Good job!

39/40

# Exercise 04

Sunday, November 5, 2023 6:13 PM

**Exercise 12** (10 points) 神经病

**(a)** (6 points) Show that the spectral norm of a matrix $A$ is equal to its largest singular value.

[Remark: Here you are allowed to use without proof that for real numbers $a_1, \ldots, a_n$, a vector $x = (x_1, \ldots, x_n)^T$ with $\|x\| = 1$ that maximizes $\sum_{i=1}^n a_i x_i^2$ is the standard basis vector $e_k$ where $k$ is the index with $a_k = \max\{a_i \mid i = 1 \ldots n\}$. All the necessary arguments ocurred in the proof of Theorem 1.7.12.]

**(b)** (4 points, tricky) Prove that the matrix $\Sigma$ in a singular value decomposition $A = U\Sigma V^T$ is unique (if one demands that the diagonal entries are all positive and ordered by size).

[Hint: Use the Eckart-Young-Mirsky theorem in its version for the spectral norm, together with (a).]

$\|A\| = \max\{\|Av\|_2 \mid \forall v \in V, \|v\|_2 = 1\}$

Assume $A \in Mat_{m\times n}$, it is then a map $V \to W$,

where $\dim V = m$ and $\dim W = n$

(a) Suppose $v \in W$ is arbitrary with $\|v\| = 1$

$U \in Mat_{m\times m}$, $\Sigma \in Mat_{m\times n}$, $V \in Mat_{n\times n}$ } $\Sigma^T\Sigma \in Mat_{n\times n}$

$\|Av\|_2^2 = \langle Av, Av\rangle = v^T A^T A v$

$\underset{SVD}{=} v^T(V\Sigma^T U^T U \Sigma V^T)v$

$= v^T(V\Sigma^T\Sigma V^T)v$

$= (V^T v)^T \cdot (\Sigma^T\Sigma)\cdot(V^T v)$

let $x = V^T v = \{x_1, \ldots, x_n\}$

since $V^T$ is orthogonal, it preserves length and angle.

$\|v\|_2 = 1 \Rightarrow \|V^T v\|_2 = 1 \Rightarrow \|x\|_2 = 1$

so $\|Av\|^2 = x^T(\Sigma^T\Sigma)x = \sum_{i=1}^n a_i x_i^2$, where $a_i$ $(i = 1, \ldots n)$

are the diagonal entries of $\Sigma^T\Sigma$, $a_1 > a_2 > \cdots > a_n$

According to the remark, for $x$ which

maximizes $\|Av\| \geq 0 \Rightarrow \max\{\|Av\|^2\}$

$= \max\{\sum_{i=1}^n a_i x_i^2\}$

$x = e_1$, $\|Av\|_{max} = \sqrt{a_1 \cdot 1 + a_2 \cdot 0 + \cdots + a_n \cdot 0}$

$= \sqrt{a_1} = \sigma_1$

and $\sqrt{a_1}$ is the largest singular value of $A$ $\square$ ✓

**(b)** (4 points, tricky) Prove that the matrix $\Sigma$ in a singular value decomposition $A = U\Sigma V^T$ is unique (if one demands that the diagonal entries are all positive and ordered by size).

[Hint: Use the Eckart-Young-Mirsky theorem in its version for the spectral norm, together with (a).]

Suppose for $A \in Mat_{m\times n}$, $\exists\, \Sigma' \neq \Sigma$, s.t. $A = U'\Sigma' V'^T$ is another SVD

We assume diagonal entries of $\Sigma$ : $\sigma_1, \ldots, \sigma_r$ $\sigma_1, \ldots, \sigma_r \in \mathbb{R}_{>0}$, $\sigma_1 > \cdots > \sigma_r$, $r = \min\{n, m\}$

$\Sigma'$ : $\sigma_1', \ldots, \sigma_r'$ $\sigma_1', \ldots, \sigma_r' \in \mathbb{R}_{>0}$, $\sigma_1' > \cdots > \sigma_r'$

1. from (a) we know $\|A\| = \sigma_1$ and $\|A\| = \sigma_1' \Rightarrow \sigma_1 = \sigma_1'$

2. for $\forall i$, $1 \leq i \leq r-1$ : $\|A - A^{(i)}\| = \|\Sigma - \Sigma^{(i)}\| = \|\Sigma' - \Sigma'^{(i)}\|$

we define $B := \Sigma - \Sigma^{(i)}$, $B' = \Sigma' - \Sigma'^{(i)}$

from (a) we know $\|B\| = \max\{\sigma_{i+1}, \ldots, \sigma_r\} = \sigma_{i+1}$

Similarly $\|B'\| = \sigma_{i+1}' = \|A - A^{(i)}\| = \|B\| = \sigma_{i+1}$

Therefore, for $\forall k$, $1 \leq k \leq r$ : $\sigma_k = \sigma_k'$

$\Rightarrow \Sigma = \Sigma'$ $\perp$ contradiction! ✓

Other ideas

$\|A - A^{(k)}\| \leq \|A - U\Sigma'^{(k)}V^T\|$

$\sigma_{k+1} \leq \|\Sigma - \Sigma'^{(k)}\|$

$\Downarrow$

$\max\left(\sum_{i=1}^k (\sigma_i - \sigma_i')^2 v_i^2 + \sum_{i=k+1}^r \sigma_i^2 v_i^2\right)$

This might not work.

10/10

**Exercise 13** (10 points)

Compute a singular value decomposition of the following matrix:

$$A = \begin{pmatrix} 2 & -2 & 1 \\ -4 & -8 & -8 \end{pmatrix}$$

$$A^T = \begin{pmatrix} 2 & -4 \\ -2 & -8 \\ 1 & -8 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & -2 & 1 \\ -4 & -8 & -8 \end{pmatrix}$$

we find SVD of $A^T$ and then transpose the result to reduce calculation in eigenvalues

$$(A^T)^T A^T = AA^T = \begin{pmatrix} 4+4+1 & -8+16-8 \\ -8+16-8 & 16+64+64 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 0 \\ 0 & 144 \end{pmatrix} \Rightarrow \lambda_{1,2} = 144, 9$$

$\lambda_2 = 9$: eigenvector $V_2 = (a, b)^T$

$$\begin{pmatrix} 0 & 0 \\ 0 & 135 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$b = 0$, let $a = 1$ $\Rightarrow V_2 = (1, 0)^T$

$\lambda_1 = 144$: $V_1 = (c, d)^T$

$$\begin{pmatrix} -135 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$c = 0$, Let $d = 1$ $\Rightarrow V_1 = (0, 1)^T$

$\langle V_1, V_2 \rangle = 0$

$\|V_1\| = \|V_2\| = 1$

$$\Sigma = \begin{pmatrix} \sqrt{144} & 0 \\ 0 & \sqrt{9} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 12 & 0 \\ 0 & 3 \\ 0 & 0 \end{pmatrix}$$

so the right orthogonal matrix for $A^T$ can be

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad A^T = \begin{pmatrix} 2 & -4 \\ -2 & -8 \\ 1 & -8 \end{pmatrix}$$

$$A^T V_1 = \begin{pmatrix} -4 \\ -8 \\ -8 \end{pmatrix} \qquad u_1 = \frac{A^T V_1}{\|A^T V_1\|} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} \qquad \langle e_2, u_1 \rangle = \langle e_2, u_2 \rangle = -\frac{2}{3}$$

$$A^T V_2 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \qquad u_2 = \frac{A^T V_2}{\|A^T V_2\|} = \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

$$\tilde{u}_3 = e_2 - \langle e_2, u_1 \rangle u_1 - \langle e_2, u_2 \rangle u_2$$

$$= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{2}{9} \\ \frac{4}{9} \\ \frac{4}{9} \end{pmatrix} - \begin{pmatrix} -\frac{4}{9} \\ \frac{4}{9} \\ -\frac{2}{9} \end{pmatrix} = \begin{pmatrix} \frac{2}{9} \\ \frac{1}{9} \\ -\frac{2}{9} \end{pmatrix}$$

$$\|\tilde{u}_3\| = \sqrt{\frac{4}{81} + \frac{1}{81} + \frac{4}{81}} = \frac{1}{3} \qquad u_3 = \frac{\tilde{u}_3}{\|\tilde{u}_3\|} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{pmatrix}$$

so left orthogonal matrix $U = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$

So left orthogonal matrix $U = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}$

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

$\Rightarrow \quad A = (A^T)^T = V \Sigma^T U^T$

$$= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 12 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix} \quad \checkmark$$

easy to forget the 0s

10/10

### Exercise 14 (10 points)

Find the polynomial of degree 2 in one variable that best approximates (in the sense of least squares) a function whose graph passes through the points $(-1,0), (2,1), (1,-1)$ and $(0,1)$ in $\mathbb{R}^2$. Use the idea of exercise 2 for this!

[Remark: It is up to you how you solve the exercise. One way is to use that the best approximate solution $x$ to the equation $Ax = y$ is given by $A^+y$, where $A^+$ denotes the pseudoinverse of $A$ – we will discuss this on Tuesday, it's Theorem 1.8.4. If you want to solve the exercise before Tuesday, you can simply use that. In the lecture I briefly showed the proof of Thm 1.8.2, i.e. how to compute a pseudoinverse in general. Alternatively, you have already seen how to compute a pseudoinverse for a full rank matrix in exercise 5. You can do it either way, but if you use the approach of exercise 5, don't forget to check that you actually have a full rank matrix. The method of exercise 5 then involves inverting a matrix – you don't need to show your calculations on how to find that inverse (e.g. you can let a computer find it). You can simply write down the inverse, preferredly in the form $\frac{1}{20}$ times an integer matrix.]

A polynomial of degree 2 can be written as

$y = w_2 x^2 + w_1 x + b \quad , \quad w_1, w_2, b \in \mathbb{R}$

Insert the known points into the equation:

$$\begin{cases} 1 \cdot w_2 + (-1) \cdot w_1 + b = 0 \\ 4 \cdot w_2 + 2 \cdot w_1 + b = 1 \\ 1 \cdot w_2 + 1 \cdot w_1 + b = -1 \\ 0 \cdot w_2 + 0 \cdot w_1 + b = 1 \end{cases}$$

$$\Rightarrow \underbrace{\begin{pmatrix} 1 & -1 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{= A} \underbrace{\begin{pmatrix} w_2 \\ w_1 \\ b \end{pmatrix}}_{= x} = \underbrace{\begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}}_{= y} \quad \checkmark$$

Now perform SVD on matrix $A$:

$$A^T = \begin{pmatrix} 1 & 4 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 1+16+1 & -1+8+1 & 1+4+1 \\ & 1+4+1 & -1+2+1 \\ \text{sym} & & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{pmatrix}$$

| | $B_1$ | $B_2$ | $B_3$ |
|---|---|---|---|
| 1 | 1/4 | -1/4 | -1/4 |
| 2 | -1/4 | 9/20 | 3/20 |
| 3 | -1/4 | 3/20 | 11/20 |

$$\det(A^T A) = 18 \cdot 6 \cdot 4 + 8 \cdot 2 \cdot 6 + 8 \cdot 2 \cdot 6 - 18 \cdot 2 \cdot 2 - 8 \cdot 8 \cdot 4 - 6 \cdot 6 \cdot 6$$

$$= 432 + 96 + 96 - 72 - 256 - 216$$

$= 80 \neq 0$, therefore $A^T A$ is invertible $\checkmark$

$$(A^T A)^{-1} = \frac{1}{20} \begin{pmatrix} 5 & -5 & -5 \\ -5 & 9 & 3 \\ -5 & 3 & 11 \end{pmatrix}$$

$$A^+ = (A^T A)^{-1} A^T = \frac{1}{20} \begin{pmatrix} 5 & -5 & -5 \\ -5 & 9 & 3 \\ -5 & 3 & 11 \end{pmatrix} \begin{pmatrix} 1 & 4 & 1 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \frac{1}{20} \begin{pmatrix} 5+5-5 & 20-10-5 & 5-5-5 & -5 \\ -5-9+3 & -20+18+3 & -5+9+3 & 3 \\ \end{pmatrix}$$

$$= \frac{1}{20}\begin{pmatrix} 5+5-5 & 20-10-5 & 5-5-5 & -5 \\ -5-9+3 & -20+18+3 & -5+9+3 & 3 \\ -5-3+11 & -20+6+11 & -5+3+11 & 11 \end{pmatrix}$$

$$= \frac{1}{20}\begin{pmatrix} 5 & 5 & -5 & -5 \\ -11 & 1 & 7 & 3 \\ 3 & -3 & 9 & 11 \end{pmatrix} \checkmark \qquad \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

so $\hat{x} = A^{+}y$

$$= \frac{1}{20}\begin{pmatrix} 5+5-5 \\ 1-7+3 \\ -3-9+11 \end{pmatrix} = \frac{1}{20}\begin{pmatrix} 5 \\ -3 \\ -1 \end{pmatrix} \checkmark$$

10/10

Ex15:
a) and b) 6/6
c) You should have used the train and test sets provided by the instructions
3/4
xtrain, xtest, ytrain, ytest = train_test_split(X, y, test_size=0.25)