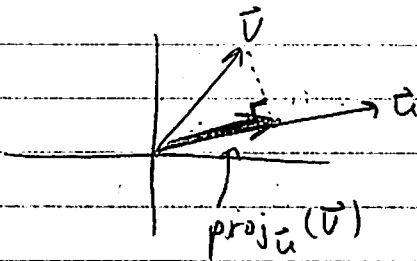


# Notes for Math 204-2, Fall 2008

## Orthonormal Bases and Projections (Section 4.9)

In Chapter 3, we learned how to project vectors

in  $\mathbb{R}^2$  onto lines:  $\text{proj}_{\vec{u}} \vec{v} = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|^2} \cdot \vec{u}$ .

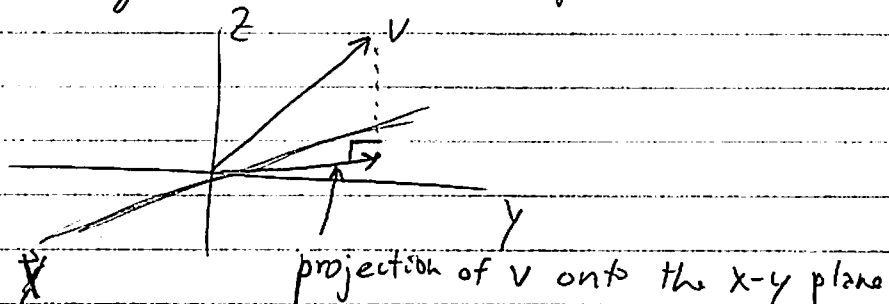


Recall that if  $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ ,  $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  then

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

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

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  $S = \{u_1, \dots, u_n\} \in \mathbb{R}^n$  is orthogonal if  $u_i \cdot u_j = 0$  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  $e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$  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)

$$\begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = 4 - 4 = 0. \quad \text{If we want to add a}$$

third vector to form a larger orthonormal set, this

vector  $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$  must satisfy the 2 linear eq's

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

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

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

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  $A = (\vec{v}_1 \dots \vec{v}_k)$ . Then

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

So  $A^T A$  is an invertible square matrix. Hence

$N(A) = \{0\}$ , and the columns of  $A$  are linearly indep.  $\square$

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

Start with a basis  $\{\vec{v}_i\}$  for  $\mathbb{R}^n$ , e.g.  $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 7 \\ 0 \\ 1 \end{pmatrix} \right\} \in \mathbb{R}^3$ .

Step 1: Normalize  $\vec{v}_1$ : replace  $\vec{v}_1$  by  $\vec{v}'_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$ .

This doesn't affect  $\text{Span}\{\vec{v}_i\}$ , and now  $\|\vec{v}'_1\|^2 = \frac{\vec{v}_1}{\|\vec{v}_1\|} \cdot \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{\vec{v}_1 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} = 1$

- In example, we replace  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  by  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} / \sqrt{1} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}$  so  
get the new basis  $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 7 \\ 0 \\ 1 \end{pmatrix} \right\}$ .

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

$$\vec{v}'_2 = \vec{v}_2 - \text{proj}_{\vec{v}'_1}(\vec{v}_2) = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{v}'_1}{\|\vec{v}'_1\|} \vec{v}'_1 = \vec{v}_2 - (\vec{v}_2 \cdot \vec{v}'_1) \vec{v}'_1$$

$$\text{Now } \vec{v}'_2 \cdot \vec{v}'_1 = (\vec{v}_2 - (\vec{v}_2 \cdot \vec{v}'_1) \vec{v}'_1) \cdot \vec{v}'_1 = \vec{v}_2 \cdot \vec{v}'_1 - (\vec{v}_2 \cdot \vec{v}'_1) \underbrace{(\vec{v}'_1 \cdot \vec{v}'_1)}_1 = 0$$

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

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

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

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

## Orthogonal Projections: (Def'n 4)

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

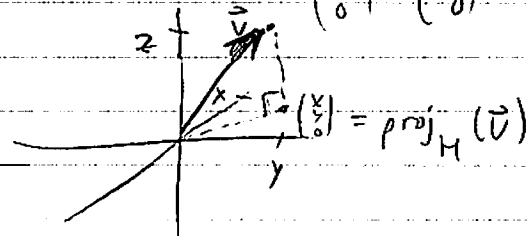
$$\begin{aligned} \text{by } \text{proj}_H(\vec{v}) &= (\vec{v} \cdot \vec{u}_1) \vec{u}_1 + \dots + (\vec{v} \cdot \vec{u}_k) \vec{u}_k \\ &= \sum_{i=1}^k \text{proj}_{\vec{u}_i}(\vec{v}). \end{aligned}$$

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

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



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

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

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

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

Theorem 8: 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  $\vec{h} = \text{proj}_H(\vec{v})$ .

Example: 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 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

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

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

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

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