Lecture 24 | 23 Nov 2010

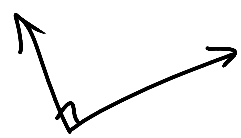
Course evaluations now available at:
https://utdirect.utexas.edu/ctl/ecis/
(Please fill them out!)

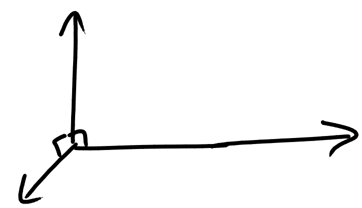
HW 11 due today
HW 12 will be due next Thu
HW 13 will be assigned but not collected

Prof. Ray Heitmann will give the last 2 lectures; I will be away 11/29–12/7

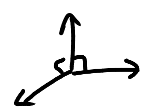
Last time:

Orthogonal sets, orthogonal bases





Orthonormal sets, orthonormal bases



Fact: An $m \times n$ matrix A has orthonormal columns if and only if $A^T A = I$.

Fact: If $A^T A = I$, then

- $\|A\vec{x}\| = \|\vec{x}\|$
- $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$

(In particular, if $\vec{x} \cdot \vec{y} = 0$ then $(A\vec{x}) \cdot (A\vec{y}) = 0$)

Ex
$$A = \begin{bmatrix} \frac{1}{3} & 0 \\ \frac{2}{3} & 1/\sqrt{2} \\ \frac{2}{3} & -1/\sqrt{2} \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} 2 \\ \sqrt{2} \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\|\vec{x}\| = \sqrt{6}$ $\qquad$ $\|\vec{y}\| = 1$ $\qquad$ $\vec{x} \cdot \vec{y} = 2$

$$A\vec{x} = \begin{bmatrix} 2/3 \\ 7/3 \\ 1/3 \end{bmatrix} \qquad A\vec{y} = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

$$\|A\vec{x}\| = \sqrt{\tfrac{54}{9}} = \sqrt{6} \qquad \|A\vec{y}\| = 1 \qquad A\vec{x}\cdot A\vec{y} = 2$$

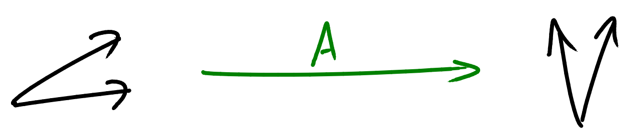
$(AB)^T = B^T A^T$

$$\left[\text{Why? } \|A\vec{x}\| = \sqrt{(A\vec{x})\cdot(A\vec{x})} = \sqrt{(A\vec{x})^T A\vec{x}} = \sqrt{\vec{x}^T A^T A \vec{x}} = \sqrt{\vec{x}^T \cdot I \cdot \vec{x}} = \sqrt{\vec{x}^T\vec{x}} = \sqrt{\vec{x}\cdot\vec{x}} = \|\vec{x}\|\right]$$

If A is $n\times n$ matrix, then A has orthonormal columns if and only if $\underline{A^T = A^{-1}}$.
We call such a matrix an <u>orthogonal matrix</u>. (Bad terminology: should be called "orthonormal matrix")

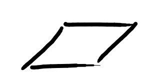
<u>Ex</u> $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is an <u>orthogonal matrix</u>.

$$\left(\text{why? } A^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \text{ has } A^TA = \begin{bmatrix} \cos^2+\sin^2 & 0 \\ 0 & \cos^2+\sin^2 \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I\right)$$

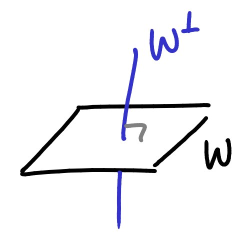


<u>Orthogonal Projection</u> (Sec 6.3)

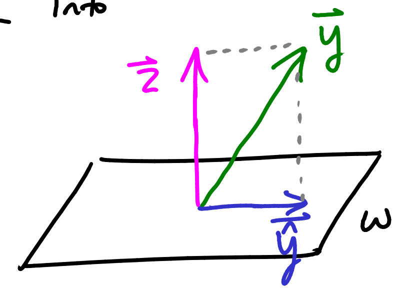
Say $W = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \text{ any real \#s} \right\}$. W is a subspace of $\mathbb{R}^3$.



$$W^\perp = \left\{ \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} : x_3 \text{ any real \#} \right\}$$



Any vector $\vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ in $\mathbb{R}^3$ can be <u>decomposed</u> into



$$\vec{y} = \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}}_{\text{in } W} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix}}_{\text{in } W^\perp}$$

$$= \hat{\vec{y}} + \vec{z}$$

Fact Say $W \subset \mathbb{R}^n$ is a <u>subspace</u> and $\vec{y} \in \mathbb{R}^n$.

Then there is a <u>unique</u> decomposition

$$\vec{y} = \hat{\vec{y}} + \vec{z}$$

where $\hat{\vec{y}}$ is in W and $\vec{z}$ is in $W^\perp$.

If $\{\vec{u}_1, \dots, \vec{u}_p\}$ is an <u>orthogonal basis</u> for W, then

$$\hat{\vec{y}} = \frac{\vec{y} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 + \cdots + \frac{\vec{y} \cdot \vec{u}_p}{\|\vec{u}_p\|^2}\vec{u}_p$$

$$\vec{z} = \vec{y} - \hat{\vec{y}}$$

We call $\hat{\vec{y}}$ the <u>orthogonal projection</u> of $\vec{y}$ to W.

Ex Say $\vec{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$ $\vec{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ $\vec{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$$

Write $\vec{y}$ as the sum of a vector in W and a vector in $W^\perp$.

$\vec{u}_1 \cdot \vec{u}_2 = 0$ so $\{\vec{u}_1, \vec{u}_2\}$ is an <u>orthogonal basis</u> for W.

So can use our formula:

$$\vec{y} = \hat{y} + \vec{z}$$

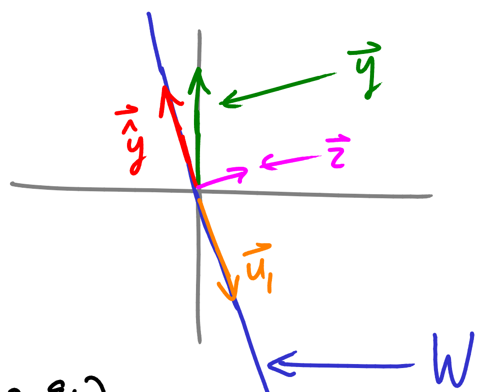
$$\hat{y} = \frac{\vec{y}\cdot\vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 + \frac{\vec{y}\cdot\vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 = \frac{9}{30}\begin{bmatrix}2\\5\\-1\end{bmatrix} + \frac{3}{6}\begin{bmatrix}-2\\1\\1\end{bmatrix} = \underline{\underline{\begin{bmatrix}-2/5\\2\\1/5\end{bmatrix}}}$$

$$\vec{z} = \vec{y} - \hat{y} = \begin{bmatrix}1\\2\\3\end{bmatrix} - \begin{bmatrix}-2/5\\2\\1/5\end{bmatrix} = \underline{\underline{\begin{bmatrix}7/5\\0\\14/5\end{bmatrix}}}$$

(Check: $\vec{z}\cdot\vec{u}_1 = 0$, $\vec{z}\cdot\vec{u}_2 = 0 \Rightarrow \vec{z} \in W^\perp$ as expected)

<u>Ex</u> $W = \left\{\begin{bmatrix}t\\-3t\end{bmatrix} : t \text{ a real \#}\right\} = \text{Span}\left\{\begin{bmatrix}1\\-3\end{bmatrix}\right\}$ $\quad \vec{u}_1 = \begin{bmatrix}1\\-3\end{bmatrix}$

$$\vec{y} = \begin{bmatrix}0\\3\end{bmatrix}$$



$$\vec{y} = \hat{y} + \vec{z}$$

$$\hat{y} = \frac{\vec{y}\cdot\vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 = \frac{-9}{10}\begin{bmatrix}1\\-3\end{bmatrix} = \begin{bmatrix}-9/10\\27/10\end{bmatrix}$$

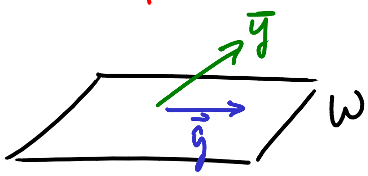
$$\vec{z} = \vec{y} - \hat{y} = \begin{bmatrix}0\\3\end{bmatrix} - \begin{bmatrix}-9/10\\27/10\end{bmatrix} = \begin{bmatrix}9/10\\3/10\end{bmatrix}$$

Fact: Say W is a subspace of $\mathbb{R}^n$

$\vec{y} \in \mathbb{R}^n$

The orthog. projection $\hat{\vec{y}}$ of $\vec{y}$ onto W is the <u>closest</u> vector to $\vec{y}$ in W.

i.e. for any $\vec{w} \in W$, $\operatorname{dist}(\vec{y}, \vec{w}) \geq \operatorname{dist}(\vec{y}, \hat{\vec{y}})$



Why? If $\vec{w} \in W$,

$$\vec{y} - \vec{w} = \underbrace{(\vec{y} - \hat{\vec{y}})}_{\text{in } W^\perp} + \underbrace{(\hat{\vec{y}} - \vec{w})}_{\text{in } W}$$

So

$$\|\vec{y} - \vec{w}\|^2 = \|\vec{y} - \hat{\vec{y}}\|^2 + \|\hat{\vec{y}} - \vec{w}\|^2$$
$$\geq \|\vec{y} - \hat{\vec{y}}\|^2$$

Least-Squares Problems

Consider a system of equations $A\vec{x} = \vec{b}$ A $m \times n$ matrix, $\vec{b} \in \mathbb{R}^m$

May be consistent or inconsistent.
If inconsistent, can still make an <u>approximate</u> solution:

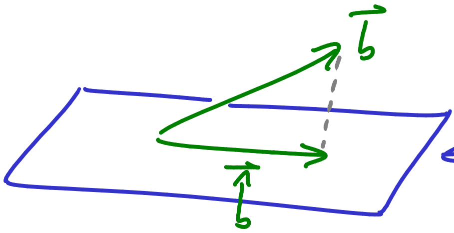
We say a <u>least-squares solution</u> of $A\vec{x} = \vec{b}$ is a vector $\hat{\vec{x}} \in \mathbb{R}^n$ such that $\|A\hat{\vec{x}} - \vec{b}\|$ is "as small as possible", i.e.

$$\|A\hat{\vec{x}} - \vec{b}\| \leq \|A\vec{x} - \vec{b}\| \text{ for every } \vec{x} \in \mathbb{R}^n.$$

Fact: The least-squares solutions of $A\vec{x} = \vec{b}$ are the same as the (ordinary) solutions of

$$A^T A \vec{x} = A^T \vec{b}.$$

Why?



$\mathrm{Col}\, A = \{\text{all vectors of the form } A\vec{x}\}$

Can't solve $A\vec{x} = \vec{b}$ but at least we can solve $A\hat{\vec{x}} = \hat{\vec{b}}$.

For this $\hat{\vec{x}}$, $\vec{b} - \hat{\vec{b}}$ lies in $(\mathrm{Col}\, A)^{\perp}$

But $(\mathrm{Col}\, A)^{\perp} = \mathrm{Nul}\, A^T$. i.e.

$$A^T(\vec{b} - \hat{\vec{b}}) = \vec{0}$$

$$A^T(\vec{b} - A\hat{\vec{x}}) = \vec{0}$$

$$\underline{A^T A \hat{\vec{x}} = A^T \vec{b}}$$

<u>Ex</u>

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

The system $A\vec{x} = \vec{b}$ is <u>inconsistent</u>. But we can look for a <u>least-squares solution</u>: i.e. solve $A^T A \vec{x} = A^T \vec{b}$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

This system is automatically consistent: solve it by row reduction, get $\hat{\vec{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$A\hat{\vec{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \qquad A\hat{\vec{x}} - \vec{b} = \begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix}$$

The difference $A\hat{\vec{x}} - \vec{b}$ has $\|A\hat{\vec{x}} - \vec{b}\| = \sqrt{84}$

("least-squares error")

and it's orthogonal to the columns of A.

If we have a subspace W of $\mathbb{R}^n$ and $\vec{y}$, and want to calculate $\hat{\vec{y}}$ we know how to do it if we have an orthogonal basis for W.

How can we get an orthog. basis if not given?

There's a standard algorithm for doing this: "Gram-Schmidt orthogonalization"

— start with any basis $\{\vec{v}_1, \ldots, \vec{v}_p\}$ for W

Our orthog basis is then $\{\vec{u}_1, \ldots, \vec{u}_p\}$

$$\vec{u}_1 = \vec{v}_1$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1$$

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2$$

$$\vdots$$

$$\vec{u}_p = \vec{v}_p - \frac{\vec{v}_p \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\vec{u}_1 - \frac{\vec{v}_p \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\vec{u}_2 - \cdots - \frac{\vec{v}_p \cdot \vec{u}_p}{\|\vec{u}_p\|^2}\vec{u}_p$$

[So in particular, <u>every</u> vector space (or every subspace of $\mathbb{R}^n$) <u>has</u> an orthogonal basis!]