Homework 13

This homework is due on Sunday, May 1, 2022 at 11:59PM. Self-grades and HW Resubmissions are due the following Sunday, May 8, 2022 at 11:59PM.

1. **Study Groups**

   If you are a student who participated in the study group survey we gave in the early weeks of the semester, we would really appreciate your feedback on the group you were matched with. If you did not participate in a study group, we would appreciate your input on what factors went into this decision. Please fill out this form to provide any feedback.

   This is optional, so to have something to write for the question, please tell us whether you filled out the survey or not.

2. **Reading Lecture Notes**

   Staying up to date with lectures is an important part of the learning process in this course. Here are links to the notes that you need to read for this week: Note 2j.

   (a) Explain (by way of an example) why the complex inner product cannot be the same as the real inner product $\langle \vec{x}, \vec{y} \rangle_{\mathbb{R}^n} = \vec{y}^\top \vec{x}$.

   Solution: If we used the real inner product (denoted as $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$) to define the real norm (denoted as $\|\cdot\|_{\mathbb{R}^n}$) by $\|\vec{x}\|^2_{\mathbb{R}^n} = \langle \vec{x}, \vec{x} \rangle_{\mathbb{R}^n}$, then we would have things like the following:

$$\left\| \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \right\|^2_{\mathbb{R}^2} = \left\langle \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix}, \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \right\rangle_{\mathbb{R}^2} \tag{1}$$

$$= \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix}^\top \begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix} \tag{2}$$

$$= 1 \cdot 1 + \mathrm{j} \cdot \mathrm{j} \tag{3}$$

$$= 1 + (-1) \tag{4}$$

$$= 0 \tag{5}$$

   so this would tell us that the norm (length) of $\begin{bmatrix} 1 \\ \mathrm{j} \end{bmatrix}$ is 0; however, this vector is nonzero and hence should have positive norm (again, length). So, the real inner product leads to paradoxical geometry on $\mathbb{C}^n$, and so we must use a different notion of complex inner product.

3. Extending Orthonormality to Complex Vectors

So far in the course, we have only dealt with real vectors. However, it is often useful to also think about complex vectors, as it allows for useful signal processing tools like the Discrete Fourier Transform (DFT) which you will learn in later courses. In this problem, we will extend several important properties of orthonormal matrices to the complex case. You are already introduced to complex inner products in the lectures (refer to Note 2j.)

(a) To get some practice computing complex inner products, **what are the complex inner products $\left\langle \begin{bmatrix} 1+j \\ 2 \end{bmatrix}, \begin{bmatrix} -3-j \\ 2+j \end{bmatrix} \right\rangle$ and $\left\langle \begin{bmatrix} -3-j \\ 2+j \end{bmatrix}, \begin{bmatrix} 1+j \\ 2 \end{bmatrix} \right\rangle$? Does the order of the vectors in the complex inner product matter i.e. is it commutative?**

Solution:

$$\left\langle \begin{bmatrix} 1+j \\ 2 \end{bmatrix}, \begin{bmatrix} -3-j \\ 2+j \end{bmatrix} \right\rangle = (1+j)\overline{(-3-j)} + (2)\overline{(2+j)} \tag{6}$$
$$= (1+j)(-3+j) + 2(2-j) \tag{7}$$
$$= -3+j-3j-1+4-2j = -4j \tag{8}$$
$$\left\langle \begin{bmatrix} -3-j \\ 2+j \end{bmatrix}, \begin{bmatrix} 1+j \\ 2 \end{bmatrix} \right\rangle = (-3-j)\overline{(1+j)} + (2+j)\overline{(2)} \tag{9}$$
$$= (-3-j)(1-j) + (2+j)(2) \tag{10}$$
$$= -3+3j-j-1+4+2j = 4j \tag{11}$$

The two inner products are complex conjugates of each other; in fact, this is a general property of complex inner products: $\langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$.

(b) Let $U = \begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_n \\ | & & | \end{bmatrix}$ be an n by n *unitary* matrix, whose columns $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ are orthonormal, i.e.

$$\langle \vec{u}_j, \vec{u}_i \rangle = \vec{u}_i^* \vec{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{12}$$

Show that $U^{-1} = U^*$, where U^* is the conjugate transpose of U.

Solution: By definition, $U^{-1}U = I$. We want to show that U^* satisfies it so let's write down U^* first:

$$U^* = \begin{bmatrix} - & \vec{u}_1^* & - \\ & \vdots & \\ - & \vec{u}_n^* & - \end{bmatrix}, \tag{13}$$

where the $\vec{u}_i$ column vectors of U turn into the conjugated row vectors $\vec{u}_i^*$. Then, the entry at the i-th row and j-th column of U^*U should be $\vec{u}_i^* \vec{u}_j$. If we write down the general form for each element of U^*U:

$$(U^*U)_{ij} = \vec{u}_i^* \vec{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{14}$$

which is the identity matrix, since $\vec{u}_1, \ldots, \vec{u}_n$ is an orthonormal basis.
Now for any square matrices A, B such that $AB = I$, right multiplying by B^{-1} gives $ABB^{-1} = IB^{-1}$ so $A = B^{-1}$. B^{-1} must exist since $\det(A)\det(B) = \det(I) \neq 0$ so $\det(B) \neq 0$. Thus since we showed $U^*U = I$, then $U^* = U^{-1}$.

(c) **Show that U as defined in (b) preserves complex inner products, i.e. if $\vec{v}, \vec{w} \in \mathbb{C}^n$ are vectors of length n, then**

$$\langle \vec{v}, \vec{w} \rangle = \langle U\vec{v}, U\vec{w} \rangle . \tag{15}$$

(HINT: Note that $(AB)^* = B^*A^*$.)

Solution: For this question, we want to show that:

$$\langle \vec{v}, \vec{w} \rangle = \vec{w}^*\vec{v} = \langle U\vec{v}, U\vec{w} \rangle \tag{16}$$

Using the definition of complex inner products we can write:

$$\langle U\vec{v}, U\vec{w} \rangle = (U\vec{w})^* U\vec{v} \tag{17}$$

Using the form for the complex conjugate of a matrix-vector product as stated in the problem:

$$(U\vec{w})^* U\vec{v} = \vec{w}^* U^* U\vec{v} \tag{18}$$

From the previous problem we know that $U^*U = I$. Therefore:

$$\langle U\vec{v}, U\vec{w} \rangle = \vec{w}^* U^* U\vec{v} = \vec{w}^*\vec{v} = \langle \vec{v}, \vec{w} \rangle . \tag{19}$$

(d) **Show that if $\vec{u}_1, \dots, \vec{u}_n$ are the columns of a unitary matrix U, they must be linearly independent.**

(HINT: Suppose $\vec{w} = \sum_{i=1}^{n} \alpha_i \vec{u}_i$, then first show that $\alpha_i = \langle \vec{w}, \vec{u}_i \rangle$. From here ask yourself whether a nonzero linear combination of the $\vec{u}_i$ could ever be identically zero.)

This fact shows how orthogonality is a very nice special case of linear independence.

Solution: Suppose the $\vec{u}_i$ are not linearly independent, then there exist $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ such that $\vec{w} = \sum_{i=1}^{n} \alpha_i \vec{u}_i = \vec{0}$, while at least one of α_i is non-zero. We can then take the inner product of both sides with $\vec{u}_j$, for all j:

$$\langle \vec{w}, \vec{u}_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i \vec{u}_i, \vec{u}_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle \vec{u}_i, \vec{u}_j \rangle = \alpha_j \tag{20}$$

Since $\vec{u}_1, \dots, \vec{u}_n$ form an orthonormal basis, we know that $\langle \vec{u}_i, \vec{u}_j \rangle$ will be 1 when $i = j$ and 0 otherwise, which is why only α_j survives in the above summation.

Since $\vec{w} = \vec{0}$, then α_j should be 0 for all inner products $\langle \vec{u}_j, \vec{w} \rangle$. However, this is a contradiction to our assumption that at least one of the α_i is non-zero. Therefore, $\vec{u}_1, \dots, \vec{u}_n$ are linearly independent.

This confirms what we know — that orthonormality is a particularly robust guarantee of linear independence.

(e) Now let V be another unitary matrix. **Show that UV is also an unitary matrix.**

Solution: Since V is a unitary matrix, we have $V^*V = I$. To show that the columns of UV also form an orthonormal basis, we could write down its conjugate transpose, $(UV)^*$, and apply it to UV:

$$(UV)^*(UV) = V^*U^*UV = V^*V = I \tag{21}$$

which means UV is unitary.

4. Gram-Schmidt on Complex Vectors

Consider the three complex vectors

$$\vec{a}_1 = \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{22}$$

Compute an orthonormal basis from this list of vectors with complex Gram-Schmidt.

(HINT: The complex version of Gram-Schmidt is Algorithm 45 of Note 2j.)

Solution: Running Gram-Schmidt, we get

$$\vec{z}_1 = \vec{a}_1 \tag{23}$$

$$\vec{p}_1 = \frac{\vec{z}_1}{\|\vec{z}_1\|} \tag{24}$$

$$\vec{z}_2 = \vec{a}_2 - \langle \vec{a}_2, \vec{p}_1 \rangle \vec{p}_1 \tag{25}$$

$$\vec{p}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} \tag{26}$$

$$\vec{z}_3 = \vec{a}_3 - \langle \vec{a}_3, \vec{p}_1 \rangle \vec{p}_1 - \langle \vec{a}_3, \vec{p}_2 \rangle \vec{p}_2 \tag{27}$$

$$\vec{p}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|}, \tag{28}$$

assuming that none of the $\vec{z}_i$ are $\vec{0}$ (otherwise we replace $\vec{p}_i$ by $\vec{0}$).

Now, we can compute

$$\vec{z}_1 = \vec{a}_1 \tag{29}$$

$$= \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix} \tag{30}$$

$$\|\vec{z}_1\| = \sqrt{\langle \vec{z}_1, \vec{z}_1 \rangle} \tag{31}$$

$$= \sqrt{\left\langle \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix} \right\rangle} \tag{32}$$

$$= \sqrt{\begin{bmatrix} 1 & -\mathrm{j} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix}} \tag{33}$$

$$= \sqrt{2} \tag{34}$$

$$\vec{p}_1 = \frac{\vec{z}_1}{\|\vec{z}_1\|} \tag{35}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \mathrm{j} \\ 0 \end{bmatrix} \tag{36}$$

$$= \begin{bmatrix} 1/\sqrt{2} \\ \mathrm{j}/\sqrt{2} \\ 0 \end{bmatrix}. \tag{37}$$

And

$$\vec{z}_2 = \vec{a}_2 - \langle \vec{a}_2, \vec{p}_1 \rangle \vec{p}_1 \tag{38}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} \tag{39}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 1/\sqrt{2} & -j/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} \tag{40}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \frac{j}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} \tag{41}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} j/2 \\ -1/2 \\ 0 \end{bmatrix} \tag{42}$$

$$= \begin{bmatrix} -j/2 \\ -1/2 \\ 0 \end{bmatrix} \tag{43}$$

$$\|\vec{z}_2\| = \sqrt{\langle \vec{z}_2, \vec{z}_2 \rangle} \tag{44}$$

$$= \sqrt{\left\langle \begin{bmatrix} -j/2 \\ -1/2 \\ 0 \end{bmatrix}, \begin{bmatrix} -j/2 \\ -1/2 \\ 0 \end{bmatrix} \right\rangle} \tag{45}$$

$$= \sqrt{\begin{bmatrix} j/2 & -1/2 & 0 \end{bmatrix} \begin{bmatrix} -j/2 \\ -1/2 \\ 0 \end{bmatrix}} \tag{46}$$

$$= \sqrt{1/2} \tag{47}$$

$$= \frac{1}{\sqrt{2}} \tag{48}$$

$$\vec{p}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} \tag{49}$$

$$= \sqrt{2} \begin{bmatrix} -j/2 \\ -1/2 \\ 0 \end{bmatrix} \tag{50}$$

$$= \begin{bmatrix} -j/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}. \tag{51}$$

And

$$\vec{z}_3 = \vec{a}_3 - \langle \vec{a}_3, \vec{p}_1 \rangle \vec{p}_1 - \langle \vec{a}_3, \vec{p}_2 \rangle \vec{p}_2 \tag{52}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} \right\rangle \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -j/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \right\rangle \begin{bmatrix} -j/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \tag{53}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \left(\begin{bmatrix} 1/\sqrt{2} & -j/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} - \left(\begin{bmatrix} j/\sqrt{2} & -1/\sqrt{2} \\ 0 & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} -j/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \tag{54}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{1}{\sqrt{2}} - \frac{j}{\sqrt{2}} \right) \begin{bmatrix} 1/\sqrt{2} \\ j/\sqrt{2} \\ 0 \end{bmatrix} - \left(\frac{j}{\sqrt{2}} - \frac{1}{\sqrt{2}} \right) \begin{bmatrix} -j/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \tag{55}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{56}$$

$$\|\vec{z}_3\| = \sqrt{\left\langle \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle} \tag{57}$$

$$= 1 \tag{58}$$

$$\vec{p}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|} \tag{59}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{60}$$

Thus

$$(\vec{p}_1, \vec{p}_2, \vec{p}_3) = \left(\begin{bmatrix} 1/\sqrt{2} \\ \mathrm{j}/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\mathrm{j}/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right). \tag{61}$$

Since none of them are $\vec{0}$'s, $\vec{q}_i = \vec{p}_i$ and we have

$$(\vec{q}_1, \vec{q}_2, \vec{q}_3) = \left(\begin{bmatrix} 1/\sqrt{2} \\ \mathrm{j}/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\mathrm{j}/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right). \tag{62}$$

as our orthonormal basis.

5. Adapting Proofs to the Complex Case

At many points in the course, we have made assumptions that various matrices or eigenvalues are real while discussing various theorems. If you have noticed, this has always happened in contexts where we have invoked orthonormality during the proof or statement of the result. Now that you understand the idea of orthonormality for complex vectors, and how to adapt Gram-Schmidt to complex vectors, you can go back and remove those restrictions. This problem asks you to do exactly that.

(a) The upper-triangularization theorem (Theorem 63 of Note 2j) says that every complex matrix $A \in \mathbb{C}^{n\times n}$ can be written as $A = UTU^*$, where $U \in \mathbb{C}^{n\times n}$ is unitary and $T \in \mathbb{C}^{n\times n}$ is upper triangular.

Adapt the proof from the real case with assumed real eigenvalues to prove this theorem.

Feel free to assume that any square matrix has an (potentially complex) eigenvalue/eigenvector pair. You don't need to prove this. But you can make no other assumptions.

(HINT: Use the exact same argument as before, just use conjugate-transposes instead of transposes.)

Solution: We use a recursive approach, though you can also use the language of induction.

The recursive base case is $n = 1$. If $A \in \mathbb{C}^{1\times 1}$ then A is essentially a scalar, so the orthonormal change-of-basis matrix $U \in \mathbb{C}^{1\times 1}$ can just be $U = \begin{bmatrix}1\end{bmatrix}$, and $T = A$. Thus $A = UTU^*$ is a Schur decomposition of A.

Now consider the general recursive case. Suppose $A \in \mathbb{C}^{n\times n}$ is a matrix with the n complex eigenvalues $\lambda_1, \ldots, \lambda_n$. Pick a normalized eigenvector $\vec{q}$ of A which corresponds to λ_1; this exists by the hint. Use Gram-Schmidt to extend $\vec{q}$ to an orthonormal basis $Q = \begin{bmatrix}\vec{q} & \widetilde{Q}\end{bmatrix}$ of $\mathbb{C}^n$. Since the columns of Q are an orthonormal set and Q is square, Q is unitary. Then

$$Q^*AQ = \begin{bmatrix}\vec{q}^* \\ \widetilde{Q}^*\end{bmatrix} A \begin{bmatrix}\vec{q} & \widetilde{Q}\end{bmatrix} \tag{63}$$

$$= \begin{bmatrix}\vec{q}^*A\vec{q} & \vec{q}^*A\widetilde{Q} \\ \widetilde{Q}^*A\vec{q} & \widetilde{Q}^*A\widetilde{Q}\end{bmatrix} \tag{64}$$

$$= \begin{bmatrix}\lambda_1\vec{q}^*\vec{q} & \vec{q}^*A\widetilde{Q} \\ \lambda_1\widetilde{Q}^*\vec{q} & \widetilde{Q}^*A\widetilde{Q}\end{bmatrix}. \tag{65}$$

Now since Q is unitary, we have

$$I_n = Q^*Q = \begin{bmatrix}\vec{q}^* \\ \widetilde{Q}^*\end{bmatrix} \begin{bmatrix}\vec{q} & \widetilde{Q}\end{bmatrix} = \begin{bmatrix}\vec{q}^*\vec{q} & \vec{q}^*\widetilde{Q} \\ \widetilde{Q}^*\vec{q} & \widetilde{Q}^*\widetilde{Q}\end{bmatrix}. \tag{66}$$

But we also know

$$I_n = \begin{bmatrix}1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & I_{n-1}\end{bmatrix} \tag{67}$$

so $\widetilde{Q}^*\vec{q} = \vec{0}_{n-1}$, and also $\vec{q}^*\vec{q} = 1$. Thus

$$Q^*AQ = \begin{bmatrix}\lambda_1\vec{q}^*\vec{q} & \vec{q}^*A\widetilde{Q} \\ \lambda_1\widetilde{Q}^*\vec{q} & \widetilde{Q}^*A\widetilde{Q}\end{bmatrix} \tag{68}$$

$$= \begin{bmatrix}\lambda_1 & \vec{q}^*A\widetilde{Q} \\ \vec{0}_{n-1} & \widetilde{Q}^*A\widetilde{Q}\end{bmatrix}. \tag{69}$$

To clean up a little, we introduce the notation

$$\vec{a}_{12}^* := \vec{q}^*A\widetilde{Q} \qquad \text{and} \qquad \widetilde{A}_{22} := \widetilde{Q}^*A\widetilde{Q}. \tag{70}$$

Thus we have

$$Q^*AQ = \begin{bmatrix}\lambda_1 & \vec{q}^*A\widetilde{Q} \\ \vec{0}_{n-1} & \widetilde{Q}^*A\widetilde{Q}\end{bmatrix} \tag{71}$$

$$= \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* \\ \vec{0}_{n-1} & \widetilde{A}_{22} \end{bmatrix}. \tag{72}$$

This is where we set up for the recursive call, where we will try to recursively upper triangularize $\widetilde{A}_{22}$. We first need to show that $\widetilde{A}_{22}$ is a smaller subproblem of A.

Since $\widetilde{Q} \in \mathbb{C}^{n\times(n-1)}$, we have $\widetilde{A}_{22} \in \mathbb{C}^{(n-1)\times(n-1)}$. Thus $\widetilde{A}_{22}$ is smaller than our original matrix A and also a square matrix.

Therefore we can recursively take the Schur decomposition of $\widetilde{A}_{22}$. Write

$$\widetilde{A}_{22} := P\widetilde{T}P^* \tag{73}$$

where $P \in \mathbb{C}^{(n-1)\times(n-1)}$ is unitary and $\widetilde{T} \in \mathbb{C}^{(n-1)\times(n-1)}$ is upper triangular. Then

$$Q^*AQ = \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* \\ \vec{0}_{n-1} & \widetilde{A}_{22} \end{bmatrix} \tag{74}$$

$$= \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* \\ \vec{0}_{n-1} & P\widetilde{T}P^* \end{bmatrix} \tag{75}$$

$$= \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix} \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* P \\ \vec{0}_{n-1} & \widetilde{T} \end{bmatrix} \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P^* \end{bmatrix}. \tag{76}$$

where the motivation to reach the last line is that we want to find a matrix factorization that isolates $\widetilde{T}$ in the bottom right corner, making the middle matrix upper-triangular. Again cleaning up notation, let

$$R := \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix}. \tag{77}$$

Then, using the unitarity of P, we have

$$R^*R = \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix}^* \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix} \tag{78}$$

$$= \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P^* \end{bmatrix} \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix} \tag{79}$$

$$= \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P^*P \end{bmatrix} \tag{80}$$

$$= \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & I_{n-1} \end{bmatrix} \tag{81}$$

$$= I_n. \tag{82}$$

Thus R is unitary. Thus we have

$$Q^*AQ = \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P \end{bmatrix} \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* P \\ \vec{0}_{n-1} & \widetilde{T} \end{bmatrix} \begin{bmatrix} 1 & \vec{0}_{n-1}^* \\ \vec{0}_{n-1} & P^* \end{bmatrix} \tag{83}$$

$$= R \begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* P \\ \vec{0}_{n-1} & \widetilde{T} \end{bmatrix} R^* \tag{84}$$

$$\implies A = \underbrace{QR}_{:=U} \underbrace{\begin{bmatrix} \lambda_1 & \widetilde{\vec{a}}_{12}^* P \\ \vec{0}_{n-1} & \widetilde{T} \end{bmatrix}}_{:=T} \underbrace{R^*Q^*}_{=U^*}. \tag{85}$$

Here Q is unitary so

$$U^*U = (QR)^*(QR) = R^*Q^*QR = R^*I_nR = R^*R = I_n. \tag{86}$$

Thus $U \in \mathbb{C}^{n\times n}$ is unitary. And $T \in \mathbb{C}^{n\times n}$ is upper triangular. Also, $A = UTU^*$ by our calculation.

(b) The spectral theorem (Theorem 58 of Note 2j) for Hermitian matrices says that a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ can be written as $A = U\Lambda U^*$ where U is a unitary matrix of eigenvectors of A, and Λ is a real and diagonal matrix of corresponding eigenvalues of A.

Adapt the proof from the real symmetric case to prove this theorem.

(HINT: As before, you should just leverage upper-triangularization and use the fact that $(ABC)^* = C^*B^*A^*$. There is a reason that this part comes after the first part.)

Solution: Since A is a square matrix, taking the Schur decomposition of A outputs a unitary matrix U and upper triangular matrix T such that $A = UTU^*$. Since A is Hermitian, we have $A = A^*$, and so

$$A = A^* \tag{87}$$
$$UTU^* = (UTU^*)^* \tag{88}$$
$$UTU^* = UT^*U^* \tag{89}$$
$$T = T^* \tag{90}$$

which means that $T = T^*$. Since T is upper triangular, T^* is both lower triangular and upper triangular, so T is diagonal. Furthermore,

$$T_{ii} = (T^*)_{ii} = \overline{T_{ii}} \tag{91}$$

so T_{ii} is real for all i, hence T is a real diagonal matrix. Since U is unitary, $A = UTU^*$ is a diagonalization of A.

To show that U is a matrix of eigenvectors of A, and T is a matrix of eigenvalues of A, note that we have

$$A = UTU^* \tag{92}$$
$$AU = UT \tag{93}$$
$$A\begin{bmatrix}\vec{u}_1 & \cdots & \vec{u}_n\end{bmatrix} = \begin{bmatrix}\vec{u}_1 & \cdots & \vec{u}_n\end{bmatrix}\begin{bmatrix}T_{11} & & \\ & \ddots & \\ & & T_{nn}\end{bmatrix} \tag{94}$$
$$\begin{bmatrix}A\vec{u}_1 & \cdots & A\vec{u}_n\end{bmatrix} = \begin{bmatrix}T_{11}\vec{u}_1 & \cdots & T_{nn}\vec{u}_n\end{bmatrix} \tag{95}$$

and looking at each column confirms that $A\vec{u}_i = T_{ii}\vec{u}_i$.

6. **(OPTIONAL) Make Your Own Problem.**

   Write your own problem about content covered in the course thus far, and provide a thorough solution to it.

   NOTE: This can be a totally new problem, a modification on an existing problem, or a Jupyter part for a problem that previously didn't have one. Please cite all sources for anything (including course material) that you used as inspiration.

   NOTE: High-quality problems may be used as inspiration for the problems we choose to put on future homeworks or exams.

7. **Homework Process and Study Group**

   Citing sources and collaborators are an important part of life, including being a student!

   We also want to understand what resources you find helpful and how much time homework is taking, so we can change things in the future if possible.

   (a) **What sources (if any) did you use as you worked through the homework?**

   (b) **If you worked with someone on this homework, who did you work with?**
   List names and student ID's. (In case of homework party, you can also just describe the group.)

   (c) **Roughly how many total hours did you work on this homework? Write it down here where you'll need to remember it for the self-grade form.**

Contributors:

- Druv Pai.
- Ashwin Vangipuram.
- Anant Sahai.
- Nathan Lambert.

