



# 大数据可视化

北京理工大学计算机学院 计卫星

2019年1月



北京理工大学计算机学院  
School of Computer Science and Technology, BIT



# 高维数据可视化



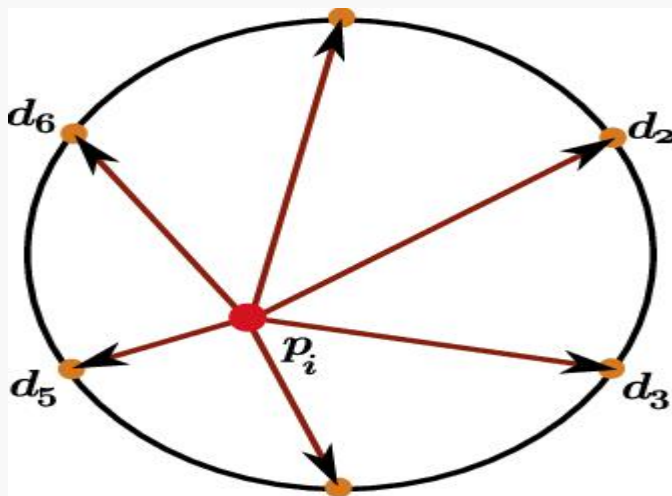
# 高维数据可视化

- 高维数据每个数据样本拥有多个数据属性
- 分析和可视化高维数据是一个极大的挑战
- 主要包括
  - RadViz
  - 平行坐标
  - 降维投影
  - ...



# 高维数据可视化

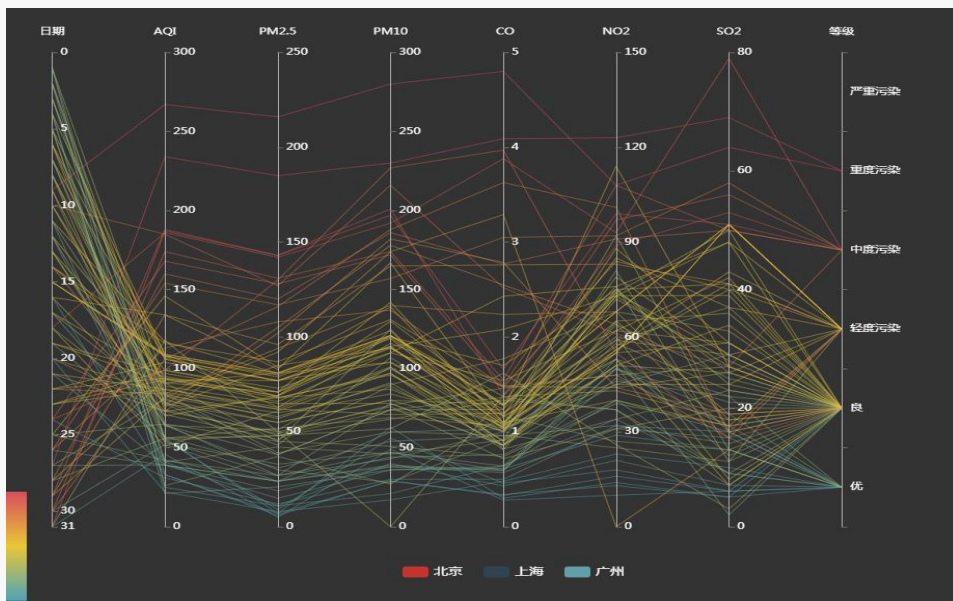
- RadViz





# 高维数据可视化

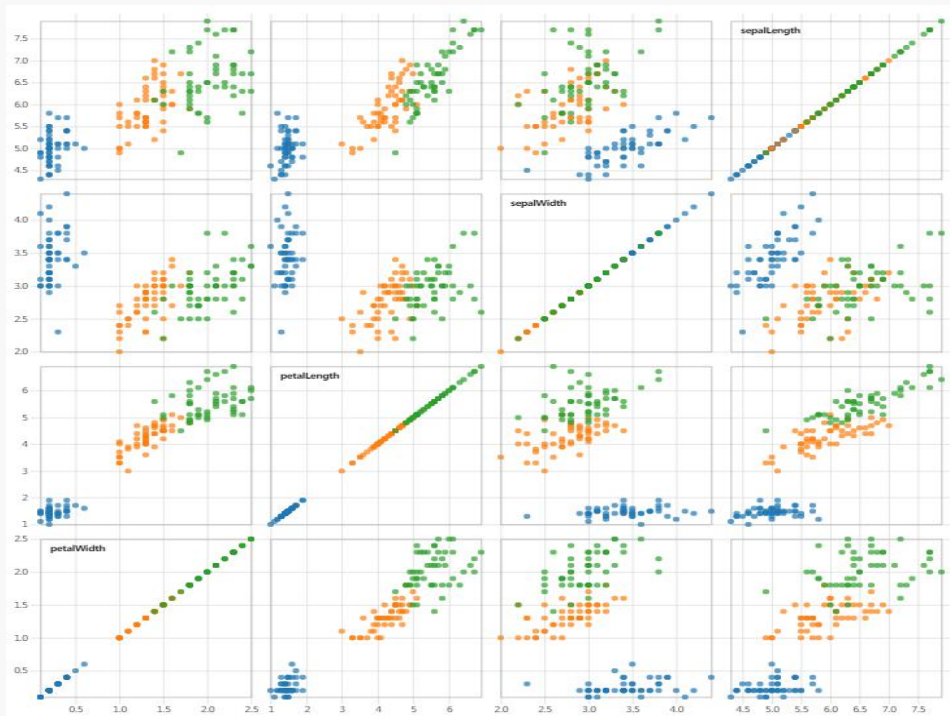
- 平行坐标





# 高维数据可视化

- 散点图矩阵

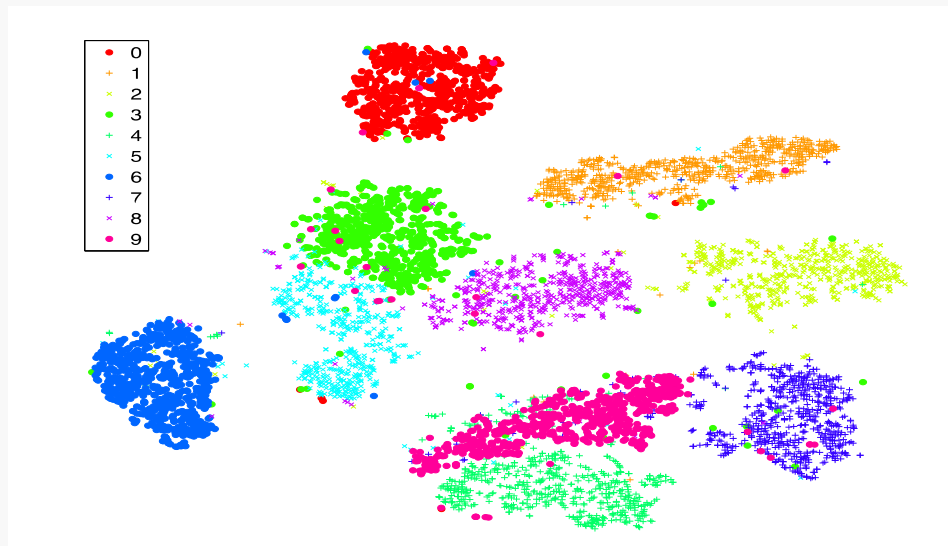




# 高维数据可视化

- 降维投影（手写数字可视化）

8	9	0	1	2	3	4	7	8	9	0	1	2	3	4	5	6	7	8	6
4	2	6	4	7	5	5	4	7	8	9	2	9	3	9	3	8	2	0	5
0	1	8	4	2	6	8	3	5	3	8	0	3	4	1	5	3	0	8	
3	0	6	2	7	1	1	8	1	7	1	3	8	9	7	6	7	4	1	6
7	5	1	7	1	9	8	0	6	9	4	9	3	7	1	9	2	2	5	
3	7	8	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	0
1	2	3	4	5	6	7	8	9	8	1	0	5	5	1	9	0	4	1	9
3	8	4	7	7	8	3	0	6	5	5	3	3	3	9	8	1	4	0	6
1	0	0	6	2	1	1	3	2	8	8	7	8	4	6	0	2	0	3	6
8	7	1	5	9	9	3	2	4	9	4	6	5	3	2	3	5	9	4	1
6	5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
8	9	0	1	2	3	4	5	6	7	8	9	6	4	2	6	4	7	5	5
4	7	8	9	2	9	3	9	3	8	2	0	9	8	0	5	6	0	1	0
4	2	6	5	5	5	4	3	4	1	5	3	0	8	3	0	6	2	7	1
1	8	1	7	1	3	8	5	4	2	0	9	7	6	7	4	1	6	8	4
7	5	1	2	6	7	1	9	8	0	6	9	4	9	9	6	2	3	7	1
9	2	2	5	3	7	8	0	1	2	3	4	5	6	7	8	0	1	2	3
4	5	6	7	8	0	1	2	3	4	5	6	7	8	9	2	1	2	1	3
9	9	8	5	3	7	0	7	7	5	7	9	9	4	7	0	3	4	1	4
4	7	5	8	1	4	8	4	1	8	6	6	4	3	5	7	2	5	9	





# 高维数据可视化

- 当维数低于10时
  - RadViz, 散点图矩阵、平行坐标等
- 当维度在10-100之间时
  - 平行坐标, 降维投影图
- 当100维以上时
  - 降维投影