Lecture 23 Math 225

Note Title 4/25/2008

6.2. & 6.3 Orthogonal sets and orthogonal Projections

<u>Def</u> A set of vectors in $\mathbb{R}^n$, $\{v_1, v_2, \ldots, v_p\}$ is called an orthogonal set if

$$v_i \cdot v_j = 0 \text{ for all } i \neq j$$
$$\|v_i\| \neq 0 \text{ for all } i$$

<u>Example</u> $\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} \right\}$ is an orthogonal set in $\mathbb{R}^3$.

<u>Theorem 1</u> If $\{v_1, v_2, \ldots, v_p\}$ is an orthogonal set then $\{v_1, v_2, \ldots, v_p\}$ is linearly independent.

<u>Proof</u> Let $c_1, c_2, \ldots, c_p$ be numbers such that

$$c_1 v_1 + c_2 v_2 + \ldots + c_p v_p = 0$$

Take the dot product of the above identity with v_1:

$$c_1 (v_1 \cdot v_1) + c_2 (v_1 \cdot v_2) + \ldots + c_p (v_1 \cdot v_p) = v_1 \cdot 0 = 0$$

But $v_1 \cdot v_2 = 0 = v_1 \cdot v_3 = \ldots = v_1 \cdot v_p$ because the set $\{v_1, v_2, \ldots, v_p\}$ is orthogonal. We get

$$c_1 \|v_1\|^2 = 0 \Rightarrow c_1 = 0 \text{ since } \|v_1\| \neq 0.$$

Similarly we find that $c_1 = c_2 = \dots = c_p = 0$
So $\{v_1, v_2, \dots, v_p\}$ is linearly indep.

<u>Theorem 2</u> If $\{v_1, v_2, \dots v_p\}$ is an orthogonal set and y is in span $\{v_1, v_2, \dots, v_p\}$ then:

$$y = \frac{v_1 \cdot y}{v_1 \cdot v_1} v_1 + \frac{v_2 \cdot y}{v_2 \cdot v_2} v_2 + \dots + \frac{v_p \cdot y}{v_p \cdot v_p} v_p$$

<u>Remark</u> If $W = \text{span}\{v_1, v_2, \dots, v_p\}$ then $\{v_1, v_2, \dots, v_p\}$ is a basis for W and the above theorem shows how to write any element in W as a linear combination of the vectors in the basis.

<u>Example</u> $\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} \right\}$ is an orthogonal set

and it spans $\mathbb{R}^3$ (why?). Using Theorem 2 we can write

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ as.}$$

$$y = \frac{8}{11} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + \frac{6}{6} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} + \frac{6}{33/2} \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

because $[3,1,1]^T \cdot y = 8$, $\|[3,1,1]^T\|^2 = 11$ ....

<u>Proof of Theorem 2</u> y in span $\{v_1, v_2, \dots, v_p\}$ $\Rightarrow$ there exists weights $c_1, c_2, \dots, c_p$ such that

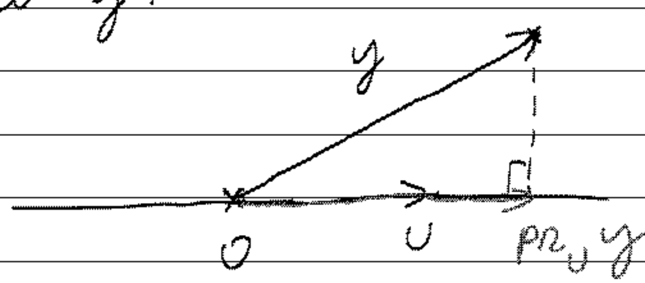
$$y = c_1 v_1 + c_2 v_2 + \dots + c_p v_p.$$

Take the dot product of the above with v_1 and use $v_1 \cdot v_2, v_1 \cdot v_3, \dots, v_1 \cdot v_p$ are all zero:

$$v_1 \cdot y = c_1 (v_1 \cdot v_1) \Rightarrow c_1 = \frac{v_1 \cdot y}{v_1 \cdot v_1}$$

Similarly $c_2 = \dfrac{v_2 \cdot y}{v_2 \cdot v_2}, \dots, c_p = \dfrac{v_p \cdot y}{v_p \cdot v_p}$.

<u>Projection on a line</u> Let y and v be vectors in $\mathbb{R}^n$. Consider the line span$\{v\}$. We are seeking the vector on the line that is closest to y:



Def $\mathrm{Pr}_v\, y$ is a vector in span$\{v\}$ such that

$$(y - \mathrm{Pr}_v\, y) \cdot v = 0$$

How to find $\mathrm{Pr}_v\, y$:

Let $z = y - \mathrm{Pr}_v\, y$ and $\mathrm{Pr}_v\, y = \alpha v$. Then

$$y = \alpha v + z \quad \text{and} \quad z \cdot v = 0$$

$$\Rightarrow v \cdot y = \alpha (v \cdot v) + \cancel{v \cdot z}^{\,0} \Rightarrow \alpha = \frac{v \cdot y}{v \cdot v}$$

So

$$\mathrm{Pr}_v\, y = \frac{v \cdot y}{v \cdot v}\, v$$

Example $y = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$\mathrm{Pr}_v\, y = \frac{8}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Remark $\mathrm{Pr}_v\, y$ is the closest vector in span$\{v\}$ to y. Any other vector βv will have

$$\| y - \beta v \|^2 = \| y - \mathrm{Pr}_v\, y + \mathrm{Pr}_v\, y - \beta v \|^2$$

$$\overset{\text{Pytagora}}{=} \| y - \mathrm{Pr}_v\, y \|^2 + \| \mathrm{Pr}_v\, y - \beta v \|^2$$

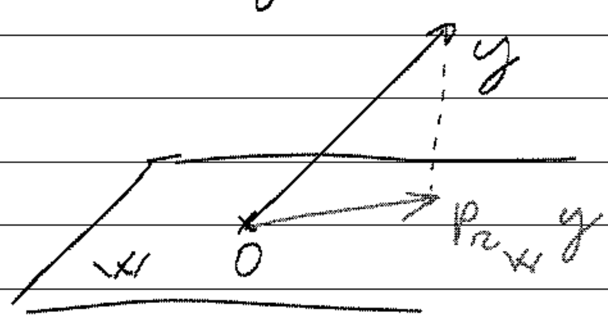
$$(y - \mathrm{Pr}_v\, y) \cdot (\mathrm{Pr}_v\, y - \beta y) = 0$$

Hence $\| y - \beta v \|^2 \geq \| y - \mathrm{Pr}_v\, y \|^2$ with equality iff $\beta v = \mathrm{Pr}_v\, y$.

Remark Theorem 2 can be restated as $y \in \text{span}\{v_1, v_2, \ldots, v_p\}$, $\{v_1, v_2, \ldots, v_p\}$ orthogonal set then

$$y = \text{Pr}_{v_1} y + \text{Pr}_{v_2} y + \ldots + \text{Pr}_{v_p} y$$

Projection on a subspace of $\mathbb{R}^n$ Let y be a vector in $\mathbb{R}^n$ and W a subspace of $\mathbb{R}^n$. As before we are looking for the vector in W closest to y:



Def $\text{Pr}_W y$ is the vector in W such that

$y - \text{Pr}_W y$ is orthogonal to all vectors in W

How to find $\text{Pr}_W y$:

Let $\{v_1, v_2, \ldots, v_p\}$ be an orthogonal basis in W.

Then, for some weights $c_1, c_2, \dots, c_p$ we have

$$Pr_W\, y = c_1 v_1 + c_2 v_2 + \dots + c_p v_p$$

Denote $z = y - Pr_W\, y$. We have

$$y = Pr_W\, y + z = c_1 v_1 + c_2 v_2 + \dots + c_p v_p + z$$

and $z \cdot v_1 = 0 = z \cdot v_2 = \dots = z \cdot v_p$.

$$v_1 \cdot y = c_1 (v_1 \cdot v_1) + c_2 \cancelto{0}{(v_1 \cdot v_2)} + \dots + c_p \cancelto{0}{(v_1 \cdot v_p)} + \cancelto{0}{v_1 \cdot z}$$

$$\Rightarrow c_1 = \frac{v_1 \cdot y}{v_1 \cdot v_1}$$

Similarly $c_2 = \dfrac{v_2 \cdot y}{v_2 \cdot v_2}, \dots, c_p = \dfrac{v_p \cdot y}{v_p \cdot v_p}$.

Hence

$$Pr_W\, y = \frac{v_1 \cdot y}{v_1 \cdot v_1} v_1 + \frac{v_2 \cdot y}{v_2 \cdot v_2} v_2 + \dots + \frac{v_p \cdot y}{v_p \cdot v_p} v_p$$

$$= Pr_{v_1} y + Pr_{v_2} y + \dots + Pr_{v_p} y$$

where $\{v_1, v_2, \dots, v_p\}$ is an orthogonal basis in W

Example $W = \text{span}\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$ $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Find $\mathrm{Pr}_W y$!

$$\mathrm{Pr}_W y = \frac{8}{11} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + \frac{6}{6} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

Because $[3,1,1]^T \cdot y = 8$, $\|[3,1,1]^T\|^2 = 11$

$[-1,2,1]^T \cdot y = 6$ $\|[-1,2,1]^T\|^2 = 6$

Remark $\mathrm{Pr}_W y$ is the closest vector in W to y because for any other vector v in W:

$$\|y - v\|^2 = \|y - \mathrm{Pr}_W y + \underbrace{\mathrm{Pr}_W y - v}_{\text{in } W}\|^2$$

$$\overset{\text{Pytagoras}}{=} \|y - \mathrm{Pr}_W y\|^2 + \|\mathrm{Pr}_W y - v\|^2$$

So $\|y-v\|^2 \geq \|y - \mathrm{Pr}_W y\|^2$ with equality iff

$v = \mathrm{Pr}_W y$!