For discussion / Week 7 – Solutions

4.28 Recall the Frobenius inner product

$$\langle A, B\rangle_F = \operatorname{tr}(A^H B)$$

Then the Cauchy-Schwarz inequality gives

$$|\langle A, B\rangle_F| \le \|A^H\|_F \, \|B\|_F$$

If A, B are pos. semi-definite, they are orthonormally diagonalizable, i.e.

$$A = UDU^H$$

where the diagonal matrix D has non-negative entries.

$$\Rightarrow \quad \|A^H\|_F^2 = \|A\|_F^2 = \operatorname{tr}(A^H A) = \operatorname{tr}(UD^2U^H) = \operatorname{tr} D^2$$

$$\le (\operatorname{tr} D)^2$$ (the last expression contains additional non-negative crossterms)

$$\Rightarrow \quad \|A^H\|_F \le \operatorname{tr} D = \operatorname{tr} A$$

Similarly, $\|B\|_F \le \operatorname{tr} B$

$$\Rightarrow \quad \underbrace{|\langle A, B\rangle|_F}_{= \operatorname{tr} AB} \le \operatorname{tr} A \ \operatorname{tr} B$$

To prove that $tr(AB) \geq 0$, note that $A^{\frac{1}{2}}$ is well-defined since A is pos. semidefinite.

Thus, $tr(AB) = tr(A^{\frac{1}{2}} B A^{\frac{1}{2}})$.

But $A^{\frac{1}{2}} B A^{\frac{1}{2}}$ is positive semidefinite:

$$\langle x, A^{\frac{1}{2}} B A^{\frac{1}{2}} x\rangle = \langle \underbrace{A^{\frac{1}{2}} x}_{=:y}, B \underbrace{A^{\frac{1}{2}} x}_{=y} \rangle$$ since $A^{\frac{1}{2}}$ also Hermitian

$$\geq 0$$ since B pos. semidefinit

But the trace of a pos. semidefinite matrix is non-negative.

$$\Rightarrow tr(AB) \geq 0$$

4.30: From earlier, the matrix representation is

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow A^H A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

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

Set $\Sigma_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Then, since $A^H A$ is already diagonal, we only need to permute to write

$$A^H A = \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}}_{=: V_1} \Sigma_1^2 \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{= V_1^H}$$

Now set $U_1 = A V_1 \Sigma_1^{-1}$

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

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

We easily complete U_1 and V_1 to full orthonormal matrices

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Then

$$A = U \Sigma V^H \qquad \text{with } \Sigma = \begin{pmatrix} \Sigma_1 & \\ & 0 \end{pmatrix} = \begin{pmatrix} 2 & & \\ & 1 & \\ & & 0 \end{pmatrix}$$

<u>Check:</u>

$$U \Sigma V^H = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{= \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}$$

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

4.31: (i) Let $A^H A = U \Sigma^2 U^H$

$$\Rightarrow \|A\|_2^2 = \sup_{\|x\|_2=1} \|Ax\|_2^2 = \sup_{\|x\|_2=1} x^H A^H A x$$

$$= \sup_{\|x\|_2=1} \underbrace{x^H U}_{=(U^H x)^H = y^H} \Sigma^2 \underbrace{U^H x}_{=: y}$$

$$= \sup_{\|y\|_2=1} y^H \Sigma^2 y$$ since U^H is a length-preserving isomorphism

$$= \sup_{\|y\|_2=1} \underbrace{\sum_{i=1}^{n} |y_i|^2 \sigma_i^2}_{\text{This is largest for } \|y\|_2=1 \text{ when } y=e_1 \text{ because } \sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2 \geq 0.}$$

$$\Rightarrow \|A\|_2^2 = \sigma_1^2$$

(ii) If A is invertible and $A = U \Sigma V^H$,

we must have that Σ has full rank, i.e. $\sigma_1 \geq \dots \geq \sigma_n > 0$,

so that Σ is invertible and

$$A^{-1} = V \Sigma^{-1} U^H .$$

Then σ_n^{-1} is the largest singular value of A^{-1}, so that, by (i), $\|A^{-1}\|_2 = \sigma_n^{-1}$.

(iii) Write $A = U \Sigma V^H$,

then $AA^H = U \tilde{\Sigma}^2 U^H$ where $\Sigma, \tilde{\Sigma}$ have the same non-zero diagonal entries

$A^H A = V \Sigma^2 V^H$

$A^T = \bar{V} \Sigma^T \bar{U}^H$

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

$\Rightarrow$ The singular values of A, A^T, A^H are the same, and their squares are the singular values of AA^H and $A^H A$ (except for possible a different number of zero singular values if A is not square)

$\Rightarrow$ By (i), their 2-norms are the same resp. the square of $\|A\|_2$.

(iv) If U, V are orthonormal, then A and UAV have the same singular values, hence the same 2-norm.

5.1 (i) $d_1 + d_2$ is a metric:

- $\underbrace{d_1(x,y)}_{\geq 0} + \underbrace{d_2(x,y)}_{\geq 0} \geq 0$

  If $x \neq y$, $d_1(x,y) > 0 \Rightarrow d_1(x,y) + d_2(x,y) > 0$

  $x = y$, $d_1(x,x) = d_2(x,x) = 0 \Rightarrow d_1(x,x) + d_2(x,x) = 0$

- symmetry is obvious
- $\big(d_1(x,y) + d_2(x,y)\big) + \big(d_1(y,z) + d_2(y,z)\big)$

$$= \underbrace{\big(d_1(x,y) + d_1(y,z)\big)}_{\geq d_1(x,z)} + \underbrace{\big(d_2(x,y) + d_2(y,z)\big)}_{\geq d_2(x,z)}$$

$$\geq d_1(x,z) + d_2(x,z)$$

(ii) not a metric in general, as, e.g., positivity may fail.

But could be a metric in special cases, e.g. if $d_2 = \frac{1}{2} d_1$

(iii) For the triangle inequality to hold, need

$$\min\{d_1(x,z), d_2(x,z)\} \leq \min\{d_1(x,y) + d_1(y,z), d_2(x,y) + d_2(y,z)\}$$

$$\cancel{\leq} \min\{d_1(x,y), d_2(x,y)\} + \min\{d_1(y,z), d_2(y,z)\}$$

↑ not true in general, e.g. $\min\{0+1, 1+0\} = 1$

$\min\{0,1\} + \min\{1,0\} = 0$

So not a metric in general, but could be in special cases (e.g. $d_1 = d_2$)

(iv) The same argument as in (iii) does work for the maximum, though. So the triangle inequality holds and pos. definiteness and symmetry are obvious.

$\Rightarrow \max\{d_1, d_2\}$ is a metric.

5.9: (X,d) metric space, $B \subset X$ nonempty.

$$\delta(B,x) := \inf\{d(x,b) : b \in B\}$$

Let $b \in B$. Then, by the triangle inequality,

$$d(b,x) - d(x,y) \leq d(b,y) \leq d(b,x) + d(x,y)$$

Taking the inf over all $b \in B$ preserves the inequalities, so

$$\inf_{b\in B}(d(b,x) - d(x,y)) \leq \inf_{b\in B} d(b,y) \leq \inf_{b\in B}(d(b,x) + d(x,y))$$

$$\Rightarrow \delta(B,x) - d(x,y) \leq \delta(B,y) \leq \delta(B,x) + d(x,y)$$

$$\left.\begin{array}{l} \Rightarrow \delta(B,x) - \delta(B,y) \leq d(x,y) \\ \text{and } \delta(B,y) - \delta(B,x) \leq d(x,y) \end{array}\right\} |\delta(B,x) - \delta(B,y)| \leq d(x,y)$$

$\Rightarrow x \mapsto \delta(B,x)$ is continuous.