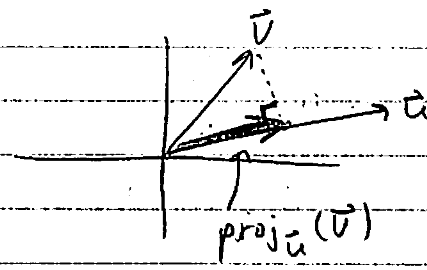
Notes for Math 204-2, Fall 2008

Orthonormal Bases and Projections (Section 4.9)

In Chapter 3, we learned how to project vectors in $\mathbb{R}^2$ onto lines: $\operatorname{proj}_{\vec{u}} \vec{v} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|^2} \cdot \vec{u}$

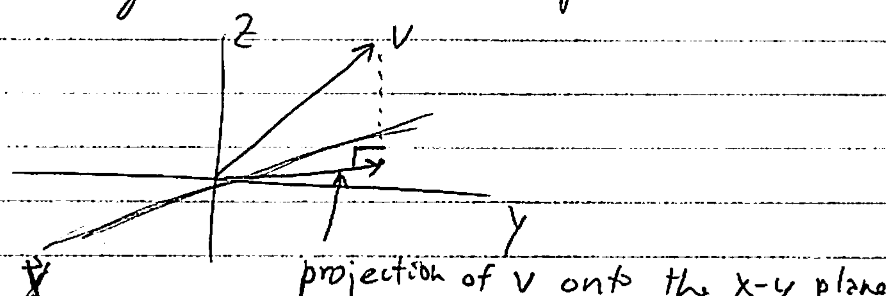


Recall that if $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ then

$$u \cdot v = \sum_{i=1}^{n} u_i v_i$$

and $\|u\| = \sqrt{u \cdot u} = \sqrt{\sum_{i=1}^{n} u_i^2}$.

In this section, we'll study projections of vectors onto higher dimensional subspaces:



To understand this notion, we need to consider orthogonality:

Def'n 1: A set $S = \{u_1, \dots, u_n\} \subseteq \mathbb{R}^n$ is *orthogonal* if

$$u_i \cdot u_j = 0 \quad \text{for } i \neq j$$

S is called *orthonormal* if it also satisfies $\|u_i\| = 1$ for all i, i.e. $u_i \cdot u_i = 1$.

Examples: • The standard basis vectors $e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ form an orthonormal subset of $\mathbb{R}^n$.

• The vectors $\begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ are orthogonal (but not orthonormal)

$\begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = 4 - 4 = 0.$ If we want to add a third vector to form a larger orthonormal set, this vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ must satisfy the 2 linear eq'ns

$$\begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \quad \text{i.e.} \quad 2a + 4b + 2c = 0$$

$$\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \quad \text{i.e.} \quad 2a - b = 0$$

This means $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in N\begin{pmatrix} 2 & 4 & 2 \\ 2 & -1 & 0 \end{pmatrix} = N\begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & -2 \end{pmatrix} = N\begin{pmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 2/5 \end{pmatrix} = \text{span}\begin{pmatrix} -1/5 \\ -2/5 \\ 1 \end{pmatrix}$.

So, for example, $\left\{ \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix} \right\}$ is orthogonal.

Theorem: If $S = \{v_1, \dots, v_k\}$ is an orthogonal set of non-zero vectors, then S is linearly independent.

(So in the above example, $\left\{ \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -5 \end{pmatrix} \right\}$ is a basis)

Pf: Consider the matrix $A = (\vec{v}_1 \cdots \vec{v}_k)$. Then

$$A^T = \begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_k \end{pmatrix} \quad \text{and} \quad A^T A = \begin{pmatrix} \vec{v}_1 \cdot \vec{v}_1 & 0 & & 0 \\ 0 & \vec{v}_2 \cdot \vec{v}_2 & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & & & \vec{v}_k \cdot \vec{v}_k \end{pmatrix}.$$

So $A^T A$ is an invertible square matrix. Hence $N(A) = 0$, and the columns of A are linearly indep. □

There is a process for converting a given basis into an <u>orthonormal</u> basis; this is called Gram-Schmidt orthogonalization. We'll illustrate with an example.

Start with a basis $\{\vec{v}_i\}$ for $\mathbb{R}^n$, e.g. $\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right\} \subseteq \mathbb{R}^3$.

<u>Step 1</u>: Normalize $\vec{v}_1$: replace $\vec{v}_1$ by $\vec{v}_1' = \dfrac{\vec{v}_1}{\|\vec{v}_1\|}$.

This doesn't affect $\operatorname{Span}\{v_i\}$, and now $\|v_1'\|^2 = \dfrac{\vec{v}_1}{\|v_1\|} \cdot \dfrac{\vec{v}_1}{\|v_1\|} = \dfrac{\vec{v}_1 \cdot \vec{v}_1}{\|v_1\|^2} = 1$

So $\|v_1'\| = 1$.

— In example, we replace $\begin{pmatrix}1\\1\\0\end{pmatrix}$ by $\begin{pmatrix}1\\1\\0\end{pmatrix}\Big/\sqrt{2} = \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}$ so get the new basis $\left\{\begin{pmatrix}1\\1\\0\end{pmatrix} \begin{pmatrix}0\\1\\1\end{pmatrix} \begin{pmatrix}1\\0\\1\end{pmatrix}\right\}$.

<u>Step 2</u>: Replace $\vec{v}_2$ by a vector orthogonal to $\vec{v}_1'$:

$$\vec{v}_2' = \vec{v}_2 - \operatorname{proj}_{\vec{v}_1'}(\vec{v}_2) = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{v}_1'}{\underbrace{\|\vec{v}_1'\|}_{=1}} \vec{v}_1' = \vec{v}_2 - (\vec{v}_2 \cdot \vec{v}_1')\, \vec{v}_1'.$$

Now $\vec{v}_2' \cdot \vec{v}_1' = \left(\vec{v}_2 - (\vec{v}_2 \cdot \vec{v}_1')\,\vec{v}_1'\right) \cdot \vec{v}_1' = \vec{v}_2 \cdot \vec{v}_1' - (v_2 \cdot \vec{v}_1')\,(\underbrace{\vec{v}_1' \cdot \vec{v}_1'}_{1})$

$= \vec{0}$,

So $\vec{v}_2'$ is orthogonal to $\vec{v}_1'$.

— In example, we replace $\begin{pmatrix}0\\1\\1\end{pmatrix}$ by $\begin{pmatrix}0\\1\\1\end{pmatrix} - \left(\begin{pmatrix}0\\1\\1\end{pmatrix} \cdot \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}\right) \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}$

$$= \begin{pmatrix}0\\1\\1\end{pmatrix} - 1/\sqrt{2} \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix} = \begin{pmatrix}-1/2\\1/2\\1\end{pmatrix}$$

<u>Step 3</u>: Replace $\vec{v}_2'$ by $v_2'/\|\vec{v}_2'\|$ to make $\vec{v}_2'$ a unit vector.

— In example, our basis is now $\left\{\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}, \begin{pmatrix}-1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right\}$

Orthogonal Projections: (Def'n 4)

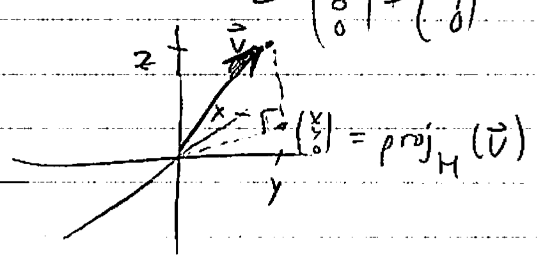
If $H \subseteq \mathbb{R}^n$ has o.n. basis $\{\vec{u}_1, \dots, \vec{u}_k\}$, then the orthogonal proj. of $\vec{v} \in \mathbb{R}^n$ onto H is given by

$$\text{proj}_H(\vec{v}) = (\vec{v}\cdot\vec{u}_1)\vec{u}_1 + \cdots + (v \cdot u_k) u_k$$
$$= \sum_{i=1}^{k} \text{proj}_{\vec{u}_i}(\vec{v}).$$

Since the $\vec{u}_i$ are in H, so is $\text{proj}_H(\vec{v})$.

Example: If $\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ and $H = \text{Span}\left(\underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\text{o.n. basis}}\right)$ is the x-y plane, then

$$\text{proj}_H(\vec{v}) = \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$



This def'n of projection relies on a choice of basis, but in fact the resulting vector is well-defined.

Theorem 5: Let $H \subseteq \mathbb{R}^n$ be a subspace with o.n. bases $\{\vec{u}_1, \dots, \vec{u}_k\}$, $\{\vec{w}_1, \dots, \vec{w}_k\}$. Then for any $\vec{v} \in \mathbb{R}^n$,

$$\sum_{i=1}^{k} (\vec{v}\cdot\vec{u}_i)\vec{u}_i = \sum_{i=1}^{k} (\vec{v}\cdot w_i)\vec{w}_i .$$

In fact, this projection can be characterized as the closest point in H to the vector $\vec{v}$:

<u>Theorem 8:</u> Let $H \subseteq \mathbb{R}^n$ be a subspace, $\vec{v} \in \mathbb{R}^n$ any vector. Then for all $\vec{h} \in H$,

$$|\vec{v} - \text{proj}_H(\vec{v})| \leq |\vec{v} - \vec{h}|$$

with equality only when $h = \text{proj}_H(\vec{v})$.

<u>Example:</u> We saw that the projection of $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ onto the x-y plane is just $\begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$. Say we take a different o.n. basis for $\text{Span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\} = x$-$y$ plane, such as $\left\{\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}\right\}$, then we compute

$$\text{proj}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}\right)\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} + \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}\right)\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}$$

$$= \left(\frac{x+y}{\sqrt{2}}\right)\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} + \left(\frac{x-y}{\sqrt{2}}\right)\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{x+y}{2} + \frac{x-y}{2} \\ \frac{x+y}{2} - \frac{x-y}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \text{ as expected.}$$

The key to the proofs of Thms 5 + 8 is the fact that coordinates wrt an o.n. basis are easy to compute:

<u>Thm 4:</u> If $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ is an o.n. basis for $\mathbb{R}^n$, then for any $\vec{v} \in \mathbb{R}^n$, $[v]_{\mathcal{B}} = \begin{pmatrix} v \cdot \vec{u}_1 \\ \vdots \\ v \cdot \vec{u}_n \end{pmatrix}$.

<u>Pf:</u> Say $\vec{v} = \sum_{i=1}^{n} c_i \vec{u}_i$. We must show $c_j = \vec{v} \cdot \vec{u}_j$. But $\vec{v} \cdot u_j = \left(\sum c_i u_i\right) \cdot u_j = \sum c_i u_i \cdot u_j = c_j$ (b/c $\{u_i\}$ is o.n.). $\square$