1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：从单词向量空间到话题向量空间的线性变换是否唯一？

讨论后的理解：单纯从未知数角度来说，对于给定的k，方程数量等于未知数的数量，所以说在非退化的情况下，我认为还是有唯一解的。但出现退化现象以后可能就会有无穷多解了。

1. 提出的问题2：潜在语义分析这个过程会不会出现什么问题？

讨论后的理解：（1）LSA在进行信息提取时，忽略词语的语法信息(甚至是忽略词语在句子中出现顺序)，认为语法结构在文本的语义表达中处于次要的地位。它仍然是一种简单地通过所有词语向量的线性总和来产生文本向量，表示文本的含义的方法。然而，句子的语法结构包含了词语之间的更深层次的语义关联信息，忽略这种语义关联关系影响了LSA对文本内容的把握能力。因此，现在的一些研究开始将LSA模型和语法信息相结合，以提高LSA的性能。（2）LSA处理的对象是可见变量(文本集中出现的词语、文本)，它不能通过计算得到词语的暗喻含义，以及类比推论含义。为解决这些因素对LSA获取知识能力影响，需要在LSA的基础上结合其他的理解模型。

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：怎么选择话题数，即k的值？

自己的理解：个人感觉这一步需要人工或者说根据经验值，或者根据实际应用中的某种指标进行计算。

1. 问题4：为什么通过奇异值分解的方法就能够直接将左奇异矩阵作为话题矩阵，后面两项作为文本在话题空间的表示？这么做有什么根据吗？

自己的理解：感觉奇异值分解只是提供了潜在语义分析的一种特定方法，在多数情况下，奇异值分解的方法是适用的，在某种情况下奇异值分解给出的只是某一个解。

1. 问题5：觉得这一章节所谓单词和话题上的相似某种意义上来说就是与文本其它部分的相对关系相似，可以做这样的理解吗？

自己的理解：我觉得可以这样理解，即单词-话题向量表示了某个单词和同一文本或者其他文本的相似相关程度。

1. （必填）读书计划

1、本周完成的内容章节：17.1 17.2

2、下周计划：第十七章剩余部分