(VII.A) Review of Orthogonality

At the beginning of our study of linear transformations in we briefly discussed projections, rotations and projections. In §III.A, projections were treated in the abstract and without regard as to whether they were “orthogonal”; while in §III.B, the examples of rotations and projections in \( \mathbb{R}^2 \) and \( \mathbb{R}^3 \) made use of an orthogonal basis. The idea was that if you want (say) to rotate about \( \hat{e}_1 \), or project “perpendicularly” to \( \text{span}\{\hat{e}_2, \hat{e}_3\} \) in \( \mathbb{R}^3 \), you can write immediately

\[
[R]_{\hat{e}} = \begin{pmatrix}
1 & \cos \theta & -\sin \theta \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}
\quad \text{and} \quad
P\vec{x} = (\vec{x} \cdot \hat{e}_2)\hat{e}_2 + (\vec{x} \cdot \hat{e}_3)\hat{e}_3.
\]

The same formulas hold for any basis \( \mathcal{B} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\} \) of \( \mathbb{R}^3 \) “like \( \hat{e} \)” in the sense that \( \vec{v}_i \cdot \vec{v}_j = \delta_{ij} \) (i.e. the vectors are of unit length and satisfy \( \vec{v}_1 \perp \vec{v}_2, \vec{v}_2 \perp \vec{v}_3, \vec{v}_1 \perp \vec{v}_3 \)). But for an arbitrary basis (not like \( \hat{e} \)), these formulas produce elliptical rotations and skew projections.

So what to do when you need to rotate about \( \begin{pmatrix} 2 \\ -2 \end{pmatrix} \) in \( \mathbb{R}^3 \), or project “perpendicularly” to the span of \( \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \) in \( \mathbb{R}^4 \)? You need to construct the right basis, one “like \( \hat{e} \)”. We shall now standardize these ideas rather than continuing in the ad hoc vein of §§III.A-B. In this section, we will stick with the dot product on \( \mathbb{R}^n \) (and its generalization to \( \mathbb{C}^n \)), while subsequent ones will consider more general bilinear forms.
Orthonormal bases and Projections.

**Definition 1.** Given \( \vec{v}, \vec{w} \in \mathbb{R}^n \), we write
- \( ||\vec{v}|| = \sqrt{(\vec{v} \cdot \vec{v})} \) = norm (or length) of \( \vec{v} \); and
- \( \vec{v} \perp \vec{w} \iff \vec{v} \cdot \vec{w} = 0 \iff \vec{v} \) and \( \vec{w} \) are orthogonal.

**Definition 2.** A basis \( \mathcal{B} = \vec{v}_1, \ldots, \vec{v}_n \) for \( \mathbb{R}^n \) is called orthogonal if \( \vec{v}_i \cdot \vec{v}_j = 0 \) for \( i \neq j \). If in addition \( \vec{v}_i \cdot \vec{v}_i = 1 \) (\( i = 1, \ldots, n \)) then the basis is called orthonormal.

For such a basis the “rotation” about span\( \{\vec{v}_3, \ldots, \vec{v}_n\} \), for example, is then given by

\[
[R]_B = S_B [R]_B S_B^{-1}
\]

where \( [R]_B = \begin{pmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 1 \\
0 & \cdots & 1
\end{pmatrix} \).

More importantly, we claim that the orthogonal projection to \( W = \text{span}\{\vec{v}_r, \ldots, \vec{v}_n\} \) is given by

\[
P_r \vec{x} = (\vec{v}_r \cdot \vec{x}) \vec{v}_r + \ldots + (\vec{v}_n \cdot \vec{x}) \vec{v}_n = \sum_{i=r}^{n} (\vec{v}_i \cdot \vec{x}) \vec{v}_i.
\]

Clearly \( P_r \vec{x} \in W \), but we must also check that \( (\vec{x} - P \vec{x}) \perp W \) :
It suffices, of course, to show 
\((\vec{x} - P\vec{x}) \perp \vec{v}_j\) for each \(j = r, \ldots, n\):

\[
(\vec{x} - P\vec{x}) \cdot \vec{v}_j = \left( \vec{x} - \sum_{i=r}^{n} (\vec{v}_i \cdot \vec{x}) \vec{v}_i \right) \cdot \vec{v}_j \\
= \vec{x} \cdot \vec{v}_j - \sum_{i=r}^{n} (\vec{v}_i \cdot \vec{x}) (\vec{v}_i \cdot \vec{v}_j) = \vec{x} \cdot \vec{v}_j - \vec{v}_j \cdot \vec{x} = 0.
\]

Setting \(r = 1\) so that \(W = \mathbb{R}^n\) (and \(P_1 \vec{x} = \vec{x}\)) we have

\[
(1) \quad \vec{x} = (\vec{v}_1 \cdot \vec{x}) \vec{v}_1 + \cdots + (\vec{v}_n \cdot \vec{x}) \vec{v}_n
\]
whenever \(\{\vec{v}_1, \ldots, \vec{v}_n\}\) is an orthonormal basis. (This is a sort of finite-dimensional Fourier expansion, and the \(\{\vec{v}_i \cdot \vec{x}\}\) are sometimes called Fourier coefficients.) An immediate consequence of (1) is the “Pythagorean theorem”

\[
||\vec{x}||^2 = \vec{x} \cdot \vec{x} = \sum_{i,j=1}^{n} (\vec{v}_i \cdot \vec{x})(\vec{v}_j \cdot \vec{x}) = \sum_{i=1}^{n} (\vec{v}_i \cdot \vec{x})^2.
\]

which implies \(||\vec{x}||^2 = (\vec{v}_1 \cdot \vec{x})^2 + \cdots + (\vec{v}_n \cdot \vec{x})^2 \geq (\vec{v}_r \cdot \vec{x})^2 + \cdots + (\vec{v}_n \cdot \vec{x})^2 = ||P_r \vec{x}||^2;\)

that is, orthogonal projection cannot increase the norm of \(\vec{x} \in \mathbb{R}^n\).

In particular, given any \(\vec{y} \in \mathbb{R}^n\) with span \(L\),

\[
||\vec{x}||^2 \geq ||P_L \vec{x}||^2 = \left\| \left( \vec{x} \cdot \frac{\vec{y}}{||\vec{y}||} \right) \frac{\vec{y}}{||\vec{y}||} \right\|^2 = \left( \vec{x} \cdot \frac{\vec{y}}{||\vec{y}||} \right)^2 \frac{\vec{y} \cdot \vec{y}}{||\vec{y}||^2} = \frac{(\vec{x} \cdot \vec{y})^2}{||\vec{y}||^2};
\]

that is, the “Cauchy-Schwarz” inequality

\[||\vec{x}|| \cdot ||\vec{y}|| \geq |\vec{x} \cdot \vec{y}|\]

holds for the dot product. Thus we may extend the notion of angle between \(\vec{x}\) and \(\vec{y}\) from \(\mathbb{R}^2\) to \(\mathbb{R}^n\): taking

\[
\theta_{\vec{x},\vec{y}} = \arccos \left( \frac{\vec{x} \cdot \vec{y}}{||\vec{x}|| \cdot ||\vec{y}||} \right)
\]
makes sense, since the argument is between \(\pm 1\).
Gram-Schmidt Orthogonalization. Suppose we are given an arbitrary basis \( \{ \vec{w}_1, \ldots, \vec{w}_k \} \) for a subspace \( W \subseteq \mathbb{R}^n \). Here is how to turn it into an orthonormal one: Begin by normalizing \( \vec{w}_1 \): set

\[
\hat{\vec{v}}_1 := \frac{\vec{w}_1}{||\vec{w}_1||}
\]

(the hat indicates a unit vector).

Referring to the above picture, we would like to make \( \vec{w}_2 \perp \hat{\vec{v}}_1 \) by getting rid of its horizontal component \( \vec{w}_2 \cdot \hat{\vec{v}}_1 \) :

\[
\vec{w}_2' := \vec{w}_2 - (\vec{w}_2 \cdot \hat{\vec{v}}_1) \hat{\vec{v}}_1
\]

Referring to the above picture, we would like to make \( \vec{w}_2 \perp \hat{\vec{v}}_1 \) by getting rid of its horizontal component \( \vec{w}_2 \cdot \hat{\vec{v}}_1 \) :

\[
\hat{\vec{v}}_2 := \frac{\vec{w}_2'}{||\vec{w}_2'||}
\]

To make \( \vec{w}_3 \perp \hat{\vec{v}}_1 \) and \( \hat{\vec{v}}_2 \), we take

\[
\vec{w}_3' := \vec{w}_3 - (\vec{w}_3 \cdot \hat{\vec{v}}_1) \hat{\vec{v}}_1 - (\vec{w}_3 \cdot \hat{\vec{v}}_2) \hat{\vec{v}}_2
\]

and so on.

Specialize to the case \( k = n \) (\( W = \mathbb{R}^n \)) and rewrite the equations relating the \( \hat{\vec{v}} \)’s and \( \vec{w} \)’s as follows:

\[
\vec{w}_1 = ||\vec{w}_1|| \hat{\vec{v}}_1
\]

\[
\vec{w}_2 = (\vec{w}_2 \cdot \hat{\vec{v}}_1) \hat{\vec{v}}_1 + \vec{w}_2' = (\vec{w} \cdot \hat{\vec{v}}_1) \hat{\vec{v}}_1 + ||\vec{w}_2'|| \hat{\vec{v}}_2
\]

\[
\vec{w}_3 = (\vec{w}_3 \cdot \hat{\vec{v}}_1) \hat{\vec{v}}_1 + (\vec{w}_3 \cdot \hat{\vec{v}}_2) \hat{\vec{v}}_2 + ||\vec{w}_3'|| \hat{\vec{v}}_3 \text{, etc.}
\]

This looks very nice in matrix terms: \( \mathcal{W} = \mathcal{V} \cdot M \) or

\[
\begin{pmatrix}
\uparrow & \uparrow \\
\vec{w}_1 & \cdots & \vec{w}_n \\
\downarrow & \downarrow \\
\end{pmatrix}
= \begin{pmatrix}
\uparrow & \uparrow \\
\vec{v}_1 & \cdots & \vec{v}_n \\
\downarrow & \downarrow \\
\end{pmatrix}
\begin{pmatrix}
||\vec{w}_1|| & \vec{w}_2 \cdot \vec{v}_1 & \vec{w}_3 \cdot \vec{v}_1 & \cdots & \vec{w}_r \cdot \vec{v}_1 \\
||\vec{w}_2'|| & \vec{w}_3 \cdot \vec{v}_2 & \vec{w}_4 \cdot \vec{v}_2 & \cdots & \vec{w}_r \cdot \vec{v}_2 \\
||\vec{w}_3'|| & \vec{w}_4 \cdot \vec{v}_3 & \vec{w}_5 \cdot \vec{v}_3 & \cdots & \vec{w}_r \cdot \vec{v}_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
||\vec{w}_n'|| \\
\end{pmatrix}.
\]
Here we have taken any invertible matrix (= the left-hand side) and written it as a product of a matrix with orthonormal columns and an upper-triangular matrix. This is called the **QR-decomposition** (where “$Q$” is $V$ and “$R$” is $M$), or (in a broader context) the *Iwasawa decomposition* for $GL_n(\mathbb{R})$ [= invertible matrices].

We will establish some properties of $V$ in a bit; in the meantime, provided you believe that $\det V = \pm 1$ (another way in which an orthonormal basis is “like $\hat{e}$”), this decomposition of $W$ given a very nice second proof that

$$|\det W| = \text{vol}\{\text{parallelepiped with edges } \bar{w}_1, \ldots, \bar{w}_n\}.$$  

What needs to be shown is that the product $(\det M = )$

$$||\bar{w}_1|| \cdot ||\bar{w}'_2|| \cdot ||\bar{w}'_3|| \cdots ||\bar{w}'_n||$$  

gives the parallelepiped’s volume. Writing $W_r = \text{span}\{\bar{w}_1, \ldots, \bar{w}_r\} = \text{span}\{\hat{v}_1, \ldots, \hat{v}_r\}$ for the orthogonal projection to $W_r$, this product becomes

$$||\bar{w}_1|| \cdot ||\bar{w}_2 - P_1\bar{w}_2|| \cdot ||\bar{w}_3 - P_2\bar{w}_3|| \cdots ||\bar{w}_n - P_{n-1}\bar{w}_n||.$$  

But this is just the generalization of

$$\text{Volume} = \text{Base} \cdot \text{Height} = (||\bar{w}_1|| \cdot ||\bar{w}_2 - P_1\bar{w}_2||) \cdot ||\bar{w}_3 - P_2\bar{w}_3||$$

as shown in this picture

![Diagram](image-url)
to higher dimensions.\footnote{The formal proof is by induction, where the volume formula for \( n - 1 \) dimensions (the inductive hypothesis) takes care of the base, and the added dimension is characterized as height\((= ||\vec{w}_n - P_{n-1}\vec{w}_n||)||.\)}

**Orthogonal matrices.** There is a nice algebraic condition on a matrix equivalent to the statement that its columns form an orthonormal basis (of \( \mathbb{R}^n \)). One notices that if \( \hat{v}_1, \ldots, \hat{v}_n \) are orthonormal then

\[
\begin{pmatrix}
\vdots \\
\hat{v}_1 & \cdots & \hat{v}_n \\
\vdots 
\end{pmatrix}
\begin{pmatrix}
\uparrow & \uparrow \\
\hat{v}_1 & \cdots & \hat{v}_n \\
\downarrow & \downarrow 
\end{pmatrix}
= \begin{pmatrix}
1 & 0 \\
\ddots & \\
0 & 1
\end{pmatrix};
\]

that is, \( \mathcal{V} \cdot \mathcal{V} = \mathbb{I}_n \). Clearly the converse also holds.

**Definition 3.** \( Q \in M_n(\mathbb{R}) \) is called **orthogonal** if

\[ tQQ = \mathbb{I} = Q^t Q. \]

For such a matrix,

\[ 1 = \det \mathbb{I} = \det(tQ) \cdot \det Q = (\det Q)^2 \implies \det Q = \pm 1. \]

**Remark 4.** This is, of course, the kind of matrix appearing in the QR-decomposition above. Notice that if one applies Gram-Schmidt to the columns of a matrix \( A \) to find (the columns of) \( Q \), we can obtain \( R \) immediately from \( A = QR \implies R = tQQR = tQA. \)

**Orthogonal transformations.** One can show that the corresponding linear transformations are compositions of rotations and reflections. But here is the standard formal

**Definition 5.** A linear transformation \( T : \mathbb{R}^n \to \mathbb{R}^n \) is **orthogonal** if it preserves length,

\[ ||T\vec{x}|| = ||\vec{x}|| \quad \text{for all } \vec{x} \in \mathbb{R}^n. \]

Now for any \( T \), let \( A = [T]_\ell \), so that the columns of \( A \) are the \( T\hat{e}_i \). If these are an orthonormal basis for \( \mathbb{R}^n \), then

\[ ||T\vec{x}||^2 = ||T(x_1\hat{e}_1 + \ldots + x_n\hat{e}_n)||^2 = ||x_1T\hat{e}_1 + \ldots + x_nT\hat{e}_n||^2 \]
\[ x_1^2 ||T\hat{e}_1||^2 + \ldots + x_n^2 ||T\hat{e}_n||^2 = x_1^2 + \ldots + x_n^2 = ||\vec{x}||^2. \]

So if \([T]\hat{e}\) is orthogonal then \(T\) is. Conversely if \(T\) is orthogonal then \(||T\hat{e}_i|| = ||\hat{e}_i|| = 1\), while

\[ 2 = ||\hat{e}_i||^2 + ||\hat{e}_j||^2 = (\hat{e}_i + \hat{e}_j) \cdot (\hat{e}_i + \hat{e}_j) = ||\hat{e}_i + \hat{e}_j||^2 = \]

\[ = ||T(\hat{e}_i + \hat{e}_j)||^2 = ||T\hat{e}_i + T\hat{e}_j||^2 = (T\hat{e}_i + T\hat{e}_j) \cdot (T\hat{e}_i + T\hat{e}_j) \]

\[ = ||T\hat{e}_i||^2 + 2T\hat{e}_i \cdot T\hat{e}_j + ||T\hat{e}_j||^2 = 2 + 2(T\hat{e}_i \cdot T\hat{e}_j) \]

\[ \implies T\hat{e}_i \cdot T\hat{e}_j = 0. \]

**Conclusion.** \(T\) an orthogonal transformation \(\iff\) \(A = [T]\hat{e}\) an orthogonal matrix.

We now address the computational problems (rotation and projection) pointed out at the beginning of the section.

**Example 6.** How to find the matrix (with respect to \(\hat{e}\)) of rotation by 30° about \(\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}\) in \(\mathbb{R}^3\). First of all,

\[ \ker \begin{pmatrix} 2 & 1 & -2 \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\}. \]

Perform Gram-Schmidt on these last two vectors:

\[ \vec{w}_1 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \quad \implies \quad \hat{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \]

\[ \vec{w}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \quad \implies \quad \vec{w}_2' = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} - \left[ \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right] \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \]

\[ \implies \hat{v}_2 = \frac{1}{3\sqrt{5}} \begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix}. \]
Now simply normalize the rotation axis to get \( \hat{\vartheta}_3 = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \), and put \( B = \{ \hat{\vartheta}_1, \hat{\vartheta}_2, \hat{\vartheta}_3 \} \) so that \( [R]_{\hat{\vartheta}} = S_B[R]_BS_B^{-1} \), where

\[
S_B = \begin{pmatrix} 0 & 5 & 2\sqrt{5} \\ 6 & -2 & \sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix}
\]

and \( [R]_B = \begin{pmatrix} \frac{\sqrt{5}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \).

**Example 7.** How to find the matrix (with respect to \( \hat{\vartheta} \)) of the projection to

\[
W = \text{span} \left\{ \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \right\} \subseteq \mathbb{R}^4.
\]

Apply Gram-Schmidt to the spanning vectors to get an orthonormal basis for \( W \):

\[
\bar{w}_1 = \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix} \quad \rightarrow \quad \hat{\vartheta}_1 = \frac{1}{10} \begin{pmatrix} 1 \\ 7 \\ 7 \end{pmatrix}
\]

\[
\bar{w}_2 = \begin{pmatrix} 0 \\ 7 \\ 2 \\ 7 \end{pmatrix} \quad \rightarrow \quad \hat{\vartheta}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}
\]

\[
\bar{w}_3 = \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \quad \rightarrow
\]
\[ \vec{w}_3' = \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} - \left[ \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \cdot \frac{1}{10} \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix} \right] \frac{1}{10} \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix} - \left[ \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right] \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ 1 \\ 10 \end{pmatrix} \begin{pmatrix} 1 \\ 10 \\ 1 \\ 10 \end{pmatrix} - \begin{pmatrix} 1 \\ 10 \\ 1 \\ 10 \end{pmatrix} \begin{pmatrix} 1 \\ 10 \\ 1 \\ 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right) \]

\[ \rightarrow \hat{v}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \]

From the projection formula

\[ P_W\vec{x} = (\vec{x} \cdot \hat{v}_1)\hat{v}_1 + (\vec{x} \cdot \hat{v}_2)\hat{v}_2 + (\vec{x} \cdot \hat{v}_3)\hat{v}_3 \]

we can now evaluate \( P_W \) on \( \hat{v}_1, \hat{v}_2, \hat{v}_3, \hat{v}_4 \) to get the columns of the matrix \([P_W]_{\hat{v}}\). But there is a shortcut: writing \( \mathcal{V} \) for the \( 4 \times 3 \) matrix with columns \( \hat{v}_1, \hat{v}_2, \hat{v}_3, \hat{v}_4 \),

\[ [P_W\vec{x}]_{\hat{v}} = \mathcal{V} \begin{pmatrix} \hat{v}_1 \cdot \vec{x} \\ \hat{v}_2 \cdot \vec{x} \\ \hat{v}_3 \cdot \vec{x} \end{pmatrix} = \mathcal{V}^t \mathcal{V} \vec{x} \implies \]

\[ (2) \quad [P_W]_{\hat{v}} = \mathcal{V}^t \mathcal{V} \]

\[ = \begin{pmatrix} \frac{1}{10} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{10} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{10} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{10} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{10} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{10} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{10} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{10} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{100} \begin{pmatrix} 51 & 7 & -49 & 7 \\ 7 & 99 & 7 & -1 \\ -49 & 7 & 51 & 7 \\ 7 & -1 & 7 & 99 \end{pmatrix}. \]

Notice that from (2) (which is valid quite generally) it is transparent that orthogonal projections are always given (in \( \hat{v} \), or at least some o.n. basis) by symmetric matrices. As we shall see in §VII.D, this is closely related to the fact that they are diagonalizable.

**Unitary transformations.** All of the above generalizes to \( \mathbb{C}^n \). Recall that in \( \mathbb{R}^n \) we had the following equivalent ways of writing the dot product:

\[ \vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i = \vec{x} \vec{y}^t, \]
where the last is matrix multiplication. If \( \vec{x}, \vec{y} \in \mathbb{C}^n \) we have the following complex “dot product”

\[
\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i = i\bar{\vec{x}} \vec{y} = \bar{\vec{x}}^* \vec{y}
\]

where for any matrix (or vector) “\( \ast \)” indicates the conjugate transpose. The resulting norm \(||\vec{x}|||^2 = \sum x_i^2 = \sum |x_i|^2 \) coincides with the “absolute value” of a complex number in case \( n = 1 \):

\[
||a + bi||^2 = (a - bi)(a + bi) = a^2 + b^2.
\]

Note also that \( \vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x} \) and \( ||\alpha \vec{x}|| = |\alpha| \cdot ||\vec{x}|| \).

**Definition 8.** A transformation

\[ T : \mathbb{C}^n \rightarrow \mathbb{C}^n \]

is called **unitary** if \( ||T\vec{v}|| = ||\vec{v}|| \) for all \( \vec{v} \in \mathbb{C}^n \); this is the complex version of “orthogonal”.

I claim (for such a \( T \)) that the columns \( T\hat{e}_i \) of \( [T]_{\hat{e}} \) satisfy \( T\hat{e}_i \cdot T\hat{e}_j = \delta_{ij} \). First of all since \( T \) is unitary

\[ T\hat{e}_i \cdot T\hat{e}_i = ||T\hat{e}_i||^2 = ||\hat{e}_i||^2 = 1, \tag{3} \]

while (taking \( i \neq j \)) for any \( \alpha \in \mathbb{C} \)

\[ ||\hat{e}_i + \alpha \hat{e}_j||^2 = ||T(\hat{e}_i + \alpha \hat{e}_j)||^2. \tag{4} \]

Let’s consider the left- and right-hand sides of (4):

**l.h.s.** \[ = t(\bar{\hat{e}}_i + \alpha \bar{\hat{e}}_j)(\hat{e}_i + \alpha \hat{e}_j) = \\
| |\hat{e}_i||^2 + |\alpha|^2 |\hat{e}_j||^2 + \alpha (\hat{e}_j \cdot \hat{e}_i) + \bar{\alpha}(\hat{e}_i \cdot \hat{e}_j) \]

0 if \( i \neq j \)

**r.h.s.** \[ = t(\overline{T\hat{e}_i + \alpha T\hat{e}_j})(T\hat{e}_i + \alpha T\hat{e}_j) = \\
||T\hat{e}_i||^2 + |\alpha|^2 ||T\hat{e}_j||^2 + \bar{\alpha}(T\hat{e}_j \cdot T\hat{e}_i) + \alpha(T\hat{e}_i \cdot T\hat{e}_j). \]
Now using (3) to cancel $||\hat{e}_i||^2 + |\alpha|^2||\hat{e}_j||^2$ with $||T\hat{e}_i||^2 + |\alpha|^2||T\hat{e}_j||^2$, we are left with

$$\bar{\alpha}(T\hat{e}_j \cdot T\hat{e}_i) + \alpha(T\hat{e}_i \cdot T\hat{e}_j) = 0$$

for any $\alpha \in \mathbb{C}$. Plug in $\alpha = 1, i$ to get the two equations

$$T\hat{e}_j \cdot T\hat{e}_i = -T\hat{e}_i \cdot T\hat{e}_j, \quad T\hat{e}_j \cdot T\hat{e}_i = T\hat{e}_i \cdot T\hat{e}_j$$

which of course imply $T\hat{e}_i \cdot T\hat{e}_j = 0 (i \neq j)$.

What we have shown is that the matrix of $T$ satisfies $t^\dagger([T]_\mathbb{C}) [T]_\mathbb{C} = \mathbb{I}_n$, which motivates the following generalization of orthogonal matrices to $\mathbb{C}$:

**Definition 9.** A matrix $U \in M_n(\mathbb{C})$ is **unitary** if $U^*U = \mathbb{I}$.

Notice that

$$1 = \det U^* \det U = \overline{\det U} \det U = |\det U|^2 \implies |\det U| = 1$$

which says $\det U$ lies on the unit circle in the complex plane. A **unitary basis** is one like the columns of $U = [T]_\mathbb{C}$: it satisfies

$$(t^\dagger \hat{v}_i \hat{v}_j =) \bar{\hat{v}}^*_i \hat{v}_j = \delta_{ij}.$$
(3) Apply Gram-Schmidt to the set
\[
\left( \begin{array}{c}
1 \\
1 \\
1 \\
1
\end{array} \right), \quad \left( \begin{array}{c}
0 \\
2 \\
0 \\
2
\end{array} \right), \quad \left( \begin{array}{c}
-1 \\
1 \\
3 \\
-1
\end{array} \right)
\]
in \( \mathbb{R}^4 \). Writing \( W \) for their span, find the matrix of the orthogonal projection to \( W \) in the standard basis \( \hat{e} \).

(4) What value of \( b \) (if any) will make the matrix
\[
A = \left( \begin{array}{cc}
\frac{1+i}{2} & b \\
\frac{1-i}{2} & \frac{1-i}{2}
\end{array} \right)
\]
unitary?