VII. Inner product spaces

VII.A. Review of orthogonality

At the beginning of our study of linear transformations we briefly discussed rotations, reflections and projections. In §III.A, projections were treated in the abstract and without regard 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 *span*$\{\hat{e}_2, \hat{e}_3\}$ in $\mathbb{R}^3$, you can write immediately

$$[R]_{\hat{e}} = \begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix} \quad \text{and} \quad P\vec{x} = (\underbrace{\vec{x}\cdot\hat{e}_2}_{=x_2})\hat{e}_2 + (\underbrace{\vec{x}\cdot\hat{e}_3}_{=x_3})\hat{e}_3 \,.$$

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 $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ in $\mathbb{R}^3$, or project "perpendicularly" to

$$\mathit{span}\left\{ \begin{pmatrix} 1 \\ 7 \\ 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 7 \\ 2 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 8 \\ 1 \\ 6 \end{pmatrix} \right\}$$

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-III.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. These definitions are likely familiar to you — first, for vectors:

VII.A.1. DEFINITION. 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**.

— and second, for bases:

VII.A.2. DEFINITION. 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 we have $\vec{v}_i \cdot \vec{v}_i = 1$ $(i = 1, \ldots, n)$, then the basis is called **orthonormal**.

Finally, if $W \subset \mathbb{R}^n$ is a subspace, we write $\vec{v} \perp W$ or $\vec{v} \in W^\perp$ if $\vec{v} \cdot \vec{w} = 0$ for every $\vec{w} \in W$ (equivalently, for each $\vec{w}_i$ in a basis of W).

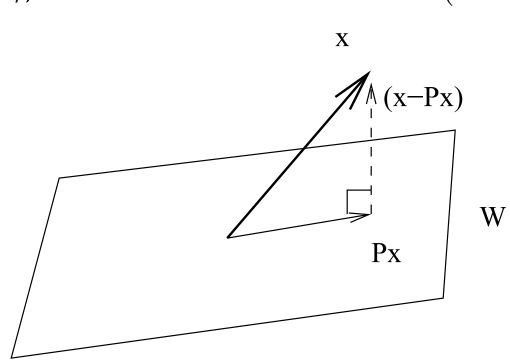
When $\mathcal{B}$ is orthonormal, the *rotation by* θ about *span*$\{\vec{v}_3, \ldots, \vec{v}_n\}$ is given simply by

$$[R]_{\hat{e}} = P_{\mathcal{B}}[R]_{\mathcal{B}}P_{\mathcal{B}}^{-1} \quad \text{where} \quad [R]_{\mathcal{B}} = \begin{pmatrix} \cos\theta & -\sin\theta & & & 0 \\ \sin\theta & \cos\theta & & & \\ & & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}.$$

More importantly, we claim that the *orthogonal projection* to $W_r :=$ *span*$\{\vec{v}_r, \ldots, \vec{v}_n\}$ is given by

$$\mathbb{P}_r\vec{x} = \mathbb{P}_{W_r}(\vec{x}) := (\vec{v}_r \cdot \vec{x})\vec{v}_r + \ldots + (\vec{v}_n \cdot \vec{x})\vec{v}_n = \sum_{i=r}^{n}(\vec{v}_i \cdot \vec{x})\vec{v}_i.$$

Clearly $\mathbb{P}_r\vec{x} \in W_r$, but we must also check that $(\vec{x} - \mathbb{P}_r\vec{x}) \perp W_r$:



It suffices, of course, to show $(\vec{x} - \mathbb{P}_r\vec{x}) \perp \vec{v}_j$ for each $j = r, \ldots, n$:

$$\begin{aligned}(\vec{x} - \mathbb{P}_r\vec{x}) \cdot \vec{v}_j &= (\vec{x} - \textstyle\sum_{i=r}^{n} (\vec{v}_i \cdot \vec{x})\vec{v}_i) \cdot \vec{v}_j \\ &= \vec{x} \cdot \vec{v}_j - \textstyle\sum_{i=r}^{n} (\vec{v}_i \cdot \vec{x}) \underbrace{(\vec{v}_i \cdot \vec{v}_j)}_{=\delta_{ij}} = \vec{x} \cdot \vec{v}_j - \vec{v}_j \cdot \vec{x} = 0.\end{aligned}$$

Notice that for $\vec{x} \in W_r$, this means that $\vec{x} - \mathbb{P}_r\vec{x} \in W_r \cap W_r^{\perp}$. But $W_r \cap W_r^{\perp} = \{0\}$ (cf. Exercise (5)), and so $\mathbb{P}_r\vec{x} = \vec{x}$. Thus the projection restricts to the identity on W_r itself, as one would expect.

Setting $r = 1$ gives $W_1 = \mathbb{R}^n$ and $\mathbb{P}_1\vec{x} = \vec{x}$ for any $\vec{x} \in \mathbb{R}^n$. Hence

$$\vec{x} = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + \cdots + (\vec{v}_n \cdot \vec{x})\vec{v}_n \tag{VII.A.3}$$

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 (VII.A.3) is the "Pythagorean theorem"

$$\|\vec{x}\|^2 = \vec{x} \cdot \vec{x} = \sum_{i,j=1}^{n} (\vec{v}_i \cdot \vec{x})(\vec{v}_j \cdot \vec{x})\underbrace{(\vec{v}_i \cdot \vec{v}_j)}_{\delta_{ij}} = \sum_{i=1}^{n} (\vec{v}_i \cdot \vec{x})^2.$$

which implies (for any r)

$$\|\vec{x}\|^2 = (\vec{v}_1 \cdot \vec{x})^2 + \ldots + (\vec{v}_n \cdot \vec{x})^2 \geq (\vec{v}_r \cdot \vec{x})^2 + \ldots + (\vec{v}_n \cdot \vec{x})^2 = \|\mathbb{P}_r\vec{x}\|^2;$$

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

In particular, for any $\vec{y} \in \mathbb{R}^n$, projection to $L := \mathit{span}(\vec{y})$ satisfies

$$\|\vec{x}\|^2 \geq \|\mathbb{P}_L\vec{x}\|^2 = \left\| \left(x \cdot \frac{\vec{y}}{\|\vec{y}\|} \right) \frac{\vec{y}}{\|\vec{y}\|} \right\|^2 = \left(\vec{x} \cdot \frac{\vec{y}}{\|\vec{y}\|} \right)^2 \frac{\vec{y} \cdot \vec{y}}{\|\vec{y}\|^2} = \frac{(\vec{x} \cdot \vec{y})^2}{\|\vec{y}\|^2};$$

that is, the "Cauchy-Schwarz" inequality

$$\|\vec{x}\| \, \|\vec{y}\| \geq |\vec{x} \cdot \vec{y}|$$

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

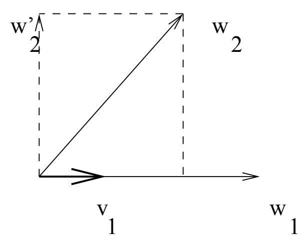
$$\theta_{\vec{x},\vec{y}} = \arccos\left(\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|} \right)$$

makes sense, since the argument of arccos is between 1 and -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

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

(where the hat indicates a unit vector).



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

$$\vec{w}_2' := \vec{w}_2 - (\vec{w}_2 \cdot \hat{v}_1)\hat{v}_1 \qquad \underset{\text{normalize}}{\longrightarrow} \qquad \hat{v}_2 := \frac{\vec{w}_2'}{\|\vec{w}_2'\|}.$$

To make $\vec{w}_3 \;\perp$ to both $\hat{v}_1$ and $\hat{v}_2$, we take

$$\vec{w}_3' := \vec{w}_3 - (\vec{w}_3 \cdot \hat{v}_2)\hat{v}_2 - (\vec{w}_3 \cdot \hat{v}_1)\hat{v}_1 \qquad \longrightarrow \qquad \hat{v}_3 := \frac{\vec{w}_3'}{\|\vec{w}_3'\|},$$

and so on.

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

$$\begin{aligned}
\vec{w}_1 &= \|\vec{w}_1\|\, \hat{v}_1 \\
\vec{w}_2 &= (\vec{w}_2 \cdot \hat{v}_1)\hat{v}_1 + \vec{w}_2' = (\vec{w} \cdot \hat{v}_1)\hat{v}_1 + \|\vec{w}_2'\|\, \hat{v}_2 \\
\vec{w}_3 &= (\vec{w}_3 \cdot \hat{v}_1)\hat{v}_1 + (\vec{w}_3 \cdot \hat{v}_2)\hat{v}_2 + \|\vec{w}_3'\|\, \hat{v}_3\,, \quad \text{etc.}
\end{aligned}$$

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

$$\begin{pmatrix} \uparrow & & \uparrow \\ \vec{w}_1 & \cdots & \vec{w}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \hat{v}_1 & \cdots & \hat{v}_n \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} \|\vec{w}_1\| & \vec{w}_2 \cdot \hat{v}_1 & \vec{w}_3 \cdot \hat{v}_1 & \cdots & \vec{w}_r \cdot \vec{v}_1 \\ & \|\vec{w}_2'\| & \vec{w}_3 \cdot \hat{v}_2 & \cdots & \vec{w}_r \cdot \hat{v}_2 \\ & & \|\vec{w}_3'\| & & \vdots \\ & & & \ddots & \vec{w}_r \cdot \hat{v}_{r-1} \\ \mathbf{0} & & & & \|\vec{w}_n'\| \end{pmatrix}.$$

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 *Iwasawa decomposition for* $GL_n(\mathbb{R})$ [= invertible matrices].)

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

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

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

$$\|\vec{w}_1\|\,\|\vec{w}_2'\|\,\|\vec{w}_3'\|\,\cdots\,\|\vec{w}_n'\|$$

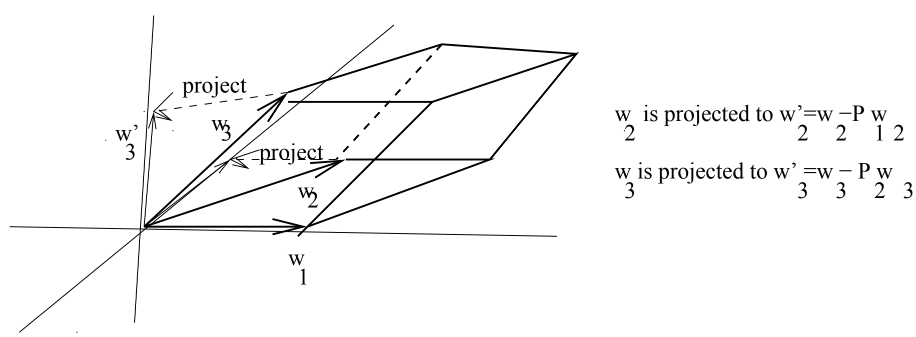
gives the parallelepiped's volume. Writing $V_r = span\{\vec{w}_1, \ldots, \vec{w}_r\} = span\{\hat{v}_1, \ldots, \hat{v}_r\}$ and $\mathbb{P}_r$ for the orthogonal projection to V_r, this product becomes

$$\|\vec{w}_1\|\,\|\vec{w}_2 - \mathbb{P}_1\vec{w}_2\|\,\|\vec{w}_3 - \mathbb{P}_2\vec{w}_3\|\,\cdots\,\|\vec{w}_n - \mathbb{P}_{n-1}\vec{w}_n\|.$$

But this is just the generalization of

$$\text{Volume} = \text{Base} \cdot \text{Height} = (\|\vec{w}_1\|\,\|\vec{w}_2 - \mathbb{P}_1\vec{w}_2\|)\,\|\vec{w}_3 - \mathbb{P}_2\vec{w}_3\|$$

as shown in this picture



to higher dimensions.[1]

[1] 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 ($= \|\vec{w}_n - \mathbb{P}_{n-1}\vec{w}_n\|$).

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

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

that is, ${}^t\mathcal{V}\mathcal{V} = \mathbb{I}_n$. Clearly the converse also holds.

VII.A.4. DEFINITION. $A \in M_n(\mathbb{R})$ is called **orthogonal** if

$${}^tAA = \mathbb{I} = A\,{}^tA.$$

For such a matrix,

$$1 = \det \mathbb{I} = \det({}^tA) \cdot \det A = (\det A)^2 \quad \Longrightarrow \quad \det A = \pm 1.$$

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

VII.A.5. DEFINITION. A linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is **orthogonal** if it preserves length, i.e.

$$\|T\vec{x}\| = \|\vec{x}\| \qquad \text{for all } \vec{x} \in \mathbb{R}^n.$$

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

$$\begin{aligned} \|T\vec{x}\|^2 &= \|T(x_1\hat{e}_1 + \cdots + x_n\hat{e}_n)\|^2 = \|x_1 T\hat{e}_1 + \cdots + x_n T\hat{e}_n\|^2 \\ &= x_1^2\|T\hat{e}_1\|^2 + \cdots + x_n^2\|T\hat{e}_n\|^2 = x_1^2 + \cdots + x_n^2 = \|\vec{x}\|^2. \end{aligned}$$

So if A is orthogonal then T is. Conversely if T is orthogonal then $\|T\hat{e}_i\| = \|\hat{e}_i\| = 1$, and

$$\begin{aligned} 2 &= \|\hat{e}_i\|^2 + \|\hat{e}_j\|^2 = (\hat{e}_i + \hat{e}_j) \cdot (\hat{e}_i + \hat{e}_j) = \|\hat{e}_i + \hat{e}_j\|^2 \\ &= \|T(\hat{e}_i + \hat{e}_j)\|^2 = \|T\hat{e}_i + T\hat{e}_j\|^2 = (T\hat{e}_i + T\hat{e}_j) \cdot (T\hat{e}_i + T\hat{e}_j) \\ &= \|T\hat{e}_i\|^2 + 2T\hat{e}_i \cdot T\hat{e}_j + \|T\hat{e}_j\|^2 = 2 + 2(T\hat{e}_i \cdot T\hat{e}_j) \end{aligned}$$

gives $T\hat{e}_i \cdot T\hat{e}_j = 0$. This yields the

VII.A.6. PROPOSITION. *T is an orthogonal transformation if and only if $A = [T]_{\hat{e}}$ is an orthogonal matrix.*

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

VII.A.7. EXAMPLE. How to find the matrix (with respect to $\hat{e}$) of rotation by 30° about $\left(\begin{smallmatrix} 2 \\ 1 \\ -2 \end{smallmatrix}\right)$ in $\mathbb{R}^3$: First of all,

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

Perform Gram-Schmidt on these last two vectors:

- $\vec{w}_1 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \longrightarrow \hat{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$
- $\vec{w}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \longrightarrow$

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

$$\longrightarrow \hat{v}_2 = \frac{1}{3\sqrt{5}} \begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix}.$$

Now simply normalize the rotation axis to get $\hat{v}_3 = \frac{1}{3}\left(\begin{smallmatrix} 2 \\ 1 \\ -2 \end{smallmatrix}\right)$, and put $\mathcal{B} = \{\hat{v}_1, \hat{v}_2, \hat{v}_3\}$ so that $[R]_{\hat{e}} = P_{\mathcal{B}}[R]_{\mathcal{B}}P_{\mathcal{B}}^{-1}$, where

$$P_{\mathcal{B}} = \frac{1}{3\sqrt{5}} \begin{pmatrix} 0 & 5 & 2\sqrt{5} \\ 6 & -2 & \sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix} \text{ and } [R]_{\mathcal{B}} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $\mathcal{B}$ is orthonormal, $P_{\mathcal{B}}^{-1} = {}^tP_{\mathcal{B}}$, and multiplying everything together yields

$$[R]_{\hat{e}} = \frac{1}{9} \begin{pmatrix} 4 + \frac{5\sqrt{3}}{2} & 5 - \sqrt{3} & -\frac{5}{2} + 2\sqrt{3} \\ -1 - \sqrt{3} & 1 + 4\sqrt{3} & -5 + \sqrt{3} \\ -\frac{11}{2} + 2\sqrt{3} & 1 + \sqrt{3} & 4 + \frac{5\sqrt{3}}{2} \end{pmatrix}.$$

That's an orthogonal matrix, as you can check!

VII.A.8. EXAMPLE. How to find the matrix (with respect to $\hat{e}$) of the projection to

$$W = span\left\{ \begin{pmatrix} 1\\7\\1\\7 \end{pmatrix}, \begin{pmatrix} 0\\7\\2\\7 \end{pmatrix}, \begin{pmatrix} 1\\8\\1\\6 \end{pmatrix} \right\} \subseteq \mathbb{R}^4 :$$

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

- $\vec{w}_1 = \begin{pmatrix} 1\\7\\1\\7 \end{pmatrix} \longrightarrow \hat{v}_1 = \frac{1}{10}\begin{pmatrix} 1\\7\\1\\7 \end{pmatrix}$
- $\vec{w}_2 = \begin{pmatrix} 0\\7\\2\\7 \end{pmatrix} \longrightarrow$

$$\vec{w}_2' = \begin{pmatrix} 0\\7\\2\\7 \end{pmatrix} - \left[\begin{pmatrix} 0\\7\\2\\7 \end{pmatrix} \cdot \frac{1}{10}\begin{pmatrix} 1\\7\\1\\7 \end{pmatrix} \right] \frac{1}{10}\begin{pmatrix} 1\\7\\1\\7 \end{pmatrix} = \begin{pmatrix} -1\\0\\1\\0 \end{pmatrix}$$

$$\longrightarrow \hat{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1\\0\\1\\0 \end{pmatrix}$$

- $\vec{w}_3 = \begin{pmatrix} 1\\8\\1\\6 \end{pmatrix} \longrightarrow$

$$\vec{w}_3' = \begin{pmatrix} 1\\8\\1\\6 \end{pmatrix} - \left[\begin{pmatrix} 1\\8\\1\\6 \end{pmatrix} \cdot \frac{1}{10}\begin{pmatrix} 1\\7\\1\\7 \end{pmatrix} \right] \frac{1}{10}\begin{pmatrix} 1\\7\\1\\7 \end{pmatrix} -$$

$$\left[\begin{pmatrix} 1\\8\\1\\6 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} -1\\0\\1\\0 \end{pmatrix} \right] \frac{1}{\sqrt{2}}\begin{pmatrix} -1\\0\\1\\0 \end{pmatrix} = \begin{pmatrix} 0\\1\\0\\-1 \end{pmatrix}$$

$$\longrightarrow \hat{v}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0\\1\\0\\-1 \end{pmatrix}.$$

Now write the projection formula

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

and evaluate P_W on $\hat{e}_1$, $\hat{e}_2$, $\hat{e}_3$, $\hat{e}_4$ to get the columns of the matrix

$$[P_W]_{\hat{e}} = \frac{1}{100}\begin{pmatrix} 51 & 7 & -49 & 7 \\ 7 & 99 & 7 & -1 \\ -49 & 7 & 51 & 7 \\ 7 & -1 & 7 & 99 \end{pmatrix}.$$

As you may verify directly, it has rank 3.

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:

$$\vec{x}\cdot\vec{y} = \sum_{i=1}^{n} x_i y_i = {}^t\vec{x}\,\vec{y},$$

where the last is matrix multiplication. If $\vec{x}, \vec{y} \in \mathbb{C}^n$ we have the following complex "dot product"

$$\vec{x}\cdot\vec{y} = \sum_{i=1}^{n} \bar{x}_i y_i = {}^t\bar{\vec{x}}\,\vec{y} = \vec{x}^*\vec{y}$$

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

$$\|a + b\mathbf{i}\|^2 = (a - b\mathbf{i})(a + b\mathbf{i}) = a^2 + b^2.$$

Note also that $\vec{x}\cdot\vec{y} = \overline{\vec{y}\cdot\vec{x}}$ and $\|\alpha\vec{x}\| = |\alpha|\|\vec{x}\|$.

VII.A.9. DEFINITION. A transformation

$$T\colon \mathbb{C}^n \to \mathbb{C}^n$$

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

$$T\hat{e}_i \cdot T\hat{e}_i = \|T\hat{e}_i\|^2 = \|\hat{e}_i\|^2 = 1, \tag{VII.A.10}$$

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

$$\|\hat{e}_i + \alpha\hat{e}_j\|^2 = \|T(\hat{e}_i + \alpha\hat{e}_j)\|^2. \tag{VII.A.11}$$

Let's consider the left- and right-hand sides of (VII.A.11):

$$\text{l.h.s.} = {}^t(\overline{\hat{e}_i + \alpha\hat{e}_j})(\hat{e}_i + \alpha\hat{e}_j) =$$
$$\|\hat{e}_i\|^2 + |\alpha|^2\|\hat{e}_j\|^2 + \underbrace{\bar{\alpha}(\hat{e}_j \cdot \hat{e}_i) + \alpha(\hat{e}_i \cdot \hat{e}_j)}_{0 \text{ if } i \neq j}$$

$$\text{r.h.s.} = {}^t(\overline{T\hat{e}_i + \alpha T\hat{e}_j})(T\hat{e}_i + \alpha T\hat{e}_j)$$
$$= \|T\hat{e}_i\|^2 + |\alpha|^2\|T\hat{e}_j\|^2 + \bar{\alpha}(T\hat{e}_j \cdot T\hat{e}_i) + \alpha(T\hat{e}_i \cdot T\hat{e}_j).$$

Using (VII.A.10) 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 $\alpha = 1$, $\mathbf{i}$ 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 ${}^t(\overline{[T]_{\hat{e}}})\,[T]_{\hat{e}} = \mathbb{I}_n$, which motivates the following generalization of orthogonal matrices to $\mathbb{C}$:

VII.A.12. DEFINITION. $U \in M_n(\mathbb{C})$ is **unitary** if $U^*U = \mathbb{I}$.

Notice that

$$1 \;=\; \det U^* \det U \;=\; \overline{\det U}\det U \;=\; |\det U|^2 \quad\Longrightarrow\quad |\det U| = 1$$

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

$$({}^t\bar{\vec{v}}_i\,\vec{v}_j =)\;\vec{v}_i^*\vec{v}_j = \delta_{ij}.$$

Exercises

(1) Show that an orthogonal transformation of $\mathbb{R}^n$ preserves all angles. [Hint: use the fact that $[T]_{\hat{e}} = A$ is an orthogonal matrix.]

(2) Apply Gram-Schmidt to the columns of

$$A = \begin{pmatrix} 2 & 1 & 6 \\ -1 & 1 & 3 \\ 2 & 4 & 9 \end{pmatrix}$$

to write $M = AB$ where A is orthogonal and B is upper triangular.

(3) Apply Gram-Schmidt to the set

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 3 \\ -1 \end{pmatrix}$$

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

$$A = \begin{pmatrix} \frac{1+\mathbf{i}}{2} & b \\ \frac{1-\mathbf{i}}{2} & \frac{1-\mathbf{i}}{2} \end{pmatrix}$$

unitary?

(5) If $W \subset \mathbb{R}^n$ is a subspace, show that $W \cap W^{\perp} = \{0\}$. [Hint: let $\vec{w}_1, \ldots, \vec{w}_k$ be an o.n. basis, and write an element in terms of it.]