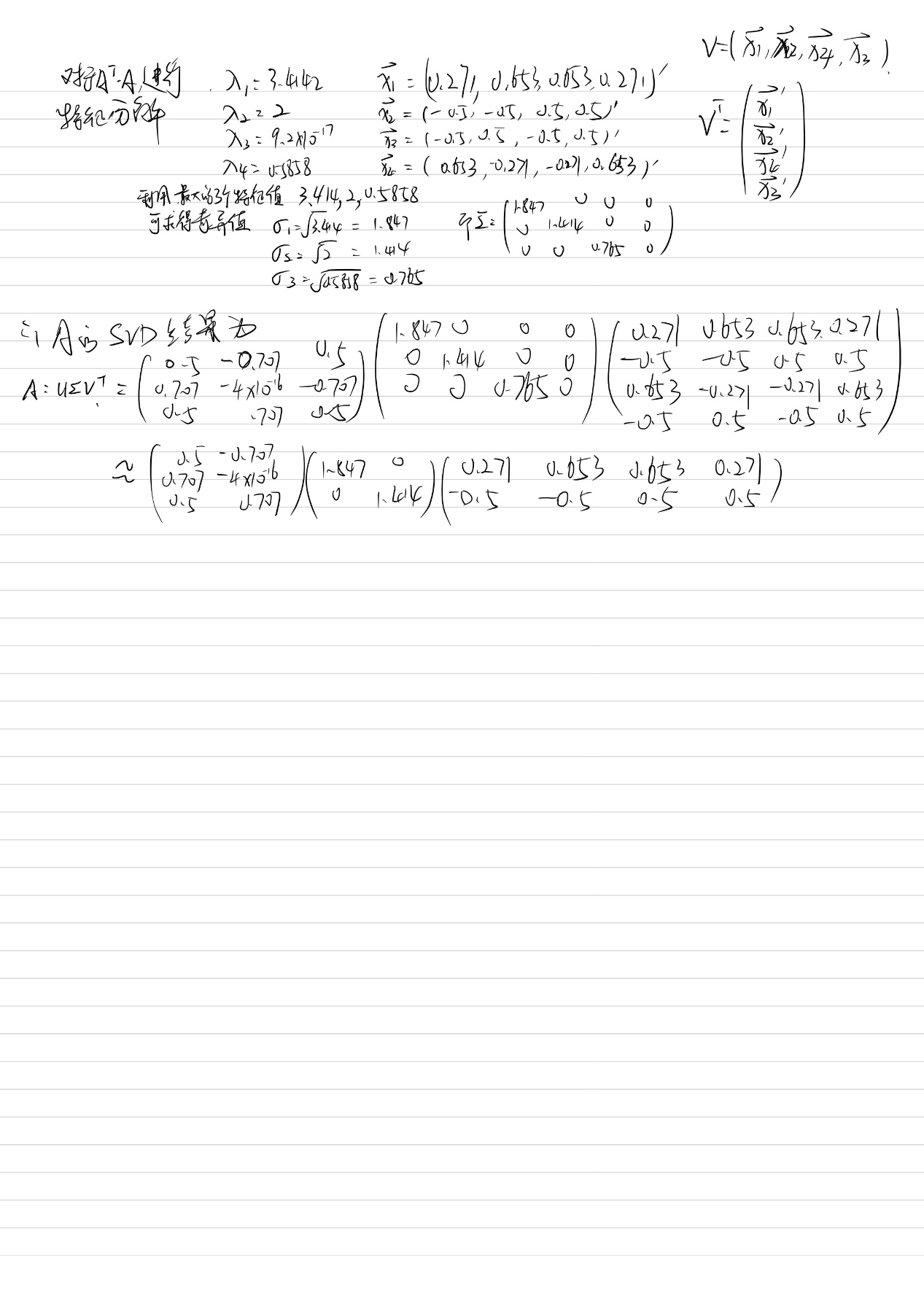
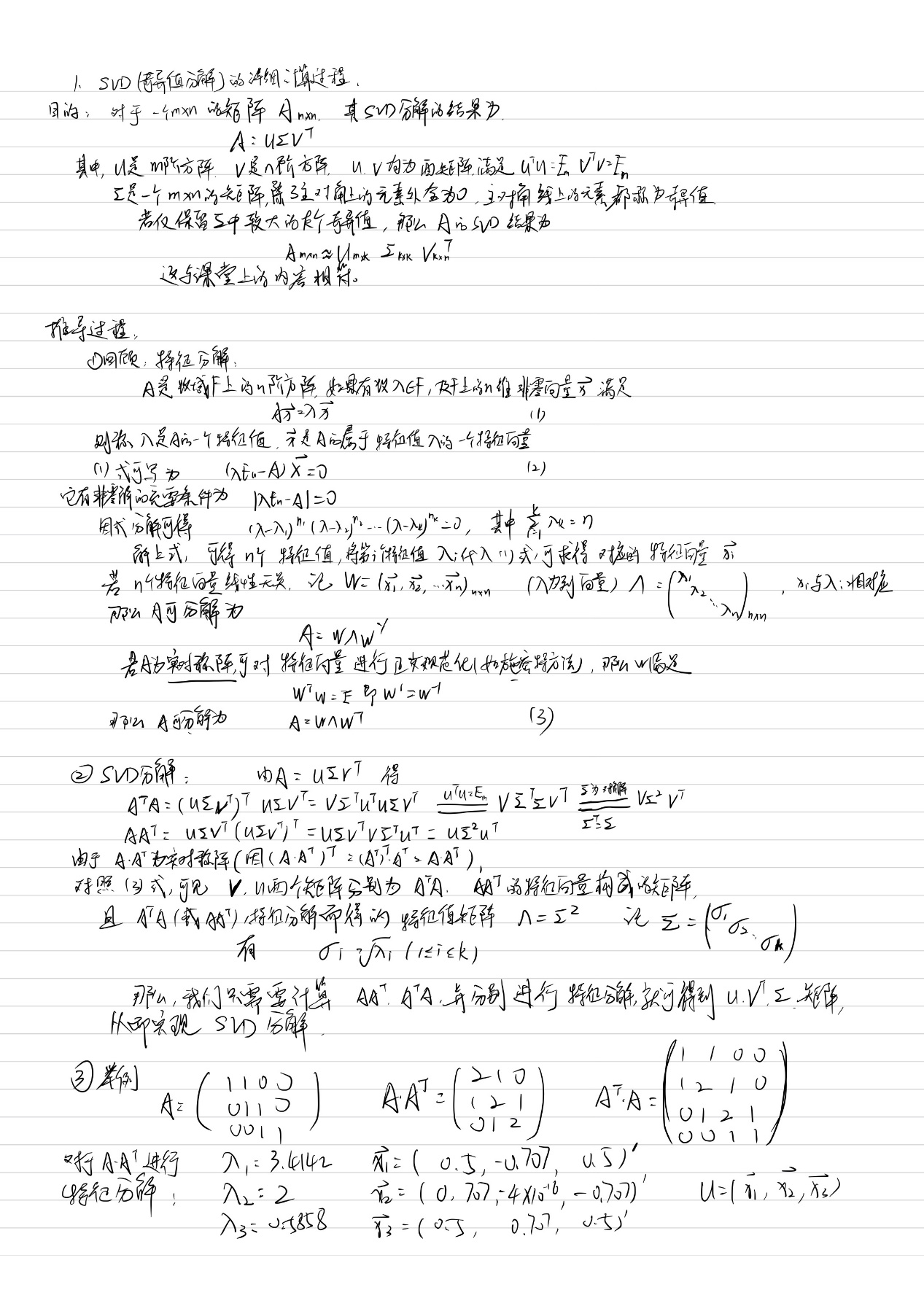
ir-hw03

1. SVD



1. 什么是PCA？PCA与SVD的区别是什么？

PCA（Principal Component Analysis） 是一种常见的数据分析方式，常用于高维数据的降维，可用于提取数据的主要特征分量。给定个 维样本，通过变换 (其中为变换矩阵)，将样本从维降到维，同时最大程度的减少降维带来的信息损失。其目标是：（1）降维后的各个维度之间相互独立，即去除降维之前样本各个维度之间的相关性。（2）最大程度保持降维后的每个维度数据的多样性，即最大化每个维度内的方差

区别：

首先，必须明确PCA和SVD是两个不同的概念。PCA的目的是通过变换 (其中为变换矩阵),实现样本的降维；而SVD是将分解为，二者的直接目的不同。

但是，在标准做法中，PCA 关键在于对协方差矩阵进行特征值分解； SVD 关键在于是对进行特征值分解。两者在计算的过程中有许多相似的地方。如果取，则二者等价。

另外，一方面，SVD也同样常常作为一种降维的方法；另一方面，由于矩阵维度高， 计算量大，特征分解效率较低，而SVD有更高效的求解法，可以避免计算该项。事实上，基于这些好处，PCA可以转化为SVD求解。