 POLITECNICO DI MILANO

Dipartimento di
Elettronica e Informazione

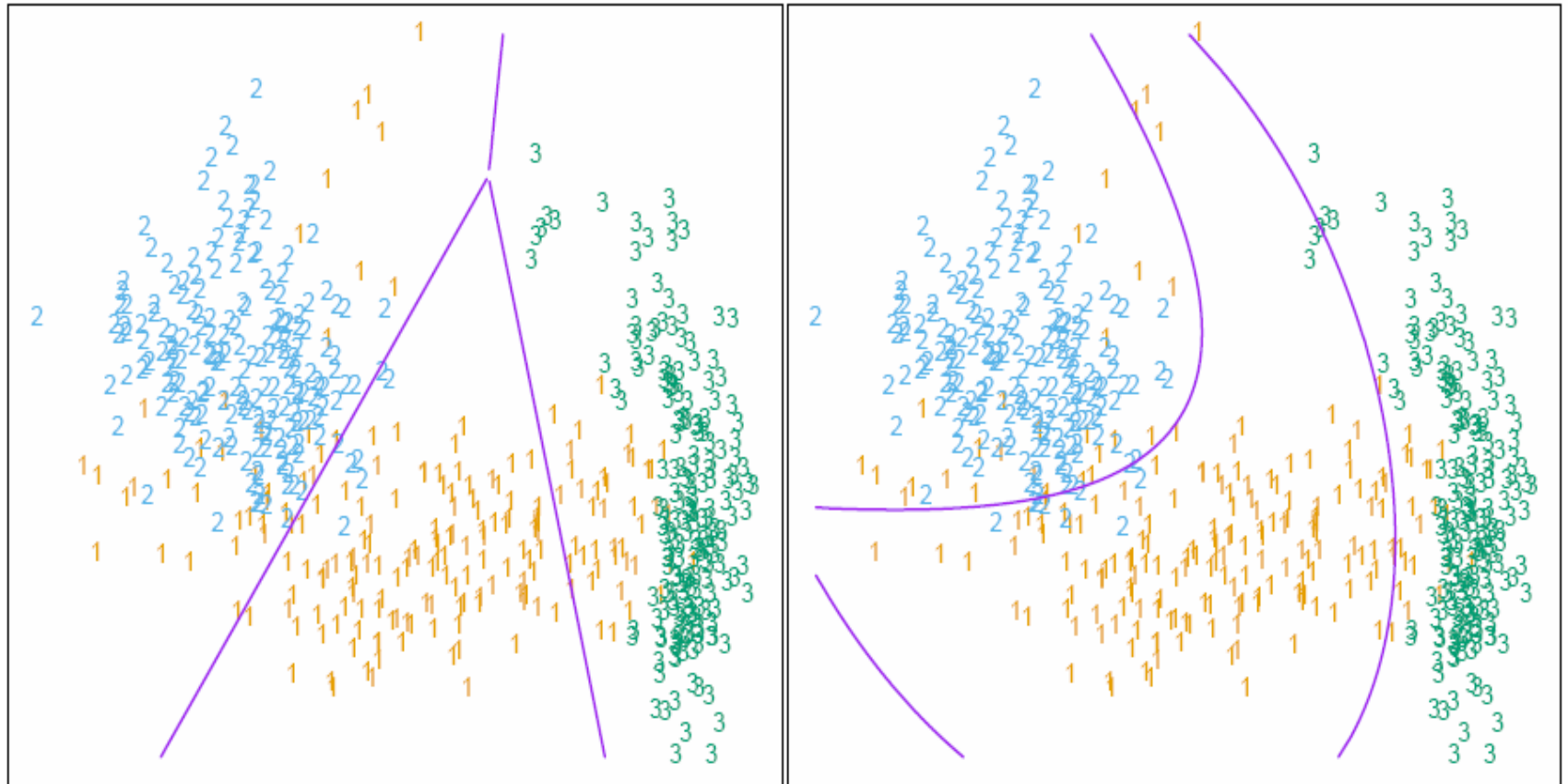
Pattern Analysis and Machine Intelligence

M. Matteucci, Luigi Malagò, Davide Eynard

matteucci@elet.polimi.it

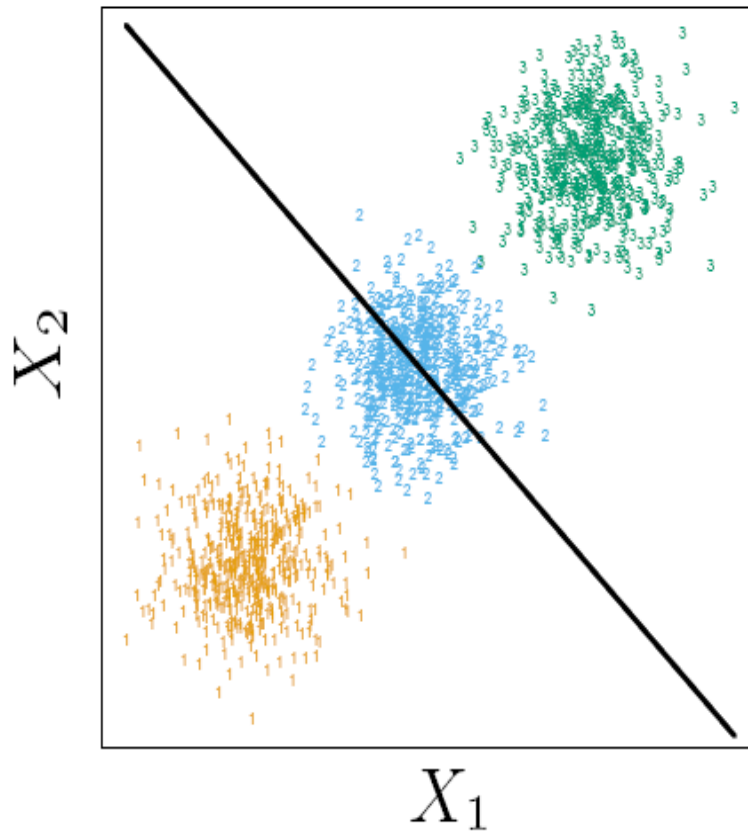
*Dipartimento di Elettronica e Informazione, Politecnico di Milano
Artificial Intelligence and Robotics Lab*

Linear Regression on the Indicator Matrix

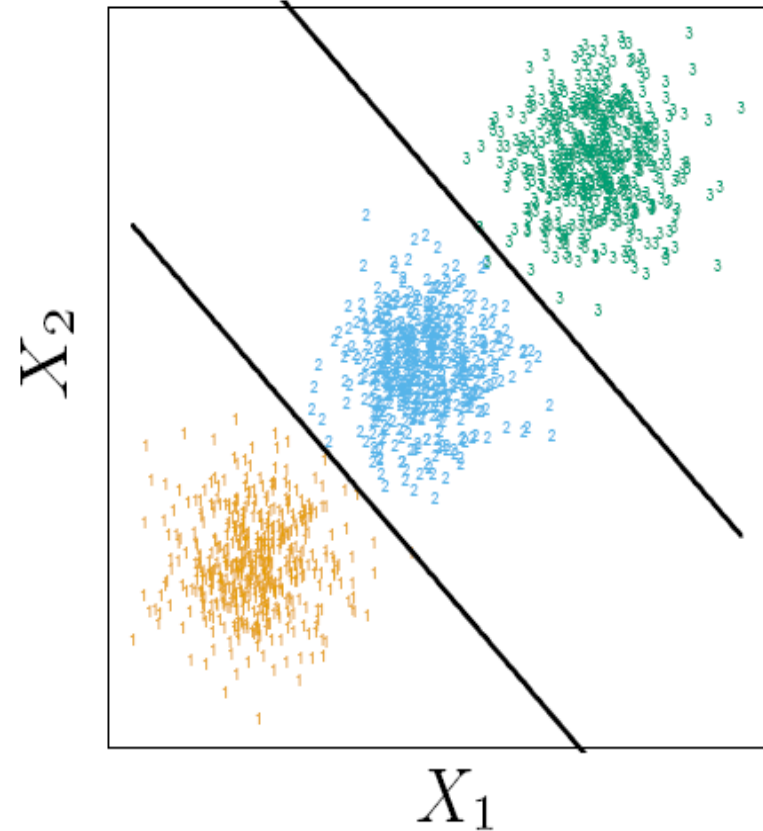


Masking Effect of Linear Regression

Linear Regression

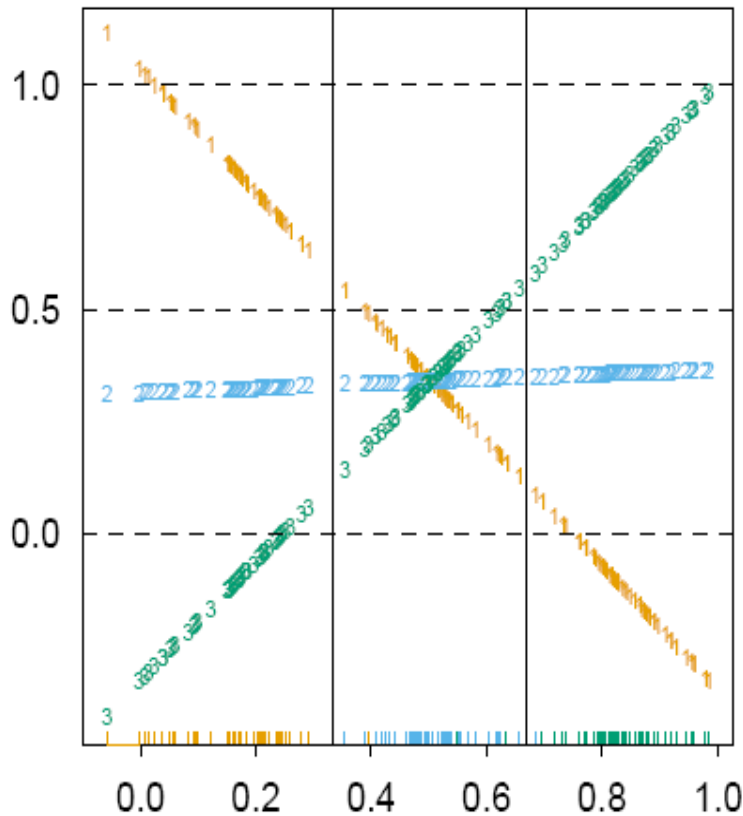


Linear Discriminant Analysis

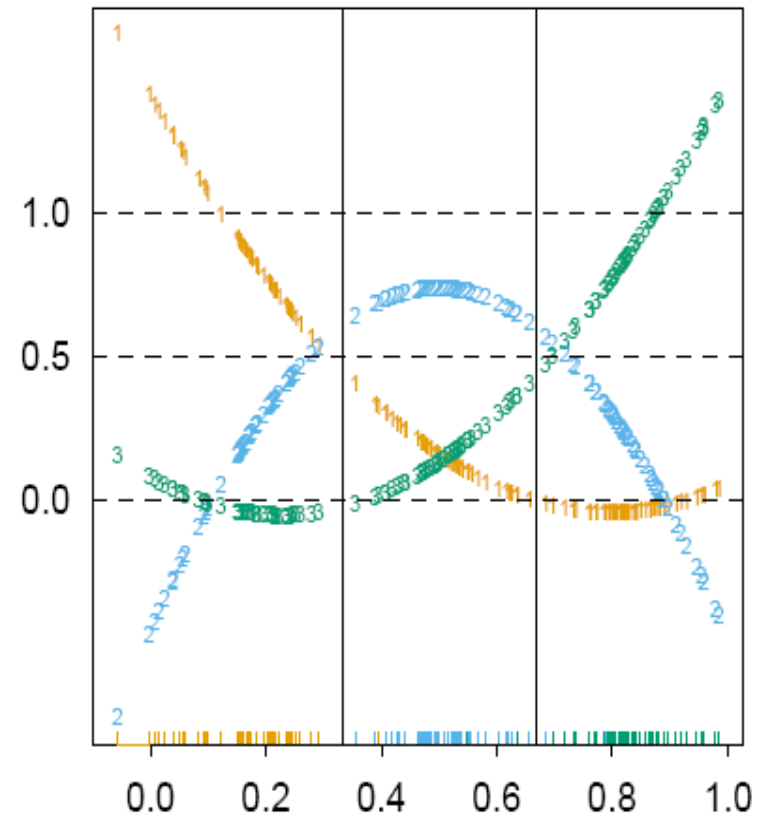


Masking Effect of Linear Regression (II)

Degree = 1; Error = 0.33



Degree = 2; Error = 0.04



The Vowel Data

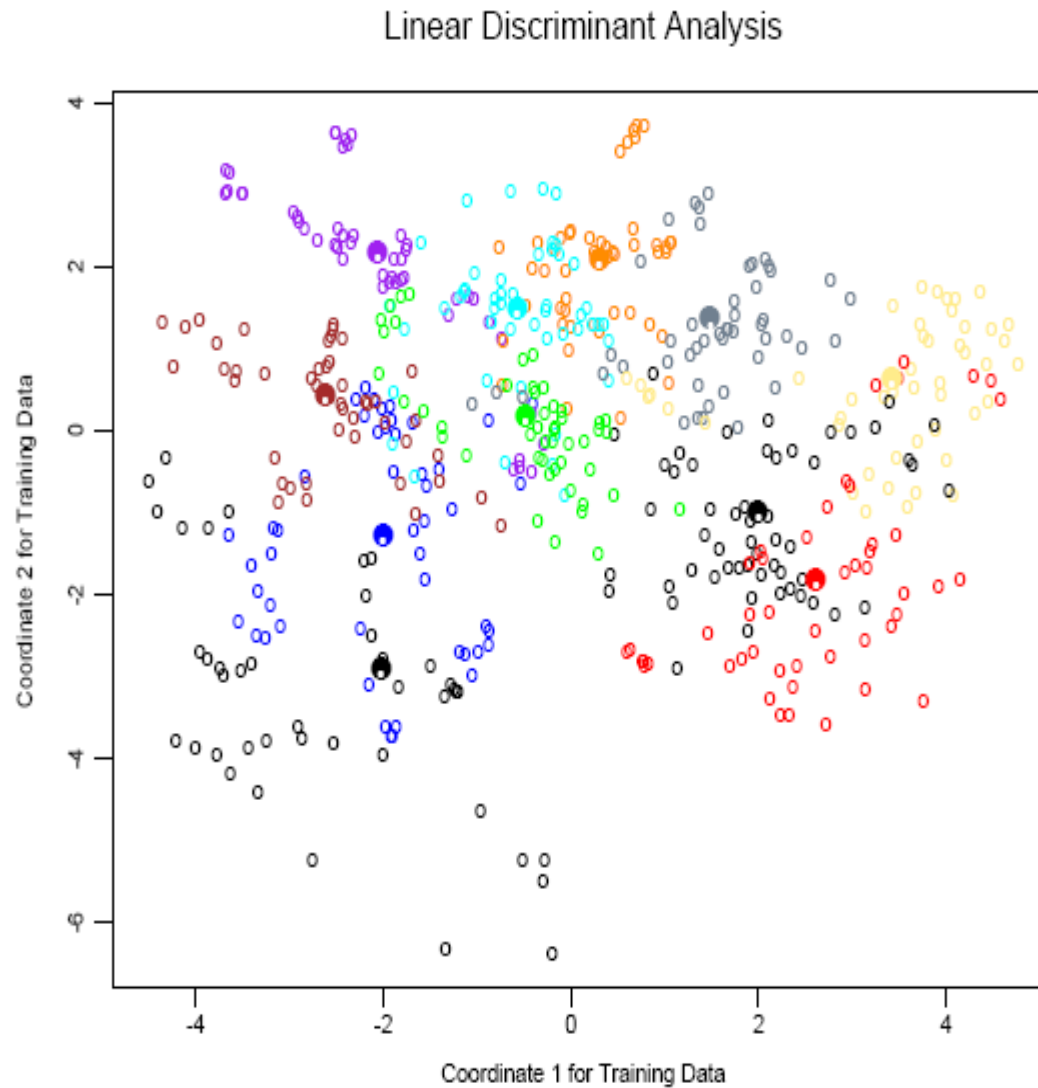
Eleven vowel sounds were recorded. The word was uttered once by each of the fifteen speakers.

Four male and four female speakers were used to train the networks, and the other four male and three female speakers were used for testing the performance.

I hit	i: tree	e pen	æ flat
ʌ cup	a: laugh	ə mother	ɜ: bird
ʊ look	u: you	ɒ not	ɔ: door

The speech signals were low pass filtered at 4.7kHz and then digitised to 12 bits with a 10kHz sampling rate. Twelfth order linear predictive analysis was carried out on six 512 sample Hamming windowed segments from the steady part of the vowel. The reflection coefficients were used to calculate 10 log area parameters, giving a 10 dimensional input space. For a general introduction to speech processing and an explanation of this technique see Rabiner and Schafer [RabinerSchafer78].

LDA Decision Boundaries on the Vowel Problem

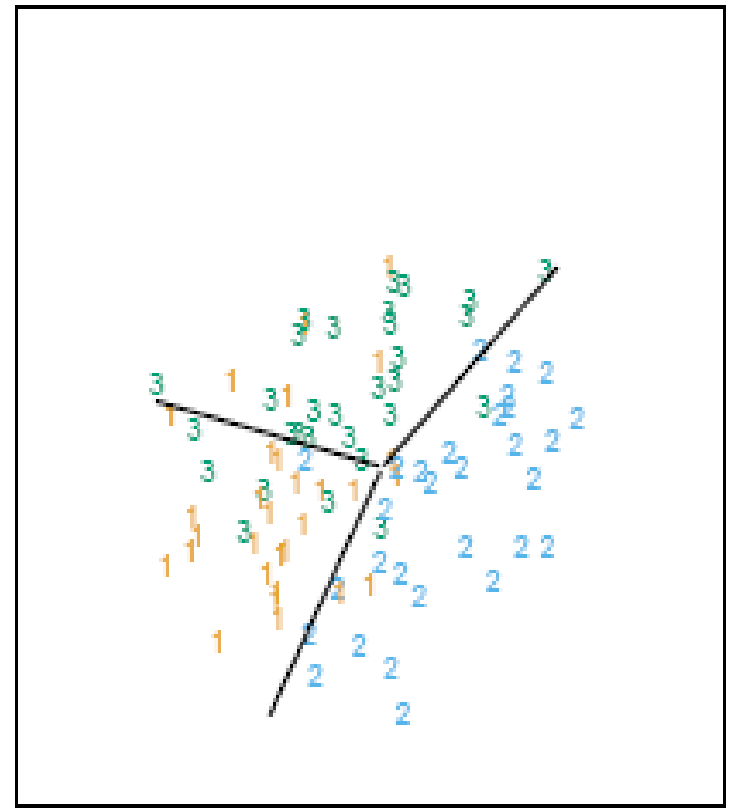
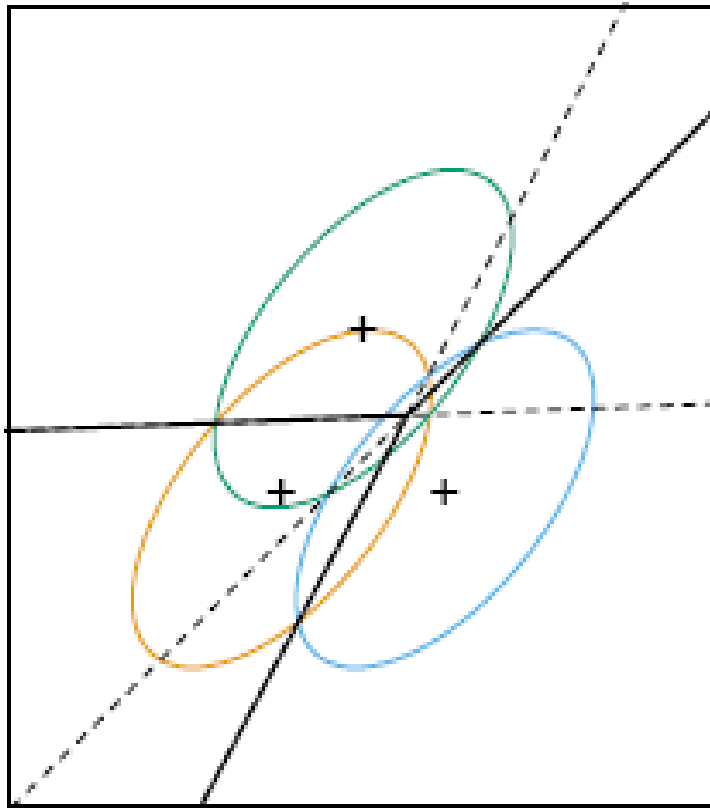


Linear Techniques Compared on the vowel data

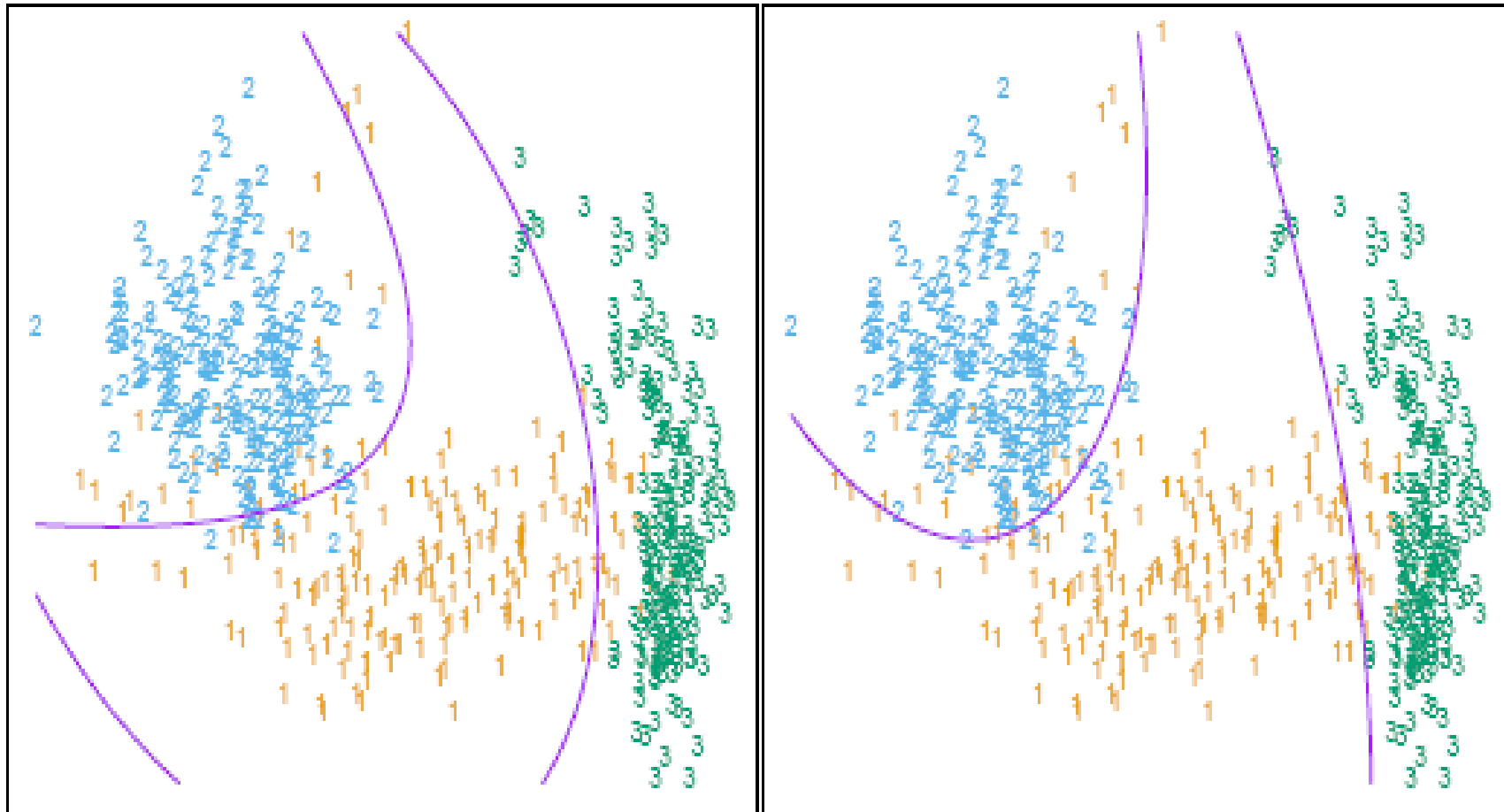
TABLE 4.1. *Training and test error rates using a variety of linear techniques on the vowel data. There are eleven classes in ten dimensions, of which three account for 90% of the variance (via a principal components analysis). We see that linear regression is hurt by masking, increasing the test and training error by over 10%.*

Technique	Error Rates	
	Training	Test
Linear regression	0.48	0.67
Linear discriminant analysis	0.32	0.56
Quadratic discriminant analysis	0.01	0.53
Logistic regression	0.22	0.51

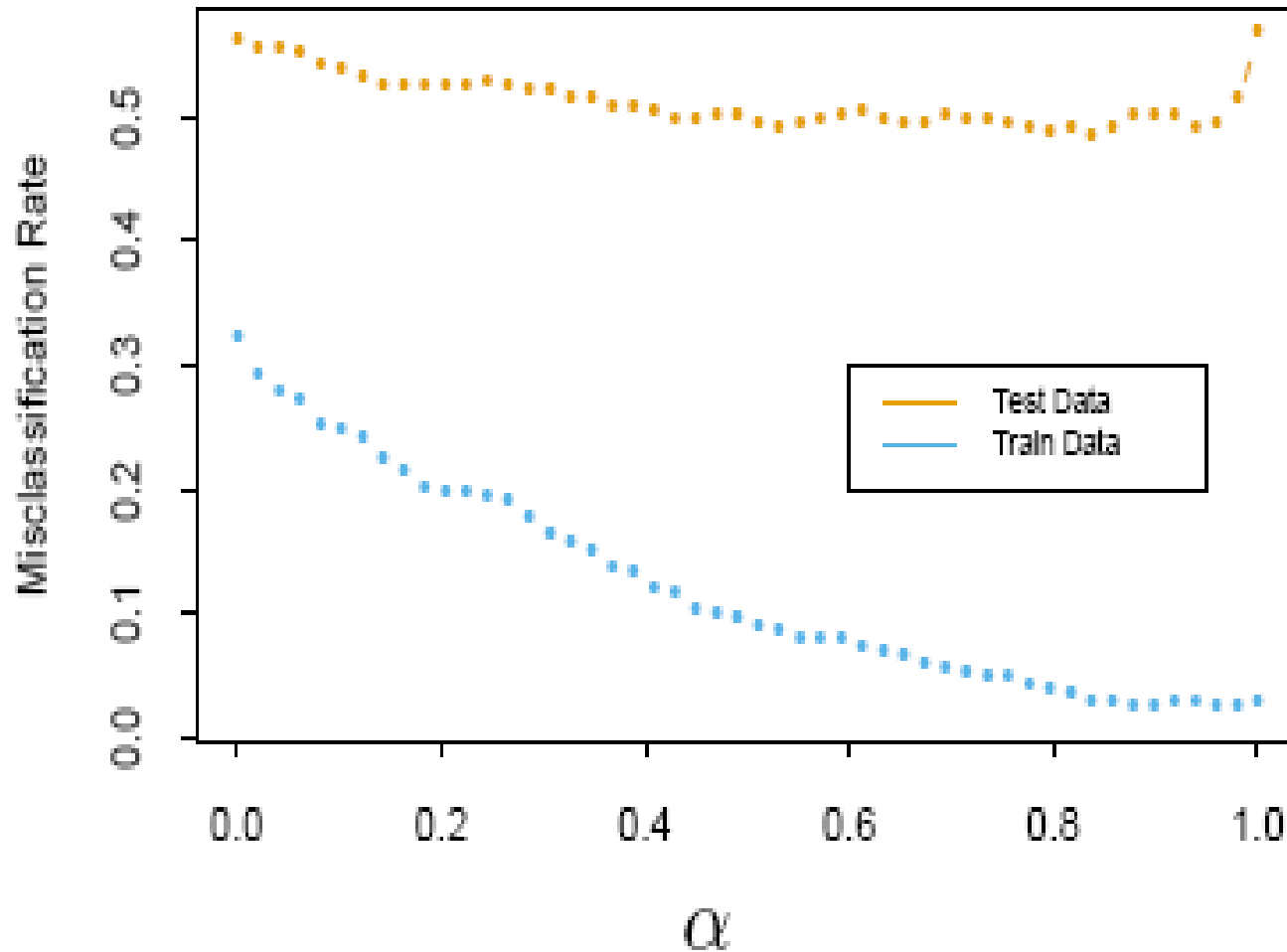
LDA Decision Boundary



Two Ways for QDA

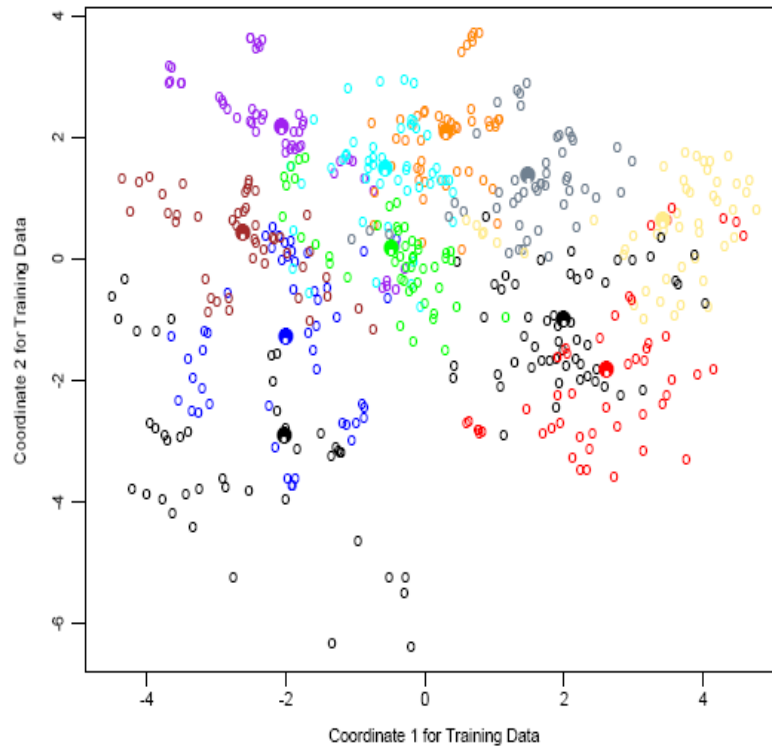


Regularized Discriminant Analysis

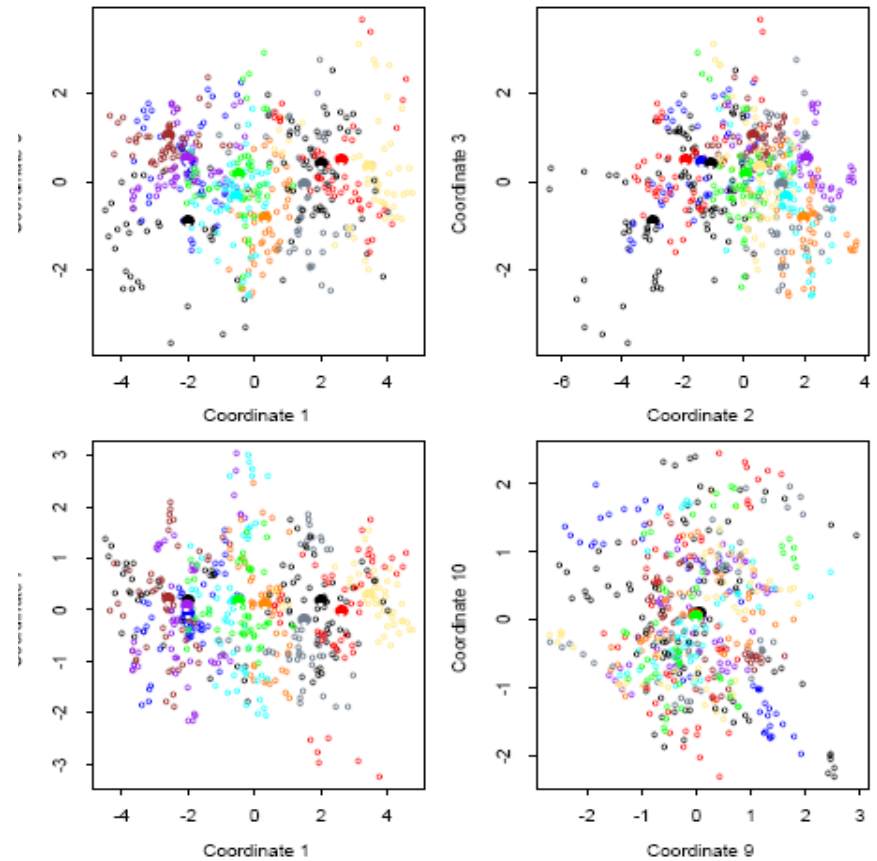


LDA Projections for the Vowel Problem

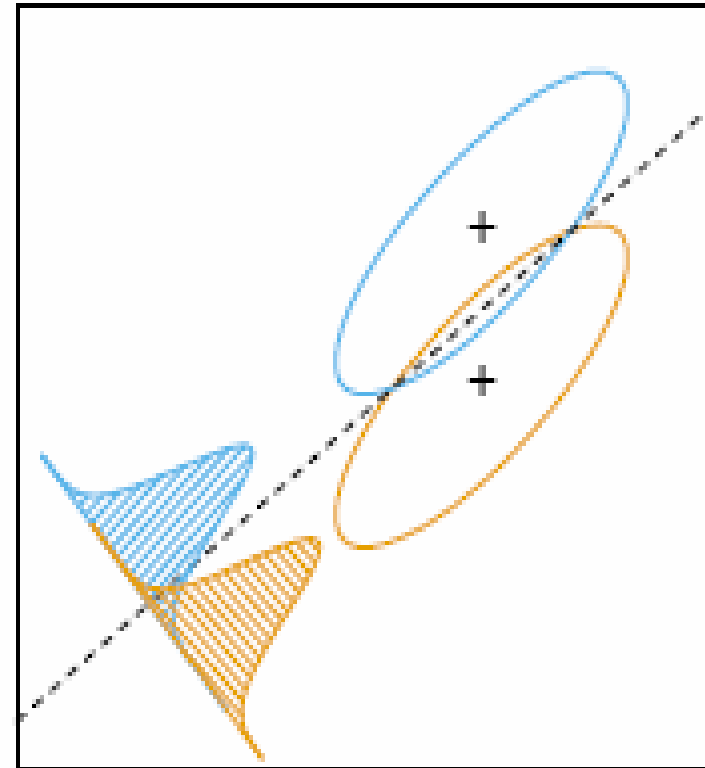
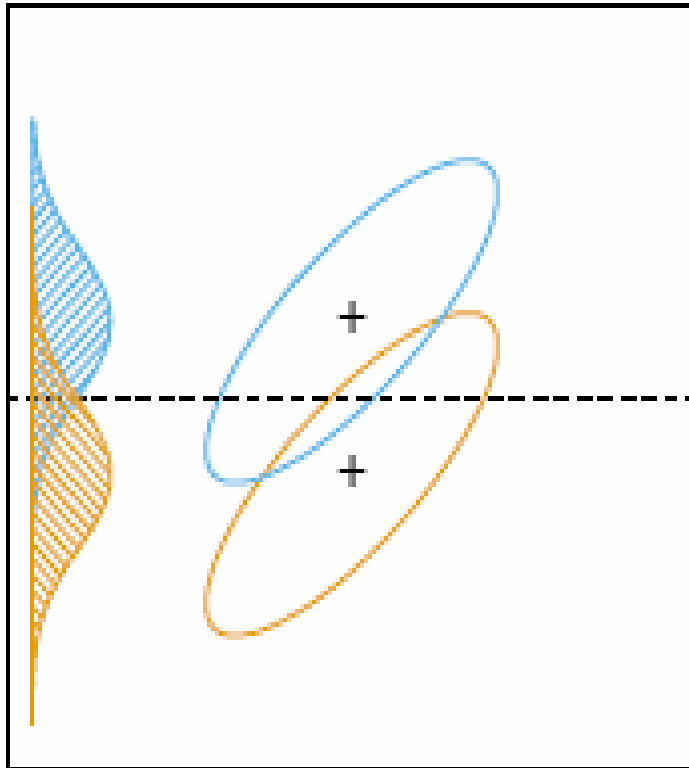
Linear Discriminant Analysis



Linear Discriminant Analysis

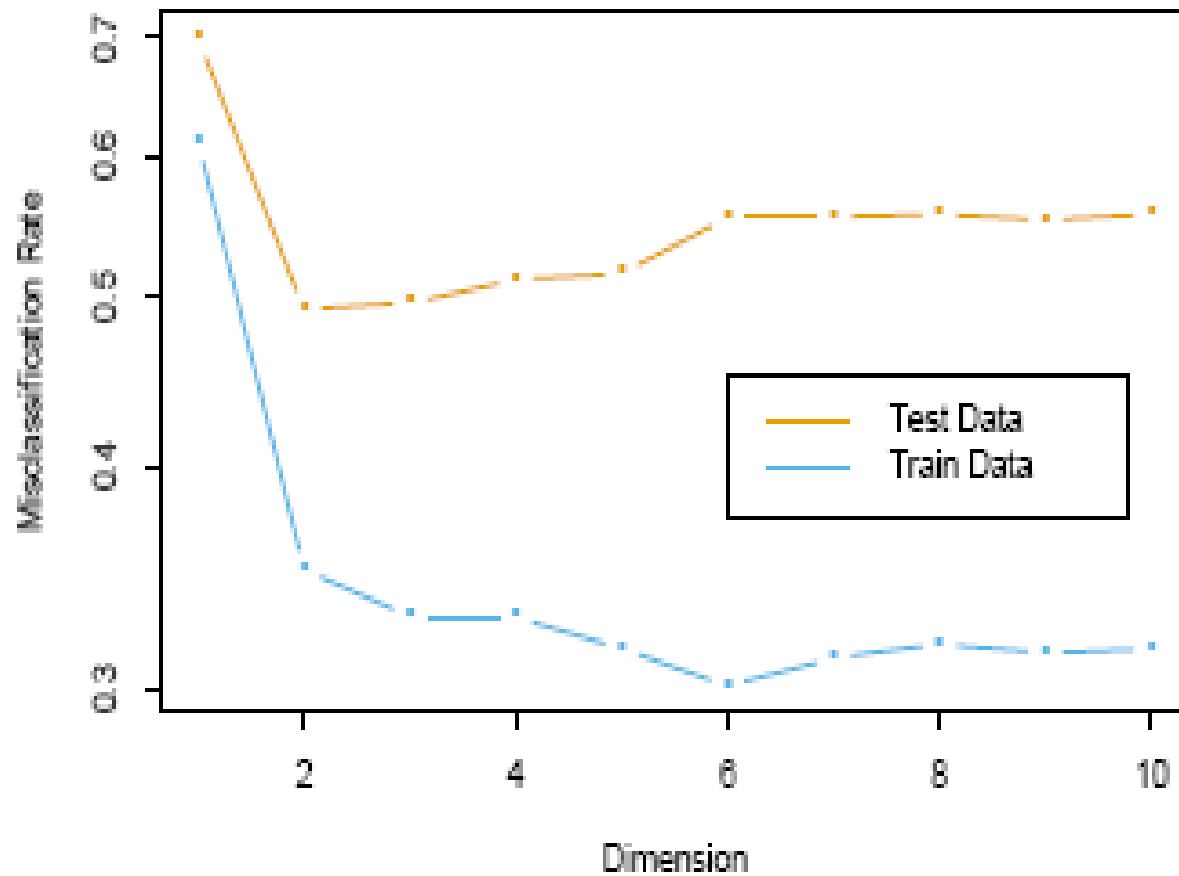


LDA Projections

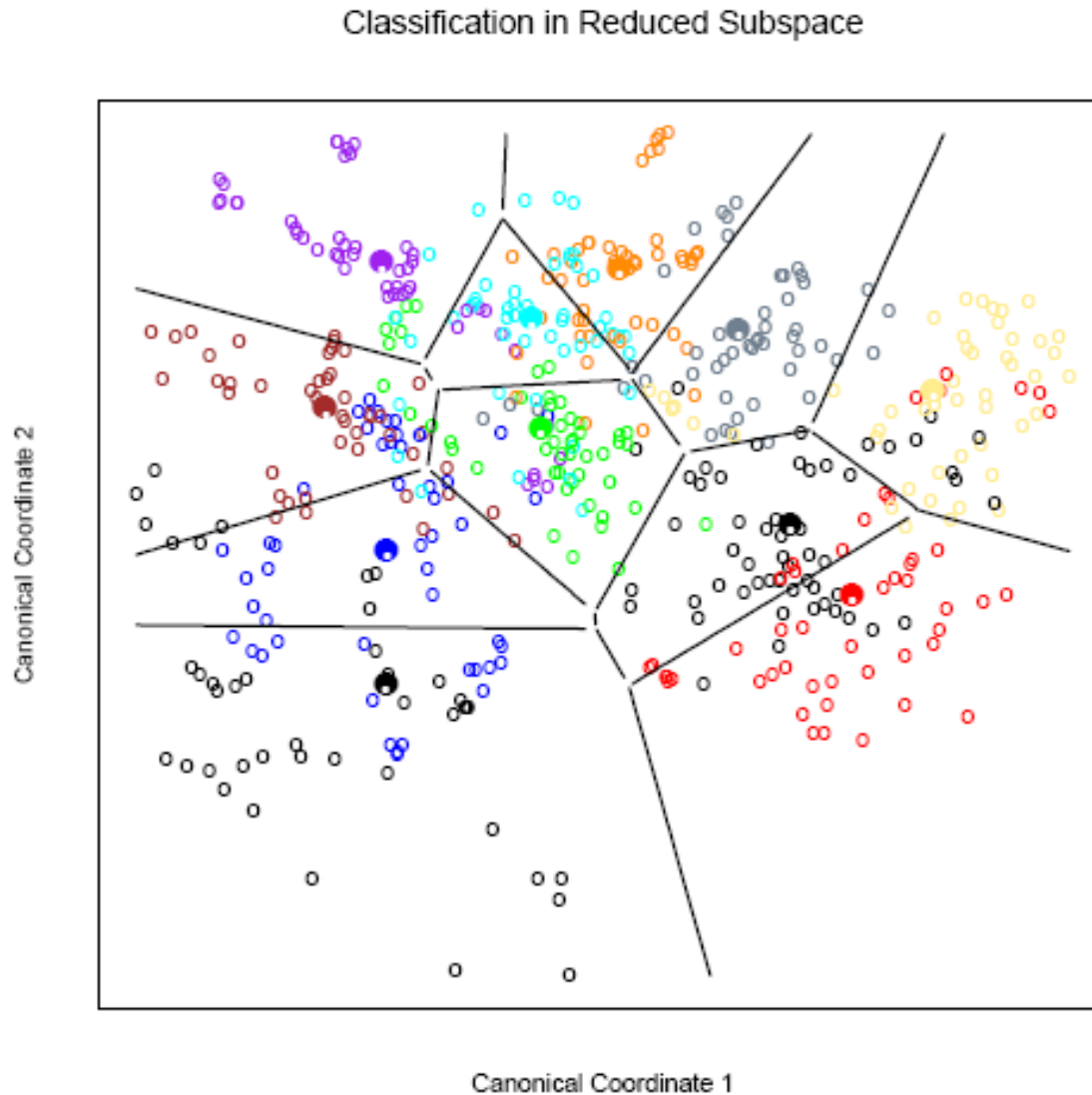


LDA in Reduced Spaces

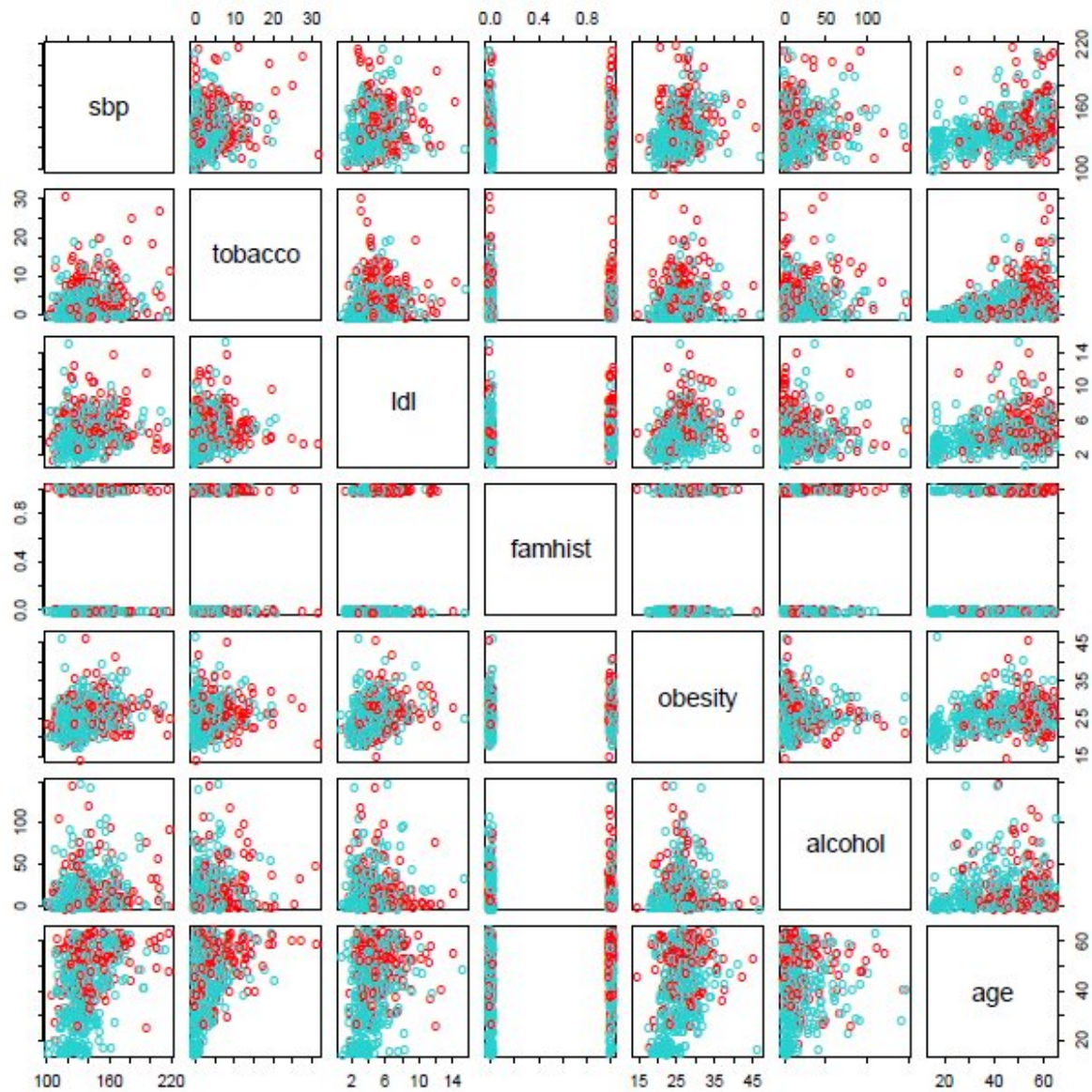
LDA and Dimension Reduction on the Vowel Data



LDA Decision Boundary for the Vowel data



South African Heart Disease



Full Logistic Regression

TABLE 4.2. *Results from a logistic regression fit to the South African heart disease data.*

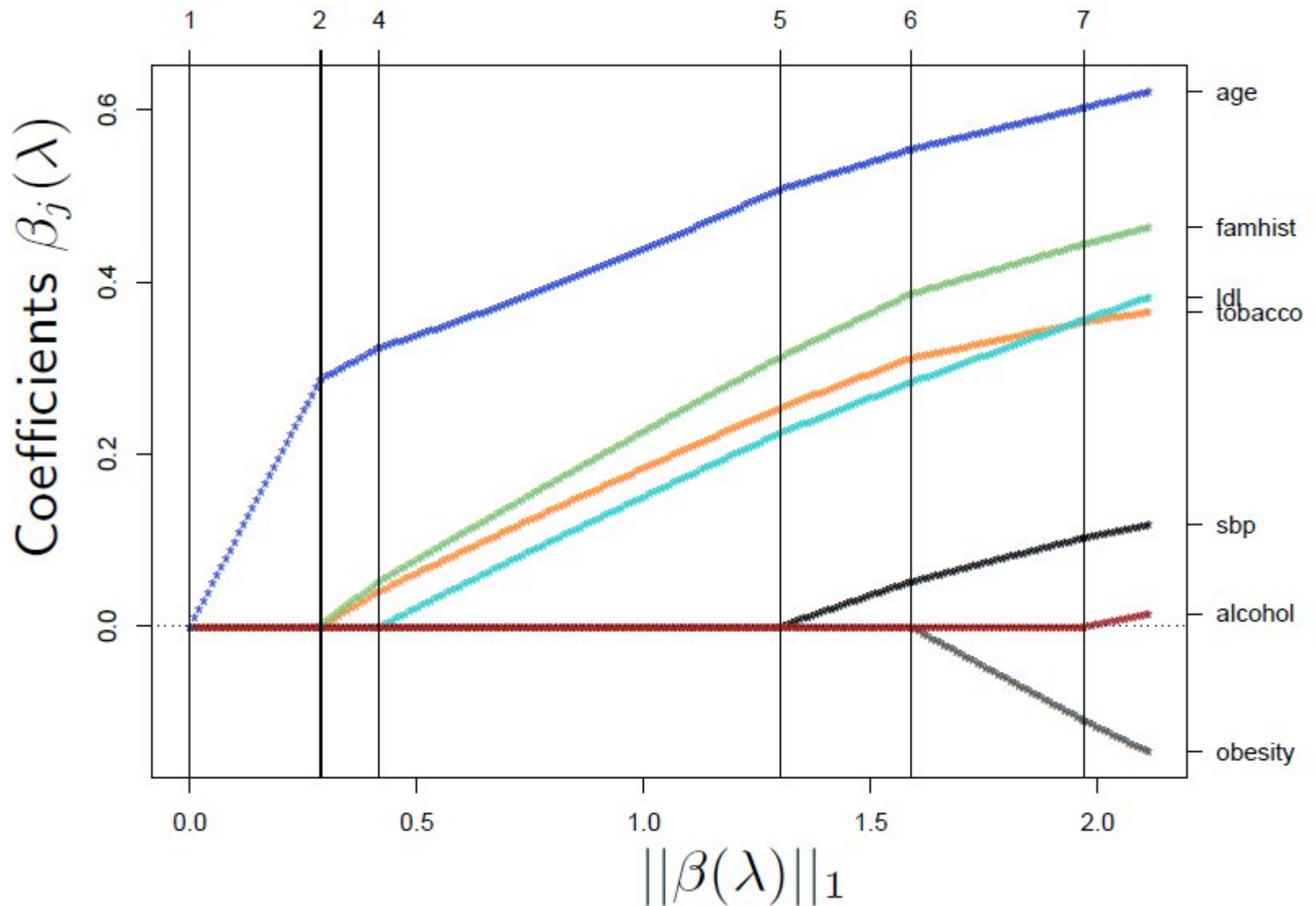
	Coefficient	Std. Error	Z Score
(Intercept)	-4.130	0.964	-4.285
sbp	0.006	0.006	1.023
tobacco	0.080	0.026	3.034
ldl	0.185	0.057	3.219
famhist	0.939	0.225	4.178
obesity	-0.035	0.029	-1.187
alcohol	0.001	0.004	0.136
age	0.043	0.010	4.184

Stepwise Logistic Regression (SBS)

TABLE 4.3. *Results from stepwise logistic regression fit to South African heart disease data.*

	Coefficient	Std. Error	Z score
(Intercept)	-4.204	0.498	-8.45
tobacco	0.081	0.026	3.16
ldl	0.168	0.054	3.09
famhist	0.924	0.223	4.14
age	0.044	0.010	4.52

Regularized Logistic Regression

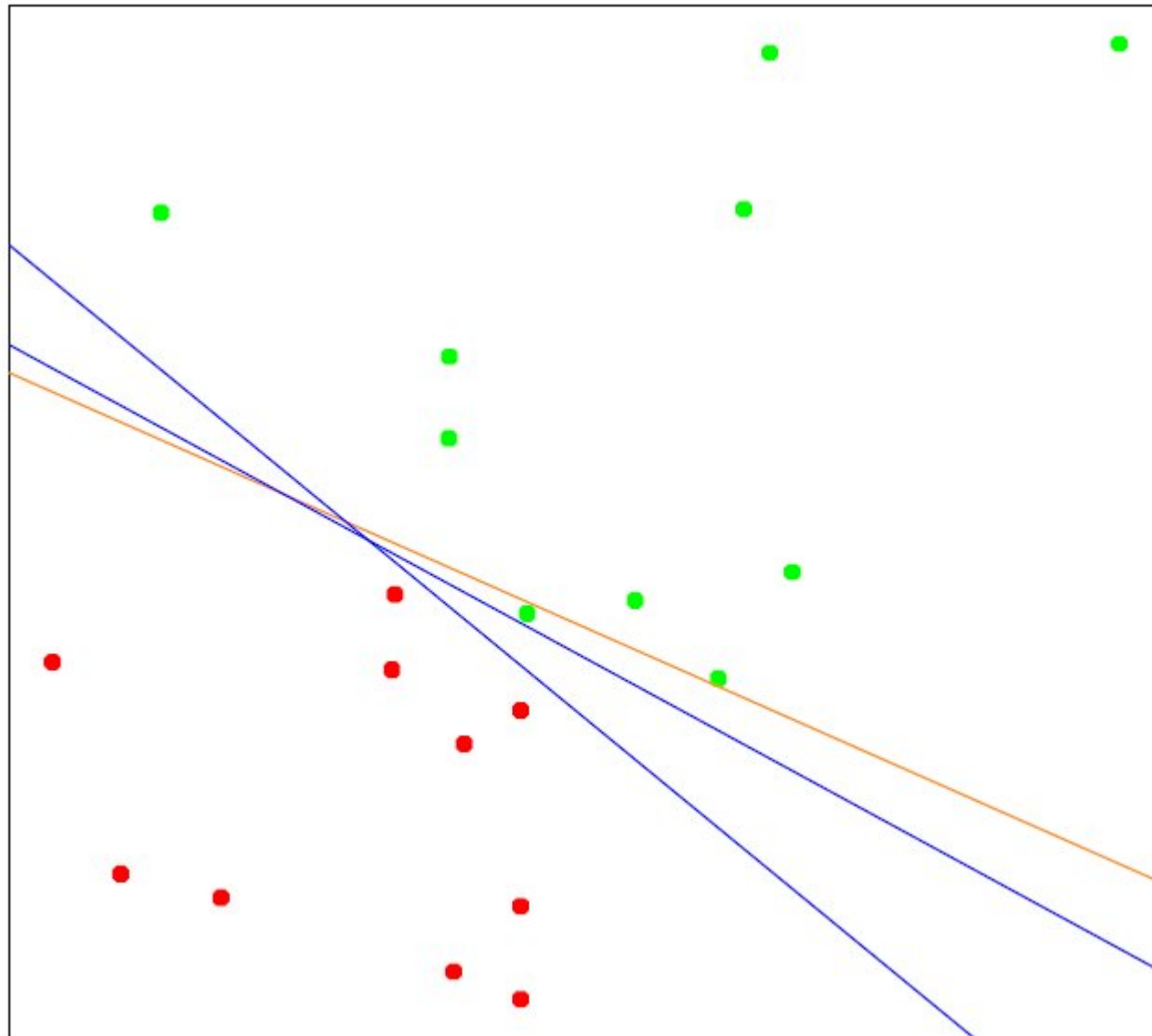


Linear Techniques Compared on the vowel data

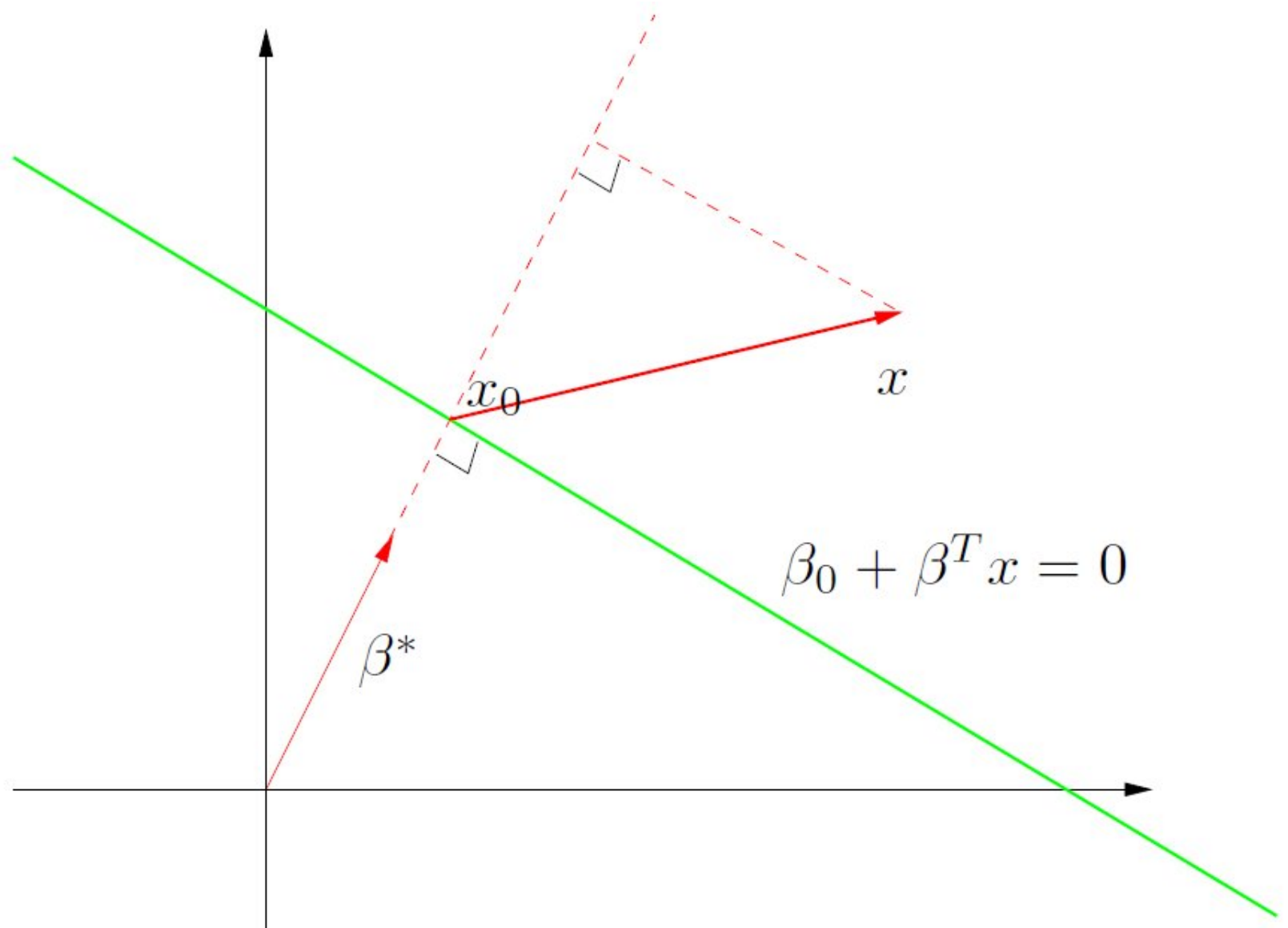
TABLE 4.1. *Training and test error rates using a variety of linear techniques on the vowel data. There are eleven classes in ten dimensions, of which three account for 90% of the variance (via a principal components analysis). We see that linear regression is hurt by masking, increasing the test and training error by over 10%.*

Technique	Error Rates	
	Training	Test
Linear regression	0.48	0.67
Linear discriminant analysis	0.32	0.56
Quadratic discriminant analysis	0.01	0.53
Logistic regression	0.22	0.51

Separating Hyperplanes



Hyperplane Linear Algebra



Maximum Margine Hyperplane

