Linear & Non-Linear Discriminant Analysis

Hugh R. Wilson

- PCA Review
- Supervised learning
- Fisher linear discriminant analysis
- Nonlinear discriminant analysis
- Research example
- Multiple Classes
- Unsupervised learning

YORK UNIVERSITÉ UNIVERSITY

Centre for Vision Research

Reading

- Monday: Principal components: Kutz, pp. 387 (bottom) - 393
- Today: Linear discriminant analysis: Kutz, pp. 442 - 445

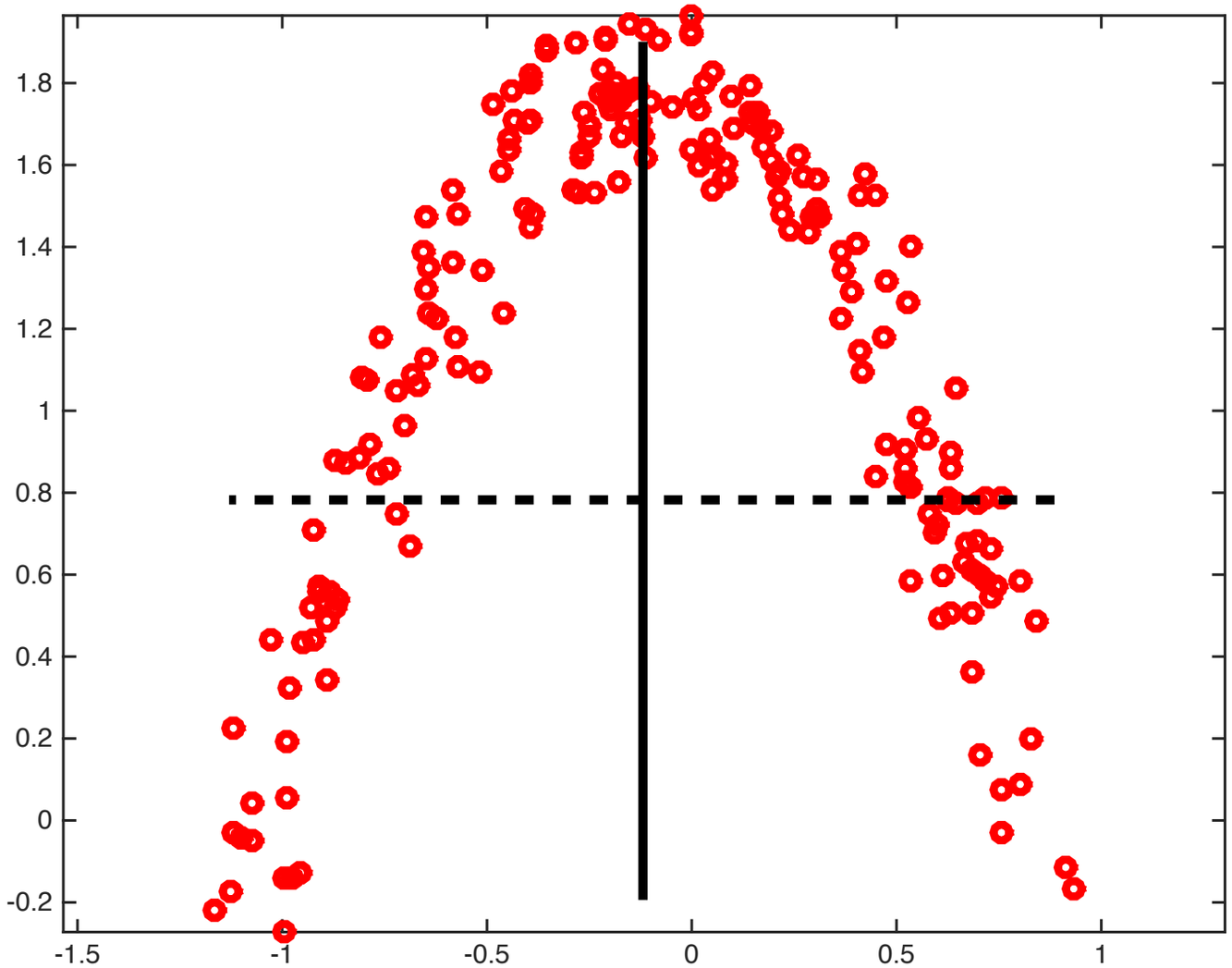
PCA Summary

- PCA reveals data structure determined by covariance
- Calculation for M measurements, N samples (N > M):
  - Subtract means from measurements
  - Data covariance matrix CV
  - [Vector, EV] = eig(CV)
- Dimensionality reduction (e.g. retain 95% of variance)
- Can be implemented by brain learning networks

YORK UNIVERSITÉ UNIVERSITY

Nonlinear PCA

- Nonlinear axes (a manifold)
- Problem: not know shape of data in multi-dimensions
- Similar problem in Discriminant Analysis



Supervised Category Learning

- Issue: learn to discriminate category to which each individual belongs based on measurements
- M measurements on each individual
- Each individual labeled by category (e.g. dog vs. cat)
- Supervised: learning algorithm knows the correct category in each case
- Simplest example: Fisher linear discriminant analysis
- Two categories (can be generalized)
- Use result to categorize new data (cross validation)

YORK UNIVERSITÉ UNIVERSITY

Fisher Linear Discriminant Analysis

- Maximize ratio of covariance between classes to covariance within classes by projection onto vector V
- Covariance Between: Cov_{Bet}
- Covariance Within: Cov_{Win}
- $Cov_{Win}*V = \lambda\ Cov_{Bet}*V$ (generalized eigenvalue problem)
- Solution: $V = eig(inv(Cov_{Win})*Cov_{Bet}))$
- V = vector for maximum class separation
- Project onto vector & find point for maximum discrimination
- Matlab demo

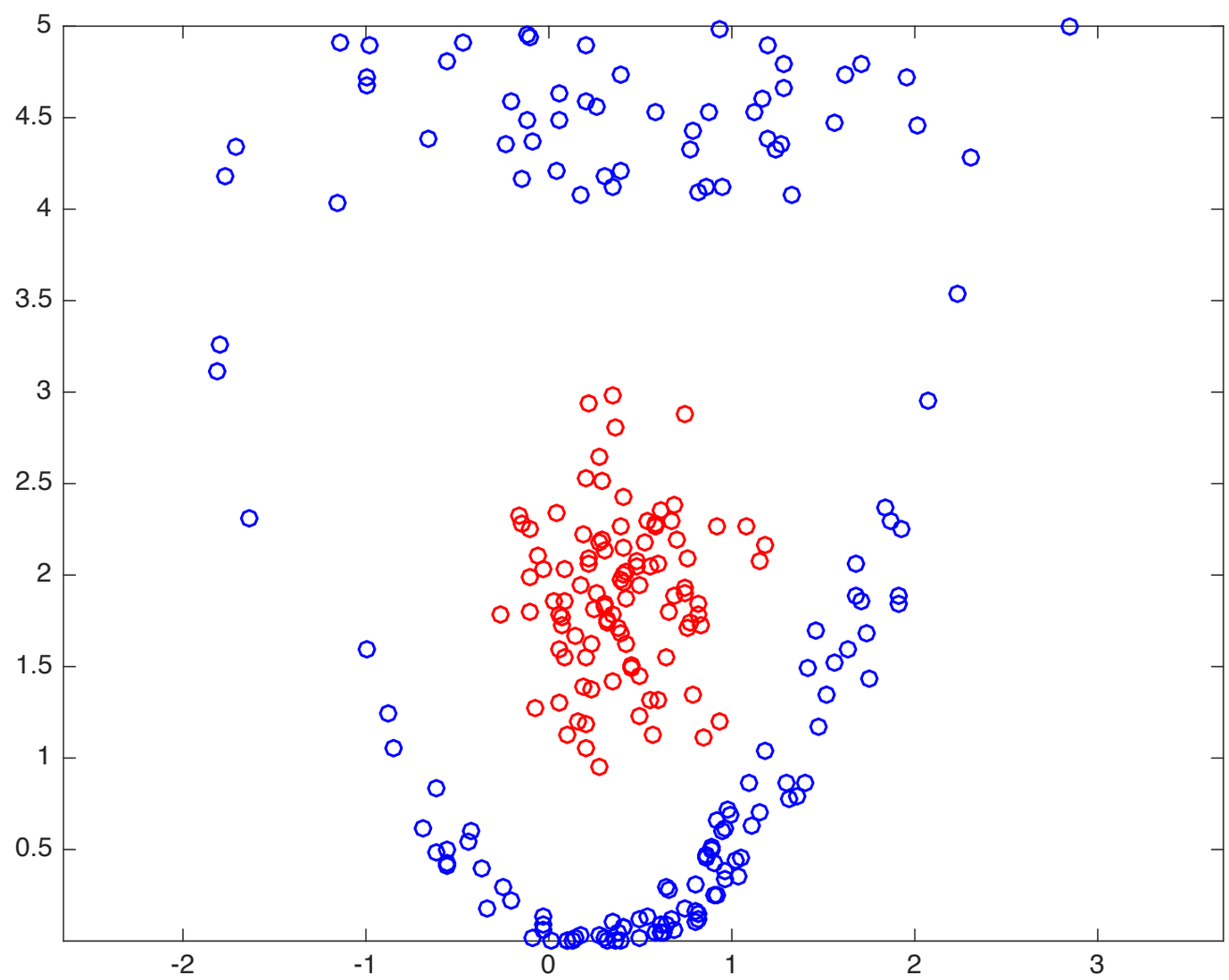
Fisher LDA Matlab Code

- Calculate between and within covariances
- Eigenvalue problem
- Maximize ratio of CovBetween/CovBetween

```
CovBetween = cov([Dog; Cat]);
CovDog = cov(Dog);
CovCat = cov(Cat);
CovWithin = (CovDog + CovCat)/2; (average category cov)
[FLD, FV] = eig(inv(CovWithin)*CovBetween); (eigenvector)
```

Extension to Non-Linear Class Boundaries

- Expand dimensionality of measurement space
- Include nonlinear functions of individual measurements
- x^2, y^2, xy, $(x^2 + y^2)$, other nonlinear combinations
- Problem: don't use too many or overfit data
- Matlab demo



Non-Linear Rationale

- Eigenvalue problem: linear algebra
- How can linear algebra deal with nonlinear boundaries?
- Increase dimensionality of data space with nonlinear measurement combinations x^2, xy, etc.
- Discriminant Analysis is still linear in finding the weighting of each of those dimensions
- Solution: $Ax + By + C x^2 + Dxy + E y^2$
- Linear in A, B, C, D, E

YORK UNIVERSITÉ UNIVERSITY

Discriminant Analysis of Child Faces

- Does face shape reveal whether child is male or female?
- Used 45 child faces
- PCA to reduce geometry from 28 to 9 dimensions (94% var)
- Non-linear discriminant analysis (included each PC^2)
- Predicted gender of 23 new child faces (cross-validation)
- Results: 80% correct on training faces
- 78% correct on novel test faces

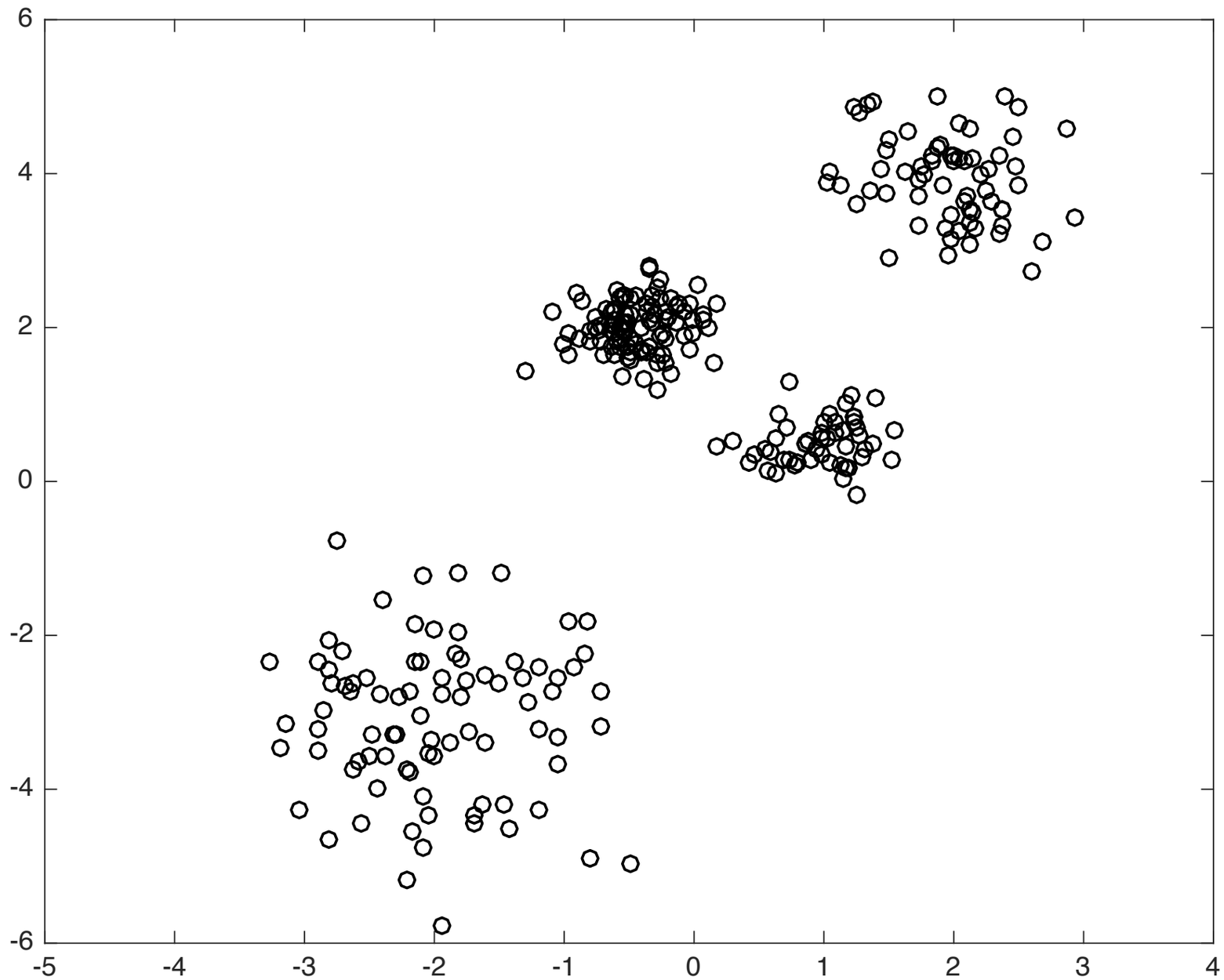
YORK UNIVERSITÉ UNIVERSITY

Extension to more Classes

- Example: 3 classes A, B, C
- Do multiple 2-class discriminant analyses:
    - A versus (B + C)
    - B versus (A + C)
    - C versus (A + B)
- Combine results
- Effective for both linear & non-linear discriminants
- Generalizes to many classes

Unsupervised Learning: Cluster Analysis

- Given data without category labels
- Must learn number of categories & category membership
- Analysis: find clusters of nearest neighbours



Summary

- PCA reveals data structure determined by eigenvalues of covariance matrix
- Fisher LDA (Linear Discriminant Analysis) reveals best axis for data projection to separate two classes
    - Eigenvalue problem for matrix $(Cov_{Bet})/(Cov_{Win})$
    - Generalizes to multiple classes
    - Non-linear Discriminant Analysis: add nonlinear combinations of measurements (extra dimensions)
- Dimensionality reduction by PCA often used first to simplify Discriminant Analysis (child face example)
- Computer simulations of brain networks can compute PCA, LDA & NLDA

YORK UNIVERSITÉ UNIVERSITY