24.3.15

Recap:

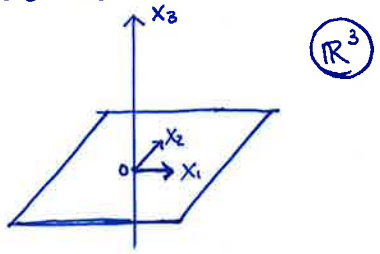
Two vectors are orthogonal if $\langle \underline{u}, \underline{v}\rangle = \underline{u}^T \underline{v} = 0$ (basically perpendicular).

Definition: Two subspaces U and W of vector space V are orthogonal, if for $\forall \underline{u} \in U$ and $\forall w \in W$, we have $\langle \underline{u}, \underline{w}\rangle = 0$.

Definition: Orthogonal complement of subspace M of vector space V contains every vector orthogonal to M. This subspace is usually denoted by $M^\perp$.

Remark: $\dim M + \dim M^\perp = \dim V$.

Example: Consider $\mathbb{R}^3$



line & plane - orthogonal subspace

orthogonal complement of each other.

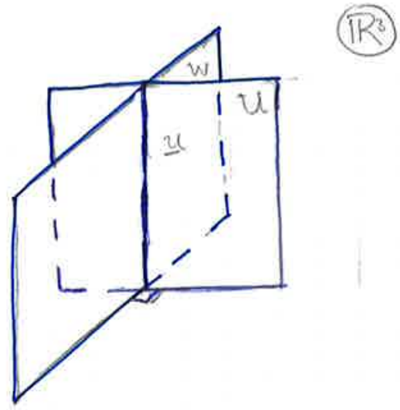
Example:

- Not orthogonal subspace!

$\underline{u} \neq 0$

$\underline{u} \in U$ & $\underline{u} \in W$

$\langle \underset{\in U}{\underline{u}}, \underset{\in W}{\underline{u}} \rangle \neq 0$



Note: if vector $\underline{u}$ belongs to 2 orthogonal subspaces, this vector is necessarily a zero vector, $\underline{u} = 0$.
Because we should have $\langle \underline{u}, \underline{u}\rangle = \underline{u}^T \underline{u} = 0 \Rightarrow \underline{u} = \underline{0}$.

Linear Mapping

definition: lets consider 2 vector spaces V and W. A function $L: V \to W$ is called a linear mapping, if:

① For any $v \in V$ and $\underline{v}' \in V$

$$L(\underline{v}+\underline{v}') = L(\underline{v}) + L(\underline{v}')$$

② For any $\underline{v} \in V$ and any scalar α

$$L(\alpha \underline{v}) = \alpha \cdot L(\underline{v}).$$

Example: Lets consider matrix $A \in \mathbb{R}^{n,m}$. We can define linear mapping L_A as follows:

$$L_A(\underline{v}) = A\underline{v}. \qquad L_A : \mathbb{R}^m \to \mathbb{R}^n$$

Is L_A a linear mapping? Yes!

① $\forall \underline{v}, v' \in \mathbb{R}^m \quad L_A(\underline{v}+\underline{v}') = A(\underline{v}+\underline{v}') = A\underline{v} + A\underline{v}'$
$= L_A(\underline{v}) + L_A(\underline{v}$

② $\forall \underline{v} \in \mathbb{R}^m, \ \forall \alpha$ - scalar

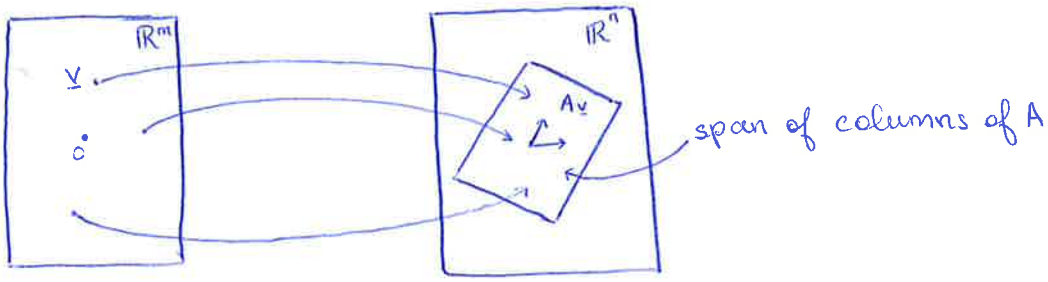
$$L_A(\alpha\underline{v}) = A(\alpha\underline{v}) = \alpha \cdot A\underline{v} = \alpha L_A(\underline{v}).$$

$\square$

Lets consider matrix $A \in \mathbb{R}^{n,m}$, $A : \mathbb{R}^m \to \mathbb{R}^n$.

Lets consider vector $\underline{v} \in \mathbb{R}^m$.

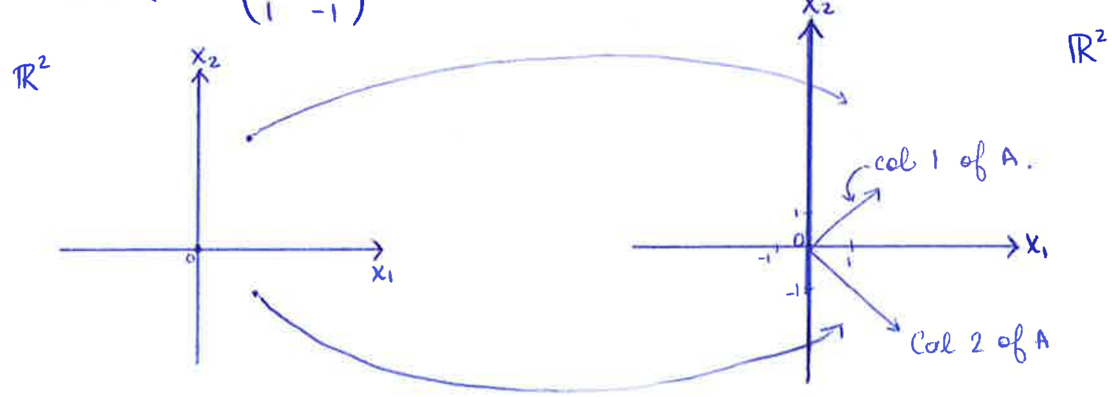
$$A\underline{v} = \underbrace{v_1 \cdot \begin{pmatrix} a_{11} \\ \\ a_{n1} \end{pmatrix} + v_2 \cdot \begin{pmatrix} a_{12} \\ \\ a_{n2} \end{pmatrix} + \dots + v_m \begin{pmatrix} a_{1m} \\ \\ a_{nm} \end{pmatrix}}_{\text{linear combination of columns of } A}$$



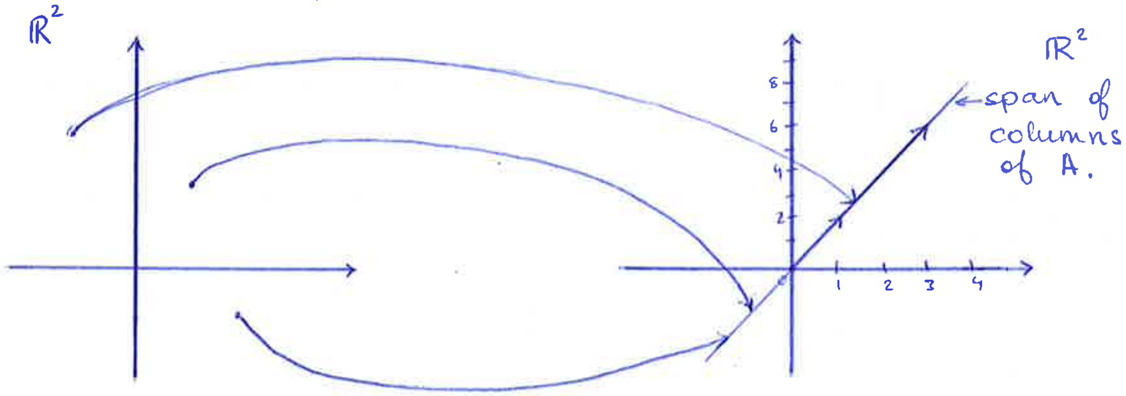
24.3.15

Example

$A \in \mathbb{R}^{2,2} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}$



Example: $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$



Note: In order for solution of $A\underline{x} = b$ to exist; $\underline{b}$ should belong to a span of columns of matrix A.

☆ Definition: The span of columns of matrix A is called a column space of A, denoted by C(A).

$$\boxed{C(A) \subset \mathbb{R}^n}$$

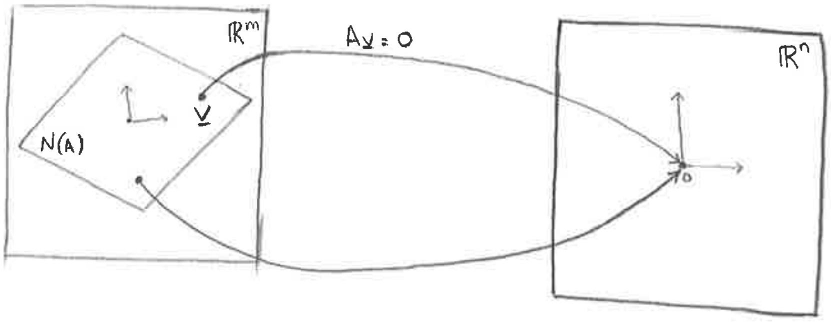
24.3.15

Definition:

Lets consider Matrix $A \in \mathbb{R}^{n,m}$, $A: \mathbb{R}^m \to \mathbb{R}^n$.

The null space of A is defined as $N(A) = \{ \underline{v} \in \mathbb{R}^m \mid A\underline{v} = \underline{0} \}$,

$N(A) \subset \mathbb{R}^m$.



Example: $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$

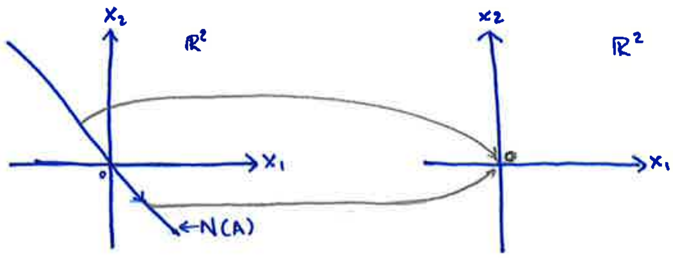
What is N(A) - ?

We should find all solutions of $A\underline{x} = \underline{0}$, this will give us N(A).

$\begin{cases} x_1 + 3x_2 \\ 2x_1 + 6x_2 \end{cases}$ $\begin{cases} x_1 + 3x_2 = 0 \\ 0 = 0 \end{cases}$ $\to$ line - $\alpha \times \begin{pmatrix} -3 \\ 1 \end{pmatrix}$ for all possible α.

$x_1 = -3x_2$, $x_2 = \alpha$ $\to$ $\alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix}$, $\alpha \begin{pmatrix} -6 \\ 2 \end{pmatrix}$

$= -3\alpha$.



24.3.15

Theorem: The nullspace, $N(A)$, of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^m$.

Proof: lets assume that $\underline{x}, \underline{x}' \in N(A)$ and α is arbitrary scalar.

$$A(\underline{x}+\underline{x}') = A\underline{x} + A\underline{x}' = \underline{0}+\underline{0} = \underline{0} \Rightarrow (\underline{x}+\underline{x}') \in N(A)$$

$$A(\alpha\underline{x}) = \alpha(A\underline{x}) = \alpha \times \underline{0} = \underline{0} \Rightarrow \alpha\underline{x} \in N(A) \qquad \boxtimes$$

Theorem: The column space, $C(A)$ of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^n$.

Proof: In Homework.

Definition: The row space of matrix A is a span of rows of A.

Clearly, $R(A) = C(A^T)$, $R(A) \subset \mathbb{R}^m$

Definition: The left null space of A is defined as $N(A^T)$.

$N(A^T) \subset \mathbb{R}^n$.

Theorem: $R(A)$ is a subspace of $\mathbb{R}^m$.

Proof: same as for $C(A)$ but for A^T.

Theorem: $N(A^T)$ is a subspace of $\mathbb{R}^n$.

Proof: as for $N(A)$ but replace A with A^T

Theorem: $R(A)$ and $N(A)$ are orthogonal subspaces in $\mathbb{R}^m$ for $A \in \mathbb{R}^{n,m}$.

Proof: Lets consider $\forall \underline{x} \in N(A)$, $A\underline{x} = \underline{0}$

$$A\underline{x} = \begin{pmatrix} \text{-row 1 of } A \rightarrow \\ \vdots \\ \text{-row } n \text{ of } A \rightarrow \end{pmatrix} \begin{pmatrix} \downarrow \\ \underline{x} \\ \downarrow \end{pmatrix} = \begin{pmatrix} \langle \text{row 1 of } A, \underline{x} \rangle \\ \vdots \\ \langle \text{row } n \text{ of } A, \underline{x} \rangle \end{pmatrix} \overset{\underline{x} \in N(A)}{=} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\Rightarrow \underline{x}$ is orthogonal to every row of A.

$\underline{x}$ is orthogonal to every linear combination of rows of A.

$\therefore \underline{x}$ is orthogonal to $R(A)$.

In fact what we just showed, is that $N(A)$ & $R(A)$ are orthogonal complements.

Theorem: $N(A^T)$ & $C(A) = R(A^T)$ are orthogonal complements in $\mathbb{R}^n$

Proof: similar to previous one.

26·3·15

$A \in \mathbb{R}^{n,m} : \mathbb{R}^m \to \mathbb{R}^n$

Row rank of A = column rank of A = rank $A = \dim(R(A)) = \dim(C(A))$.

$N(A) : A\underline{x} = \underline{0} \quad \forall x \in \mathbb{R}^m$.

$C(A) : A\underline{v}$ = linear combinations of columns of A

$= v_1 \times \text{col}1 A + \dots + v_n \times \text{col}n A \in \mathbb{R}^n$.

Theorem: $N(A)$ is an orthogonal complement of $R(A)$ in $\mathbb{R}^m$,

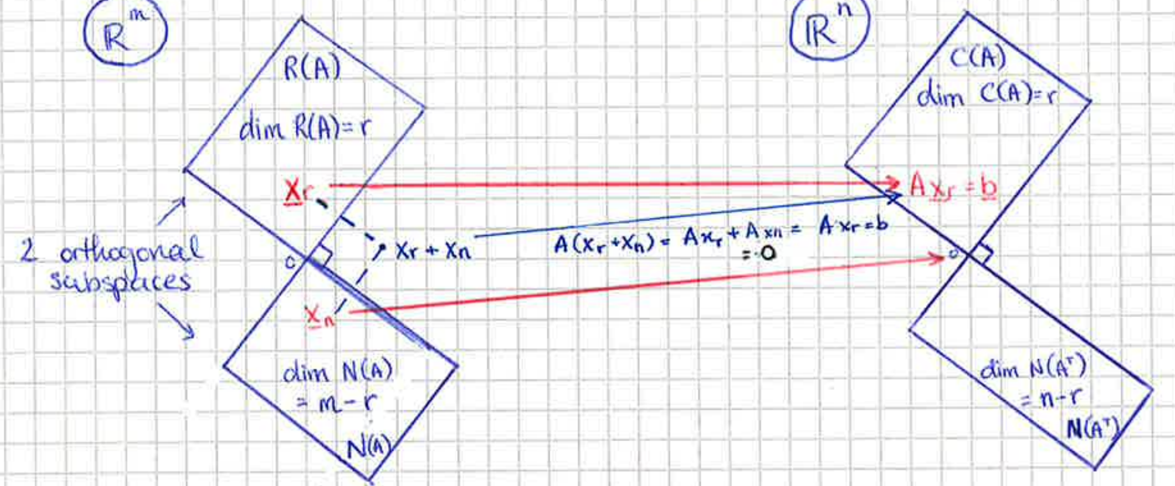
$\dim N(A) + \underbrace{\dim R(A)}_{= \text{rank} A} = m$.

Theorem: $N(A^T)$ is an orthogonal complement of $R(A^T) = C(A)$ in $\mathbb{R}^n$,

$\dim N(A^T) + \underbrace{\dim C(A)}_{= \text{rank} A} = n$.

26.3.15

Lets consider $A \in \mathbb{R}^{n,m}$, $A: \mathbb{R}^m \to \mathbb{R}^n$, rank $A = r$



Lemma: For any vector $\underline{b}$ in C(A), there exists one and only one vector $\underline{x}_r \in R(A)$ such that $A\underline{x}_r = \underline{b}$.

Proof: Lets assume that x_r and x_r' are in the row space, R(A). Lets also assume $Ax_r = A\underline{x}_r'$.

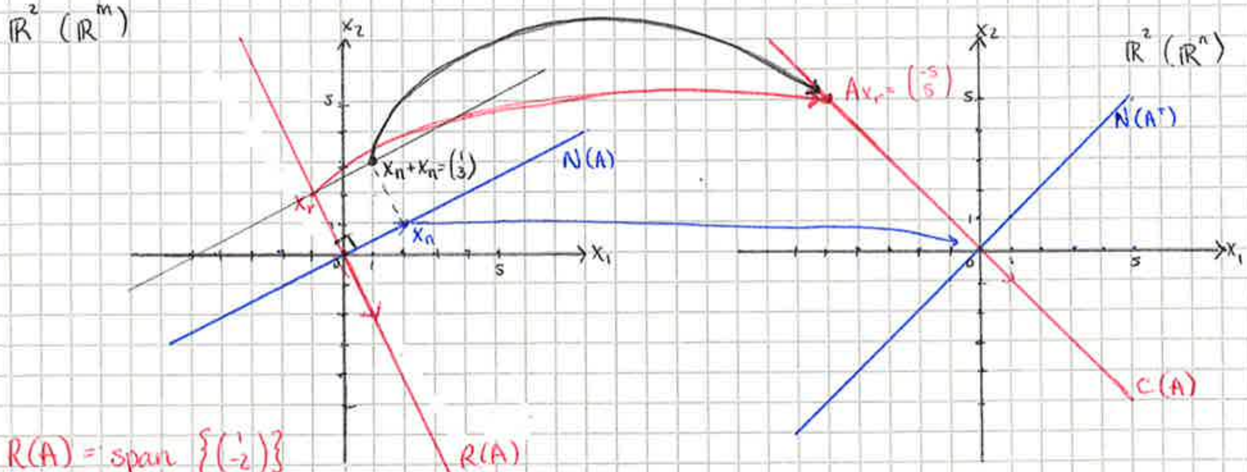
we have $\overset{\in R(A)}{\underline{x}_r} - \overset{\in R(A)}{\underline{x}_r'} \in R(A)$

But also we have $A\underline{x}_r - A\underline{x}_r' = \underbrace{A(\underline{x}_r - \underline{x}_r')}_{\in N(A)} = \underline{0}$

It means that $(\underline{x}_r - \underline{x}_r')$ is in R(A) and N(A), but they are orthogonal subspaces, therefore $\underline{x}_r - \underline{x}_r' = \underline{0} \Rightarrow \underline{x}_r = \underline{x}_r'$.

Example:

Lets consider $A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \in \mathbb{R}^{2,2}$



R(A) = span $\left\{ \binom{1}{-2} \right\}$

Row space: Rank $A = 1 \Rightarrow \dim R(A) = 1$

$R(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\} = \text{span}\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$

Null space: $\dim N(A) = \underset{\text{"m"}}{2} - \underset{\dim R(A)}{1} = 1$

$A\underline{x} = \underline{0}$

$\begin{cases} x_1 - 2x_2 = 0 \\ -x_1 + 2x_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1 - 2x_2 = 0 \\ 0 = 0 \end{cases} \Rightarrow x_1 = 2x_2$ - line.

Column space: $\dim C(A) = \dim R(A) = \underline{1}$

$C(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix} \right\} = \text{span}\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$

Left Null space: $\dim N(A^T) = 2 - 1 = 1$

Consider $x_r = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $A\underline{x}_r = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \underline{\underline{\begin{pmatrix} -5 \\ 5 \end{pmatrix}}}$

$x_n = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow Ax_n = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

31.3.15

Orthogonal Basis and Gram-Schmidt process

definition: Vectors $q_1 \cdots q_m$ are orthogonal if:

$$\langle q_i, q_j \rangle = q_i^T q_j = 0 \qquad \text{if } i \neq j$$

definition: Vectors $q_1 \cdots q_m$ are orthonormal if:

$$\langle q_i ; q_j \rangle = q_i^T q_j = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

If the columns of matrix are orthonormal vectors, then this matrix is usually denoted by Q.

In this case, we have $Q^T Q = I$

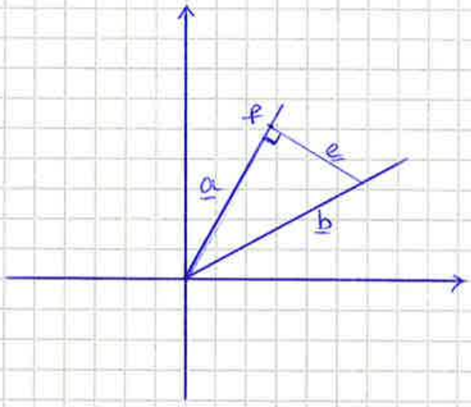
If Q is not a square matrix then QQ^T is not necessarily I.

Warning: Confusing

definition: A square matrix is called orthogonal (If its columns are orthonormal vectors)
If $Q^T Q = I$. In this case, since it is a square matrix $QQ^T = I$.

Projection on the line

Lets assume we have a line given by vector $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ and vector $\underline{b} \in \mathbb{R}^n$. We want to find vector $\underline{p}$ belonging to the line, closest to vector $\underline{b}$. In other words, we are looking for $\underline{p}$ which is orthogonal projection of $\underline{b}$ onto the line given by $\underline{a}$.



$\underline{p}$ is proportional to $\underline{a}$,
$\underline{p} = \hat{x}\,\underline{a}$, $\hat{x}$ is some scalar

lets define vector $\underline{e} = \underline{b} - \underline{p}$
$= \underline{b} - \hat{x}\,\underline{a}$
(error vector)

$\underline{e}$ is orthogonal to the line,
therefore $\langle \underline{a}, \underline{e} \rangle = 0$.

$$\langle a, \underline{e} \rangle = \underline{a}^T(\underline{b} - \hat{x}\,\underline{a}) = a^T\underline{b} - \hat{x}\,a^Ta = 0$$

$$\Rightarrow \hat{x} = \frac{a^T b}{\underline{a}^T\underline{a}}$$

$$\Rightarrow p = \hat{x}\,\underline{a} = \underline{a}\,\hat{x} = a\frac{a^T b}{a^T a}$$

$$= \boxed{\frac{a a^T}{a^T a}} \times \underline{b}$$

$P \in \mathbb{R}^{n,n}$
(projection matrix)

<u>Example</u>

Lets consider $\underline{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \in \mathbb{R}^3$

$$P = \frac{a a^T}{a^T a} = \left\langle \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, (1 \ 2 \ 2) \right\rangle \times \frac{1}{9}$$

$$= \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}$$

lets take $\underline{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $p = P\underline{b} = \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 5 \\ 10 \\ 10 \end{pmatrix}$

<u>Note</u>: $P^2 = P$

<u>Note</u>: $(I - P)$ - projection onto subspace orthogonal to the line given by $\underline{a}$.

31.3.15

Gram-Schmidt process

Given linear independant vectors $\underline{a}, \underline{b}, \underline{c}, \cdots$, we first find orthogonal vectors $\underline{a}', \underline{b}', \underline{c}', \cdots$ which span the same subspace as $a, b, c \ldots$, and then we normalise them,

$$q_1 = \frac{\underline{a}'}{\|\underline{a}'\|}, \quad q_2 \ \frac{b'}{\|\underline{b}'\|}, \quad q_3 = \frac{\underline{c}'}{\|\underline{c}'\|}, \cdots$$

So, Gram-Schmidt process allows us to construct an orthogonal basis of span $\{\underline{a}, \underline{b}, \underline{c}, \cdots\} \in$

① Choose $\underline{a}' = \underline{a}$.

② It is likely that $\underline{b}$ is not orthogonal to $\underline{a}'$, so we need to subtract its projection on the line defined by $\underline{a}'$:

$$\underline{b}' = \underline{b} - \frac{a^T \underline{b}}{a^T a} a'$$

③ c' is likely not orthogonal to a' and b'.

Again, subtract its projections

$$c' = \underline{c} - \frac{a^T c}{a^T a} \underline{a}' - \frac{b'^T c}{b'^T b'^T} b'$$

and so on...

and finally normalise. $q_1, q_2, q_3 \cdots$

Example

$\underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $\underline{b} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$, $\underline{c} = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$, find $\underline{a}', \underline{b}', \underline{c}'$, q_1, q_2, q_3

① $a' = \underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

② $b' = \underline{b} - \frac{a'^T b}{a'^T a'} \underline{a}' = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \frac{\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \rangle}{\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \rangle} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$

(quick check: $\langle \underline{a}', b' \rangle = \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \rangle = 0$)

③ $\underline{c}' = \underline{c} - \frac{a'^T c}{a'^T a'} a' - \frac{b'^T c}{b'^T b'} b' = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$\langle \underline{a}', \underline{c}' \rangle = 0$, $\langle \underline{b}', \underline{c}' \rangle = 0$.

Finally normalise:

$q_1 = \frac{a'}{\|\underline{a}'\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$; $q_2 = \frac{b'}{\|b'\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$, $q_3 = \frac{\underline{c}'}{\|\underline{c}'\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Projection onto subspace

Assume we have linearly independant vectors $a_1, \dots, a_m \in \mathbb{R}^n$.
We want to project vector $\underline{b} \in \mathbb{R}^n$ onto subspace spanned by $a_1, \dots, a_m$.

Subspace consists of all linear combinations

$$x_1 a_1 + \dots + a_m = \underbrace{\begin{pmatrix} | & & | \\ a_1 & \cdots & a_m \\ \downarrow & & \downarrow \end{pmatrix}}_{A \in \mathbb{R}^{n,m}} \underset{\mathbb{R}^m}{\overset{\bar{x}}{\Cap}}$$

We are looking for projection $\underline{p}$ of $\underline{b}$ onto this subspace.

31.3.15

We can define $\underline{e} = \underline{b} - \underline{p}$, $\underline{e}$ should be orthogonal to all $a_1, \dots, a_m$.

$$\begin{aligned} \langle a_1, e \rangle &= a_1^T \times (\underline{b} - A\hat{x}) = 0 \\ &\vdots \\ \langle a_m, e \rangle &= a_m^T \times (\underline{b} - A\hat{x}) = 0 \end{aligned} \Rightarrow \underbrace{\begin{pmatrix} \text{---} \; a_1^T \rightarrow \\ \vdots \\ \text{---} \; a_m^T \rightarrow \end{pmatrix}}_{A^T} (b - A\hat{x}) = 0$$

$$\therefore \quad A^T(\underline{b} - A\hat{\underline{x}}) = 0$$

$$A^Tb - A^TA\hat{x} = 0$$

<u>Theorem</u>: A has linearly independant columns.

Then $\underbrace{A^TA}_{\substack{\text{square} \\ \text{symmetric} \\ \text{invertible}}}$ is invertible

Proof: later

$$\underline{\hat{x}} = (A^TA)^{-1}A^T\underline{b}$$

$$\underline{p} = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}_{P\text{ - projection matrix}}b. \quad \text{- projection vector.}$$

x———————— MIDTERM ————————x