图像处理与机器学习

Digital Image Processing and Machine Learning

主讲人：黄琳琳

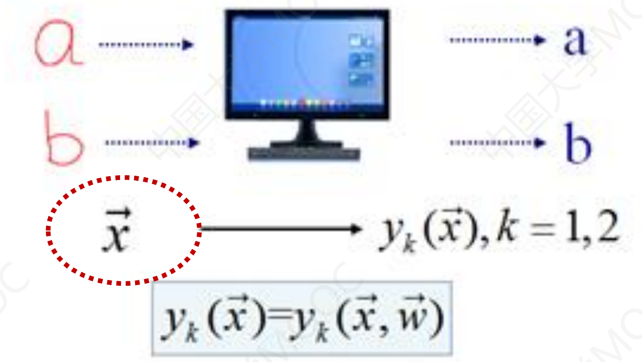
电子信息工程学院

第八章　非监督学习

- 引言
- 聚类算法
- 主成份分析

主成份分析

➤ 举例：手写字符识别



- 维度过高引起过拟合
- 特征之间具有相关性

降 维

主成份分析

(Principal Component Analysis, PCA)

主成份分析

- PCA又称为K-L变换 (Karhunen–Lo ève transform)

  -- 1901 年Pearson 在研究回归分析时附带提出的

  -- 1933 年由Hotelling 奠定数学基础。

  -- 1947年美国统计学家斯通(stone)应用到经济分析

主成份分析

- 1947年美国的统计学家斯通，对经济研究时
  - -- 利用美国1929—1938年各年的数据
  - -- 得到了17个反映国民收入与支出的变量

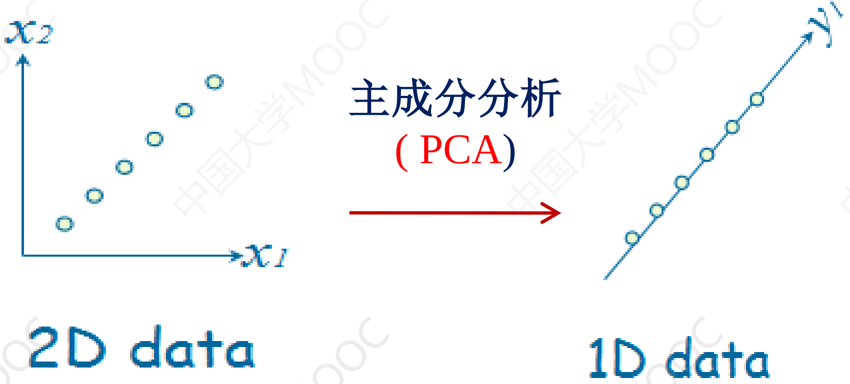
✓ 雇主补贴
✓ 消费资料和生产资料
✓ 纯公共支出
✓ 净增库存
✓ 股息
✓ 利息
✓ 外贸平衡等等
……

$\xrightarrow[\text{(PCA)}]{\text{主成分分析}}$ 三个新变量

97.4% 分析精度

主成份分析

- PCA 试图在力保数据信息丢失最少的原则下
  - -- 对多变量数据进行最佳综合简化
  - -- 对高维变量空间进行 降 维 处理



主成份分析

- PCA可以通过求解特征矢量来得到主成份

实对称矩阵 Σ

-- 特征矢量及特征值 $\Sigma\phi_i = \lambda_i\phi_i$

-- 特征矢量构成**正交**特征空间

$$\Phi = [\phi_1, \phi_2, \cdots \phi_d] \qquad \Lambda = diag[\lambda_1, \lambda_2, \cdots, \lambda_d]$$

$$\Phi^T\Phi = I \qquad \Longleftrightarrow \qquad \phi_i^t\phi_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$\Phi^T = \Phi^{-1}$$

主成份分析

- 实对称矩阵 Σ

$$\Sigma\phi_i = \lambda_i\phi_i \qquad \Phi = [\phi_1, \phi_2, \cdots \phi_d]$$

$$\Lambda = diag[\lambda_1, \lambda_2, \cdots, \lambda_d]$$

$$\Sigma\Phi = \Phi\Lambda \qquad \Sigma = \Phi\Lambda\Phi^T$$

$$\Phi^T\Sigma\Phi = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_d \end{bmatrix}$$

矩阵对角化

主成份分析

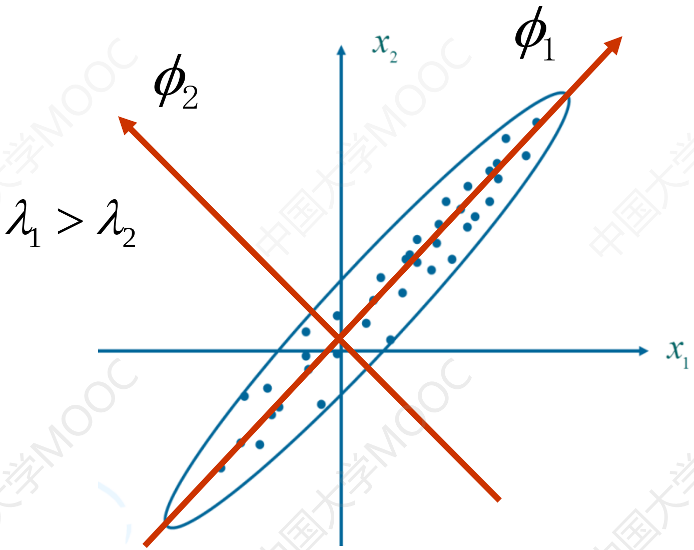
- 应用：Principal component analysis (PCA)
  - 利用n个样本矢量求取协方差矩阵
  - 求取协方差矩阵的特征矢量及特征值
  - 将特征矢量按照特征值由大到小排序
  - 选择特征值最大的 m ($m<d$)个特征向量构成低维子空间
  - 将随机矢量投影到低维子空间
  - 使子空间投影的重建误差最小

主成份分析

$$\Sigma = \Phi\Lambda\Phi^T \quad \Lambda = diag[\lambda_1, \lambda_2, \cdots, \lambda_d]$$

特征矢量按照特征值由大到小排序

$$\Phi = [\phi_1, \phi_2, \cdots \phi_d] \quad \lambda_1 > \lambda_2, \cdots, > \lambda_d$$

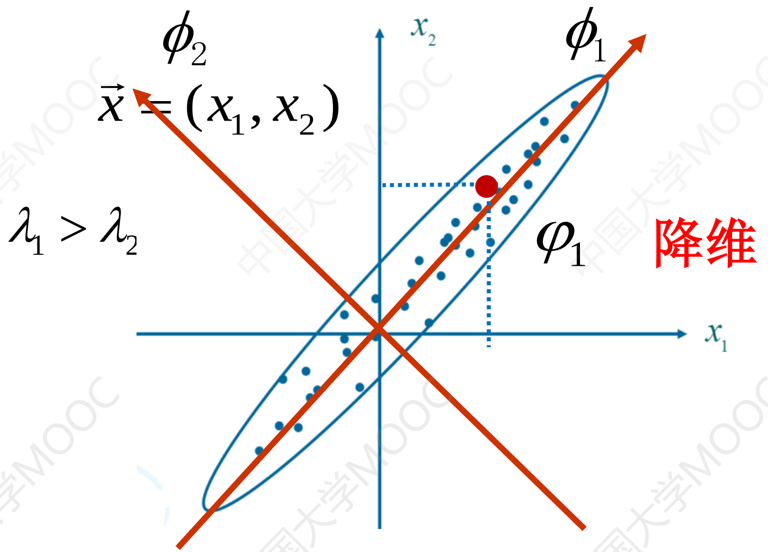


特征矢量构成

正交特征空间

主成份分析

- 应用：Principal component analysis (PCA)
  - 将随机矢量投影到低维子空间
  - 投影即为该矢量特征表达（低维）



主成份分析

线性空间中正交变换($\Phi^T\Phi = I$) 不影响欧氏距离

$$\sum_{j=1}^{d}[(\mathbf{x}-\mu)^T\phi_j]^2 = \|\mathbf{x}-\mu\|^2$$

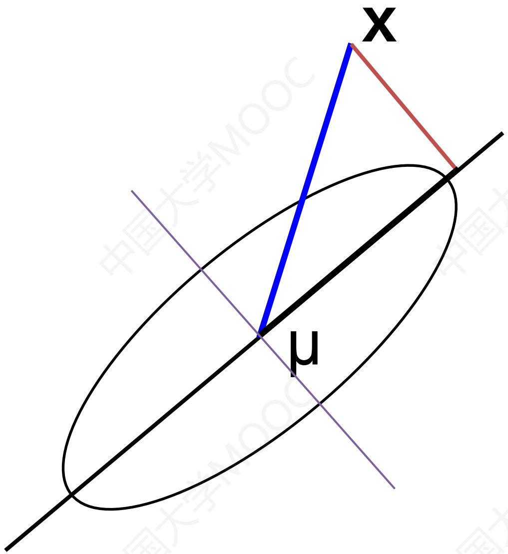
子空间投影 $\mu+\sum_{j=1}^{m}[(\mathbf{x}-\mu)^T\phi_j]\phi_j$

投影重建误差

$$r_E = \|\mathbf{x}-\mu\|^2 - \sum_{j=1}^{m}[(\mathbf{x}-\mu)^T\phi_j]^2 = \sum_{j=m+1}^{d}[(\mathbf{x}-\mu)^T\phi_j]^2$$

期望 $$\mathscr{E}(r_E) = \mathscr{E}\left\{\sum_{j=m+1}^{d}[(\mathbf{x}-\mu)^T\phi_j]^2\right\} = \mathscr{E}\left\{\sum_{j=m+1}^{d}\phi_j^T(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T\phi_j\right\}$$

$$= \sum_{j=m+1}^{d}\phi_j^T\mathscr{E}[(\mathbf{x}-\mu)(\mathbf{x}-\mu)^T]\phi_j = \sum_{j=m+1}^{d}\phi_j^T\Sigma\phi_j = \boxed{\sum_{j=m+1}^{d}\lambda_j}$$



谢 谢

本课程所引用的一些素材为主讲老师多年的教学积累，来源于多种媒体及同事和同行的交流，难以一一注明出处，特此说明并表示感谢！