PRINCIPAL COMPONENT ANALYSIS

1

Principal Component Analysis (PCA)

2

- PCA: Data Analysis Technique

PCA

2

PCA: Introduction

3

- 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

PCA

3

PCA (optional)

4

- Covariance matrix
  - Let $x_1\, x_2 \dots x_M$ be a set of M $N \times 1$ vectors and let be their $\overline{x}$ mean vector

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iN} \end{bmatrix} \qquad \overline{x} = \frac{1}{M}\sum_{i=1}^{M}\begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iN} \end{bmatrix}$$

  - Construct a matrix $X = [x_1 - \overline{x}\;\; x_2 - \overline{x} \dots x_M - \overline{x}]$
  - For the above set of observations, N x N covariance matrix can be computed as follows

$$C = XX^T = [x_1 - \overline{x}\;\; x_2 - \overline{x} \dots x_M - \overline{x}\;]\begin{bmatrix} (x_1 - \overline{x})^T \\ (x_2 - \overline{x})^T \\ \vdots \\ (x_M - \overline{x})^T \end{bmatrix}$$

PCA

4

PCA

5

- 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

PCA

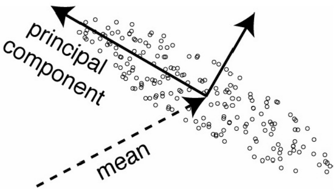
5

PCA (optional)

6

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

Find the direction w^* for which the variance is maximized



$$w^* = \arg\max_w \quad \mathrm{var}\left(w^T X\right)$$

$$\text{subject to: } w^T w = 1$$

$$\mathrm{var}\left(w^T X\right) = \left(w^T X\right)\left(w^T X\right)^T = w^T\left(XX^T\right)w = w^T C w$$

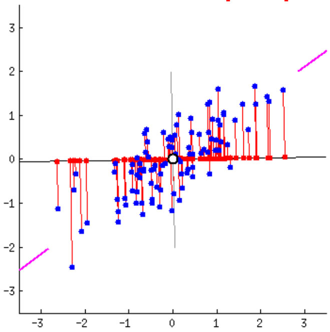
PCA

6

PCA

7

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



7

PCA (optional)

8

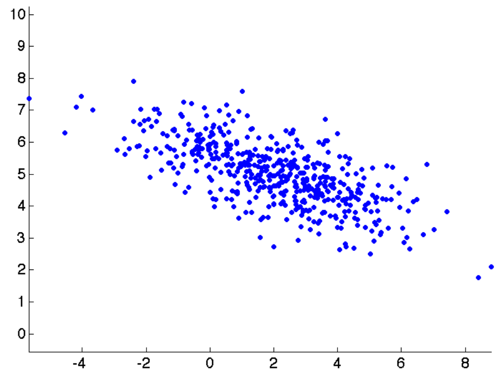
- Solving the constrained optimization (Lagrangian Multiplier Method)
$$L(\mathrm{w},\lambda_1) = \mathrm{w}^T C \mathrm{w} + \lambda_1 (1 - \mathrm{w}^T \mathrm{w})$$
- Computing the derivative of L relative to w and setting to zero
$$L(\mathrm{w},\lambda_1) = \mathrm{w}^T C \mathrm{w} + \lambda_1 (1 - \mathrm{w}^T \mathrm{w})$$
$$2C\mathrm{w} - 2\lambda_1 \mathrm{w} >> C\mathrm{w} = \lambda_1 \mathrm{w}$$
- The above is the eigenvalue problem
- Multiplying both the eqn. by w^T : $\mathrm{w}^T C \mathrm{w} = \lambda_1$
- Eigenvalue determines the projection variance of the corresponding eigenvector

8

PCA

9

- Distribution of a set of correlated 2D data



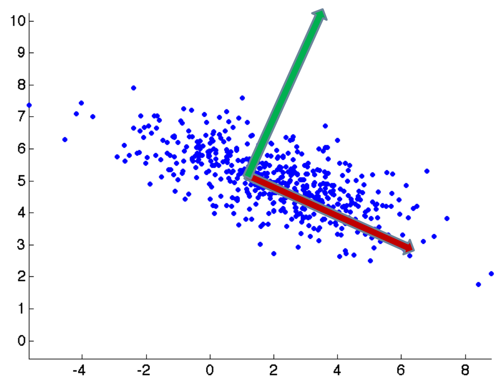
PCA

9

PCA

10

- Eigenvectors of the covariance matrix



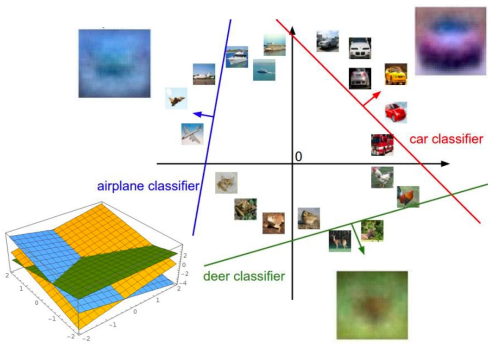
PCA

10

PCA: Applications

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

PCA for Feature Space Visualization



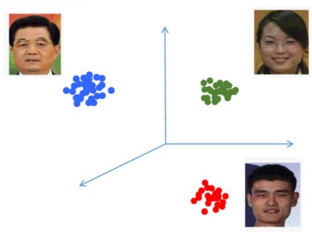
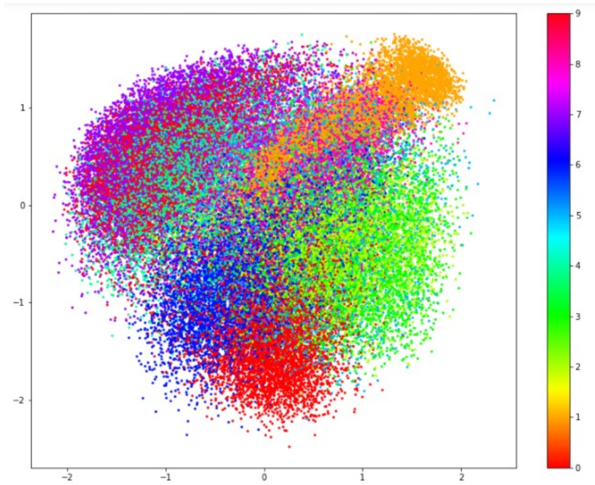


Image source: Stanford University CS 231n

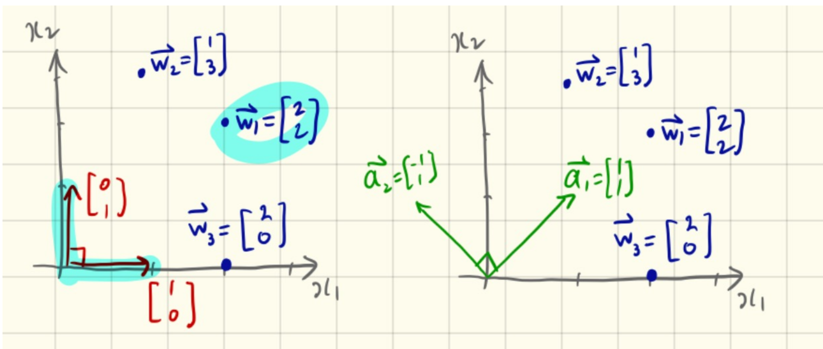
PCA for Feature Space Visualization



13

Change of Basis

14



PCA

14

Thank You!

15