

Good value of  $k = \sqrt{\text{samples}}$

The variance of the MLE method is higher than MAP

XOR example: Polynomial kernel of degree 2

Perceptron advantages and disadvantages:

## Lecture 5

PCA : projection : min reconstruction error, max variance

Goal

Change of Basis

Find an appropriate basis

Noise

Rotation

Redundancy

Covariance Matrix

Algorithm

Dimensionality Reduction

Feature selection

Greedy Forward feature selection

Lower Dimensional Projections

SVD  $\rightarrow X = VSU^t$

EigenFaces

Size of covariance matrix

Computation of Eigenvectors

Choosing the dimension  $K$

Singular value decomposition  $\rightarrow$  example

## Lecture 6

FDA

Derivation

- Example
- MDA
- Drawbacks
- Generative vs discriminative
- Parametric methods vs Discriminant Functions
- Approaches
- LDF two class, multi class

## Lecture 7

- Linear regression
- Logistic regression
- Gradient descent/ascent - min/max
- Perceptron
- Linear regression example
- Assumptions
- LSE
- Optimisation
- GD
- LDF criterion function
- Perceptron
  - Example
- LDF
  - Separable
  - Non separable
- Convergence perceptron rule

## Lecture 8

- MSE
- SVM

## Lecture 9

- Ensemble methods

