Linear Discriminant Analysis

K. Gibert [1,2,3], ***A. Palomino*** [1,2,4]

[1] *Department of Statistics and Operation Research*

[2] *Knowledge Engineering and Machine Learning group at Intelligent Data Science and Artificial Intelligence Specific Research Center*

[3] *Institut Universitari de Recerca en Ciència I Tecnologia de la Sostenibilitat*

[4] *Kantar*

karina.gibert@upc.edu, https://www.eio.upc.edu/en/homepages/karina

Universitat Politècnica de Catalunya, Barcelona

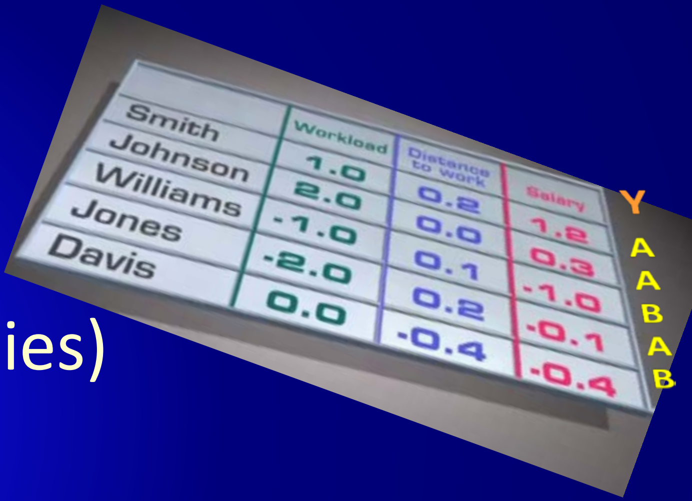
Linear Discriminant Analysis

- Find the function that better predicts the class of an object

- Based on Factorial methods

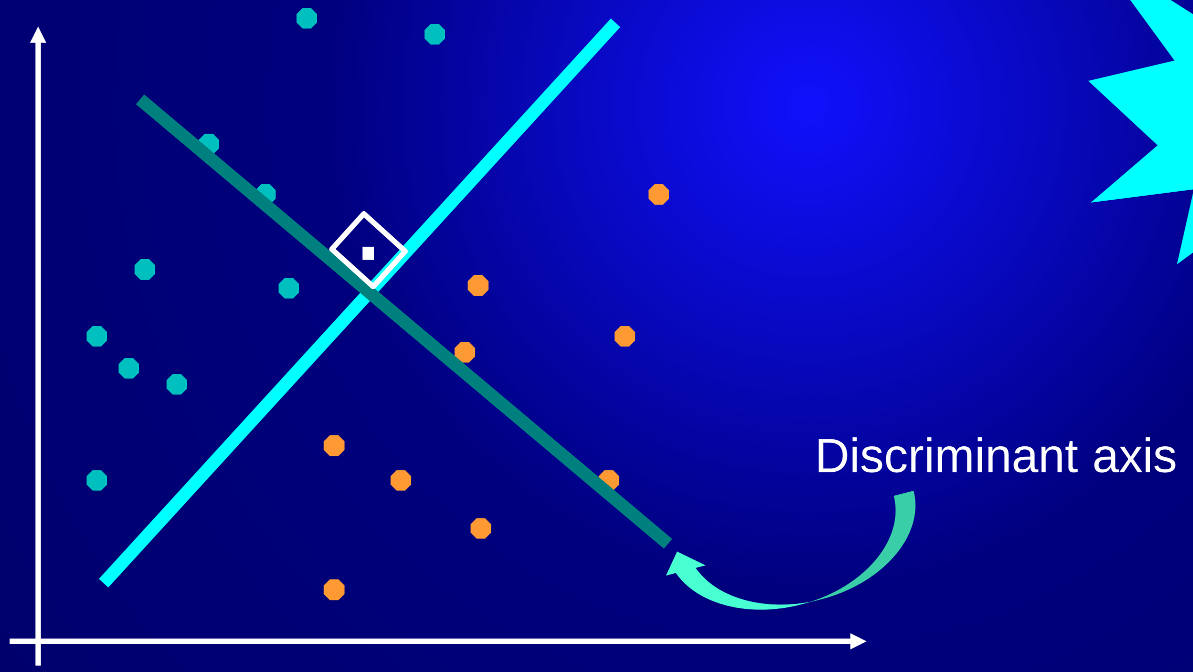
Linear Discriminant Analysis

- Given
  - A data matrix X (nxp) centered
    - n individuals
    - p variables
  - Y categorical variable (m modalities)



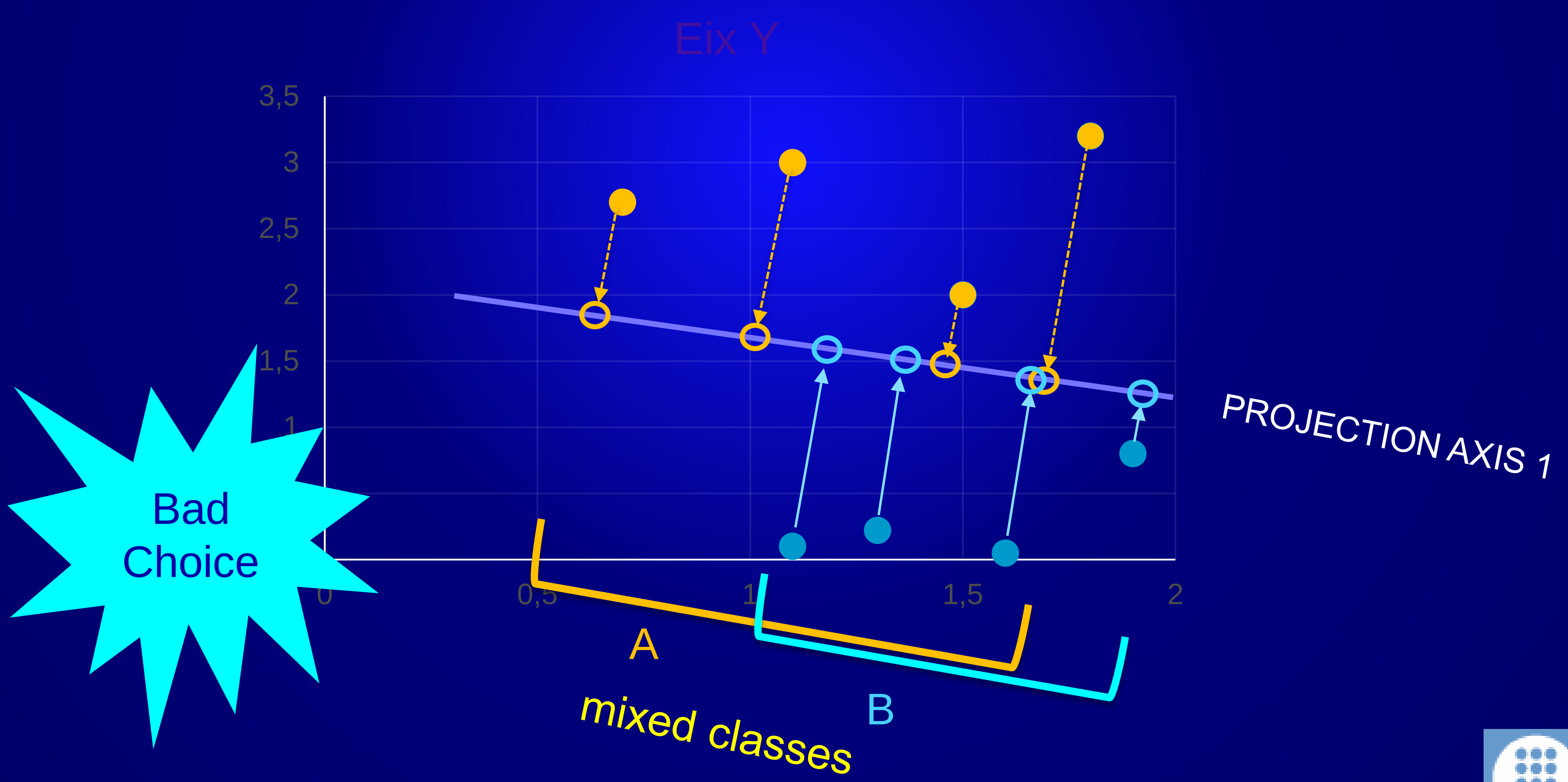
Linear Discriminant Analysis

- Find the axis that better discriminates



Linear Discriminant Analysis

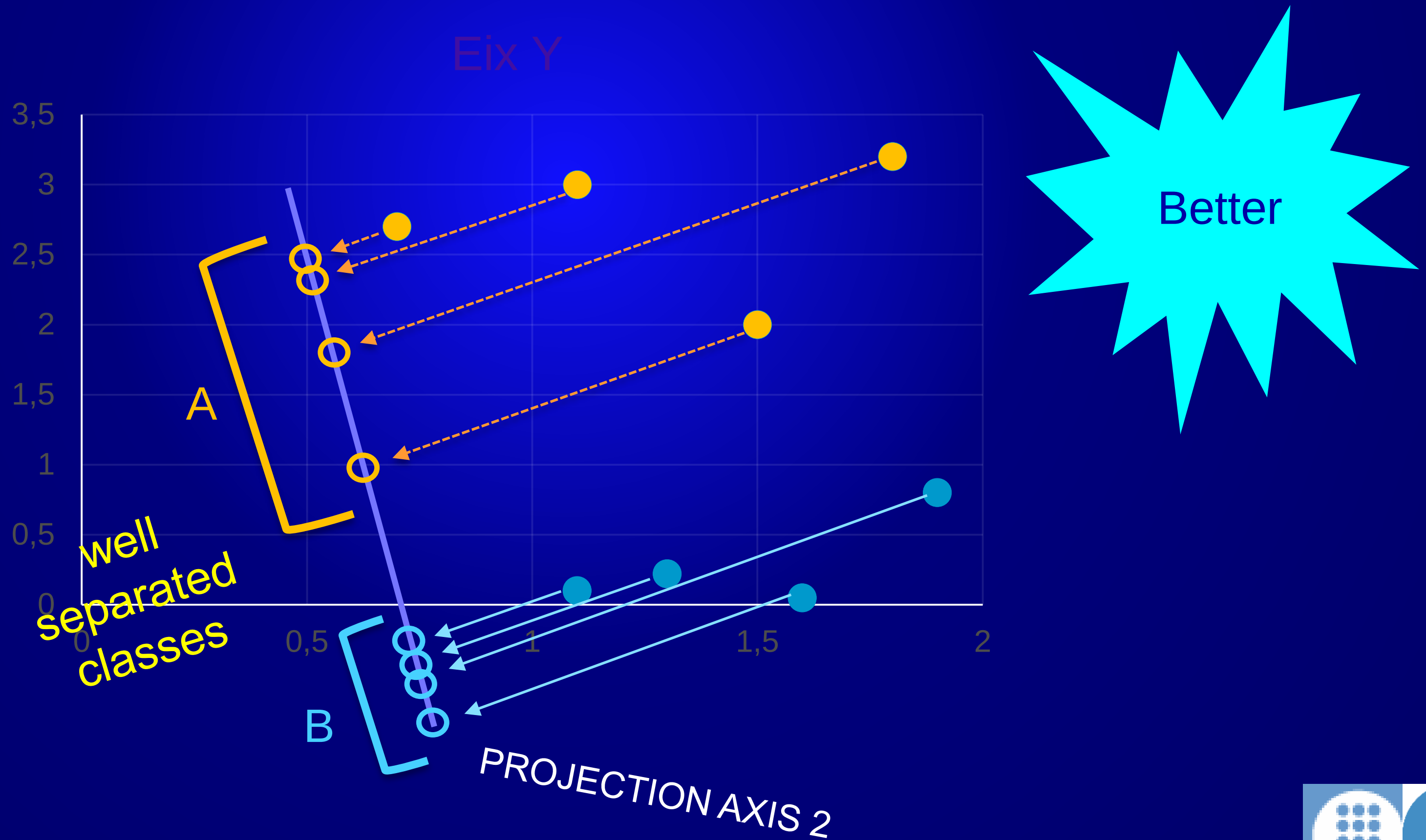
Find the projection space that better separates classes
Find the optimal solution



Linear Discriminant Analysis

Find the projection space that better separates classes

Find the optimal solution



Linear Discriminant Analysis

- Given
  - A data matrix X (nxp) centered
    - n individuals
    - p variables
  - Y categorical variable (m modalities)

- Total Variability of X:
  – $T_{pp} = X'PX$
  – P = diagonal(1/n)

$$T_{pp} = V(X) = \frac{(X - \bar{X})^2}{n}$$

Matricial notation. Care, centered data means $\bar{X}$=0

Linear Discriminant Analysis

- Variability decomposition
  - Total variability $T_{pp} = X^t P X$
  - $T = E + D$ (total)
  - $E = G^t H G$ (between classes)
    - G centroids matrix
    - $H = \mathrm{diag}\,(n_k/n)$
  - $D = X'^t P X$ (within classes)
    - $X' = (x_{ij} - g_{kj})$
    - $P = \mathrm{diag}\,(1/n)$

Linear Discriminant Analysis

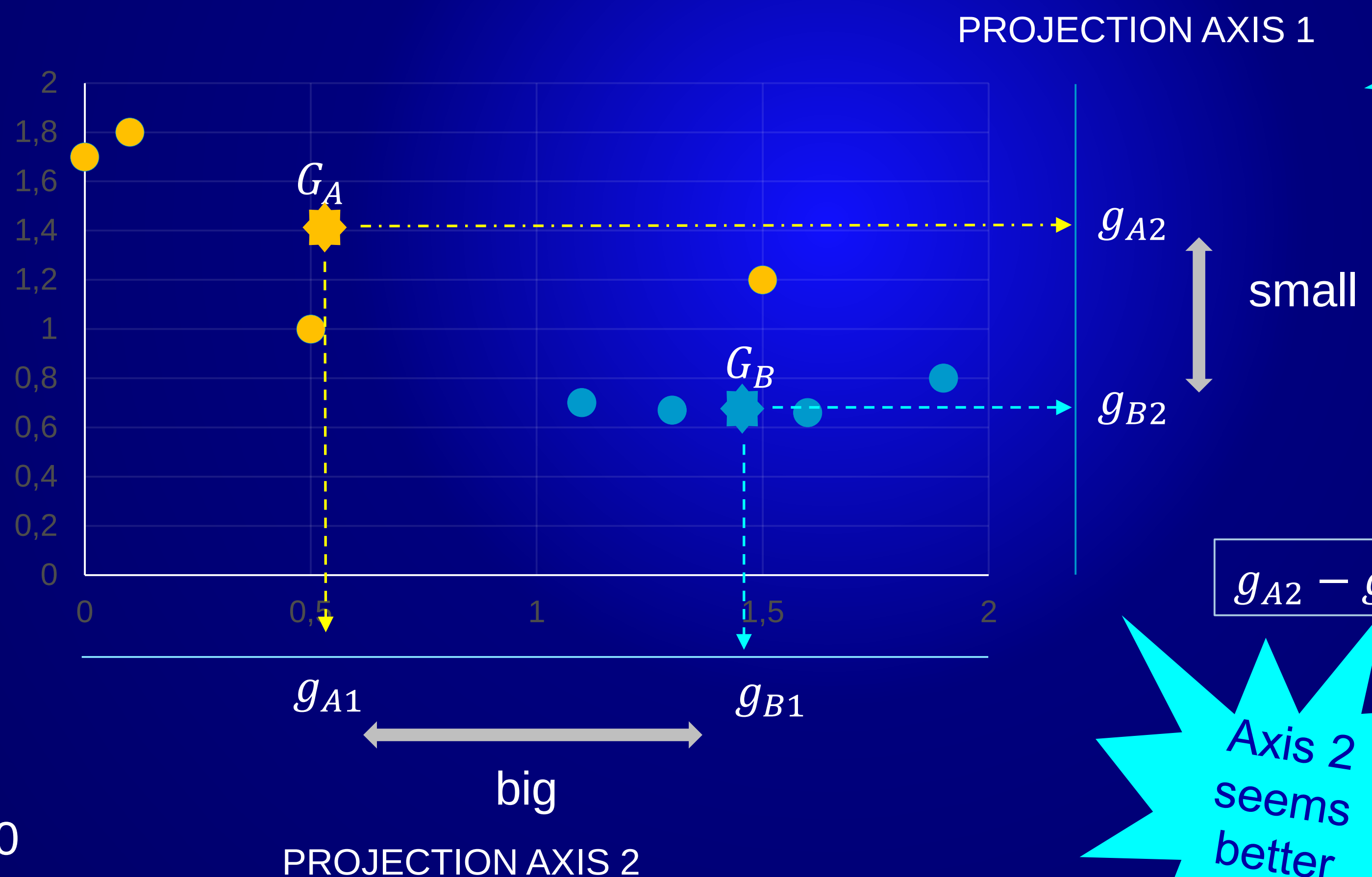
- Factorial spaces
  - $<X_{nxp}, \mathbb{I}_p, P_{nxn}>$
  - $<G_{mxp}, \mathbb{I}_p, H_{mxm}>$
  - $<G_{mxp}, D_{pxp}, H_{mxm}>$

  - *Projected variabilities on direction* $\vec{w}$
  - *Total variability projected* $\vec{w}^t T \vec{w}$
  - Between Classes variability projected $\vec{w}^t E \vec{w}$
  - Within Classes variability projected $\vec{w}^t D \vec{w}$

Linear Discriminant Analysis

Distances must be big between projected centroids

G_A=(g_{A1}, g_{12}) G_B =(g_{B1}, g_{B2})



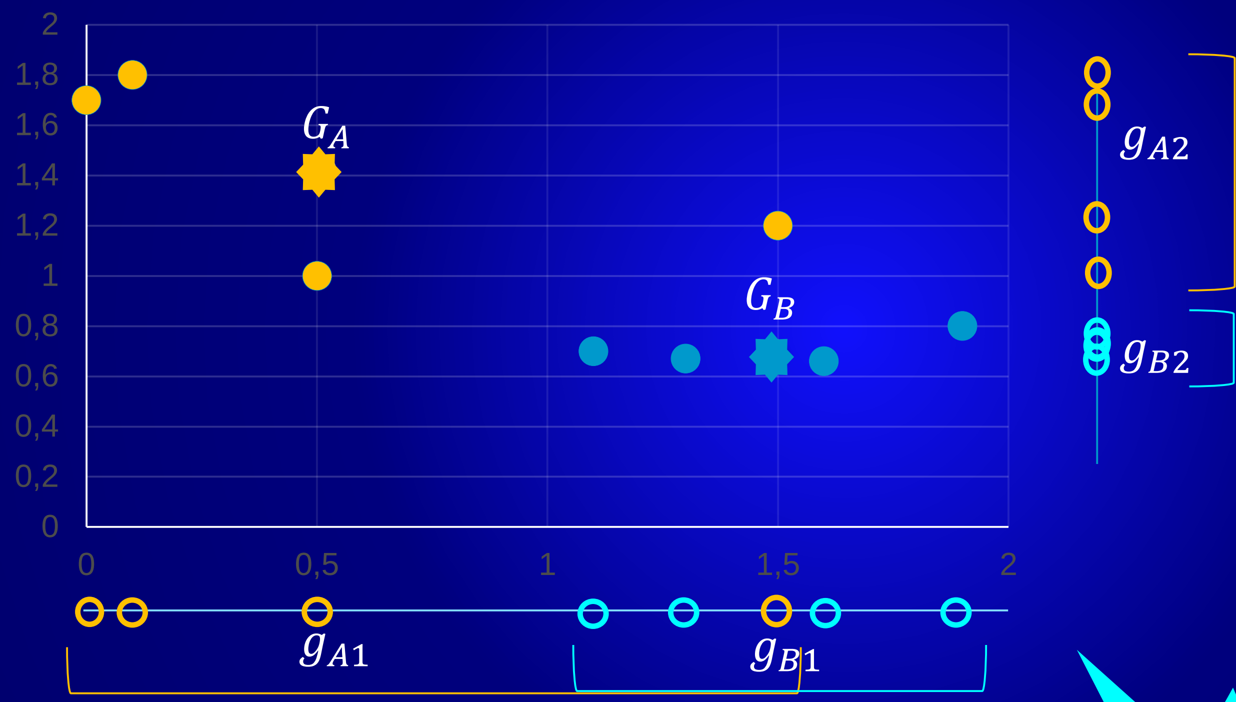
$$\boxed{g_{A2} - g_{B2}| < |g_{A1} - g_{B1}|}$$

Trade-Off

Axis 2 seems better

Linear Discriminant Analysis

Distances must be big between projected centroids
But also variance must be small on the projection space



Compact groups
Within variance small
No overlapping

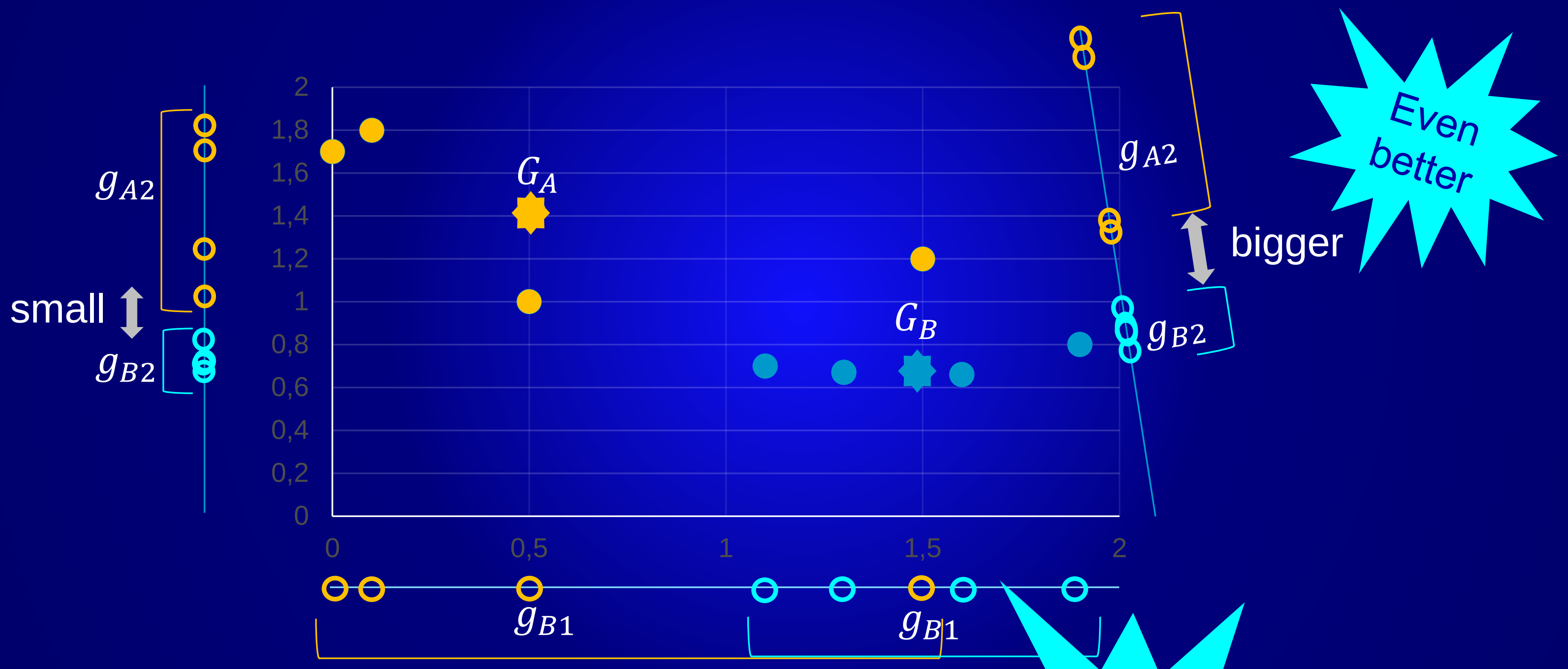
No compact groups
Within variance big
Class overlapping

Axis 1 is better

Trade-Off E vs D

Linear Discriminant Analysis

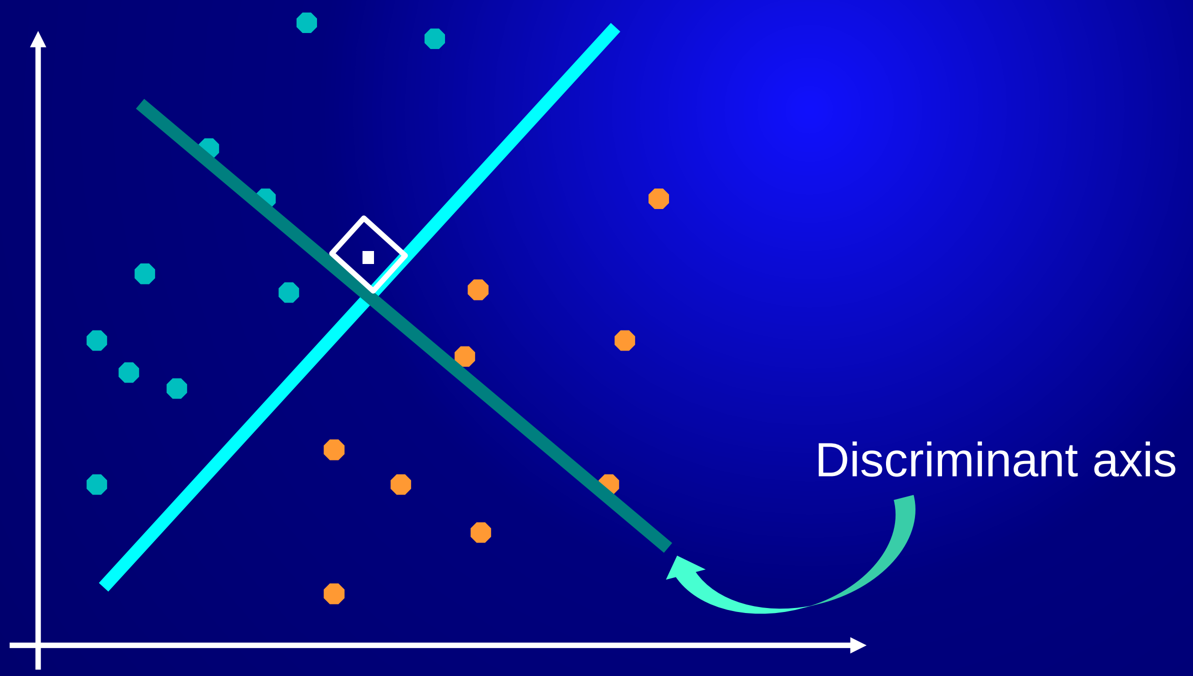
Distances must be big between projected centroids
But also variance must be small on the projection space



No compact groups
Within variance big
Class overlapping

Linear Discriminant Analysis

- Find the axis that better discriminates
- Find the axis that maximizes $\vec{w}^{t} E \vec{w}$



Linear Discriminant Analysis

- <G,D^{-1}, H>
- Search direction $\vec{w}$ such that $\vec{w}^t E \vec{w}$ is maximized
- Diagonalization of

$$\vec{u}^t D^{-1} G^t \; H \; G \; D^{-1} \vec{u}$$

subject to $\vec{u}^t D \, \vec{u} = 1$

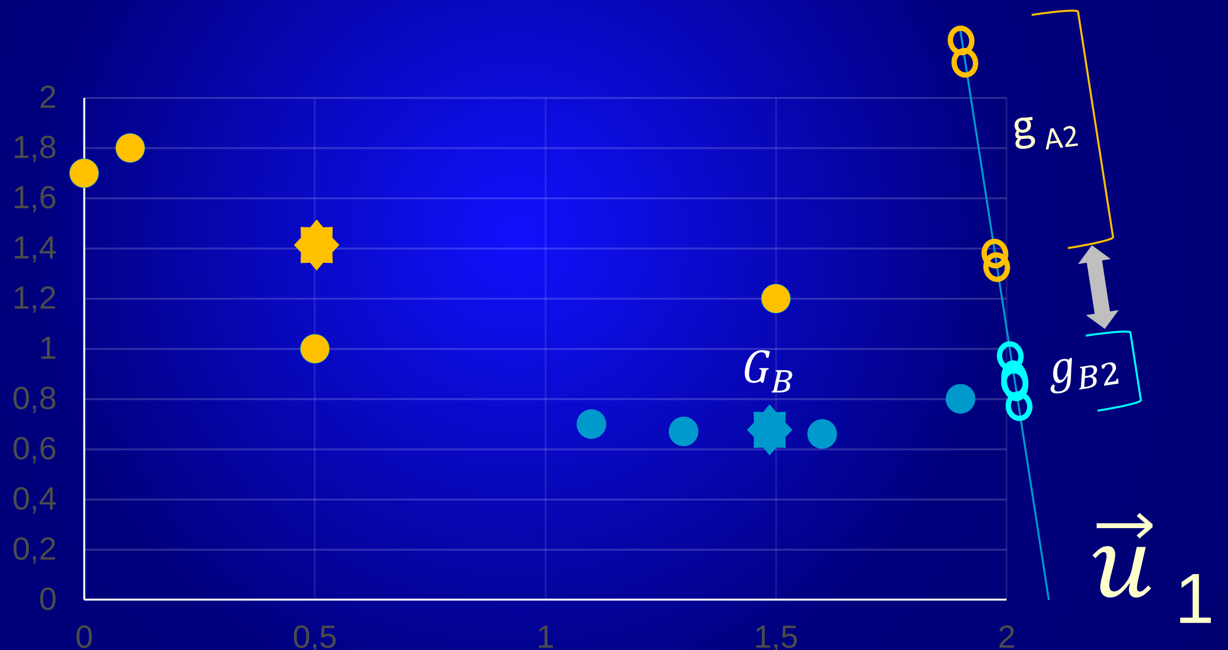
- The first eigen value λ_1 points to the eigen vector $\vec{u}_1$ that gives solution

  u_α , α =(1:p), principal factors of X : good rotation directions

$$Projection\text{: } \Psi_\alpha = X\, u_\alpha$$

Linear Discriminant Analysis

Distances must be maximal between projected centroids
But also variance must be minimum on the projection space



Linear Discriminant Analysis

Solution

Linear discriminant function ψ_1

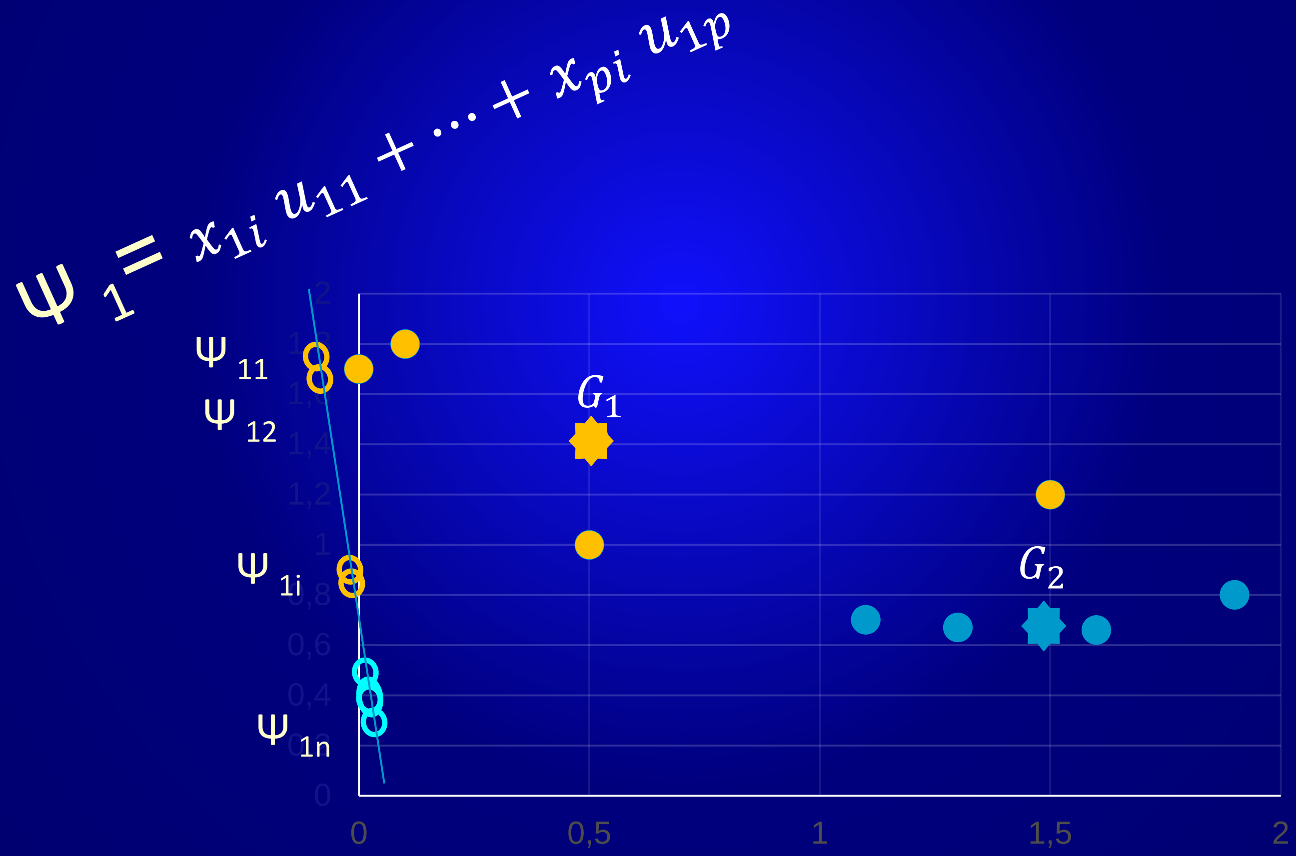
- Ψ_1 : Linear combination of original variables

$$\Psi_{1i} = x_{1i}\, u_{11} + x_{2i}\, u_{12} + \cdots + x_{pi}\, u_{1p}$$

$\Psi_{1\mathrm{i}}$ Coordinates of objects in First factorial plane

Linear Discriminant Analysis

Distances must be maximal between projected centroids
But also variance must be minimum on the projection space



Linear Discriminant Analysis

Solution

Linear discriminant function ψ_1

- Ψ_1 : Linear combination of original variables

$$\Psi_{1i} = x_{1i}\, u_{11} + x_{2i}\, u_{12} + \cdots + x_{pi}\, u_{1p}$$

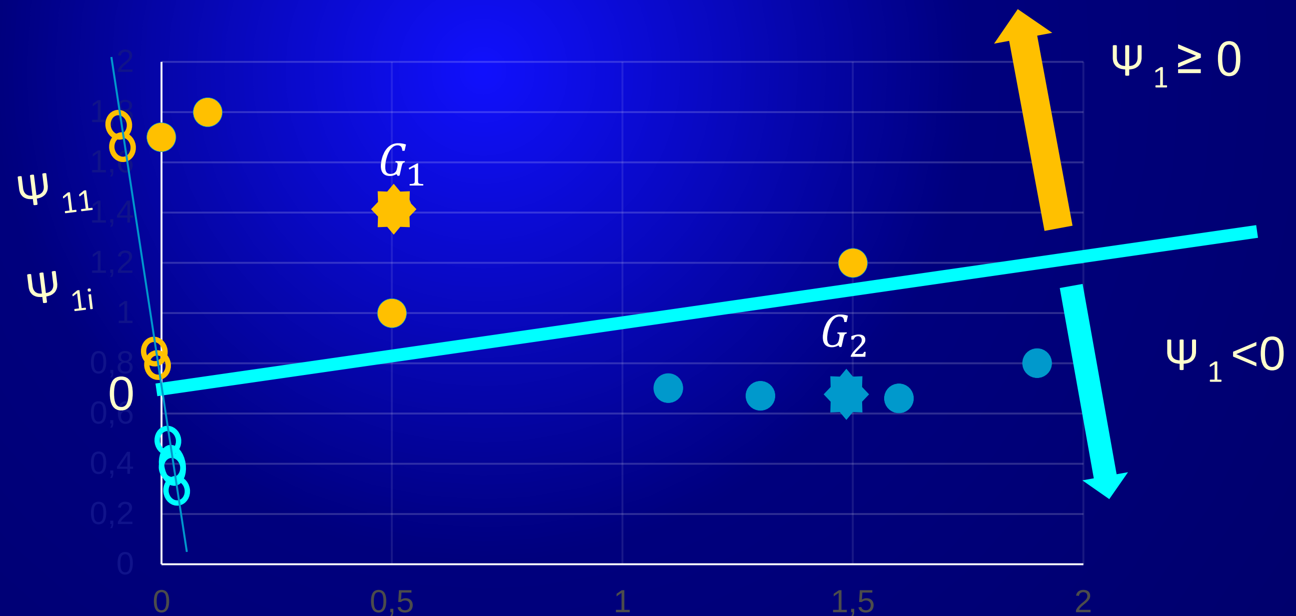
Model Use

- Given an object i$=(x_{1i}, \ldots, x_{pi})$ decision rule is:
  - If $\Psi_{1i} \geq 0 \Rightarrow$ Class A
  - If $\Psi_{1i} < 0 \Rightarrow$ Class B

Linear Discriminant Analysis

The discriminant function divides space in two areas

$$\Psi_1 = x_{1i}\, u_{11} + \dots + x_{pi}\, u_{1p}$$

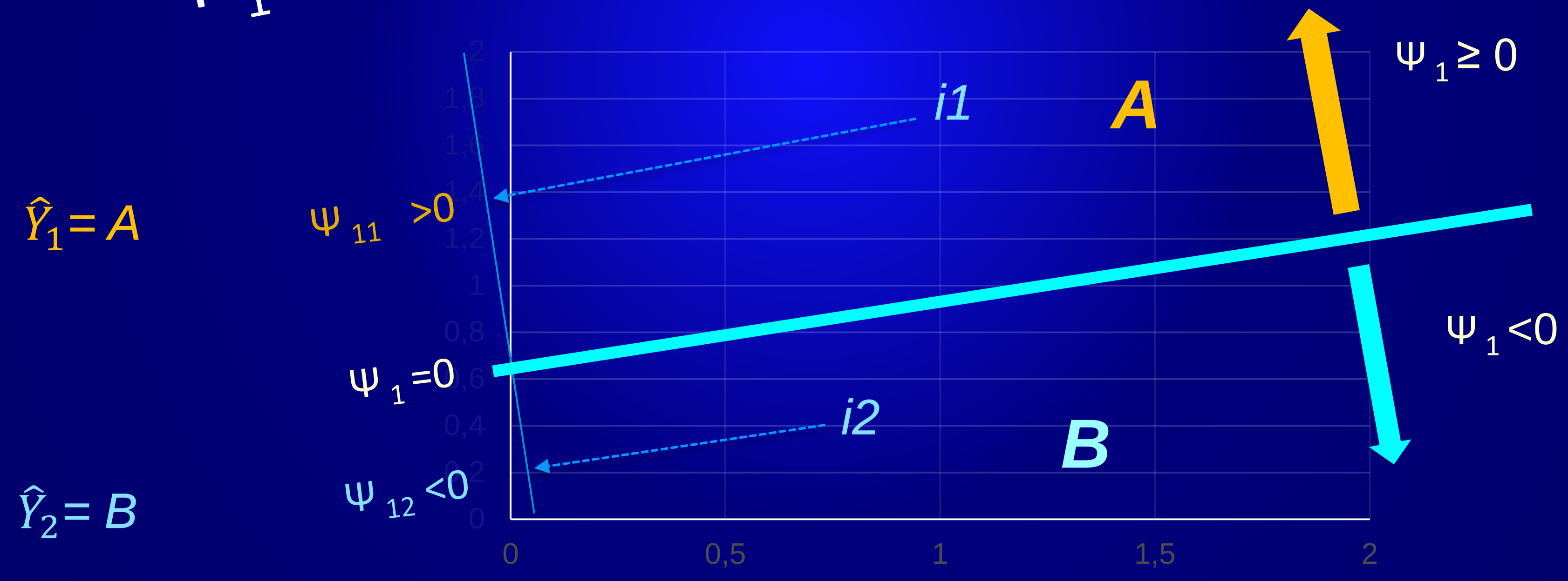


Linear Discriminant Analysis

USE OF MODEL: Class of new objects is predicted easily

- If $\Psi_{1\,i} \geq 0 \Rightarrow$ Class A
- If $\Psi_{1\,i} < 0 \Rightarrow$ Class B

$$\Psi_1 = x_{1i}\, u_{11} + \dots + x_{pi}\, u_{1p}$$



Linear Discriminant Analysis

Process to interpret the linear discriminant function

- Project individuals in first factorial map
- Project numerical explanatory variables
- Eventually Project illustrative numerical variables
- Project centroids of qualitative variables (only illustrative)
- Forget about variables/centroids bad represented in the factorial plan
- Which are the variables/centroid with relevant direct contribution to Ψ_1?
- Which are the variables with relevant inverse contribution to Ψ_1=
- Analyze profiles opposed in two extremes of Axis X
- Induce a label for Ψ_1
- See which class corresponds to $\Psi_1>0$ or $\Psi_1<0$ and describe

Linear Discriminant Analysis

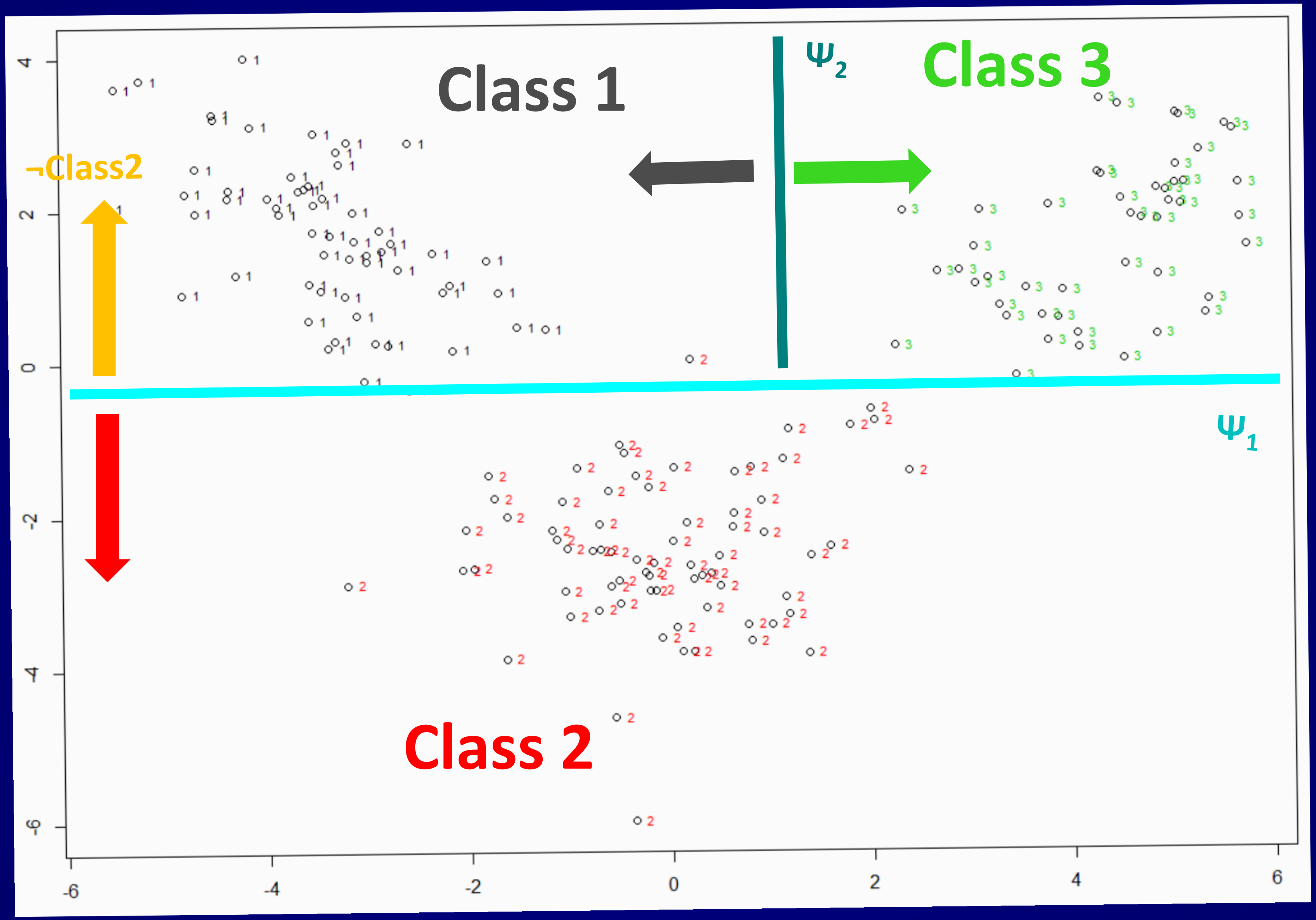
For response variables with n_Y modalities

- Check if all the clases are separable with Ψ_1
  - Find corresponding Cut-off
- If not, second principal component (Ψ_2) provides the second discriminant function.

Interpret the map and find the discrimination rule, i.e.

  - If phi1>0 then A
  - If phi1<=0 then
    - If phi2>0 then B
    - If phi2<=0 then C

Linear Discriminant Analysis



Linear Discriminant Analysis

Further uses

Use projection space to
project Variables and individuals
to see clusters or individuals,
to interpret discriminant functions,
to find relationship between variables,
to see main explanatory variables for the clases,
etc

Linear Discriminant Analysis

K. Gibert[1,2,3], A. Palomino[1,2,4]

[1] Department of Statistics and Operation Research

[2] Knowledge Engineering and Machine Learning group at
Intelligent Data Science and Artificial Intelligence Specific Research Center

[3] Institut Universitari de Recerca en Ciència I Tecnologia de la Sostenibilitat

[4] Kantar

karina.gibert@upc.edu, https://www.eio.upc.edu/en/homepages/karina

Are there any questions?...