OSM Boot Camp: **Math ProbSet3**

Harrison Beard

9 July 2018

Exercise 4.2.

Question. Let $V = \text{span}\left(\{1, x, x^2\}\right)$ be a subspace of the inner product space $L^2\left([0,1];\mathbb{R}\right)$. Let $\mathbf{D}$ be the derivative operator $\mathbf{D} : V \to V$ given by $\mathbf{D}[p](x) = p'(x)$. Find all the eigenvalues and eigenspaces of $\mathbf{D}$. What are their algebraic and geometric multiplicities?

Answer.

We can see that $\mathbf{D} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ is upper-triangular, so all its eigenvalues are $\boxed{0}$. In other words, the algebraic multiplicity of $\lambda = 0$ is $\boxed{3}$. Moreover, $\Sigma_0(\mathbf{D}) = \boxed{\text{span}\left(\{1\}\right)}$, so the geometric multiplicity of $\lambda = \dim \Sigma_0(\mathbf{D}) = 0$ is $\boxed{1}$.

Exercise 4.4.

Question. Recall that a matrix $A \in M_n(\mathbb{F})$ is Hermitian if $A^{\mathsf{H}} = A$ and skew-Hermitian if $A^{\mathsf{H}} = -A$. Using Exercise 4.3, prove that

1. a Hermitian 2×2 matrix has only real eigenvalues;
2. a skew-Hermitian 2×2 matrix has only imaginary eigenvalues.

Answer.

1. $A \in M_{2\times 2}(\mathbb{F})$ is Hermitian so a_{11} and a_{22} are their own conjugates ($\in \mathbb{R}$) and $a_{12} = \bar{a_{21}}$. Thus, $a_{12}a_{21} = \|a_{21}\|^2 \in \mathbb{R}$, and the char poly is

$$\lambda^2 - \text{tr}\,(A)\,\lambda + \det A = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - \|a_{21}\|^2\,,$$

so

$$\begin{aligned} \lambda &= \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} + a_{22})^2 - 4\left(a_{11}a_{22} - \|a_{21}\|^2\right)}}{2} \\ &= \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + \|a_{21}\|^2}}{2}, \end{aligned}$$

where $(a_{11} - a_{22})^2 + \|a_{21}\|^2 \geq 0$ so $\lambda \in \mathbb{R}$. ■

2. $A \in M_{2\times 2}(\mathbb{F})$ is *skew-Hermitian*, so $a_{11} = -\bar{a}_{11}$, $a_{22} = -\bar{a}_{22}$, $a_{12} = -\bar{a}_{21}$, so $a_{11}, a_{22} \in \mathbb{C} \setminus \mathbb{R}$ and $a_{12}a_{21} = -\|a_{21}\|^2$, and $(a_{11}a_{22})/i < 0$. The char poly is

$$\lambda^2 - \operatorname{tr}(A)\,\lambda + \det(A) = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} + \|a_{21}\|^2,$$

so

$$\begin{aligned}\lambda &= \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} + \|a_{21}\|^2)}}{2} \\ &= \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + \|a_{21}\|^2}}{2},\end{aligned}$$

where $(a_{11} - a_{22})^2 + \|a_{21}\|^2 < 0$, so $a_{11}, a_{12}, a_{21}, a_{22}, \lambda \in \mathbb{C} \setminus \mathbb{R}$. ■

Exercise 4.6.

Question. Prove Proposition 4.1.22.
Prop 4.1.22. *The diagonal entries of an upper-triangular (or a lower-triangular) matrix are its eigenvalues.*

Answer.

- *WLOG*, suppose A is upper-triangular, so then $\lambda I - A$ is also upper-triangular with diagonals $\{\lambda - a_{i,i}\}_{i=1}^{n}$. For the case $n = 1$, the char poly is

$$\det(\lambda - A) = \prod_{i=1}^{1} \lambda - a_{i,i}.$$

For the induction step,

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}).$$

Note $a_{i,1} = \delta_{i,1}$ where δ is the Kronecker delta, and $A_{1,1} \in M_{n-1}(\mathbb{F})$ is upper-triangular. So,

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det(A_{i,1}) \\ &= a_{1,1} \det(A_{1,1}) \\ &= (\lambda - a_{1,1}) \prod_{i=1}^{n-1} \left(\lambda - (a_{A_{11},A_{11}}) \cdot i\right) \\ &= (\lambda - a_{1,1}) \prod_{i=2}^{n} (\lambda - a_{i,i}) \\ &= \prod_{i=1}^{n} (\lambda - a_{i,i}).\end{aligned}$$

■

Exercise 4.8.

Question. Let V be the span of the set $S = \{\sin(x), \cos(x), \sin(2x), \cos(2x)\}$ in the vector space $\mathcal{C}^\infty(\mathbb{R};\mathbb{R})$.

1. Prove that S is a basis for V.
2. Let $\mathbf{D}$ be the derivative operator. Write the matrix representation of $\mathbf{D}$ in the basis S.
3. Find two complementary $\mathbf{D}$-invariant subspaces in V.

Answer.

1. We showed in the last homework that S is orthonormal as defined under the inner product $\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,\mathrm{d}t$. That is, *all elements in the spanning set are independent*, so they form a basis for the span. ∎
2. We know from elementary calculus that $\begin{cases} \mathbf{D}\sin(x) = \cos(x) \\ \mathbf{D}\cos(x) = -\sin(x) \\ \mathbf{D}\sin(2x) = 2\cos(x) \\ \mathbf{D}\cos(2x) = -2\sin(x) \end{cases}$, so it follows straightforwardly that
$$\mathbf{D} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 2 & 0 \end{pmatrix}.$$ ∎
3. $\boxed{\operatorname{span}(\{\sin(x), \cos(x)\})}$ and $\boxed{\operatorname{span}(\{\sin(2x), \cos(2x)\})}$.

Exercise 4.13.

Question. Let
$$A = \begin{pmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{pmatrix}.$$
Compute the transition matrix P such that $P^{-1}AP$ is diagonal.

Answer.

Note $\det(\lambda I - A) = \lambda^2 - 1.4\lambda + 0.4$, so $\lambda \in \{1, 0.4\}$. Thus, $P = \begin{pmatrix} \vec{\mathbf{v}}_1 & \vec{\mathbf{v}}_{0.4} \end{pmatrix} = \boxed{\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}}$.

Exercise 4.15.

Question. Prove Theorem 4.3.12.
Thm 4.3.12. (Semisimple Spectral Mapping). *If $(\lambda_i)_{i=1}^n$ are the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$ and $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial, then $(f(\lambda_i))_{i=1}^n$ are the eigenvalues of $f(A) = a_0 I + a_1 A + \cdots + a_n A^n$.*

Answer.

Note that

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= a_0 P P^{-1} + a_1 P \Lambda P^{-1} + \cdots + a_n P \Lambda^n P^{-1} \\ &= P f(\Lambda) P^{-1}, \end{aligned}$$

where every term in $f(\Lambda)$ is diagonal itself, so the diagonals entries are simply $(f(\lambda_i))_{i=1}^n$. But $f(\Lambda)$ is similar to $f(A)$, so their eigenvalues are both $(f(\lambda_i))_{i=1}^n$. ■

Exercise 4.16.

Question. Let A be the matrix in Exercise 4.13 above.

1. Compute $\lim_{n\to\infty} A^n$ with respect to the 1-norm; that is, find a matrix B such that for any $\varepsilon > 0$ there exists an $N > 0$ with $\left\|A^k - B\right\| < \varepsilon$ whenever $k < N$. *Hint*: Use Proposition 4.3.10. **Prop 4.3.10.** *If matrices* $A, B \in M_n(\mathbb{F})$ *are similar, with* $A = P^{-1}BP$, *then* $A^k = P^{-1}B^kP$ *for all* $k \in \mathbb{N}$.
2. Repeat part (1) for the ∞-norm and the Frobenius norm. Does the answer depend on the choice of norm? We discuss this further in Section 5.8.
3. Find all the eigenvalues of the matrix $3I + 5A + A^3$. *Hint*: Consider using Theorem 4.3.12 (see above for the theorem).

Answer.

1. $A^n = P\Lambda^n P^{-1}$, where

$$\Lambda^n = \begin{pmatrix} 1 & 0 \\ 0 & 0.4^n \end{pmatrix},$$

such that

$$\begin{aligned} A^k &= \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.4^k \end{pmatrix} \cdot \frac{1}{3} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \\ &= \frac{1}{3} \cdot \begin{pmatrix} 2 + 0.4^k & 2 - 2\cdot 0.4^k \\ 1 - 0.4^k & 1 + 2 \cdot 0.4^k \end{pmatrix} \end{aligned}$$

with limit

$$B = \boxed{\frac{1}{3} \cdot \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}}.$$

Also note that $A^k - B = \frac{1}{3}\begin{pmatrix} 0.4^k & -2\cdot 0.4^k \\ -0.4^k & 2 \cdot 0.4^k \end{pmatrix}$ converges entry-wise by the 1-norm.

2. $\boxed{\text{The answer does } \textit{not} \text{ depend on choice of norm}}$: The ∞-norm goes to $\boxed{\text{zero}}$, as above, since the

maximum entry goes to zero; the Frobenius norm is

$$\begin{aligned}\|A^k - B\| &= \frac{1}{3}\sqrt{\operatorname{tr}\left(\begin{pmatrix} 0.4^k & -0.4^k \\ -2\cdot 0.4^k & 2\cdot 0.4^k \end{pmatrix}\begin{pmatrix} 0.4^k & -2\cdot 0.4^k \\ -0.4^k & 2\cdot 0.4^k \end{pmatrix}\right)} \\ &= \frac{1}{3}\sqrt{\operatorname{tr}\left(\begin{pmatrix} 2\cdot 0.4^{2k} & -4\cdot 0.4^{2k} \\ -4\cdot 0.4^{2k} & 8\cdot 0.4^{2k} \end{pmatrix}\right)} \\ &= \sqrt{10\cdot 0.4^{2k}},\end{aligned}$$

which also goes to $\boxed{\text{zero}}$.

3. We have $\boxed{f(1) = 9}$ and $\boxed{f(0.4) = 5.064}$, since the original eigenvalues were 1 and 0.4, and $f(x) = 3 + 5x + x^3$ where $f(A)$ are simply the eigenvalues of A.

Exercise 4.18.

Question. Prove: If λ is an eigenvalue of the $A \in M_n(\mathbb{F})$, then there exists a nonzero row vector $\vec{\mathbf{x}}^{\mathsf{T}}$ such that $\vec{\mathbf{x}}^{\mathsf{T}} A = \lambda \vec{\mathbf{x}}^{\mathsf{T}}$.

Answer.
$\det(\lambda I - A) = \det(\lambda I - A^{\mathsf{T}})$, so all eigenvalues $\{\lambda_i\}$ of A are an eigenvalues of A^{T}. So we know there must be a $\vec{\mathbf{v}} \in \mathbb{F}^n$ such that $A^{\mathsf{T}}\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, so it follows that $\vec{\mathbf{v}}^{\mathsf{T}} A = \lambda\vec{\mathbf{v}}^{\mathsf{T}}$. ■

Exercise 4.20.

Question. Prove Lemma 4.4.2.
Lemma 4.4.2. *If A is Hermitian and orthonormally similar to B, then B is also Hermitian.*

Answer.
From the given we know that $B = PAP^{-1}$ for some orthonormal P. Note that A is Hermitian so we have

$$\begin{aligned} B^{\mathsf{H}} &= \left(PAP^{-1}\right)^{\mathsf{H}} \\ &= \left(PAP^{\mathsf{H}}\right)^{\mathsf{H}} \\ &= \left(P^{\mathsf{H}}\right) A^{\mathsf{H}} P^{\mathsf{H}} \\ &= PAP^{\mathsf{H}} \\ &= B. \end{aligned}$$

■

Exercise 4.24.

Question. Given $A \in M_n(\mathbb{C})$, define the *Rayleigh quotient* as

$$\rho(\vec{\mathbf{x}}) = \frac{\langle \vec{\mathbf{x}}, A\vec{\mathbf{x}} \rangle}{\|\vec{\mathbf{x}}\|^2},$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{F}^n$. Show that the Rayleigh quotient can only take on real values for Hermitian matrices and only imaginary values for skew-Hermitian matrices.

Answer.
$\|\vec{\mathbf{x}}\|^2 \in \mathbb{R}$, so it suffices to show that $\langle \vec{\mathbf{x}}, A\vec{\mathbf{x}} \rangle = \vec{\mathbf{x}}^{\mathsf{H}} A\vec{\mathbf{x}}$:

$$\begin{aligned} \overline{\vec{\mathbf{x}}^{\mathsf{H}} A\vec{\mathbf{x}}} &= \left(\vec{\mathbf{x}}^{\mathsf{H}} A\vec{\mathbf{x}}\right)^{\mathsf{H}} \\ &= \vec{\mathbf{x}}^{\mathsf{H}} A^{\mathsf{H}} \vec{\mathbf{x}} \\ &= \begin{cases} \vec{\mathbf{x}}^{\mathsf{H}} A\vec{\mathbf{x}} & (\text{Hermitian, } \langle \vec{\mathbf{x}}, A\vec{\mathbf{x}} \rangle \in \mathbb{R}) \\ -\vec{\mathbf{x}}^{\mathsf{H}} A\vec{\mathbf{x}} & (\text{skew-Hermitian, } \langle \vec{\mathbf{x}}, A\vec{\mathbf{x}} \rangle \in \mathbb{C} \setminus \mathbb{R}). \end{cases} \end{aligned}$$

■

Exercise 4.25.

Question. Let $A \in M_n(\mathbb{C})$ be a normal matrix with eigenvalues $(\lambda_1, \dots \lambda_n)$ and corresponding orthonormal eigenvectors $(\vec{\mathbf{x}}_1, \dots \vec{\mathbf{x}}_n)$.

1. Show that the identity matrix can be written $I = \vec{\mathbf{x}}_1 \vec{\mathbf{x}}_1^{\mathsf{H}} + \cdots + \vec{\mathbf{x}}_n \vec{\mathbf{x}}_n^{\mathsf{H}}$. *Hint*: What is $\left(\vec{\mathbf{x}}_1 \vec{\mathbf{x}}_1^{\mathsf{H}} + \cdots + \vec{\mathbf{x}}_n \vec{\mathbf{x}}_n^{\mathsf{H}}\right) \vec{\mathbf{x}}_j$?
2. Show that A can be written $A = \lambda_1 \vec{\mathbf{x}}_1 \vec{\mathbf{x}}_1^{\mathsf{H}} + \cdots + \lambda_n \vec{\mathbf{x}}_n \vec{\mathbf{x}}_n^{\mathsf{H}}$. This is called an *outer product expansion.*

Answer.

1. By orthonormality, $\vec{\mathbf{x}}_j^{\mathsf{H}} \vec{\mathbf{x}}_i = 0$ if $i \neq j$, so consider $\left(\sum_{i=1}^n \vec{\mathbf{x}}_i \vec{\mathbf{x}}_i^{\mathsf{H}}\right) \vec{\mathbf{y}} = \vec{\mathbf{y}}$ for some $\vec{\mathbf{y}}$:

$$\begin{aligned} \left(\sum_{i=1}^n \vec{\mathbf{x}}_i \vec{\mathbf{x}}_i^{\mathsf{H}}\right)\left(\sum_{i=1}^n \alpha_i \vec{\mathbf{x}}_i\right) &= \sum_{i=1}^n \vec{\mathbf{x}}_i \vec{\mathbf{x}}^{\mathsf{H}} \vec{\mathbf{x}}_i \alpha_i \\ &= \sum_{i=1}^n \vec{\mathbf{x}}_i \alpha_i \\ &= \underbrace{\sum_{i=1}^n \alpha_i \vec{\mathbf{x}}_i}_{\vec{\mathbf{y}}}, \end{aligned}$$

i.e., $\sum_{i=1}^n \vec{\mathbf{x}}_i \vec{\mathbf{x}}_i^{\mathsf{H}} = I$. ■

2.

$$\begin{aligned} A\vec{\mathbf{y}} &= A\left(\sum_{i=1}^{n}\alpha_i\vec{\mathbf{x}}_i\right) \\ &= \sum_{i=1}^{n}\alpha_i A\vec{\mathbf{x}}_i \\ &= \sum_{i=1}^{n}\alpha_i\lambda_i\vec{\mathbf{x}}_i \end{aligned}$$

and

$$\begin{aligned} \underbrace{\left(\sum_{i=1}^{n}\lambda_i\vec{\mathbf{x}}_i\vec{\mathbf{x}}_i^{\mathsf{H}}\right)}_{A}\vec{\mathbf{y}} &= \left(\sum_{i=1}^{n}\lambda_i\vec{\mathbf{x}}_i\vec{\mathbf{x}}_i^{\mathsf{H}}\right)\left(\sum_{i=1}^{n}\alpha_i\vec{\mathbf{x}}_i\right) \\ &= \sum_{i=1}^{n}\lambda_i\vec{\mathbf{x}}_i\vec{\mathbf{x}}_i^{\mathsf{H}}\vec{\mathbf{x}}_i\alpha_i \\ &= \sum_{i=1}^{n}\alpha_i\lambda_i\vec{\mathbf{x}}_i \end{aligned}$$

■

Exercise 4.27.

Question. Assume $A \in M_n(\mathbb{F})$ is positive definite. Prove that all its diagonal entries are real and positive.

Answer.
By positive definiteness of A, we have that $\underbrace{\vec{\mathbf{e}}_i^{\mathsf{H}} A\vec{\mathbf{e}}_i}_{a_{ii}\in\mathbb{R}} > 0$ for $\vec{\mathbf{e}}_i \in \operatorname{basis}(A)$, orthonormal. ■

Exercise 4.28.

Question. Assume $A, B \in M_n(\mathbb{F})$ are positive semidefinite. Prove that

$$0 \leq \operatorname{tr}(AB) \leq \operatorname{tr}(A)\operatorname{tr}(B),$$

and use this result to prove that $\|\cdot\|_{\mathrm{F}}$ is a matrix norm.

Answer.
By positive semidefiniteness, there exist $S_A, S_A^{\mathsf{H}}, S_B, S_B^{\mathsf{H}}$ (where $S_A S_B^{\mathsf{H}}$ is Hermitian) such that

$$S_A^{\mathsf{H}} S_A = A$$

and

$$S_B^{\mathsf{H}} S_B = B,$$

so

$$\begin{aligned}
\operatorname{tr}(AB) &= \operatorname{tr}\left(S_A^{\mathsf{H}} S_A S_B^{\mathsf{H}} S_B\right) \\
&= \operatorname{tr}\left(S_B S_A^{\mathsf{H}} S_A S_B^{\mathsf{H}}\right) \\
&= \operatorname{tr}\left(\left(S_A S_B^{\mathsf{H}}\right)^{\mathsf{H}} S_A S_B^{\mathsf{H}}\right) \\
&\geq 0.
\end{aligned}$$

Diagonalizing A and B as $A = P_A D_A P_A^{-1}$, we have

$$\begin{aligned}
\operatorname{tr}(A) &= \operatorname{tr}\left(P_A D_A P_A^{-1}\right) \\
&= \operatorname{tr}\left(P_A^{-1} P_A D_A\right) \\
&= \operatorname{tr}(D_A) \\
&= \sum_i \lambda_{A_i}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\operatorname{tr}(AB) &= \operatorname{tr}\left(P_A D_A P_A^{-1} P_B D_B P_B^{-1}\right) \\
&= \operatorname{tr}\left(P_A P_A^{-1} P_B D_A D_B P_B^{-1}\right) \\
&= \operatorname{tr}\left(P_B^{-1} P_B D_A D_B\right) \\
&= \operatorname{tr}(D_A D_B) \\
&= \sum_i \lambda_{A_i} \lambda_{B_i} \\
&\leq \left(\sum_i \lambda_{A_i}\right)\left(\sum_i \lambda_{B_i}\right) \\
&= \operatorname{tr}(A) \operatorname{tr}(B).
\end{aligned}$$

$\|\cdot\|_F$ is a matrix norm because $\|A\| = \sqrt{\operatorname{tr}(A^{\mathsf{H}} A)} \geq 0$ ("=" only if all diagonals of $A^{\mathsf{H}} A$ are 0, i.e., $A = \vec{\mathbf{0}}$), and

$$\begin{aligned}
\|\alpha A\| &= \sqrt{\operatorname{tr}(\alpha^{\mathsf{H}} A^{\mathsf{H}} A \alpha)} \\
&= \alpha \sqrt{\operatorname{tr}(A^{\mathsf{H}} A)} \\
&= \alpha \|A\|,
\end{aligned}$$

and finally:

$$\begin{aligned}
\|A + B\|_{\mathrm{F}}^2 &= \operatorname{tr}\left((A+B)^{\mathsf{H}}(A+B)\right) \\
&= \operatorname{tr}\left(A^{\mathsf{H}} A + B^{\mathsf{H}} B + A^{\mathsf{H}} B + A^{\mathsf{H}} B\right) \\
&= \operatorname{tr}\left(A^{\mathsf{H}} A\right) + \operatorname{tr}\left(B^{\mathsf{H}} B\right) + \operatorname{tr}\left(A^{\mathsf{H}} B + A^{\mathsf{H}} B\right) \\
&\leq \operatorname{tr}\left(A^{\mathsf{H}} A\right) + \operatorname{tr}\left(B^{\mathsf{H}} B\right) + 2\|A\|_{\mathrm{F}} \|B\|_{\mathrm{F}} \\
&= \|A\|_{\mathrm{F}}^2 + \|B\|_{\mathrm{F}}^2 + 2\|A\|_{\mathrm{F}} \|B\|_{\mathrm{F}} \\
&= \left(\|A\|_{\mathrm{F}} + \|B\|_{\mathrm{F}}\right)^2.
\end{aligned}$$

■

Exercise 4.31.

Question. Assume $A \in M_{m\times n}(\mathbb{F})$ and A is not identically zero. Prove that

1. $\|A\|_2 = \sigma_1$, where σ_1 is the largest singular value of A;
2. If A is invertible, then $\left\|A^{-1}\right\|_2 = \sigma_n^{-1}$;
3. $\left\|A^{\mathsf{H}}\right\|_2^2 = \left\|A^{\mathsf{T}}\right\|_2^2 = \left\|A^{\mathsf{H}}A\right\|_2 = \|A\|_2^2$;
4. if $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_2 = \|A\|_2$.

Answer.

1. $A^{\mathsf{H}}A$ normal, so there exist orthonormal eigenvectors $\{\vec{\mathbf{v}}_i\}_{i=1}^n$ with eigenvalues $\left\{\sigma_i^2\right\}_{i=1}^n$:

$$\begin{aligned}
\|A\|_2 &= \sup_{\|\vec{\mathbf{u}}\|=1} \|A\vec{\mathbf{u}}\| \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\langle A\vec{\mathbf{u}}, A\vec{\mathbf{u}}\rangle} \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\langle \vec{\mathbf{u}}, A^{\mathsf{H}}A\vec{\mathbf{u}}\rangle} \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\left\langle \sum_{i=1}^n \alpha_i\vec{\mathbf{v}}_i, \sum_{i=1}^n \alpha_i\sigma_i^2\vec{\mathbf{v}}_i \right\rangle} \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\left\langle \sum_{i=1}^n \alpha_i\vec{\mathbf{v}}_i, \sum_{i=1}^n \alpha_i\sigma_i^2\vec{\mathbf{v}}_i \right\rangle} \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\sum_{i=1}^n |\alpha_i|^2\,\sigma_i^2} \\
&= \sqrt{\sigma_i^2} \\
&= \sigma_1,
\end{aligned}$$

where $\vec{\mathbf{u}} = \vec{\mathbf{v}}_1$. ■

2. $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$, so $A^{-1}\vec{\mathbf{v}} = \frac{1}{\lambda}\vec{\mathbf{v}}$:

$$\begin{aligned}
\left\|A^{-1}\right\|_2 &= \sup_{\|\vec{\mathbf{u}}\|=1} \left\|A^{-1}\vec{\mathbf{u}}\right\| \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\langle A^{-1}\vec{\mathbf{u}}, A^{-1}\vec{\mathbf{u}}\rangle} \\
&= \sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\left\langle \vec{\mathbf{u}}, \left(A^{\mathsf{H}}A\right)^{-1}\vec{\mathbf{u}}\right\rangle},
\end{aligned}$$

so

$$\sup_{\|\vec{\mathbf{u}}\|=1} \sqrt{\sum_{i=1}^n |\alpha_i|^2\,\frac{1}{\sigma_i^2}} = \sigma_n^{-1},$$

maximizing with respect to $\alpha_n = 1$, $\vec{\mathbf{u}} = \vec{\mathbf{v}}_n$. ■

3. $A = U\Sigma V^{\mathsf{H}}$, for U, V orthonormal:

$$\begin{aligned} A^{\mathsf{H}} &= \left(U\Sigma V^{\mathsf{H}}\right)^{\mathsf{H}} \\ &= V\Sigma^{\mathsf{H}}U^{\mathsf{H}} \\ &= V\Sigma U^{\mathsf{H}}, \end{aligned}$$

$$\begin{aligned} A^{\mathsf{T}} &= \left(U\Sigma V^{\mathsf{H}}\right)^{\mathsf{T}} \\ &= \bar{V}\Sigma^{\mathsf{T}}U^{\mathsf{T}} \\ &= \bar{V}\Sigma U^{\mathsf{T}}, \end{aligned}$$

for $\bar{V}, U^{\mathsf{T}}$ orthonormal. Moreover,

$$\begin{aligned} A^{\mathsf{H}}A &= \left(U\Sigma V^{\mathsf{H}}\right)^{\mathsf{H}} U\Sigma V^{\mathsf{H}} \\ &= V\Sigma^{\mathsf{H}}U^{\mathsf{H}}U\Sigma V^{\mathsf{H}} \\ &= V\Sigma^{\mathsf{H}}U^{\mathsf{H}}U\Sigma V^{\mathsf{H}} \\ &= V\tilde{\Sigma}^2 V^{\mathsf{H}}, \end{aligned}$$

where $\tilde{\Sigma}$ is a square matrix with the singular values along the diagonal and zeros elsewhere. So by (4) we have $\left\|A^{\mathsf{H}}\right\|_2^2 = \left\|A^{\mathsf{T}}\right\|_2^2 = \|A\|_2^2$, and

$$\begin{aligned} \left\|A^{\mathsf{H}}A\right\|_2 &= \left\|V\tilde{\Sigma}^2V^{\mathsf{H}}\right\| \\ &= \sup_{\|\vec{\mathbf{v}}\|=1} \left\|V\tilde{\Sigma}^2V^{\mathsf{H}}\vec{\mathbf{v}}\right\| \\ &= \sup_{\|V^{\mathsf{H}}\vec{\mathbf{v}}\|=1} \left\|V^{\mathsf{H}}V\tilde{\Sigma}^2V^{\mathsf{H}}\vec{\mathbf{v}}\right\| \\ &= \sup_{\|\tilde{\vec{\mathbf{v}}}\|=1} \left\|\Sigma^2\tilde{\vec{\mathbf{v}}}\right\| \\ &= \sup_{\|\vec{\mathbf{v}}\|=1} \sqrt{\sum_{i=1}^{n} |\alpha_i|^2 \sigma_i^4} \\ &= \|A\|_2^2. \end{aligned}$$

■

4. W orthonormal with full rank, so $W^{-1} = W^{\mathsf{H}}$ and W bijective:

$$\begin{aligned} \|UAV\|_2 &= \sup_{\|\vec{\mathbf{v}}\|=1} \sqrt{\langle UAV\vec{\mathbf{v}}, UAV\vec{\mathbf{v}}\rangle} \\ &= \sup_{\|\vec{\mathbf{v}}\|=1} \sqrt{\langle AV\vec{\mathbf{v}}, U^{\mathsf{H}}UAV\vec{\mathbf{v}}\rangle} \\ &= \sup_{\|\vec{\mathbf{v}}\|=1} \sqrt{\langle AV\vec{\mathbf{v}}, AV\vec{\mathbf{v}}\rangle} \\ &= \sup_{\|\tilde{\vec{\mathbf{v}}}\|=1} \sqrt{\left\langle A\tilde{\vec{\mathbf{v}}}, A\tilde{\vec{\mathbf{v}}}\right\rangle} \\ &= \|A\|_2. \end{aligned}$$

■

Exercise 4.32.

Question. Assume $A \in M_{m\times n}(\mathbb{F})$ is of rank r. Prove that

1. if $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal, then $\|UAV\|_F = \|A\|_F$;
2. $\|A\|_{\mathrm{F}} = \left(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2\right)^{1/2}$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the singular values of A.

Answer.

1.

$$\begin{aligned}
\|UAV\|_{\mathrm{F}} &= \sqrt{\operatorname{tr}\left(V^{\mathsf{H}}A^{\mathsf{H}}U^{\mathsf{H}}UAV\right)} \\
&= \sqrt{\operatorname{tr}\left(V^{\mathsf{H}}A^{\mathsf{H}}AV\right)} \\
&= \sqrt{\operatorname{tr}\left(A^{\mathsf{H}}AVV^{\mathsf{H}}\right)} \\
&= \sqrt{\operatorname{tr}\left(A^{\mathsf{H}}A\right)} \\
&= \|A\|_{\mathrm{F}}\,.
\end{aligned}$$

■

2.

$$\begin{aligned}
\|A\|_{\mathrm{F}} &= \left\|U\Sigma V^{\mathsf{H}}\right\|_{\mathrm{F}} \\
&= \|\Sigma\|_{\mathrm{F}} \\
&= \sqrt{\operatorname{tr}\left(\Sigma^{\mathsf{H}}\Sigma\right)} \\
&= \sqrt{\sum_{i=1}^{n}\sigma_i^2}.
\end{aligned}$$

■

Exercise 4.33.

Question. Assume $A \in M_n(\mathbb{F})$. Prove that

$$\|A\|_2 = \sup_{\substack{\|\vec{\mathbf{x}}\|_2 = 1 \\ \|\vec{\mathbf{y}}\|_2 = 1}} \left|\vec{\mathbf{y}}^{\mathsf{H}}A\vec{\mathbf{x}}\right|.$$

Hint: Use Exercise 4.31 (above).

Answer.

$$\begin{aligned}\sup_{\substack{\|\vec{\mathbf{x}}\|_2=1\\ \|\vec{\mathbf{y}}\|_2=1}} \left|\vec{\mathbf{y}}^{\mathsf{H}}A\vec{\mathbf{x}}\right| &= \sup_{\substack{\|\vec{\mathbf{x}}\|_2=1\\ \|\vec{\mathbf{y}}\|_2=1}} \left|\vec{\mathbf{y}}^{\mathsf{H}}U\Sigma V^{\mathsf{H}}\vec{\mathbf{x}}\right| \\ &= \sup_{\substack{\left\|\tilde{\vec{\mathbf{x}}}\right\|_2=1\\ \left\|\tilde{\vec{\mathbf{y}}}\right\|_2=1}} \left|\tilde{\vec{\mathbf{y}}}^{\mathsf{H}}\Sigma\tilde{\vec{\mathbf{x}}}\right| \\ &= \|A\|_2\,,\end{aligned}$$

with maximum occurring when $\tilde{\vec{\mathbf{x}}} = \tilde{\vec{\mathbf{y}}} = \vec{\mathbf{e}}_1$. ■

Exercise 4.36.

Question. Give an example of a 2×2 matrix whose determinant is nonzero and whose singular values are not equal to any of its eigenvalues.

Answer.
Consider

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

where $a > b > 0$. ■

Exercise 4.38.

Question. Prove Proposition 4.6.2.
Prop 4.6.2. *If $A \in M_{m\times n}(\mathbb{F})$, then the Moore-Penrose pseudoinverse of A satisfies the following:*

1. $AA^\dagger A = A$.
2. $A^\dagger AA^\dagger = A$.
3. $(AA^\dagger)^{\mathsf{H}} = AA^\dagger$.
4. $(A^\dagger A)^{\mathsf{H}} = A^\dagger A$.
5. $AA^\dagger = \operatorname{proj}_{\mathscr{R}(A)}(\cdot)$ *is the orthogonal projection onto* $\mathscr{R}(A)$.
6. $A^\dagger A = \operatorname{proj}_{\mathscr{R}(A^{\mathsf{H}})}(\cdot)$ *is the orthogonal projection onto* $\mathscr{R}(A^{\mathsf{H}})$.

Answer.

1. U, V full rank so

$$\begin{aligned} U^{\mathsf{H}}U &= I \\ &= V^{\mathsf{H}}V, \end{aligned}$$

but $VV^{\mathsf{H}} \neq I$, so

$$\begin{aligned} A^{\dagger}A &= V_1\Sigma_1^{-1}U_1^{\mathsf{H}}U_1\Sigma_1V_1^{\mathsf{H}} \\ &= V_1\Sigma_1^{-1}\Sigma_1V_1^{\mathsf{H}} \\ &= V_1V_1^{\mathsf{H}}, \end{aligned}$$

so

$$\begin{aligned} AA^{\dagger}A &= AV_1V_1^{\mathsf{H}} \\ &= U_1\Sigma_1V_1^{\mathsf{H}}V_1V_1^{\mathsf{H}} \\ &= U_1\Sigma_1V_1^{\mathsf{H}} \\ &= A. \end{aligned}$$

■

2.

$$\begin{aligned} A^{\dagger}AA^{\dagger} &= V_1V_1^{\mathsf{H}}V_1\Sigma_1^{-1}U_1^{\mathsf{H}} \\ &= V_1\Sigma_1^{-1}U_1^{\mathsf{H}} \\ &= A^{\dagger}. \end{aligned}$$

■

3.

$$\begin{aligned} \left(AA^{\dagger}\right)^{\mathsf{H}} &= \left(U_1U_1^{\mathsf{H}}\right)^{\mathsf{H}} \\ &= \left(U_1^{\mathsf{H}}\right)^{\mathsf{H}}U_1^{\mathsf{H}} \\ &= U_1U_1^{\mathsf{H}} \\ &= AA^{\dagger}. \end{aligned}$$

■

4.

$$\begin{aligned} \left(A^{\dagger}A\right)^{\mathsf{H}} &= \left(V_1V_1^{\mathsf{H}}\right)^{\mathsf{H}} \\ &= \left(V_1^{\mathsf{H}}\right)^{\mathsf{H}}V_1^{\mathsf{H}} \\ &= V_1V_1^{\mathsf{H}} \\ &= A^{\dagger}A. \end{aligned}$$

■

5. Consider

$$\begin{aligned} \left\langle A\vec{\mathbf{v}}, \vec{\mathbf{v}} - AA^{\dagger}\vec{\mathbf{v}}\right\rangle &= \left\langle \vec{\mathbf{v}}, \left(A^{\mathsf{H}} - A^{\mathsf{H}}AA^{\dagger}\right)\vec{\mathbf{v}}\right\rangle \\ &= \left\langle \vec{\mathbf{v}}, \left(A^{\mathsf{H}} - V_1\Sigma_1U_1^{\mathsf{H}}U_1U_1^{\mathsf{H}}\right)\vec{\mathbf{v}}\right\rangle \\ &= \left\langle \vec{\mathbf{v}}, \left(A^{\mathsf{H}} - V_1\Sigma_1U_1^{\mathsf{H}}\right)\vec{\mathbf{v}}\right\rangle \\ &= \left\langle \vec{\mathbf{v}}, \left(A^{\mathsf{H}} - A^{\mathsf{H}}\right)\vec{\mathbf{v}}\right\rangle \\ &= \langle \vec{\mathbf{v}}, 0\rangle \\ &= 0, \end{aligned}$$

for fixed $\vec{\mathbf{v}} \in \mathbb{F}^n$. $\vec{\mathbf{v}} \in \mathscr{R}(A)$ is *fixed* under this given mapping, so the mapping is surjective, and we also have that $\dim\left(\mathscr{R}(A)\right) = \operatorname{rank}(\underbrace{U_1 U_1^{\mathsf{H}}}_{AA^\dagger})$. We know $AA^\dagger$ is a projection since we have that $\left(AA^\dagger\right)^2 = AA^\dagger$ and (1). ■

6.

$$\begin{aligned} A^{\mathsf{H}}\left(A^{\mathsf{H}}\right)^\dagger &= V_1 \Sigma_1 U_1^{\mathsf{H}} U_1 \left(\Sigma_1^{\mathsf{H}}\right)^{-1} V_1^{\mathsf{H}} \\ &= V_1 \Sigma_1 \Sigma_1^{-1} V_1^{\mathsf{H}} \\ &= V_1 \Sigma_1 \Sigma_1^{-1} V_1^{\mathsf{H}} = V_1 V_1^{\mathsf{H}} \\ &= A^\dagger A, \end{aligned}$$

so the proof in the case of A^{H} follows from (5), above. ■