Math 3191
Applied Linear Algebra

Lecture 21: Inner Products, Length, Orthogonality

Stephen Billups

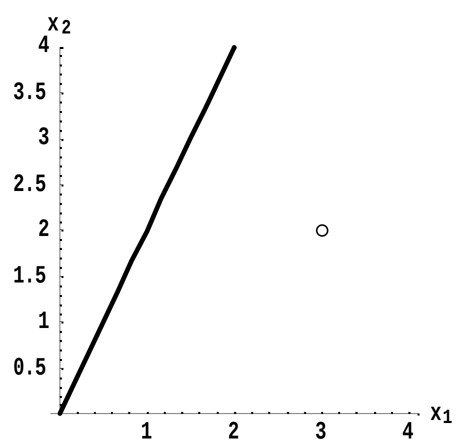
University of Colorado at Denver

Motivation

Not all linear systems have solutions.

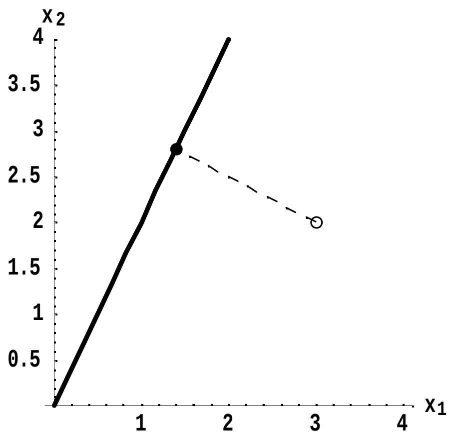
EXAMPLE: No solution to $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ exists. Why?

$A\mathbf{x}$ is a point on the line spanned by $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ and $\mathbf{b}$ is not on the line. So $A\mathbf{x} \neq \mathbf{b}$ for all $\mathbf{x}$.



Approximate Solutions

Instead find $\widehat{\mathbf{x}}$ so that $A\widehat{\mathbf{x}}$ lies ”closest” to $\mathbf{b}$.



Using techniques in this chapter, we will find that $\widehat{\mathbf{x}} = \begin{bmatrix} 1.4 \\ 0 \end{bmatrix}$, so that $A\widehat{\mathbf{x}} = \begin{bmatrix} 1.4 \\ 2.8 \end{bmatrix}$.

Observation: Segment joining $A\widehat{\mathbf{x}}$ and $\mathbf{b}$ is *perpendicular (* or *orthogonal)* to the set of solutions to $A\mathbf{x} = \mathbf{b}$.

Need to develop fundamental ideas of

- length
- orthogonality
- orthogonal projections

The key to all of these concepts is the **Inner Product**.

The Inner Product

Inner product or **dot product** of

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}:$$

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Note that $\mathbf{v} \cdot \mathbf{u} = v_1 u_1 + \cdots + v_n u_n = u_1 v_1 + \cdots + u_n v_n = \mathbf{u} \cdot \mathbf{v}$

THEOREM 1

Let $\mathbf{u}, \mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbf{R}^n$, and let c be any scalar. Then

a. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

b. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

c. $(c\mathbf{u}) \cdot \mathbf{v} = c\,(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$

d. $\mathbf{u} \cdot \mathbf{u} \geq \mathbf{0}$, and $\mathbf{u} \cdot \mathbf{u} = \mathbf{0}$ if and only if $\mathbf{u} = \mathbf{0}$.

Combining parts b and c, one can show

$$\boxed{(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{w} = c_1\,(\mathbf{u}_1 \cdot \mathbf{w}) + \cdots + c_p\,(\mathbf{u}_p \cdot \mathbf{w})}$$

Length of a Vector

For $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, the **length** or **norm of** $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defi ned by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \qquad \text{and} \qquad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}.$$

For example, if $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, then $\|\mathbf{v}\| = \sqrt{a^2 + b^2}$ (distance between $\mathbf{0}$ and $\mathbf{v}$)

Observation: For any scalar c,

$$\boxed{\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|}$$

Distance in $\mathbf{R}^n$

The **distance between $\mathbf{u}$ and $\mathbf{v}$** in $\mathbf{R}^n$:

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

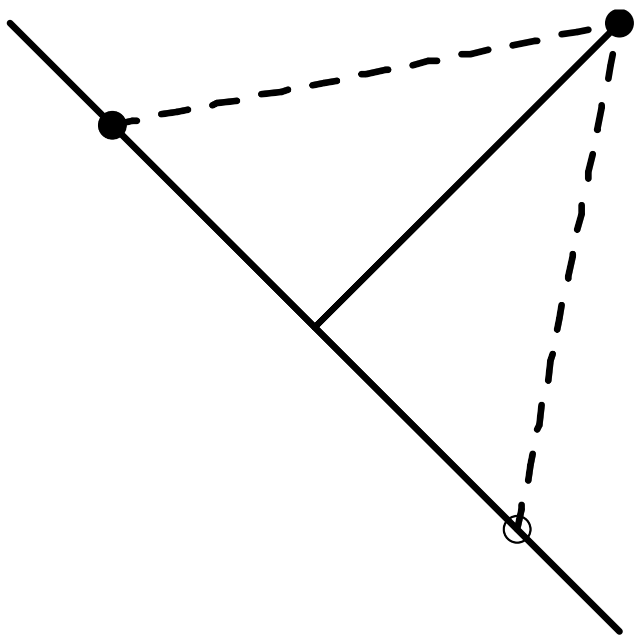
This agrees with the usual formulas for $\mathbf{R}^2$ and $\mathbf{R}^3$. Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$.

Then $\mathbf{u} - \mathbf{v} = (u_1 - v_1, u_2 - v_2)$ and

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|(u_1 - v_1, u_2 - v_2)\|$$

$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$$

Orthogonal Vectors



$$[\text{dist}\,(\mathbf{u},\mathbf{v})]^2 = \|\mathbf{u}-\mathbf{v}\|^2 = (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})$$

$$= (\mathbf{u})\cdot(\mathbf{u}-\mathbf{v}) + (-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) =$$

$$= \mathbf{u}\cdot\mathbf{u} - \mathbf{u}\cdot\mathbf{v} + -\mathbf{v}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v}$$

$$= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}$$

$$\Rightarrow \quad [\text{dist}\,(\mathbf{u},\mathbf{v})]^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}$$

Previous slide showed that $[\text{dist}\,(\mathbf{u},\mathbf{v})]^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}$

Similarly, we can show that $[\text{dist}\,(\mathbf{u},-\mathbf{v})]^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u}\cdot\mathbf{v}$

Since $[\text{dist}\,(\mathbf{u},-\mathbf{v})]^2 = [\text{dist}\,(\mathbf{u},\mathbf{v})]^2$, $\mathbf{u}\cdot\mathbf{v} = \underline{\quad\quad}$.

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthogonal** (to each other) if $\mathbf{u}\cdot\mathbf{v} = \mathbf{0}$.

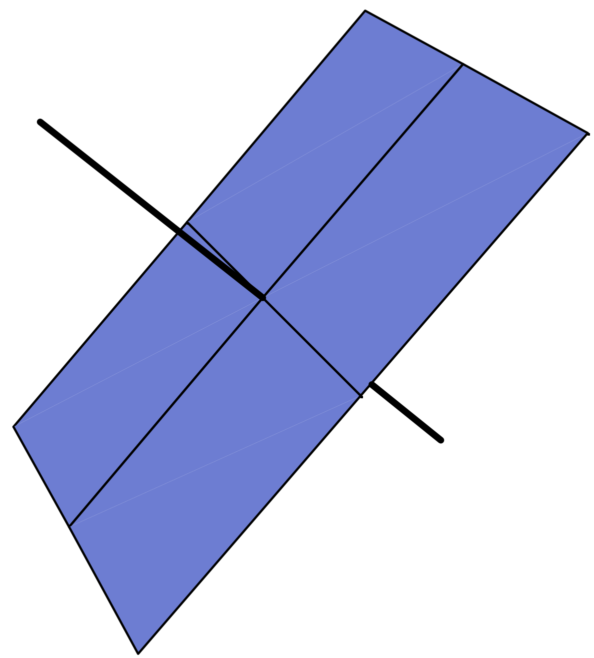
Also note that if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

THEOREM 2 THE PYTHAGOREAN THEOREM

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Orthogonal Complements

If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbf{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (read as “W perp”).



Row, Null and Columns Spaces

THEOREM 3

Let A be an $m \times n$ matrix. Then the orthogonal complement of the row space of A is the nullspace of A, and the orthogonal complement of the column space of A is the nullspace of A^T:

$$(\text{Row } A)^{\perp} = \text{Nul } A, \qquad (\text{Col } A)^{\perp} = \text{Nul } A^T.$$

Why? (See complete proof in the text) Consider $A\mathbf{x} = \mathbf{0}$:

Note that $A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ where $\mathbf{r}_1, \ldots, \mathbf{r}_m$ are the rows of A. Thus, $\mathbf{x}$ is orthogonal to each row of A. So $\mathbf{x}$ is orthogonal to Row A.

EXAMPLE

Let $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & 2 \end{bmatrix}$.

Basis for Nul $A = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ and therefore Nul A is a plane in $\mathbf{R}^3$.

Basis for Row $A = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$ and therefore Row A is a line in $\mathbf{R}^3$.

Basis for Col $A = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ and therefore Col A is a line in $\mathbf{R}^2$.

Basis for Nul $A^T = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ and therefore Nul A^T is a line in $\mathbf{R}^2$.

Section 6.2 Orthogonal Sets

A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ in $\mathbf{R}^n$ is called an **orthogonal set** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE: Is $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ an orthogonal set?

Solution: Label the vectors $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3$ respectively. Then

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = \underline{\hspace{3cm}}, \quad \mathbf{u}_1 \cdot \mathbf{u}_3 = \underline{\hspace{3cm}}, \quad \mathbf{u}_2 \cdot \mathbf{u}_3 = \underline{\hspace{3cm}}$$

Therefore, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set.

THEOREM 4

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbf{R}^n$ and $W =$span$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$. Then S is a linearly independent set and is therefore a basis for W.

Partial Proof: Suppose

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p = \mathbf{0}$$

$$(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p) \cdot \underline{\quad} = \mathbf{0}\cdot$$

$$(c_1\mathbf{u}_1) \cdot \mathbf{u}_1 + (c_2\mathbf{u}_2) \cdot \mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p) \cdot \mathbf{u}_1 = \mathbf{0}$$

$$c_1\left(\mathbf{u}_1 \cdot \mathbf{u}_1\right) + c_2\left(\mathbf{u}_2 \cdot \mathbf{u}_1\right) + \cdots + c_p\left(\mathbf{u}_p \cdot \mathbf{u}_1\right) = \mathbf{0}$$

$$c_1\left(\mathbf{u}_1 \cdot \mathbf{u}_1\right) = \mathbf{0}$$

Since $\mathbf{u}_1 \neq \mathbf{0}$, $\mathbf{u}_1 \cdot \mathbf{u}_1 > 0$ which means $c_1 =$ ____.

In a similar manner, c_2,. . .,c_p can be shown to by all 0. So S is a linearly independent set.■

Orthogonal Basis

An **orthogonal basis** for a subspace W of $\mathbf{R}^n$ is a basis for W that is also an orthogonal set.

Question: Why would we want to have an orthogonal basis?

Ans: It makes it easy to calculate the coordinates relative to the basis.

EXAMPLE: Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for a subspace W of $\mathbf{R}^n$ and suppose $\mathbf{y}$ is in W. Find c_1,. . .,c_p so that $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p$.

Solution:

$$\mathbf{y}\cdot \qquad = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot$$

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1$$

$$\mathbf{y}\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1)$$

$$\mathbf{y}\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

$$c_1 = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}$$

Similarly, $c_2 =$, $c_3 =$,. . . , $c_p =$

THEOREM 5

Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbf{R}^n$. Then each $\mathbf{y}$ in W has a unique representation as a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p$. In fact, if

$$\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p$$

then

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \qquad (j = 1, \ldots, p)\,.$$

EXAMPLE:

Express $\mathbf{y} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ as a linear combination of the orthogonal basis

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Solution:

$$\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \qquad \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \qquad \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3} =$$

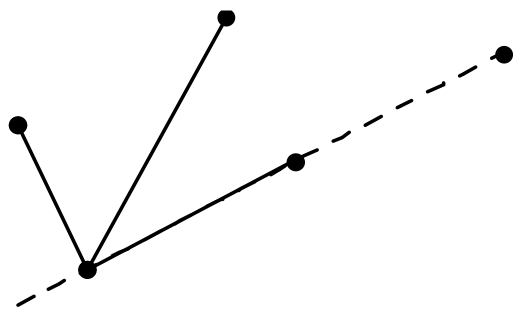
Hence

$$\mathbf{y} = \underline{\qquad\qquad}\mathbf{u}_1 + \underline{\qquad\qquad}\mathbf{u}_2 + \underline{\qquad\qquad}\mathbf{u}_3$$

Orthogonal Projections

For a nonzero vector $\mathbf{u}$ in $\mathbf{R}^n$, suppose we want to write $\mathbf{y}$ in $\mathbf{R}^n$ as the the following

$$\mathbf{y} = (\text{multiple of } \mathbf{u}) + (\text{multiple a vector } \perp \text{ to } \mathbf{u}) .$$



$$(\mathbf{y} - \alpha \mathbf{u}) \cdot \mathbf{u} = 0$$

$$\mathbf{y} \cdot \mathbf{u} - \alpha (\mathbf{u} \cdot \mathbf{u}) = 0 \qquad \Longrightarrow \qquad \alpha =$$

$$\widehat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \qquad \textbf{(orthogonal projection of } \mathbf{y} \textbf{ onto } \mathbf{u}\textbf{)}$$

and

$$\mathbf{z} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \qquad \textbf{(component of } \mathbf{y} \textbf{ orthogonal to } \mathbf{u}\textbf{)}$$

Example

Let $\mathbf{y} = \begin{bmatrix} -8 \\ 4 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Compute the distance from $\mathbf{y}$ to the line through $\mathbf{0}$ and $\mathbf{u}$.

Solution:

$$\widehat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} = \_\_\_\_\_\_$$

Distance from $\mathbf{y}$ to the line through $\mathbf{0}$ and $\mathbf{u}$ $=$ distance from $\widehat{\mathbf{y}}$ to $\mathbf{y}$

$= \|\widehat{\mathbf{y}} - \mathbf{y}\| = \_\_$

Orthonormal Sets

A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ in $\mathbf{R}^n$ is called an **orthonormal set** if

1. It is *orthogonal*.
2. Each vector has length 1.

If the orthonormal set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ spans a vector space W, then $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is called an **orthonormal basis** for W.

Orthogonal Matrices

Recall that $\mathbf{v}$ is a unit vector if $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\mathbf{v}^T\mathbf{v}} = 1$.

Suppose $U = [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \mathbf{u}_3]$ where $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set.

Then

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \mathbf{u}_3] = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} =$$

It can be shown that $UU^T = I$ also. So $U^{-1} = U^T$ (such a matrix is called an **orthogonal matrix**). (NOTE: U must be square to be orthogonal).

THEOREM 6 An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

THEOREM 7 Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbf{R}^n$. Then

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$

b. $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$

c. $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Proof of part b: $(U\mathbf{x}) \cdot (U\mathbf{y}) =$