1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：奇异值和主成分有什么内在关联吗？

讨论后的理解：16.3.2给出了利用数据矩阵奇异值分解方法进行主成分分析的算法，从算法中可以看到对矩阵X’进行截断奇异值分解得到的矩阵V的前k列构成了k个样本主成分。

1. 提出的问题2：例16.1中的两个聚类的含义是什么？

讨论后的理解：由整体成绩y1和文科成绩y2为轴描点，每个点都代表一门学科。该题中一个聚类表示一类学科，例如语文和外语的聚类表示文科，数学和物理的聚类表示理科。学科表示的点所具有的坐标的正负体现出该学科对于坐标轴元素的影响的积极与否。

1. （必填）别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. 问题3：例题中，文科成绩提高都可使 y2 提高，理科成绩提高都可使 y2 降低，为什么不是恰好相反，这是由什么决定的？

自己的理解：这个问题里的现象是样本主成分分析的结论，与原有数据有关，是从表16.1 样本相关矩阵R所体现出的。其实从表中也可以粗略观察出，同类型学科的相关度明显比不同类型学科的相关性要高。

1. 问题4：P315实例中给出了第一、第二主成分代表的含义，但似乎是事后从观察得到的，从实际意义上怎么理解主成分对问题的意义呢？

自己的理解：个人感觉样本主成分分析只是一种聚类所采用的常见方法，而聚类的结果是人为赋予的。因为这是无监督学习，我们事先并不知道有多少个类，这些类的标签又是什么。而主成分分析提供了对于样本的一种聚类方法。

5、问题5：主成分分析得出的结果除了可以用于聚类分析还有什么常见应用吗？

自己的理解：PCA 是一种强大的数据转换技术，可以先应用此技术，再进行进一步的分析工作。这种方法在您遇到高维数据集时非常实用，在文本挖掘、图像分析、生物数据分析、客户偏好分析以及口味分析（协同过滤）之类的领域中尤其如此。

1. （必填）读书计划

1、本周完成的内容章节：第十六章

2、下周计划：第十七章