Ch. 6 - Inner product spaces & orthogonality

§1. Applications of the dot product

Recall that the dot product of two vectors $\underset{\sim}{X}$ & $\underset{\sim}{Y}$ in $\mathbb{R}^n$ was defined by

$$\underset{\sim}{X} \circ \underset{\sim}{Y} = X_1Y_1 + X_2Y_2 + \dots + X_nY_n.$$

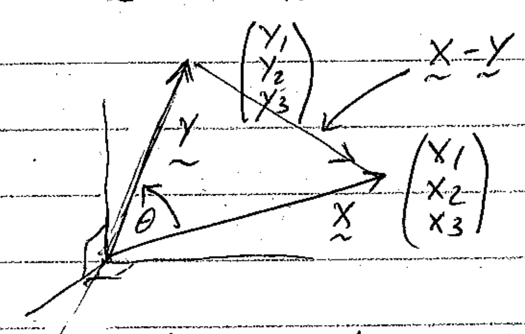
We can also write $\underset{\sim}{X} \cdot \underset{\sim}{Y}$ as $\underset{\sim}{X}^T\underset{\sim}{Y}$ where X^TY is the matrix product of $\underset{\sim}{X}^T$ and $\underset{\sim}{Y}$.

We also defined the length of a vector $\underset{\sim}{x}$ in $\mathbb{R}^n$ by

$$\|\underset{\sim}{X}\| = \sqrt{\underset{\sim}{X} \cdot \underset{\sim}{X}} = \sqrt{(x_1^2 + x_2^2 + \dots + x_n^2)}.$$

Prop. 1 Let $\underset{\sim}{X}$ & $\underset{\sim}{Y}$ be non-zero vectors in $\mathbb{R}^3$ and θ be the angle from $\underset{\sim}{x}$ to Y. Then $\cos\theta = \dfrac{\underset{\sim}{X} \cdot \underset{\sim}{Y}}{\|\underset{\sim}{X}\|\|\underset{\sim}{Y}\|}$

Proof:



Consider the triangle determined the origin and the position vectors $\underset{\sim}{x}$ & $\underset{\sim}{Y}$. Then by the law of cosines

$$\|\underset{\sim}{x} - \underset{\sim}{Y}\| = \|\underset{\sim}{X}\|^2 + \|\underset{\sim}{Y}\|^2 - 2\|\underset{\sim}{X}\|\,\|\underset{\sim}{Y}\| \cos\theta.$$

So $2\|\underset{\sim}{X}\|.\|\underset{\sim}{Y}\|.\cos\theta = \|\underset{\sim}{X}\|^2 + \|\underset{\sim}{Y}\|^2 - \|\underset{\sim}{X} - \underset{\sim}{Y}\|^2$

$$= (X_1^2 + X_2^2 + X_3^2) + (Y_1^2 + Y_2^2 + Y_3^2) - \{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + (X_3 - Y_3)^2\}$$

$$= 2X_1Y_1 + 2X_2Y_2 + 2X_3Y_3 = 2(\underset{\sim}{X} \cdot \underset{\sim}{Y})$$

$$\therefore \cos\theta = \frac{2(\underset{\sim}{X} \cdot \underset{\sim}{Y})}{2\|\underset{\sim}{X}\|\,\|\underset{\sim}{Y}\|} = \frac{\underset{\sim}{X} \cdot \underset{\sim}{Y}}{\|\underset{\sim}{X}\|\,\|\underset{\sim}{Y}\|}.$$

Note: The results holds in $\mathbb{R}^n$ also for any n, but it is not easy to visualize in $\mathbb{R}^n$ for $n \geq 4$.

Def. Let $\underset{\sim}{v}$ & $\underset{\sim}{w}$ be any two vectors in $\mathbb{R}^n$. We say that $\underset{\sim}{v}$ is <u>orthogonal</u> to $\underset{\sim}{w}$, and write $\underset{\sim}{v} \perp \underset{\sim}{w}$ if $\underset{\sim}{v} \cdot \underset{\sim}{w} = 0$.

We define the component of $\underset{\sim}{w}$ that is parallel to $\underset{\sim}{v}$ by
$\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \{(\underset{\sim}{w} \cdot \underset{\sim}{v}) / \|\underset{\sim}{v}\|^2\} \underset{\sim}{v}$, if $\underset{\sim}{v} \neq \underset{\sim}{0}$.
We define the component of $\underset{\sim}{w}$ that is orthogonal to $\underset{\sim}{v}$ by $\text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \underset{\sim}{w} - \text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w})$.

<u>Fact</u>: (a) If $\underset{\sim}{v} \perp \underset{\sim}{w}$, then the angle from $\underset{\sim}{v}$ to $\underset{\sim}{w}$ is $\pm 90°$.
(b) $\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) \perp \text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w})$.

<u>Proof</u> (a) Let θ be the angle from $\underset{\sim}{v}$ to $\underset{\sim}{w}$. Then
$\cos(\theta) = (\underset{\sim}{v} \cdot \underset{\sim}{w}) / \{\|\underset{\sim}{v}\| \, \|\underset{\sim}{w}\|\} = 0$. So $\theta = \pm 90°$.

(b)
$$\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) \cdot \text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \frac{(\underset{\sim}{w} \cdot \underset{\sim}{v})}{\|\underset{\sim}{v}\|^2}\left(\underset{\sim}{v} \cdot \left\{\underset{\sim}{w} - \frac{\underset{\sim}{w} \cdot \underset{\sim}{v}}{\|\underset{\sim}{v}\|^2}\underset{\sim}{v}\right\}\right)$$
$$= \frac{\underset{\sim}{w} \cdot \underset{\sim}{v}}{\|\underset{\sim}{v}\|^2}\left(\underset{\sim}{v} \cdot \underset{\sim}{w} - \frac{\underset{\sim}{w} \cdot \underset{\sim}{v}}{\|\underset{\sim}{v}\|^2}\|\underset{\sim}{v}\|^2\right)$$
$$= \frac{\underset{\sim}{w} \cdot \underset{\sim}{v}}{\|\underset{\sim}{v}\|^2}\left(\underset{\sim}{v} \cdot \underset{\sim}{w} - \underset{\sim}{v} \cdot \underset{\sim}{w}\right) = 0.$$
So $\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) \perp \text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w})$.

<u>Ex. 1</u> Let $\underset{\sim}{v} = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$ & $\underset{\sim}{w} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$. Find (a) $\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w})$ &
(b) $\text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w})$.

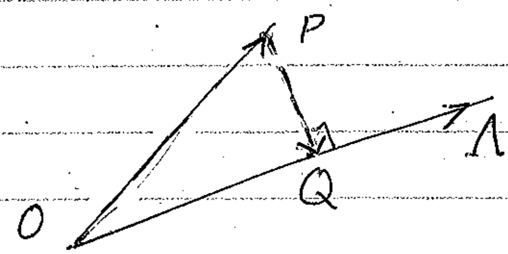
<u>Sol.</u> (a)
$$\text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \frac{(\underset{\sim}{w} \cdot \underset{\sim}{v})}{\|\underset{\sim}{v}\|^2}\underset{\sim}{v} = \frac{1}{4+4+1}\left\{\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}\right\}\underset{\sim}{v} = \frac{-3}{9}\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -2/3 \\ -2/3 \\ 1/3 \end{pmatrix}$$

(b)
$$\text{orthog}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \underset{\sim}{w} - \text{proj}_{\underset{\sim}{v}}(\underset{\sim}{w}) = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix} - \begin{pmatrix} -2/3 \\ -2/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 5/3 \\ -1/3 \\ 8/3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 5 \\ -1 \\ 8 \end{pmatrix}$$

Ex.2 (a) Find the point Q on the line $\Lambda = \{\alpha(1,2,-2)^T : \alpha \in \mathbb{R}\}$ that is closest to the point $P = (-1, 2, 1)^T$

(b) Find the shortest distance between P and the line Λ.

Sol. (a) Suppose $Q = c(1,2,-2)^T$. Then $\vec{PQ} = \vec{PO} + \vec{OQ} = \vec{OQ} - \vec{OP}$ must be orthogonal to the direction of the line Λ.



Thus $\begin{pmatrix}1\\2\\-2\end{pmatrix} \cdot (\vec{OQ} - \vec{OP}) = 0$.

Now $\begin{pmatrix}1\\2\\-2\end{pmatrix} \cdot (\vec{OQ} - \vec{OP}) = \begin{pmatrix}1\\2\\-2\end{pmatrix} \cdot \left\{\begin{pmatrix}c\\2c\\-2c\end{pmatrix} - \begin{pmatrix}-1\\2\\1\end{pmatrix}\right\} = \begin{pmatrix}1\\2\\-2\end{pmatrix} \cdot \begin{pmatrix}c+1\\2c+2\\-2c-1\end{pmatrix}$.

So $1.(c+1) + 2(2c-2) + (-2)(-2c+1) = 0$.

$\therefore \quad 9c - 3 = 0$. Hence $c = 1/3$.

Thus $Q = \frac{1}{3}\begin{pmatrix}1\\2\\-2\end{pmatrix} = \begin{pmatrix}1/3\\2/3\\-2/3\end{pmatrix}$. Q is called the projection of P onto Λ.

(b) Shortest distance between P and Λ $= \|\vec{PQ}\|$. Now

$\vec{PQ} = \vec{OQ} - \vec{OP} = \begin{pmatrix}1/3\\2/3\\-2/3\end{pmatrix} - \begin{pmatrix}1\\2\\1\end{pmatrix} = \begin{pmatrix}-2/3\\-4/3\\-5/3\end{pmatrix} = \frac{1}{3}\begin{pmatrix}-2\\-4\\-5\end{pmatrix}$. So

Shortest dist. $= \frac{1}{3}\sqrt{4+16+25} = \frac{1}{3}\sqrt{45} = \frac{1}{3}\sqrt{9}\sqrt{5} = \sqrt{5}$

Ex 3 (a) Find the point Q in the plane $\Pi = \left\{\alpha\begin{pmatrix}1\\1\\0\end{pmatrix} + \beta\begin{pmatrix}1\\0\\-1\end{pmatrix} : \alpha, \beta \in \mathbb{R}\right\}$ that is closest to the point $P = (1,2,-1)^T$

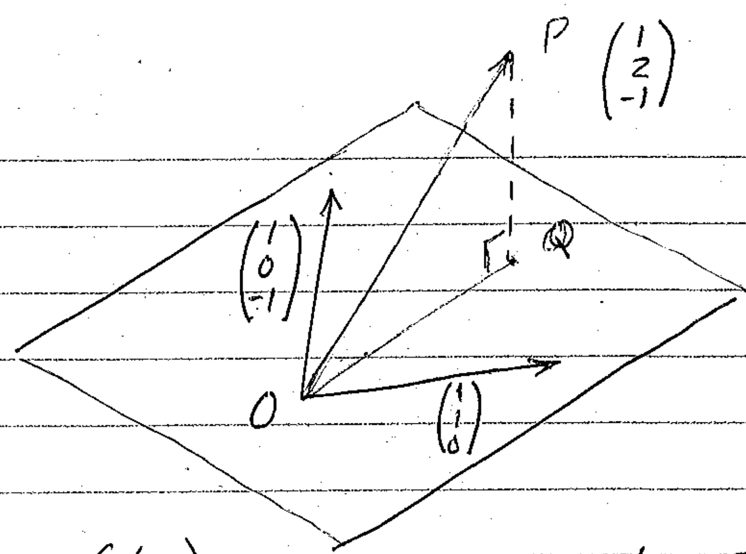
(b) Find the shortest distance between the point P and the plane Π.

Sol. (a) Suppose $Q = a\begin{pmatrix}1\\1\\0\end{pmatrix} + b\begin{pmatrix}1\\0\\-1\end{pmatrix}$. Then $\vec{PQ} = \vec{OQ} - \vec{OP}$ must be orthogonal to the plane Π. So $\vec{PQ}$ must be orthogonal to both of the vectors generating Π.

Ex. 3(a) $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ (4)

$= a\begin{pmatrix}1\\1\\0\end{pmatrix} + b\begin{pmatrix}1\\0\\-1\end{pmatrix} - \begin{pmatrix}1\\2\\-1\end{pmatrix}$

$= \begin{pmatrix}a+b-1\\a-2\\-b+1\end{pmatrix}$



So $\begin{pmatrix}1\\1\\0\end{pmatrix} \cdot \overrightarrow{PQ} = 0$

$\Rightarrow 1(a+b-1) + 1(a-2) + 0(-b+1) = 0 \quad \therefore\ 2a+b = 3$ (1)

And $\begin{pmatrix}1\\0\\-1\end{pmatrix} \cdot \overrightarrow{PQ} = 0 \Rightarrow 1(a+b-1) + 0(a-2) + (-1)(-b+1) = 0$

$\Rightarrow a + 2b = 2$ (2)

From (1) $b = 3 - 2a$. Sub. in (2) gives us

$a + 2(3-2a) = 2 \Longrightarrow 4 = 3a \Rightarrow a = 4/3$.

$\therefore\ b = 3 - 2(4/3) = 9/3 - 8/3 = 1/3$.

$\therefore\ Q = \frac{4}{3}\begin{pmatrix}1\\1\\0\end{pmatrix} + \frac{1}{3}\begin{pmatrix}1\\0\\-1\end{pmatrix} = \begin{pmatrix}5/3\\4/3\\-1/3\end{pmatrix}$. Q is called the projection of P onto Π.

(b) Shortest distance from P to $\Pi = \|\overrightarrow{PQ}\|$. Now $\overrightarrow{PQ} =$

$= \begin{pmatrix}5/3\\4/3\\-1/3\end{pmatrix} - \begin{pmatrix}1\\2\\-1\end{pmatrix} = \begin{pmatrix}2/3\\-2/3\\2/3\end{pmatrix} = \frac{2}{3}\begin{pmatrix}1\\-1\\1\end{pmatrix}$. So shortest

distance $= \frac{2}{3}\sqrt{1+1+1} = \frac{2}{3}\sqrt{3}$.

Remark: Ex. 2 & 3 shows how we can find the point Q in a subspace S of $\mathbb{R}^n$ that is nearest to a given point P in $\mathbb{R}^n$. An affine space is any set of vectors $T = \{\underline{a} + \underline{v} : \underline{v} \in S\}$ where $\underline{a}$ is a fixed vector in $\mathbb{R}^n$ and S is a subspace of $\mathbb{R}^n$. The same method (from Ex. 2 & 3) can be used to find the point Q in an affine space T of $\mathbb{R}^n$ that is nearest to a given point P in $\mathbb{R}^n$.

Theorem 2 : Let $\mathcal{U}$ be a subspace of the vector space V and $\underset{\sim}{v}$ be any vector in V. Then we can find unique vectors $\underset{\sim}{u} \in \mathcal{U}$ & $w \in V$ such that $v = \underset{\sim}{u} + w$ and $\underset{\sim}{w} \cdot \underset{\sim}{x} = 0$ for each $\underset{\sim}{x} \in \mathcal{U}$.

Proof: Let $\langle \underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_k \rangle$ be a basis of $\mathcal{U}$. Put $\underset{\sim}{u} = c_1 u_1 + c_2 \underset{\sim}{u}_2 + \cdots + c_k \underset{\sim}{u}$ where $c_1, \ldots, c_k$ are chosen such that

$$\underset{\sim}{u} \cdot \underset{\sim}{u}_i = \underset{\sim}{v} \cdot \underset{\sim}{u}_i \quad \text{for each } i = 1, \ldots, k.$$

This choice is possible because this is a system of k linear equations in k unknowns and since $\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_k$ are linearly independent, it will have a unique solution. Now let $\underset{\sim}{w} = \underset{\sim}{v} - \underset{\sim}{u}$. Then $\underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w}$ and for each $i = 1, \ldots, k$;

$$\underset{\sim}{w} \cdot \underset{\sim}{u}_i = (\underset{\sim}{v} - \underset{\sim}{u}) \cdot \underset{\sim}{u}_i = \underset{\sim}{v} \cdot \underset{\sim}{u}_i - \underset{\sim}{u} \cdot \underset{\sim}{u}_i = 0$$

Now if $\underset{\sim}{x} \in \mathcal{U}$, then $\underset{\sim}{x} = \alpha_1 \underset{\sim}{u}_1 + \alpha_2 \underset{\sim}{u}_2 + \cdots + \alpha_k \underset{\sim}{u}_k$ for some scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Thus

$$\begin{aligned} \underset{\sim}{w} \cdot \underset{\sim}{x} &= \underset{\sim}{w} \cdot (\alpha_1 \underset{\sim}{u}_1 + \cdots + \alpha_k \underset{\sim}{u}_k) \\ &= \alpha_1 (\underset{\sim}{w} \cdot \underset{\sim}{u}_1) + \cdots + \alpha_k (\underset{\sim}{w} \cdot \underset{\sim}{u}_k) \\ &= \alpha_1 (0) + \cdots + \alpha_k (0) = 0. \end{aligned}$$

So $\underset{\sim}{w} \cdot \underset{\sim}{x} = 0$ for each $\underset{\sim}{x} \in \mathcal{U}$.

Def. Let $\mathcal{U}$ be a subspace of the vector space V and $\underset{\sim}{v} \in V$. We define the projection of $\underset{\sim}{v}$ onto the subspace $\mathcal{U}$ by

$$\mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v}) = \text{the unique vector } \underset{\sim}{u} \in \mathcal{U} \text{ such that } \underset{\sim}{v} = \underset{\sim}{u} + \underset{\sim}{w} \text{ and } \underset{\sim}{w} \cdot \underset{\sim}{x} = 0 \text{ for each } \underset{\sim}{x} \in \mathcal{U}.$$

§2. Orthogonal subspaces

Def. Let S_1 and S_2 be two subspaces of $\mathbb{R}^n$. We say that S_1 is *orthogonal* to S_2, and write $S_1 \perp S_2$, if $\underset{\sim}{x} \cdot \underset{\sim}{y} = 0$ for each $\underset{\sim}{x} \in S_1$ & each $\underset{\sim}{y} \in S_2$.

Def. Let S be a subspace of $\mathbb{R}^n$. We define the *orthogonal complement* of S by

$$S^\perp = \{\underset{\sim}{x} \in \mathbb{R}^n : \underset{\sim}{x} \cdot \underset{\sim}{y} = 0 \text{ for each } \underset{\sim}{y} \in S\}.$$

Prop. 3 : If S is a subspace of $\mathbb{R}^n$, then (a) $S^\perp$ is also a subspace of $\mathbb{R}^n$ & (b) $S \cap S^\perp = \{\underset{\sim}{0}\}$.

Proof (a) Let S be a subspace of $\mathbb{R}^n$. Then $\underset{\sim}{0} \in S^\perp$ because $\underset{\sim}{0} \cdot \underset{\sim}{y} = 0$ for each $\underset{\sim}{y} \in S$. So $S^\perp \neq \emptyset$.
Now suppose $\alpha \in \mathbb{R}$ and $\underset{\sim}{x}_1, \underset{\sim}{x}_2 \in S^\perp$. Then for each $\underset{\sim}{y} \in S$,

$$(\underset{\sim}{x}_1 + \underset{\sim}{x}_2) \cdot \underset{\sim}{y} = \underset{\sim}{x}_1 \cdot \underset{\sim}{y} + \underset{\sim}{x}_2 \cdot \underset{\sim}{y} = 0 + 0 = 0$$

$$\& \quad (\alpha \underset{\sim}{x}_1) \cdot \underset{\sim}{y} = (\alpha \underset{\sim}{x}_1) \cdot \underset{\sim}{y} = \alpha(\underset{\sim}{x}_1 \cdot \underset{\sim}{y}) = \alpha(0) = 0.$$

So $x_1 + x_2 \in S^\perp$ and $\alpha \underset{\sim}{x}_1 \in S^\perp$. So $S^\perp$ is a subspace of $\mathbb{R}^n$.

(b) Now suppose $\underset{\sim}{x} \in S \cap S^\perp$. Then $\underset{\sim}{x} \cdot \underset{\sim}{x} = 0$ because $\underset{\sim}{x} \in S^\perp$ & $\underset{\sim}{x} \in S$. So $\|\underset{\sim}{x}\|^2 = 0 \Rightarrow \underset{\sim}{x} = \underset{\sim}{0}$.
Since $\underset{\sim}{0} \in S$ & $\underset{\sim}{0} \in S^\perp$, it follows that $S \cap S^\perp = \{\underset{\sim}{0}\}$

Ex. 1 (a) Let $S = \left\{\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R}\right\}$. Then $S^\perp = \left\{\beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} : \beta \in \mathbb{R}\right\}$

(b) Let $S = \left\{\alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} : \alpha \in \mathbb{R}\right\}$. Then $S^\perp = \left\{\beta \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : \beta, \gamma \in \mathbb{R}\right\}$.

Note: There will be many ways of expressing $S^\perp$.

Def. Let $\mathcal{U}$ & W be subspaces of the vector space V, We define $\mathcal{U}\cap W$ and $\mathcal{U}+W$ by

$$\mathcal{U}\cap W = \{\underset{\sim}{v}\in V : \underset{\sim}{v}\in \mathcal{U} \ \& \ \underset{\sim}{v}\in W\}$$
$$\mathcal{U}+W = \{\underset{\sim}{u}+\underset{\sim}{w} : \underset{\sim}{u}\in \mathcal{U} \ \& \ \underset{\sim}{w}\in W\}.$$

If $\mathcal{U}\cap W = \{\underset{\sim}{0}\}$, then we will write $\mathcal{U}\oplus W$ for $\mathcal{U}+W$ to indicate that $\mathcal{U}$ & W have no non-zero vector in common.

Prop. 4 Let $\mathcal{U}$ & W be subspace of the vector space V. Then (a) $\mathcal{U}\cap W$ is a subspace of V.
(b) $\mathcal{U}+W$ is a subspace of V.

Proof: Do for H.W.

Theorem 5: Let $\mathcal{U}$ be any subspace of $\mathbb{R}^n$ Then $\mathbb{R}^n = \mathcal{U}\oplus \mathcal{U}^\perp$.

Proof: Let v be any vector in $\mathbb{R}^n$. Then we can write $\underset{\sim}{v}=\underset{\sim}{u}+\underset{\sim}{w}$ where $v = \mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v})$, the projection of $\underset{\sim}{v}$ onto the subspace $\mathcal{U}$, and $w = v-u$. Since $\underset{\sim}{u}$ is the vector in $\mathcal{U}$ that is closest to $\underset{\sim}{v}$, it follows that $\underset{\sim}{v}-\underset{\sim}{u}$ is orthogonal to each vector in $\mathcal{U}$. Thus $w = \underset{\sim}{v}-\underset{\sim}{u}\in \mathcal{U}^\perp$. So $\mathbb{R}^n = \mathcal{U}+\mathcal{U}^\perp$. Since $\mathcal{U}\cap\mathcal{U}^\perp=\{\underset{\sim}{0}\}$, it follows that $\mathbb{R}^n = \mathcal{U}\oplus\mathcal{U}^\perp$.

Remark: The projection of $\underset{\sim}{v}$ onto $\mathcal{U}$ can be written as $u = (\underset{\sim}{v}\cdot\underset{\sim}{u}_1)\underset{\sim}{u}_1 + (\underset{\sim}{v}\cdot\underset{\sim}{u}_2)\underset{\sim}{u}_2 + \cdots + (\underset{\sim}{v}\cdot\underset{\sim}{u}_k)\underset{\sim}{u}_k$ where $\underset{\sim}{u}_1,\ldots,\underset{\sim}{u}_k$ is an orthonormal bases of $\mathcal{U}$. (See Sec. 6.4)

Corollary 6: Let $\mathcal{U}$ be a subspace of $\mathbb{R}^n$. Then
$(\mathcal{U}^\perp)^\perp = \mathcal{U}$.

Proof: Suppose $\underset{\sim}{y} \in \mathcal{U}$. Then $\underset{\sim}{x}.\underset{\sim}{y} = 0$ for each $\underset{\sim}{x} \in \mathcal{U}^\perp$.
So $\underset{\sim}{y}.\underset{\sim}{x} = 0$ for each $\underset{\sim}{x} \in \mathcal{U}^\perp$. Hence $\underset{\sim}{y} \in (\mathcal{U}^\perp)^\perp$,
because $(\mathcal{U}^\perp)^\perp = \{\underset{\sim}{y} \in \mathbb{R}^n : \underset{\sim}{y}.\underset{\sim}{x} = 0 \text{ for each } \underset{\sim}{x} \in \mathcal{U}^\perp\}$.
Thus $\mathcal{U} \subseteq (\mathcal{U}^\perp)^\perp$ ... (*)
Now suppose $\underset{\sim}{y} \in (\mathcal{U}^\perp)^\perp$. Then we can write $\underset{\sim}{y}$
in the form $\underset{\sim}{y} = \underset{\sim}{u} + \underset{\sim}{w}$ where $\underset{\sim}{u} \in \mathcal{U}$ & $\underset{\sim}{w} \in \mathcal{U}^\perp$
because $\mathbb{R}^n = \mathcal{U} \oplus \mathcal{U}^\perp$. So
w : $\underset{\sim}{y} - \underset{\sim}{u} = \underset{\sim}{w} \in \mathcal{U}^\perp$
Since $\underset{\sim}{y} \in (\mathcal{U}^\perp)^\perp$ and $\underset{\sim}{u} \in \mathcal{U} \subseteq (\mathcal{U}^\perp)^\perp$, it follows
that $\underset{\sim}{y} - \underset{\sim}{u} \in (\mathcal{U}^\perp)^\perp$. Thus $\underset{\sim}{y} - \underset{\sim}{u} \in (\mathcal{U}^\perp) \cap (\mathcal{U}^\perp)^\perp$.
But $(\mathcal{U}^\perp) \cap (\mathcal{U}^\perp)^\perp = \{\underset{\sim}{0}\}$. So $\underset{\sim}{y} - \underset{\sim}{u} = \underset{\sim}{0}$. Hence
$\underset{\sim}{y} = \underset{\sim}{u} \in \mathcal{U}$. Thus $(\mathcal{U}^\perp)^\perp \subseteq \mathcal{U}$ ... (**).
From (*) & (**), it follows that $(\mathcal{U}^\perp)^\perp = \mathcal{U}$.

Theorem 7 Let A be any $m \times n$ matrix. Then
(a) $[\mathrm{RowSp}(A)]^T = [\mathrm{Null}(A)]^\perp$ & $\mathbb{R}^n = \mathrm{Null}(A) \oplus [\mathrm{RowSp}(A)]^T$.
(b) $[\mathrm{CoNull}(A)]^T = [\mathrm{ColSp} A]^\perp$ & $\mathbb{R}^m = \mathrm{ColSp}(A) \oplus [\mathrm{CoNull}(A)]^T$.

Proof:
(a) $\underset{\sim}{x} \in \mathrm{Null}(A) \Leftrightarrow A\underset{\sim}{x} = \underset{\sim}{0} \Leftrightarrow \vec{r}_i(A)\,\underset{\sim}{x} = 0$ for each $i = 1, \dots, m$
$\Leftrightarrow [\vec{r}_i(A)]^T . \underset{\sim}{x} = 0$ for each row $\vec{r}_i$ of A
$\Leftrightarrow \underset{\sim}{y}.\underset{\sim}{x} = 0$ for each $\underset{\sim}{y} \in [\mathrm{RowSp}(A)]^T$
$\Leftrightarrow \underset{\sim}{x} \in \{[\mathrm{RowSp}(A)]^T\}^\perp$.
Thus $\mathrm{Null}(A) = \{[\mathrm{RowSp}(A)]^T\}^\perp$. Hence
$[\mathrm{Null}(A)]^\perp = \{[\mathrm{RowSp}(A)]^T\}^{\perp\perp} = [\mathrm{RowSp}(A)]^T$.

(b) $\underset{\sim}{x} \in [\mathrm{CoNull}(A)]^T \iff (\underset{\sim}{x})^T \in \mathrm{CoNull}(A) \iff (\underset{\sim}{x})^T A = \underset{\sim}{0}$

$\iff (x^T)\, \underset{\sim}{c}_j(A) = \underset{\sim}{0}$ for each column j of A.

$\iff \underset{\sim}{x} \cdot \underset{\sim}{c}_i(A) = 0$ for each column of A

$\iff \underset{\sim}{x} \cdot \underset{\sim}{y} = 0$ for each $y \in \mathrm{ColSp}(A)$

$\iff \underset{\sim}{x} \in [\mathrm{ColSp}(A)]^\perp$

Thus $[\mathrm{CoNull}(A)]^T = [\mathrm{ColSp}(A)]^\perp$.

The second part of the results follow immediately from Theorem 5.

Ex.2 Let $A = \begin{bmatrix} 1 & 2 & -3 & 2 \\ 1 & 2 & -2 & 0 \\ -2 & -2 & 4 & 0 \end{bmatrix}$, the matrix from Ex. 1 Sec 4.6.

a) Then $[\mathrm{RowSp}(A)]^T = \mathrm{span}\{(1,2,0,-4), (0,0,1,-2)\}^T$

$= \mathrm{span}\left\{\begin{pmatrix}1\\2\\0\\-4\end{pmatrix}, \begin{pmatrix}0\\0\\1\\-2\end{pmatrix}\right\}$.

Also $\mathrm{Null}(A) = \mathrm{span}\left\{\begin{pmatrix}-2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}4\\0\\2\\1\end{pmatrix}\right\}$. Now we can check that

$[\mathrm{RowSp}(A)]^T = [\mathrm{Null}(A)]^\perp$

(b) Also $[\mathrm{CoNull}(A)]^T = \mathrm{span}\{(0,2,1)\}^T = \mathrm{span}\left\{\begin{pmatrix}0\\2\\1\end{pmatrix}\right\}$

and $\mathrm{ColSp}(A) = \mathrm{span}\left\{\begin{pmatrix}1\\1\\-2\end{pmatrix}, \begin{pmatrix}-3\\-2\\4\end{pmatrix}\right\}$.

Now we can check that $[\mathrm{CoNull}(A)]^T = [\mathrm{ColSp}(A)]^\perp$.

(c) $\mathrm{span}\left\{\begin{pmatrix}-2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}4\\0\\2\\1\end{pmatrix}\right\} \oplus \mathrm{span}\left\{\begin{pmatrix}1\\2\\0\\-4\end{pmatrix}, \begin{pmatrix}0\\0\\1\\-2\end{pmatrix}\right\} = \mathbb{R}^4$ & $\mathrm{span}\left\{\begin{pmatrix}0\\2\\1\end{pmatrix}\right\} \oplus \mathrm{span}\left\{\begin{pmatrix}1\\1\\-2\end{pmatrix}, \begin{pmatrix}-3\\-2\\4\end{pmatrix}\right\} = \mathbb{R}$

§3. Least squares problems

Suppose we have n variables $x_1, x_2, \ldots, x_n$ and each experiment we do gives us a linear equation in $x_1, \ldots, x_n$. Then by doing n experiments, we will get a system of n linear equations in n unknowns. Since each experiment is independently done, we would expect the n equations to be linearly independent. So we can expect a unique solution set $x_1 = \hat{x}_1, \ldots, x_n = \hat{x}_n$. Unfortunately, because of experimental errors, this solution set might not be very accurate. So how can we improve the accurracy?

Well, we can do m experiments where $m > n$. Unfortunately this will inevitably lead to an inconsistent $m \times n$ system. So what can we do now? Well, we can find the solution of the system that is in some sense the "best" approximation. We will explain what "best" means soon.

Let A be an $m \times n$ matrix with rank n and $\underset{\sim}{b} \in \mathbb{R}^m$. If $m > n$, then the system of linear equations $A\underset{\sim}{x} = \underset{\sim}{b}$ will, in general, be inconsistent. [Recall that $A\underset{\sim}{x} = \underset{\sim}{b}$ is a consistent system iff $\underset{\sim}{b} \in \mathrm{ColSp}(A)$.] Now if $\underset{\sim}{b} \notin \mathrm{ColSp}(A)$, then we can find the vector $\underset{\sim}{u} \in \mathrm{ColSp}(A) \subseteq \mathbb{R}^m$ that is closest to $\underset{\sim}{b}$. So $\hat{\underset{\sim}{u}} = \mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{b})$ where $\mathcal{U} = \mathrm{ColSp}(A)$. Then we know that $\underset{\sim}{u} = A\hat{\underset{\sim}{x}}$ for a unique $\hat{\underset{\sim}{x}} \in \mathbb{R}^n$ & this $\hat{\underset{\sim}{x}}$ is our best approximation.

Qu: Let A be an $m \times n$ matrix and suppose the system $A\underset{\sim}{x} = \underset{\sim}{b}$ is inconsistent. How can we find a point $\hat{\underset{\sim}{x}} \in \mathbb{R}^n$ such that $\|A\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|$ is a minimum?

Sol. To find $A\hat{\underset{\sim}{x}}$ all we have to do is to find the projection of $\underset{\sim}{b}$ onto the subspace $ColSp(A)$ of $\mathbb{R}^m$. Now let $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n$ be the columns of A. Then $A = [\underset{\sim}{a}_1 \ \cdots \ \underset{\sim}{a}_n]$. Now in order to find $\hat{\underset{\sim}{x}}$, we must solve the equations

$$\begin{aligned}
\underset{\sim}{a}_1 \cdot (A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) &= 0, \quad \text{i.e.,} \quad \underset{\sim}{a}_1^T(A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) = 0 \\
\underset{\sim}{a}_2 \cdot (A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) &= 0, \quad \text{i.e.,} \quad \underset{\sim}{a}_2^T(A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) = 0 \\
&\vdots \\
\underset{\sim}{a}_n \cdot (A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) &= 0, \quad \text{i.e.,} \quad \underset{\sim}{a}_n^T(A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) = 0.
\end{aligned}$$

So
$$\begin{bmatrix} \underset{\sim}{a}_1^T \\ \vdots \\ \underset{\sim}{a}_n^T \end{bmatrix} (A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{i.e.,} \quad A^T(A\hat{\underset{\sim}{x}} - \underset{\sim}{b}) = \underset{\sim}{0}$$

So $\underset{n\times m}{A^T}\,\underset{m\times n}{A}\,\underset{n\times 1}{\hat{\underset{\sim}{x}}} = \underset{n\times m}{A^T}\,\underset{m\times 1}{\underset{\sim}{b}}$. This system of equations is called the <u>normal system</u> of equations of the system $A\underset{\sim}{x} = \underset{\sim}{b}$.

If $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n$ are linearly independent, the normal system $A^TA\underset{\sim}{x} = A^T\underset{\sim}{b}$ will have a unique solution $\hat{\underset{\sim}{x}}$. This solution $\hat{\underset{\sim}{x}}$ is called the <u>least squares solution</u> because $\|A\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|^2 = [(\vec{r}_1\hat{\underset{\sim}{x}} - b_1)^2 + \cdots + (\vec{r}_m\hat{\underset{\sim}{x}} - b_m)^2]$ will be a minimum. Here $\vec{r}_i$ = row i of A.

Theorem 8 Let A be any $m\times n$ matrix of rank n. Then the normal system $A^TAx = A^T\underset{\sim}{b}$ has a unique solution $\hat{\underset{\sim}{x}} = (A^TA)^{-1}A^T\underset{\sim}{b}$ which is the unique least squares solution of the system $A\underset{\sim}{x}=\underset{\sim}{b}$

Proof: We will first show that A^TA is non-singular. Suppose $(A^TA)\underset{\sim}{x} = \underset{\sim}{0}$. Let $\underset{\sim}{z} = A\underset{\sim}{x}$. Then $A^T\underset{\sim}{z} = A^T(A\underset{\sim}{x}) = (A^TA)\underset{\sim}{x} = \underset{\sim}{0}$. So $\underset{\sim}{z} \in Null(A^T) = [Co\,Null(A)]^T$. But $\underset{\sim}{z} = A\underset{\sim}{x} \in ColSp(A)$. But $[Co\,Null(A)]^T = [ColSp(A)]^{\perp}$. So $\underset{\sim}{z} \in ColSp(A) \cap [ColSp(A)]^{\perp} = \{\underset{\sim}{0}\}$. Hence $\underset{\sim}{z} = \underset{\sim}{0}$. So $A\underset{\sim}{x} = \underset{\sim}{0}$. Since A has rank n, $\underset{\sim}{x} = 0$. So (A^TA) is non-singular & hence is invertible. Thus $(A^TA)\underset{\sim}{x} = A^T\underset{\sim}{b}$ has a unique solution $\hat{\underset{\sim}{x}} = (A^TA)^{-1} A^T\underset{\sim}{b}$

Ex. 1 Find the least squares solution of the inconsistent system.

$$x_1 + 2x_2 = 0$$
$$x_1 + x_2 = 1$$
$$x_1 + 2x_2 = 0.1$$

Sol. We have $\begin{bmatrix}1 & 2\\ 1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}0\\ 1\\ 0.1\end{bmatrix}$. So $A = \begin{bmatrix}1 & 2\\ 1 & 1\\ 1 & 2\end{bmatrix}$ & $\underset{\sim}{b} = \begin{bmatrix}0\\ 1\\ .1\end{bmatrix}$

The normal system is $A^TA\hat{\underset{\sim}{x}} = A^T\underset{\sim}{b}$. So

$$\begin{bmatrix}1 & 1 & 1\\ 2 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & 2\\ 1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix} = \begin{bmatrix}1 & 1 & 1\\ 2 & 1 & 2\end{bmatrix}\begin{bmatrix}0\\ 1\\ 0.1\end{bmatrix}$$

$$\therefore \begin{bmatrix}3 & 5\\ 5 & 9\end{bmatrix}\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix} = \begin{bmatrix}1.1\\ 1.2\end{bmatrix}, \quad \therefore \begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix} = \frac{1}{27-25}\begin{bmatrix}9 & -5\\ -5 & 3\end{bmatrix}\begin{bmatrix}1.1\\ 1.2\end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix}9.9-6.0\\ -5.5+3.6\end{bmatrix} = \begin{pmatrix}1.95\\ -0.95\end{pmatrix}.$$

Ex. 1 Checking: $\hat{x}_1 + 2\hat{x}_1 - b_1 = 0.05 - 0 = 0.05$

$\hat{x}_1 + \hat{x}_2 - b_2 = 1.00 - 1 = 0$

$\hat{x}_1 + 2\hat{x}_2 - b_3 = 0.05 - 0.1 = -0.05$

So $\|A\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|^2 = (0.05)^2 + (0)^2 + (-0.05)^2 = 0.005$.

Now if we were to take any other vector $\underset{\sim}{y} \in R^2$, we will find $\|A\underset{\sim}{y} - \underset{\sim}{b}\|^2 > \|A\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|^2$. Let us try $\underset{\sim}{y} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. Then

$$A\underset{\sim}{y} - \underset{\sim}{b} = \begin{cases} 2 + 2(-1) - 0 = 0 \\ 2 + (-1) - 1 = 0 \\ 2 + 2(-1) - .1 = .1 \end{cases}$$

So $\|A\underset{\sim}{y} - \underset{\sim}{b}\|^2 = 0^2 + 0^2 + (.1)^2 = 0.010 > 0.005 = \|A\hat{\underset{\sim}{x}} - \underset{\sim}{b}\|^2$.

Ex. 2 Three experiments were done to estimate the value of x_1 and the following system was obtained

$$3x_1 = 2.7$$
$$-4x_1 = -4.0$$
$$5x_1 = 5.1$$

Find the least squares estimate $\hat{x}_1$ of x_1.

Sol. $\underbrace{\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}}_{A} x_1 = \underbrace{\begin{bmatrix} 2.7 \\ -4.0 \\ 5.1 \end{bmatrix}}_{\underset{\sim}{b}}$. So $\begin{bmatrix} 3 & -4 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix} \hat{x}_1 = \begin{bmatrix} 3 & -4 & 5 \end{bmatrix} \begin{bmatrix} 2.7 \\ -4.0 \\ 5.1 \end{bmatrix}$

$\therefore\ 50\hat{x}_1 = 8.1 + 16 + 25.5 = 49.6$. So $\hat{x}_1 = \frac{49.6}{50} = 0.992$

Checking $3\hat{x}_1 - 2.7 = 2.976 - 2.7 = 0.276$

$-4\hat{x}_1 - (-4) = -3.968 + 4 = 0.032$

$5x - 5.1 = 4.960 - 5.1 = -0.140$

$\|A\hat{x} - b\|^2 = 0.076176 + 0.001024 + 0.0196 = 0.0968$

Ex.2 Now if we try any other value $Y = 0.99$ say, we we see that

$$3Y - 2.7 = 0.27 \qquad \|AY - b\|^2 = 0.0729$$
$$-4Y - (-4) = 0.04 \qquad + 0.0016 + 0.0225 = 0.0971$$
$$5Y - (5.1) = -0.15$$

Ex.3 Find the least squares fit by a linear function $Y = c_0 + c_1 X$ of the data:

X	1	2	3
Y	-1	0	2

Sol. If we substitute the three values of x into the equation $Y = c_0 + c_1 X$ we get

$$c_0 + 1.c_1 = -1$$
$$c_0 + 2.c_1 = 0$$
$$c_0 + 3.c_1 = 2$$

So $\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \underbrace{\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}}_{b}$. $\therefore \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$

$\therefore \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$. $\therefore \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \frac{1}{42-36} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} -8/3 \\ 3/2 \end{bmatrix}$

$\therefore \; Y = -\frac{8}{3} + \frac{3}{2} X$. Check.

X	1	2	3
$\hat{c}_0 + \hat{c}_1 x$	-7/6	-2/6	11/6

Ex.4 Find the least squares fit by a quadratic function $y = c_0 + c_1 X + c_2 X^2$ of the data:

X	-1	0	1	2
Y	0	-1	0	2

(for masochists only)

Sol.

$$c_0 + c_1(-1) c_2 \cdot 1 = 0$$
$$c_0 + c_1(0) + c_2 \cdot 0 = -1$$
$$c_0 + c_1 \cdot 1 + c_3 \cdot 1 = 0$$
$$c_0 + c_1 \cdot 2 + c_3 \cdot 4 = 2$$

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \\ \hat{c}_2 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 11 & 3 & -5 \\ 3 & 9 & -5 \\ -5 & -5 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 8 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} -17 \\ -1 \\ 15 \end{bmatrix}$$

$\therefore \; Y = \frac{3X^2}{4} - \frac{X}{20} - \frac{17}{20}$

§4. Orthogonal sets of vectors & orthonormal bases

Def. Let $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. We say that $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k\}$ is an orthogonal set if $\underset{\sim}{v}_i \cdot \underset{\sim}{v}_j = 0$ for all i & j with $i \neq j$.

Ex. 1 (a) $\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ is an orthogonal set of vectors in $\mathbb{R}^2$.

(b) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$ is an orthogonal set of vectors in $\mathbb{R}^3$.

Prop. 9 If $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, then $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k\}$ is lin. indep.

Proof: Suppose $c_1 \underset{\sim}{v}_1 + c_2 \underset{\sim}{v}_2 + \cdots + c_k \underset{\sim}{v}_k = \underset{\sim}{0}$. Then by tak the dot product of both sides of this equation with $\underset{\sim}{v}_i$, we get

$$(c_1 \underset{\sim}{v}_1 + c_2 \underset{\sim}{v}_2 + \cdots + c_k \underset{\sim}{v}_k) \cdot \underset{\sim}{v}_i = \underset{\sim}{0} \cdot \underset{\sim}{v}_i$$

So $\quad 0 + 0 + \cdots + c_i (\underset{\sim}{v}_i \cdot \underset{\sim}{v}_i) + 0 + \cdots + 0 = 0$

Thus $c_i \cdot \|\underset{\sim}{v}_i\|^2 = 0$. Since $\underset{\sim}{v}_i$ is nonzero, $\|\underset{\sim}{v}_i\|^2 \neq 0$, it follows that $c_i = 0$ for each $i = 1, \ldots, k$.
Hence $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k\}$ is linearly independent.

Def. A basis $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\}$ of $\mathbb{R}^n$ is said to be an orthogonal basis if $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\}$ is an orthogonal set. We say that $\{\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_n\}$ is an orthonormal basis if in addition $\|\underset{\sim}{v}_i\| = 1$ for each $i = 1, \ldots, n$.

Ex. 2 (a) The set $\left\{ \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \end{pmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^2$.

(b) The set $\left\{ \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}, \begin{pmatrix} -3/5 \\ -4/5 \end{pmatrix} \right\}$ is an orthonormal basis of $\mathbb{R}^2$.

Prop. 10 Let $B = \langle \underset{\sim}{u}_1, u_2, \ldots, u_n \rangle$ be an ordered orthonormal basis of $\mathbb{R}^n$ and $\underset{\sim}{v}$ be a vector in $\mathbb{R}^n$. Then $[\underset{\sim}{v}]_B = \begin{bmatrix} \underset{\sim}{v} \cdot \underset{\sim}{u}_1 \\ \vdots \\ \underset{\sim}{v} \cdot \underset{\sim}{u}_n \end{bmatrix}$.

Proof: Suppose $[\underset{\sim}{v}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$. Then $\underset{\sim}{v} = c_1 \underset{\sim}{u}_1 + \cdots + c_n \underset{\sim}{u}_n$.

$$
\begin{aligned}
\text{So } \underset{\sim}{v} \cdot \underset{\sim}{u}_i &= (c_1 \underset{\sim}{u}_1 + c_2 \underset{\sim}{u}_2 + \cdots + c_n \underset{\sim}{u}_n) \cdot \underset{\sim}{u}_i \\
&= c_1 (u_1 \cdot u_i) + \cdots + c_i \, \underset{\sim}{u}_i \cdot \underset{\sim}{u}_i + \cdots + c_n (u_n \cdot u_i) \\
&= c_1 (0) + \cdots + c_{i-1}(0) + c_i \|\underset{\sim}{u}_i\|^2 + \cdots + 0 \\
&= c_i \|\underset{\sim}{u}\|^2 = c_i
\end{aligned}
$$

Hence $[\underset{\sim}{v}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \underset{\sim}{v} \cdot \underset{\sim}{u}_1 \\ \vdots \\ \underset{\sim}{v} \cdot \underset{\sim}{u}_n \end{bmatrix}$

Prop. 11 Let $\mathcal{U}$ be a subspace of $\mathbb{R}^n$ and $\{\underset{\sim}{u}_1, \ldots, u_k\}$ be an orthonormal basis of $\mathcal{U}$. Then for any vector v in $\mathbb{R}^n$, $\mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v}) = (\underset{\sim}{v} \cdot u_1) u_1 + \cdots + (\underset{\sim}{v} \cdot \underset{\sim}{u}_k) \underset{\sim}{u}_k$. Consequently, $\mathrm{Orthog}_{\mathcal{U}}(\underset{\sim}{v}) = \underset{\sim}{v} - \mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v})$.

Proof: Let $\underset{\sim}{u} = (\underset{\sim}{v} \cdot u_1) \underset{\sim}{u}_1 + \cdots + (\underset{\sim}{v} \cdot \underset{\sim}{u}_k) \underset{\sim}{u}_k$. Then $\underset{\sim}{u} \in \mathcal{U}$. Now put $\underset{\sim}{w} = \underset{\sim}{v} - \underset{\sim}{u}$. Then for each $i = 1, \ldots, k$ we have

$$
\begin{aligned}
\underset{\sim}{w} \cdot \underset{\sim}{u}_i &= (\underset{\sim}{v} - \underset{\sim}{u}) \cdot \underset{\sim}{u}_i \\
&= \underset{\sim}{v} \cdot u_i - \{(\underset{\sim}{v} \cdot \underset{\sim}{u}_1) \underset{\sim}{u}_1 + \cdots + (\underset{\sim}{v} \cdot \underset{\sim}{u}_k) \underset{\sim}{u}_k\} \cdot \underset{\sim}{u}_i \\
&= \underset{\sim}{v} \cdot u_i - \{0 + 0 + \cdots + \underset{\sim}{v} \cdot u_i + \cdots + 0 \\
&= (\underset{\sim}{v} \cdot \underset{\sim}{u}_i) - (\underset{\sim}{v} \cdot \underset{\sim}{u}_i) = 0.
\end{aligned}
$$

So $\underset{\sim}{w} \perp \underset{\sim}{u}_i$ for each $\underset{\sim}{u}_i$. Thus $\underset{\sim}{u} = \mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v})$ and $\underset{\sim}{w} = \underset{\sim}{v} - \underset{\sim}{u} = \underset{\sim}{v} - \mathrm{Proj}_{\mathcal{U}}(\underset{\sim}{v}) = \mathrm{Orthog}_{\mathcal{U}}(\underset{\sim}{v})$.

The Gram-Schmidt orthonormalization process.

Let $\{\underset{\sim}{x}_1, \dots, \underset{\sim}{x}_k\}$ be a basis of the subspace $\mathcal{U}$ of $\mathbb{R}^n$. The Gram-Schmidt process produces

(a) an orthogonal bases $\underset{\sim}{v}_1, \dots, v_k$ of $\mathcal{U}$ and

(b) an orthonormal basis $\underset{\sim}{u}_1, \dots, u_k$ of $\mathcal{U}$.

1. Let $\underset{\sim}{v}_1 = \underset{\sim}{x}_1$. 2. Put $\underset{\sim}{v}_2 = \underset{\sim}{x}_2 - \dfrac{(x_2 \cdot v_1)}{\|\underset{\sim}{v}_1\|^2} v_1$

3. Put $\underset{\sim}{v}_3 = x_3 - \dfrac{(x_3 \cdot v_1)}{\|\underset{\sim}{v}_1\|^2} v_1 - \dfrac{x_3 \cdot v_2}{\|\underset{\sim}{v}_2\|^2} \underset{\sim}{v}_2$

⋮

k Put $\underset{\sim}{v}_k = \underset{\sim}{x}_k - \dfrac{x_k \cdot v_1}{\|\underset{\sim}{v}_1\|^2} v_1 - \dfrac{x_k \cdot v_2}{\|v_2\|^2} v_2 - \cdots - \dfrac{x_k \cdot v_{k-1}}{\|\underset{\sim}{v}_{k-1}\|^2} v_{k-1}$.

Finally, put $\underset{\sim}{u}_i = \underset{\sim}{v}_i / \|\underset{\sim}{v}_i\|$ for each $i = 1, \dots, k$.

Ex. 3 Find an orthogonal basis & an orthogonal basis of the Rowspace of the matrix A on the right.

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Sol. Let $\underset{\sim}{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$, $\underset{\sim}{x}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix}$ and $\underset{\sim}{x}_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$. Then

$\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \underset{\sim}{x}_3\}$ is a basis for RowSp(A).

(a) $\underset{\sim}{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}$, $\underset{\sim}{v}_2 = \underset{\sim}{x}_2 - \dfrac{(x_2 \cdot v_1)}{\|\underset{\sim}{v}_1\|^2} \underset{\sim}{v}_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 2 \end{pmatrix} - \dfrac{6}{9}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$,

$\underset{\sim}{v}_3 = \underset{\sim}{x}_3 - \dfrac{\underset{\sim}{x}_3 \cdot \underset{\sim}{v}_1}{\|\underset{\sim}{v}_1\|^2} v_1 - \dfrac{\underset{\sim}{x}_3 \cdot \underset{\sim}{v}_2}{\|\underset{\sim}{v}_2\|^2} \underset{\sim}{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} - \dfrac{3}{9}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} - \dfrac{2}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$

Ex 3 So an orthogonal basis of RowSp(A) will be

$$\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \underset{\sim}{v}_3\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}. \quad \|\underset{\sim}{v}_1\| = 3,\ \|\underset{\sim}{v}_2\| = \sqrt{2} \text{ and } \|\underset{\sim}{v}_3\| = \sqrt{3}$$

(b) An orthormal basis for RowSp(A) with thus be:

$$\{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \underset{\sim}{u}_3\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Ex.4 Let $\underset{\sim}{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\underset{\sim}{x}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ & $\underset{\sim}{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Find an orthonormal basis of span$\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \underset{\sim}{x}_3\}$

a) $\underset{\sim}{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\quad \underset{\sim}{v}_2 = \underset{\sim}{x}_2 - \dfrac{\underset{\sim}{x}_2 \cdot \underset{\sim}{v}_1}{\|\underset{\sim}{v}_1\|^2}\underset{\sim}{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \dfrac{5}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$

$$\underset{\sim}{v}_3 = \underset{\sim}{x}_3 - \frac{\underset{\sim}{x}_3 \cdot \underset{\sim}{v}_1}{\|\underset{\sim}{v}_1\|^2}\underset{\sim}{v}_1 - \frac{\underset{\sim}{x}_3 \cdot \underset{\sim}{v}_2}{\|\underset{\sim}{v}_2\|^2}\underset{\sim}{v}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{(-1/3)}{6/9}\begin{pmatrix} -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2/3 \\ 2/3 \\ 2/3 \end{pmatrix} + \begin{pmatrix} -2/6 \\ 1/6 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \\ -1/2 \end{pmatrix}$$

$$\therefore \{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \underset{\sim}{v}_3\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

(b) $\|\underset{\sim}{v}_1\| = \sqrt{1^2+1^2+1^2} = \sqrt{3}$, $\|\underset{\sim}{v}_2\| = \frac{1}{3}\sqrt{2^2+1^2+1^2} = \frac{\sqrt{6}}{3}$ and $\|\underset{\sim}{v}_3\| = \frac{1}{2}\sqrt{0+1^2+(-1)^2} = \frac{\sqrt{2}}{2}$.

$$\therefore \{\underset{\sim}{u}_1, \underset{\sim}{u}_2, \underset{\sim}{u}_3\} = \left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{3}{\sqrt{6}} \cdot \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \frac{2}{\sqrt{2}} \cdot \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

$$= \left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

Inner product & normed spaces.

Def An inner product on a complex vector space V is any binary operation on V that assigns a complex number $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ to any $x, y \in V$ and is such that

(a) $\langle \underset{\sim}{x}, \underset{\sim}{x} \rangle \geq 0$ with equality iff $x = \underset{\sim}{0}$

(b) $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle = \overline{\langle \underset{\sim}{y}, \underset{\sim}{x} \rangle}$ for any $x, y \in V$ where $\bar{z}$ = conjugate of z.

(c) $\langle \alpha x, \underset{\sim}{y} \rangle = \alpha \langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ for all $\alpha \in \mathbb{R}$ & $x, y \in V$.

(d) $\langle \underset{\sim}{w} + \underset{\sim}{x}, \underset{\sim}{y} \rangle = \langle \underset{\sim}{w}, \underset{\sim}{y} \rangle + \langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ for all $\underset{\sim}{w}, \underset{\sim}{x}, \underset{\sim}{y} \in V$

Ex. 5 (a) Let $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle = \frac{1}{2}(x_1 y_1) + \frac{1}{3}(x_2 y_2) + \frac{1}{4}(x_3 y_3)$. Then $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ is an inner product on $\mathbb{R}^3$.

(b) Let $\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} A[i,j] \cdot B[i,j]$. Then $\langle A, B \rangle$ is an inner product on $\mathbb{R}^{m \times n}$.

(c) Let $\langle p(x), q(x) \rangle = \int_0^1 p(x) \cdot q(x)\, dx$. Then $\langle p(x), q(x) \rangle$ is an inner product on $C[0,1]$, the set of continuous functions on $[0,1]$.

Def. A norm on a vector space V is any unary operation on V that assigns a real number $\|\underset{\sim}{x}\|$ to each $\underset{\sim}{x} \in V$ such that

(a) $\|\underset{\sim}{x}\| \geq 0$ with equality iff $\underset{\sim}{x} = \underset{\sim}{0}$

(b) $\|\alpha \underset{\sim}{x}\| = |\alpha| \|\underset{\sim}{x}\|$ for all $\alpha \in \mathbb{C}$ & $\underset{\sim}{x} \in V$.

(c) $\|\underset{\sim}{x} + \underset{\sim}{y}\| \leq \|\underset{\sim}{x}\| + \|\underset{\sim}{y}\|$ for all $\underset{\sim}{x}, \underset{\sim}{y} \in V$.

Now every inner product $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ produces a norm $\|\underset{\sim}{x}\|_2 = \sqrt{\langle \underset{\sim}{x}, \underset{\sim}{x} \rangle}$. But there are norms that do not come from any inner products such as the ones in Ex. 6

Ex. 6 Let $\underset{\sim}{x}$ be any vector in $\mathbb{R}^n$. Put

$\|\underset{\sim}{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$ & $\|\underset{\sim}{x}\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$.

Then $\|x\|_1$ & $\|\underset{\sim}{x}\|_\infty$ are both norms on $\mathbb{R}^n$