In summary we have

\[ \begin{align*}
\text{All eigenvalues are real} & \rightarrow U^T U = I \quad \text{``self-adjoint''} \quad U^T U = I \quad \text{``isometric''} \\
\text{real} & \rightarrow \text{symmetric} \quad \text{complex} \rightarrow \text{Hermitian} \quad \text{orthogonal} \rightarrow \text{unitary}
\end{align*} \]

Also satisfies since \( \langle u, w \rangle = \langle w, u \rangle = 0 \) and it can be derived from a common condition satisfied by both (or a few others).

\[ A \text{ is normal if } A^* A = AA^* \text{ (or } [A^*, A] = 0) \]

Thus, what else is normal? Anti-self-adjoint \( A^* = -A \)? Transforms by real or imaginary.

What about anti-isometric, i.e. \( U^* U = I \)?

Now here is the powerful result that is derived from the definition of normal and hence applies to all.

\[ \text{If } A \text{ is normal, then the eigenvectors belonging to distinct eigenvalues are orthogonal.} \]

**Proof:**

If \( A \) is normal then
\[ \| A x \|_2 = \langle A x, A x \rangle = \langle A^* A x, x \rangle = \langle A A^* x, x \rangle = (Ax, Ax) \]

If \( A \) is normal then so is \( A - \lambda \) since
\[ (A - \lambda)^* (A - \lambda) = A^* A - \lambda A^* - \lambda A + \lambda^2 \]
\[ = A A^* - \lambda (A^* + A) + \lambda^2 \]
\[ = (A - \lambda)(A - \lambda)^* \]

Plugging \( A - \lambda \) into previous result:
\[ \| A x - \lambda x \|_2 = \| A^* x - \lambda^* x \|_2 \quad \text{or} \quad A x = \lambda x \iff A^* x = \lambda^* x \]

Then if \( A x = \lambda x \) and \( A x = \lambda x \) we have:
\[ \langle x_i, A x_i \rangle = \lambda \langle x_i, x_i \rangle \]
\[ \langle x_i, A x_i \rangle = \langle A^* x_i, x_i \rangle = \langle x_i, x_i \rangle \]
\[ = \lambda (x_i, x_i) \]

Then \( (\lambda_i - \lambda_j) \langle x_i, x_j \rangle = 0 \) \( \forall \) \( (x_i, x_j) = 0 \) if \( \lambda_i \neq \lambda_j \).
Consider: \( A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \)

First of all, \( A^* = A \) so \( A^* A = AA^* = A \) is normal.

The eigenvalues are \( \lambda_i = \{3, 5, -1, -1\} \)

The eigenvectors are \( X_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \) and \( \lambda_i \neq 0 \Rightarrow \hat{X}_i = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \)

\( \langle \hat{X}_i, \hat{X}_j \rangle = 0 \) if \( i \neq j \)

\( \langle \hat{X}_i, \hat{X}_i \rangle = 1 \)

\( \hat{X}_1, \hat{X}_2, \hat{X}_3, \hat{X}_4 \)

\( X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \) for \( \lambda_1 = 3 \)

\( X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \) for \( \lambda_2 = 5 \)

\( X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \) for \( \lambda_3 = -1 \)

\( X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \) for \( \lambda_4 = -1 \)

Note that \( X_3 \) and \( X_4 \) need not be orthogonal, e.g. \( X_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \) and \( X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \) \( \langle X_3, X_4 \rangle = 1 \)

And the eigenvalues are \( \lambda_i \), e.g., \( \lambda_1 = 3 \)

\( X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \) for \( \lambda_1 = 3 \)

\( X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \) for \( \lambda_2 = 5 \)

\( X_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \) for \( \lambda_3 = -1 \)

\( X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \) for \( \lambda_4 = -1 \) \( \langle X_3, X_3 \rangle = 0 \)
Let's now turn to the question of when and how we diagonalize.

Recall that if we manage to get n eigenvectors of an n×n matrix $A$, then we can use them to form $P$ s.t. $D = P^{-1}AP$ where $D$ is a diagonal matrix whose elements are the eigenvalues of $A$. Note: All matrices are n×n.

One caveat in this construction is that the eigenvectors used must be linearly independent and in fact must span the space.

Now one may wonder if the transformation matrices $P$ and $P^{-1}$ above are by chance self-adjoint or isometric. Well it turns out:

A matrix $A$ can be diagonalized by a unitary similarity transformation $P$ iff $A$ is normal.

This actually contains lesser results (A being Hermitian) and its proof uses some of these.

Proof with a proof with a proof:

Any matrix $A$ can be expressed as: $A = \frac{1}{2} (A + A^*) + \frac{i}{2} (A - A^*) = B + iC$ where $B$ and $C$ are Hermitian.

Now if $B$ and $C$ can be simultaneously diagonalized by the same unitary transformation, then obviously $A$ can be since $P^{-1}AP = P^{-1}BP + iP^{-1}CP$ is diagonal.

So we have two things to show:

1) That Hermitian matrices are diagonalizable by a unitary transform.

2) That two Hermitian matrices are simultaneously diagonalizable and any conditions that must be met in those.
For (a) we have:

Any Hermitian matrix $A$ may be diagonalized by a unitary similarity transformation.

Recall that if $A$ is normal (as Hermitian is) then the eigenvectors belonging to distinct eigenvalues are orthogonal. But this means if we form $P$ as usual as a matrix, each column of which is a normalized eigenvector then $P^*P = I$. Clearly, symmetric matrices with distinct eigenvalues will $P^*P = I$.

Now when this gets tricky is if the eigenvalue spectrum is degenerate, i.e., different eigenvectors have the same eigenvalue. In this case the orthogonality of eigenvectors (belonging to the same eigenvalue) is not obvious.

So let us a 2x2: $A = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$ and det $(a-c) = (a-\lambda)(c-\lambda) - 16b^*b = (c-\lambda)(c-\lambda) - 16b^*b = 0$.

$\lambda^2 - (a+c)\lambda + ac - 16b^*b = 0$ \Rightarrow $\lambda = \frac{a+c \pm \sqrt{(a+c)^2 - 4ac + 64|b|^2}}{2}$

$A^*A = \begin{pmatrix} a^* & c \\ c^* & a \end{pmatrix}$ so $(c-\lambda) = (a^* - \lambda)(a - \lambda) = 0$ for $\lambda = \lambda_2 \Rightarrow a = \lambda_2, b = 0$.

We “need” to $A\vec{x} = \lambda_2 \vec{x} \Rightarrow A(\vec{x}) = 2\vec{x}$, the corresponding eigenvector $\lambda = \lambda_2 = 2$ for any $c$ and $d$ are eigenvectors.

We do this if you want and write $S$ as its unitary!

Now obviously for larger matrices, we can have algebraic multiplicity which is less than the dimension of the matrix, you will deal with in homework. Let’s proceed in an arbitrary case.

Via induction is good practice.

Suppose $A$ is $n \times n$, and let us assume that any $(n-1) \times (n-1)$ Hermitian matrix is unitarily diagonalizable.

Let $\lambda_i$ be any eigenvalue of $A$ with normalized $X_i = (x_i, x_{i1}, \ldots, x_{in})$, the corresponding eigenvector.

Form $V = (X_1, X_2, \ldots, X_n)$ w/ $X_i = \vec{e}_j$.

Note this means $V^*V = I$, i.e., $V$ is unitary.

Then $(V^*AV)_{ij} = (V^*AU)_{ij} = \sum_{k} x_{ki}^* \lambda_k x_{kj} \Rightarrow (V^*AV)_{ij} = \sum_{k} X_k \lambda_k \delta_{ij}$

But $(V^*AV) = V^*A^*V = V^*^AV$ (so $V^*AV$ is Hermitian) thus $(V^*AV)_{ij} = (\lambda_i \delta_{ij}) \Rightarrow (V^*AV)_{ij} = \lambda_i \delta_{ij}$.

This means $V^*AV = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \Rightarrow B$ but recall our assumption that an $(n-1) \times (n-1)$ is diagonalizable be a unitary, $V$, s.t. $(V^*A^*V)^* = \lambda_i \delta_{ij}$. This means $V^*AV = \begin{pmatrix} A_{(\text{diagonal})} \\ \vdots \\ A_{(\text{diagonal})} \end{pmatrix}$.

Thus $B$ is diagonalizable via $W^*BV$ w/ $W = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ and since $V$ is unitary, so is $W$.

But what about $A$? Well $W^*BW = W^*V^*AV = W^*V^*V = (WV)^{-1}A(WV) = U^*AU$ w/ $U = VW$.

But $U^* = (WV)^* = W^*U = W^*V^*V = (WV)^{-1} = U^{-1}$, so $U$ is unitary!

So an non-Hermitian $A$ is diagonalizable via a unitary $U$ if $(n-1) \times (n-1)$ is. But start $w$ and $b$!