Dot / Inner Products, Adjoints, Orthogonal / Unitary & Symmetric / Skew-Symmetric (Hermitian / Skew-Hermitian) Matrices

Let's compare the standard product in $\mathbb{R}^n$, a bilinear function

$$\boxed{\begin{aligned} &g_0 : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \\ &g_0(\vec{u}, \vec{v}) \overset{\text{def}}{=} \vec{u} \cdot \vec{v} \overset{\text{def}}{=} u_1 v_1 + \cdots + u_n v_n \end{aligned}}$$

with the standard dot product in $\mathbb{C}^n$,

$$\boxed{\begin{aligned} &g_0 : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} \\ &g_0(\vec{z}, \vec{\omega}) \overset{\text{def}}{=} \vec{z} \cdot \vec{\omega} \overset{\text{def}}{=} z_1 \overline{\omega}_1 + \cdots + z_n \overline{\omega}_n \end{aligned}}$$

where

$$\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} a_1 + ib_1 \\ \vdots \\ a_n + ib_n \end{pmatrix} \in \mathbb{C}^n$$

$$\vec{\omega} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix} = \begin{pmatrix} c_1 + id_1 \\ \vdots \\ c_n + id_n \end{pmatrix} \in \mathbb{C}^n$$

and

$$\boxed{\overline{z} = \overline{a + ib} \overset{\text{def}}{=} a - ib} \quad (\text{complex conjugation})$$

ex. 1 Consider $\vec{z} = \begin{pmatrix} -i \\ 4+i \\ -2-2i \end{pmatrix}$, $\vec{w} = \begin{pmatrix} 2+4i \\ -7i \\ -1 \end{pmatrix} \in \mathbb{C}^3$.

Then, their Hermitian inner product is.

$$g_0(\vec{z}, \vec{w}) \equiv \vec{z} \cdot \vec{w} = \begin{pmatrix} -i \\ 4+i \\ -2-2i \end{pmatrix} \cdot \begin{pmatrix} 2+4i \\ -7i \\ -1 \end{pmatrix}$$

$$= -i\,\overline{(2+4i)} + (4+i)\overline{(-7i)} + (-2-2i)\overline{(-1)}$$

$$= -i(2-4i) + (4+i)(7i) + (-2-2i)(-1)$$

$$= -2i + 4i^2 + 28i + 7i^2 + 2 + 2i$$

$$= (-4-7+2) + (\cancel{-2}+28+\cancel{2})i$$

$$= -9+28i$$

Also,

$$g_0(\vec{z}, \vec{z}) \equiv \vec{z} \cdot \vec{z} = (-i)\overline{(-i)} + (4+i)\overline{(4+i)} + (-2-2i)\overline{(-2-2i)}$$

$$= -i \cdot i + (4+i)(4-i) + (-2-2i)(-2+2i$$

$$= 1 + 16 + \cancel{4i} - \cancel{4i} - i^2 + (-2)^2 + \cancel{4i} - \cancel{4i} + (-4)i^2$$

$$= 17+1+4+4 = \boxed{26}$$

The complex dot product is also called the Hermitian inner product on $\mathbb{C}^n$, and unlike the real g_0, is sesquilinear. Let us explain:

Def: A bilinear function

$$\boxed{g : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n}$$

or more generally

$$\boxed{g : V \times W \longrightarrow Z}$$

for any vector spaces V, W, Z over the same field of numbers $F = \mathbb{R}$ or $\mathbb{C}$, is simply linear in each of its two variables,

$$\boxed{\begin{aligned} g(\vec{u}_1 + \vec{u}_2, \vec{v}) &= g(\vec{u}_1, \vec{v}) + g(u_2, \vec{v}) \\ g(c\vec{u}, \vec{v}) &= c\, g(\vec{u}, \vec{v}) \end{aligned}}$$

& similarly with the other slot. $\boxtimes$

ex. The standard dot product in $\mathbb{R}^n$ is bilinear:

$$g_0(\vec{u}_1+\vec{u}_2,\vec{v}) = (\vec{u}_1+\vec{u}_2)\cdot\vec{v}$$

$$= \sum_{i=1}^{n}(u_{1i}+u_{2i})v_i$$

$$= \sum_{i=1}^{n}u_{1i}v_i + \sum_{i=1}^{n}u_{2i}v_i$$

$$= \vec{u}_1\cdot\vec{v} + \vec{u}_2\cdot\vec{v}$$

and

$$g_0(c\vec{u},\vec{v}) = (c\vec{u})\cdot\vec{v}$$

$$= \sum_{i=1}^{n}(cu_i)v_i$$

$$= \sum_{i=1}^{n}c(u_iv_i)$$

$$= c\sum_{i=1}^{n}u_iv_i$$

$$= c(\vec{u}\cdot\vec{v})$$

$$= cg_0(\vec{u},\vec{v})$$

And since $\vec{u}\cdot\vec{v} = \sum_{i=1}^{n}u_iv_i = \sum_{i=1}^{n}v_iu_i = \vec{v}\cdot\vec{u}$

(i.e. g_0 is symmetric, $g_0(\vec{u},\vec{v}) = g_0(\vec{v},\vec{u})$), we see that

$$g_0(\vec{u}, \vec{v}_1+\vec{v}_2) = g_0(\vec{v}_1+\vec{v}_2, \vec{u}) \quad \text{(by symmetry)}$$
$$= g_0(\vec{v}_1,\vec{u}) + g_0(\vec{v}_2,\vec{u}) \quad \text{(linearity in 1st variable)}$$
$$= g_0(\vec{u},\vec{v}_1) + g_0(\vec{u},\vec{v}_2) \quad \text{(symmetry)}$$

&

$$g_0(\vec{u}, c\vec{v}) = g_0(c\vec{v}, \vec{u}) \quad \text{(symmetry)}$$
$$= c\,g_0(\vec{v},\vec{u}) \quad \text{(linearity in 1st var.)}$$
$$= c\,g_0(\vec{u},\vec{v}) \quad \text{(symmetry)}$$

so g_0 is also linear in its second variable. $\boxtimes$

By contrast g_0 on $\mathbb{C}^n$ is <u>sesquilinear</u>, meaning it is <u>linear in its 1st variable</u>

$$\boxed{\begin{array}{ll} (1) & g_0(\alpha\vec{z},\vec{w}) = \alpha\, g_0(\vec{z},\vec{w}), \quad \alpha\in\mathbb{C},\ \vec{z},\vec{w}\in\mathbb{C}^n \\ (2) & g_0(\vec{z}_1+\vec{z}_2,\vec{w}) = g_0(z_1,\vec{w}) + g_0(\vec{z}_2,\vec{w}) \end{array}}$$

but <u>conjugate-linear in its 2nd variable</u>

complex conjugate of $\alpha\in\mathbb{C}$

$$\boxed{\begin{array}{ll} (3) & g_0(\vec{z},\alpha\vec{w}) = \boxed{\overline{\alpha}}\, g_0(\vec{z},\vec{w}) \\ (4) & g_0(\vec{z},\vec{w}_1+\vec{w}_2) = g_0(\vec{z},\vec{w}_1) + g_0(\vec{z},\vec{w}) \end{array}}$$

We prove (3) & leave the rest as an exercise:

complex conjugate

$$g_0(\vec{z},\alpha\vec{w}) \overset{\text{def}}{=} \sum_{i=1}^{n} z_i\,\overline{(\alpha w_i)}$$

$$= \sum_{i=1}^{n} z_i\,\overline{\alpha}\,\overline{w_i}$$

$$= \overline{\alpha}\sum_{i=1}^{n} z_i\,\overline{w_i}$$

$$= \overline{\alpha}\,\vec{z}\cdot\vec{w} = \overline{\alpha}\, g_0(\vec{z},\vec{w}).$$

bec. $\overline{zw} = \overline{(a+ib)(c+id)}$

$= \overline{(ac-bd) + i(ad+bc)}$

$= (ac-bd) - i(ad+bc)$

$= (a-ib)(c-id)$

$= \overline{(a+ib)}\,\overline{(c+id)}$

$= \overline{z}\,\overline{w}$

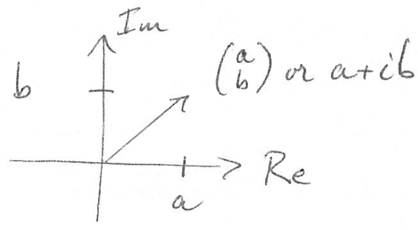
<u>Remark</u>: The reason for the appearance of complex conjugation in the Hermitian dot product on $\mathbb{C}^n$ is because, in $\mathbb{C}$, for any

$$z = a + ib \in \mathbb{C}$$

we have

$$|z| \overset{\text{def}}{=} \sqrt{a^2+b^2} \qquad \text{(complex \underline{modulus})}$$

which is the length of the corresponding vector $\binom{a}{b} \in \mathbb{R}^2$



which satisfies

$$\boxed{\begin{aligned} z\bar{z} &= (a+ib)\overline{(a+ib)} \\ &= (a+ib)(a-ib) \\ &= a^2+b^2+\cancel{iab}-\cancel{iab} \\ &= a^2+b^2 = |z|^2 \end{aligned}}$$

☒

Symmetric / Skew-Symmetric & Hermitian / Skew-Hermitian (coming for matrix addition)

Def. We call a matrix $A \in M_n(\mathbb{R})$ symmetric if

$$\boxed{A^T = A}$$

and skew-symmetric if

$$\boxed{A^T = -A}$$

In the first case, $\boxed{A^T = A \iff a_{ij} = a_{ji} \text{ for all } i,j}$

& in the second case, $A^T = -A \iff a_{ij} = -a_{ji}$ for all i,j

note, if $i=j$, $a_{ii} = -a_{ii}$ $\iff$ all $a_{ii} = 0$

⊠

ex. $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$ is symmetric, while

$B = \begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{pmatrix}$ is skew-symmetric.

⊠

<u>Def</u> We call a complex matrix $A \in M_n(\mathbb{C})$ <u>Hermitian</u> if

$$\boxed{A^* = A}$$

and <u>skew-Hermitian</u> if

$$\boxed{A^* = -A}$$

Note again,

$A^* = A \iff a_{ij} = \overline{a}_{ji}$ (hence $a_{ii} = \overline{a}_{ii}$, i.e. $a_{ii} \in \underline{\mathbb{R}}$)

$A^* = -A \iff a_{ij} = -\overline{a}_{ji}$ (hence $a_{ii} = -\overline{a}_{ii}$

$\iff \underbrace{a_{ii} + \overline{a}_{ii}}_{= 2\mathrm{Re}\, a_{ii}} = 0$

$\iff a_{ii} = bi$ is purely imaginary)

index

$\sqrt{-1}$

<u>ex.</u> $A = \begin{pmatrix} 2 & 2+i & 4 \\ 2-i & 3 & i \\ 4 & -i & 1 \end{pmatrix}$ is Hermitian.

$B = \begin{pmatrix} -i & 2+i \\ -2+i & 0 \end{pmatrix}$ is skew-Hermitian.

$\boxtimes$

Def. Let

$$\boxed{\begin{aligned} Sym_n(\mathbb{R}) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{R}) \mid A^T = A\} \quad \leftarrow \text{set of all symm. } n\times n \text{ real matrices} \\ Skew_n(\mathbb{R}) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{R}) \mid A^T = -A\} \quad \leftarrow \text{set of all skew-sym. } n\times n \text{ matrices} \end{aligned}}$$

and analogously

$$\boxed{\begin{aligned} Sym_n(\mathbb{C}) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{C}) \mid A^* = A\} \quad \leftarrow \text{set of all Hermitian matrices} \\ Skew_n(\mathbb{C}) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{C}) \mid A^* = -A\} \quad \leftarrow \text{set of all skew-Herm. matrices} \end{aligned}}$$

$\boxtimes$

Prop. $Sym_n(\mathbb{R})$ & $Skew_n(\mathbb{R})$ are (real) subspaces of the vector space $M_n(\mathbb{R})$, and

$$\boxed{M_n(\mathbb{R}) = Sym_n(\mathbb{R}) \oplus Skew_n(\mathbb{R})}$$

Similarly, $Sym_n(\mathbb{C})$ and $Skew_n(\mathbb{C})$ are real subspaces of the (complex n^2-dim, & ∴ real $2n^2$-dim real) vector space $M_n(\mathbb{C})$, &

$$\boxed{M_n(\mathbb{C}) = \underbrace{Sym_n(\mathbb{C}) \oplus Skew_n(\mathbb{C})}_{\text{real subspaces}}}$$

Exercise: Prove that $Sym_n(F)$ & $Skew_n(F)$ are real subspaces of $M_n(F)$, for each of $F = \mathbb{R}$ or $\mathbb{C}$. $\boxtimes$

pf: Will prove the other half, that

$$M_n(F) = Sym_n(F) \oplus Skew_n(F)$$

namely

$$A = \underbrace{\tfrac{1}{2}(A + A^*)}_{A_{sym}} + \underbrace{\tfrac{1}{2}(A - A^*)}_{A_{skew}}$$

for

$$\begin{aligned} A^*_{sym} = \left(\tfrac{1}{2}(A+A^*)\right)^* &= \tfrac{1}{2}(A^* + A^{**}) \\ &= \tfrac{1}{2}(A^* + A) \\ &= \tfrac{1}{2}(A + A^*) \\ &= A_{sym} \end{aligned}$$

and

$$\begin{aligned} A^*_{skew} = \left(\tfrac{1}{2}(A - A^*)\right)^* &= \tfrac{1}{2}(A^* - A^{**}) \\ &= \tfrac{1}{2}(A^* - A) \\ &= -\tfrac{1}{2}(A - A^*) = -A_{skew} \end{aligned}$$

This shows that

$$M_n(F) \subseteq \mathrm{Sym}_n(F) \oplus \mathrm{Skew}_n(F)$$

The reverse inclusion $\supseteq$ is the previous exercise plus the fact that

$$\mathrm{Sym}_n(F) \cap \mathrm{Skew}_n(F) = \{\underbrace{O}_{\text{zero matrix}}\}$$

since any $A \in \mathrm{Sym}_n(F) \cap \mathrm{Skew}_n(F)$ is both symmetric <u>and</u> skew-symmetric, so satisfies

$$A = A^* = -A$$

$$\Longrightarrow \quad 2A = A + A = 0$$

$$\Longrightarrow \quad A = \tfrac{1}{2}(2A) = \tfrac{1}{2}O = O \qquad \underline{\text{QED}}$$

<u>ex.</u> $$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \tfrac{1}{2}(A + A^T) + \tfrac{1}{2}(A - A^T)$$

$$= \tfrac{1}{2}\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\right) + \tfrac{1}{2}\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\right)$$

$$= \underbrace{\tfrac{1}{2}\begin{pmatrix} 2 & 5 \\ 5 & 8 \end{pmatrix}}_{\text{symmetric}} + \underbrace{\tfrac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\text{skew}} \qquad \boxtimes$$

<u>Exercise:</u> Decompose $A = \begin{pmatrix} 1+2i & 3+4i \\ 5+6i & 7+8i \end{pmatrix}$ into a sum of Hermitian & skew-Hermitian parts

$$A = A_{sym} + A_{skew}$$

$\boxtimes$

Orthogonal & Unitary (coming from matrix product)

Consider a real matrix $A \in M_n(\mathbb{R})$.

Then

$$A^T A = \begin{pmatrix} -\vec{a}_1- \\ \vdots \\ -\vec{a}_n- \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix}$$

A^T has rows the old columns of A

$$= \begin{pmatrix} \vec{a}_1 \cdot \vec{a}_1 & \cdots & \vec{a}_1 \cdot \vec{a}_n \\ \vdots & \ddots & \vdots \\ \vec{a}_n \cdot \vec{a}_1 & \cdots & \vec{a}_n \cdot \vec{a}_n \end{pmatrix}$$

$$= \begin{pmatrix} g_0(\vec{a}_1, \vec{a}_1) & \cdots & g_0(\vec{a}_1, \vec{a}_n) \\ \vdots & \ddots & \vdots \\ g_0(\vec{a}_n, \vec{a}_1) & \cdots & g_0(\vec{a}_n, \vec{a}_n) \end{pmatrix}$$

let's make the presence of the standard dot product felt

and compare with a complex matrix $A \in M_n(\mathbb{C})$,

$$\bar{A}^T A = \begin{pmatrix} -\bar{\vec{a}}_1- \\ \vdots \\ -\bar{\vec{a}}_n- \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix}$$

real dot

$$= \begin{pmatrix} \bar{\vec{a}}_1 \cdot \vec{a}_1 & \cdots & \bar{\vec{a}}_1 \cdot \vec{a}_n \\ \vdots & \ddots & \vdots \\ \bar{\vec{a}}_n \cdot \vec{a}_1 & \cdots & \bar{\vec{a}}_n \cdot \vec{a}_n \end{pmatrix}$$

Hermitian dot product

$$\overset{df}{=} \begin{pmatrix} g_0(\vec{a}_1, \vec{a}_1) & \cdots & g_0(\vec{a}_1, \vec{a}_n) \\ \vdots & & \vdots \\ g_0(\vec{a}_n, \vec{a}_1) & \cdots & g_0(\vec{a}_n, \vec{a}_n) \end{pmatrix}$$

<u>Def</u> The complex conjugate transpose of a matrix $A \in M_n(\mathbb{C})$ is called the <u>adjoint matrix</u> of A, & denoted

complex conjugates of the entries in A

$$\boxed{\begin{aligned} A^* &\overset{def}{=} \overline{A}^T = \begin{pmatrix} \overline{a}_{11} & \cdots & \overline{a}_{n1} \\ \vdots & & \vdots \\ \overline{a}_{1n} & \cdots & \overline{a}_{nn} \end{pmatrix} \\ &= \begin{pmatrix} - & \overline{\vec{a}}_1 & - \\ & \vdots & \\ - & \overline{\vec{a}}_n & - \end{pmatrix} = \begin{pmatrix} | & & | \\ \overline{\vec{A}}_1 & \cdots & \overline{\vec{A}}_n \\ | & & | \end{pmatrix} \end{aligned}}$$ ⊠

conjugates of old ~~rows~~ cols. of A, now rows

complex conjugates of old rows of A, now col's.

From our work above, we see that, in terms of the dot products g_0, respectively over $\mathbb{R}$ & $\mathbb{C}$, we have:

ex. Consider $A = \begin{pmatrix} i & 1-2i \\ 2 & 3+4i \end{pmatrix} \in M_2(\mathbb{C})$.

Then,

$$A^* = \bar{A}^T = \begin{pmatrix} \bar{i} & \bar{2} \\ \overline{1-2i} & \overline{3+4i} \end{pmatrix}$$

$$= \begin{pmatrix} -i & 2 \\ 1+2i & 3-4i \end{pmatrix}$$

Observe, incidentally, that

$$A^*A = \begin{pmatrix} -i & 2 \\ 1+2i & 3-4i \end{pmatrix}\begin{pmatrix} i & 1-2i \\ 2 & 3+4i \end{pmatrix}$$

$$= \begin{pmatrix} \langle -i, 2\rangle \cdot \begin{pmatrix} i \\ 2 \end{pmatrix} & \langle -i, 2\rangle \cdot \begin{pmatrix} 1-2i \\ 3+4i \end{pmatrix} \\ \langle 1+2i, 3-4i\rangle \cdot \begin{pmatrix} i \\ 2 \end{pmatrix} & \langle 1+2i, 3-4i\rangle \begin{pmatrix} 1-2i \\ 3+4i \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} -i^2+4 & -i+2i^2+6+8i \\ i+2i^2+6-8i & 1+\cancel{2i}-\cancel{2i}-2i^2+9+\cancel{12i}-\cancel{12i}-16i^2 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 4+7i \\ 4-7i & 28 \end{pmatrix}$$

which is the complex analog of a symmetric matrix ($A^T = A$), namely a <u>self-adjoint</u> matrix

$(A^*A)^* = A^*A$

<u>Thm. 1</u> For any matrix $A \in M_n(\mathbb{R})$ we have

$$\boxed{\begin{aligned} A^T A = I_n &\iff \text{the \underline{columns} of } A \text{ are \underline{orthonormal in } } \mathbb{R}^n \\ &\iff \text{the \underline{rows} of } A \text{ are \underline{orthonormal in } } \mathbb{R}^n \end{aligned}}$$

<u>with respect to the standard dot product</u> in $\mathbb{R}^n$.

Similarly, for any $A \in M_n(\mathbb{C})$,

$$\boxed{\begin{aligned} A^* A = I_n &\iff \text{the \underline{columns} of } A \text{ are \underline{orthonormal in } } \mathbb{C}^n \\ &\iff \text{the \underline{rows} of } A \text{ are \underline{orthonormal in } } \mathbb{C}^n \end{aligned}}$$

<u>with respect to the Hermitian dot product in $\mathbb{C}^n$.</u> $\boxtimes$

<u>Def</u> The term <u>orthonormal</u> means two things:

$$\boxed{\begin{array}{ll} (1) \quad \vec{u}\cdot\vec{v} = 0 & \left(\vec{u} \;\&\; \vec{v} \in \mathbb{R}^n \text{ are } \underline{\text{orthogonal}} \; (\underline{\text{with}} \text{ resp. to } g_0)\right) \\ (2) \quad \|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} = 1 \;\&\; \|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = 1 & \left(\vec{u} \;\&\; \vec{v} \text{ are } \underline{\text{normal}}(\underline{\text{-ized}}), \text{ i.e. are } \underline{\text{unit length}}, \text{ again, } \underline{\text{with}} \text{ resp. to } g_0\right) \end{array}}$$

We say <u>$\vec{u}$ & $\vec{v}$ are orthonormal</u> in this case.

Similarly with <u>$\vec{z}, \vec{w} \in \mathbb{C}^n$</u>; but now <u>with respect to the Hermitian dot product</u>:

$$\boxed{\begin{array}{ll} (1) \quad g_0(\vec{z}, \vec{w}) \left(= \sum_{i=1}^{n} z_i \overline{w_i}\right) = 0 & (\underline{g_0\text{-orthogonality}}) \\ (2) \quad \|\vec{z}\| = \|\vec{w}\| = 1 & (\underline{g_0\text{-normality}}) \\ \quad (\text{i.e. } g_0(\vec{z}, \vec{z}) = g_0(\vec{w}, \vec{w}) = 1) & \end{array}}$$

$\boxtimes$

We can of course say any collection of vectors $\vec{u}_1, \dots, \vec{u}_k \in F^n$ ($\mathbb{R}^n$ or $\mathbb{C}^n$) is orthonormal if

$$\boxed{\vec{u}_i \cdot \vec{u}_j \equiv g_0(\vec{u}_i, \vec{u}_j) = \underbrace{\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}}_{\text{Kronecker delta}}}$$

but: there can't be too many of them, because

Prop. If $\{\vec{u}_1, \dots, \vec{u}_k\}$ is any collection of orthonormal vectors in F^n (w.r.t. g_0), then $\vec{u}_1, \dots, \vec{u}_k$ are linearly independent. (so $k \leq n$)

pf: Suppose $a_1\vec{u}_1 + \cdots + a_k\vec{u}_k = \vec{0}$, $a_i \in F$. Then

true for each i, so all $a_i = 0$. QED

$$\begin{aligned} 0 = \vec{u}_i \cdot \vec{0} &= \vec{u}_i \cdot (a_1\vec{u}_1 + \cdots + a_k\vec{u}_k) \\ &= a_1\underbrace{(\vec{u}_i \cdot \vec{u}_1)}_{=0} + \cdots + a_i\underbrace{(\vec{u}_i \cdot \vec{u}_i)}_{=1} + \cdots + a_k\underbrace{(\vec{u}_i \cdot \vec{u}_k)}_{=0} \\ &= a_i \end{aligned}$$

<u>Def.</u> A square matrix $A \in M_n(\mathbb{R})$ satisfying

$$\boxed{A^T A = I_n}$$

is called <u>orthogonal</u>. The set of all orthogonal matrices in $M_n(\mathbb{R})$ is called the <u>orthogonal group</u> and denoted

$$\boxed{\begin{aligned} O(n) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{R}) \mid A^T A = I_n\} \\ &\overset{\text{Thm. 1}}{=} \{A \in M_n(\mathbb{R}) \mid \underbrace{\vec{a}_i \cdot \vec{a}_j = \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}}_{\text{the columns } \vec{a}_i \text{ of } A \text{ are orthonormal}} \} \end{aligned}}$$

Similarly, a square complex matrix $A \in M_n(\mathbb{C})$ satisfying

$$A^* A = I_n$$

is called <u>unitary</u>, & the set of all unitary matrices in $M_n(\mathbb{C})$ is called the <u>unitary group</u> and is denoted

$$\boxed{\begin{aligned} U(n) &\overset{\text{def}}{=} \{A \in M_n(\mathbb{C}) \mid A^*A = I_n\} \\ &= \{A \in M_n(\mathbb{C}) \mid \underbrace{\vec{a}_i \cdot \vec{a}_j = \delta_{ij}}\} \end{aligned}}$$

the col's. of A are orthonormal w.r.t. the <u>Hermitian</u> inner prod.

<u>Exercise:</u> Use property (9) of the adjoint in Thm 3, p. 12.5 to show that

Hermitian

$$A \in U(n) \iff (A\vec{z}) \cdot (A\vec{w}) = \vec{z} \cdot \vec{w}$$

for all $\vec{z}, \vec{w} \in \mathbb{C}^n$.

$\boxtimes$

<u>Remark</u> Thus, $A \in U(n) \iff \underbrace{\|A\vec{z}\| = \|\vec{z}\|}$

A is an <u>isometry</u>, i.e. preserves lengths

Exercise: Show that
$$\boxed{\begin{array}{l} U(n) \subseteq GL(n, \mathbb{C}) \\ O(n) \subseteq GL(n, \mathbb{C}) \end{array}}$$
and

In fact,

and
$$\boxed{\begin{array}{lcl} A \in U(n) & \Longleftrightarrow & A^{-1} = A^* \\ A \in O(n) & \Longleftrightarrow & A^{-1} = A^T \end{array}}$$
$\boxtimes$

Exercise: Show that U(n) and O(n) are groups, that is they are closed under multiplication & inversion:

$$\boxed{\begin{array}{lll} (1) & A \in O(n)/U(n) \implies A^{-1} \in O(n)/U(n) \\ (2) & A, B \in O(n)/U(n) \implies AB \in O(n)/U(n) \\ (3) & I_n \in O(n) \ \& \ U(n) \end{array}}$$

Thus, O(n) is a subgroup of $GL(n, \mathbb{R})$, & U(n) of $GL(n, \mathbb{C})$. $\boxtimes$

<u>Exercise:</u> Use the properties of the determinant

$$\det(AB) = (\det A)(\det B)$$

$$\det I_n = 1$$

to show that

$$\boxed{\det A = \pm 1 \quad \text{for all } A \in O(n)}$$

and

$$\boxed{\underbrace{|\det A|}_{\substack{\text{complex}\\ \text{modulus}\\ |z| = |a+ib| \\ = \sqrt{a^2+b^2}}} = 1 \quad \text{for all } A \in U(n)}$$

Indeed, letting $\boxed{\mathbb{R}^* \overset{\text{def}}{=} \mathbb{R} \setminus \{0\}}$, $\boxed{\mathbb{C}^* = \mathbb{C} \setminus \{0\}}$,
which are <u>groups under multiplication</u>,

$$\boxed{\begin{aligned} \det &: GL(n,\mathbb{R}) \longrightarrow \mathbb{R}^* \\ \det &: GL(n,\mathbb{C}) \longrightarrow \mathbb{C}^* \end{aligned}}$$

getting ahead of myself.

~~1×1 unitary matrices = unit length cpx. #s = unit circle in $\mathbb{C}$.~~

are group homomorphisms, with

$$\boxed{\underbrace{\ker(\det)}_{\text{kernel}} \overset{\text{def}}{=} \{A \in GL(n,\mathbb{R}) \mid \det A = 1\} \overset{\text{def}}{=} SL(n,\mathbb{R})}$$

<u>special linear group</u> (in $\mathbb{R}$)

and

$$\boxed{\ker(\det) \overset{\text{def}}{=} \{A \in GL(n,\mathbb{C}) \mid \det A = 1\} = SL(n,\mathbb{C})}$$

<u>complex special linear group</u>

The <u>special linear orthogonal matrices</u> are called the <u>special orthogonal matrices</u>, & the <u>special orthogonal group</u> is their set,

$$\boxed{SO(n) \overset{\text{def}}{=} SL(n,\mathbb{R}) \cap O(n)}$$

Similarly the <u>special ~~orthogonal~~ or unitary group</u> is

$$\boxed{SU(n) = SL(n,\mathbb{C}) \cap U(n)}$$

ex. <u>All rotations</u> $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ <u>are orthogonal</u>, in fact <u>special orthogonal</u>.

In fact,

$$\boxed{\text{All rotations} = SO(2)}$$

pf: Clearly $R_\theta \in O(2)$ because its columns are orthonormal,

$$\vec{r}_1 \cdot \vec{r}_2 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \cdot \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

$$= -\sin\theta\cos\theta + \sin\theta\cos\theta$$

$$= 0$$

$$\vec{r}_1 \cdot \vec{r}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \sin^2\theta + \cos^2\theta = 1$$

$$\vec{r}_2 \cdot \vec{r}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \cdot \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = \sin^2\theta + \cos^2\theta = 1$$

or alternatively bec.

$$R_\theta^T R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

Moreover,

$$\det R_\theta = \det\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$$= \sin^2\theta + \cos^2\theta$$

$$= 1$$

so $R_\theta \in O(2) \cap SL(2,\mathbb{R}) = SO(2)$.

<u>Exercise</u> Show that <u>all reflections R_ℓ in $M_2(\mathbb{R})$ are orthogonal,</u>

$$\boxed{R_\ell \in O(2)}$$

<u>but that</u>

$$\boxed{\det R_\ell = -1} \quad (\text{so } R_\ell \notin SO(2))$$

In fact, $\boxed{O(2) = \underbrace{SO(2)}_{\text{rotations}} \cup \underbrace{O^-(2)}_{\text{reflections}}}$.

For example the reflection $R_\ell = \frac{-1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}$ across $y = -3x$ we computed in example 7.19 ("Bases, Coords, & Reps."): $\vec{r}_1 \cdot \vec{r}_2 = \frac{-1}{5}\begin{pmatrix} 4 \\ 3 \end{pmatrix} \cdot \frac{-1}{5}\begin{pmatrix} 3 \\ -4 \end{pmatrix} = \left(\frac{-4}{5}\right)\left(\frac{-3}{5}\right) + \left(\frac{-3}{5}\right)\left(\frac{4}{5}\right)$

& $\vec{r}_i \cdot \vec{r}_i = 1$. $\qquad = \frac{12-12}{25} = 0$, $\boxtimes$

Thm. 2 Properties of the Transpose & Adjoint

For any $A, B \in M_n(\mathbb{R})$ (or $M_{m,n}(\mathbb{R})$ more generally)

(1) $(A+B)^T = A^T + B^T$

(2) $(cA)^T = cA^T \quad (c \in \mathbb{R})$

(3) $(AB)^T = B^T A^T$

(4) $(A^T)^T = A \quad (A \in GL(n,\mathbb{R}))$

(5) $(A^T)^{-1} = (A^{-1})^T$

(6) $\operatorname{rank}(A^T A) = \operatorname{rank} A$

$N(A^T A) = N(A)$

(7) $\operatorname{rank} A^T = \operatorname{rank} A$

(8) $R(A^T) = N(A)^\perp$

$N(A^T) = R(A)^\perp$

(9) $(A\vec{x}) \cdot \vec{y} = \vec{x} \cdot (A^T \vec{y})$

For any $A, B \in M_n(\mathbb{C})$ (or $M_{m,n}(\mathbb{C})$ more generally)

(1) – (9) apply to A^* (instead of A^T), w/ (2) $(cA)^* = \bar{c}A^*$

Thm. 3 For any $A, B \in M_n(\mathbb{C})$ (or $M_{m,n}(\mathbb{C})$ more generally, whenever AB makes sense),

$$(1)\quad (A+B)^* = A^* + B^*$$

$$(2)\quad (cA)^* = \bar{c}A^* \qquad (c \in \mathbb{C})$$

$$(3)\quad (AB)^* = B^*A^*$$

$$(4)\quad (A^*)^* = A$$

$$(5)\quad (A^*)^{-1} = (A^{-1})^*, \qquad A \in GL(n, \mathbb{F})$$

$$(6)\quad \operatorname{rank}(A^*A) = \operatorname{rank} A$$

$$\operatorname{null}(A^*A) = \operatorname{null} A$$

In fact

$$N(A^*A) = N(A)$$

$$(7)\quad \operatorname{rank}(A^*) = \operatorname{rank} A$$

$$(8)\quad R(A^*) = N(A)^{\perp}$$

w.r.t. g_0 Hermitian

$$N(A^*) = R(A)^{\perp}$$

$$(9)\quad (A\vec{z}) \cdot \vec{w} = \vec{z} \cdot (A^*\vec{w}) \quad \text{for all } \vec{z}, \vec{w} \in \mathbb{C}^n$$

pf: We'll prove the complex case only, since for any real $A \in M_n(\mathbb{R})$, can be viewed as a complex matrix with no imaginary part, &

$$A^* = A^T$$

in that case. Let's start at the top:

(1) $$(A+B)^* = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nn}+b_{nn} \end{pmatrix}^*$$

$$= \begin{pmatrix} \overline{a_{11}+b_{11}} & \cdots & \overline{a_{n1}+b_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}+b_{1n}} & \cdots & \overline{a_{nn}+b_{nn}} \end{pmatrix}$$

Exercise: Prove that for any $z = a+ib$, $w = c+id \in \mathbb{C}$,

$$\boxed{\overline{z+w} = \overline{z} + \overline{w}.}$$

$$= \begin{pmatrix} \overline{a_{11}}+\overline{b_{11}} & \cdots & \overline{a_{n1}}+\overline{b_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}}+\overline{b_{1n}} & \cdots & \overline{a_{nn}}+\overline{b_{nn}} \end{pmatrix}$$

$$= \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \cdots & \overline{a_{nn}} \end{pmatrix} + \begin{pmatrix} \overline{b_{11}} & \cdots & \overline{b_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{b_{1n}} & \cdots & \overline{b_{nn}} \end{pmatrix}$$

$$= A^* + B^*$$

(2) $$(cA)^* = \left(c \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \right)^*$$

$$= \begin{pmatrix} \overline{ca_{11}} & \cdots & \overline{ca_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{ca_{1n}} & \cdots & \overline{ca_{nn}} \end{pmatrix}$$

$$= \begin{pmatrix} \overline{c}\,\overline{a_{11}} & \cdots & \overline{c}\,\overline{a_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{c}\,\overline{a_{1n}} & \cdots & \overline{c}\,\overline{a_{nn}} \end{pmatrix}$$

Exercise: Prove that for any $z = a+ib$, $\omega = c+id \in \mathbb{C}$, $\boxed{\overline{z\omega} = \overline{z}\,\overline{\omega}}$.

$$= \overline{c} \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \cdots & \overline{a_{nn}} \end{pmatrix}$$

$$= \overline{c} A^*$$

(3) $$(AB)^* = \begin{pmatrix} | & & | \\ A\vec{b}_1 & \cdots & A\vec{b}_n \\ | & & | \end{pmatrix}^*$$

$$= \begin{pmatrix} - & \overline{A\vec{b}_1} & - \\ & \vdots & \\ - & \overline{A\vec{b}_n} & - \end{pmatrix}$$

$$= \begin{pmatrix} - & \overline{A}\,\overline{\vec{b}_1} & - \\ & \vdots & \\ - & \overline{A}\,\overline{\vec{b}_n} & - \end{pmatrix}$$

Exercise: Prove that for any $A \in M_{m,n}(\mathbb{C})$ & any $\vec{z} \in \mathbb{C}^n$,

$$\boxed{\underbrace{\overline{A\vec{z}}}_{\text{conjugate every element in the column vector } A\vec{z}} = \underbrace{\overline{A}}_{\text{conjugate all components, } a_{ij} \text{ of } A.}\ \underbrace{\overline{\vec{z}}}_{\text{conjugate all } z_i}}$$

$$= \begin{pmatrix} | & & | \\ \overline{A}\,\overline{\vec{b}_1} & \cdots & \overline{A}\,\overline{\vec{b}_n} \\ | & & | \end{pmatrix}^T$$

$$= (\overline{A}\,\overline{B})^T = \overline{B}^T \overline{A}^T = B^* A^*$$

$(AB)^T = B^T A^T$

pf: $(AB)^T = \begin{pmatrix} | & & | \\ A b_1 & \cdots & A b_n \\ | & & | \end{pmatrix}^T$

$= \begin{pmatrix} - & A\vec{b}_1 & - \\ & \vdots & \\ - & A\vec{b}_n & - \end{pmatrix}$

exercise

$= \begin{pmatrix} - & \vec{b}_1^T A^T & - \\ & \vdots & \\ - & \vec{b}_n^T A^T & - \end{pmatrix}$

exercise

$= B^T A^T$

1st transpose $\bar{A}$, then conjugate

Exercise

$$(4)\quad (A^*)^* = \overline{(A^*)}^T = \overline{(\bar{A}^T)}^T = (\bar{\bar{A}}^T)^T = A$$

same as conjugating before transposing

$\left((\bar{A})^T = \overline{(A^T)} \right)$

Exercise: For any $z = a + ib \in \mathbb{C}$, show that $\bar{\bar{z}} = z$. (very easy).

(5) Exercise.

(6) First, we note that

$$N(A) \subseteq N(A^*A)$$

since

$$\underbrace{A\vec{z} = \vec{0}}_{\iff z \in N(A)} \implies \underbrace{(A^*A)\vec{z} = A^*(A\vec{z}) = A^*\vec{0} = \vec{0}}_{\iff \vec{z} \in N(A^*A)}$$

We prove $N(A) = N(A^*A)$ by a double inclusion.

Conversely,

$$N(A^*A) \subseteq N(A)$$

since if $(A^*A)\vec{z} = \vec{0}$, then, using (9),

$$(A\vec{z})\cdot(A\vec{z}) = \vec{z}\cdot(A^*A\vec{z})$$
$$= \vec{z}\cdot\vec{0} \quad \text{by assump.}$$
$$= 0$$

Hermitian

which shows that

$$\|A\vec{z}\|^2 = (A\vec{z}) \cdot (A\vec{z}) = 0$$

But: Exercise: for any $\vec{w} \in \mathbb{C}^n$,

$$\|\vec{w}\|^2 = \vec{w} \cdot \vec{w} = 0 \iff \vec{w} = \vec{0}$$

Hermitian

The proof is very similar to the real case, $\vec{x} \cdot \vec{x} = 0 \iff \vec{x} = \vec{0}$ in $\mathbb{R}^n$.

so we conclude that $A\vec{z} = \vec{0}$, and $\vec{z} \in N(A)$. Thus,

$$N(A^*A) = N(A).$$

Then, by the Rank-Nullity theorem, since

$$A : \mathbb{C}^n \to \mathbb{C}^m$$

$$A^* : \mathbb{C}^m \to \mathbb{C}^n$$

$$A^*A : \mathbb{C}^n \to \mathbb{C}^n$$

we get

$$\operatorname{rank}(A^*A) = n - \dim N(A^*A) = n - \dim N(A)$$
$$= \operatorname{rank} A$$

Exercise: For any $A \in M_{m,n}(F)$, $B \in M_{p,m}(F)$, $F = \mathbb{R}$ or $\mathbb{C}$, show that

$$\boxed{\operatorname{rank}(AB) \le \min\{\operatorname{rank} A, \operatorname{rank} B\}}$$

Hint: Consider a basis for $R(B)$ and see whether B is onto or not.

(7) By the previous exercise & (6),

$$\operatorname{rank} A \overset{(6)}{=} \operatorname{rank}(A^*A) \overset{ex.}{\le} \operatorname{rank} A^*$$

and

$$\operatorname{rank} A^* \overset{(6)}{=} \operatorname{rank}((A^*A)^*) \overset{(3)}{=} \operatorname{rank}(A^*A^{**}) \overset{(4)}{=} \operatorname{rank}(A^*A) \overset{ex.}{\le} \operatorname{rank} A$$

so $\operatorname{rank} A^* = \operatorname{rank} A$.

(8) Follows from (9):

$$\vec{z} \in N(A^*) \overset{\text{def}}{\iff} A^*\vec{z} = \vec{0}$$

$$\iff 0 = \vec{0} \cdot \vec{w} \quad \text{for all } \vec{w} \in \mathbb{C}^n$$

$$\overset{\text{assump}}{=} (A^*\vec{z}) \cdot \vec{w}$$

$$\overset{(4),(9)}{=} \vec{z} \cdot \underbrace{(A\vec{w})}_{\in R(A)}$$

$$\iff \vec{z} \perp R(A)$$

or

$$\vec{z} \in R(A)^{\perp}$$

with resp. to the Hermitian inner prod.

Thus, $N(A^*) = R(A)^{\perp}$. From this and (4) we get

$$N(A) = N(A^{**}) \overset{\text{above result applied to } A^*}{=} R(A^*)^{\perp}$$

(9) $$(A\vec{z}) \cdot \vec{w} = \vec{b} \cdot \vec{w} \qquad (\vec{b} \overset{\text{def}}{=} A\vec{z})$$

Hermitian

$$= \sum_{i=1}^{m} b_i \bar{w}_i$$

$$= \sum_{i,j=1}^{m,n} a_{ij} z_j \bar{w}_i \qquad \vec{b} = A\vec{z} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j} z_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} z_j \end{pmatrix}$$

$$= \sum_{i,j=1}^{m,n} z_j \overline{(\bar{a}_{ij} w_i)}$$

$$= \sum_{j=1}^{m,n} z_j \overline{(\bar{A}^T \vec{w})_j} = \vec{z} \cdot (A^*\vec{w}) \qquad \boxtimes$$

ith slot

$$\begin{pmatrix} \vdots \\ \sum_{i=1}^{m} \bar{a}_{ij} w_i \\ \vdots \end{pmatrix} = \begin{pmatrix} \bar{a}_{11} & \cdots & \bar{a}_{m1} \\ \vdots & & \vdots \\ \bar{a}_{1n} & \cdots & \bar{a}_{mn} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}$$

Diagonalizability of Each Class Above

(orthogonal diagonalizability)

Prop. 1 For any $A \in M_n(\mathbb{R})$ TFAE:

real case, with resp. to standard dot prod. on $\mathbb{R}^n$

(1) A is orthogonally diagonalizable

$$\left(\boxed{A = PDP^{-1} \ \& \ P \in O(n)}\right)$$

(2) $\mathbb{R}^n$ has an orthonormal basis β of eigenvectors of A

(3) $\boxed{\mathbb{R}^n = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}}$

$\oplus$ & also $\perp$, i.e. $E_{\lambda_i} \perp E_{\lambda_j}$ orthogonal to each other

(unitary diagonalizability)

Prop. 2 For any $A \in M_n(\mathbb{C})$, TFAE

complex case, w.r.t. the Hermitian inner prod. on $\mathbb{C}^n$

(1) A is unitarily diagonalizable

$$\boxed{(A = PDP^{-1} \ \& \ P \in U(n))}$$

(2) $\mathbb{C}^n$ has an orthonormal basis β of eigenvectors of A (w.r.t. the Hermitian prod. $\langle\,,\rangle$ in $\mathbb{C}^n$)

(3) $\boxed{\mathbb{C}^n = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}}$

$\oplus$ & $\perp$ with resp. to the Hermitian prod.

pf: Will prove the complex case since the real case proceeds entirely analogously. Now, by the change-of-coordinates formula, which proceeds exactly the same in the complex case,

$$[A]_\gamma = M_{\beta\gamma}[A]_\beta M_{\beta\gamma}^{-1}$$

we have, for a diagonalizing basis β for $\mathbb{C}^n$;

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = [A]_\beta$$

$$= M_{\sigma\beta}[A]_\sigma M_{\sigma\beta}^{-1}$$

$$= M_{\beta\sigma}^{-1} A M_{\beta\sigma}$$

where

$$M_{\beta\sigma} = (\vec{b}_1 \cdots \vec{b}_n), \quad \vec{b}_i \in \mathbb{C}^n \text{ eigenvectors of } A$$

& conversely $D = P^{-1}AP \implies \boxed{P = M_{\beta\sigma} = (\vec{b}_1 \cdots \vec{b}_n)}$ where $A\vec{b}_i = \lambda\vec{b}_i$. Thus, it's clear that

(1) $\Longleftrightarrow$ (2), since the $\vec{b}_i$ are mutually orthogonal & of unit length,

$$\vec{b}_i \cdot \vec{b}_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$\Longleftrightarrow$ $M_{\beta\sigma} = P \in U(n)$ by def. $P \in U(n) \Longleftrightarrow$ $P^*P = I \Longleftrightarrow$ the cols of P are orthonormal

$\Longleftrightarrow$ A is unitarily diagonalizable

To see that (~~1~~) $\Longleftrightarrow$ (3), recall Thm. 6, p. 14, 'Eigenvalues & Eigenvectors II,' which says

A is diagonalizable $\Longleftrightarrow$ $[A]_\beta = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

for a basis $\beta = (\vec{b}_1, \ldots, \vec{b}_n)$ of eigenvectors of A

$\Longleftrightarrow$ for the distinct λ_i's,

$\mathbb{C}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$

so if we add the orthonormality condition on β, then (1) $\iff$ (3) because

$$\vec{b}_i \cdot \vec{b}_j = \delta_{ij} \iff E_{\lambda_i} \perp E_{\lambda_j}$$

$E_{\lambda_i} =$ span β_i, β_i a basis for E_{λ_i} of eigenvectors of A corresp. to λ_i

$E_{\lambda_j} =$ span β_j, β_j a basis for E_{λ_j} of eigenvecs. corresp. to λ_j

QED

Both symmetric & skew-symmetric matrices $A \in M_n(\mathbb{R})$ satisfy

$$\boxed{A^T A = A A^T}$$ ← key fact

in the first case bec.

$$A^T = A \implies A^T A = AA = AA^T$$

& in the second bec.

$$A^T = -A \implies A^T A = -AA$$
$$= A(-A)$$
$$= AA^T$$

proof

Similarly with Hermitian & skew-Hermitian matrices $A \in M_n(\mathbb{C})$,

$$\boxed{A^* A = A A^*}$$ ← key fact

$(A^* = A \implies A^* A = AA = AA^*)$

$(A^* = -A \implies A^* A = -AA = AA^*)$

proof

<u>Def:</u> Any matrix $A \in M_n(\mathbb{R})$ satisfying

$$\boxed{A^T A = A A^T}$$

or any $A \in M_n(\mathbb{C})$ satisfying

$$\boxed{A^* A = A A^*}$$

is called <u>normal</u>.

The <u>keypoint here is that both symmetric/ skew-symmetric and orthogonal matrices are normal.</u>

<u>Thm. 4</u> Let A be a normal matrix (in $M_n(\mathbb{R})$ or $M_n(\mathbb{C})$, understanding A^* as A^T in the case of $\mathbb{R}$). Then,

(1) $\boxed{(A\vec{u})\cdot(A\vec{v}) = (A^*\vec{u})\cdot(A^*\vec{v}),}$ for all $\vec{u},\vec{v} \in F^n$

(2) $\boxed{\|A\vec{u}\| = \|A^*\vec{u}\|}$ for all $\vec{u} \in F^n$

(3) $\boxed{A\vec{u} = \lambda\vec{u} \iff A^*\vec{u} = \bar{\lambda}\vec{u}}$

<u>pf:</u> <u>Exercise.</u>

Thm. 5 (Spectral Thm. for Complex Normal Matrices)

Let $A \in M_n(\mathbb{C})$ and let $\boxed{\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}}$ be its spectrum (= all of A's distinct eigenvalues).

Then the following are equivalent, logically:

(1) A is normal,

$$\boxed{A^*A = AA^*}$$

(2) A is unitarily diagonalizable

$$\boxed{\begin{aligned}[A]_\beta &= \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\\ &= M_{\beta\sigma}^{-1} A M_{\beta\sigma},\end{aligned}}$$

where $\boxed{M_{\beta\sigma} = (\vec{b}_1, \ldots, \vec{b}_n) \in U(n)}$ ← orthonormal eigenvector basis

or equivalently, by Prop. 2, p. 13 above

$$\boxed{\mathbb{C}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}}$$

Thus, every Hermitian, skew-Hermitian, and unitary matrix is unitarily diagonalizable,

$$\boxed{\begin{aligned} D &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ &= P^{-1}AP \quad \text{for some } P \in U(n) \\ &\qquad (\text{namely } P = M_{\beta\sigma} = (\vec{b}_1, \ldots, \vec{b}_n) \text{ w/ orthonormal } \underset{\text{eigenvectors}}{\vec{b}_i}) \end{aligned}}$$

and characteristically so, i.e. iff (well, there may potentially be other normal matrices)

Similarly, viewing $O(n) \subseteq U(n)$ w/ imaginary part O, i.e. $\underset{\in M_n(\mathbb{C})}{X} = A + iB$, $A, B \in M_n(\mathbb{R})$, $B = O$,

we may say: every symmetric, skew-symmetric, and orthogonal matrix is unitarily diagonalizable & characteristically so (iff)

over $\mathbb{C}$!!! not $\mathbb{R}$ — pay attention here!

What about over $\mathbb{R}$?

Thm. 6 (Spectral Theorem for Real Normal Matrices) Let $A \in M_n(\mathbb{R})$ and let $\boxed{\sigma(A) = \{\lambda_1, \dots, \lambda_n\}}$ be its spectrum (its distinct eigenvalues). Then TFAE: the following are equivalent

(1) A is normal,

$$\boxed{A^T A = A A^T}$$

iff

(2) A is (not necessarily orthogonally, see below) diagonalizable up to 2×2 rotation blocks,

n_1 of them $\iff$ $\dim E_{\lambda_1} = n_1$ = alg. mult of n_1 etc.

$$\boxed{[A]_\beta = \begin{pmatrix} \lambda_1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \ddots & & & & & \\ & & & & \lambda_k & & & & \\ & & & & & \ddots & & & \\ & & & & & & \lambda_k & & \\ & & & & & & & R_{\theta_1} & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & & R_{\theta_\ell} \end{pmatrix} = P^{-1}AP,}$$

$R_{\theta_1} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \in SO(2)$

$\boxed{P = M_{\beta\sigma} = (\vec{b}_1, \dots, \vec{b}_n)}$ eigenvectors

for some β of eigenvecs.

What about orthogonally diagonalizable matrices?

A: Those are exactly the symmetric matrices.

Thm. 7 $A \in M_n(\mathbb{R})$ is symmetric iff A is orthogonally diagonalizable,

$$\boxed{\begin{aligned} A^T = A \iff [A]_\beta &= D \\ &= P^{-1}AP \\ &\text{for some } P = M_{\beta 0} \in O(n) \end{aligned}}$$

namely $M_{\beta 0} = (\vec{b}_1 \cdots \vec{b}_n)$ orthonormal basis of eigenvectors of A

combining, $A \in Sym_n(\mathbb{R}) \iff$ A is orthog. diag. & $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ all real, no R_θ's

pf: Soon!

Thm. 8 $\boxed{A \in \underset{\text{Symmetric}}{Sym_n(\mathbb{R})} \text{ or } \underset{\text{Hermitian}}{Sym_n(\mathbb{C})} \iff \sigma(A) \subseteq \mathbb{R}, \text{ i.e. } \lambda_i \in \mathbb{R} \text{ in Thm. 6.}}$

all eigenvalues of A are real

pf: Soon!

Thm. 9 (Spectral Thm. For Orthogonal Matrices)

Let $A \in O(n)$. Then, since A is normal, in the diagonalization

$$\boxed{[A]_\beta = \begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_k & & & & \\ & & & R_{\theta_1} & & & \\ & & & & \ddots & & \\ 0 & & & & & R_{\theta_\ell} & \end{pmatrix}}$$

of Thm. 6,

$$\boxed{\lambda_i = \pm 1}$$

And conversely, if $A^TA = AA^T$ & A diagonalizes as above with $\lambda_i = \pm 1$, then $A \in O(n)$. ⊠

For $A^TA = AA^T$
$\lambda_i = \pm 1$
$\iff$
$A \in O(n)$

Remark $\lambda_i = +1 \implies A$ acts as the identity in the E_{λ_i} directions

$\lambda_i = -1 \implies A$ acts as a reflection in the E_{λ_i} direction,

R_{θ_i} in the other 2-dimensional planes of rotation. ⊠

Thm. 10 (Spectral Thm. for Unitary Matrices)

$A \in U(n) \implies$ in the diagonalization fr. Thm. 5 (since $A^*A = I = AA^*$)

$$[A]_\beta = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

all $\lambda_i \in \mathbb{C}$ have modulus 1, $|\lambda_i| = 1$, i.e. λ_i lie on the unit circle $U(1)$ in $\mathbb{C}$.

and conversely if $A^*A = AA^*$ & all $\lambda_i \in U(1)$, then $A \in U(n)$.

$A \in U(n)$ $\iff A^*A = AA^*$ & A is unitarily diagonalizable and all $\lambda_i \in U(1)$

$\boxtimes$

Thm. 11 $A \in U(n)$ or $O(n) \iff \|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in F^n$ — A is an isometric isomorphism

$\iff A$ takes orthonormal bases to orthonormal bases.

pf: Exercise.