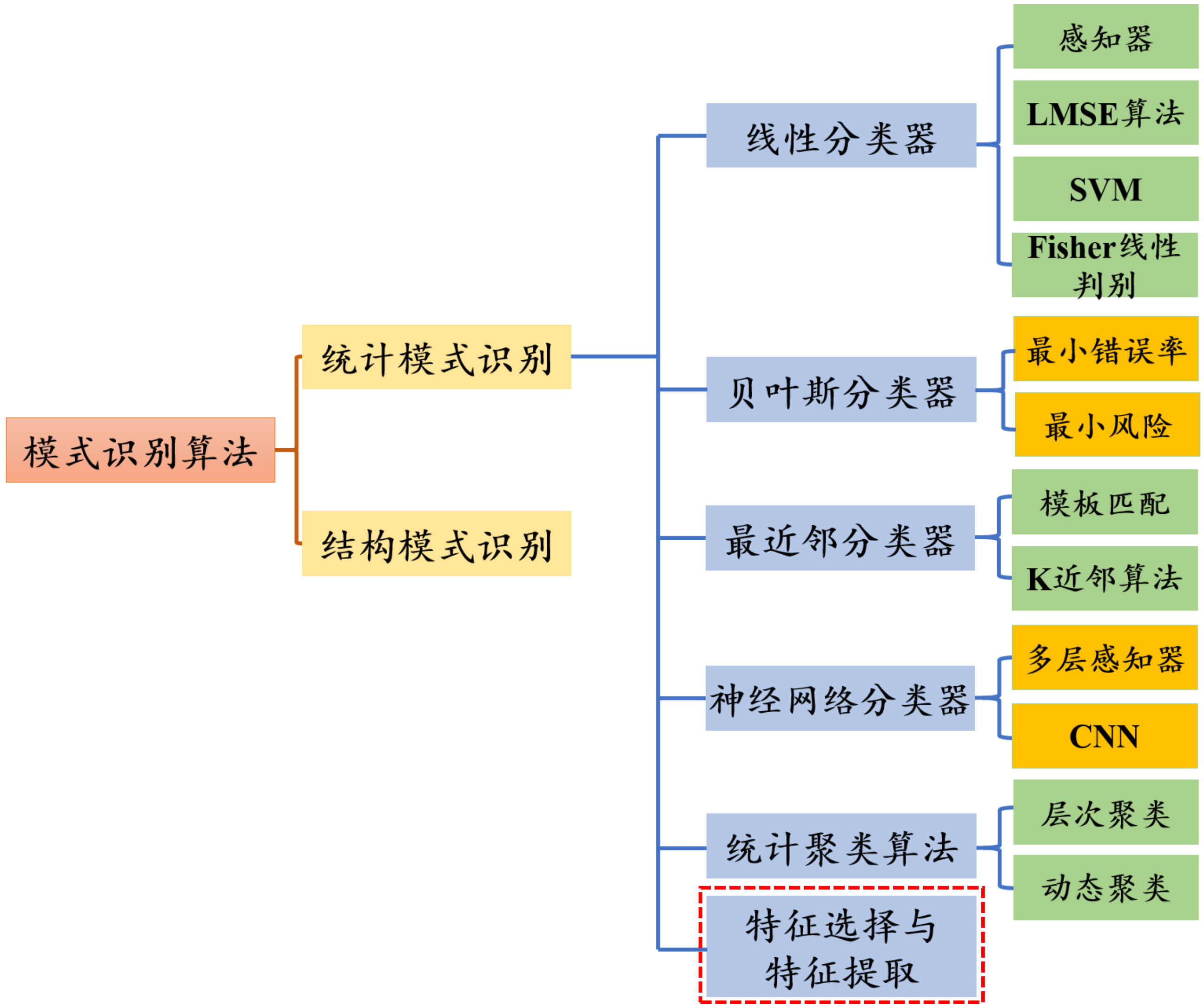
模式识别课程

特征选择与提取

张俊超

中南大学
航空航天学院

模式识别-特征选择与提取



特征选择与提取– KL变换

- Karhunen-Loeve (KL)变换

对D维随机向量$\mathbf{x}\in\mathbb{R}^D$,可以用一个完备的正交归一向量系$\mathbf{u}_j$,

$j=1,2,\cdots,\infty$来展开：

$$\mathbf{x}=\sum_{j=1}^{\infty}c_j\mathbf{u}_j$$

其中，$\mathbf{u}_i^T\mathbf{u}_j=\begin{cases}1, i=j\\0, i\neq j\end{cases}$

c_j:线性组合的系数

特征选择与提取- KL变换

$$\mathbf{x}=\sum_{j=1}^{\infty} c_j \mathbf{u}_j$$

$$\mathbf{u}_j^T \mathbf{x}=\sum_{j=1}^{\infty} c_j \mathbf{u}_j^T \mathbf{u}_j$$

$$c_j=\mathbf{u}_j^T \mathbf{x}$$

如果只用有限的d项($d < D$)来逼近$\mathbf{x}$,即

$$\bar{\bar{\mathbf{x}}}=\sum_{j=1}^{d} c_j \mathbf{u}_j$$

$\bar{\bar{\mathbf{x}}}$与原向量$\mathbf{x}$的均方差是：

$$c_j = \mathbf{u}_j^T \mathbf{x}$$

$$e = E\left[\left(\mathbf{x}-\bar{\bar{\mathbf{x}}}\right)^T\left(\mathbf{x}-\bar{\bar{\mathbf{x}}}\right)\right]$$

$$= E\left[\left(\sum_{j=d+1}^{\infty} c_j \mathbf{u}_j\right)^T\left(\sum_{j=d+1}^{\infty} c_j \mathbf{u}_j\right)\right] = E\left[\sum_{j=d+1}^{\infty} c_j^2\right]$$

$$= E\left[\sum_{j=d+1}^{\infty} \mathbf{u}_j^T \mathbf{x}\mathbf{x}^T \mathbf{u}_j\right] = \sum_{j=d+1}^{\infty} \mathbf{u}_j^T E\left[\mathbf{x}\mathbf{x}^T\right]\mathbf{u}_j$$

$$= \sum_{j=d+1}^{\infty} \mathbf{u}_j^T \Phi \mathbf{u}_j$$

特征选择与提取- KL变换

$$\min e = \sum_{j=d+1}^{\infty} \mathbf{u}_j^T \boldsymbol{\Phi} \mathbf{u}_j, s.t. \mathbf{u}_j^T \mathbf{u}_j = 1, \forall j$$

拉格朗日乘子

$$f(\mathbf{u}) = \sum_{j=d+1}^{\infty} \mathbf{u}_j^T \boldsymbol{\Phi} \mathbf{u}_j - \sum_{j=d+1}^{\infty} \lambda_j \left(\mathbf{u}_j^T \mathbf{u}_j - 1 \right)$$

$$\frac{\partial f(\mathbf{u})}{\partial \mathbf{u}_j} = \left(\boldsymbol{\Phi} - \lambda_j \mathbf{I} \right) \mathbf{u}_j = 0, j = d+1, \cdots, \infty$$

特征值分解

$$\boxed{\boldsymbol{\Phi} \mathbf{u}_j = \lambda_j \mathbf{u}_j}$$

特征选择与提取- KL变换

$$e=\sum_{j=d+1}^{\infty}\lambda_j\mathbf{u}_j^T\mathbf{u}_j=\sum_{j=d+1}^{\infty}\lambda_j$$

要用d项$(d<D)$来逼近$\mathbf{x}$，使误差最小：则应该把矩阵$\mathbf{\Phi}$的特征值从大到小的顺序排列，选择前d个特征值对应的特征向量。

$\mathbf{u}_j, j=1,2,\cdots,d$组成了新的特征空间，样本$\mathbf{x}$在新空间上的展开系数$c_j=\mathbf{u}_j^T\mathbf{x}$，$j=1,2,\cdots,d$就组成了样本的新的特征向量。（KL变换）

$$\begin{bmatrix}\mathbf{u}_1^T\\ \mathbf{u}_2^T\\ \vdots\\ \mathbf{u}_d^T\end{bmatrix}\mathbf{x}=\begin{bmatrix}c_1\\ c_2\\ \vdots\\ c_d\end{bmatrix}$$

特征选择与提取–PCA

- 主成分分析(Principle Component Analysis)

样本做去均值的操作 $\mathbf{x}=\mathbf{x}\text{-}\boldsymbol{\mu}=\sum_{j=1}^{\infty} c_j \mathbf{u}_j$

$$e = E\left[\left(\mathbf{x}-\bar{\bar{\mathbf{x}}}\right)^T\left(\mathbf{x}-\bar{\bar{\mathbf{x}}}\right)\right]$$

$$= E\left[\left(\sum_{j=d+1}^{\infty} c_j \mathbf{u}_j\right)^T\left(\sum_{j=d+1}^{\infty} c_j \mathbf{u}_j\right)\right] = E\left[\sum_{j=d+1}^{\infty} c_j^2\right]$$

$$= E\left[\sum_{j=d+1}^{\infty} \mathbf{u}_j^T \mathbf{x}\mathbf{x}^T \mathbf{u}_j\right] = \sum_{j=d+1}^{\infty} \mathbf{u}_j^T E\left[\mathbf{x}\mathbf{x}^T\right]\mathbf{u}_j$$

$$= \sum_{j=d+1}^{\infty} \mathbf{u}_j^T \boldsymbol{\Phi} \mathbf{u}_j$$

Φ： 协方差矩阵

特征选择与提取–PCA

- 协方差矩阵的特征值分解：

$$\mathbf{\Phi}=\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$$

Q：正交矩阵（每一列是一个特征向量）

Λ：对角阵（主对角线上的元素是特征值）

PCA变换后： $\mathbf{y}=\mathrm{Q}^T\mathbf{x}$

特征选择与提取–PCA

- PCA流程：
  - ➢样本进行去均值的操作
  - ➢估计协方差矩阵
  - ➢对协方差矩阵进行特征值分解，并按照特征值从大到小进行排序，对应地调整特征向量→**Q**
  - ➢变换矩阵**P=Q’**(转置)
  - ➢降维时：**选择变换矩阵的前*d*行**


谢谢聆听！