CS475/CS675
Lecture 23: July 19, 2016

Principal Component Analysis, Eigenfaces

Principal Component Analysis (PCA)

- Data exploration technique:
  - Dimensionality reduction
  - Principal components are axes that preserve most of the variance in the data

- Picture

Empirical Variance

- Data: X

- Empirical mean: $\mu = \frac{1}{n}\sum_{i=1}^{n} X_i$

- Empirical covariance:
  $\mathrm{C} = \frac{1}{n-1}\sum_{i=1}^{n}[X_i - \mu][X_i - \mu]^T$

Principal Component

- Axis v that preserves the most variance

$$\max_v \frac{1}{n-1} \sum_{i=1}^{n} \left[X_i^T v - \mu^T v\right]^T \left[X_i^T v - \mu^T v\right]$$
$$= \max_v \frac{1}{n-1} \sum_{i=1}^{n} v^T [X_i - \mu][X_i - \mu]^T v$$
$$= \max_v v^T \left[\frac{1}{n-1} \sum_{i=1}^{n} [X_i - \mu][X_i - \mu]^T\right] v$$
$$= \max_v v^T C v$$

When $\left||v|\right| = 1$, then

$\max_v v^T C v$ is

v is

Principal Component Analysis

- Eigendecomposition of the empirical covariance matrix

$$C = Q\Lambda Q^T$$

- Eigenvector: dimension (or basis function)
- Eigenvalue: amount of variance preserved in that dimension

Dimensionality Reduction

- Problem: what is the smallest linear subspace (i.e., fewest dimensions) that captures 95% of the variance?
- Solution: retain eigenvectors of the k largest eigenvalues such that k is the smallest integer that satisfies

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{n} \lambda_j} \geq 0.95$$

- Picture:

Example: Eigenfaces

- Turk and Pentland (1991):
  - Image compression
  - Face detection

- Solution:
  - embed images in low dimensional eigenspace
  - Face detection: nearest neighbour in eigenspace

Principal Component Analysis

- Data: X (m pixels $\times$ n images)
- Covariance matrix: C ($n \times n$)
- Eigendecomposition: $\mathrm{C} = Q\Lambda Q^T$

Eigenfaces

Dataset

Mean image

Eigenfaces

Face Detection

- Project each image in the space if eigenfaces
  - I.e., approximate each image as a linear combination of the eigenfaces

$$\tilde{v} \leftarrow Q^T v$$

- Face detection: find matching image in a database
  - Nearest neighbour in space of eigenfaces

$$w^* = argmin_{\tilde{w}} \left|\left|\tilde{v} - \tilde{w}\right|\right|_2$$