

# 《矩阵分析与应用》大作业 1

姓名：谷绍伟      学号：202418020428007

1 利用现在已有的 AI 模型，例如 ChatGPT 等，尝试学习一个和矩阵 SVD 相关的内容 (如定义，快速算法以及相关应用等)，让 AI 模型帮你答疑解惑，希望做到从一个概念开始，到最后的熟悉掌握。要求：

- 了解相关定义和一些重要的结果；
- 学习相关理论知识内容；
- 学习相关的应用内容，最好能举出 1-2 例子；

把整个学习的过程导出一个文档，并总结自己整个通过 AI 学习的心得体会，例如：有哪些启发，AI 模型有哪些有效的帮助，有哪些缺点等。

## SVD 的定义

使用字节跳动的豆包大模型作为 AI 模型，学习 SVD 分解的相关定义和实现方法。结果如下：

### 提问

什么是矩阵的 SVD 分解，他的定义是什么？

### 回答

#### 奇异值分解 (SVD) 的定义

- 对于一个  $m \times n$  的矩阵  $A$  (其中  $m$  和  $n$  是正整数)，存在如下分解：

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

- 这里  $U$  是一个  $m \times m$  的正交矩阵，也就是说  $U^T U = I_m$  ( $I_m$  是  $m$  阶单位矩阵)。它的列向量被称为矩阵  $A$  的左奇异向量。
- $V$  是一个  $n \times n$  的正交矩阵，即  $V^T V = I_n$ ，它的列向量被称为矩阵  $A$  的右奇异向量。

- $\Sigma$  是一个  $m \times n$  的对角矩阵, 其对角线上的元素  $\sigma_i$  ( $i = 1, 2, \dots, \min\{m, n\}$ ) 是非负实数, 并且按照从大到小的顺序排列, 这些对角元素被称为矩阵  $A$  的奇异值。
- 例如, 对于一个简单的  $2 \times 2$  矩阵  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , 它的 SVD 分解为  $A = U\Sigma V^T$ , 其中  $U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$ ,  $\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$  (假设  $\sigma_1 \geq \sigma_2 \geq 0$ ),  $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ 。
- 从几何意义上来说, 矩阵  $A$  可以看作是从  $\mathbb{R}^n$  空间到  $\mathbb{R}^m$  空间的一个线性变换。SVD 分解将这个线性变换分解成了三个简单的线性变换的组合: 一个是  $V^T$  在  $\mathbb{R}^n$  空间中的旋转变换 (因为  $V$  是正交矩阵), 一个是  $\Sigma$  在  $\mathbb{R}^n$  空间中的伸缩变换 (沿着坐标轴方向进行伸缩), 最后一个是  $U$  在  $\mathbb{R}^m$  空间中的旋转变换。这有助于我们更好地理解矩阵  $A$  所代表的线性变换的本质。

询问 SVD 分解的计算流程, 结果如下:

### 提问

矩阵的 SVD 分解怎么计算, 请你给出流程并举一个例子。

### 回答

#### 1. 计算矩阵 SVD 分解的流程

- 步骤一: 计算  $A^T A$  和  $AA^T$ 
  - 对于给定的  $m \times n$  矩阵  $A$ , 先计算  $A^T A$  (这是一个  $n \times n$  矩阵) 和  $AA^T$  (这是一个  $m \times m$  矩阵)。
- 步骤二: 求特征值和特征向量
  - 计算  $A^T A$  的特征值  $\lambda_i$  和对应的单位特征向量  $v_i$ 。 $A^T A$  是对称矩阵, 所以它的特征向量是正交的。这些特征向量将构成  $V$  的列向量。
  - 同样地, 计算  $AA^T$  的特征值和对应的单位特征向量  $u_i$ , 这些特征向量将构成  $U$  的列向量。
- 步骤三: 计算奇异值
  - 矩阵  $A$  的奇异值  $\sigma_i = \sqrt{\lambda_i}$ , 其中  $\lambda_i$  是  $A^T A$  (或  $AA^T$ ) 的非负特征值。 $\Sigma$  是一个  $m \times n$  的对角矩阵, 其对角元素为  $\sigma_i$ , 其他元素为 0。

#### 2. 示例

- 设矩阵  $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
- 步骤一：
  - 计算  $A^T A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ ,  $AA^T = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ 。
- 步骤二：
  - 对于  $A^T A$ , 求其特征值。特征方程为  $\det(A^T A - \lambda I) = 0$ , 即  $\begin{vmatrix} 2-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} = 0$ 。
  - 展开行列式得到  $(2-\lambda)^2 - 4 = 0$ , 即  $\lambda^2 - 4\lambda = 0$ , 解得  $\lambda_1 = 4, \lambda_2 = 0$ 。
  - 当  $\lambda_1 = 4$  时, 代入  $(A^T A - \lambda_1 I)x = 0$ , 即  $\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , 得到特征向量为  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ , 单位化后为  $v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 。
  - 当  $\lambda_2 = 0$  时, 代入  $(A^T A - \lambda_2 I)x = 0$ , 即  $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , 得到特征向量为  $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ , 单位化后为  $v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ 。
  - 所以  $V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ 。
  - 对于  $AA^T$  (因为和  $A^T A$  相同), 特征向量也相同,  $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ 。
- 步骤三：
  - 奇异值  $\sigma_1 = \sqrt{\lambda_1} = 2, \sigma_2 = \sqrt{\lambda_2} = 0$ , 则  $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ 。
  - 所以  $A$  的 SVD 分解为  $A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ 。

经过验证, 上述的计算过程是正确的。进一步询问矩阵奇异值分解是否具有唯一性:

### 提问

矩阵的 SVD 分解是唯一的吗?

## 回答

### 奇异值的唯一性

- 矩阵  $A$  的奇异值是唯一的。在  $A = U\Sigma V^T$  的奇异值分解 (SVD) 中,  $\Sigma$  是一个  $m \times n$  (设  $A$  是  $m \times n$  矩阵) 的对角矩阵, 其对角线上的元素  $\sigma_i$  (奇异值) 是  $A^T A$  (当  $m \geq n$ ) 或者  $AA^T$  (当  $m < n$ ) 的非负特征值的平方根, 并且通常按照从大到小的顺序排列, 即  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ , 其中  $r$  是矩阵  $A$  的秩。
- 例如, 对于矩阵  $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ , 计算  $A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ , 其特征值为  $\lambda_1 = 4$ ,  $\lambda_2 = 0$ , 那么奇异值  $\sigma_1 = \sqrt{\lambda_1} = 2$ ,  $\sigma_2 = \sqrt{\lambda_2} = 0$ , 这些奇异值是由矩阵  $A$  本身的性质决定的, 不会因为分解方式的不同而改变 (只要按照上述从大到小的顺序排列奇异值)。

### 奇异向量的非唯一性 (存在相同奇异值的情况)

- 当有重复的奇异值时, 左右奇异向量不唯一。假设  $\sigma_i = \sigma_j$  ( $i \neq j$ ), 对于左奇异向量  $u_i$  和  $u_j$  ( $U$  的列向量) 以及右奇异向量  $v_i$  和  $v_j$  ( $V$  的列向量), 可以通过对它们进行线性组合 (只要保持  $U$  和  $V$  是正交矩阵) 来得到不同的分解。
- 例如, 对于单位矩阵  $I_n$ , 它的奇异值分解为  $I_n = I_n \times I_n \times I_n$  ( $U = V = \Sigma = I_n$ ), 但如果我们对  $U$  和  $V$  的列向量进行任意正交变换 (例如, 旋转这些向量), 仍然可以得到  $I_n$  的有效 SVD 分解。
- 不过, 如果所有奇异值都是不同的, 并且规定  $U$  和  $V$  是列正交矩阵, 且  $\Sigma$  对角线上的奇异值按降序排列, 那么在相差一个对角元素为  $\pm 1$  的对角矩阵的意义下, SVD 分解是唯一的。即如果  $A = U_1 \Sigma_1 V_1^T$  和  $A = U_2 \Sigma_2 V_2^T$  是满秩矩阵  $A$  的两个 SVD 分解, 且奇异值都不同, 那么存在对角矩阵  $D$  (对角元素为  $\pm 1$ ), 使得  $U_2 = U_1 D$ ,  $V_2 = V_1 D$ , 并且  $\Sigma_1 = \Sigma_2$ 。

根据回答, 矩阵的奇异值是唯一的, 但  $U$ 、 $V$  不一定唯一, 这是符合定理的。再询问奇异值分解的意义:

## 提问

奇异值分解在几何上有什么意义?

### 从线性变换角度理解

- 矩阵可以看作是一种线性变换。对于一个  $m \times n$  的矩阵  $A$ ，它将  $n$  维向量空间  $\mathbb{R}^n$  中的向量变换到  $m$  维向量空间  $\mathbb{R}^m$  中。奇异值分解  $A = U\Sigma V^T$  把这个线性变换分解为三个基本的线性变换。
- 首先是  $V^T$ ，它是一个  $n \times n$  的正交矩阵。从几何上看， $V^T$  代表了在  $n$  维空间  $\mathbb{R}^n$  中的一个旋转（或者反射）操作。因为正交矩阵的列向量是单位正交向量组，所以它不会改变向量的长度和夹角的大小，只是改变了向量的方向，就像在空间中把坐标轴进行旋转一样。
- 接着是  $\Sigma$ ，它是一个  $m \times n$  的对角矩阵。 $\Sigma$  代表了在  $n$  维空间（如果  $m \geq n$ ）中的拉伸（或者压缩）操作。对角线上的元素  $\sigma_i$  就是在对应的坐标轴方向上的拉伸（压缩）比例。例如，如果  $\sigma_1 = 2$ ， $\sigma_2 = 1$ ，那么在第一个坐标轴方向上向量的长度会变为原来的 2 倍，在第二个坐标轴方向上长度不变。如果某个奇异值为 0，则表示在对应的方向上向量被压缩到了 0 长度，也就是这个维度被“消除”了。
- 最后是  $U$ ，它是一个  $m \times m$  的正交矩阵。 $U$  在  $m$  维空间  $\mathbb{R}^m$  中也代表一个旋转（或者反射）操作，将经过拉伸后的向量旋转到最终的位置。
- 例如，考虑一个  $2 \times 2$  矩阵  $A$  对平面上单位圆  $x^2 + y^2 = 1$  的作用。通过  $A$  的 SVD 分解，我们可以看到单位圆先在  $\mathbb{R}^2$  空间中被  $V^T$  旋转，然后被  $\Sigma$  拉伸成一个椭圆，最后被  $U$  旋转到最终的位置。

### 数据分布和维度的几何意义

- 在数据处理方面，假设我们有一组数据点在  $n$  维空间中构成一个数据云。奇异值分解可以帮助我们理解数据的分布情况。
- 较大的奇异值对应的方向（由  $V$  的列向量确定）表示数据在这些方向上的变化较大，也就是数据分布比较“分散”。这些方向构成了数据的主要特征方向。而较小的奇异值对应的方向表示数据在这些方向上变化较小，可能是噪声或者不太重要的信息。
- 从维度的角度看，通过奇异值分解可以实现降维。如果我们只保留前  $k$  个较大的奇异值（ $k < n$ ）以及对应的左右奇异向量，那么实际上是将数据从  $n$  维空间投影到一个  $k$  维子空间中。这个投影过程在几何上就是将数据点沿着与被舍去的奇异值对应的方向进行“压缩”，使得数据在低维空间中得到表示，并且尽可能保留了数据在主要特征方向上的信息。例如，在三维空间中有一个近似平面的数据分布，通过奇异值分解可以发现其中一个奇异值很

小，然后将数据投影到由另外两个较大奇异值确定的二维平面上，实现降维并保留主要信息。

## SVD 分解的应用

### 提问

SVD 分解有什么应用价值？

### 回答

奇异值分解（SVD）具有多方面的应用价值，包括：

#### 数据压缩：

- **图像压缩：**图像可以表示为像素值的矩阵，通过 SVD 分解后，可以提取出图像的主要特征。只保留较大的奇异值以及对应的左右奇异向量，舍去较小的奇异值和相关向量，能在保持图像主要信息的基础上实现压缩。例如，对于灰度图像，经过 SVD 分解后，根据前几个较大的奇异值重构图像，人眼可能无法分辨出与原始图像的明显差异，但数据量却大大减少。
- **音频压缩：**音频信号也可以转换为矩阵形式进行处理。SVD 可以提取音频数据的主要特征，去除一些次要信息，从而实现音频的压缩存储和传输，在不影响听觉效果的前提下降低数据量。
- **文本数据压缩：**在文本处理中，文档-词项矩阵可以通过 SVD 进行分解，提取出文本的主要主题和特征，去除一些噪声和不太重要的信息，实现文本数据的压缩和简化表示。

#### 降维与特征提取：

- **机器学习与数据挖掘：**在高维数据处理中，SVD 可用于降维。将高维数据映射到低维空间，提取出最具代表性的特征，降低计算复杂度，同时避免过拟合问题。例如，在人脸识别中，原始的人脸图像数据维度很高，通过 SVD 提取出关键的特征向量（即特征脸），可以有效地对人脸进行识别和分类。
- **推荐系统：**推荐系统中，用户-物品评分矩阵是稀疏的。SVD 可以对这个矩阵进行分解，得到用户和物品的潜在特征向量，从而预测用户对未评分物品的评分，实现个性化的推荐。像 Netflix 等公司就广泛应用 SVD 等技术来改进其推荐系统。

#### 信号处理：

- **信号去噪**：在信号处理中，信号往往会受到噪声的干扰。通过对信号进行 SVD 分解，可以将信号分解为不同的成分，其中较小的奇异值对应的成分可能是噪声。去除这些噪声成分，保留主要的信号成分，就可以实现信号的去噪。
- **信号重构**：对于部分缺失或损坏的信号，SVD 可以根据已知的部分信号数据进行分解，然后利用分解得到的奇异值和奇异向量来重构原始信号，恢复缺失或损坏的部分。

### 流体力学与工程领域

- **流场分析**：在流体力学中，SVD 可用于流场模态分解，提取流场的主要特征和模态，帮助理解流场的结构和动态特性。例如，在圆柱绕流等问题中，通过对流场数据进行 SVD 分解，可以得到流场的主要模态，为研究流场的稳定性和流动特性提供重要信息。
- **结构动力学**：在工程结构的动力学分析中，SVD 可以用于对结构的振动响应数据进行分解，提取结构的主要振动模态和特征，帮助工程师了解结构的动态特性，为结构的设计和优化提供依据。

### 数值分析与求解线性方程组：

- **求解线性最小二乘问题**：对于线性最小二乘问题，即找到一个向量使得线性方程组的误差平方和最小，SVD 可以提供一种有效的求解方法。通过对系数矩阵进行 SVD 分解，可以将线性最小二乘问题转化为更容易求解的形式，得到更稳定和准确的解。
- **矩阵求逆的近似计算**：对于一些接近奇异或大规模的矩阵，直接求逆可能非常困难或不准确。利用 SVD 分解，可以得到矩阵的 Moore-Penrose 广义逆的近似，从而为解决相关的矩阵方程提供一种可行的途径。

### 生物学与化学领域

- **蛋白质结构分析**：在生物信息学中，蛋白质的结构可以用矩阵来表示。SVD 可以用于分析蛋白质结构的特征和变化，帮助研究人员理解蛋白质的功能和相互作用。
- **化学光谱分析**：在化学分析中，光谱数据可以通过 SVD 进行处理，提取出光谱的主要特征和成分，用于物质的定性和定量分析。

## 提问

你能具体讲讲在数据压缩上的应用吗？可以实现一段简单的代码来演示压缩的效果

## 回答

### SVD 在数据压缩中的原理

- 对于一个矩阵  $A$ （假设它代表图像、音频或其他数据），进行奇异值分解  $A = U\Sigma V^T$ 。奇异值  $\sigma_i$  ( $\Sigma$  对角线上的元素) 按照从大到小的顺序排列。通常，前几个较大的奇异值包含了矩阵  $A$  的大部分信息。通过保留前  $k$  个较大的奇异值（以及对应的  $U$  和  $V$  中的部分向量），并将其余较小的奇异值设为 0，可以得到一个近似矩阵  $A_k$ ，从而实现数据压缩。

### 以图像数据压缩为例的 Python 代码演示

- 首先，需要安装 ‘numpy’ 和 ‘pillow’ 库。‘numpy’ 用于进行数值计算，包括奇异值分解，‘pillow’ 用于图像的读取和保存。
- 代码如下：

```
1 import numpy as np
2 from PIL import Image
3
4 # 读取图像并转换为灰度图
5 image = Image.open("logo2.jpg").convert("L")
6 image.save("gray_image.jpg")
7 image_array = np.array(image)
8
9 # 进行奇异值分解
10 U, s, V = np.linalg.svd(image_array, full_matrices=False)
11
12 # 选择保留的奇异值数量
13 k = 50
14 # 构建近似的奇异值矩阵
15 Sigma = np.diag(s[:k])
16 # 重建近似图像
17 approx_image_array = U[:, :k] @ Sigma @ V[:, :k]
18 approx_image_array = approx_image_array.astype(np.uint8)
19
20 # 保存近似图像
21 approx_image = Image.fromarray(approx_image_array)
22 approx_image.save("approx_image_k=50.jpg")
```

- 在上述代码中：
  - 首先读取了一张图像并将其转换为灰度图像，存储为 `image_array`（一个二维的 numpy 数组）。



- 然后对 `image_array` 进行奇异值分解，得到  $U$ 、 $s$ （奇异值向量）和  $V$ 。
  - 选择保留前  $k = 50$  个奇异值，构建新的近似奇异值矩阵  $\Sigma$ ，并通过  $U$ 、 $\Sigma$  和  $V$  的部分向量重建近似图像 `approx_image_array`。
  - 最后将重建的近似图像保存为 `approx_image.jpg`。
- 请注意：
    - 上述代码中的 `your_image.jpg` 需要替换为你自己的图像文件名。
    - 根据图像的内容和复杂程度，你可能需要调整  $k$  的值来获得更好的压缩效果和图像质量平衡。 $k$  值越小，压缩率越高，但图像质量可能越差； $k$  值越大，图像质量越好，但压缩效果可能不明显。

利用实现的代码进行测试，选取输入大小为  $1773 \times 1182$  的灰度图像，分别选择保留前 500 个、100 个、50 个奇异值，进行图像压缩和重建，结果如图 1 所示。



Figure 1: 保留不同奇异值的压缩效果

代码实现了对图像的压缩，程序计算得到的奇异值个数为 1182 个，根据实际效果观察，压缩掉接近一半的特征值时，图像还原信息影响较小，可见奇异值分解是一种有效的图像压缩方法。

# 总结

与大模型对话过程见连接<https://www.doubao.com/thread/w9735347f1f75367b>

## AI 学习过程的启发

- **信息获取便捷。**使用 AI 可以方便地获取信息，特别目前部分大模型能够联网搜索，有较好的信息搜集能力。在学习 SVD 分解时，可以快速获取定义、应用等信息，减少了学习成本。
- **有一定发散性。**AI 具有一定发散性，能根据提问整合信息并回答，在关于图像压缩的应用中，既能清晰地解释应用，又能提供简单的验证代码，有多学科的能力。

## AI 学习中的不足

- **缺乏引导。**尽管 AI 能够提供大量的知识和信息，但在学习过程中，我发现它有时缺乏对知识的系统性引导。学习内容可能会显得有些零散，缺乏一个从基础到深入、逐步递进的完整学习路径规划。这就需要我自己在获取的信息中进行梳理和整合，构建一个适合自己的学习框架，对于初学者来说可能会有一定的难度。
- **数学计算错误较多。**AI 在概念性和文字性任务上表现较好，但在数学计算上性能较差，让模型举例 SVD 分解的计算时， $2 \times 2$  的矩阵计算正确，但对  $2 \times 3$  的一个例子就计算错误，因此在使用大模型学习时，还需要仔细检查计算结果。
- **逻辑性较差。**在处理证明相关的问题或者需要长序列思考的问题上表现较差，在 SVD 分解学习中，豆包大模型无法给出证明。