Author: 陆星宇 Date: 2021/7/9

第15章 决策树

一、交流讨论

问题1：在奇异值分解中，紧奇异值分解和截断奇异值分解的区别是什么

紧奇异值分解仅保留非零奇异值及其左右奇异向量，属于无损压缩；二而截断奇异值分解保留k个非零奇异值及其左右奇异向量，k 往往小于矩阵阶数，属于有损压缩。

问题2：奇异值分解的几何解释是什么？

矩阵可以被奇异值分解为三个矩阵的乘积，乘以原矩阵相当于连续左乘这三个矩阵，相当于连续做了一个旋转变换，一个伸缩变换，一个旋转变换。

问题3：奇异值为什么能对应矩阵中某一行列的"重要性"？

奇异值相当关于将矩阵中对应于奇异向量所对应分量的部分乘以奇异值，可以近似为零。

二、内容概要

1. 奇异值分解的定义

$$A = U\Sigma V^T$$

其中$A \in R^{m\times n}$，U 是 m 阶正交矩阵，V 是 n 阶正交矩阵，Σ 是由降序排序的非负的对角线元素组成的矩形对角矩阵。

> 若 A 为正交矩阵，则 $AA^T = E$.

A 的对角线元素 σ_i 为 A 的奇异值，U 与 V 的列向量分别称为左奇异向量与右奇异向量。

奇异值分解可以看作是对方阵对角化的推广。

矩阵的奇异值分解是一定存在且不唯一的。

2. 紧奇异值分解与截断奇异值分解

3. 几何解释

矩阵的奇异值分解也可以看作是将其对应的线性变换分解为旋转变换，缩放变换及旋转变换的组合。

4. 主要性质

- $A^TA = V(\Sigma^T\Sigma)V^T$
  $AA^T = U(\Sigma\Sigma^T)U^T$
- 若 $m > n$，则
  $Av_j = \sigma_j u_j,\ \ j = 1, 2, \ldots, n$
  $A^Tu_j = \sigma_j v_j,\ \ j = 1, 2, \ldots, n$
  $A^Tu_j = 0,\ \ j = n+1, n+2, \ldots, m$
- 矩阵 A 的奇异值分解中，奇异值是唯一的，而矩阵 U 和 V 不是唯一的
- 矩阵 A 和 Σ 的秩相等，等于正奇异值的个数(包含重复的奇异值)

- 矩阵 A 的 r 个右奇异值 $v_1, v_2, \ldots, v_r$ 构成 A^T 的值域 $R(A^T)$ 的一组标准正交基
  矩阵 A 的 r 个左奇异值 $u_1, u_2, \ldots, u_r$ 构成 A 的值域 $R(A)$ 的一组标准正交基
  矩阵 A 的 $n-r$ 个右奇异值 $v_{r+1}, v_{v+2}, \ldots, v_n$ 构成 A 的 零空间 $N(A)$ 的一组标准正交基
  矩阵 A 的 $m-r$ 个左奇异值 $u_{r+1}, u_{r+2}, \ldots, u_m$ 构成 A^T 零空间 $N(A^T)$ 的一组标准正交基

5. 奇异值分解的计算

可以通过性质1和性质2求解

6. 奇异值分解与矩阵近似