PRINCIPAL COMPONENT ANALYSIS

1

Principal Component Analysis (PCA)

2

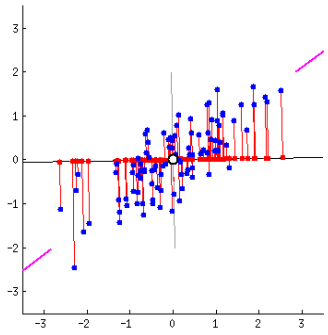
- PCA: Data Analysis Technique

2

PCA

3

- The optimal direction: The direction of maximum variance of the projected data



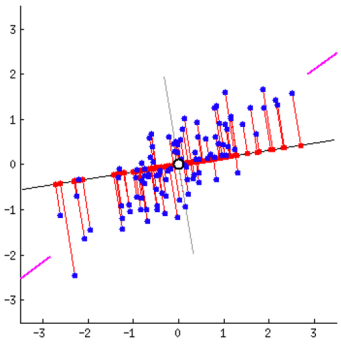
PCA

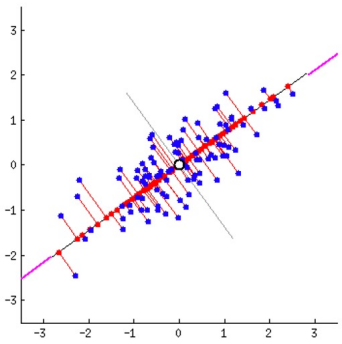
3

PCA

4

- The optimal direction: The direction of maximum variance of the projected data





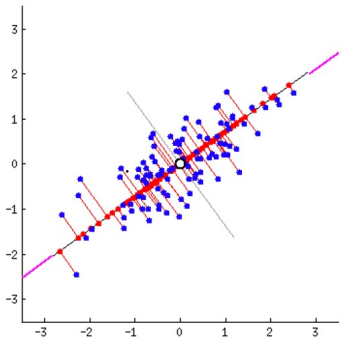
PCA

4

PCA

5

- The optimal direction: The direction of maximum variance of the projected data



Principal eigenvector of the covariance matrix of the data >> Direction of maximum variance of data

Eigenvalue determines the projection variance of the corresponding eigenvector

5

PCA: Fundamental Concepts

6

- Variance of a random variable fluctuating about its mean value
  - Average of the square of the fluctuations
- Covariance for a pair of random variables, each fluctuating about its mean value
  - Average of product of fluctuations
- N random variables: Covariance matrix
  - N x N symmetric matrix
  - Diagonal elements are variances of individual random variables

6

PCA: Fundamental Concepts

7

- Eigenvalues and eigenvectors of a square matrix A satisfy the following

$$Av_i = \lambda_i v_i$$

- Principal eigenvector
  - Eigen vector corresponding to maximal eigenvalue

Principal eigenvector of the covariance matrix of the data >> Direction of maximum variance of data

Eigenvalue determines the projection variance of the corresponding eigenvector

7

PCA Example

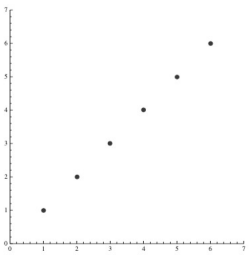
8

- Dataset

$$(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)$$

- Covariance matrix

	Column 1	Column 2
Column 1	2.91666667	
Column 2	2.91666667	2.91666667



- Eigen values

$\lambda_1 = \frac{2917}{500}$

$\lambda_2 = 0$

- Eigen vectors

$v_1 = (1, 1)$

$v_2 = (-1, 1)$

8

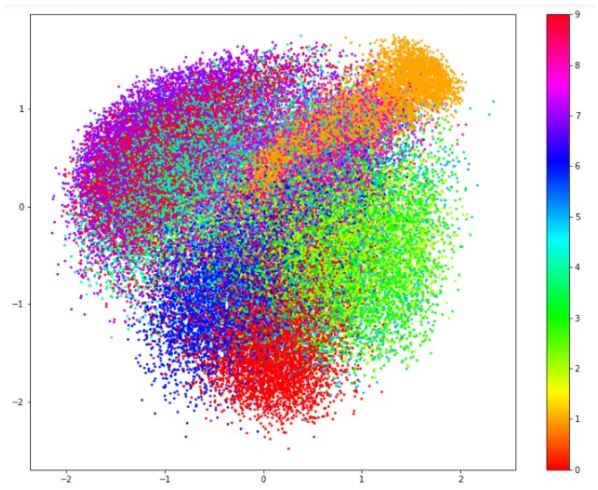
PCA: Applications

9

- General Dimensionality Reduction
  - Feature vector
- Visualization of High-dimensional Data
  - Feature space visualization

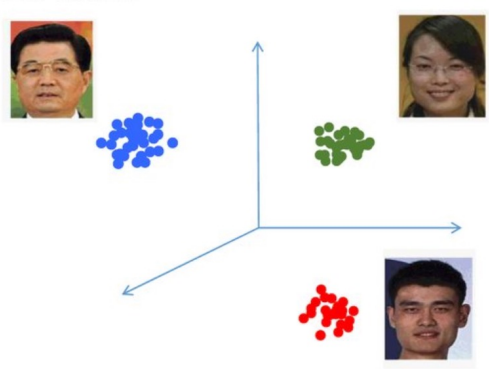
9

PCA for Feature Space Visualization



10

PCA for Feature Space Visualization



11

Thank You!

12