

Table Of Contents:

- [今日工作](#)
- [复杂数据密度建模](#)
 - [t分布](#)
 - [因子分析](#)
 - [组合模型](#)
 - [总结](#)
- [编程练习](#)

今日工作

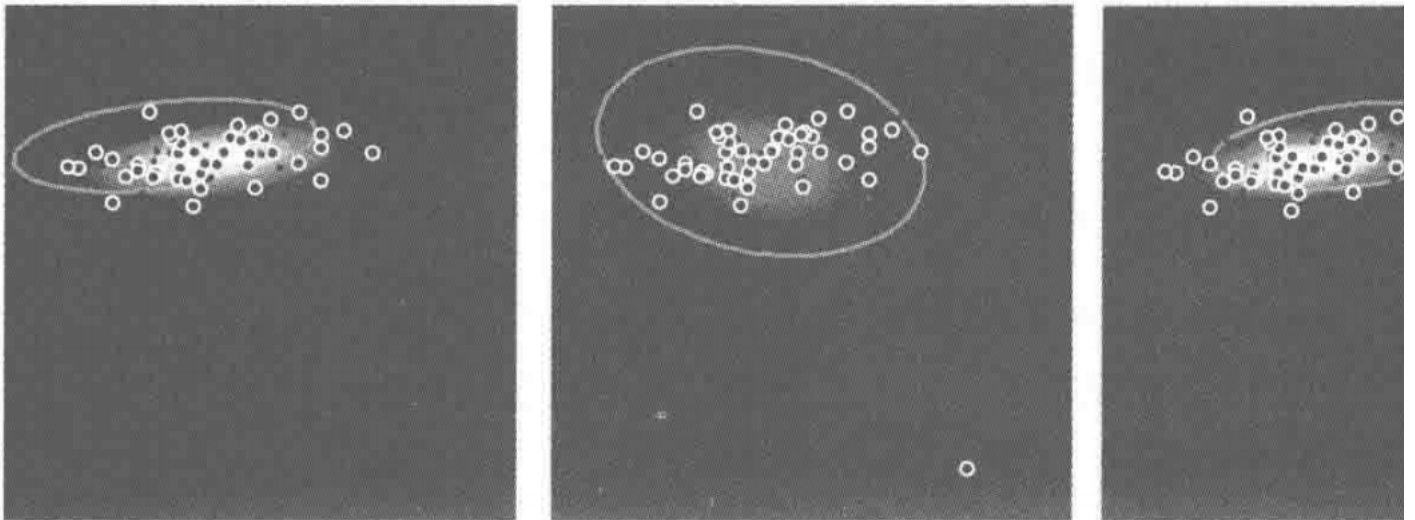
- 1、阅读《计算机视觉 模型、学习和推理》第七章的5-9节，完成第七章的阅读。
- 2、编程实现二元正态分布的可视化。
- 3、尝试实现混合高斯模型的拟合算法。

复杂数据密度建模

t分布

正态分布的缺陷:

不鲁棒: 边缘数据概率密度很低, 密度函数在边缘下降得很快。单个极端数据对参数影响很大。



- a) 拟合数据得多元正态分布模型。
- b) 增加单个极端值得拟合模型。模型完全改变。
- c) 增加极端点情况下的t分布的拟合。

t分布比正态分布多了一个参数——自由度, ν 。当 ν 无限大其成为正态分布。

因子分析

因子分析解决视觉数据维度太高的问题:

60*60的RGB图像要用60*60*3=10800维度的向量表示。

如果用多元正态分布建模, 协方差矩阵维度为10800*10800。

协方差矩阵描述了x内不同维度数据的关系。

如果假设各维度独立, 那么协方差矩阵就是一个对角矩阵, 只有10800个参数, 但是这样做丢失了太多信息。

因子分析提供的方式: 使用子空间描述高维空间。

二维空间内的点可以用一维函数描述, D维空间内的点可以用1, 2, ..., D-1维子空间函数描述。

组合模型

混合高斯模型: K个具有不同均值和方差的正态分布的加权和。

t分布: 具有相同均值不同方差的正态分布的**无限**加权和。

因子分析模型: 具有**不同**均值有相同对角协方差的正态分布的无限加权和。

混合因子分析模型(MoFA): 混合高斯+因子分析

混合t分布(鲁棒混合模型): 混合高斯+t分布

鲁棒子空间模型: t分布+因子分析

混合鲁棒子空间模型: 混合高斯+t分布+因子分析

各种模型加权得到最终密度函数 $Pr(x)$ 。

总结

本章介绍了隐变量的思想, 进而引出五种模型结构。

学习这种模型的主要方法是期望值最大化算法, 该算法只能找到局部最优解。

并说明了期望值最大算法的两个重要步骤: E步和M步。以及每个模型参数更新时的公式。

编程练习

二维高斯模型的可视化:

