Homework 2: Report

Dante Buhl

Jan 20^{th} 2024

1. Let $A \in C^{m\times m}$ be both upper-triangular and unitary. Show that A is a diagonal matrix. Does the same hold if $A \in C^{m\times m}$ is both lower-triangular and unitary?

   Proof. **(Upper Triangular, by Induction)**

   Assume matrix $A \in \mathbb{C}^{m\times m}$ is unitary and is upper triangular such that,

$$A^*A = \mathrm{I}_m = AA^*$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ 0 & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{mm} \end{bmatrix}$$

   Where A^* is the complex transpose matrix of A. We have then that A^* is of the form,

$$A^* = \begin{bmatrix} \overline{a_{11}} & 0 & \cdots & 0 \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1m}} & \overline{a_{2m}} & \cdots & \overline{a_{mm}} \end{bmatrix}$$

   Then the product of matrix multiplication of A^* and A is then defined as B, (i.e. $B = A^*A$), and because A is a Unitary matrix, is equal to I.

   (Base Case):

   Now assume that the elements above the diagonal in A are non-zerp. Next examine the $(1,1)$ and $(2,1)$ cells of the matrix product, B. By the operation of Matrix multiplication we should have,

$$B(2,1) = a_{11} \cdot \overline{a_{12}} = \mathrm{I}_m(2,1) = 0$$

$$B(1,1) = a_{11} \cdot \overline{a_{11}} = \mathrm{I}_m(1,1) = 1$$

   From this, we know that $a_{11} \neq 0$ and $\overline{a_{11}} \neq 0$. But we have that the product of $a_{11} \cdot \overline{a_{12}} = 0$. Since we have that $a_{11} \neq 0$, we must therefore have that $\overline{a_{12}} = 0$ and by the definition of a complex conjugate, $a_{12} = 0$.

   (Inductive Step)

   We need to show that for a integer $k \leq m-1$ all of the columns of matrix A, $\vec{C}_i$, up to $\vec{C}_k$ is of the form,

$$\vec{C}_i = \begin{bmatrix} 0 \\ \vdots \\ a_{ii} \\ \vdots \\ 0 \end{bmatrix}$$

then $\vec{C}_{k+1}$ is also of the same form. We have that in the matrix product between A^* and A, B, then the i-th row of B is defined as the inner produt between the i-th row of A^*, $\vec{r}_i^*$ and the j-th column of A, $\vec{c}_j$.

$$B(i,:) = [(\vec{r}_i^*, \vec{c}_1), (\vec{r}_2^*, \vec{c}_2), \cdots, (\vec{r}_i^*, \vec{c}_j)] = \mathrm{I}_m(i,:) = [0, \cdots, 1, \cdots, 0]$$

Look at the $(k+1)$-th row of B. We have from the given form of the columns, $\{\vec{c}_1, \cdots, \vec{c}_k\}$,

$$(\vec{r}_{k+1}^*, \vec{c}_j) = \overline{a_{j(k+1)}} \cdot a_{jj} = \left\{ \begin{array}{lll} 0 & \text{if,} & j \neq k+1 \\ 1 & \text{if,} & j = k+1 \end{array} \right\}, i, j < k+1$$

We also have that each $a_{ii} \neq 0$. Thereby, for all $j < k+1$,

$$\overline{a_{j(k+1)}} = 0 \implies a_{j(k+1)} = 0$$

We now write the column, $\vec{c}_{k+1}$.

$$\vec{c}_{k+1} = \begin{bmatrix} 0 \\ \vdots \\ a_{(k+1)(k+1)} \\ \vdots \\ 0 \end{bmatrix}$$

Therefore, we have that $\vec{c}_{k+1}$ is of the same form as $\vec{c}_i, i \leq k$. By induction, each column of A is of this form. Therefore, A is a diagonal matrix!

□

Proof. **(Lower Triangular, by case of Upper Triangular)**
Assume as before, $A \in \mathbb{C}^{m \times m}$ is unitary and is lower triangular such that,

$$A^*A = \mathrm{I}_m = AA^*$$

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

Where A^* is the complex transpose matrix of A. We have then that A^* is of the form,

$$A^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{m1}} \\ 0 & \overline{a_{22}} & \cdots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \overline{a_{mm}} \end{bmatrix}$$

We then define the matrix $C = A^*$, $C^* = A$. Notice that C is an upper triangular, unitary matrix. By the previous proof, C is a diagonal matrix. Notice all of its "off-diagonal" elements are zero. As a consequnce, all (i, j)-elements of C which are zero imply that (j, i)-elements of C^* are zero. Therefore, $C^* = A$ is a diagonal matrix.

□

2. Prove the following in each problem.

   (a) Let $A \in \mathbb{C}^{m \times m}$ be invertible and $\lambda \neq 0$ is an eigenvalue of A. Showthat λ^{-1} is an eigenvalue of A^{-1}.

Proof. Take any $A \in \mathbb{C}^{m\times m}$ to be invertible. Then we have inverse, A^{-1} exists such that,

$$AA^{-1} = \mathrm{I}_m = A^{-1}A$$

We also have by the fact that $\lambda \neq 0$ is an eigenvalue of A that,

$$\det(A - \lambda \mathrm{I}_m) = 0$$

We can substitute for I_m.

$$\det(A - \lambda \mathrm{I}_m) = \det(A - \lambda(A^{-1}A)) = \det(A)\det(\mathrm{I}_m - \lambda A^{-1}) = 0$$

$$\det(I_m - \lambda A^{-1}) = -\det(A^{-1} - \frac{1}{\lambda}\mathrm{I}_m) = 0$$

$$\det(A^{-1} - \frac{1}{\lambda}\mathrm{I}_m) = 0$$

Therefore, λ^{-1} is an eigenvalue of A^{-1}.

□

(b) Let $A, B \in \mathbb{C}^{m\times m}$. Show that AB and BA have the same eigenvalues.

Proof. Let A, B be square matrices as shown above. Now look at some eigenvalue of the matrix product AB, λ. We have by definition of an eigenvalue the following equality.

$$AB\vec{v} = \lambda\vec{v}$$

Now we multiply both vectors by the matrix B.

$$B(AB\vec{v}) = B(\lambda\vec{v})$$

$$(BA)(B\vec{v}) = \lambda(B\vec{v})$$

$$BA\vec{w} = \lambda\vec{w}$$

Therefore λ is also an eigenvalue of the matrix product BA. Since our choice of λ was arbitrary, we have that all eigenvalues of AB are eigenvalues of BA. □

(c) Let $A \in \mathbb{R}^{m\times m}$. Show that A and A^* have the same eigenvalues. (Hint 1: Use $\det(M) = \det(M^T)$ for any square matrix $M \in \mathbb{R}^{m\times m}$ in connection to the definition of characteristic polynomials. Hint 2: When a real-valued matrix A has a complex eigenvalue λ, then $\overline{\lambda}$ is also an eigenvalue of A.)

Proof. First look at an an arbitrary eigenvalue, λ, of A.

$$\det(A - \lambda\mathrm{I}) = 0$$

We look at two cases, 1. $\lambda \in \mathbb{R}$, and 2. $\lambda \in \mathbb{C}$.

Case 1: $\lambda \in \mathbb{R}$. We have since A, I, λ are all real-valued, that the conjugate transpose of $(A - \lambda\mathrm{I})^*$ is equal to the transpose of the same matrix quantity. i.e.

$$(A - \lambda\mathrm{I})^* = (A - \lambda\mathrm{I})^T = A^T - \lambda\mathrm{I} = A^* - \lambda\mathrm{I}$$

Therefore we can write,

$$\det(A - \lambda\mathrm{I}) = \det(A^* - \lambda\mathrm{I}) = 0$$

We can immediately see that any real eigenvalue of A is also an eigenvalue of A^*.

Case 2: $\lambda \in \mathbb{C}$, We again look at the conjugate transpose,

$$(A - \lambda\mathrm{I})^* = (A^* - \overline{\lambda}\mathrm{I})$$

□

3. Let $A \in \mathbb{C}^{m\times m}$ be hermitian. Suppose that for nonzero eigenvectors of A, there exist corresponding eigenvalues λ satisfying $Ax = \lambda x$.

   a Prove that all eigenvalues of A are real.

   Proof. We look at an arbitary eigenvalue of A.

$$Ax = \lambda x, x \in \mathbb{C}^m$$

   we multiple both sides by the conjugate transpose of x.

$$x^*(Ax) = x^*(\lambda x)$$

$$x^*Ax = \lambda(x^*x)$$

   We should notice that x^*x is a scalar with a real value. This is because each component of x is multiplied against its complex conjugate. Next we look at the dimensions and hermitian quantity of x^*Ax. We have that $x^* \in \mathbb{C}^{1\times m}$, otherwise known as a row vector. We also have, $Ax \in \mathbb{C}^m$. Thereby, the matrix product of x^* and Ax is a 1×1 quantity, a scalar! More importantly we have,

$$(x^*Ax)^* = x^*A^*(x^*)^* = x^*Ax$$

   So, x^*Ax is hermitian, or rather, x^*Ax is a real-valued scalar. We then have,

$$x^*Ax = \lambda(x^*x)$$

   Where both x^*Ax and x^*x are real valued, so consequently $\lambda \in \mathbb{R}$. □

   b. Let x and y be eigenvectors corresponding to distinct eigenvalues. Show that (x, y) = 0, i.e., they are orthogonal. (Hint: Use the result of Part (a).)

   Proof. By the quaality that A is hermition, we have for any two vectors, $x, y \in \mathbb{C}^m$, that

$$(Ax, y) = x^*A^*y = x^*Ay = (x, Ay)$$

   Therefore we can say for distinct eigenvectors, v_1, v_2 $(v_1 \neq v_2)$, with distinct eigenvalues, λ_1, λ_2 $(\lambda_1 \neq \lambda_2)$,

$$(Av_1, v_2) - (v_1, Av_2) = 0$$
$$= (\lambda_1 v_1, v_2) - (v_1, \lambda_2 v_2) = \overline{\lambda_1} v_1^* v_2 - v_1^* \lambda_1 v_2$$
$$= (\overline{\lambda_1} - \lambda_2) v_1^* v_2 = 0$$

   There are two things to notice, first since all eigenvalues are real, $\overline{\lambda_1} = \lambda_1$. Second, by our construction of the problem, $\lambda_1 \neq \lambda_2$. Thereby, $(\overline{\lambda_1} - \lambda_2) \neq 0$. So,

$$v_1^* v_2 = 0 = (v_1, v_2)$$

□

4. A matrix A is called positive definite if and only if $(Ax, x) > 0$ for all $x \neq 0$ in $\mathbb{C}^m$. Suppose A is Hermitian. Show that A is positive definite if and only if $\lambda_i > 0, \forall \lambda_i \in \Lambda(A)$, the spectrum of A.

Proof. By the property of A being hermitian, that we can write any vector, $x \in \mathbb{C}_m, x \neq \vec{0}$ as the linear combination of the orthonormal eigenvectors of A, u_i.

$$x = \alpha_1 u_1 + \cdots + \alpha_m u_m$$

We then look the inner product, (Ax, x).

$$Ax = A(\alpha_1 u_1 + \cdots + \alpha_m u_m) = \lambda_1 \alpha_1 u_1 + \cdots + \lambda_m \alpha_m u_m$$

$$(Ax)^* = \overline{\lambda_1 \alpha_1} u_1^* + \cdots + \overline{\lambda_m \alpha_m} u_m^*$$
$$(Ax, x) = (\overline{\lambda_1 \alpha_1} u_1^* + \cdots + \overline{\lambda_m \alpha_m} u_m^*)(\alpha_1 u_1 + \cdots + \alpha_m u_m)$$

Here by the property of an orthonormal vector set, we have that $u_i^* u_j = 0$ if $i \neq j$ and $= 1$ if $i = j$.

$$(Ax, x) = \overline{\lambda_1 \alpha_1}\alpha_1 + \cdots + \overline{\lambda_m \alpha_m}\alpha_m = \sum_{i=1}^{m} \lambda_i |\alpha_i|^2$$

Of course, $|\alpha_i|^2$ is a strictly positive value. So for $(Ax, x) < 0$ we need at least one $\lambda_i < 0$. In fact, it is the case that if even one $\lambda_i < 0$ that $(Ax, x) \not> 0$ for all $x \in \mathbb{C}^m$. To prove that $(Ax, x) > 0, \forall x \in \mathbb{C}^m$, we take the case of only the smallest λ_i, $\lambda_k < 0$ (i.e. $|\lambda_k| < |\lambda_i|, \forall \lambda_i \in (\Lambda(A) - \{\lambda_k\})$). We can show by counter-example

$$\lambda_k < 0, x \in \mathbb{C}^m, x = \alpha_1 u_1 + \cdots + \alpha_m u_m$$

$$(Ax, x) = \lambda_k |\alpha_k|^2 + \sum_{i=1, i \neq k}^{m} \lambda_i |\alpha_i|^2$$

$$\exists x_* \in \mathbb{C}^m, \text{ such that } |\alpha_k|^2 = \frac{1}{\lambda_k} \sum_{i=1, i \neq k}^{m} \lambda_i |\alpha_i|^2 + 1$$

$$(Ax_*, x_*) < 0, \text{ by construction.}$$

□

5. Suppose A is unitary.

   (a) Let (λ, x) be an eigenvalue-vector pair of A. Show λ satisfies $|\lambda| = 1$.

   Proof. Since A is unitary, we have that it preserves the angle and length of vectors under transformations. (i.e $(Ax, Ax) = (x, x)$ for any vector $x \in \mathbb{C}^m$). Thereby we have,

$$(Ax, Ax) = (\lambda x, \lambda x) = \overline{\lambda} x^* \lambda x = |\lambda|^2 x^* x = |\lambda|^2 (x, x)$$

$$(x, x) = (Ax, Ax) = |\lambda|^2 (x, x) \implies |\lambda|^2 = 1$$

$$|\lambda| = 1$$

□

   (b) Prove or disprove $||A||_F = 1$

   Proof. We have from the definition of the Frobenius Norm and since A is unitary,

$$||A||_F = \sqrt{\text{Tr}(A^* A)} = \sqrt{\text{Tr(I)}}$$

   Assume now that $\text{I} \in \mathbb{R}^{m \times m}$. Then, $\text{Tr(I)} = m$

$$||A||_F = \sqrt{m}$$

   Therefore, $||A||_F \neq 1$ unless, $A \in \mathbb{C}^{1 \times 1}$ i.e. A is a scalar. In general though, for any $A \in \mathbb{C}^{m \times n}$ where $m, n > 1$, $||A||_F \neq 1$. □

6. Let $A \in \mathbb{C}^{m\times m}$ be skew-hermitian, i.e., $A^* = -A$.

   (a) Show that the eigenvalues of A are pure imaginary.

   Proof. We look at the skew-hermition matrix A with (λ, x) being an eigenvalue-eigenvector pair $(Ax = \lambda x)$. We start by looking at the inner products $(Ax, x), (x, Ax)$.

$$\begin{aligned}&(Ax,x)-(x,Ax)\\&= x^*A^*x - x^*Ax\\&= -2x^*(Ax)\\&= -2\lambda x^*x\end{aligned}$$

$$\begin{aligned}&(Ax,x)-(x,Ax)\\&= (\lambda x)^*x - x^*\lambda x\\&= (\overline{\lambda}-\lambda)x^*x\end{aligned}$$

$$-2\lambda x^*x = (\overline{\lambda}-\lambda)x^*x$$

$$-2\lambda = \overline{\lambda}-\lambda$$

$$-\lambda = \overline{\lambda} \implies \mathbb{R}(\lambda) = 0$$

   Since our choice of x and lambda were arbitrary, we have that all eigenvalues of A are purely imaginary.

   □

   (b) Show that I A is nonsingular

   Proof. □

7. Show that $\rho(A) \le ||A||$, where $\rho(A)$ is the spectral radius of A.

   Proof. Start by taking the eigenvalue-eigenvector pair (λ_*, v) such that, $|\lambda_*| \ge |\lambda_i|, \lambda_i \in \Lambda(A)$. We have from the definition of the a matrix norm,

$$||A|| = \sup_{x\in\mathbb{C}^m}\frac{||Ax||}{||x||} \ge \frac{||Av||}{||v||}$$

$$\frac{||Av||}{||v||} = \frac{||\lambda_* v||}{||v||} = \frac{|\lambda_*|||v||}{||v||} = |\lambda_*|$$

$$||A|| = \sup_{x\in\mathbb{C}^m}\frac{||Ax||}{||x||} \ge |\lambda_*| = \rho(A)$$

$$||A|| \ge \rho(A)$$

   □

8. Let $A \in \mathbb{R}^{m\times m}$ and $Av_i = \alpha_i v_i, i = 1, ..., m$, where (α_i, v_i) is the eigenvalue-eigenvector pair of A for each i. Assume that A is symmetric, $A = A^T$ and the eigenvalues α_i are all distinct. Show that the solution to $Ax = b, x \ne 0$, can be written as,

$$x = \sum_{i=1}^{m}\frac{v_i^T b}{v_i^T A v_i}v_i$$

   (Hint 1: Use the fact that symmetric matrices are non-defective, and non-defective matrices are diagonalizable. Hint 2: Use the fact that, for realsymmetric matrices, the eigenvectors corresponding to distinct eigenvalues are orthogonal to each other, i.e., $(v_i, v_j) = 0, i \ne j$.)

Proof. Since we have that A is a real, symmetric matrix, it is true that its eigenvectors form an orthogonal basis which spans $\mathbb{R}^m$. Then we could write b as a linear combination of the eigenvalues of A.,

$$b = c_1 v_1 + \cdots + c_m v_m = A(d_1 v_1 + \cdots + d_m v_m), c_i = \alpha_i d_i$$

We now need to find the scalar coefficients d_i to obtain the correct linear combination. We next look at the inner products $(v_i, b), (v_i, Av_i)$. We have,

$$(v_i, b) = v_i^T(\alpha_1 d_1 v_1 + \cdots + \alpha_m d_m v_m) = \alpha_i d_i v_i^T v_i$$

The inner product of v_i with any $v_j, i \neq j$ is zero by orthogonality, so only the $v_i^T v_i$ term remains.

$$(v_i, Av_i) = v_i^T(Av_i) = v_i^T(\alpha_i v_i) = \alpha_i v_i^T v_i$$

We have then that, $\frac{(v_i,b)}{(v_i,Av_i)} = d_i$. Therefore we can now write,

$$x = \sum_{i=1}^{m} \frac{(v_i, b)}{(v_i, Av_i)} v_i = \sum_{i=1}^{m} \frac{v_i^T b}{v_i^T A v_i} v_i$$

□

9. Let A be defined as an outer product $A = uv^*$, where $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$.

   (a) Prove or disprove $||A||_2 = ||u||_2||v||_2$

   Proof. We have from the definition for the p-norm of a matrix $A \in \mathbb{C}^{\boldsymbol{\mu}_* \times n}$,

$$||A||_2 = \sup\left\{\frac{||Ax||_2}{||x||_2}\middle| x \in \mathbb{C}^n\right\}$$

$$||A||_2 = ||uv^*||_2 = \sup\left\{\frac{||uv^*x||_2}{||x||_2}\middle| x \in \mathbb{C}^n\right\}$$

   Notice here, that v^*a produces a 1×1 matrix, a scalar. We have the by the property of matrix and vector norms,

$$||A||_2 = \sup\left\{\frac{||u\alpha||_2}{||x||_2}\middle| x \in \mathbb{C}^n\right\} = \sup\left\{\frac{||u||_2|\alpha|}{||x||_2}\middle| x \in \mathbb{C}^n\right\}, \ \alpha = |v^*x|$$

$$||A||_2 = ||u||_2 \sup\left\{\frac{|\alpha|}{||x||_2}\middle| x \in \mathbb{C}^n\right\}$$

   We now look at the vector norm equality, the Hölder Inequality, for $p = q = 2$, with x per

$$|\alpha| = |v^*x| \leq ||v||_2||x||_2$$

   So,

$$\sup\left\{\frac{|\alpha|}{||x||_2}\middle| x \in \mathbb{C}^n\right\} \leq ||v||_2$$

   We now chose $x = cv, c \in \mathbb{R}$, c constant. Therefore,

$$|(v, x)| = |v^*x| = |\alpha| = |v^*cv| = |c|||v||_2^2 = ||v||_2||x||_2$$

   We have then that any x colinear to v gives us the supremum case. Therefore,

$$||A||_2 = ||u||_2\left(\frac{||v||_2||x||_2}{||x|_2}\right) = ||u||_2||v||_2$$

□

(b) Prove or disprove $||A||_F = ||u||_F||v||_F$

Proof. We begin with the definition of the Frobenius Norm.

$$||A||_F = \sqrt{\mathrm{Tr}(AA^*)} = \sqrt{\mathrm{Tr}(uv^*vu^*)}$$

$$||A||_F = \sqrt{\mathrm{Tr}(u||v||_2^2u^*)} = ||v||_2\sqrt{\mathrm{Tr}(uu^*)}$$

$$||A||_F = ||v||_2\sqrt{||u||_2^2} = ||v||_2||u||_2 = ||v||_F||u||_F$$

□

10. Let $A, Q \in \mathbb{C}^{m\times m}$, where A is arbitrary and Q is unitary

(a) Show that $||AQ||_2 = ||A||_2$

Proof. We begin with the definition of a 2-norm for matrices.

$$||AQ||_2 = \sup\left\{\frac{||AQx||_2}{||x||_2}\Big| x \in \mathbb{C}^m\right\}$$

$$||AQ||_2 = \sup\left\{\frac{||Ay||_2}{||x||_2}\Big| x \in \mathbb{C}^m\right\},\ y = Qx, ||y||_2 = ||x||_2$$

We have by the property of unitary matrices that the length and angles of vectors are preserved under transformations. Thus, $||Qx||_2 = ||x||_2 = ||y||_2$. We then substitute,

$$||AQ||_2 = \sup\left\{\frac{||Ay||_2}{||y||_2}\Big| y \in \mathbb{C}^m\right\} = ||A||_2$$

□

(b) Show that $||AQ||_F = ||QA||_F = ||A||_F$.

Proof. We start with the definition of the frobenius norm.

$$||AQ||_F = \sqrt{\mathrm{Tr}((AQ)^*AQ)} = \sqrt{\mathrm{Tr}((QA)^*QA)} = ||QA||_F$$

$$||AQ||_F = \sqrt{\mathrm{Tr}(A^*Q^*QA)} = \sqrt{\mathrm{Tr}(A^*A)} = ||A||_F$$

□

11. We say that $A, B \in \mathbb{C}^{m\times m}$ are unitarily equivalent if $A = QBQ^*$ for some unitary $Q \in \mathbb{C}^{m\times m}$.

(a) Show that if A and B are unitarily equivalent, then they have the same singular values.

Proof. We start with the SVD of B.

$$B = U_B\Sigma V_B^T$$

$$A = QBQ^* = QU_B\Sigma V_B^TQ^* = U_A\Sigma V_A^T$$

□

(b) Show that the converse of Part (a) is not necessarily true

Proof. Assume that A, B share the same singular value matrix Σ. We have then that,

$$A = U_A\Sigma V_A^T,\ B = U_B\Sigma V_B^T$$

In order for A and B to be unitarily equivalent, we must have that $U_A = QU_B$ and $V_A^T = V_B^TQ^*$

$$V_A^TU_A = V_B^TQ^*QU_B = V_B^TU_B$$

$$(V_A^TU_A)^* = Q^*U_A^*$$

□