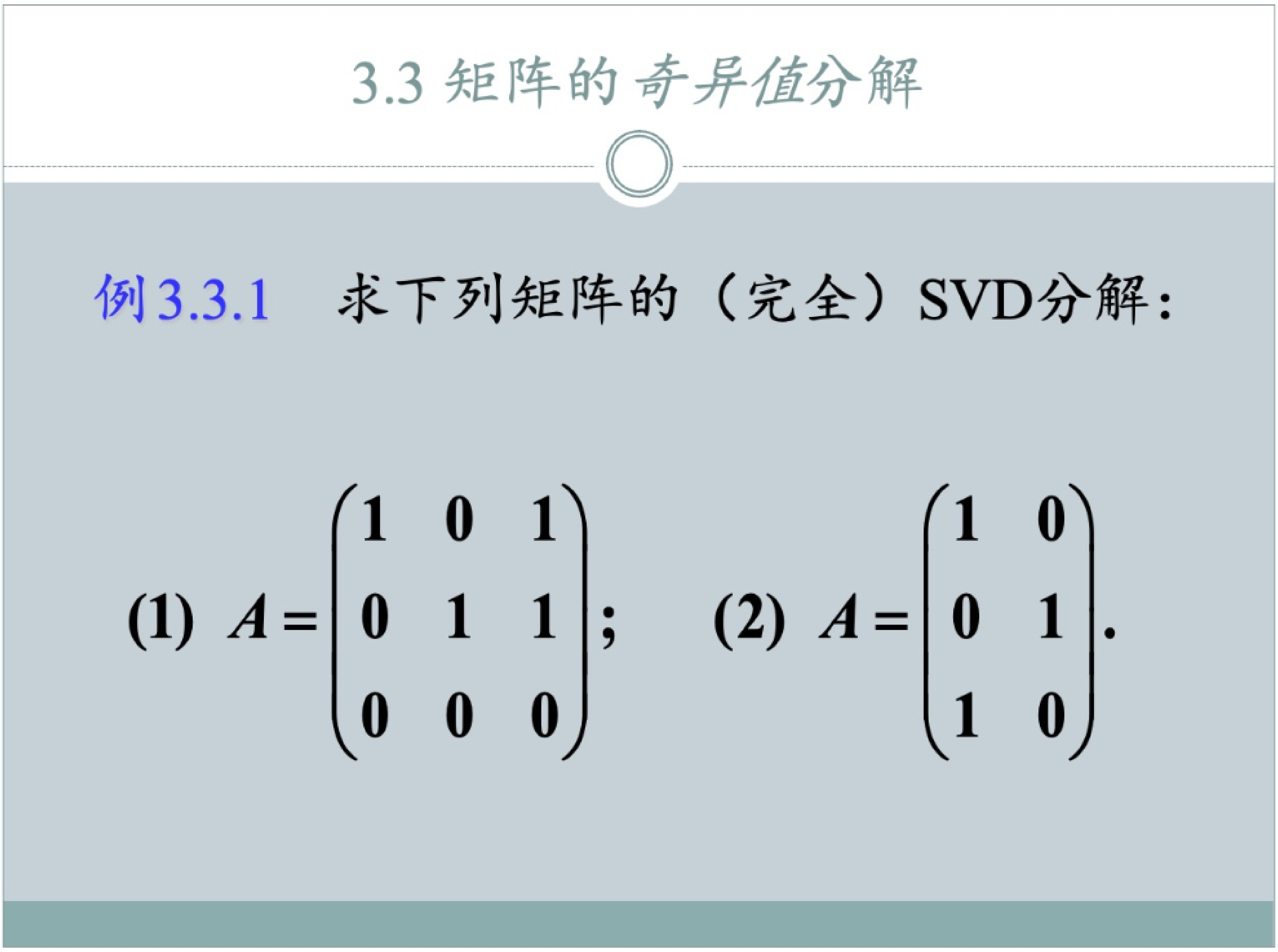
题型：矩阵的完全和约化SVD分解



SVD 分解流程 ($A = U\Sigma V^T$)

1. **求 V**: 计算 A^TA，求其特征值 λ_i 和对应的**单位**特征向量 v_i。将这些 v_i 作为列，组成正交矩阵 **V**。
2. **求 Σ**: 将特征值开根号得到奇异值 $\sigma_i = \sqrt{\lambda_i}$，并**从大到小排序**。将它们放在对角矩阵 **Σ** 的对角线上。
3. **求 U**:
   - 对于**非零**奇异值 $\sigma_i > 0$，直接计算 $u_i = \frac{1}{\sigma_i}Av_i$。
   - 如果 U 的列数不够（因为存在零奇异值），则通过求解 AA^T 的零空间（即特征值为0的单位特征向量）来补全剩余的 u_i。
   - 将所有 u_i 作为列，组成正交矩阵 **U**。

关于约化SVD

- **约化SVD分解**：直接**丢弃**所有零奇异值，以及 U 和 V 中与它们对应的向量。
  - 即：只保留 **r** 个非零奇异值，以及 U 和 V 的前 **r** 列。

例 3.3.1: 求下列矩阵的 (完全) SVD 分解。
第一问 (1): 对以下矩阵 A 进行奇异值分解 (SVD)。

A =

$$\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$$

我们将按照要求，计算该矩阵的 **约化SVD (Reduced SVD)** 和 **完全SVD (Full SVD)**。

解题步骤

SVD 分解的形式为 $A = U\Sigma V^T$。

第1步：计算 A^TA

$A^T =$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

$A^TA =$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

第2步：求 A^TA 的特征值和特征向量，以构建 Σ 和 V

计算 A^TA 的特征值 (λ)：
$\det(A^TA - \lambda I) = 0$

$$\begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (1-\lambda)((1-\lambda)(2-\lambda)-1) + 1(0-(1-\lambda))$$

$$= (1-\lambda)(\lambda^2 - 3\lambda + 1) - (1-\lambda)$$

$$= (1-\lambda)(\lambda^2 - 3\lambda + 1 - 1)$$

$$= (1-\lambda)(\lambda^2 - 3\lambda) = (1-\lambda)\lambda(\lambda - 3) = 0$$

所以特征值为 $\lambda_1 = 3, \lambda_2 = 1, \lambda_3 = 0$。

构建 Σ (奇异值矩阵)：
奇异值 $\sigma_i = \sqrt{\lambda_i}$。按降序排列：
$\sigma_1 = \sqrt{3}$
$\sigma_2 = \sqrt{1} = 1$
$\sigma_3 = \sqrt{0} = 0$

矩阵的秩 r = 2 (非零奇异值的数量)。

构建 V (右奇异向量矩阵)：
求解对应特征值的单位特征向量。

- **对于 $\lambda_1 = 3$:** $(A^TA - 3I)x = 0$

$$\begin{pmatrix} -2 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  得到 z = 2x, z = 2y，所以 x = y。取 x=1, 则 y=1, z=2。
  特征向量为 $(1, 1, 2)^T$。单位化后：
  $\mathbf{v}_1 = (1/\sqrt{6}, 1/\sqrt{6}, 2/\sqrt{6})^T$

- **对于 $\lambda_2 = 1$:** $(A^TA - 1I)x = 0$

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  得到 z = 0, x + y = 0。取 x=1, 则 y=-1, z=0。
  特征向量为 $(1, -1, 0)^T$。单位化后：
  $\mathbf{v}_2 = (1/\sqrt{2}, -1/\sqrt{2}, 0)^T$

- **对于 $\lambda_3 = 0$:** $(A^TA - 0I)x = 0$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  得到 x = -z, y = -z。取 z=1, 则 x=-1, y=-1。
  特征向量为 $(-1, -1, 1)^T$。单位化后：
  $\mathbf{v}_3 = (-1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3})^T$

3.3 矩阵的奇异值分解

解： （1）$A^TA=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}, AA^T=\begin{pmatrix}2&1&0\\1&2&0\\0&0&0\end{pmatrix}$

矩阵 A^TA 的特征值为 3,1,0 ，对应的特征向量为 $\beta_1=(1,1,2)^T,\beta_2=(1,-1,0)^T,\beta_3=(1,1,-1)^T$

从而

$$V=\left(\begin{array}{cc:c}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{1}{\sqrt6}&\frac{-1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&\frac{-1}{\sqrt3}\end{array}\right)\equiv(V_1,V_2),\quad \Sigma=\begin{pmatrix}\sqrt3&0\\0&1\end{pmatrix}$$

第3步：构建 U (左奇异向量矩阵)

利用公式 $\mathbf{u}_i = (1/\sigma_i)A\mathbf{v}_i$ 计算与非零奇异值对应的 u_i。

- **对于 u_1 (对应 $\sigma_1 = \sqrt3$):**

$$\mathbf{u}_1=\frac{1}{\sqrt3}A\mathbf{v}_1=\frac{1}{\sqrt3}\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}1/\sqrt6\\1/\sqrt6\\2/\sqrt6\end{pmatrix}=\frac{1}{\sqrt{18}}\begin{pmatrix}3\\3\\0\end{pmatrix}=\frac{1}{3\sqrt2}\begin{pmatrix}3\\3\\0\end{pmatrix}=\begin{pmatrix}1/\sqrt2\\1/\sqrt2\\0\end{pmatrix}$$

- **对于 u_2 (对应 $\sigma_2 = 1$):**

$$\mathbf{u}_2=\frac{1}{1}A\mathbf{v}_2=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}\begin{pmatrix}1/\sqrt2\\-1/\sqrt2\\0\end{pmatrix}=\begin{pmatrix}1/\sqrt2\\-1/\sqrt2\\0\end{pmatrix}$$

至此，我们已经可以写出约化SVD。对于完全SVD，我们需要找到 $\mathbf{u}_3$，它需要与 u_1 和 u_2 正交。
我们可以通过 Gram-Schmidt 方法或直接观察，找到一个与 u_1 和 u_2 正交的单位向量。
u_1 和 u_2 都在 xy 平面，一个与它们都正交的向量是 z 轴方向的向量。
所以，$\mathbf{u}_3 = (0, 0, 1)^T$。

3.3 矩阵的奇异值分解

计算 $U_1 = A V_1 \Sigma^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}$

解 $AA^T y = \theta$ ，得其基础解系为 $\beta_3 = (0,0,1)^T$

从而

$$U_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad U = (U_1, U_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3.3 矩阵的奇异值分解

因此所求（完全）SVD为

$$A = U \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T$$

由于是方阵，约化SVD同上面的结果一样。

分解结果

1. 约化 SVD (Reduced SVD)

使用秩 r=2 对应的奇异值和奇异向量。
$A = U_r\Sigma_r V_r^T$

U_r (m x r) =

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \end{pmatrix}$$

Σ_r (r x r) =

$$\begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix}$$

V_r (n x r) =

$$\begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \end{pmatrix}$$

所以约化 SVD 分解为：

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}$$

2. 完全 SVD (Full SVD)

使用完整的 U, Σ, V 矩阵。
$A = U\Sigma V^T$

U (m x m) =

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Σ (m x n) =

$$\begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

V (n x n) =

$$\begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}$$

所以完全 SVD 分解为：

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$$

为什么约化SVD分解要逆序排放奇异值

层面一：为了方便进行约化分解

将奇异值从大到小排序，使得所有非零的奇异值都集中在矩阵 Σ 的左上角，而所有的零奇异值（如果存在的话）都被推到了右下角。

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & \Large 0 & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & \Large 0 & & & 0 & \\ & & & & & \ddots \end{pmatrix}$$

这个约定俗成的排序带来了巨大的便利：

- **清晰的界限**：我们能立刻知道矩阵的秩 (rank) 就是非零奇异值的个数 `r`。
- **简单的切片操作**：要从完全SVD得到约化SVD，我们不需要去搜索哪些奇异值是零。我们只需要简单地取矩阵的前 `r` 部分即可：
  - 取 U 的前 `r` 列。
  - 取 Σ 的左上角 `r x r` 块。
  - 取 V^T 的前 `r` 行。

这使得算法实现起来非常干净和高效。**所以，这是一个重要的程序性原因。**

层面二：更深刻的理由 —— 信息的层次与矩阵近似