**OSM Lab 2017** **Name:** HUNG HO
**Math Problem Set #3** **Email:** hqdhftw@uchicago.edu

**4.2.** We have:

$$D(ax^2+bx+c)=2ax+b.$$

Thus, the corresponding matrix of $D$ is

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

The characteristic polynomial of $A$ is

$$p_A(\lambda)=\lambda^3.$$

Hence, the only eigenvalue of $D$ is 0, with algebraic multiplicity 3. Since $\lambda=0$, its geometric multiplicity is also the dimension of $\mathrm{Ker}(A)$. Suppose $\mathbf{v}=(v_1,v_2,v_3)\in \mathrm{Ker}(A)$, we know that $A\mathbf{v}=0 \Leftrightarrow v_2=2v_3=0 \Leftrightarrow \mathbf{v}\in \mathrm{span}(\{\mathbf{e}_1\})$. This implies that the geometric multiplicity of $\lambda$ is 1.

**4.4.** Any $2\times 2$ matrix has the form

$$\begin{pmatrix} a & b \\ d & c \end{pmatrix}$$

(i) The characteristic polynomial of $A$ is $p_A(\lambda)=\lambda^2-\mathrm{tr}(A)\lambda+\det(A)$. Since $A$ is Hermitian, we have

$$A^H=\begin{pmatrix} \bar{a} & \bar{d} \\ \bar{b} & \bar{c} \end{pmatrix}.$$

$A^H=A$ implies $a=\bar{a}, b=\bar{d}$. Therefore,

$$\Delta=\mathrm{tr}(A)^2-4\det(A)=(a+c)^2-4(ac-bd)=(a-c)^2+4bd=a\bar{a}+c\bar{c}-a\bar{c}-\bar{a}c+4b\bar{b}=(a-c)(\bar{a}-\bar{c})+4b\bar{b}\geq 0.$$

Hence, $A$ has only real eigenvalues.

(ii) Since $A^H=-A$, now we deduce that $\bar{a}=-a, \bar{c}=-c, \bar{d}=-b$. Hence,

$$\Delta=a^2-2ac+c^2+4bd=-a\bar{a}+\bar{a}c+a\bar{c}-c\bar{c}-4b\bar{b}=-(a-c)(\bar{a}-\bar{c})-4b\bar{b}\leq 0.$$

Hence, $A$ has only imaginary eigenvalues.

**4.6.** Note that the determinant of an upper-triangular matrix or a diagonal matrix is just the product of the entries on the diagonal. Hence, assuming the matrix has dimension $n$, its characteristic polynomial $p$ has the form $p(\lambda)=\prod_{i=1}^{n}(\lambda-a_{ii})$, whose roots are exactly $a_{11},\dots,a_{nn}$.

**4.8.**

(i) It suffices to show that $\sin(x),\cos(x),\sin(2x),\cos(2x)$ are linearly independent. By exercise 3.8, we know that this is an orthonormal set, and thus linearly independent.

(ii) We know that $D$ is a $4\times 4$ matrix, whose columns (in that order) are

$$D(\sin(x)),D(\cos(x)),D(\sin(2x)),D(\cos(2x)).$$

We have

$$\begin{aligned} D(\sin(x)) &= \cos(x) \\ D(\cos(x)) &= -\sin(x) \\ D(\sin(2x)) &= 2\cos(2x) \\ D(\cos(2x)) &= -2\sin(2x). \end{aligned}$$

Thus, the matrix representation of $D$ is

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{pmatrix}.$$

(iii) Consider $U = \text{span}\{\sin(x), \cos(x)\}$ and $U^c = \text{span}\{\sin(2x), \cos(2x)\}$.

**4.13.** By some calculations, we got

$$P = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}, P^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}.$$

**4.15.** Since $A$ is semisimple and thus diagonalizable, there exists an invertible matrix $P$ such that $A = P^{-1}DP$, where $P$ is a diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$, the eigenvalues of $A$. Then, clearly for any non-negative integer $m$, we have $A^m = P^{-1}D^mP$. Thus, if $f(x) = a_0 + a_1x + \ldots + a_nx^n$, then $f(A) = a_0I + a_1P^{-1}DP + \ldots + a_nP^{-1}D^nP = P^{-1}f(D)P$. Note that for any $m$, $D^m$ is also a diagonal matrix whose entries are $\lambda_1^m, \ldots, \lambda_n^m$, thus we deduce that $f(D)$ is also a diagonal matrix whose entries are $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$. This shows that $\{f(\lambda_i)\}_{i=1}^n$ are the eigenvalues of $f(A)$.

**4.16.**

(i) We know that $A = P^{-1}DP$, where

$$P = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & 0.4 \end{pmatrix}, P^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}.$$

Hence,

$$\lim_{n\to\infty} A^n = \lim_{n\to\infty} P^{-1}D^nP = \lim_{n\to\infty} P^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 0.4^n \end{pmatrix}P = P^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}P = \frac{1}{3}\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}.$$

(ii) The answer is still the same (does not depend on the choice of norm).

(iii) Since $A$ is similar to $D$, we know that the eigenvalues of $A$ are 1 and 0.4. Thus by Theorem 4.3.12, the eigenvalues of $f(A) = 3I + 5A + A^3$ are $f(\lambda) = 3 + 5\lambda + \lambda^3$ for $\lambda = 1, 0.4$, which are 9 and 5.064.

**4.18.** We know that $\lambda$ is also an eigenvalue of $A^T$, so there exists $\mathbf{x}$ such that $A^T\mathbf{x} = \lambda\mathbf{x}$. Taking transposes of both sides yields $\mathbf{x}^TA = \lambda\mathbf{x}$.

**4.20.** Suppose $A$ is Hermitian and is similar to $B$. We can write $A = Q^HBQ$ for some $Q$ orthonormal, thus $B = QAQ^H$ and $B^H = (QAQ^H)^H = (Q^H)^HA^HQ^H = QAQ^H = B$ (note that $(Q^H)^H = Q$ and $A^H = A$). Hence $B$ is also Hermitian.

**4.24.** Suppose $A$ is a Hermitian matrix such that $A^H = A$. It suffices to show $\rho_A(\mathbf{x}) = \overline{\rho_A(\mathbf{x})}$ for any $\mathbf{x}$. Indeed, We have

$$\overline{\rho_A(\mathbf{x})} = \frac{\overline{\langle \mathbf{x}, A\mathbf{x}\rangle}}{\|\mathbf{x}\|^2} = \frac{\langle A\mathbf{x}, \mathbf{x}\rangle}{\|\mathbf{x}\|^2} = \frac{\langle A^H\mathbf{x}, \mathbf{x}\rangle}{\|\mathbf{x}\|^2} = \frac{\langle \mathbf{x}, A\mathbf{x}\rangle}{\|\mathbf{x}\|^2} = \rho_A(\mathbf{x}).$$

Similarly, if $A$ is skew-Hermitian, or $A^H = -A$, we have

$$\overline{\rho_A(\mathbf{x})} = \frac{\overline{\langle \mathbf{x}, A\mathbf{x}\rangle}}{\|\mathbf{x}\|^2} = \frac{\langle A\mathbf{x}, \mathbf{x}\rangle}{\|\mathbf{x}\|^2} = -\frac{\langle A^H\mathbf{x}, \mathbf{x}\rangle}{\|\mathbf{x}\|^2} = -\frac{\langle \mathbf{x}, A\mathbf{x}\rangle}{\|\mathbf{x}\|^2} = -\rho_A(\mathbf{x}).$$

Hence, the Rayleigh quotient can only take real values if $A$ is Hermitian and only imaginary values if $A$ is skew-Hermitian.

**4.25.**

(i) For any $j$, consider $(\mathbf{x_1}\mathbf{x}_1^H + \ldots + \mathbf{x}_n\mathbf{x}_n^H)\mathbf{x_j} = T\mathbf{x}_j$. We know that $\mathbf{x}_i^H\mathbf{x}_j = 0$ for $i \neq j$ and $\mathbf{x}_j^h\mathbf{x}_j = 1$ (because $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are orthonormal). This implies that $T\mathbf{x}_j = \mathbf{x}_j$ for all $j = 1, \ldots, n$. Hence, the kernel of $(T - I)$ has dimension $n$, (because $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly independent). Therefore, $T - I$ is the zero matrix, or $T = I$.

(ii) Let $T = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{x}_i^H$. By a similar calculation as in part (i), we know that $T\mathbf{x}_i = \lambda_i \mathbf{x}_i$ for all $i = 1, \ldots, n$. Since $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$ for all $i$ also, we deduce that $(T - A)\mathbf{x}_i = 0$ for all $i$, thus the kernel of $T - A$ has dimension $n$, and therefore $T = A$.

**4.27.** For any $i = 1, \ldots, n$, we know that $\mathbf{e}_i^H A \mathbf{e}_i > 0$ and is real-valued because $A$ is positive definite (note that $\mathbf{e}_i$ has 1 at the $i$-th entry and 0 everywhere else). Let $\mathbf{x}_i$ be the $i$-th column of $A$. Then $A\mathbf{e}_i = \mathbf{x}_i$. Hence, $\mathbf{e}_i^H A \mathbf{e}_i = \mathbf{e}_i^H \mathbf{x}_i = x_{ii} > 0$ and $x_{ii}$ is real-valued. This holds for any $i$, thus we deduce that all diagonal entries are real-valued and positive.

**4.28.** By the Cauchy-Schwarz's inequality with respect to the Frobenius norm, we have:

$$\operatorname{tr}(AB) = \operatorname{tr}(A^H B) = \langle A, B \rangle \leq \|A\|_F \|B\|_F = \sqrt{\operatorname{tr}(A^2)\operatorname{tr}(B^2)}.$$

Now we will show that $\operatorname{tr}(A^2) = \operatorname{tr}(A)^2$. Assume $A = (a_{ij})$, with $a_{ij} = \overline{a_{ji}}$. The $(kk)$-th entry of $A^2$ is therefore $\sum_{j=1}^{n} a_{kj} a_{jk} = \sum_{j=1}^{n} a_{jk} \overline{a_{jk}}$.

**4.31.**

(i) We can write the SVD decomposition of $A$ as

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

Hence we have $\|A\| \leq \|U\| \|\Sigma\| \left\|V^H\right\| = \|\Sigma\|$ (because $U$ and $V^H$ are orthonormal so their norms is 1). Now, for any $\mathbf{x} = (x_1, \ldots, x_n)$ such that $\|\mathbf{x}\|_2 = 1$, we have

$$\|\Sigma \mathbf{x}\|_2^2 = \sum_{i=1}^{r} \sigma_i^2 x_i^2 \leq \sigma_1^2 (\sum_{i=1}^{n} x_i^2) = \sigma_1^2.$$

Hence $\|A\|_2 \leq \|\Sigma\|_2 \leq \sigma_1$. On the other hand, we know that $\|A\|_2^2 = \max_{\|\mathbf{x}\|_2} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{\|\mathbf{x}\|_2} \frac{\mathbf{x}^H A^H A \mathbf{x}}{\|\mathbf{x}\|_2}$. Picking $\mathbf{x} = \mathbf{v}_1$, the eigenvector corresponding to $\lambda_1 = \sigma_1^2$ of $A^H A$, we know that $\mathbf{v}_1^H A^H A \mathbf{v}_1 = \mathbf{v}_1^H \lambda_1 \mathbf{v} = \lambda_1 \|\mathbf{v}\|_2^2$. This shows that $\|A\mathbf{v}_1\|_2 = \sigma_1^2 \|\mathbf{v}_1\|$. Therefore, $\left\|A^H A\right\|_2 = \|A\|_2^2 = \sigma_1^2$.

(ii) The proof is similar to that of (i), observing that the largest singular value of $A^{-1}$ is the inverse of the smallest singular value of $A$, which is $\sigma_n^{-1}$.

(iii) Follows directly from part (i).

(iv) Apply part (iii) for the matrix $UAV$, we know that $\|UAV\|_2^2 = \left\|(UAV)^H UAV\right\|_2 = \left\|V^H A^H AV\right\|_2$. It suffices to show that $\left\|V^H A^H AV\right\|_2 = \left\|A^H A\right\|_2$ for any orthonormal matrix $V$. But this is obvious because since $V$ is orthonormal, $V^H A^H AV$ and $A^H A$ are orthonormally similar, thus sharing the same eigenvalues. Therefore, their 2-norms is just the largest eigenvalue. From this we deduce that $\|UAV\|_2 = \sqrt{\|A^H A\|_2} = \|A\|_2$.

**4.32.**

(i) We have

$$\|UAV\|_F^2 = \operatorname{tr}((UAV)^H UAV) = \operatorname{tr}(V^H AU^H UAV) = \operatorname{tr}(V^H A^H AV) = \operatorname{tr}(VV^H A^H A) = \operatorname{tr}(A^H A) = \|A\|_F^2.$$

Hence, $\|UAV\|_F = \|A\|_F$.

(ii) We can apply the SVD decomposition and write $A = U\Sigma V^H$. By part (i), we know that $\|A\|_F = \left\|U^H AV\right\|_F = \|\Sigma\|_F = \sqrt{\operatorname{tr}(\Sigma^H \Sigma)} = \sqrt{\operatorname{tr}(\Sigma^2)}$. But $\Sigma^2$ is just a diagonal matrix whose diagonal entries are $\sigma_1^2, \ldots, \sigma_n^2$, hence $\|A\|_F = \sqrt{\operatorname{tr}(\Sigma^2)} = (\sigma_1^2 + \ldots + \sigma_n^2)^{\frac{1}{2}}$.

**4.33.** It is clear that there exists $\mathbf{x} = \mathbf{y} = \dfrac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$, where $\mathbf{v}_1$ is the eigenvector corresponding to the largest eigenvalue of $A$, such that $\mathbf{y}^H A\mathbf{x} = \|A\|_2$ (by Exercise 4.31). Now it suffices to show that $|\mathbf{y}^H A\mathbf{x}| \leq \|A\|_2$ for all $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$. Suppose on the contrary that there exists $\mathbf{y}, \mathbf{x}$ such that $|\mathbf{y}^H A\mathbf{x}| > \|A\|_2$. Since $\mathbf{y}^H A\mathbf{x} = \overline{\mathbf{x}^H A^H \mathbf{y}}$, it follows that $|\mathbf{y}^H A\mathbf{x}|^2 = |\mathbf{y}^H A\mathbf{x}\mathbf{x}^H A^H \mathbf{y}| = |\mathbf{y}^H A A^H \mathbf{y}| = \left\|A^H \mathbf{y}\right\| > \|A\|_2 = \left\|A^H\right\|_2$, which is a contradiction because $\left\|A^H\right\|_2 = \max\limits_{\|\mathbf{y}\|=1} \left\|A^H \mathbf{y}\right\|$.

**4.36.**

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

Its eigenvalues are $\pm i$, but its singular values are 1.

**4.38.**

(i) We have $AA^\dagger = U_1\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = U_1\Sigma_1\Sigma^{-1}U_1^H$. Thus $AA^\dagger A = U_1\Sigma_1\Sigma^{-1}U_1^H U_1\Sigma_1 V_1^H = U_1\Sigma_1\Sigma_1^{-1}\Sigma_1 V_1^H = U_1\Sigma_1 V_1^H = A$.

(ii) $A^\dagger A A^\dagger = V_1\Sigma^{-1}U_1^H U_1\Sigma_1\Sigma_1^{-1}U_1^H = V_1\Sigma_1^{-1}U_1^H = A^\dagger$.

(iii) $(AA^\dagger)^H = (U_1\Sigma_1\Sigma^{-1}U_1^H)^H = (U_1^H)^H\Sigma_1^{-1}\Sigma_1 U_1^H = U_1\Sigma_1\Sigma_1^{-1}U_1^H = AA^\dagger$ (because $\Sigma_1$ is diagonal, it commutes with its inverse).

(iv) Similar to part (iv).