

第二十三讲 复习

2024.6.5

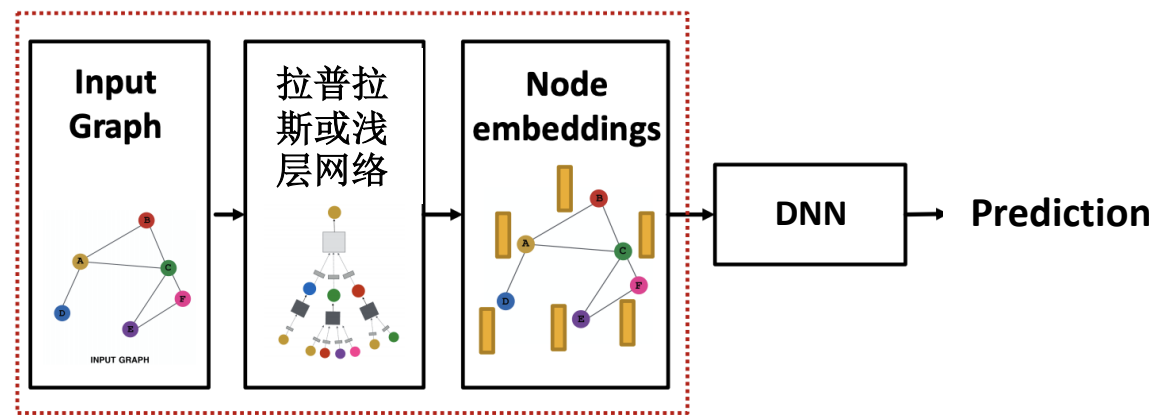
期末考试时间:2024-06-18 14:30-16:30; 地点:5204

DNN (两类): $p(\theta, \mathbf{x}_i) = P(y_i = 1 | \mathbf{x}_i) = \frac{1}{1 + \exp[-\alpha - \beta^T f^h(f^{h-1}(\dots f^1(\mathbf{x}_i)))]}$

优化目标函数: $J(\theta) = -\sum_{i=1}^n [y_i \log(p(\theta, \mathbf{x}_i)) + (1 - y_i) \log(1 - p(\theta, \mathbf{x}_i))]$,

梯度下降法求解 (反向传播): $\theta^{\text{new}} = \theta^{(0)} - \eta \frac{\partial J}{\partial \theta}(\theta^{(0)})$.

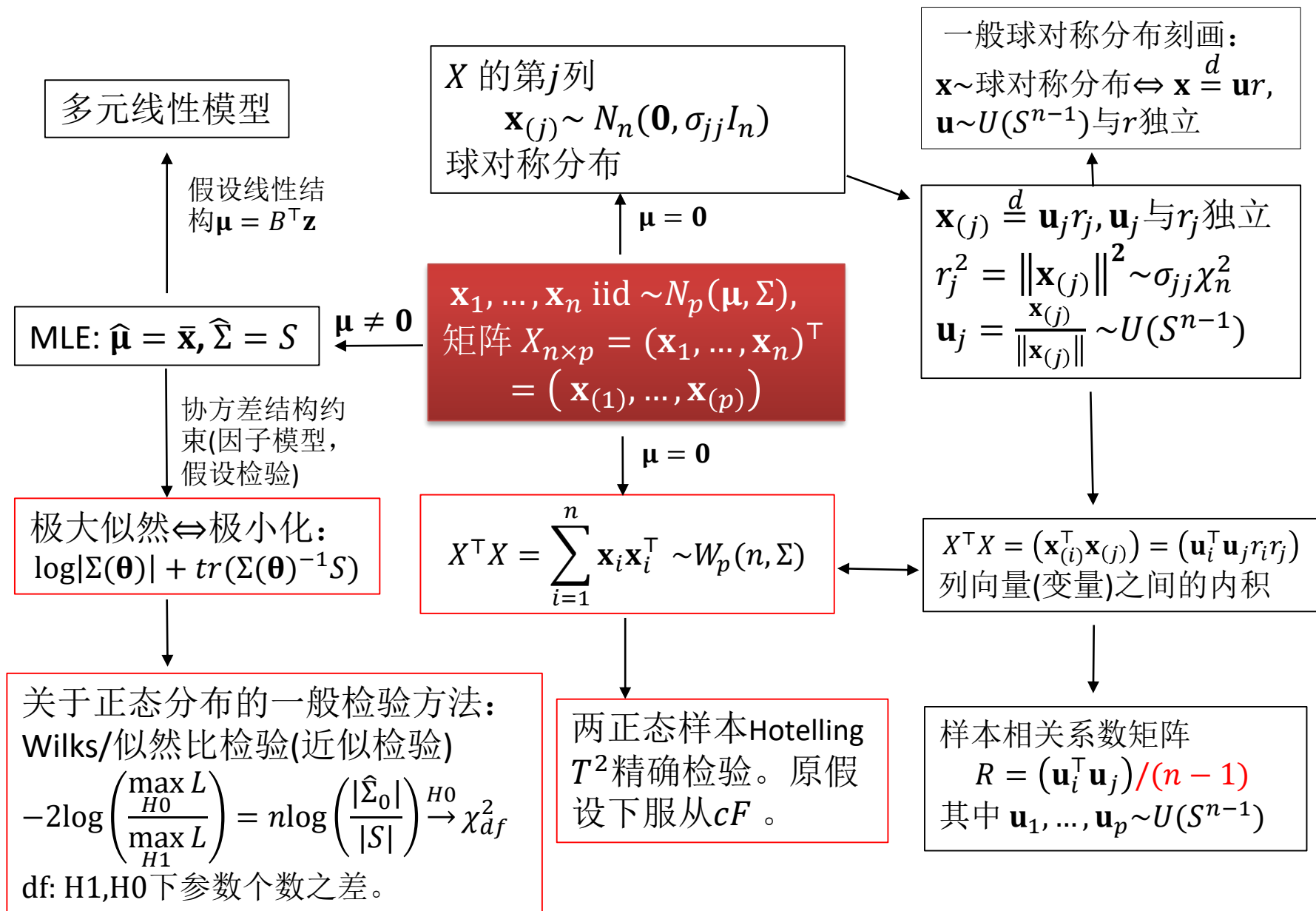
基本的深度神经网络(DNN) + 具体的embedding/encoding layer + 模型结构的改变 (相依结构, attention etc.) 生成多种DNN模型。以图神经网络Graph neural network (GNN,GCN,GAT...)为例:



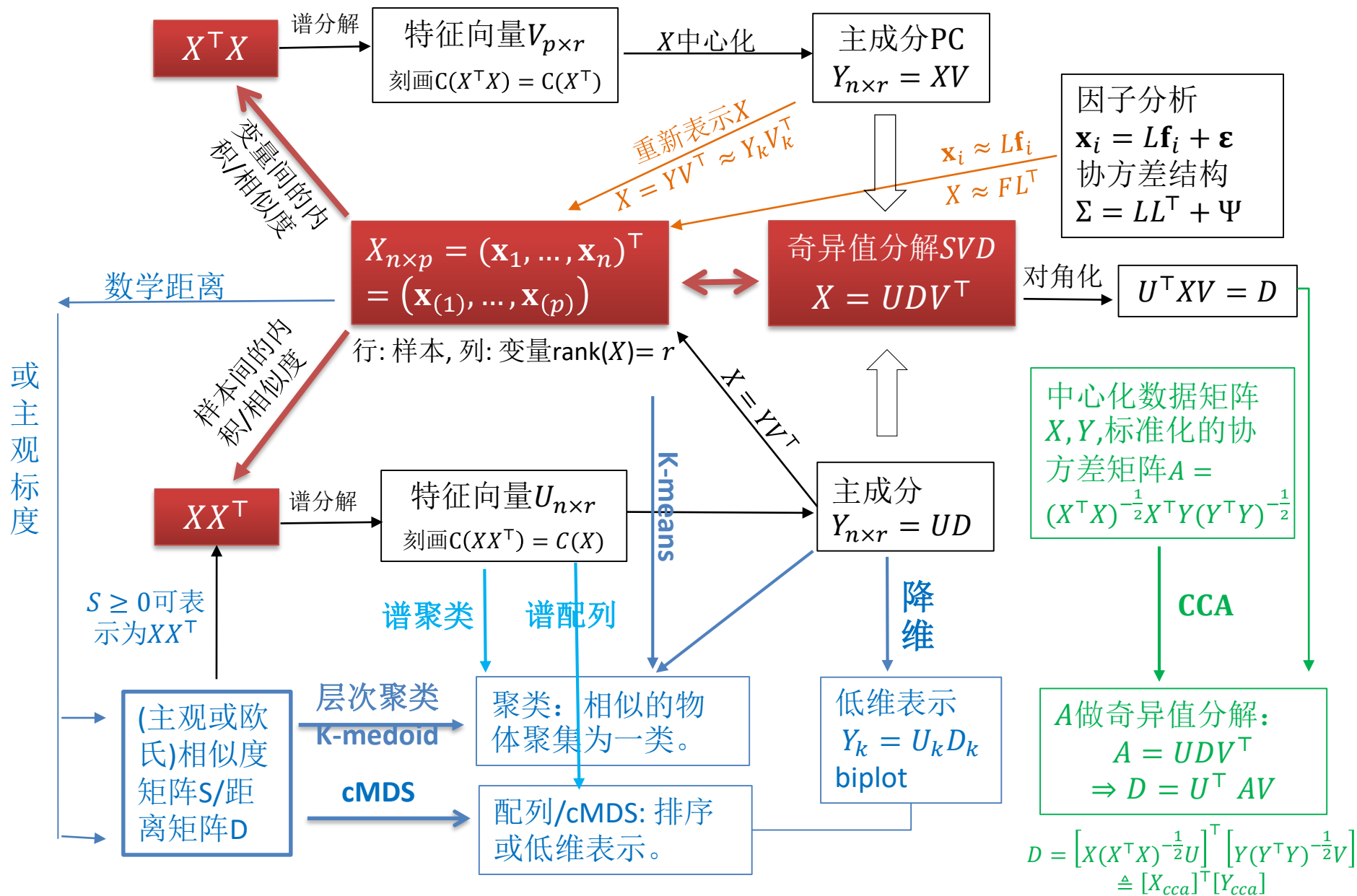
课程主要内容：前半部分和后半部分

- ❖ 第1-10讲：经典统计（normal 多元正态，微积分）
 - ❖ 第11-21讲：机器学习（singular SVD，线性代数）
-
- 7次纸质作业（hw6-7的参考解答将于今天下午上传）；
 - 5次上机作业（lab5已上传，不交）。
 - 若干阅读材料。

第1-10讲 (多元正态)



第11-21讲 (奇异值分解SVD)



复习、考试范围

- ❖ 复习范围：课件+纸质作业（忽略下面列表中删除的课件以及对应的作业题，忽略lab作业，忽略课件中的补充附录和虚线框内容）。
- ❖ 考试：与纸质作业难度、类型类似。带计算器。

课件

整行
删除

第一讲：多元分析简介 (2.26)	
第二讲：球对称分布 (2.28)	
第三讲：球对称分布 II (3.4)	球面均匀分布的边际
第四讲：多元正态 (3.11)	
第五讲：高斯图模型 (3.13)	高斯图模型、马氏随机场
第六讲：Wishart分布 (3.18)	
第七讲：Wishart分布 II (3.25)	矩阵拉直、Kronecker乘积、Wishart密度及其求解
第八讲：Hotelling's T^2 检验 (3.27)	
第九讲：多元方差分析MANOVA (4.1)	
第十讲：多元线性模型 (4.8)	

第十一讲：主成分分析 (4.11)	
第十二讲：双标图 (4.15)	
第十三讲：因子分析 (4.22)	
第十四讲： 结构方程模型 (4.24)	确认因子模型、路径模型、SEM
第十五讲：奇异值分解 (4.29)	
第十六讲： 对应分析 (5.6)	列联表背景
第十七讲： 对应分析 (续) (5.8)	列联表、丰度矩阵的对应分析
第十八讲： 典则相关分析 (5.13)	
第十九讲：距离与相似系数 (5.20)	
第二十讲：配列 (5.22)	
第二十一讲：多维标度法 (5.27)	
第二十二讲： 聚类和分类 (6.3)	K-means、高斯混合、谱聚类、分类
第二十三讲：复习 (6.5)	