Another subgroup of linear transformations which is extremely useful are called "isometries."

For a linear transformation U on an inner-product space V we define:

If V is complex and $U^+U = UU^+ = I$ then U is "unitary"

If V is real and $\tilde{U}U = U\tilde{U} = I$ then U is "orthogonal"

} In both cases U is an isometry

Note these imply: $(U^+)^+ = U = (U^{-1})^+ = (U^+)^{-1} = (U^{-1})^{-1}$ and similarly for $\sim$

Now this definition doesn't really use the inner-product, but it turns out that these transformations are particularly important for inner-product spaces. In the following, if one is true, all are true.

U on an inner-product space satisfies:

1. $U^+U = I$
2. $(Ux, Uy) = (x, y)$ for all x and y
3. $\|Ux\| = \|x\|$ for all x

To prove it, start w/ 1 being true: $U^+U = I \Rightarrow (Ux, Uy) = (x, U^+Uy) = (x, y)$ for all x, y

and if $x = y$ $(Ux, Ux) = \|Ux\|^2 = (x, x) = \|x\|^2$ for all x

So $1 \Rightarrow 2 \Rightarrow 3$, and to finish up we need $3 \Rightarrow 1$

If $\|Ux\|^2 = (Ux, Ux) = \underbrace{(U^+Ux, x)}_{\text{we're not saying } U^+U = I \text{ yet!}} = (x, x) \Rightarrow ([U^+U - I]x, x) = 0$ for all x

This is just another expression for what is on the left of it.

Recall that if $(Ax, x) = 0$ for all x and A is self-adjoint then $A = 0$.

So if $[U^+U - I]$ is self adjoint then $U^+U - I = 0 \Rightarrow U^+U = I$ and 1 holds.

Is $[U^+U - I]$ self-adjoint? $[U^+U - I]^+ = (U^+U)^+ - I^+ = U^+U^{++} - I = U^+U - I$

so yes it is!

So isometries preserve lengths of vectors, but they also preserve the angle between two vectors

$$\cos\theta = \frac{(x,y)}{\|x\|\|y\|} \Rightarrow \frac{(Ux, Uy)}{\|Ux\|\|Uy\|} = \frac{(x,y)}{\|x\|\|y\|} = \cos\theta$$

But the preservation of lengths and angles means that an isometry will carry one $\underbrace{\text{ortho}}_{\text{angle}=90°}\underbrace{\text{normal}}_{\text{length}=1}$ set into another. The only concern might be whether it takes a complete orthonormal basis into another, and this can be proven w/ Parseval's equation.

So we have: [If $\{x_i\}$ is a complete orthonormal basis, then so is $\{Ux_i\}$ for any isometry U.

Going back to eigenvalues, we can draw the results for self-adjoint and isometric transformations:

For Hermitian transformations we know the eigenvalues are all real. But the same is true for symmetric transformations. Which leads to the slightly more general:

If A is a self-adjoint transformation, then all of its eigenvalues are real.

To prove this recall eigenvalues λ of A are s.t. $Ax = \lambda x$ for $x \neq 0$.

Now $(x, Ax) = (x, \lambda x) = \lambda(x, x)$, but if $A = A^{+}$ then $(x, Ax) = (A^{+}x, x) = (Ax, x) = \lambda^{*}(x, x)$

so we have $\lambda(x, x) = \lambda^{*}(x, x)$ and $(x, x) \neq 0$ so $\lambda = \lambda^{*}$.

For isometric transformations we have:

All eigenvalues of isometric transformations have absolute values of 1.

Proof:

If U is an isometry and $Ux = \lambda x$ for $x \neq 0$ then $\|x\| = \|Ux\| = |\lambda| \|x\| \Rightarrow |\lambda| = 1$

since $\|x\| \neq 0$.

In summary we have

All eigenvalues are real

$U^+ = U$ "self-adjoint" real → "symmetric", complex → "Hermitian"

All eigenvalues have $|\lambda| = 1$

$U^+U = I$ "isometry" real → "orthogonal", complex → "unitary"

also satisfies since $[U, U^+] = [U, U] = 0$

also satisfies since $[U, U^+] = I - I = 0$

Do these have anything in common? Turns out they do, and it can be derived from a common condition satisfied by both (w/ a few others)

[A linear transformation is "normal" if $A^+A = AA^+$ (or $[A^+, A] = 0$)

What else is normal? Anti-self adjoint $A^+ = -A$ transformations (real or imaginary).

What about anti-isometric, i.e. $U^+U = -1$?

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

[If A is normal, then the eigenvectors belonging to distinct eigenvalues are orthogonal.

Proof:

If A is normal then $\|Ax\|^2 \equiv (Ax, Ax) = (A^+Ax, x) = (AA^+x, x) = (A^+x, A^+x) = \|A^+x\|^2$

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

$$= AA^+ - \lambda^*A - \lambda A^+ + \lambda^*\lambda$$

$$= (A-\lambda)(A-\lambda)^+$$

Plugging $A - \lambda$ into previous result: $\|Ax - \lambda x\|^2 = \|A^+x - \lambda^* x\|^2 \Rightarrow \underline{Ax = \lambda x \Leftrightarrow A^+x = \lambda^* x}$

Then if $Ax_1 = \lambda_1 x_1$ and $Ax_2 = \lambda_2 x_2$ we have: $(x_1, Ax_2) = \lambda_2 (x_1, x_2)$

$$(x_1, Ax_2) = (A^+x_1, x_2) = (\lambda_1^* x_1, x_2)$$

$$= \lambda_1 (x_1, x_2)$$

Then $(\lambda_2 - \lambda_1)(x_1, x_2) = 0 \Rightarrow (x_1, x_2) = 0$ if $\lambda_1 \neq \lambda_2$.

Let's now turn to the question of when and how we diagonalize.

Recall that if we manage to get n eigenvectors of an $n \times n$ 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 $n \times n$.
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 within proof within proof:
Any matrix A can be expressed as: $A = \frac{1}{2}(A + A^{\dagger}) + i\,\frac{1}{2i}(A - A^{\dagger}) = B + iC$ where $\underbrace{B \text{ and } C \text{ are Hermitian}}^{\text{Note } A \text{ may not be}}$.
Now if B and C can be simultaneously diagonalized by the same unitary transformation, then obviously A can be since $P^{-1}AP = P^{-1}BP + iP^{-1}CP = \text{diag} + \text{diag} = \text{diag}$
So we have two things to show a) That Hermitian matrices are diagonalizable by a unitary trans.
b) When two Hermitian matrices are simultaneously diagonalizable
and any conditions that must be met in these.

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 $P^{+}P = \underline{I}$. Clearly symmetric matrices w/ distinct eigenvalues $\Rightarrow \tilde{P}P = \underline{I}$

Now where 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.

Start w/ a 2×2: $A = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix} \Rightarrow \det\begin{pmatrix} a-\lambda & b \\ b^* & c-\lambda \end{pmatrix} = (a-\lambda)(c-\lambda) - |b|^2 = ac - a\lambda - c\lambda + \lambda^2 - |b|^2 = 0$

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

$A^{+} = A \qquad A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \quad \Rightarrow (a-c)^2 + 4|b|^2 = 0$ for $\lambda_+ = \lambda_- \Rightarrow a = c,\ b = 0$

No "need" to diagonalize, though $\quad Ax = \lambda x \Rightarrow A\begin{pmatrix} c \\ d \end{pmatrix} = a\begin{pmatrix} c \\ d \end{pmatrix} \Rightarrow \begin{matrix} ac = ac \\ ad = ad \end{matrix} \Rightarrow \begin{pmatrix} c \\ d \end{pmatrix}$ for any c and d are eigenvectors.

We can choose $x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ for an orthonormal set.

$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ if you want and $U^{+}U = I$ so its unitary!

Now obviously for larger matrices, we can have algebraic multiplicity which is $1 \le k \le n$ where n is the dimension of the matrix. You will deal w/ $k = n$ in homework. Let's proceed in an arbitrary case.

Via induction:

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

Let λ_1 be any eigenvalue of A w/ normalized $x_1 = (x_{11}, x_{21}, \cdots, x_{n1})$ the corresponding eigenvector.

Form $V = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & & x_{2n} \\ \vdots & \vdots & & \vdots \end{pmatrix} = (x_{i1}, x_{i2}, \cdots, x_{in})$ w/ $\langle x_k, x_j \rangle = \delta_{kj}$.

Note this means $V^{+}V = \underline{I}$, i.e. V is unitary.

Then $(V^{-1}AV)_{ij} = (V^{+}AV)_{ij} = \sum_k x_{ki}^* \sum_m a_{km} x_{mj} \Rightarrow (V^{-1}AV)_{i1} = \sum_k x_{ki}^* \lambda_1 x_{k1} = \lambda_1 \delta_{i1}$

But $(V^{-1}AV)^{+} = V^{+}A^{+}(V^{-1})^{+} = V^{-1}AV$ (so $V^{-1}AV$ is Hermitian) thus $(V^{-1}AV)_{1i} = (\lambda_1\delta_{1i})^* = \lambda_1\delta_{1i}$

This means: $V^{-1}AV = \begin{pmatrix} \lambda_1 & \cdots \\ \vdots & A'_{(n-1)\times(n-1)} \end{pmatrix} \equiv B$ but recall our assumption that an $(n-1)\times(n-1)$ is diagonalizable by a unitary V' s.t. $(V')^{-1}A'V'$ is diagonal.

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

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

But $U^{+} = (VW)^{+} = W^{+}V^{+} = W^{-1}V^{-1} = (VW)^{-1} = U^{-1}$ so U is unitary!

So an $n\times n$ Hermitian A is diagonalizable via a unitary U if $(n-1)\times(n-1)$ is. But start w/ 2×2!

For (6) we have:

Okay so arbitrary $A = B + iC$ w/ B, C Hermitian. Can we diagonalize both B and C w/ the same unitary transformation? Well...

[Two Hermitian matrices exhibit a complete orthonormal set of common eigenvectors, iff $[B,C]=0$.

If you think about it, this should have been pointed out in Quantum Mechanics. To fully pen out a set of observables, you must make sure they commute, e.g. H, L^2, L_z, since only then can you use a single orthonormal basis. If not, e.g. x and p_x, then these require different bases and hence cannot be measured simultaneously (an eigenvector of one is not an eigenvector of the other).

Okay, so we know $A = B + iC$ can be diagonalized by the unitary transformation that simultaneously diagonalizes B and C as long as $[B,C]=0$.

But recall: $A = \underbrace{\frac{1}{2}(A+A^+)}_{B} + i\underbrace{\frac{1}{2i}(A-A^+)}_{C} \Rightarrow [B,C] = 0 = [\frac{1}{2}(A+A^+), \frac{1}{2i}(A-A^+)]$

$$= \underbrace{\cancel{[A,A]}}_{0} + [A^+,A] - [A,A^+] - \underbrace{\cancel{[A^+,A^+]}}_{0}$$

$$= 2[A^+,A] \Rightarrow \underbrace{[A^+,A]=0}_{\text{i.e. } A \text{ is normal}}$$

Recall that normal includes (anti) Hermitian, (anti) symmetric, unitary and orthogonal. So all of these can be diagonalized by a unitary matrix.

But we might suspect that possibly the (anti)symmetric and orthogonal matrices might be diagonalizable w/ an orthogonal transformation since it just amounts to making everything real, i.e. $A^+ = A \xrightarrow[\text{real}]{} \tilde{A} = A$, $A^+A = I \xrightarrow[\text{real}]{} \tilde{A}A = I$, $[A^+,A]=0 \xrightarrow[\text{real}]{} [\tilde{A},A]=0$,

but

$A = \underbrace{\frac{1}{2}(A+A^+)}_{\text{Hermitian}} + i\underbrace{\frac{1}{2i}(A-A^+)}_{\text{Hermitian}} \xrightarrow[\text{real}]{} A = \underbrace{\frac{1}{2}(A+\tilde{A})}_{\text{symmetric}} + \underbrace{\frac{1}{2}(A-\tilde{A})}_{\text{antisymmetric}}$ Note: In the real case they are different.

Moreover, at times real matrices have complex eigenvalues and eigenvectors.

It turns out:

[Any real symmetric matrix can be diagonalized by an orthogonal transformation.

otherwise (for antisymmetric and orthogal matrices) the diagonalization may only be unitary.

Examples: $\underset{A=\tilde{A}}{\underbrace{\text{Symmetric}}}$ $A=\begin{pmatrix}1&2\\2&1\end{pmatrix} \Rightarrow \lambda_1=-1 \quad \hat{x}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{pmatrix}$

$\lambda_2=3 \quad \hat{x}_2=\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}$

and $A_{diag}=P^{-1}AP$ w/ $P=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}$

or

$P'=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}$

both of which are orthogonal $\tilde{P}P=I$

$$A_{diag}=P^{-1}AP=\tilde{P}AP=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1&2\\2&1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}-\frac{1}{\sqrt{2}}&\frac{3}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&\frac{3}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}-1&0\\0&3\end{pmatrix}$$

$$A_{diag}=P'^{-1}AP'=\tilde{P}'AP'=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1&2\\2&1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{3}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\\frac{3}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}3&0\\0&-1\end{pmatrix}$$

$\underset{B=-\tilde{B}}{\underbrace{\text{Antisymmetric}}}$ and $\underset{B\tilde{B}=I}{\underbrace{\text{orthogonal}}}$ $B=\begin{pmatrix}0&1\\-1&0\end{pmatrix} \Rightarrow \lambda_1=i \quad \hat{x}_1=\begin{pmatrix}\frac{1}{\sqrt{2}}\\i\frac{1}{\sqrt{2}}\end{pmatrix}$

$\lambda_2=-i \quad \hat{x}_2=\begin{pmatrix}\frac{1}{\sqrt{2}}\\-i\frac{1}{\sqrt{2}}\end{pmatrix}$

and $B_{diag}=P^{-1}BP$ w/ $P=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\-i\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\end{pmatrix}$

or

$P'=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\i\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\end{pmatrix}$

both of which are unitary $P^{+}P=I$

$$B_{diag}=P^{-1}BP=P^{+}BP=\begin{pmatrix}\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\-i\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}-i\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}-i&0\\0&i\end{pmatrix}$$

$$B_{diag}=P'^{-1}BP'=P'^{+}BP'=\begin{pmatrix}\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}0&1\\-1&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\i\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&i\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}i\frac{1}{\sqrt{2}}&-i\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}=\begin{pmatrix}i&0\\0&-i\end{pmatrix}$$

What more does adjointing (but not self-adjoint) bring us?

[If A is any linear operator w/ eigenvalues λ_i, then A^+ has eigenvalues λ_i^*.

Proof: Recall λ come from $\det(A-\lambda I)=0$, but $\det A = (\det A^+)^* \Rightarrow (\det A)^* = \det A^+$

and since $\det(A-\lambda I) = 0 = [\det(A-\lambda I)]^* \Rightarrow \det(A-\lambda I)^+ = \underbrace{0 = \det(A^+ - \lambda^* I)}_{\text{hence } \lambda^* \text{ are eigenvalues of } A^+}$

Moreover:

[If any A has an eigenvalue λ_i w/ an eigenvector x_i, and A^+ has an eigenvalue λ_j^* w/ eigenvector x_j, then $\langle x_i, x_j \rangle = 0$ whenever $\lambda_i \neq \lambda_j$.

Proof: Consider $Ax_i = \lambda_i x_i \Rightarrow \underbrace{\langle x_j, Ax_i \rangle}_{} = \lambda_i \langle x_j, x_i \rangle$ $\Rightarrow (\lambda_i - \lambda_j)\underbrace{\langle x_j, x_i \rangle}_{=0 \text{ if } \lambda_i \neq \lambda_j} = 0$

def. of adjoint $\longrightarrow$ $\langle A^+ x_j, x_i \rangle = \langle \lambda_j^* x_j, x_i \rangle = \lambda_j \langle x_j, x_i \rangle$