

# 第22章 无监督学习方法总结

- 22.1 无监督学习方法总结

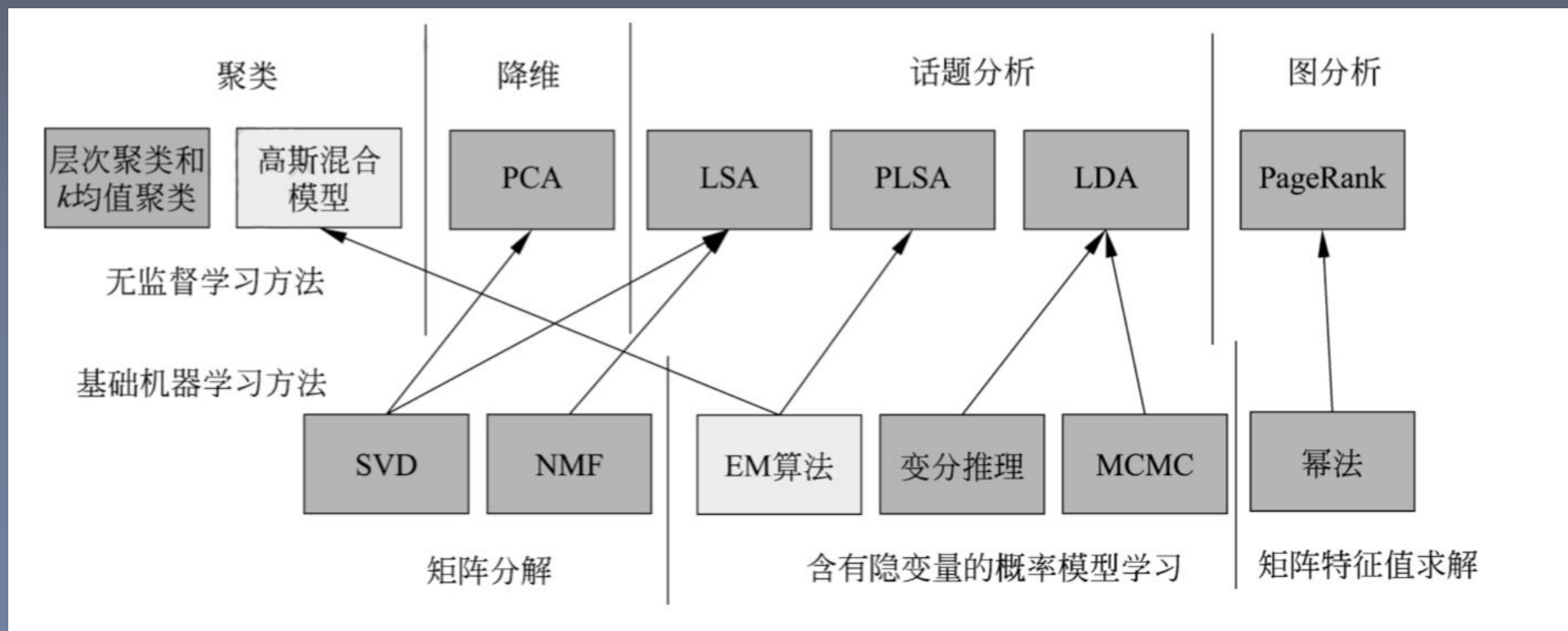
导师：Irene

---



# 各种方法之间的关系

## Correlation among Methods



# 各种方法之间的关系

## Correlation among Methods

### • 无监督学习

- 聚类 →
  - 聚类的方法
    - 层次聚类
    - K均值聚类
    - 高斯混合模型
- 降维 →
  - 降维的方法
    - PCA
- 话题分析 →
  - 话题分析的方法
    - LSA
    - PLSA
    - LDA
- 图分析 →
  - 图分析的方法
    - PageRank



# 各种方法之间的关系

## Correlation among Methods

- 基础方法

线性代数  
问题

- 矩阵分解

- 矩阵特征值求解

- 含有隐变量的概率模型估计

概率统计  
问题

- 矩阵分解的方法

- SVD
  - NMF

- 矩阵特征值求解的方法

- 幂法

- 含有隐变量的概率模型学习的方法

- EM算法
  - 变分推理
  - MCMC



# 各种方法之间的关系

## Correlation among Methods

无监督学习方法的特点				
	方法	模型	策略	算法
硬聚类	聚类	层次聚类	类内样本距离最小	启发式算法
		$k$ 均值聚类	$k$ 中心聚类	迭代算法
软聚类		高斯混合模型	似然函数最大	EM 算法
线性降维	降维	PCA	低维正交空间	SVD
非概率模型	话题分析	LSA	矩阵分解模型	SVD
		NMF	矩阵分解模型	非负矩阵分解
		PLSA	PLSA 模型	EM 算法
概率模型		LDA	LDA 模型	吉布斯抽样, 变分推理
	图分析	PageRank	有向图上的马尔可夫链	幂法

# 基础机器学习方法

## Basic Machine Learning Methods

表 22.2 含有隐变量概率模型的学习方法的特点

算法	基本原理	收敛性	收敛速度	实现难易度	适合问题
EM 算法	迭代计算、后验概率估计	收敛于局部最优	较快	容易	简单模型
变分推理	迭代计算、后验概率近似估计	收敛于局部最优	较慢	较复杂	复杂模型
吉布斯抽样	随机抽样、后验概率估计	依概率收敛于全局最优	较慢	容易	复杂模型

# 话题模型

## Topic Model

表 22.3 矩阵分解的角度看话题模型

方法	一般损失函数 $B(D\ UV)$	矩阵 $U$ 的约束条件	矩阵 $V$ 的约束条件
LSA	$\ D - UV\ _F^2$	$U^T U = I$	$V V^T = \Lambda^2$
NMF	$\ D - UV\ _F^2$	$u_{mk} \geq 0$	$v_{kn} \geq 0$
PLSA	$\sum_{mn} d_{mn} \log \frac{d_{mn}}{(UV)_{mn}}$	$U^T \mathbf{1} = \mathbf{1}$ $u_{mk} \geq 0$	$V^T \mathbf{1} = \mathbf{1}$ $v_{kn} \geq 0$

# 话题模型

## Topic Model

表 22.4 话题模型 LSA 和 NMF 的约束条件

方法	变量 $u_k$ 的约束条件	变量 $v_n$ 的约束条件
LSA	正交	正交
NMF	$u_{mk} \geq 0$	$v_{kn} \geq 0$



# 结语

## —— 结 语 ——

在这次课程中我们了解到了各种方法间的关系、基础  
机器学习方法特点等

希望在课下，大家都能

**掌握本节知识并且及时复习**



关注公众号深度之眼，后台回复 统计学，获取统计学习方法第二版电子书及其他AI必学书籍



**deepshare.net**

深度之眼

联系我们：

电话：18001992849

邮箱：service@deepshare.net

Q Q：2677693114



公众号



客服微信

关注公众号深度之眼，后台回复 统计学，获取统计学习方法第二版电子书及其他AI必学书籍