Homework II

MATH 596

Sandy Urazayev*

250; 12020 H.E.

Exercise 4.3

Part a

Let $\mathbf{A} := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. We can find the eigenvalues by solving $\det(\lambda \mathbf{I} - \mathbf{A})$ for λ. We get characteristic polynomial $(\lambda - 1)^2 = 0$, then the eigenvalue is $\lambda = 1$ with multiplicity of 2. Using this, we get

$$\begin{pmatrix} \lambda - 1 & 0 \\ -1 & \lambda - 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Therefore the eigenspace $E_1 = (\begin{pmatrix} 0 \\ 1 \end{pmatrix})$

Part b

Let $\mathbf{A} := \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix}$. We can find the eigenvalues by solving $\det(\lambda \mathbf{I} - \mathbf{A})$ for λ. We get characteristic polynomial

$$\det(\lambda \mathbf{I} - \mathbf{A}) = (\lambda + 2)(\lambda - 1) - 4 = \lambda^2 + \lambda - 6 = (\lambda - 2)(\lambda + 3) = 0$$

*University of Kansas (ctu@ku.edu)

then the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -3$

Using this, we get for λ_1

$$\begin{pmatrix} \lambda_1+2 & -2 \\ -2 & \lambda_1-1 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1/2 \\ 0 & 0 \end{pmatrix}$$

Therefore the eigenspace $E_2 = (\begin{pmatrix} 1 \\ 2 \end{pmatrix})$

Using this, we get for λ_2

$$\begin{pmatrix} \lambda_2+2 & -2 \\ -2 & \lambda_2-1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

Therefore the eigenspace $E_{-3} = (\begin{pmatrix} -2 \\ 1 \end{pmatrix})$

Exercise 4.8

Let us find the SVD of $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$. Firstly, let us find the eigenvalues and eigenvector of the matrix, so we can build a decomposition. Let $\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$, so then $\mathbf{A}^\top = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -1 \end{pmatrix}$. Further, $\mathbf{A}^\top\mathbf{A} = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$. To find the eigenvalue decomposition, we have to solve the equation

$$\det(\lambda\mathbf{I} - \mathbf{A}^\top\mathbf{A}) = \det\begin{pmatrix} \lambda-13 & -12 & -2 \\ -12 & \lambda-13 & 2 \\ -2 & 2 & \lambda-8 \end{pmatrix} = 0$$

for λ. We will get a characteristic polynomial when expanding the determinant

$$\det(\lambda\mathbf{I} - \mathbf{A}^\top\mathbf{A}) = \lambda^3 - 34\lambda^2 + 225 + 0 = \lambda(\lambda^2 - 34\lambda + 225) = \lambda(\lambda-25)(\lambda-9) = 0$$

We found our eigenvalues and let us assign them as follows, $\lambda_1 = 25$, $\lambda_2 = 9$, and $\lambda_3 = 0$. To find the respective eigenvectors, we substitute and find a solution to the resulting homogeneous system. Substituting λ_1, we get

$$\begin{pmatrix} \lambda_1 - 13 & -12 & -2 \\ -12 & \lambda_1 - 13 & 2 \\ -2 & 2 & \lambda_1 - 8 \end{pmatrix} = \begin{pmatrix} 12 & -12 & -2 \\ -12 & 12 & 2 \\ -2 & 2 & 17 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

therefore $\vec{v_1} = (1,1,0)$ and $\|\vec{v_1}\| = (1/\sqrt{2}, 1/\sqrt{2}, 0)$. Substituting λ_2, we get

$$\begin{pmatrix} \lambda_2 - 13 & -12 & -2 \\ -12 & \lambda_2 - 13 & 2 \\ -2 & 2 & \lambda_2 - 8 \end{pmatrix} = \begin{pmatrix} -4 & -12 & -2 \\ -12 & -4 & 2 \\ -2 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{pmatrix}$$

therefore $\vec{v_2} = (1,-1,4)$ and $\|\vec{v_2}\| = (1/\sqrt{18}, -1/\sqrt{18}, 4/\sqrt{18})$. Substituting λ_3, we get

$$\begin{pmatrix} \lambda_3 - 13 & -12 & -2 \\ -12 & \lambda_3 - 13 & 2 \\ -2 & 2 & \lambda_3 - 8 \end{pmatrix} = \begin{pmatrix} -13 & -12 & -2 \\ -12 & -13 & 2 \\ -2 & 2 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

therefore $\vec{v_3} = (-2,2,1)$ and $\|\vec{v_3}\| = (-2/3, 2/3, 1/3)$.

So then

$$\mathbf{V} = \begin{pmatrix} \|\vec{v_1}\| & \|\vec{v_2}\| & \|\vec{v_3}\| \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & 2/3 \\ 0 & 4/\sqrt{18} & 1/3 \end{pmatrix}$$

$\mathbf{\Sigma}$ is the a diagonal matrix composed of square roots of the eigenvalues. We know that $\lambda_3 = 0$ so $\sigma_3 = 0$, we will need to pad $\mathbf{\Sigma}$ to fit the dimensions of this SVD

$$\mathbf{\Sigma} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

We can trivially compute $\mathbf{U}$ by given

$$\mathbf{U} = \begin{pmatrix} 1/5\mathbf{A}\|\vec{v_1}\| & 1/3\mathbf{A}\|\vec{v_2}\| & \vec{0} \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}$$

Finally, we can verify that

$$\therefore \mathbf{A} = \mathbf{U\Sigma V}^\top = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

See numerical confirmation in Exercise 4.8 code

Exercise 4.9

Let us find the SVD of $\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$. Firstly, let us find the eigenvalues and eigenvector of the matrix, so we can build a decomposition. Let $\mathbf{A} = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$, so then $\mathbf{A}^\top = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}$. Further, $\mathbf{A}^\top\mathbf{A} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$. To find eigenvalues, we have to solve

$$\det(\lambda\mathbf{I} - \mathbf{A}^\top\mathbf{A}) = \det\begin{pmatrix} \lambda - 5 & -3 \\ -3 & \lambda - 5 \end{pmatrix} = 0$$

for λ. After some trivial computations, we get $\lambda_1 = 8$ and $\lambda_2 = 2$. Let us find the corresponding eigenvectors $\vec{v_1}$ and $\vec{v_2}$. Substituting λ_1, we get

$$\begin{pmatrix} \lambda_1 - 5 & -3 \\ -3 & \lambda_1 - 5 \end{pmatrix} = \begin{pmatrix} 8-5 & -3 \\ -3 & 8-5 \end{pmatrix} \sim \begin{pmatrix} 3 & -3 \\ -3 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

therefore $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Similarly, substituting λ_2, we have

$$\begin{pmatrix} \lambda_2 - 5 & -3 \\ -3 & \lambda_2 - 5 \end{pmatrix} = \begin{pmatrix} 2-5 & -3 \\ -3 & 2-5 \end{pmatrix} \sim \begin{pmatrix} -3 & -3 \\ -3 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

so $\vec{v_2} = (1, -1)$. Let us normalize the eigenvector, so $\|\vec{v_1}\| = (1/\sqrt{2}, 1/\sqrt{2})$ and $\|\vec{v_2}\| = (1/\sqrt{2}, -1/\sqrt{2})$. So then $\mathbf{P} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$. Finally, we see how $\mathbf{V} = \mathbf{P}$ and $\mathbf{\Sigma}$ is the square root of diagonal values of $\mathbf{D}$, therefore

$\boldsymbol{\Sigma} = \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$. To find U, we compute

$$\mathbf{U} = \begin{pmatrix} {}^1\!/\!_{2\sqrt{2}}\mathbf{A}\|\vec{v}_1\| & {}^1\!/\!_{\sqrt{2}}\mathbf{A}\|\vec{v}_2\| \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Finally, we can confirm that

$$\therefore \mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$$

See numerical confirmation in Exercise 4.9 code

Exercise 4.10

Recall from Exercise 4.8 that

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

Let us construct a rank-1 matrix $\mathbf{A}_1$, where

$$\mathbf{A}_1 = \vec{u}_1 \otimes \vec{v}_i$$

which is formed by outer product of ith orthogonal column vector of $\mathbf{U}$ and $\mathbf{V}$. Computing, we get

$$\mathbf{A}_1 = \vec{u}_1 \otimes \vec{v}_1 = \vec{u}_1\vec{v}_1^\top = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

We can finally find the rank approximation $\hat{\mathbf{A}}(1)$

$$\hat{\mathbf{A}}(1) = \sum_{i=1}^{1} \sigma_i \mathbf{A}_i = \sigma_1 \mathbf{A}_1 = 5 \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix} = \begin{pmatrix} 5/2 & 5/2 & 0 \\ 5/2 & 5/2 & 0 \end{pmatrix}$$

Exercise 4.11

Show that for any $A \in \mathbb{R}^{m\times n}$ the matrices $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ possess the same nonzero eigenvalues.

Proof: Let $A \in \mathbb{R}^{m\times n}$. Then by finding the eigenvalues λ of $\mathbf{A}\mathbf{A}^\top$, it satisfies

$$\mathbf{A}\mathbf{A}^\top\vec{v} = \lambda\vec{v}$$

for some associated eigenvector $\vec{v}$. Further, let us left-multiply both sides by $\mathbf{A}^\top$

$$\begin{aligned}
\mathbf{A}\mathbf{A}^\top\vec{v} = \lambda\vec{v}& \\
&\implies \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top\vec{v}) = \mathbf{A}^\top\lambda\vec{v} \\
&\implies \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top\vec{v}) = \lambda\mathbf{A}^\top\vec{v} \\
&\implies \mathbf{A}^\top\mathbf{A}(\mathbf{A}^\top\vec{v}) = \lambda(\mathbf{A}^\top\vec{v})
\end{aligned}$$

Since $\lambda \neq 0$, let $\vec{x} = \mathbf{A}^\top\vec{v} \neq 0$. So then

$$\mathbf{A}^\top\mathbf{A}(\mathbf{A}^\top\vec{v}) = \lambda(\mathbf{A}^\top\vec{v}) \implies \mathbf{A}^\top\mathbf{A}\vec{x} = \lambda\vec{x}$$

where $\vec{x}$ is the eigenvector associated with same non-zero λ of $\mathbf{A}\mathbf{A}^\top$ for $\mathbf{A}^\top\mathbf{A}$

$\therefore$ the matrices $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ possess the same nonzero eigenvalues. □

Exercise 4.12

Show that for $\mathbf{x} \neq \mathbf{0}$ Theorem 4.24 holds, i.e., show that

$$\max_x \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2 = \sigma_1}$$

where σ_1 is the largest singular value of $\mathbf{A} \in \mathbb{R}^{m\times n}$

Proof: Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Firstly, let us note that we can decompose $\mathbf{A}$ using SVD, so then $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices. Recall the properties of square orthogonal matrices $\mathbf{Q}$, when $\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$. Applied to Euclidean matrix norms, we derive such a property, where for any dimension-wise compatible $\mathbf{x} \neq \mathbf{0}$

$$\|\mathbf{Q}\mathbf{x}\|_2^2 = \mathbf{x}^\top \mathbf{Q}^\top \mathbf{Q}\mathbf{x} = \mathbf{x}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{I}\mathbf{x} = \|\mathbf{x}\|_2^2$$

When $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$, we apply the derivation above, where we see

$$\max_x \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_x \frac{\|\mathbf{U}\Sigma\mathbf{V}^\top\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_x \frac{\|\Sigma\mathbf{V}^\top\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

To simplify the max arg, let us define a new variable $\mathbf{y}$, such that $\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_2}$, where $\|\mathbf{y}\|_2 = 1$. Therefore we built an equivalent statement, where

$$\max_x \frac{\|\Sigma\mathbf{V}^\top\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \equiv \max_{\|\mathbf{y}\|_2=1} \left\|\Sigma\mathbf{V}^\top\mathbf{y}\right\|_2$$

Note that by SVD, **Sigma** is a diagonal matrix, with its main diagonal composed of n singular values, ranging from highest value to the lowest $\sigma_1 > \sigma_2 > \ldots > \sigma_n$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & \ldots \\ \vdots & \vdots & \vdots \\ \ldots & 0 & \sigma_n \end{pmatrix}$$

The main question at hand, is how can $\mathbf{y}$ result in the following $\|\Sigma\mathbf{V}^\top\mathbf{y}\|_2 = \sigma_1$. The nature of $\boldsymbol{\Sigma}$ gives us a great opportunity to extract the largest singular value when $\boldsymbol{\Sigma}\mathbf{z} = (\sigma_1, 0, \ldots, 0) \implies \|\boldsymbol{\Sigma}\mathbf{z}\|_2 = \sigma_1$, for some $\mathbf{z} = (1, 0, \ldots, 0)$. For convenience, let $\mathbf{z} = \mathbf{V}^\top\mathbf{y}$. Solving for $\mathbf{y}$, we get that $\mathbf{y} = \mathbf{V} \cdot (1, 0, \ldots, 0)$. Let us verify that this satisfies all conditions before. $\mathbf{V}$ is an orthogonal matrix, composed out of normal vectors, so the Euclidean norm of each column is 1. Writing $\mathbf{y}$ as a unique combination of $\mathbf{V}$'s columns with some linear coefficients $\alpha_1, \ldots$, we get $\mathbf{y} = \alpha_1\mathbf{V}_1 + \ldots + \alpha_n\mathbf{V}_n$ where V_k is the k^{th} column of $\mathbf{V}$. As mentioned, the following columns are normalized, therefore

$$\|\mathbf{y}\|_2^2 = (\alpha_1\mathbf{V}_1^\top + \ldots + \alpha_n\mathbf{V}_n^\top) \cdot (\alpha_1\mathbf{V}_1 + \ldots + \alpha_n\mathbf{V}_n)$$

The nature of $\mathbf{V}$ lets us see

$$\mathbf{V}_i^\top\mathbf{V}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

then

$$\|\mathbf{y}\|_2^2 = (\alpha_1\mathbf{V}_1^\top + \ldots + \alpha_n\mathbf{V}_n^\top)\cdot(\alpha_1\mathbf{V}_1 + \ldots + \alpha_n\mathbf{V}_n) = \alpha_1^2 + \ldots + \alpha_n^2$$

Therefore

$$\mathbf{\Sigma}\mathbf{V}^\top\cdot(\alpha_1\mathbf{V}_1 + \ldots + \alpha_n\mathbf{V}_n) = \mathbf{\Sigma}\left[\alpha_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \ldots + \alpha_n\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right]$$

This holds true iff $\alpha_1 = 1$ and for all $j = 2, \ldots,$ we have $\alpha_j = 0$. It is clear that the following conditions will always be satisfied for any $\mathbf{A} \in \mathbb{R}^{m\times n}$ thanks to SVD and we can safely confirm there always exists such a $\|\mathbf{y}\|_2 = 1$. Where $\|\mathbf{\Sigma}\mathbf{V}^\top\mathbf{y}\|_2$ cannot possibly exceed σ_1. Finally $\mathbf{x}$ dependent on it will maximize the spectral norm, yielding the maximum singular value σ_1. □

Appendix

Exercise 4.8 code

```
>>> U
matrix([[0.70710678, 0.70710678],
        [0.70710678, 0.70710678]])
```

```
>>> sigma
matrix([[5, 0, 0],
        [0, 3, 0]])
```

```
>>> V
matrix([[ 0.70710678,  0.23570226, -0.66666667],
        [ 0.70710678, -0.23570226,  0.66666667],
        [ 0.        ,  0.94280904,  0.33333333]])
```

```
>>> U * sigma * V.T
matrix([[3., 2., 2.],
        [3., 2., 2.]])
```

Exercise 4.9 code

```
>>> U
matrix([[ 1,  0],
        [ 0, -1]])
```

```
>>> sigma
matrix([[2.82842712, 0.        ],
        [0.        , 1.41421356]])
```

```
>>> V
matrix([[ 0.70710678,  0.70710678],
        [ 0.70710678, -0.70710678]])
```

```
>>> U * sigma * V.T
matrix([[ 2.,  2.],
        [-1.,  1.]])
```