به نام خدا

پاسخ تمرین سری دوم

۱-

الف)

$$A \overset{SVD}{=} U\Sigma V^T \qquad V = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$tr\sqrt{A^TA} = tr\sqrt{V\Sigma^T U^T U \Sigma V^T} = tr\sqrt{V\Sigma^2 V^T} \overset{*}{=} tr\big(\overbrace{V\Sigma V^T}^{\triangleq S}\big) = S_{11} + S_{22} + \dots + S_{nn}$$

$$= \sigma_1 v_1^T v_1 + \dots + \sigma_n v_n^T v_n = \sigma_1 + \dots + \sigma_n = \|A\|_* \ \boxed{\checkmark}$$

$*$: $V\Sigma^2 V^T$ is schur decom. of A^TA

ب)

$$tr(AB) = (AB)_{11} + \dots + (AB)_{nn} = \sum_{i=1}^{n} a_{1i} b_{i1} + \dots + \sum_{i=1}^{n} a_{ni} b_{in}$$

$$= \underbrace{Sum}_{\text{جمع درایه‌های ماتریس}}(A \underbrace{.*}_{\text{Hadamard}} B^T)$$

به طریق مشابه : $tr(BA) = Sum(B .* A^T)$

$\xrightarrow{\text{داریم}}$: $Sum(A .* B^T) = Sum((A .* B^T)^T) = Sum(A^T .* B) = Sum(B .* A^T)$

$\Rightarrow tr(AB) = tr(BA)$ $\boxed{\checkmark}$

پ)

$$\max_{\substack{C^TC=I \\ \text{یعنی متعامد باشد}}} tr(AC) = \max tr(U\Sigma V^T C) \overset{(ب)}{=} \max tr\big(\Sigma \underbrace{\overbrace{V^T C U}^{\triangleq Q}}_{\text{یک ماتریس متعامد}}\big) = \max \sigma_1 Q_{11} + \dots + \sigma_n Q_{nn}$$

$$\leq \sigma_1 + \dots + \sigma_n = \|A\|_*$$ (با شرط تساوی که $Q=I$ باشد)

Q متعامد است پس درایه‌های آن کمتر از یک است.

کافی است نشان دهیم چنین C ای داریم که Q برابر I شود،

$Q = I = V^TCU \Rightarrow$ VU^T شرط $Q=I$ را ارضا می‌کند $\boxed{\checkmark}$

ت)

$$\|A+B\|_* \overset{(پ)}{=} \max_{CC^T=I} tr((A+B)C) = \max_{CC^T=I} tr(AC) + tr(BC)$$

$$\leq \max_{C_1C_1^T=I} tr(AC_1) + \max_{C_2C_2^T=I} tr(BC_2) \overset{(پ)}{=} \|A\|_* + \|B\|_* \ \boxed{\checkmark}$$

-۲

الف)

$$
\begin{aligned}
&H^T H=\left(I-2vv^T/\|v\|^2\right)^T\left(I-2vv^T/\|v\|^2\right)=\\
&\left(I^T-\left(2vv^T\right)^T/\|v\|^2\right)\left(I-2vv^T/\|v\|^2\right)=\\
&\left(I^T-2vv^T/\|v\|^2\right)^T\left(I-2vv^T/\|v\|^2\right)=\\
&I-\frac{2vv^T}{\|v\|^2}-\frac{2vv^T}{\|v\|^2}+\frac{4\left(vv^T\right)\left(2vv^T\right)}{\|v\|^4}=\\
&I-\frac{4vv^T}{\|v\|^2}+\frac{4vv^T}{\|v\|^2}=I
\end{aligned}
$$

ب)

$$
\begin{aligned}
&\det(H-\lambda I)=\det\left(\left(I-2vv^T/\|v\|^2\right)-\lambda I\right)\\
&=\det\left((1-\lambda)I-2vv^T/\|v\|^2\right)\\
&=\det\left((1-\lambda)I\right)\det\left(I-2vv^T/\|v\|^2\right)\\
&=(1-\lambda)^n\det\left(I-2vv^T/\|v\|^2\right)
\end{aligned}
$$

از انجایی که مولفه det در خط آخر به مقدار دترمینان ماتریس متعامد (همانطور که در بخش الف نشان داده شد) تعلق دارد، مقدار آن مثبت، منفی یک است. بنابراین، مقادیر ویژه H برابر مثبت، منفی یک هستند.

-۳

3. The Householder matrix reflects all vectors in the direction of $\mathbf{v}$

$$\mathbf{H}(\alpha\mathbf{v})=\left(\mathbf{I}-2\frac{\mathbf{v}\mathbf{v}^{\mathbf{T}}}{\mathbf{v}^{\mathbf{T}}\mathbf{v}}\right)(\alpha\mathbf{v})=\alpha\mathbf{v}-2\alpha\frac{\mathbf{v}(\mathbf{v}^{\mathbf{T}}\mathbf{v})}{\mathbf{v}^{\mathbf{T}}\mathbf{v}}=\alpha(\mathbf{v}-2\mathbf{v})=-(\alpha\mathbf{v})$$

and leaves all vectors $\mathbf{x}$ with $\mathbf{v}^{\mathbf{T}}\mathbf{x}=\mathbf{0}$ invariant

$$\mathbf{H}\mathbf{x}=\left(\mathbf{I}-2\frac{\mathbf{v}\mathbf{v}^{\mathbf{T}}}{\mathbf{v}^{\mathbf{T}}\mathbf{v}}\right)\mathbf{x}=\mathbf{x}-2\frac{\mathbf{v}(\mathbf{v}^{\mathbf{T}}\mathbf{x})}{\mathbf{v}^{\mathbf{T}}\mathbf{v}}=\mathbf{x}$$

therefore, $\mathbf{H}$ is a reflector about the hyperplane $\{x:\mathbf{v}^{\mathbf{T}}\mathbf{x}=\mathbf{0}\}$.

-۴

Let

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

be the SVD of A. Then, because U is orthogonal,

$$||Ax|| = ||\Sigma V^* x||$$

If $V^*x = y$, $||y|| = ||x|| = 1$ we are left with

$$||Ax||^2 = ||\Sigma y||^2 = \sum_i |\sigma_i y_i|^2 \geq \sum_i |\sigma_q y_i|^2 = \sigma_q^2$$

So, the inequality holds. If x is a right singular vector of A then it is one of the columns of V. But V is orthogonal, so there i such that $V^*x = e_i$, hence$||Ax|| = ||\Sigma V^* x|| = ||\Sigma e_i|| = \sigma_i$. So, any singular vector of σ_q minimises $||Ax||$ and the result follows. Note that it is unaccurate to talk of the minimiser of $||Ax||$, as the singular vectors are not unique (we can multiply x by -1).

-۵

Lemma 1 *If the columns of an $m \times n$ matrix A are linearly independent, then the $n \times n$ matrix $A^\top A$ is non-singular. Similarly, if the rows of A are linearly independent, then the $m \times m$ matrix $AA^\top$ is non-singular.*

Proof Suppose A has linearly independent columns. It follows that

$$\begin{aligned}(A^\top A)\boldsymbol{x} = 0 \quad &\Longrightarrow \quad \boldsymbol{x}^\top A^\top A\boldsymbol{x} = 0 \\ &\Longrightarrow \quad (A\boldsymbol{x})^\top (A\boldsymbol{x}) = 0 \\ &\Longrightarrow \quad ||A\boldsymbol{x}|| = 0 \\ &\Longrightarrow \quad A\boldsymbol{x} = 0 \\ &\Longrightarrow \quad \boldsymbol{x} = \mathbf{0},\end{aligned}$$

where the last step follows from that the columns of A are linearly independent. Thus, the null space of $A^\top A$ contains only $\mathbf{0}$. The matrix is therefore non-singular.

Now suppose that A has linearly independent rows. Equivalently, $A^\top$ has linearly independent columns. Applying what we have just shown above, $(A^\top)^\top A^\top = AA^\top$ is non-singular. □

Theorem 2 *The following holds for the pseudoinverse of an $m \times n$ matrix A as defined in (4):*

$$A^\dagger = \begin{cases} (A^\top A)^{-1} A^\top & \textit{if } \operatorname{rank}(A) = n; \\ A^\top (A A^\top)^{-1} & \textit{if } \operatorname{rank}(A) = m. \end{cases} \tag{5}$$

Proof Consider the first situation where $\operatorname{rank}(A) = n$. we make use of the SVD of A:

$$A^\top A = V \Sigma^\top U^\top U \Sigma V^\top = V \Sigma^\top \Sigma V^\top,$$

where

$$\Sigma^\top \Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}.$$

By Lemma 1, $A^\top A$ is non-singular. Subsequently,

$$\begin{aligned}
(A^\top A)^{-1} A^\top &= V \left(\Sigma^\top \Sigma\right)^{-1} V^\top V \Sigma^\top U^\top \\
&= V \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{pmatrix} U^\top \\
&= V \begin{pmatrix} 1/\sigma_1 & & & \\ & \ddots & & 0 \\ & & 1/\sigma_n & \end{pmatrix} U^\top \\
&= V \Sigma^\dagger U^\top \\
&= A^\dagger.
\end{aligned}$$

Consider the second situation where $\operatorname{rank}(A) = m \leq n$. The product matrix $AA^\top$ is non-singular by Lemma 1 (if we substitute $A^\top$ for A). Now, we have

$$\begin{aligned}
A^\top(AA^\top)^{-1} &= V\Sigma^\top U^\top \left(U\Sigma V^\top V\Sigma^\top U^\top\right)^{-1} \\
&= V\Sigma^\top U^\top \left(U\Sigma\Sigma^\top U^\top\right)^{-1} \\
&= V\Sigma^\top U^\top U \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_m^2 \end{pmatrix} U^\top \\
&= V \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{pmatrix} \begin{pmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_m^2 \end{pmatrix} U^\top \\
&= V \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_m \\ & 0 & \end{pmatrix} U^\top \\
&= A^\dagger.
\end{aligned}$$

□

-۶

7.1.B Show that $\boldsymbol{\sigma_1 \geq |\lambda|_{\max}}$. The largest singular value dominates all eigenvalues.

Solution Start from $A = U\Sigma V^{\mathrm{T}}$. Remember that multiplying by an orthogonal matrix *does not change length*: $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$ because $\|Q\boldsymbol{x}\|^2 = \boldsymbol{x}^{\mathrm{T}}Q^{\mathrm{T}}Q\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} = \|\boldsymbol{x}\|^2$. This applies to $Q = U$ and $Q = V^{\mathrm{T}}$. In between is the diagonal matrix Σ.

$$\|A\boldsymbol{x}\| = \|U\Sigma V^{\mathrm{T}}\boldsymbol{x}\| = \|\Sigma V^{\mathrm{T}}\boldsymbol{x}\| \leq \sigma_1\|V^{\mathrm{T}}\boldsymbol{x}\| = \sigma_1\|\boldsymbol{x}\|. \tag{14}$$

An eigenvector has $\|A\boldsymbol{x}\| = |\lambda|\|\boldsymbol{x}\|$. So (14) says that $|\lambda|\|\boldsymbol{x}\| \leq \sigma_1\|\boldsymbol{x}\|$. Then $\boldsymbol{|\lambda| \leq \sigma_1}$.

Apply also to the unit vector $\boldsymbol{x} = (1, 0, \ldots, 0)$. Now $A\boldsymbol{x}$ is the first column of A. Then by inequality (14), this column has length $\leq \sigma_1$. Every entry must have $|a_{ij}| \leq \sigma_1$.

Equation (14) shows again that ***the maximum value of*** $\boldsymbol{||A\boldsymbol{x}||/||\boldsymbol{x}||}$ ***equals*** $\boldsymbol{\sigma_1}$.

۷-

4. 1. $\|A\|_F = \sqrt{Tr(A^TA)} = \sqrt{\sum_{j=1}^{n} (A^TA)_{j,j}} = \sqrt{\sum_{j=1}^{n} A_j^T A_j} = \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{m} |A_{ij}|^2}$

2. $A = U\Sigma V^T \Rightarrow \|A\|_F = \sqrt{Tr(A^TA)} = \sqrt{Tr(V\Sigma U^T U\Sigma V^T)}$

$= \sqrt{Tr(V\Sigma^2 V^T)} = \sqrt{Tr(\Sigma^2 V^T V)} = \sqrt{Tr(\Sigma^2)} = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$ ✓

3. $\sqrt{\sigma_1^2} \leq \sqrt{\sum_{i=1}^{r} \sigma_i^2} \leq \sqrt{\sum_{i=1}^{r} \sigma_1^2} \Rightarrow \sigma_1 \leq \|A\|_F \leq \sqrt{r\sigma_1^2} = \sqrt{r}\,\sigma_1$

$\Rightarrow \sigma_{max}(A) \leq \|A\|_F \leq \sqrt{r}\,\sigma_{max}(A)$ ✓

۸- ماتریس householder ماتریس نامعین است. زیرا

$$H = I - 2\frac{vv^T}{v^Tv}$$

$$v^THv = v^T(-v) = -\|v\|_2^2 < 0$$

از طرفی، اگر بردار w عمود بر v را در نظر بگیریم ($v^Tw = 0$):

$$w^THw = \|w\|_2^2 > 0$$

لذا ماتریس نامعین است.

۹- کتاب Golub، فصل 5.1.13