Numerical Linear Algebra

Assignment 1

Abdukhomid Nurmatov

KAUST

19-09-2023

Contents

1

Calculate the full and reduced singular value decompositions of the matrix

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

- Show all the steps and describe them.
- Add any comments you think are relevant to explain the results you get (try to be critical).

Solution:

(a) First let's revise some theory. We know that if we're given some matrix $A \in \mathbb{C}^{m\times n}$, then SVD of A will be a factorization $A = U\Sigma V^*$, where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices, and $\Sigma \in \mathbb{C}^{m\times n}$ is diagonal matrix with nonnegative entries which are located in nonincreasing order ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$), where $p = \min(m, n)$. As U, V are unitary which means that $U^*U = UU^* = I$, $V^*V = VV^* = I$, then let's calculate A^*A and AA^*:

$$A^*A = (U\Sigma V^*)^* U\Sigma V^* = V\Sigma^*U^*U\Sigma V^* = V\Sigma^*\Sigma V^*, \text{ or } A^*AV = V\Sigma^*\Sigma \tag{1.1}$$

$$AA^* = U\Sigma V^* (U\Sigma V^*)^* = U\Sigma V^*V\Sigma^*U^* = U\Sigma\Sigma^*U^*, \text{ or } AA^*U = U\Sigma\Sigma^* \tag{1.2}$$

In (1.1) $\Sigma^*\Sigma$ is a $n \times n$ diagonal matrix, with entries along the diagonal that are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$, and $n - p$ additional zeros if $n > p$, and in (1.2) $\Sigma\Sigma^*$ is a $m \times m$ diagonal matrix, with entries along the diagonal that are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$, and $m - p$ additional zeros if $m > p$.

Let's find matrix Σ:

In our problem we have $m = 2$ and $n = 3$, so $p = \min(2, 3) = 2$. Then, from (1.2) we have that $AA^*u = \sigma^2 u$, where $u \neq \vec{0}$ (because $u \in U$), which is the same as $(AA^* - \lambda I)u = 0$, where $\lambda = \sigma^2$. Now, as $u \neq \vec{0}$ then $\det(AA^* - \lambda I) = 0$, so we'll have

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \quad A^* = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \quad AA^* = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix} \quad AA^* - \lambda I = \begin{pmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{pmatrix}$$

$$\det(AA^* - \lambda I) = (17-\lambda)^2 - 8^2 = 0 \hookrightarrow \lambda_1 = 25,\ \lambda_2 = 9 \tag{1.3}$$

As $\lambda = \sigma^2$, and $\sigma \geq 0$, then from (1.3) we'll have that

$$\sigma_1 = 5,\ \sigma_2 = 3 \tag{1.4}$$

Then using (1.4) we'll have our Σ matrix, that is:

$$\boxed{\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}} \tag{1.5}$$

Next, let's find matrix V:

From (1.1) we have that $A^*Av = \sigma^2 v$, where $v \neq \vec{0}$ (because $v \in V$), which is the same as $(A^*A - \lambda I)v = 0$, where according to theory and (1.3) $\lambda_1 = 25$, $\lambda_2 = 9$ and $\lambda_3 = 0$. We have that

$$A^*A = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix} \tag{1.6}$$

Let $v_1 = (x_1, y_1, z_1)^T$, $v_2 = (x_2, y_2, z_2)^T$, $v_3 = (x_3, y_3, z_3)^T$ and $\widetilde{v}_1 = \frac{v_1}{\|v_1\|}$, $\widetilde{v}_2 = \frac{v_2}{\|v_2\|}$, $\widetilde{v}_3 = \frac{v_3}{\|v_3\|}$, then

$$(A^*A - \lambda_1 I)v_1 = 0 \hookrightarrow x_1 = y_1,\ z_1 = 0 \hookrightarrow v_1 = (1,1,0)^T,\ \widetilde{v}_1 = \frac{1}{\sqrt{2}}(1,1,0)^T \tag{1.7}$$

$$(A^*A - \lambda_2 I)v_2 = 0 \hookrightarrow x_2 = -y_2,\ z_2 = 4x_2 \hookrightarrow v_2 = (1,-1,4)^T,\ \widetilde{v}_2 = \frac{1}{3\sqrt{2}}(1,-1,4)^T \tag{1.8}$$

$$(A^*A - \lambda_3 I)v_3 = 0 \hookrightarrow x_3 = -y_3,\ y_3 = 2z_3 \hookrightarrow v_3 = (-2,2,1)^T,\ \widetilde{v}_3 = \frac{1}{3}(-2,2,1)^T \tag{1.9}$$

So, from (1.7)−(1.9) we'll have our V matrix, that is:

$$\boxed{V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \end{pmatrix}} \tag{1.10}$$

Now, let's compute U:

Let $\widetilde{u}_1 = \frac{u_1}{\|u_1\|}$, $\widetilde{u}_2 = \frac{u_2}{\|u_2\|}$. We know that $Av_i = \sigma_i u_i$, for all i from 1 to p. So, in our case $p = 2$, then using (1.7), (1.8) and (1.4) we'll have that

$$Av_1 = \sigma_1 u_1 \hookrightarrow u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{5}(5,5)^T = (1,1)^T,\ \widetilde{u}_1 = \frac{1}{\sqrt{2}}(1,1)^T \tag{1.11}$$

$$Av_2 = \sigma_2 u_2 \hookrightarrow u_2 = \frac{1}{\sigma_2}Av_2 = \frac{1}{3}(9,-9)^T = (3,-3)^T,\ \widetilde{u}_2 = \frac{1}{\sqrt{2}}(1,-1)^T \tag{1.12}$$

So, from (1.11)−(1.12) we'll have our U matrix, that is:

$$\boxed{U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}} \tag{1.13}$$

So, we've obtained the following singular value decomposition for A:

$$\underbrace{\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}}_{V^*}$$

Now, let's calculate the Reduced SVD:

In general, the reduced SVD can be presented as $A = \hat{U}\hat{\Sigma}\hat{V}^*$, where $\hat{U}$ is $m \times p$ matrix, $\hat{\Sigma}$ is $p \times p$ matrix, $\hat{V}$ is $n \times p$ matrix, and $p = \min(m, n)$. As been mentioned before in our problem we have $m = 2$ and $n = 3$, so $p = \min(2, 3) = 2$, which means that $\hat{U}$ is 2×2 matrix, so we get the whole U for $\hat{U}$ (see (1.13)), also $\hat{V}$ is 3×2 matrix, so we get 3×2 part of matrix V, and $\hat{\Sigma}$ is 2×2

matrix, so we get 2×2 part of matrix Σ. With all been explained above we'll have the following

$$\boxed{\hat{U} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}} \qquad \boxed{\hat{\Sigma} = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}} \qquad \boxed{\hat{V} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ 0 & \frac{4}{3\sqrt{2}} \end{pmatrix}} \tag{1.14}$$

So, we've obtained the following reduced singular value decomposition for A:

$$\underbrace{\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}}_{\hat{U}} \underbrace{\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}}_{\hat{\Sigma}} \underbrace{\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \end{pmatrix}}_{\hat{V}^*}$$

(b) Let's start by pointing out that in (1.10) and (1.13) in order to form matrices V and U, I've used the normalized vectors, i.e. vectors $\widetilde{v}_1$, $\widetilde{v}_2$, $\widetilde{v}_3$ for matrix V and vectors $\widetilde{u}_1$, $\widetilde{u}_2$ for matrix U, it's because matrices V and U must be unitary matrices, which means that they should satisfy $V^*V = VV^* = I$, and $U^*U = UU^* = I$. Also, let's notice that $\text{rank}(A) = 2$, which the same as the number of nonzero singular values, it's because $A = U\Sigma V^*$, and as V and U are unitary matrices, then they can be reduced by Gaussian Elimination to their respective identity matrices I_n and I_m, then we'll have $I_m \Sigma I_n^* = \Sigma$, which means that $\text{rank}(U\Sigma V^*) = \text{rank}(I_m\Sigma I_n^*) = \text{rank}(\Sigma)$, and on another hand $\text{rank}(A) = \text{rank}(U\Sigma V^*)$, so we'll have that $\text{rank}(A) = \text{rank}(\Sigma)$, which means that $\text{rank}(A)$ is the same as the number of nonzero singular values. Let's also notice that our matrix A can be presented as following

$$\begin{aligned} A &= 5 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right) + 3 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \left(\frac{1}{3\sqrt{2}}, -\frac{1}{3\sqrt{2}}, \frac{4}{3\sqrt{2}} \right) = \\ &= 5 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} + 3 \begin{pmatrix} \frac{1}{6} & -\frac{1}{6} & \frac{4}{6} \\ -\frac{1}{6} & \frac{1}{6} & -\frac{4}{6} \end{pmatrix} = \sigma_1 \widetilde{u}_1 \widetilde{v}_1^* + \sigma_2 \widetilde{u}_2 \widetilde{v}_2^* \end{aligned} \tag{1.15}$$

From (1.15) we can see that A is presented as sum of two rank-one matrices. This can be generalized for any matrices, by what, I mean that A in general can be presented as sum of $r = \text{rank}(A)$ rank-one matrices $A = \sum_{i=1}^{r} \sigma_i u_i v_i^*$, because we can present $\Sigma = \sum_{i=1}^{r} \Sigma_i$, where $\Sigma_i = \text{diag}(0, \ldots, \sigma_i, \ldots, 0)$, then as $A = U\Sigma V^*$, and as $U\Sigma V^* = \sum_{i=1}^{r} U\Sigma_i V^* = \sum_{i=1}^{r} \sigma_i u_i v_i^*$, what means that $A = \sum_{i=1}^{r} \sigma_i u_i v_i^*$.

2

Let A be an $m \times n$ matrix of full rank ($m \geq n$).

- What is the condition number of A in terms of the singular values of A? Use the 2-norm. (Hint: Use the SVD of A, i.e., $A = U\Sigma V^T$.)
- Using the SVD of A, compute the SVD of the following matrices in terms of U, Σ, and V:
  - $\left(A^T A\right)^{-1}$,
  - $\left(A^T A\right)^{-1} A^T$,
- Show all the steps and describe them.

- Add any comments you think are relevant to explain the results you get (try to be critical).

Solution:

(a)(c) As A is an $m \times n$ matrix of full rank, where $m \geq n$ then (see Problem 1)

$$\operatorname{rank}(A) = n = \min(m, n) = p \tag{2.1}$$

Now, let's compute the condition number:

We know that the condition number is the following

$$\kappa(A) = \frac{M}{m}, \text{ where } M = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}, \; m = \inf_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \tag{2.2}$$

So, as $A = U\Sigma V^*$, and as $V^*V = VV^* = I$, and $U^*U = UU^* = I$, then we'll have that

$$M = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x \neq 0} \frac{\|U\Sigma V^* x\|_2}{\|x\|_2} = \sup_{x \neq 0} \frac{\sqrt{x^* V \Sigma^* U^* U \Sigma V^* x}}{\|x\|_2} = \sup_{x \neq 0} \frac{\sqrt{x^* V \Sigma^* \Sigma V^* x}}{\|x\|_2} =$$
$$= \sup_{x \neq 0} \frac{\|\Sigma V^* x\|_2}{\|x\|_2} = \sup_{x \neq 0} \frac{\|\Sigma V^* x\|_2}{\|VV^* x\|_2}, \text{ now let's } y = V^* x, \text{ then}$$

$$M = \sup_{y \neq 0} \frac{\|\Sigma y\|_2}{\|Vy\|_2} = \sup_{y \neq 0} \frac{\|\Sigma y\|_2}{\sqrt{y^* V^* V y}} = \sup_{y \neq 0} \frac{\|\Sigma y\|_2}{\sqrt{y^* y}} = \sup_{y \neq 0} \frac{\|\Sigma y\|_2}{\|y\|_2} = \sup_{y \neq 0} \frac{\left(\sum_{i=1}^{p} \sigma_i^2 |y_i|^2\right)^{\frac{1}{2}}}{\left(\sum_{i=1}^{p} |y_i|^2\right)^{\frac{1}{2}}}, \text{ so} \tag{2.3}$$

$$M \leq \sigma_{\max} \cdot \sup_{y \neq 0} \frac{\left(\sum_{i=1}^{p} |y_i|^2\right)^{\frac{1}{2}}}{\left(\sum_{i=1}^{p} |y_i|^2\right)^{\frac{1}{2}}} = \sigma_{\max}, \text{ and as } \sigma_{\max} = \sigma_1 \text{ (see Problem 1) then } M \leq \sigma_1 \tag{2.4}$$

From (2.3) we can see that for $y = (1, 0, \ldots, 0)^T$ the supremum is achieved, and also using the result from (2.4), that means that

$$M = \sigma_1 \tag{2.5}$$

Now, let's see what is m equal to. As been showed above that

$$\|Ax\|_2 = \|U\Sigma V^* x\|_2 = \|\Sigma V^* x\|_2 = \|\Sigma y\|_2, \text{ where } y = V^* x, \text{ then as}$$
$$\|\Sigma y\|_2 = \left(\sum_{i=1}^{p} \sigma_i^2 |y_i|^2\right)^{\frac{1}{2}} \geq \sigma_{\min} \left(\sum_{i=1}^{p} |y_i|^2\right)^{\frac{1}{2}} = \sigma_{\min} \|y\|_2 \tag{2.6}$$
$$\|y\|_2 = \|V^* x\| = \sqrt{x^* V V^* x} = \sqrt{x^* x} = \|x\|_2 \tag{2.7}$$

Then using (2.6) and (2.7) from (2.2) we'll have that

$$m = \inf_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \geq \frac{\sigma_{\min} \|y\|_2}{\|y\|_2} = \sigma_{\min}, \text{ and as } \sigma_{\min} = \sigma_p \text{ (see Problem 1) then } m \geq \sigma_p \tag{2.8}$$

From (2.6) we can see that for $y = (0, 0, \ldots, 1)^T$ the infimum is achieved, and also using the result from (2.8), that means that

$$m = \sigma_p \tag{2.9}$$

So, using (2.5) and (2.9) from (2.2) we'll have that the condition number is

$$\boxed{\kappa(A) = \frac{\sigma_1}{\sigma_p}} \tag{2.10}$$

(b)(c) **SVD of $(A^*A)^{-1}$:**

As $A = U\Sigma V^*$, $V^*V = VV^* = I$, $U^*U = UU^* = I$, $V = (V^*)^{-1}$, and $V^* = V^{-1}$ then we'll have that

$$(A^*A)^{-1} = \left((U\Sigma V^*)^* U\Sigma V^*\right)^{-1} = (V\Sigma^*U^*U\Sigma V^*)^{-1} = (V\Sigma^*\Sigma V^*)^{-1} = (V^*)^{-1}(\Sigma^*\Sigma)^{-1}V^{-1}, \text{ so}$$
$$(A^*A)^{-1} = V(\Sigma^*\Sigma)^{-1}V^* = U_1\Sigma_1V_1^* \tag{2.11}$$

Let's analyze (2.11): We have that $(A^*A)^{-1}$ is $n \times n$ matrix, and $U_1 = V_1 = V$ is $n \times n$ unitary matrix, and $\Sigma^*\Sigma$ is a $n \times n$ diagonal matrix with entries along the diagonal that are $\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2$, where $p = n$ (see (2.1) and Problem 1), which means that $\Sigma^*\Sigma$ has a full rank that is equal to n, that means that $\Sigma^*\Sigma$ is invertible, so there exist $\Sigma_1 = (\Sigma^*\Sigma)^{-1}$ matrix which is $n \times n$ diagonal matrix with entries along the diagonal that are $\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_p^2}$ (because $\Sigma^*\Sigma(\Sigma^*\Sigma)^{-1} = (\Sigma^*\Sigma)^{-1}\Sigma^*\Sigma = I$). But here is one more thing left to do is that as you can notice entries of Σ_1 are located in nondecreasing order $\left(\frac{1}{\sigma_1^2} \le \frac{1}{\sigma_2^2} \le \cdots \le \frac{1}{\sigma_p^2}, \text{ because } \sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_p\right)$, while it's better when they are located in nonincreasing order. How to handle this? So, as been showed in Problem 1, for any matrix $A_x = \sum_{i=1}^{p} \overset{(x)}{\sigma_i}\overset{(x)}{u_i}\overset{(x)}{v_i^*} = \overset{(x)}{\sigma_1}\overset{(x)}{u_1}\overset{(x)}{v_1^*} + \cdots + \overset{(x)}{\sigma_p}\overset{(x)}{u_p}\overset{(x)}{v_p^*}$, then we can do the same thing for $A_x = (A^*A)^{-1}$, and we just put the singular values in descending order, i.e. $A_x = \sum_{i=1}^{p} \frac{1}{\sigma_{p-i+1}^2}\overset{(1)}{u}_{p-i+1}\overset{(1)}{v^*}_{p-i+1} = \overset{(1)}{U_1}\overset{(1)}{\Sigma_1}\overset{(1)}{V_1^*}$. Now we have that $\overset{(1)}{\Sigma_1}$ is $n \times n$ diagonal matrix with entries along the diagonal that are $\frac{1}{\sigma_p^2}, \frac{1}{\sigma_{p-1}^2}, \ldots, \frac{1}{\sigma_1^2}$, and they're located in nonincreasing order, and matrices $\overset{(1)}{U_1}$, $\overset{(1)}{V_1}$ are still unitary matrices because they are obtained from matrices U_1 and V_1 respectively, just by changing positions of their column elements (pth column with 1st column, $(p-1)$th column with 2nd column, …). Now, with all stated above we've satisfied the definition of SVD and also showed the SVD of $(A^*A)^{-1}$.

SVD of $(A^*A)^{-1}A^*$:

As $A = U\Sigma V^*$, $V^*V = VV^* = I$, $U^*U = UU^* = I$, $V = (V^*)^{-1}$, and $V^* = V^{-1}$ then we'll have that

$$(A^*A)^{-1}A^* = \left((U\Sigma V^*)^* U\Sigma V^*\right)^{-1}A^* = (V\Sigma^*U^*U\Sigma V^*)^{-1}A^* = (V\Sigma^*\Sigma V^*)^{-1}A^* =$$
$$= (V^*)^{-1}(\Sigma^*\Sigma)^{-1}V^{-1}A^* = (V^*)^{-1}(\Sigma^*\Sigma)^{-1}V^{-1}V\Sigma^*U^*, \text{ so}$$
$$(A^*A)^{-1}A^* = V(\Sigma^*\Sigma)^{-1}\Sigma^*U^* = U_2\Sigma_2V_2^* \tag{2.12}$$

Let's analyze (2.12): We have that $(A^*A)^{-1}A^*$ is $n \times m$ matrix, and $U_2 = V$ is $n \times n$ unitary matrix, $V_2 = U$ is $m \times m$ unitary matrix, and as been showed before $(\Sigma^*\Sigma)^{-1}$ is $n \times n$ diagonal matrix with entries along the diagonal that are $\frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_p^2}$, so $\Sigma_2 = (\Sigma^*\Sigma)^{-1}\Sigma^*$ is $n \times m$ diagonal matrix with entries along the diagonal that are $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_p}$, where $p = n$. Here as done before it's better to locate entries of Σ_2 in nonincreasing order. So we put the singular values of Σ_2 in descending order, i.e. $A_x = \sum_{i=1}^{p} \frac{1}{\sigma_{p-i+1}}\overset{(2)}{u}_{p-i+1}\overset{(2)}{v^*}_{p-i+1} = \overset{(2)}{U_2}\overset{(2)}{\Sigma_2}\overset{(2)}{V_2^*}$. Now we have that $\overset{(2)}{\Sigma_2}$ is $n \times m$ diagonal matrix with entries along the diagonal that are $\frac{1}{\sigma_p}, \frac{1}{\sigma_{p-1}}, \ldots, \frac{1}{\sigma_1}$, and they're located in nonincreasing order, and matrices $\overset{(2)}{U_2}$, $\overset{(2)}{V_2}$ are still unitary matrices because they are obtained from matrices U_2 and V_2 respectively, just by changing positions of their column elements (pth column with 1st column, $(p-1)$th column with 2nd column, ...). With all stated above we've satisfied the definition of SVD and also showed the SVD of $(A^*A)^{-1}A^*$.

(d) The first thing I'd like to point out is that $(A^*A)^{-1} \neq A^{-1}A^{-*}$, it's because that A is not square matrix, while A^*A is a square matrix and also has a full rank that is equal n (as been mentioned before), that's why we can apply the inverse operation on A^*A, but we cannot do that on matrices A or A^*, the same arguments apply to $(\Sigma^*\Sigma)^{-1}$. The second thing I'd like to mention is that we know that SVD of A is a factorization $A_x = U_x\Sigma_x V_x^*$, where $\Sigma_x \in \mathbb{C}^{m\times n}$ is diagonal matrix with nonnegative entries and we assume that they are located in nonincreasing order ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$), where $p = \min(m, n)$, and my emphasise is that the nonincreasing order (order of importance) is important. For that purpose, in section (b)(c) I did the ordering.

3

Write an essay discussing the differences between the eigenvalue decomposition and the singular value decomposition. Include one or more examples to illustrate your points.

Solution:

Introduction Eigenvalue decomposition (EVD) and singular value decomposition (SVD) are two fundamental techniques in linear algebra with diverse applications in fields like signal processing, machine learning, and quantum mechanics. While both methods serve the purpose of breaking down matrices into simpler components, they exhibit distinct properties and are suitable for different types of matrices and applications. This essay explores the differences between EVD and SVD, utilizing examples to illustrate their unique characteristics.

Eigenvalue Decomposition (EVD) Eigenvalue decomposition applies to square matrices and serves purposes like diagonalization and solving linear differential equations. Given a square $n \times n$ matrix A, EVD decomposes it into three matrices:

$$A = X\Lambda X^{-1}, \text{ where} \tag{3.1}$$

- A is the original $n \times n$ matrix.
- X is the $n \times n$ matrix of eigenvectors of A.
- Λ is the $n \times n$ diagonal matrix with eigenvalues of A along the diagonal.

Let's emphasize that only diagonalizable matrices can be factorized like in (3.1), which means that such matrices have enough linearly independent eigenvectors to form a complete basis for the vector space. In other words, it must have a full set of linearly independent eigenvectors. So, we also can say that any $n \times n$ matrix that has no repeated eigenvalues can be diagonalized, because suppose the eigenvalues $\lambda_1, \ldots, \lambda_n$ are different, then it means that the eigenvectors $x_1, \ldots, x_n$ are independent, and the eigenvector matrix X will be invertible. Let's see some example

$$A_1 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \hookrightarrow \det(A_1 - \lambda I) = 0 \hookrightarrow (\lambda - 1)(\lambda + 1) + 1 = 0 \hookrightarrow \lambda_{1,2} = 0 \tag{3.2}$$

So, from (3.2) we have that eigenvalues of A_1 are 0 and 0, and the problem is that 0 is repeated. For eigenvectors we'll get the following

$$\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \hookrightarrow y_1 = y_2 = c \hookrightarrow x_1 = c \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{3.3}$$

From (3.3) we can see that all eigenvectors of the matrix A_1 are multiples of $(1, 1)^T$, and there is no second linearly independent eigenvector, so matrix A_1 cannot be diagonalized.

Singular Value Decomposition (SVD) Singular value decomposition, unlike EVD, can be applied to both square and rectangular matrices, making it a cool tool in various applications. For $m \times n$ matrix A, SVD breaks it down into three matrices:

$$A = U\Sigma V^*, \text{ where} \tag{3.4}$$

- A is the $m \times n$ original matrix.
- U is the $m \times m$ unitary matrix containing left singular vectors.
- Σ is the $m \times n$ diagonal matrix with singular values (nonnegative and in nonincreasing order) along the diagonal.
- V^* is the $n \times n$ unitary matrix containing right singular vectors.

SVD vs EVD

- SVD uses two different bases (matrices U and V), while EVD uses just one (matrix X).
- Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition (can be clearly seen from proof of the theorem about existence of SVD, see e.g. Numerical Linear Algebra by Trefethen). For EVD from (3.1) we know that matrices should square, but even if they're square there's no guarantee of EVD, as been showed for matrix A_1 in (3.2).
- Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers (can be clearly seen from proof of the theorem about existence of SVD, see e.g. Numerical Linear Algebra by Trefethen). Let's give some example for this case by considering the following matrix

  $$A_2 = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix} \qquad A_2^* = \begin{pmatrix} 1 & 2 \\ -2 & -1 \end{pmatrix} \qquad A_2A_2^* = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$$

  So, for eigenvalues we have that $\det(A_2 - \lambda I) = 0 \hookrightarrow \lambda_{1,2} = \pm\sqrt{3}i$. For singular values (see Problem 1) we have that $\det(A_2A_2^* - \mu I) = 0 \hookrightarrow \mu_1 = \sigma_1^2 = 9,\ \mu_2 = \sigma_2^2 = 1$, which means that $\sigma_1 = 3,\ \sigma_2 = 1$. Or another example

  $$A_3 = \begin{pmatrix} i & -2 \\ 2 & -i \end{pmatrix} \qquad A_3^* = \begin{pmatrix} -i & 2 \\ -2 & i \end{pmatrix} \qquad A_3A_3^* = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

  So, in this for eigenvalues we have that $\det(A_3 - \lambda I) = 0 \hookrightarrow \lambda_{1,2} = \pm\sqrt{5}i$. For singular values we have that $\det(A_3A_3^* - \mu I) = 0 \hookrightarrow \mu_1 = \mu_2 = 5$, which means that $\sigma_1 = \sigma_2 = \sqrt{5}$.
- Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices) (see e.g. Introduction to Linear Algebra by Gilbert Strang), but left (or right) singular vectors are orthogonal to each other (because U and V are unitary matrices). For example let's take the following matrices

  $$A_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_5 = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$$

  So, for symmetric matrix A_4 for eigenvalues we have that $\det(A_4 - \lambda I) = 0 \hookrightarrow \lambda_1 = 1,\ \lambda_2 = -1$, then $(A_4 - \lambda_1 I)x_1 = 0 \hookrightarrow x_1 = (1,1)^T$, $(A_4 - \lambda_2 I)y_1 = 0 \hookrightarrow y_1 = (-1,1)^T$, so we can see $(x_1, y_1) = 0$, as should be for symmetric matrices. For matrix A_5 for eigenvalues we have that $\det(A_5 - \lambda I) = 0 \hookrightarrow \lambda_1 = 2,\ \lambda_2 = -2$, then $(A_5 - \lambda_1 I)x_1 = 0 \hookrightarrow x_1 = (0.5,1)^T$, $(A_5 - \lambda_2 I)y_1 = 0 \hookrightarrow y_1 = (-0.5,1)^T$, so we can see $(x_1, y_1) \neq 0$. For SVD, U and V are unitary matrices, which already means that left (or right) singular vectors are orthogonal to each other.
- In the SVD, the nondiagonal matrices U and V are not necessarily the inverse of one another. They are usually not related to each other at all. In the EVD the nondiagonal matrices X and X^{-1} are inverses of each other.
- In applications, eigenvalues are really helpful for problems like matrix powers A^k or exponentials e^{tA}, while SVD is helpful to problems involving the behavior of A itself.

Conclusion In conclusion, eigenvalue decomposition (EVD) and singular value decomposition (SVD) are essential techniques for matrix decomposition, but they differ significantly in their applicability and purpose. EVD is tailored for square matrices, enabling diagonalization and simplifying certain computations, while SVD allows to handle rectangular matrices, making it invaluable in various fields, particularly in data analysis and compression. Understanding these differences is crucial for selecting the appropriate technique based on the specific problem and matrix characteristics.