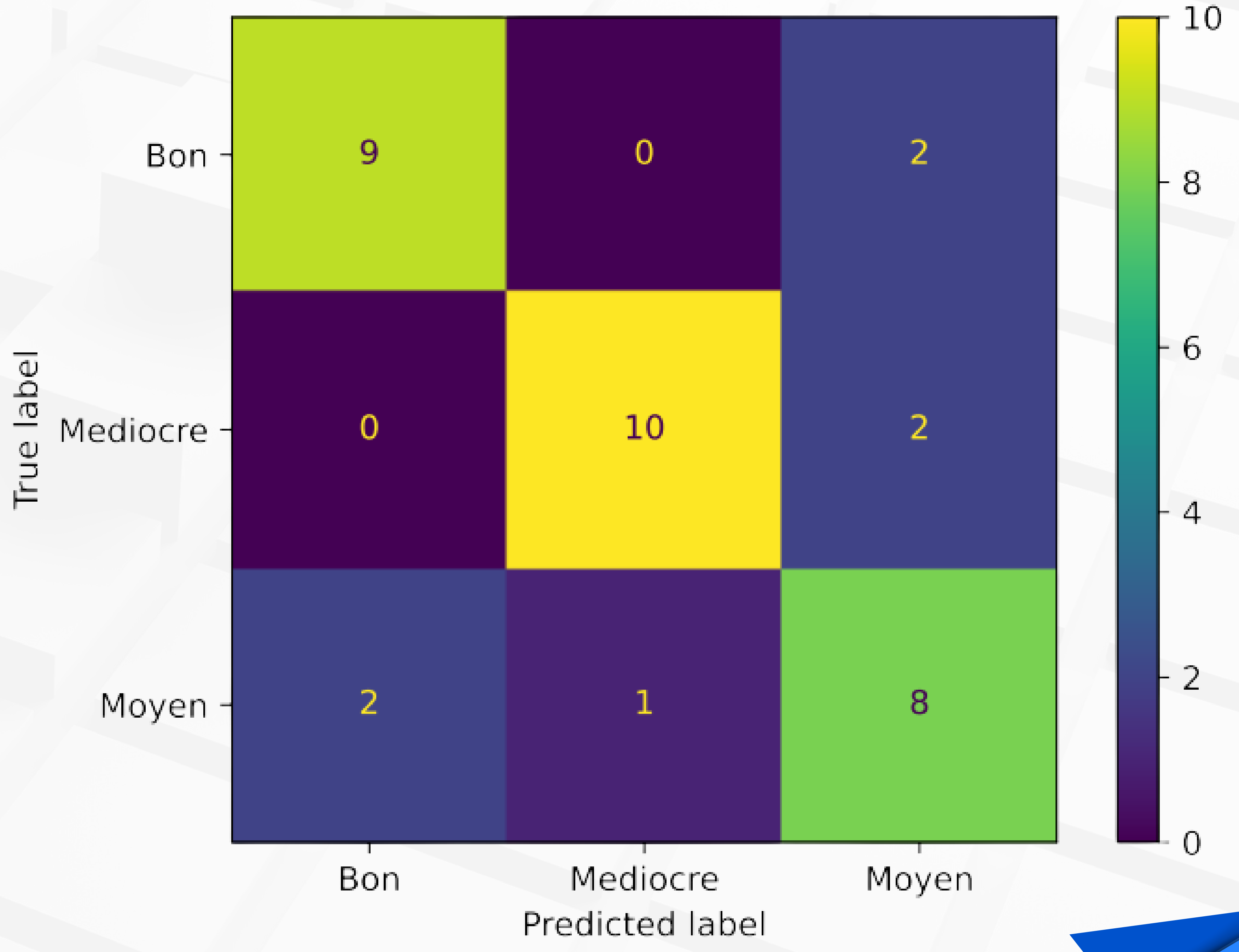
$$Y = a + b_1 \times x_1 + \ldots\ldots\ldots + b_m \times x_m$$

IV - Les étapes et démarches à suivre

IV -1 -4eme étape :

- on observe la qualité de la représentation on s'assure que la fonction discriminante classifie bien les individus en sous-groupes.

- Pour cela on analyse la matrice de confusion qui regroupe les individus bien classés et les mal classé



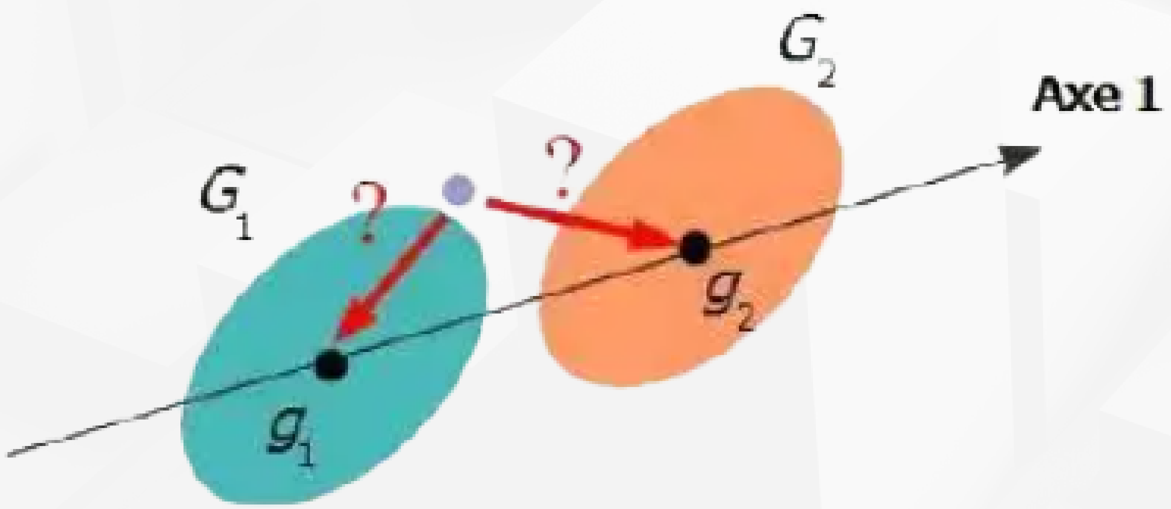
RÈGLE DE CLASSIFICATION

V - Règle de classification

Construction de la règle de classement

Principe de la règle

On affecte un individu au groupe dont il est le **plus proche**



Quelle distance choisir ?

Calcul de la distance entre un individu (i) et le groupe G_k

$$\boxed{d_{W^{-1}}{}^2(x_i, g_k) = (x_i - g_k)'\, W^{-1}\, (x_i - g_k)}$$

W^{-1} est appelée métrique de Mahalanobis

V - Règle de classification

Construction de la règle de classement

Distance entre individus (i) et le groupe Gk

Classification a priori et a posteriori et carrés des distances:

Observation	groupe d'origine	groupe d'affectation	D2(Canada)	D2(Istres)
Obs1	Canada	Canada	4,047	10,867
Obs2	Canada	Canada	2.729	6.650
Obs3	Canada	Canada	3.127	10.649
Obs4	Canada	Canada	4,654	14,876
Obs5	Canada	Istres	8,237	2,208

VI

EXEMPLE D'APPLICATION

VI - Exemple d’application

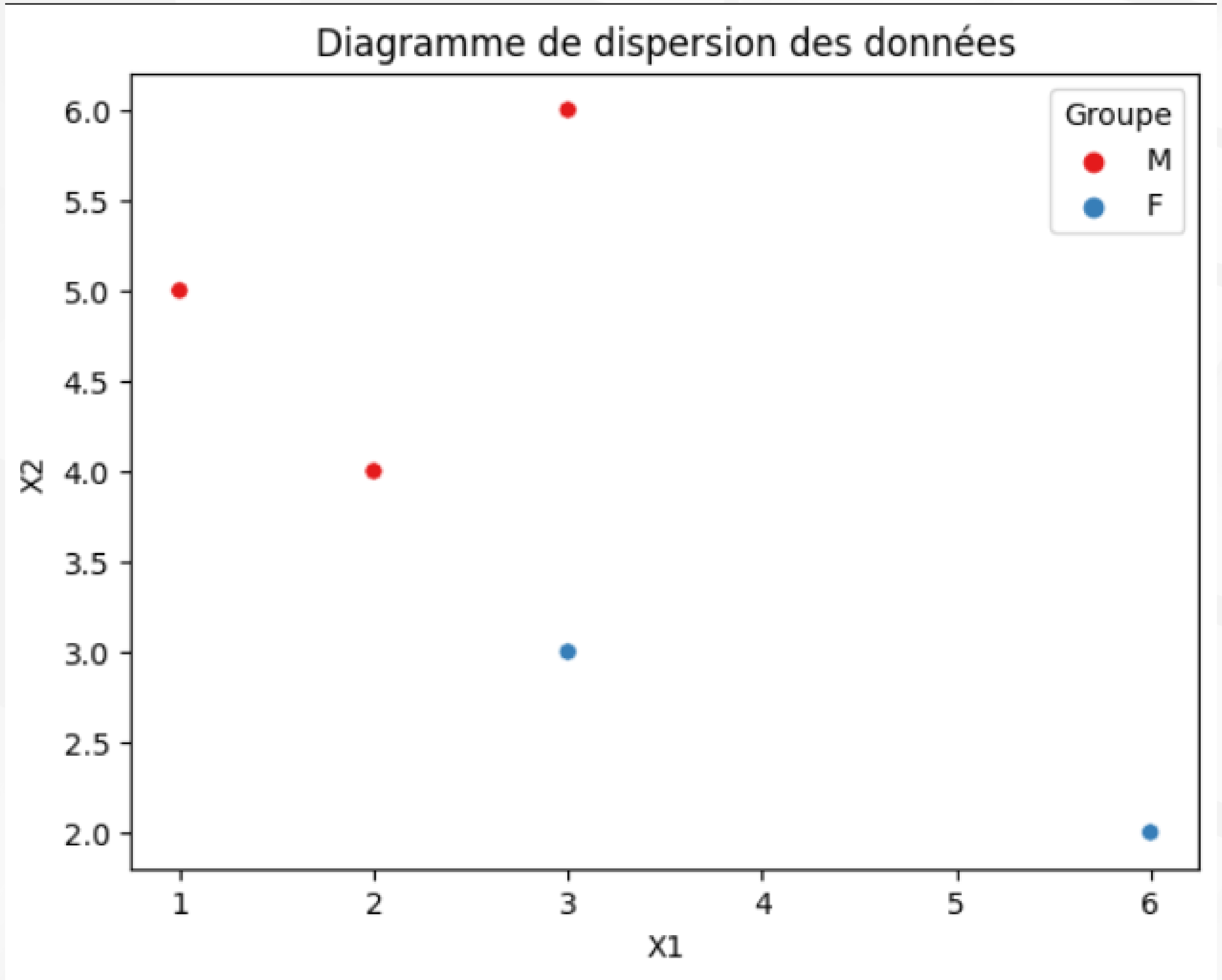
VI -1 - Deux groupes / classes:

On observe deux variables quantitatives X_1 et X_2 sur un ensemble de n=5 individus de même poids, supposés répartis en deux groupes (M : masculin et F : féminin) :

Groupe	X1	X2
M	1	5
M	3	6
M	2	4
F	3	3
F	6	2

VI - Exemple d'application

VI -1 - Deux groupes / classes:



VI - Exemple d'application

VI -1 - Deux groupes / classes:

Grandeurs d'intérêt

$n_1=3,\ n_2=2,\ n=5$,

$$X = \begin{pmatrix} -2 & 1 \\ 0 & 2 \\ -1 & 0 \\ 0 & -1 \\ 3 & -2 \end{pmatrix},$$

$$\tilde{X}_1 = \begin{pmatrix} -2 & 1 \\ 0 & 2 \\ -1 & 0 \end{pmatrix}, \quad G_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$M = \begin{pmatrix} 3/5 & 0 \\ 0 & 2/5 \end{pmatrix},$$

$$\tilde{X}_2 = \begin{pmatrix} 0 & -1 \\ 3 & -2 \end{pmatrix} \quad G_2 = \begin{pmatrix} 1.5 \\ -1.5 \end{pmatrix}$$

VI - Exemple d'application

VI -1 - Deux groupes / classes:

Recherche de l'axe discriminant :

- Matrice variance totale: V = X'X/n $$V = \frac{1}{5}\begin{pmatrix} 14 & -8 \\ -8 & 10 \end{pmatrix}$$
- Matrice de variance inter-classes :

$$B = C'MC \qquad B = \frac{1}{5}\begin{pmatrix} 7.5 & -7.5 \\ -7.5 & 7.5 \end{pmatrix}$$

- Matrice de variance intra-classes :

$$V_k = \frac{1}{n_k}\tilde{X}_k^{'}\tilde{X}_k - G_k G_k^{'} \qquad W = \frac{1}{n}\sum n_k V_k$$

$$V_1 = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad V_2 = \frac{1}{2}\begin{pmatrix} 4.5 & -1.5 \\ -1.5 & 0.5 \end{pmatrix} \quad W = \frac{1}{5}\begin{pmatrix} 6.5 & -0.5 \\ -0.5 & 2.5 \end{pmatrix}$$

VI - Exemple d'application

VI -1 - Deux groupes / classes:

Matrice à diagonaliser :

$$V^{-1}B = \frac{7.5}{76}\begin{pmatrix} 2 & -2 \\ -6 & 6 \end{pmatrix}$$

La valeur propre non nulle de $V^{-1}B$ est = 0.79, qui est le pouvoir de discriminant de l'axe

Le vecteur propre unitaire associé à cette valeur propre est donné par :

$$u = \begin{pmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix}$$

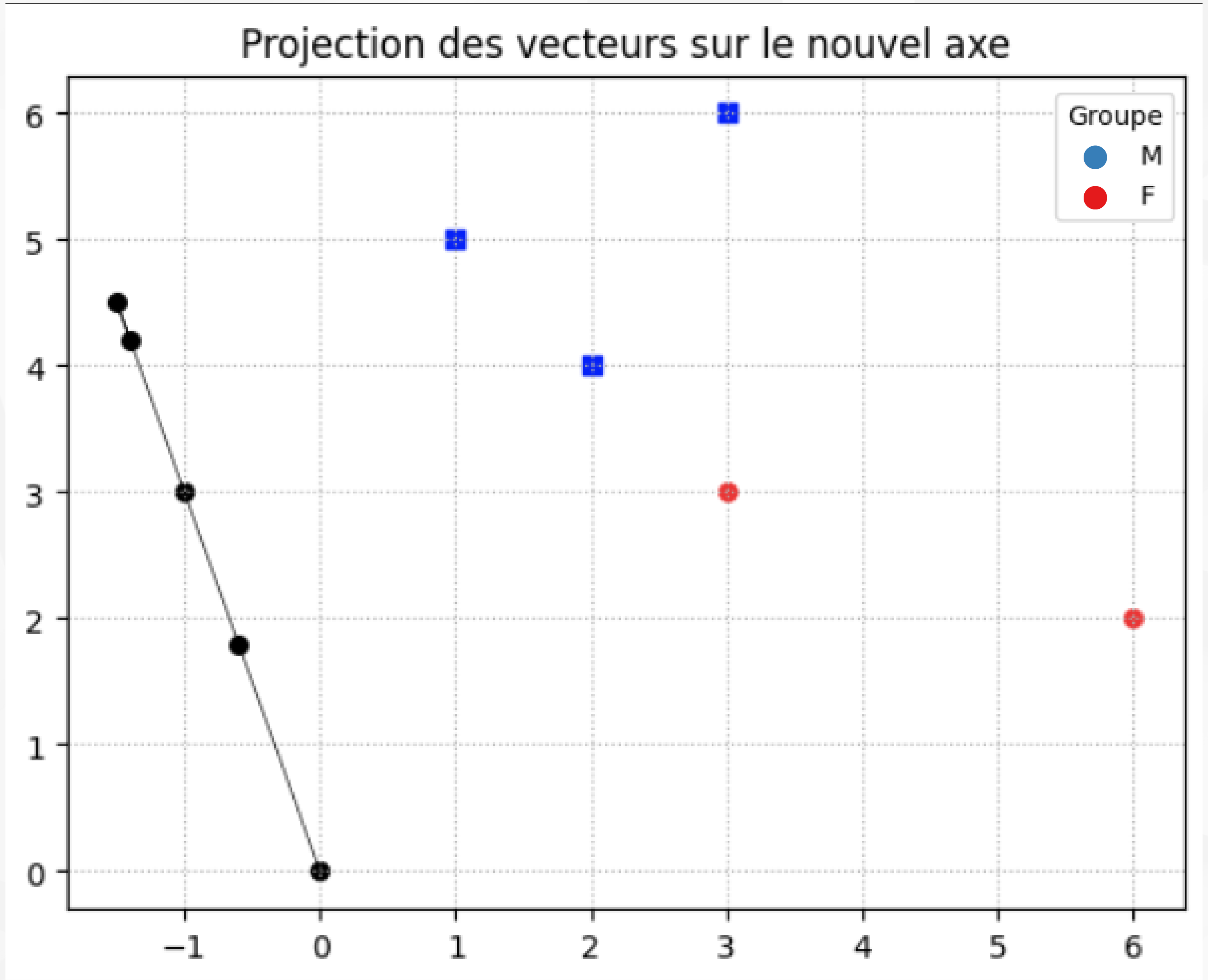
Les coordonnées sur cet axe D=Xu sont :

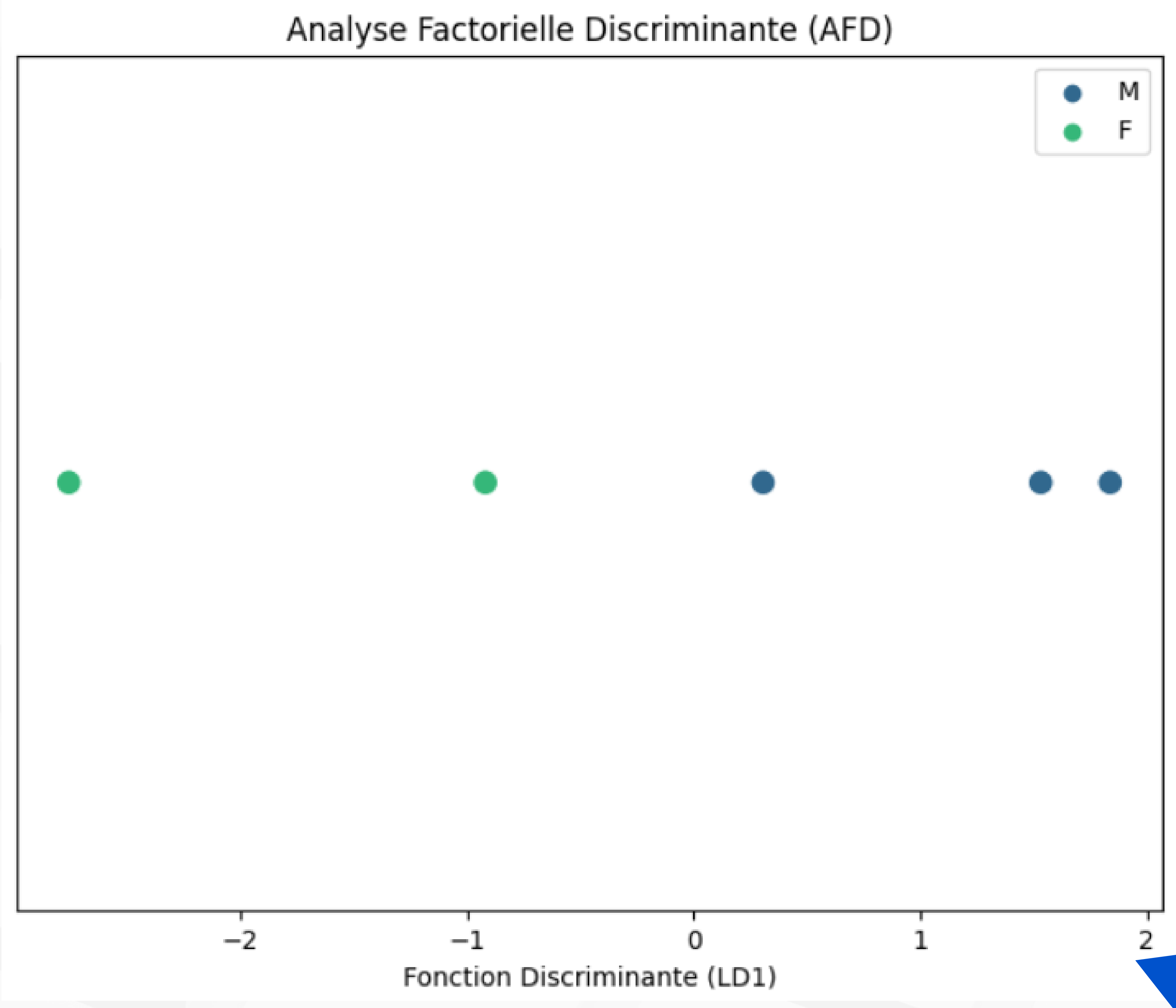
C =

M	5
M	6
M	1
F	-3
F	-9

VI - Exemple d'application

VI -1 - Deux groupes / classes:





VI - Exemple d’application

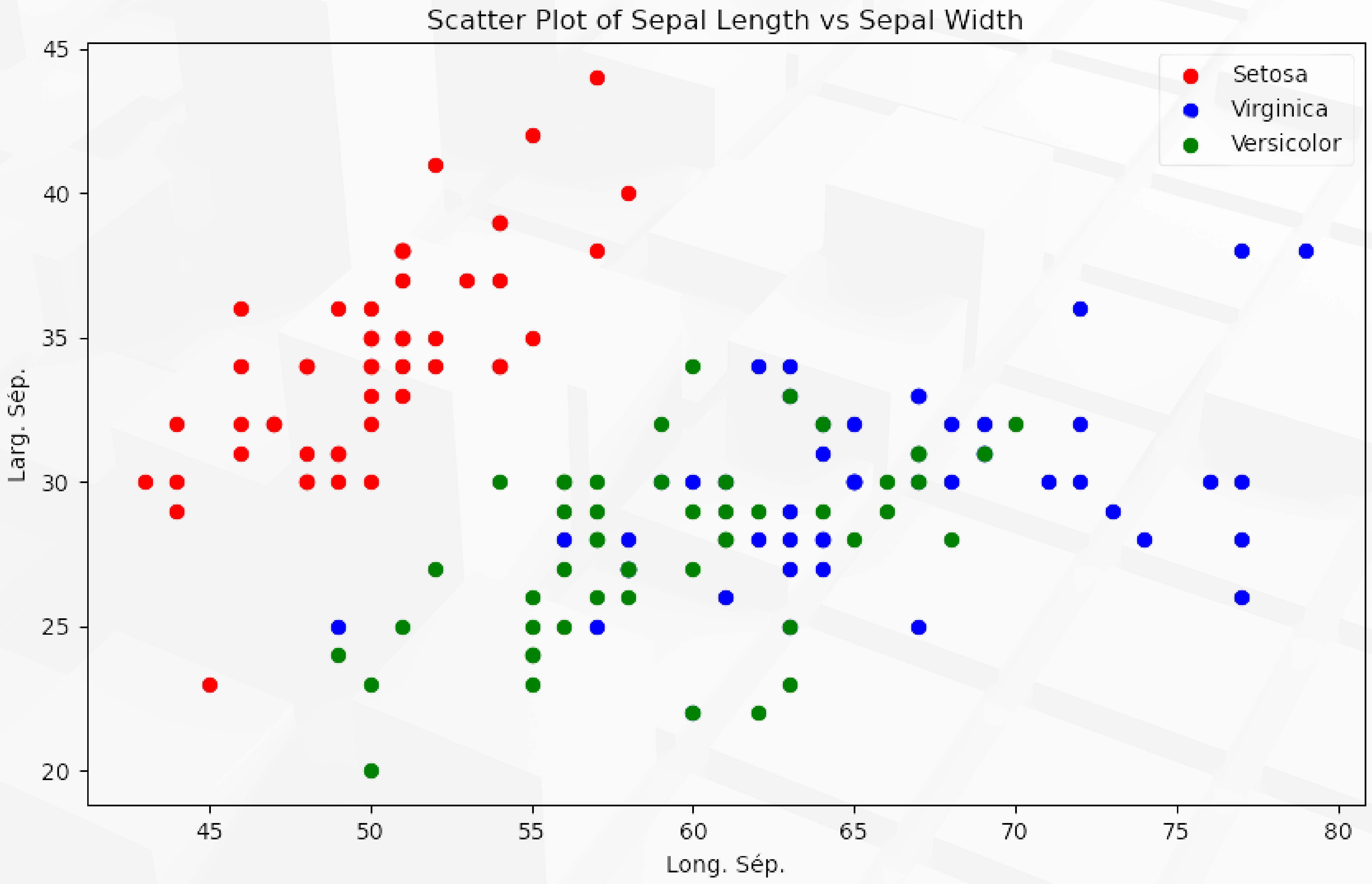
VI -2 - Plus de deux groupes

- Les données proviennent de [Fisher M. (1936). correspondent à 150 fleurs d'Iris, décrites par 4 variables quantitatives (longueur des sépales, largeur des sépales, longueur des pétales, largeur des pétales, et par leur classe.

- Trois différentes classes font partie de cette étude : setosa, versicolor and virginica. Notre but est de tester si les quatre variables descriptives permettent de distinguer les classes, puis de représenter les données dans l'espace factoriel, afin de vérifier visuellement si les classes sont bien discriminées.

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.	Classe
0	50	33	14	2	Setosa
1	64	28	56	22	Virginica
2	65	28	46	15	Versicolor
3	67	31	56	24	Virginica
4	63	28	51	15	Virginica
...	...	...	...	...	...
145	52	35	15	2	Setosa
146	58	28	51	24	Virginica
147	67	30	50	17	Versicolor
148	63	33	60	25	Virginica
149	53	37	15	2	Setosa

VI - Exemple d'application



VI - Exemple d’application

Moyennes par classe :

Classe \ Variable	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Setosa	50,06	34,28	14,62	2,46
Versicolor	59,36	27,70	42,60	13,26
Virginica	65,88	29,74	55,52	20,26

VI - Exemple d’application

Test du Lambda de Wilks :

-Interprétation du test :

- H0 : Les vecteurs moyens des 3 classes sont égaux.
- H1 : Au moins l'un des vecteurs moyens est différent d'un autre.
- Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative H1.

Lambda	0.02
F (Valeur observée)	199.15
F (Valeur critique)	1.97
DDL1	8
DDL2	288
p-value (bilatérale)	<0,0001
alpha	0.05

VI - Exemple d’application

Matrice de covariance inter-classes :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	63,21	-19,95	165,25	71,28
Larg. Sép.	-19,95	11,34	-57,24	-22,93
Long. Pét.	165,25	-57,24	437,10	186,77
Larg. Pét.	71,28	-22,93	186,77	80,41

Matrice de covariance intra-classe pour la classe Setosa :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	12,42	9,92	1,64	1,03
Larg. Sép.	9,92	14,37	1,17	0,93
Long. Pét.	1,64	1,17	3,02	0,61
Larg. Pét.	1,03	0,93	0,61	1,11

Matrice de covariance intra-classe pour la classe Versicolor :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	26,64	8,52	18,29	5,58
Larg. Sép.	8,52	9,85	8,27	4,12
Long. Pét.	18,29	8,27	22,08	7,31
Larg. Pét.	5,58	4,12	7,31	3,91

Matrice de covariance intra-classe pour la classe Virginica :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	40,43	9,38	30,33	4,91
Larg. Sép.	9,38	10,40	7,14	4,76
Long. Pét.	30,33	7,14	30,46	4,88
Larg. Pét.	4,91	4,76	4,88	7,54

VI - Exemple d’application

Matrice de covariance intra-classe totale :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	26,50	9,27	16,75	3,84
Larg. Sép.	9,27	11,54	5,52	3,27
Long. Pét.	16,75	5,52	18,52	4,27
Larg. Pét.	3,84	3,27	4,27	4,19

Matrice de covariance totale :

	Long. Sép.	Larg. Sép.	Long. Pét.	Larg. Pét.
Long. Sép	68,57	-4,24	127,43	51,63
Larg. Sép.	-4,24	19,00	-32,97	-12,16
Long. Pét.	127,43	-32,97	311,63	129,56
Larg. Pét.	51,63	-12,16	129,56	58,10

VI - Exemple d’application

Test de Box (Approximation asymptotique du F de Fisher) :

-2Log(M)	146.66
F (Valeur observée)	7.05
F (Valeur critique)	1.57
DDL1	20
DDL2	77566.75
p-value (bilatérale)	<0,0001
alpha	0.05

-Interprétation du test :

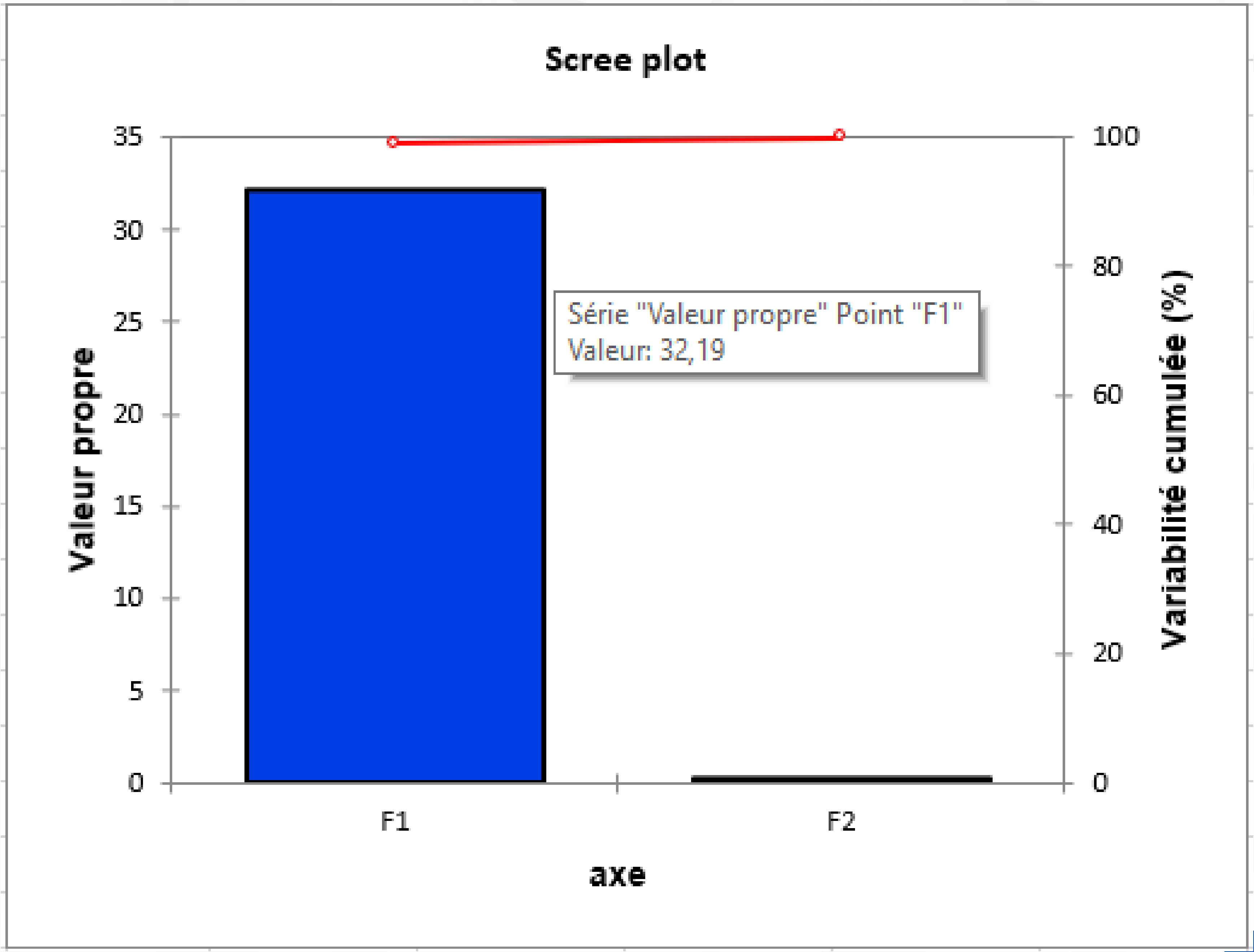
- H0 : Les matrices de covariance intra-classe sont égales.
- Ha : Les matrices de covariance intra-classe sont différentes.
- Etant donné que la p-value calculée est inférieure au niveau de signification alpha=0,05, on doit rejeter l'hypothèse nulle H0, et retenir l'hypothèse alternative Ha.

VI - Exemple d'application

Valeurs propres :

	F1	F2
Valeur propre	32.19	0.29
Discrimination(%)	99.12	0.88
% cumulé(%)	99.12	100.00

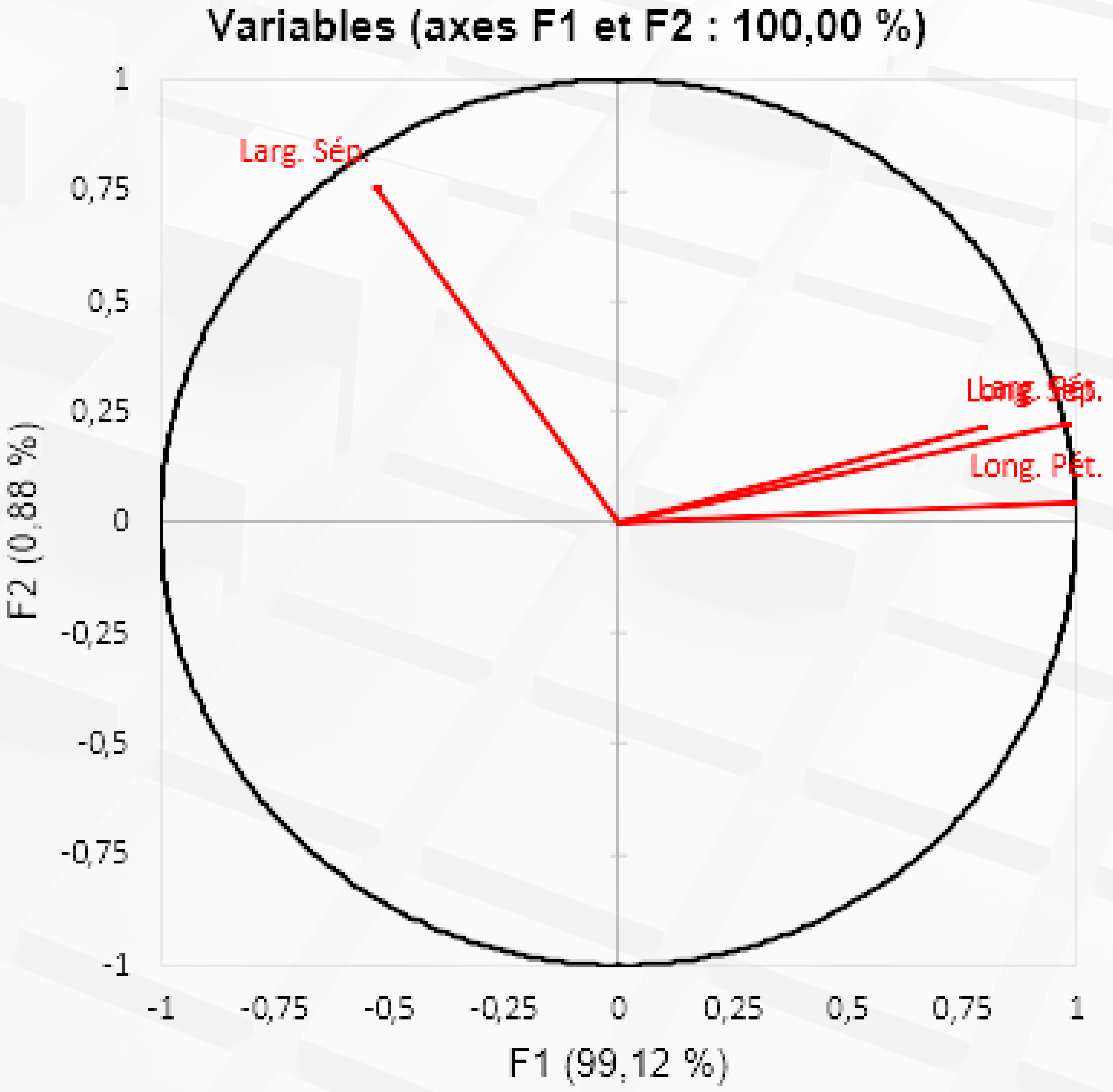
- On peut voir que 99% de la variance sont représentés par le premier facteur.



VI - Exemple d’application

Corrélations Variables/Facteurs :

	F1	F2
Long. Sép	0,79	0,22
Larg. Sép.	-0,53	0,76
Long. Pét.	0,98	0,05
Larg. Pét.	0,97	0,22



Les Intérpretations :

- Le graphique suivant montre comment les quatre variables initiales sont corrélées avec les deux facteurs obtenus .
- F1 est corrélé avec Long. Sép., Long. Pét. et Larg. Pét.
- F2 est corrélé avec Larg. Pét.
- La longueur des pétales semble être la variable la plus discriminante.

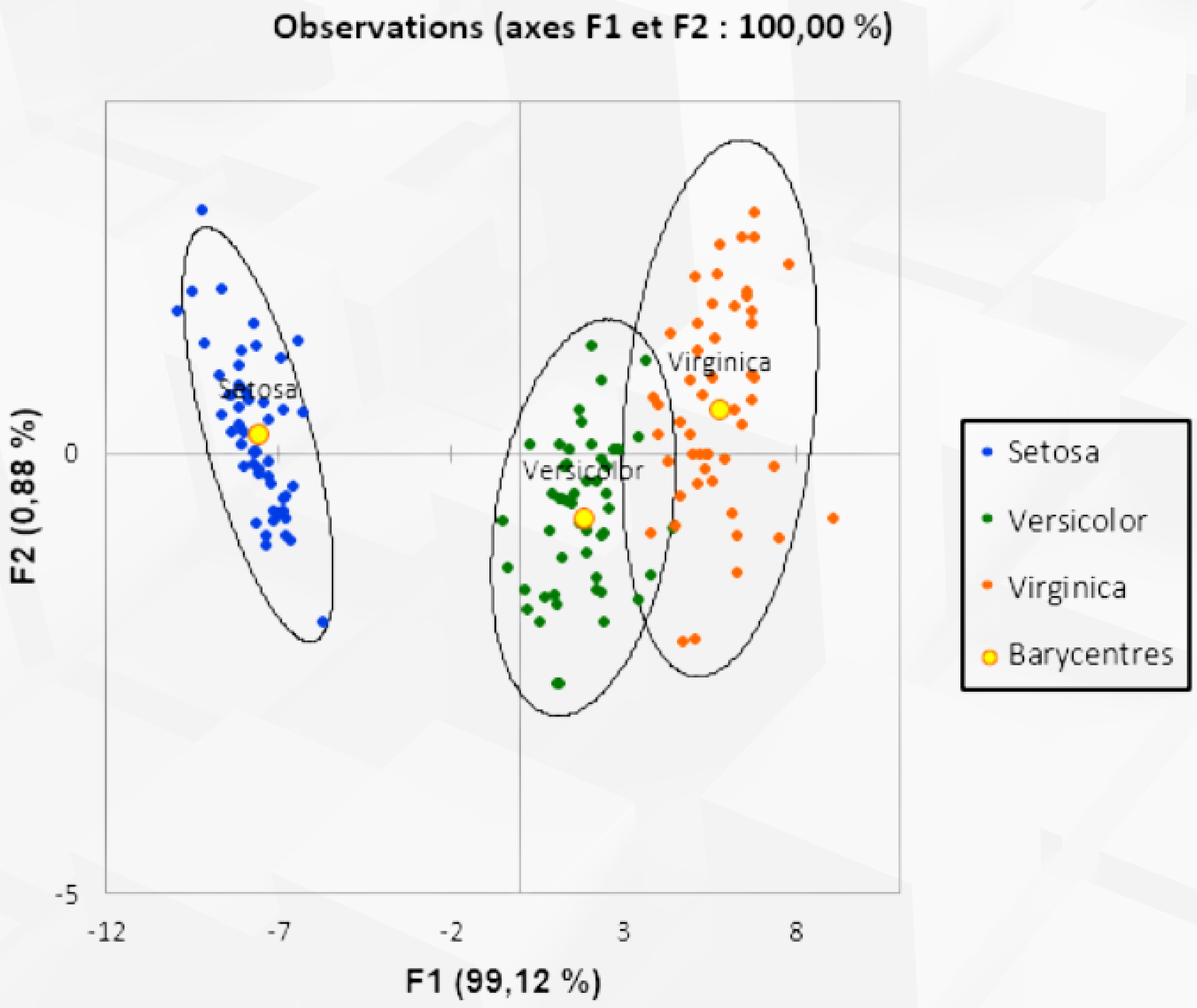
VI - Exemple d'application

Classification a priori et a posterioret carrés des distances :

- On prends les 16 premiers observations

Observation	A priori	A posteriori	F1	F2	D²(Setosa)	D²(Versicolor)	D²(Virginica)
Obs1	Setosa	Setosa	-7.67	-0.13	8.05	110.59	181.34
Obs2	Virginica	Virginica	6.80	0.58	772.50	44.46	13.44
Obs3	Versicolor	Versicolor	2.55	-0.47	425.53	12.20	24.06
Obs4	Virginica	Virginica	6.65	1.81	815.65	55.23	13.86
Obs5	Virginica	**Versicolo**	3.82	-0.94	522.26	15.12	15.98
Obs6	Setosa	Setosa	-7.21	0.36	10.97	105.75	165.71
Obs7	Virginica	Virginica	5.11	1.99	686.17	57.84	20.00
Obs8	Versicolor	Versicolor	3.50	-1.68	433.64	22.23	25.17
Obs9	Versicolor	**Virginica**	3.72	1.04	490.31	18.26	16.90
Obs10	Setosa	Setosa	-8.68	0.88	18.59	140.95	208.12
Obs11	Versicolor	Versicolor	2.29	-0.33	393.99	11.04	22.63
Obs12	Versicolor	**Virginica**	4.50	-0.88	536.26	17.83	14.43
Obs13	Virginica	Virginica	4.97	0.82	625.74	26.48	12.80
Obs14	Versicolor	Versicolor	1.09	-1.63	254.88	11.15	31.63
Obs15	Virginica	Virginica	5.07	-0.03	654.30	19.58	12.83
Obs16	Virginica	Virginica	5.51	-0.04	590.20	28.97	13.59

VI - Exemple d’application



Les interprétations:

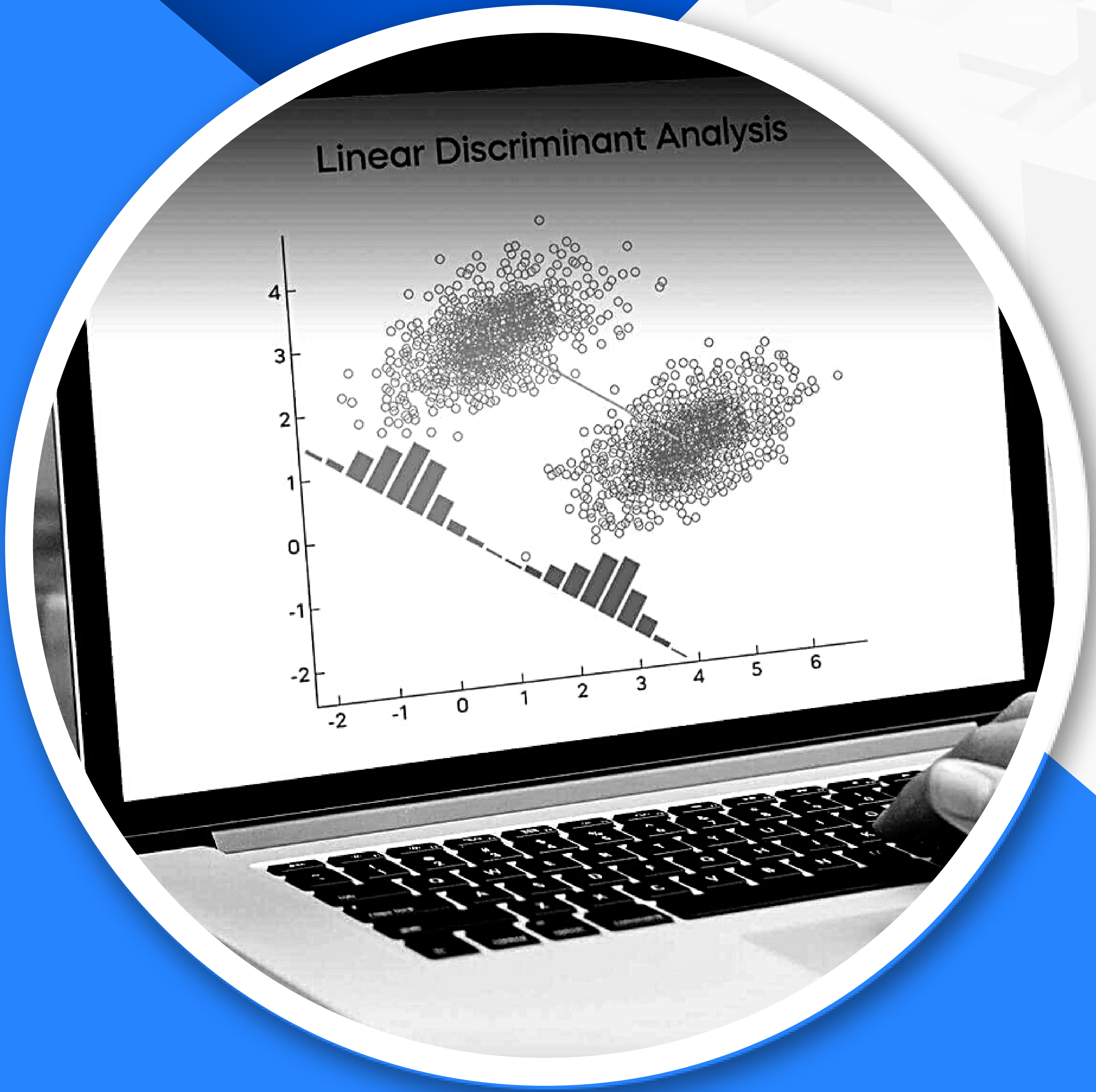
- Le graphique étant bien orthonormé, on peut constater que c'est bien le premier axe qui discrimine le mieux les trois classes.
- Les barycentres des trois classes sont affichés, ainsi que les classes de confiance.

VI - Exemple d’application

Matrice de confusion pour l'échantillon d'apprentissage :

de \ à	Setosa	Versicolor	Virginica	Total	% correct
Setosa	50	0	0	50	100.00%
Versicolor	0	48	2	50	96.00%
Virginica	0	1	49	50	98.00%
Total	50	49	51	150	98.00%

- Matrice de confusion résume l'information concernant les reclassements d'observations, et on peut en déduire les taux de bon et mauvais classement
- Le "% correct" correspond au rapport du nombre d'observations bien classées, sur le nombre total d'observations.



MERCI

POUR VOTRE ATTENTION