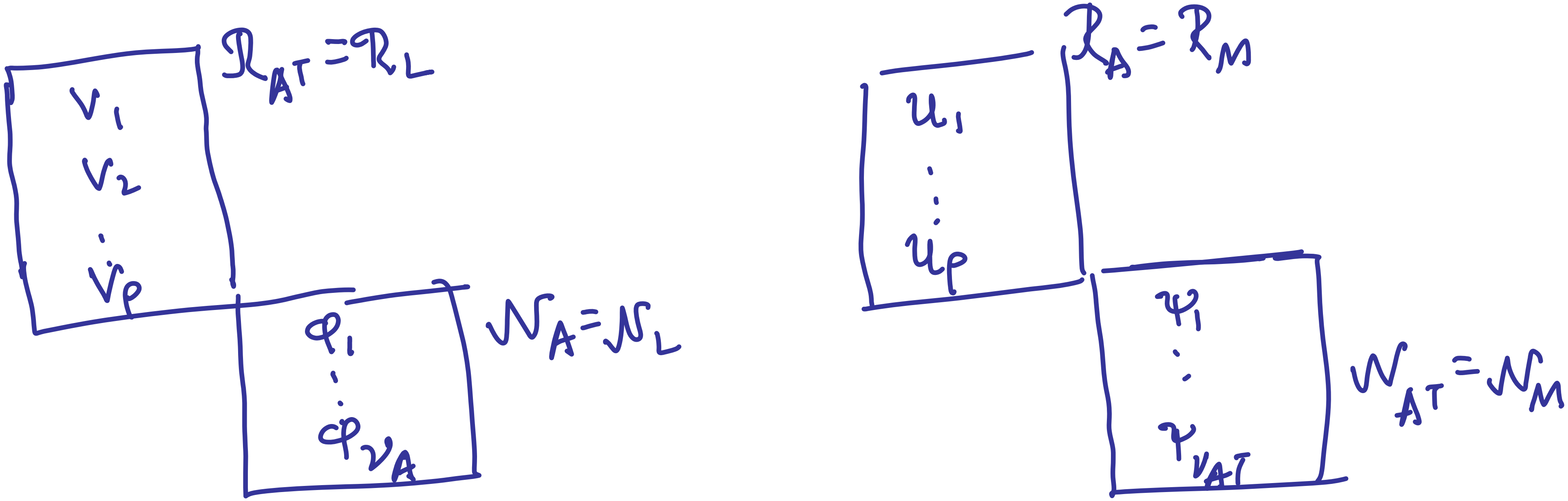
$A \in \mathbb{R}^{m \times n}$ Rank $A = \rho$, Nullity of A is ν_A

$$\left.\begin{aligned} L &= A^T A \in \mathbb{R}^{n \times n} \\ M &= A A^T \in \mathbb{R}^{m \times m} \end{aligned}\right\} \text{PSD}$$



$$\mathbb{R}^n \underset{A^T}{\overset{A}{\rightleftarrows}} \mathbb{R}^m$$

Chose

$v_1, v_2, \ldots, v_p$ as the o.n. eigenvectors corr to the positive eigenvalues of L arranged as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$$

$\varphi_1, \varphi_2, \ldots, \varphi_{\nu_A}$ as the o.n. eigenvectors corr to the 0 eigenvalue of L

$\psi_1, \psi_2, \ldots, \psi_{\nu_{A^T}}$ as the o.n. eigenvectors corr. to the 0 eigenvalue of M

$u_1, u_2, \ldots, u_p$ as

$$\left.\begin{array}{l} u_j = \dfrac{1}{s_j} A v_j \\ \text{where } s_j = \sqrt{\lambda_j} \end{array}\right\} \quad 1 \leq j \leq p$$

(s_j are called singular values of A)

$$A v_j = s_j u_j \qquad 1 \le j \le p$$

$$A^T u_j = s_j v_j \qquad 1 \le j \le p$$

We used these to get

(I) SVD — Product form:

$$A = U S_A V^T$$

where

$$U = [u_1 \cdots u_p \ \psi_1 \cdots \psi_{\nu_{A^T}}] \in \mathbb{R}^{m \times m}$$ orthogonal matrix

$$V = [v_1 \cdots v_p \; \varphi_1 \cdots \varphi_{\nu_A}] \in \mathbb{R}^{n \times n}$$ orthogonal matrix

$$S_A = \left(\begin{array}{c|c} \begin{matrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_p \end{matrix} & 0 \\ \hline 0 & 0 \end{array} \right)_{m \times n}$$

II SVD – Sum form

$$A = \sum_{j=1}^{p} s_j \, v_j \otimes u_j$$

where

$$v_j \otimes u_j = \underset{m\times n}{u_j} \underset{1\times n}{v_j^T} \in \mathbb{R}^{m\times n}$$

Use of these four o.n. bases to analyse the system

$$Ax = b \quad \text{---} \quad (\text{I})$$

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

<u>Given $b \in \mathbb{R}^m$</u> <u>To find</u> <u>$x \in \mathbb{R}^n$</u>

s.t. $Ax = b$ i.e. (I) is satisfied

The given $b \in \mathbb{R}^m$

$u_1, \dots, u_p, \psi_1, \dots, \psi_{\nu_{A^T}}$ is an onb for $\mathbb{R}^m$

We can expand b in terms of this o.n. basis as

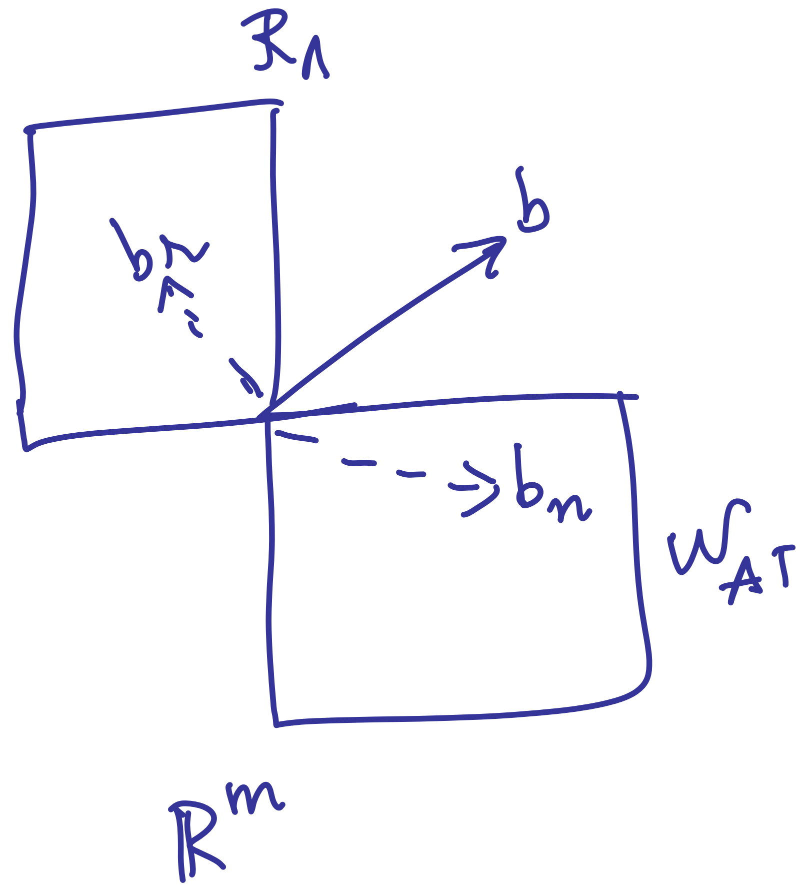
$$b = \underbrace{\sum_{j=1}^{p} (b, u_j) u_j}_{b_r} + \underbrace{\sum_{r=1}^{\nu_{A^T}} (b, \psi_r) \psi_r}_{b_n}$$

$b_r \in \mathcal{R}_A$, orthogonal proj of b onto $\mathcal{R}_A$

$b_n \in \mathcal{N}_{A^T}$ orthogonal projection of b onto $\mathcal{N}_{A^T}$

Pythagoras Theorem

$$\|b\|^2 = \|b_r\|^2 + \|b_n\|^2$$



$$b = \sum_{j=1}^{p} (b, u_j)\, u_j + \sum_{r=1}^{\nu_{A^T}} (b, \psi_r)\, \psi_r \quad \text{---} \quad (\#)$$

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

For $\mathbb{R}^n$ we have onb

$$v_1, v_2, \dots, v_p, \varphi_1, \dots, \varphi_{\nu_A}$$

The x we are looking for must be a l.c. of these basis vectors.

Hence x must be of the form

$$\boxed{x = \sum_{j=1}^{p} \alpha_j v_j + \sum_{r=1}^{\nu_A} \beta_r \varphi_r}$$

The vector x is known the moment we know $\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_{\nu_A}$

We have to find these scalars in such a way that $Ax = b$

Ax is of the form

$$Ax = A\left(\sum_{j=1}^{\rho} \alpha_j v_j + \sum_{r=1}^{\nu_r} \beta_r \varphi_r\right)$$

$$= \sum_{j=1}^{\rho} \alpha_j A v_j + \underline{\sum_{r=1}^{\nu_r} \beta_r A \varphi_r}$$

$$= \sum_{j=1}^{\rho} \alpha_j s_j u_j$$

We have

$$b = \sum_{j=1}^{p} (b, u_j)\, u_j + \sum_{r=1}^{\nu_{A^T}} (b, \psi_j)\, \psi_j .$$

and

$$Ax = \sum_{j=1}^{p} \alpha_j s_j u_j$$

To find x s.t

$$Ax = b$$

We need

$$\sum_{j=1}^{p} \alpha_j s_j u_j = \sum_{j=1}^{p} (b, u_j)\, u_j + \sum_{r=1}^{\nu_{A^T}} (b, \psi_r)\, \psi_r$$

$$\Rightarrow \underbrace{\sum_{j=1}^{p}\left\{\alpha_j s_j - (b,u_j)\right\}u_j}_{\in \mathcal{R}_A \ \because\ u_1,\dots,u_p \text{ is an o.n.basis for } \mathcal{R}_A} = \underbrace{\sum_{r=1}^{\nu_{A^T}}(b,\psi_r)\psi_r}_{\in \mathcal{N}_{A^T} \ \because\ \psi_1,\dots,\psi_{\nu_{A^T}} \text{ is an o.n.basis for } \mathcal{N}_{A^T}}$$

$$\Rightarrow \begin{cases} \displaystyle\sum_{r=1}^{\nu_{A^T}}(b,\psi_r)\psi_r = \theta_m \quad \text{and} \quad \text{--}(1) \\ \displaystyle\sum_{j=1}^{p}\left(\alpha_j s_j - (b,u_j)\right)u_j = \theta_m \quad \text{--}(2) \end{cases}$$

First of all (1) tells that we can get $Ax = b$ only if b satisfies

$$\sum_{r=1}^{\nu_{A^T}} (b, \psi_r)\, \psi_r = \theta_m$$

$$\iff (b, \psi_r) = 0, \quad 1 \le r \le \nu_{A^T}$$

(only if $b \in \mathcal{R}_A$)

[C] (Consistency Condition)

i) $b \in \mathcal{R}_A$ or same as

ii) $b \perp$ all the vectors in $\mathcal{N}_{A^T}$ or same as

iii) $b \perp$ any basis in N_{A^T} or same as

(iv) $(b, \psi_r) = 0, \quad 1 \le r \le \nu_{A^T}$

SUPPOSE b satisfies the above consistency condition

Then (2) gives us

$$\sum_{j=1}^{p} \left(\alpha_j s_j - (b, u_j) \right) u_j = \theta_m$$

$u_1, \ldots, u_p$ are o.n & hence l.i.

$$\Rightarrow \quad \left(\alpha_j s_j - (b, u_j) \right) = 0 \qquad 1 \leq j \leq p$$

$$\Rightarrow \quad \alpha_j = \frac{(b, u_j)}{s_j} \qquad 1 \leq j \leq p$$

& β_r can be chosen arbetrarils

Therefore the sol. x must be of the form

$$\boxed{x = \sum_{j=1}^{p} \frac{(b, u_j)}{s_j} v_j + \sum_{r=1}^{\nu_A} \beta_r \varphi_r}$$

where $\beta_1, \ldots, \beta_{\nu_A}$ can be chosen arbitrarily in $\mathbb{R}$

Therefore summarizing we get

When $b \in \mathbb{R}^m$ satisfies

$$(b, \psi_r) = 0 \ , \ 1 \le r \le \nu_{A^T} \quad -- (C)$$

then

$$Ax = b$$

has a sol and any solution must be of the form

$$x = \sum_{j=1}^{\rho} \frac{(b, u_j)}{s_j} v_j + \sum_{r=1}^{\nu_A} \beta_r \varphi_r .$$

where $\beta_r \in \mathbb{R}$, $1 \le r \le \nu_A$ can be chosen arbitrarily

The arbitrariness in the sol. is due to the $\beta_1, \ldots, \beta_{\nu_A}$

Hence if $\nu_A = 0$ i.e. if $\rho = n$

then there is no arbitrariness

and we have UNIQUE sol given by

$$\boxed{x = \sum_{j=1}^{\rho (=n)} \frac{(b, u_j)}{s_j} v_j}$$

If $\nu_A \neq 0$ then $\rho < n$ and we have INFINITE # of sol. :- $\beta_1, \dots, \beta_{\nu_A}$ can be chosen arbitrarily

All these inf. sol are of the form

$$x = \underbrace{\sum_{j=1}^{\rho} \frac{(b, u_j)}{s_j} v_j}_{\in \mathcal{R}_{A^T}} + \underbrace{\sum_{r=1}^{\nu_A} \beta_r \varphi_r}_{\in \mathcal{N}_A}$$

$x_r \qquad x_n$

$$x_r = \sum_{j=1}^{p} \frac{(b, u_j)}{s_j} v_j \in R_{A^T}$$

$$x_n = \sum_{r=1}^{\nu_A} \beta_r \phi_r \in N_A$$

Pythagoras theorem

$$\|x\|^2 = \|x_r\|^2 + \|x_n\|^2$$

Hence the sol. with the least length is obtained when $x_n = \theta_n$ i.e. when

$$\beta_1 = \cdots = \beta_{\nu_A} = 0$$

This sol is called the OPTIMAL sol.

$$x_{OPT} = \sum_{j=1}^{p} \frac{(b, u_j)}{s_j} v_j$$

<u>Summarize</u> : $A \in \mathbb{R}^{m \times n}$

If $b \in \mathbb{R}^m$ satisfies

$$(b, \psi_r) = 0 \quad , \quad 1 \leq r \leq \nu_{A^T} \quad \text{-- -- } (G)$$

then the system

$$Ax = b$$

HAS A SOLUTION

i) The solution is unique if

$$\rho = n$$

UNIQUE

And the unique sol is given by

$$x = \sum_{j=1}^{\rho (=n)} \frac{(b, u_j)}{s_j} v_j$$

ii) If $\rho < n$ then there are infinite number of solutions

Inf sol

And they are all of the form

$$\boxed{x = \sum_{j=1}^{\rho} \frac{(b, u_j)}{s_j} v_j + \sum_{r=1}^{V_A} \beta_r \varphi_r}$$

where $\beta_r \in \mathbb{R}$, $1 \le r \le \nu_A$.

Among these solutions the solution having least length is called the OPTIMAL sol & is denoted by x_{OPT} and we have

$$\boxed{x_{OPT} = \sum_{j=1}^{\rho\,(<n)} \frac{(b, u_j)}{s_j} v_j}$$

Thus we have a complete analysis of the system

$$Ax = b$$

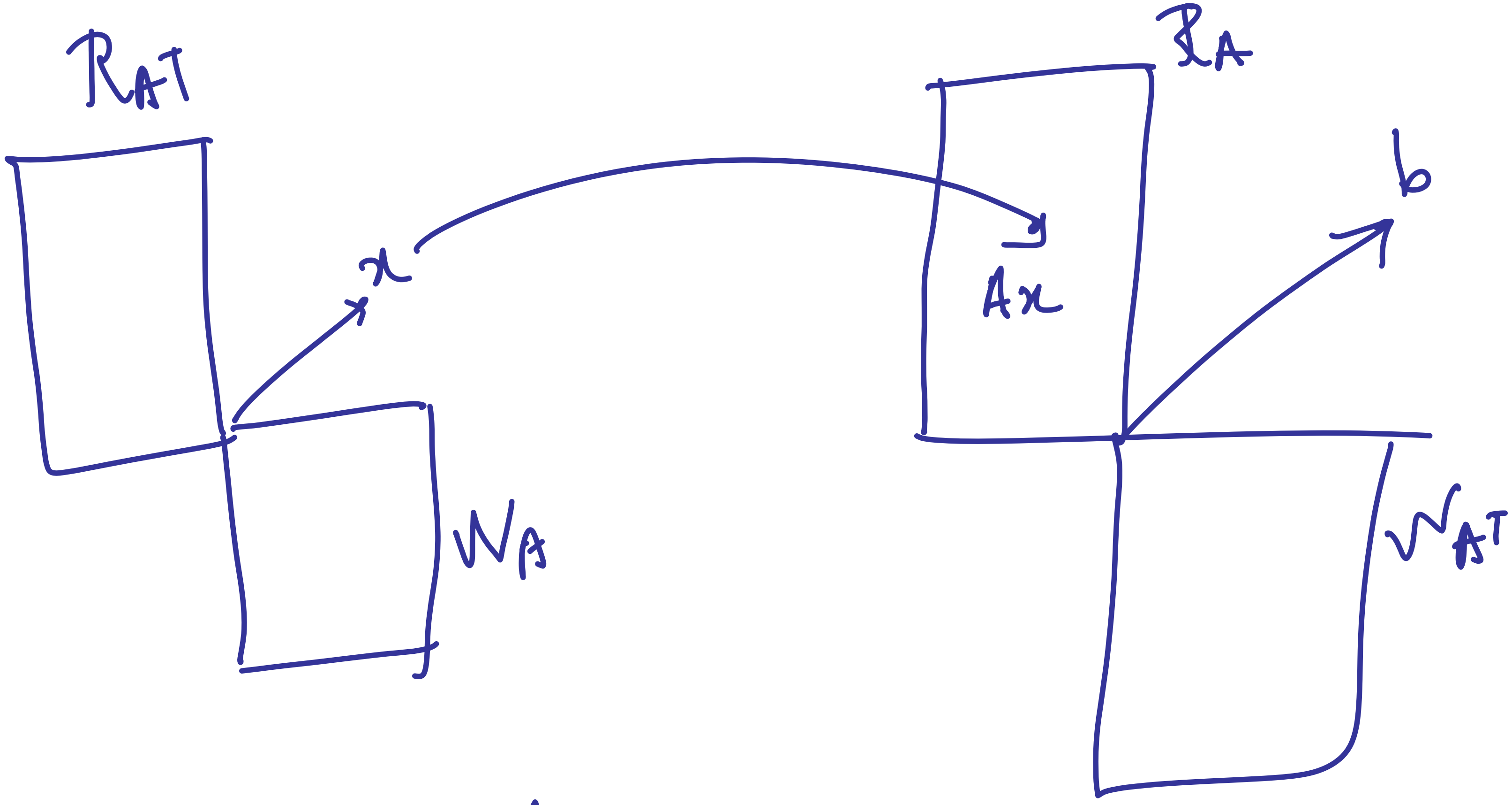
in terms of the bases we have chosen in the case when b satisfies the consistency condition

We next look at the case when $b \in \mathbb{R}^m$ <u>DOES NOT</u> satisfy the consistency condition.

What does it mean to say that b does not satisfy consistency condn.

b satisfies (C) means $b \in \mathcal{R}_A$

b does not satisfy (C) means $b \notin \mathcal{R}_A$



$$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$$

$b \notin \mathcal{R}_A$ means

for any $x \in \mathbb{R}^n$ we cannot get $Ax = b$

Therefore

$$b - Ax \neq \theta_m \text{ for any } x \in \mathbb{R}^n$$

This means

$$\|b - Ax\|^2 > 0$$

$$e_b(x) = \|b - Ax\|^2$$

square error of taking x as a possible sol for $Ax = b$.

We would like this error to be minimum.

If $x_l \in \mathbb{R}^n$ is s.t. this error is min we call it Least Square sol for $Ax = b$

<u>Def</u> $x_l \in \mathbb{R}^n$ is called Least Square sol for $Ax = b$ if

$$\|b - Ax_l\|^2 \leq \|b - Ax\|^2 \quad \forall x \in \mathbb{R}^n$$

i.e. $\quad e_b(x_l) \leq e_b(x) \quad \forall x \in \mathbb{R}^n$