**读书报告**

61518424 王贵涛

1. **问题与解答**

我提出问题：

1. 为什么左矩阵第1列向量 u1 的值均为正？

讨论结果：左奇异向量的第一列表示每一个词的出现频繁程度，虽然不是线性的，但是可以认为是一个大概的描述，比如book是0.15对应文档中出现的2次，investing是0.74对应了文档中出现了9次，rich是0.36对应文档中出现了3次，其次，右奇异向量中一的第一行表示每一篇文档中的出现词的个数的近似，比如说，T6是0.49，出现了5个词，T2是0.22，出现了2个词。

1. 为什么WH与X完全相等很难实现，当m×k+k×n>m×n的时候，仅用解未知数的方法就能得出多组解了？

讨论结果：虽然理论上当做求解方程组确实可以求出很多解，但是保证所有解均非负就很困难了。一般这种矩阵规模都应该不小。

别人提出的问题：

1. 初始的W以及H该如何选择？是不是只能随机尝试？

我的解答：应该是随机生成，然后取多种不同的初始值来选取最小的那一个。每个不同的初值都会收敛到一个局部最小值。

1. 非负矩阵分解是不是需要较多的文本数量才可达到较好的效果，如果文本数量较少是否会导致得到的矩阵不具有很好的代表性？

我的解答：是的，无论什么模型，data的数量都是很重要的，过少的数量将使模型不具有代表性。

1. **下周计划安排**

看完第十八章并参加讨论。

1. **读书收获**

非负矩阵分解

非负矩阵分解也可以用于话题分析。

对单词一文本矩阵进行非负矩阵分解，将其左矩阵作为话题向量空间，将其右矩阵作为文本在话题向量空间的表示。注意通常单词-文本矩阵是非负的。

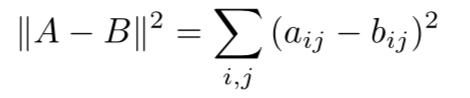
给定一个非负矩阵X≥0，找到两个非负矩阵W≥0和H≥ 0，使得X=WH，即将非负矩阵X分解为两个非负矩阵W和H的乘积的形式，称为非负矩阵分解。因为WH与X完全相等很难实现，所以只要求WH与X近似相等。

假设非负矩阵X是 m x n 矩阵，非负矩阵W和H分别为 m x k 矩阵和 k x n 矩阵。假设k<min(m, n)，即W和H小于原矩阵X，所以非负矩阵分解是对原数据的压缩。

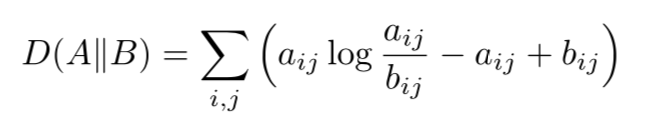
非负矩阵分解可以形式化为最优化问题求解。

首先定义损失函数或代价函数。

第一种损失函数是平方损失：



另一种损失函数是散度（divergence)：



其最优化问题分别为：

