$W^{\perp}$: orthogonal complement of a subspace W of $\mathbb{R}^k$

Matrix Situation:

$$A \in \mathbb{R}^{m \times n}$$

For $\mathbb{R}^n$: R_{A^T} and $\mathcal{N}_A$

For $\mathbb{R}^m$: R_A and $\mathcal{N}_{A^T}$

We had $\begin{cases} \mathcal{N}_A = R_{A^T}^{\perp} \\ \mathcal{N}_A^{\perp} = R_{A^T} \end{cases}$ & hence and

$$\begin{cases} N_{A^T} = R_A^{\perp} \\ N_{A^T}^{\perp} = R_A \end{cases}$$

$$\dim R_{A^T} + \dim R_{A^T}^{\perp} = \dim \mathbb{R}^n$$

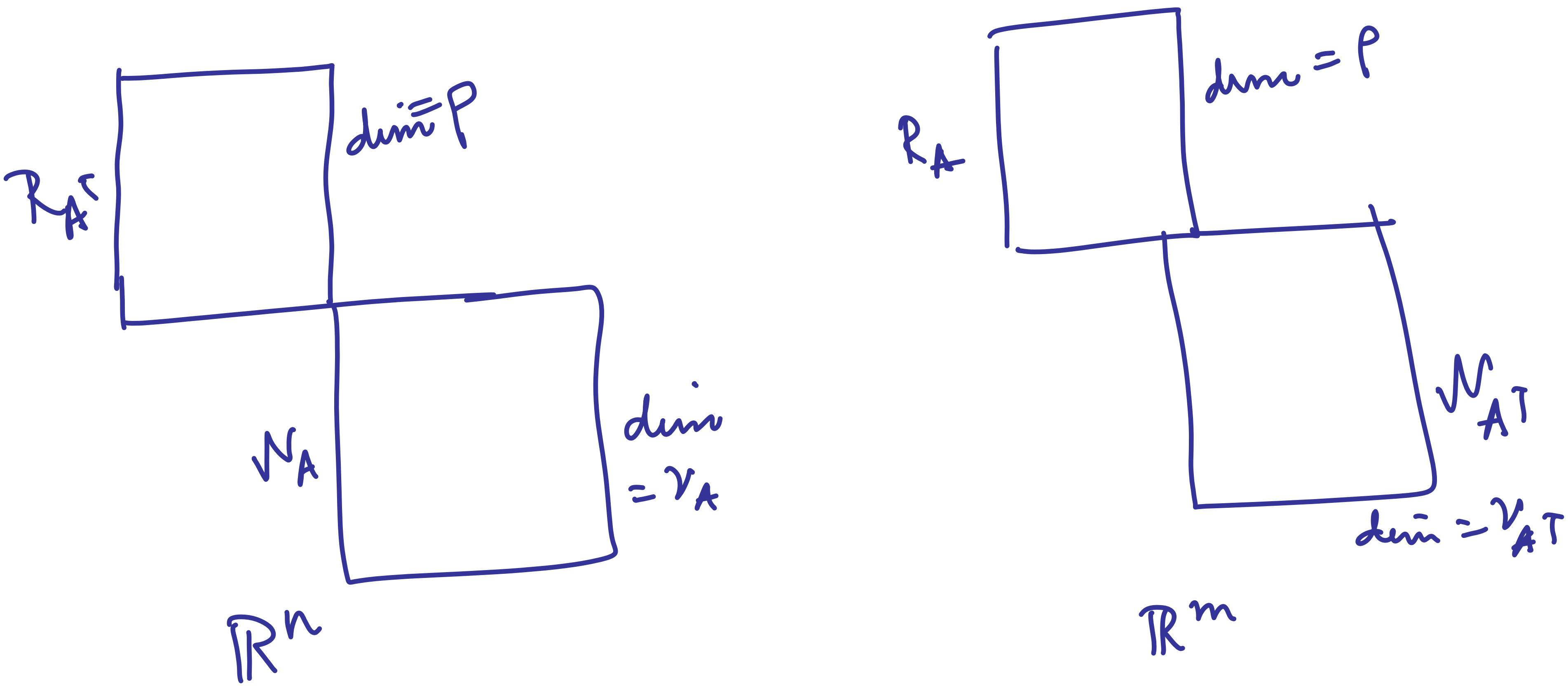
$$\Rightarrow \quad \rho_{A^T} + \nu_A = n$$

$$\rho_A + \nu_A = n \quad \text{(Rank - Nullity)}$$

$$\Rightarrow \quad \rho_A = \rho_{A^T}$$

$\Rightarrow$ Rank of a matrix

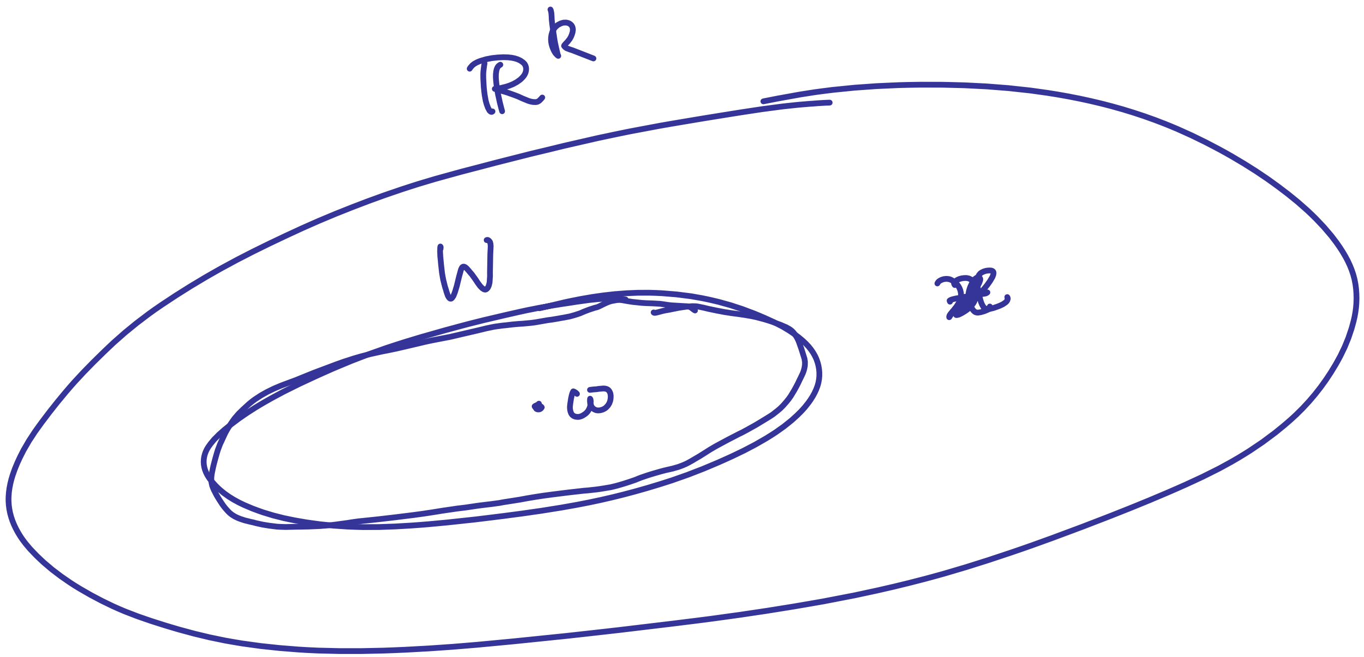
= Rank of its transpose



R_{A^T}
dim = ρ
$\mathcal{N}_A$
dim = ν_A
$\mathbb{R}^n$
R_A
dim = ρ
$\mathcal{N}_{A^T}$
dim = ν_{A^T}
$\mathbb{R}^m$

ρ = Rank of A

Best Approximation from a subspace



W is a subspace of dim d in $\mathbb{R}^k$

$x \in \mathbb{R}^k$

If we take any $w \in W$ and

consider it as an approximation of x

then the square error

$$= \|x - w\|^2$$

We would like to minimize this error

This means : We want a $w_0 \in W$ s.t

$$\|x - w_0\|^2 < \|x - \underline{w}\|^2 \quad \forall\, w \in W,\ w \neq w_0$$

Does there exist such a w_0?

W is of dim d

Let $\varphi_1, \varphi_2, \ldots, \varphi_d$ be an o.n.b. for W

Any $w \in W$ can be expressed as a l.c.

of $\varphi_1, \varphi_2, \ldots, \varphi_d$ as

$$\underline{w = \alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \cdots + \alpha_d \varphi_d}$$

Square Error

$$\underline{\|x - w\|^2} = (x - w, x - w)$$

$$= (x - (\underline{\alpha_1 \varphi_1} + \cdots + \alpha_d \varphi_d), \underline{x} - (\alpha_1 \varphi_1 + \cdots + \alpha_d \varphi_d))$$

$$= (x, x) - \alpha_1 (x, \varphi_1) - \alpha_2 (x, \varphi_2) \cdots - \alpha_k (x, \varphi_k) \cdots - \alpha_d (x, d)$$

$$- \alpha_1 (x, \varphi_1) - \alpha_2 (x, \varphi_2) - \cdots$$

$$+ \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_d^2$$

$$= \|x\|^2 + \sum_{j=1}^{d} \left\{ \alpha_j^2 - 2\alpha_j (x, \varphi_j) \right\}$$

$$= \|x\|^2 + \sum_{j=1}^{d} \left\{ \alpha_j^2 - 2\alpha_j (x, \varphi_j) + (x, \varphi_j)^2 - (x, \varphi_j)^2 \right\}$$

$$= \|x\|^2 + \sum_{j=1}^{d} \left(\alpha_j - (x,\varphi_j)\right)^2 - \sum_{j=1}^{d} (x,\varphi_j)^2$$

For any $w = \underline{\underline{\alpha_1}}\varphi_1 + \cdots + \underline{\underline{\alpha_d}}\varphi_d$ in W

the square error is

$$\|x - w\|^2 = \underline{\|x\|^2} + \underbrace{\sum_{j=1}^{d} \left(\alpha_j - (x,\varphi_j)\right)^2} - \underline{\sum_{j=1}^{d} (x,\varphi_j)^2}$$

The square error is minimum when

$$\underline{\alpha_j = (x,\varphi_j)^2} \quad , \quad 1 \le j \le d$$

Hence if we choose

$$w_0 = (x,\varphi_1)\varphi_1 + (x,\varphi_2)\varphi_2 + \cdots + (x,\varphi_d)\varphi_d$$

then we get

$$\|x - w_0\|^2 = \boxed{\|x\|^2 - \sum_{j=1}^{d} (x, \varphi_j)^2}$$

$$< \underbrace{\|x\|^2 + \sum_{j=1}^{d} \left(\alpha_j - (x, \varphi_j)\right)^2 - \sum_{j=1}^{d} (x, \varphi_j)^2}_{\|x - w\|^2}$$

for any other $w \in W$

Hence $w_0 = \boxed{\sum_{j=1}^{d} (x, \varphi_j) \varphi_j} \in W$

and approximates x from W with least square error.

This is called the best approximation of

x from W

<u>OBSERVATION</u>

1) Note $w_0 = \sum_{j=1}^{d} (x, \varphi_j)\, \varphi_j$

$= x_W$ the orthogonal projection of x onto W

$\therefore$ The best approximation of x from W is the orthogonal projection x_W of x onto W

2) Similarly $x_{W^\perp}$ is the best approximant of x from $W^\perp$

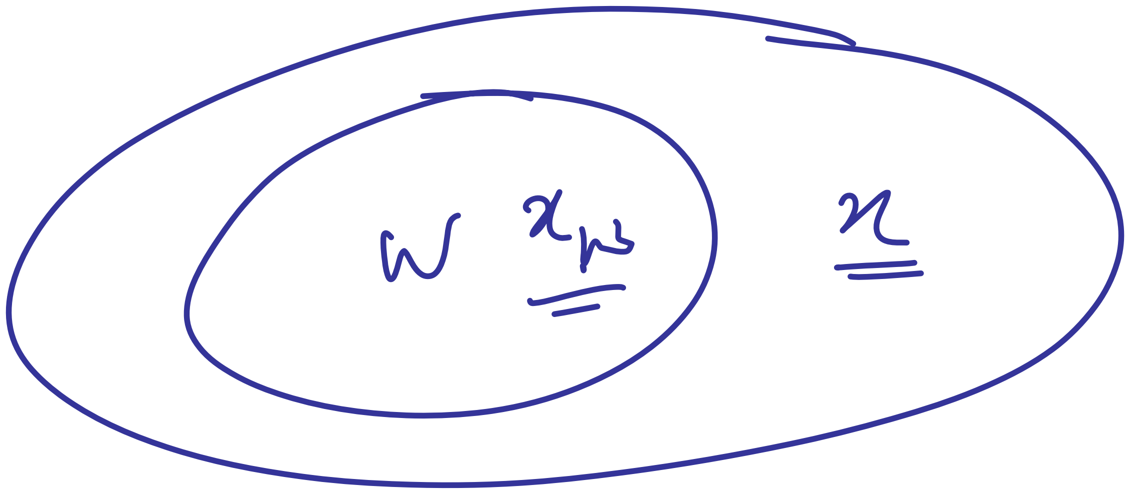
3) The least square error

We have $x = x_W + x_{W^\perp}$

Best approximant from W is x_W.

$\therefore$ least sq. error $= \|x - x_W\|^2$

$= \|x_{W^\perp}\|^2$



$\|(x_{W^\perp})\|^2$

Example:

In $\mathbb{R}^3$ consider $W = \left\{ x = \begin{pmatrix} a \\ a \\ a \end{pmatrix} : a \in \mathbb{R} \right\}$

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ in $\mathbb{R}^3$

Want to find best approximation of x from W

We know this is given by the orthogonal projection of x onto W

We have seen (earlier lecture)

$$x_W = \begin{pmatrix} (x_1+x_2+x_3)/3 \\ (x_1+x_2+x_3)/3 \\ (x_1+x_2+x_3)/3 \end{pmatrix}$$

Let us verify this directly

Take any $w \in W$

Let $w = \begin{pmatrix} t \\ t \\ t \end{pmatrix}$; $t \in R$

Square error

$$= \|x - w\|^2 = \left\| \begin{pmatrix} x_1 - t \\ x_2 - t \\ x_3 - t \end{pmatrix} \right\|^2$$

$$= (x_1-t)^2 + (x_2-t)^2 + (x_3-t)^2$$

We want to find w s.t this square error is minimum i.e. we want to find $t \in \mathbb{R}$ s.t

$$(x_1 - t)^2 + (x_2 - t)^2 + (x_3 - t)^2$$

is minimum.

$$E(t) = (x_1 - t)^2 + (x_2 - t)^2 + (x_3 - t)^2$$

$$\frac{dE}{dt} = -2(x_1 - t) - (2)(x_2 - t) - 2(x_3 - t)$$

$$= \underline{6t - 2(x_1 + x_2 + x_3)}$$

$$\frac{dE}{dt} = 0 \text{ when } t = \frac{x_1 + x_2 + x_3}{3}$$

$$\frac{d^2E}{dt^2} = 6 > 0$$

$$\left.\frac{d^2E}{dt^2}\right|_{t=\frac{x_1+x_2+x_3}{3}} = 6 > 0$$

$\therefore$ $t = \dfrac{x_1+x_2+x_3}{3}$ is the pt. at which $E(t)$ is min.

. The vector $w \in W$ corresp to this t is

$$\begin{pmatrix} (x_1+x_2+x_3)/3 \\ (x_1+x_2+x_3)/3 \\ (x_1+x_2+x_3)/3 \end{pmatrix} = x_W \text{ as found earlier}$$

<u>Note:</u> If $x \in W$ then $x_W = x$ and hence x is itself its best approximant from W

$\mathbb{R}^k$	Matrix Context $A \in \mathbb{R}^{m \times n}$
1. General Space $\mathbb{R}^k$	1. We look at $\mathbb{R}^n$ and $\mathbb{R}^m$
2. Subspace in $\mathbb{R}^k$	2. a) Two subspaces R_{A^T} & N_A in $\mathbb{R}^n$ b) Two subspaces

	R_A & N_{A^T} in R^n
3. Orthogonal complement $W^\perp$ of a subspace W	3. $R_{A^T}^\perp$, $N_A^\perp$ in R^n $R_A^\perp$, $N_{A^T}^\perp$ in R^m $N_A = R_{A^T}^\perp$, $N_{A^T} = R_A^\perp$ Had this
4. $(W^\perp)^\perp = W$	4. $R_{A^T} = N_A^\perp$ $R_A = N_{A^T}^\perp$
5. $\dim W + \dim W^\perp = k$	5 $\rho_{A^T} + \nu_A = n$ Using together with Rank-Nullity Thm gave us $\rho_A = \rho_{A^T}$

6. If $\mathcal{B}_W$ is a basis for W and $\mathcal{B}_{W^\perp}$ " " " $W^\perp$ then $\mathcal{B} = \mathcal{B}_W \cup \mathcal{B}_{W^\perp}$ is a basis for $\mathbb{R}^k$

6 If $\mathcal{B}_{R_{A^T}}$ is a basis for R_{A^T} and $\mathcal{B}_{R_{A^T}^\perp}$ i.e $\mathcal{B}_{N_A}$ is a basis for $R_{A^T}^\perp = N_A$

Then $\mathcal{B}_n = \mathcal{B}_{R_{A^T}} \cup \mathcal{B}_{N_A}$ is a basis for $\mathbb{R}^n$

& similarly

$\mathcal{B}_{R_A} \cup \mathcal{B}_{N_{A^T}} = \mathcal{B}_m$ a basis for $\mathbb{R}^n$

7. $\underline{(\text{ONB for } W)} \cup \underline{(\text{ONB for } W^\perp)} = \underline{\text{ONB for } \mathbb{R}^k}$

a) $(\text{ONB for } R_{A^T}) \cup (\text{ONB for } N_A) = \text{ONB for } \mathbb{R}^n$

b) $(\text{ONB for } R_A) \cup (\text{ONB for } R_{A^T})$
$= \text{ONB for } \mathbb{R}^m$

8 $x \in \mathbb{R}^k \Rightarrow x = x_W + x_{W^\perp}$
$x_W \in W, \; x_{W^\perp} \in W^\perp$
(unique)

8 $x \in \mathbb{R}^n$
$\Rightarrow x = x_{R_{A^T}} + x_{N_A}$
(unique

$b \in \mathbb{R}^m \Rightarrow$
$b = b_{R_A} + b_{N_{A^T}}$
(unique)

9 Best approximation of x from W is x_W

9. a) Best approximation of x from R_{A^T} is $x_{R_{A^T}}$
$\in \mathbb{R}^n$

b) Best app of $x \in \mathbb{R}^m$ from R_A is x_{R_A}

10. Least sq. error

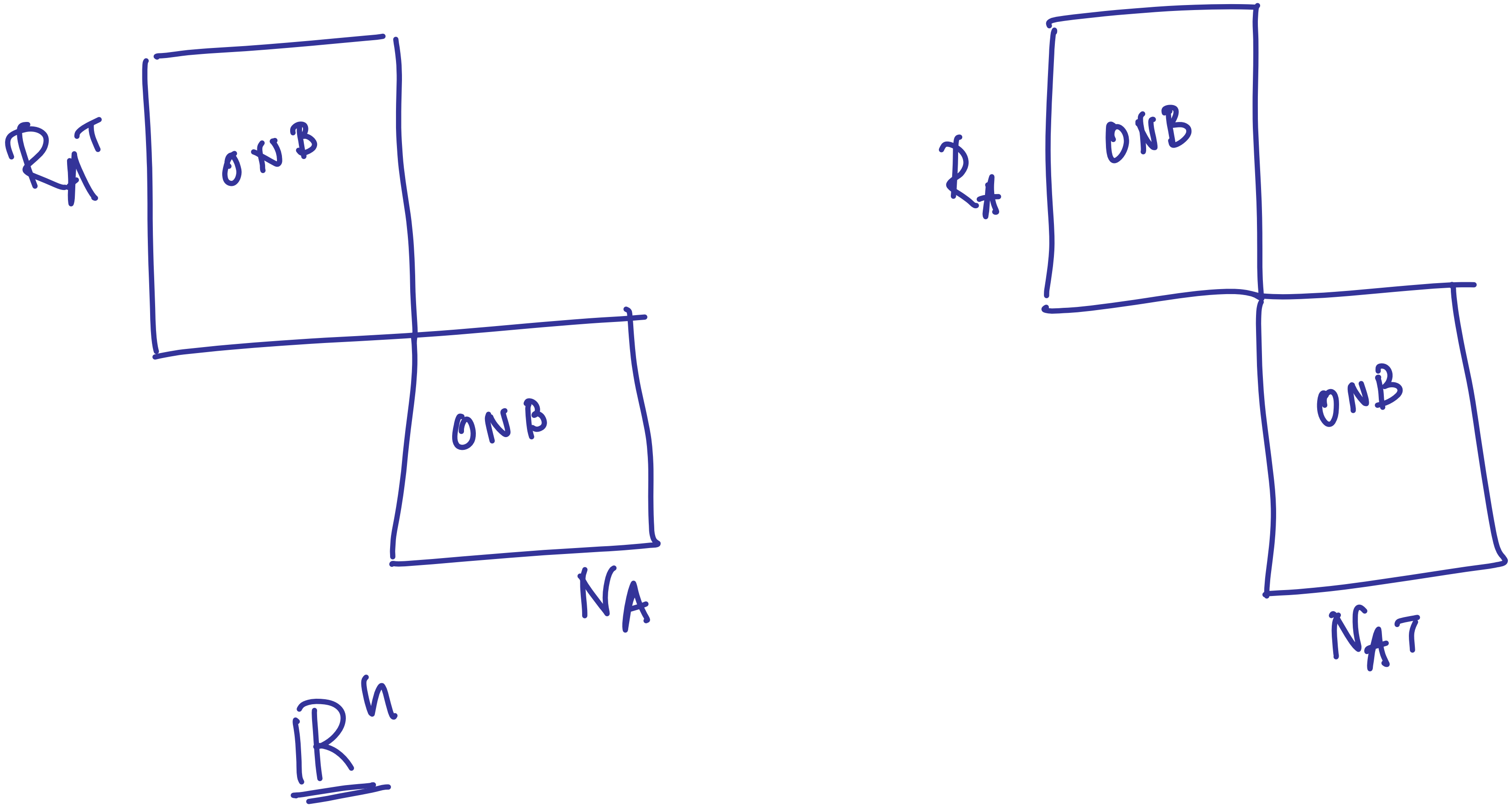
$$\|x_{W^\perp}\|^2$$

10 a) least sq. error

$$\|x_{N_A}\|^2$$

b) least sq error

$$\|x_{N_{A^T}}\|^2$$



The hunt for "suitable" bases for these four subspaces leads us to the notion of <u>eigenvalues</u>

and eigenvectors

Another Problem where eigenvalues and eigenvector notions appear

System

$$Ax = b$$

$A \in \mathbb{R}^{n \times n}$

$b \in \mathbb{R}^{n}$

To find $x \in \mathbb{R}^{n}$

Change of Variables:

$$y = Cx \qquad C \in R^{n\times n} \text{ invertible}$$

$$z = Cb$$

$$x = C^{-1}y = Py$$

$$b = C^{-1}z = Pz$$

System

$Ax = b$ becomes

$$APy = Pz$$

$$\Rightarrow \quad P^{-1}APy = z$$

If we can choose P such that $P^{-1}AP$ is diagonal then we have an easy system for y

Can we find P s.t

$P^{-1}AP$ is diagonal