**读书报告**

61518424 王贵涛

1. **问题与解答**

我提出问题：

1. 奇异值分解时使用ATA和AAT得到的结果有什么不同？为什么选取ATA的方法？

讨论结果：两种方法应该是等价的，P281有说U的列向量是AAT的特征向量，所以U和V的地位是一样的，也可以用特征值的方法求出，而书中的方法应该是直接用了构造奇异值分解的那些等式关系，也就避免了再计算一次AAT矩阵并求其特征向量的开销，但是结果应该一样。两者的特征值应该相同，但特征向量不同，一个是用来组成U矩阵一个用来组成V矩阵。

1. 奇异值分解有什么具体意义？

讨论结果：可以对图像进行压缩处理，也可以运用在PCA降维之中。

别人提出的问题：

1. 在定理15.3中，A'是阶段奇异值矩阵，奇异值分解时用的是U，V，但按照书上前面15.1的定义，A'用的U，V应是前k列，而不是A的U，V。这两个有区别吗？

我的解答：在只取前k个奇异值时，后面的奇异值都为零，U和V的后n-k列无论取什么值，乘起来都是零，所以可以只取前k列 。

1. 奇异值分解算法如果直接求解ATA的特征值，会带来哪些效率上的提升？计算的步骤与15.2中的基本方法会有很大差异吗？

我的解答：可以带来计算上效率的提升，ATA的特征值是A的奇异值的平方，而且可以通过ATA直接计算出V矩阵。

1. 矩阵的外积展开形式是否是奇异值分解对矩阵近似的一种直观表现？能否从外积展开式角度直观证明或者说明最小平方损失是如何得出的？

我的解答：矩阵的外积展开式就是矩阵的另一种表达方式，在截断式奇异值分解时对矩阵进行近似。近似时丢弃的是最小的部分奇异值，因此做到了最小平方损失。

1. **下周计划安排**

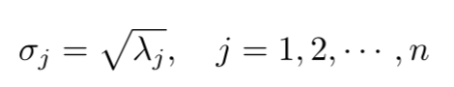
看完第十六章并参加讨论。

1. **读书收获**

矩阵A的奇异值分解可以通过求对称矩阵ATA的特征值和特征向量得到。

ATA的特征向量构成正交矩阵V的列

ATA的特征值λ的平方根为奇异值σ，即



对其由大到小排列作为对角线元素，构成对角矩阵∑

求正奇异值对应的左奇异向量，再求扩充的AT的标准正交基，构成正交矩阵U的列

从而得到A的奇异值分解：



具体的求解步骤：

（1）求ATA的特征值和特征向量

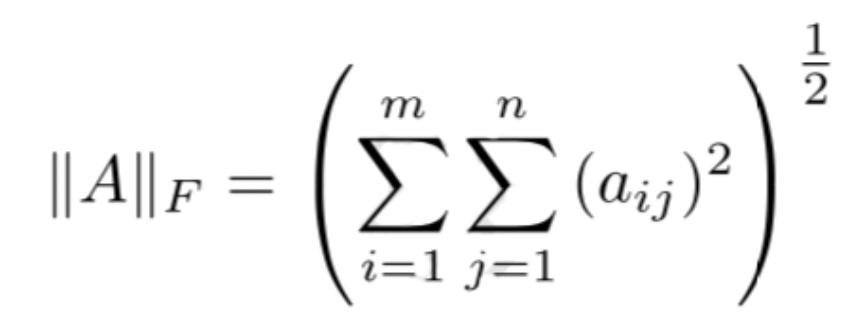
（2）求n阶正交矩阵V

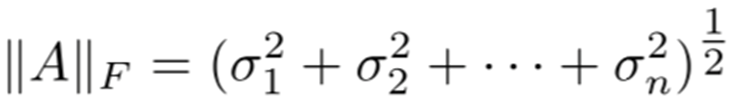
（3）求 m x n 对角矩阵∑

（4）求m阶正交矩阵U

（5）得到奇异值分解

奇异值分解也是一种矩阵近似的方法，这个近似是在弗罗贝尼乌斯范数（Frobenius norm）意义下的近似。矩阵的弗罗贝尼乌斯范数是向量的LZ范数的直接推广，对应着机器学习中的平方损失函数。定义矩阵A(m\*n)的弗罗贝尼乌斯范数为：





奇异值分解是在平方损失弗罗贝尼乌斯范数）意义下对矩阵的最优近似，即数据压缩。

紧奇异值分解是在弗罗贝尼乌斯范数意义下的无损压缩。

截断奇异值分解是有损压缩。

截断奇异值分解得到的矩阵的秩为k，通常远小于原始矩阵的秩r，所以是由低秩矩阵实现了对原始矩阵的压缩。