MATH 569 Statistical Learning

Part III: Linear Methods of Classification

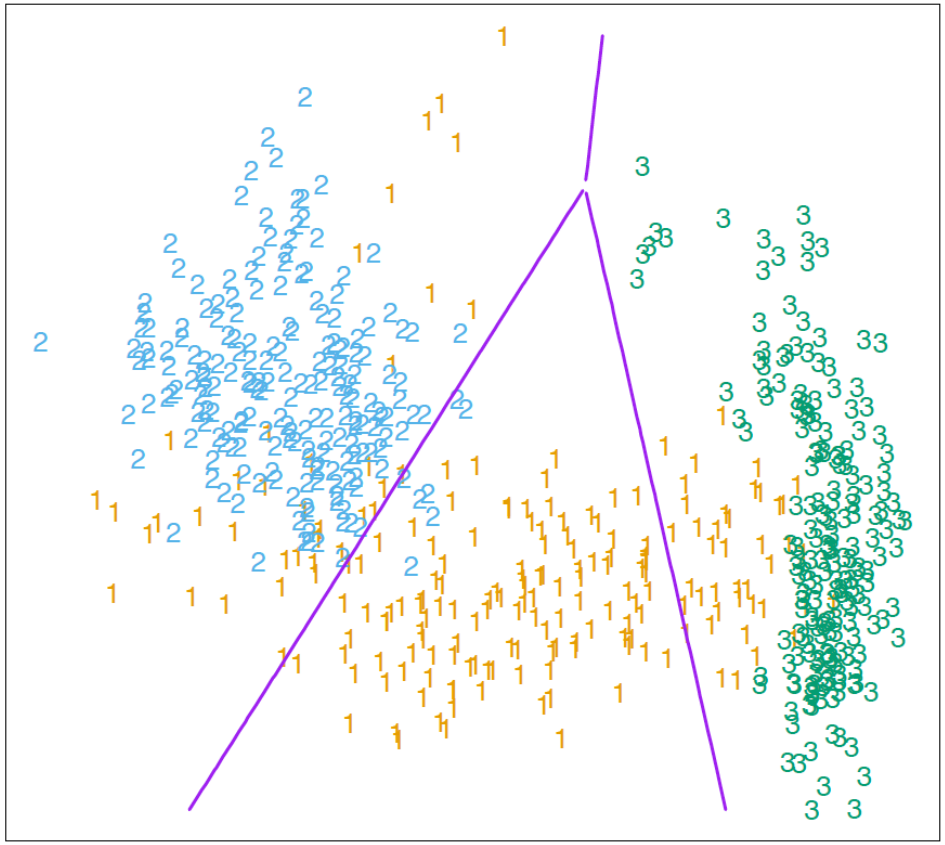
Maggie Cheng

ILLINOIS INSTITUTE OF TECHNOLOGY

Fig 4.1 Linear boundary vs quadratic boundary (found by LDA)

Linear boundary

Using only X_1 and X_2



Quadratic boundary

Using $X_1, X_2, X_1X_2, X_1^2, X_2^2$

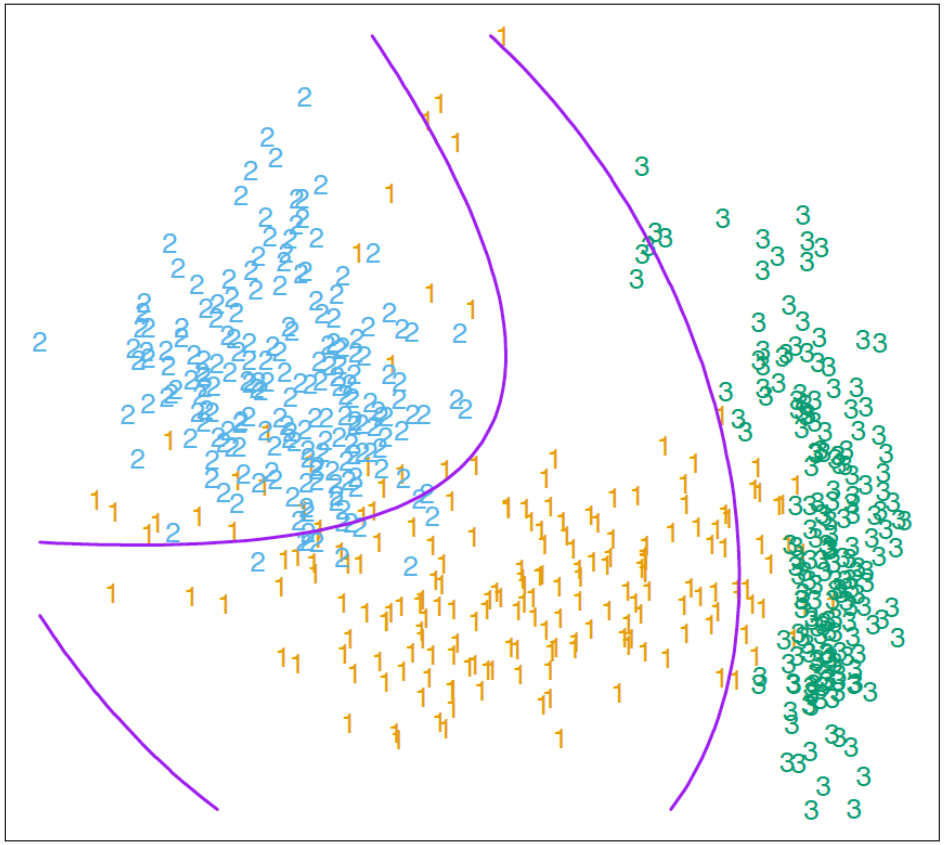


Fig 4.2 Masking effect for linear regression
Three classes in a 2-dimensional space (K=3, p=2)

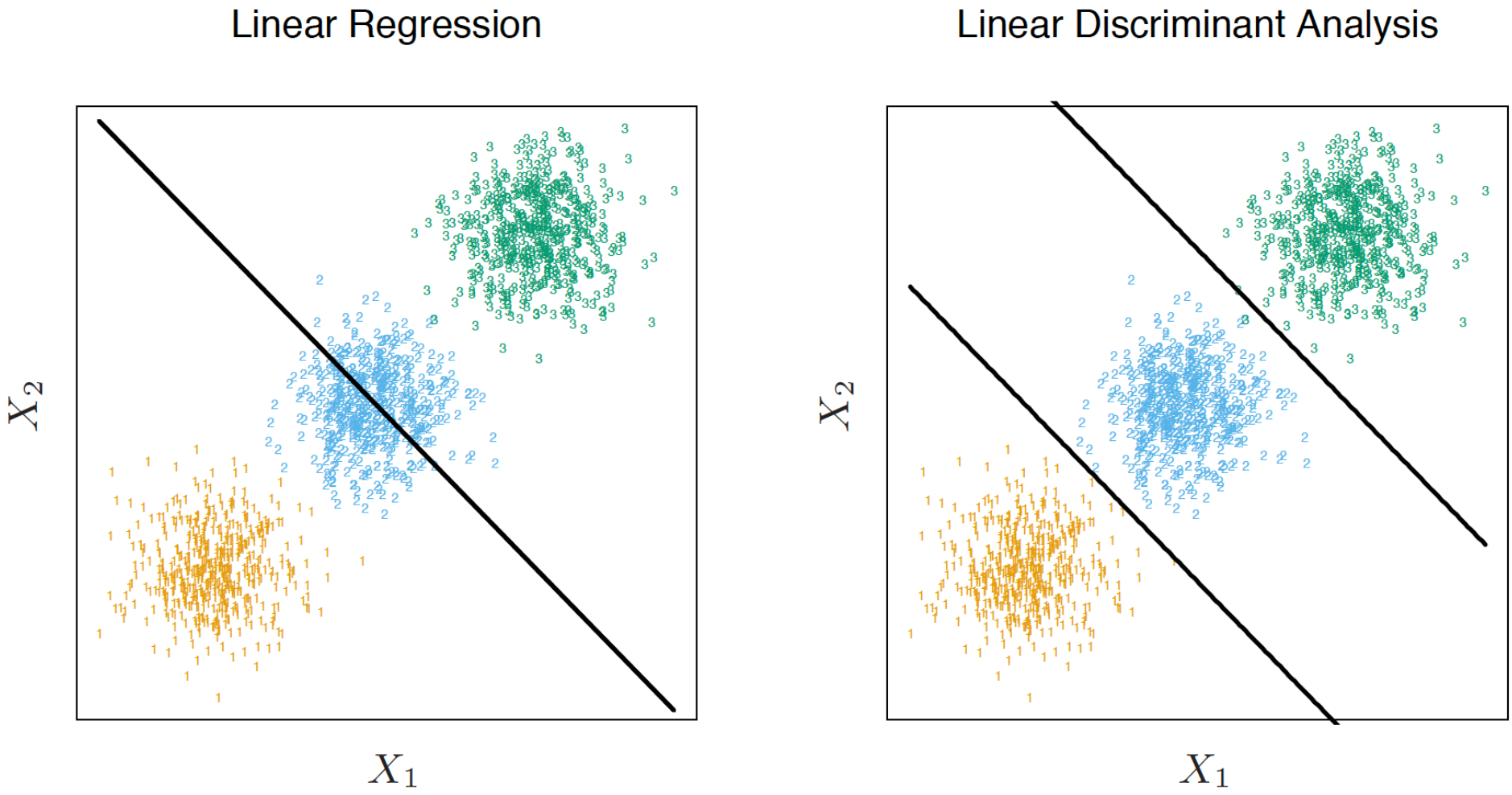
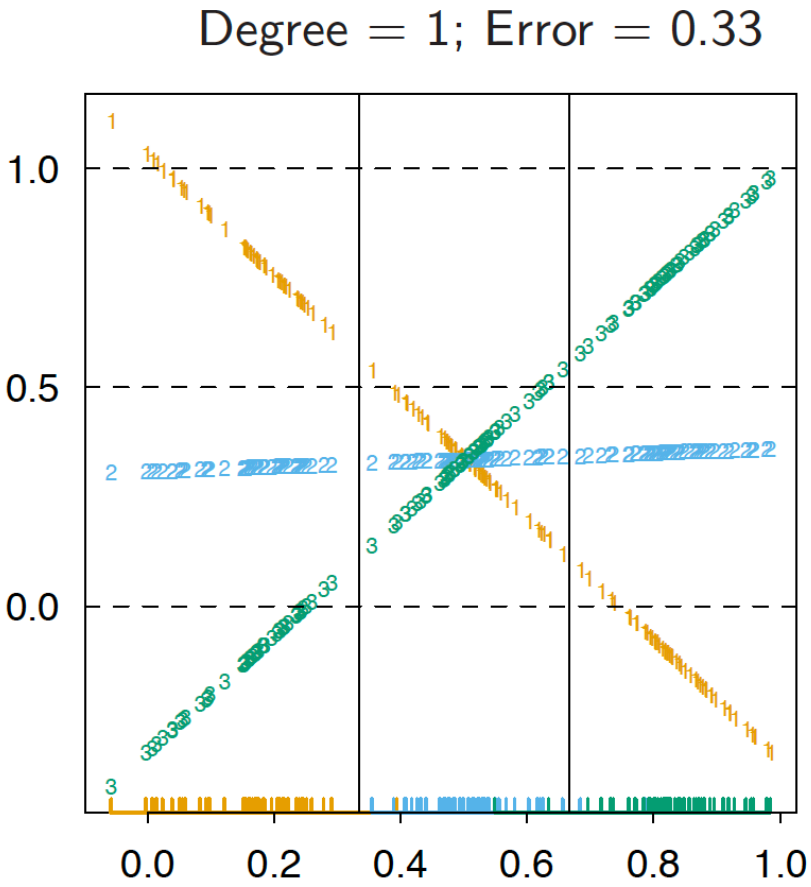
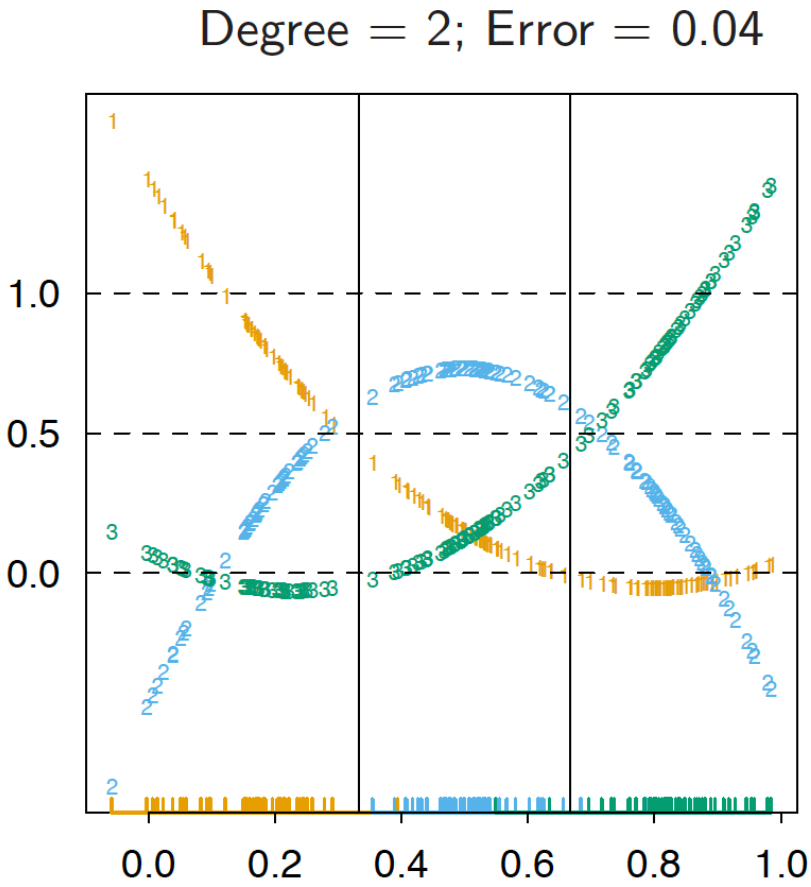


Fig 4.3 Use higher degree polynomials to fix the masking effect of linear regression

Using linear regression on linear terms,
the middle class never dominates.
Class 2 is masked by class 1 and class 3.

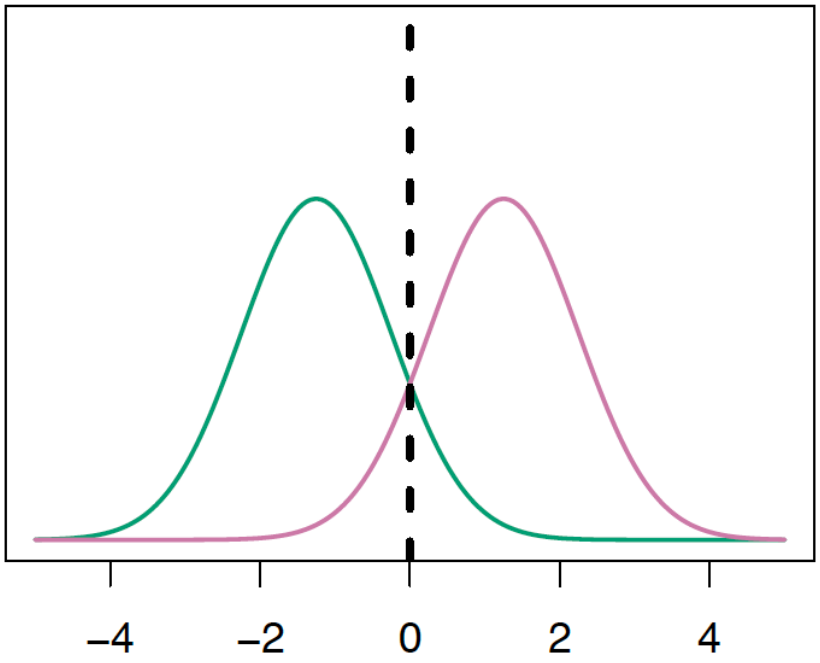


Using quadratic terms,
class 2 is no longer masked.



Bayes Decision Boundary

$\pi_1 = 0.5,\ \pi_2 = 0.5$



$\pi_1 = 0.3,\ \pi_2 = 0.7$

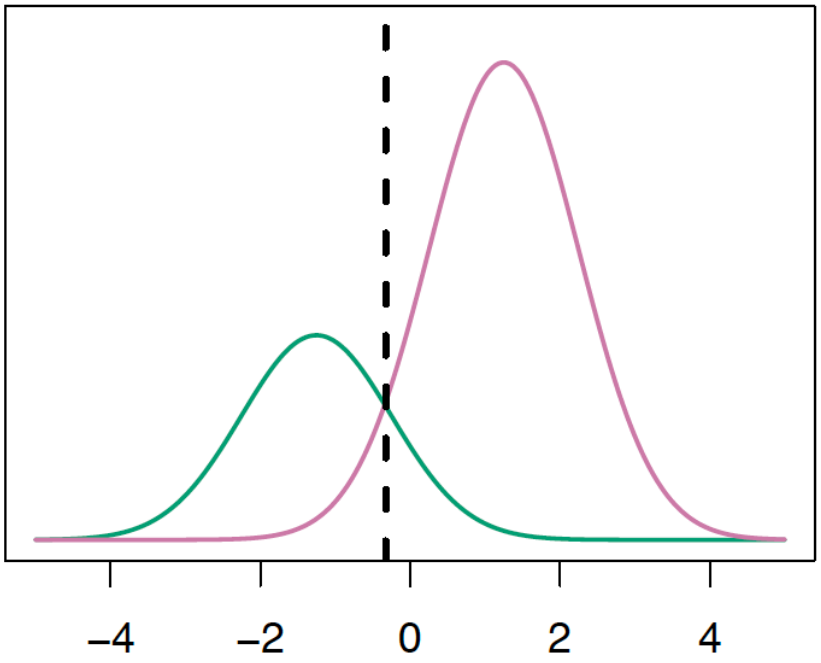


Fig 4.6 LDA and QDA both can find quadratic boundaries

Using LDA on 5 dimensional space

$X_1, X_2, X_1X_2, X_1^2, X_2^2$

Using QDA

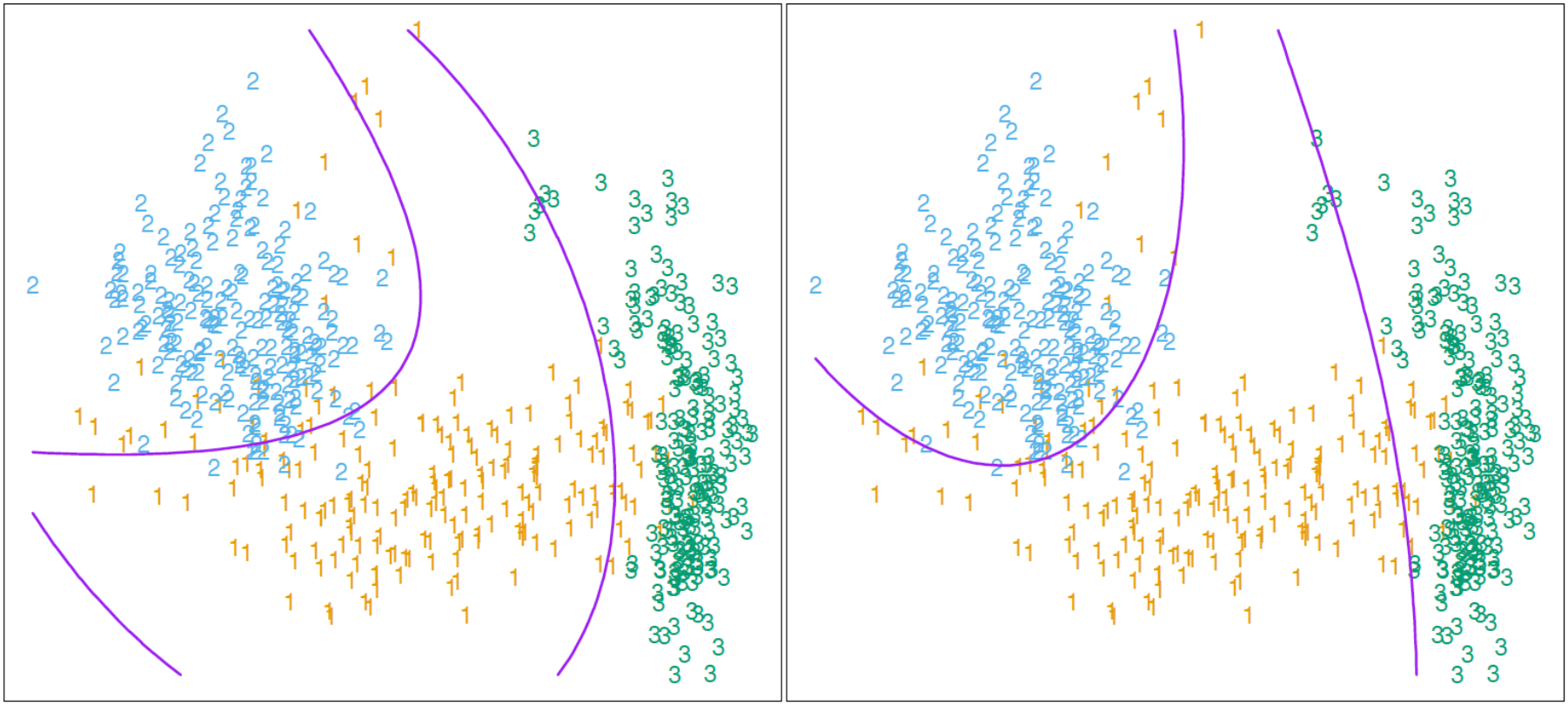
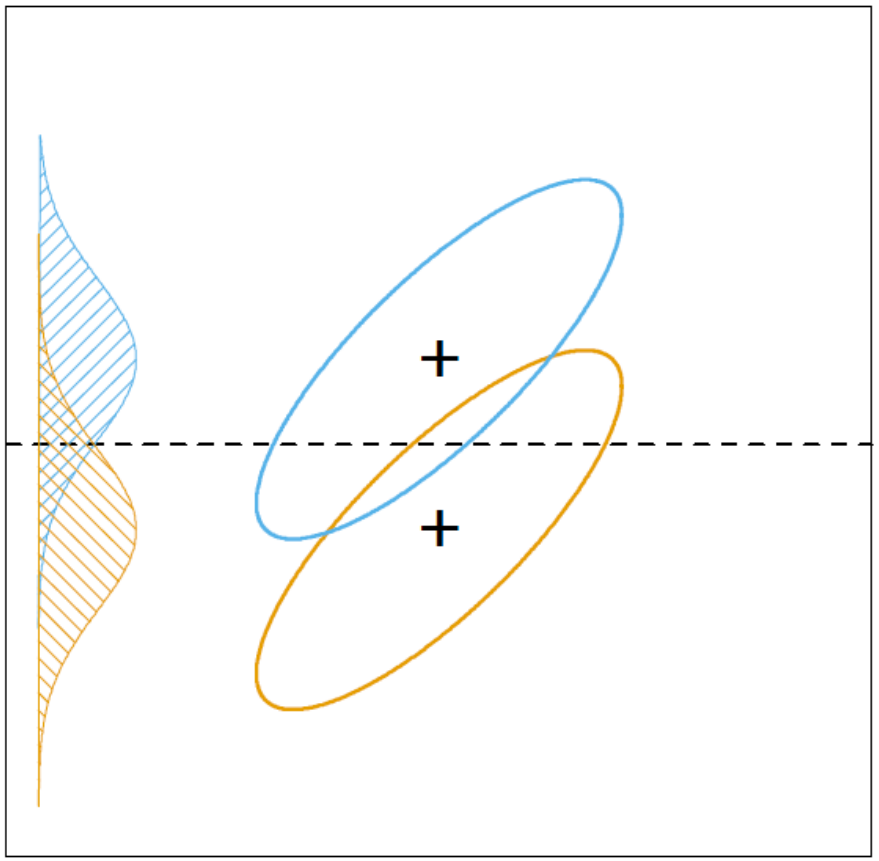


Fig 4.9 Discriminant Direction

The direction connecting two centroids may not minimize overlap



The discriminant direction minimizes this overlap for Gaussian data

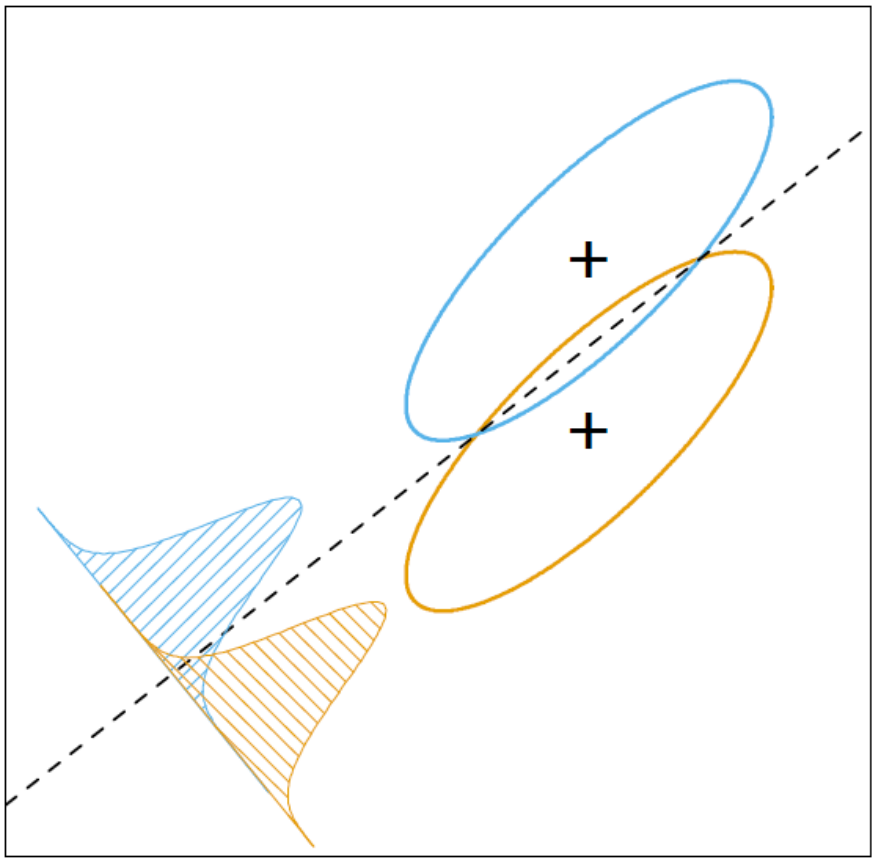


Fig 4.11 Classification in the two-dimensional subspace spanned by the first two canonical variates (Vowel training data: K=11, p=10)