Properties of PSD matrix

1) all eigenvalues are real
2) all eigenvalues are ≥ 0
3) If ν_A is the nullity of A & ρ is the rank of A (A is PSD) then its eigenvalues can be arranged as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\rho > 0 = \underbrace{\lambda_{\rho+1} = \cdots = \lambda_n}_{\nu_A \text{ of them}}$$

4) Corresponding to these we have o.n. eigenvectors

$v_1, v_2, \dots, v_p$ corresp. to $\lambda_1, \dots, \lambda_p$
(pos. eigenvalues)

$\varphi_1, \varphi_2, \dots, \varphi_{\nu_A}$ corresp to the eigenvalue 0

5) $v_1, v_2, \dots, v_p$ Provide an o.n-b
for the range of A

Analysis of a general $m \times n$
matrix

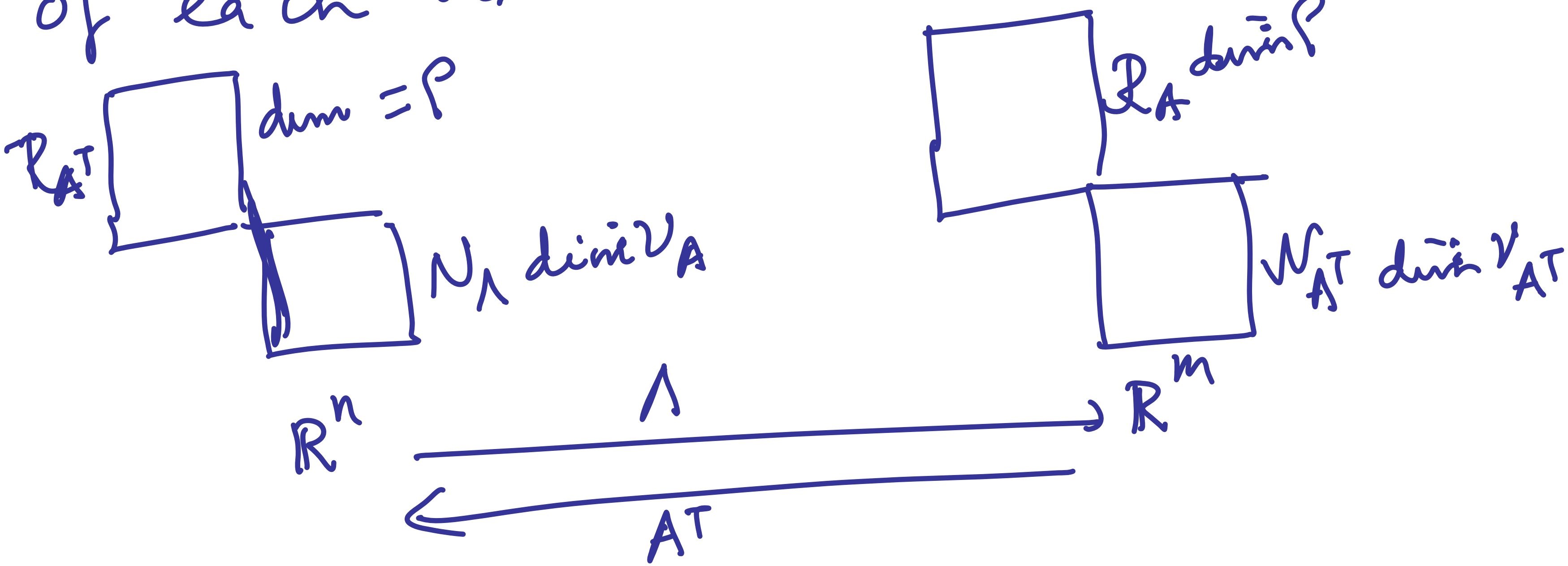
W.l.g. we shall look at $A \in \mathbb{R}^{m \times n}$

Recall that corresp. to A we have four subspaces,

R_{A^T}, N_A : subspaces of $\mathbb{R}^n$

R_A, N_{A^T} : subspaces of $\mathbb{R}^m$

These pairs are orthogonal complements of each other



The Basic Problem:

To choose o.n bases for these four subspaces in a "suitable" manner which makes the analysis of A "easy"

It is for this purpose (of suitable bases choosing) we use the ideas of PSD matrices.

How do we do this?

Given $A \in \mathbb{R}^{m \times n}$ we construct a PSD matrix $L \in \mathbb{R}^{n \times n}$ s.t the analysis of L reflects in the analysis of A

Define

$$\underset{n\times m}{A^T}\ \underset{m \times n}{A}$$

$$L = A^T A \in \mathbb{R}^{n \times n}$$

1) $L \in \mathbb{R}^{n \times n}$, $L^T = L$, $\therefore$ L is real symm.

2) $x \in \mathbb{R}^n \implies (Lx, x) = (A^T A x, x)$

$$= (Ax, Ax)$$

$$= \|Ax\|^2 \geq 0$$

$\Rightarrow (Lx, x) \geq 0 \quad \forall x \in \mathbb{R}^n$

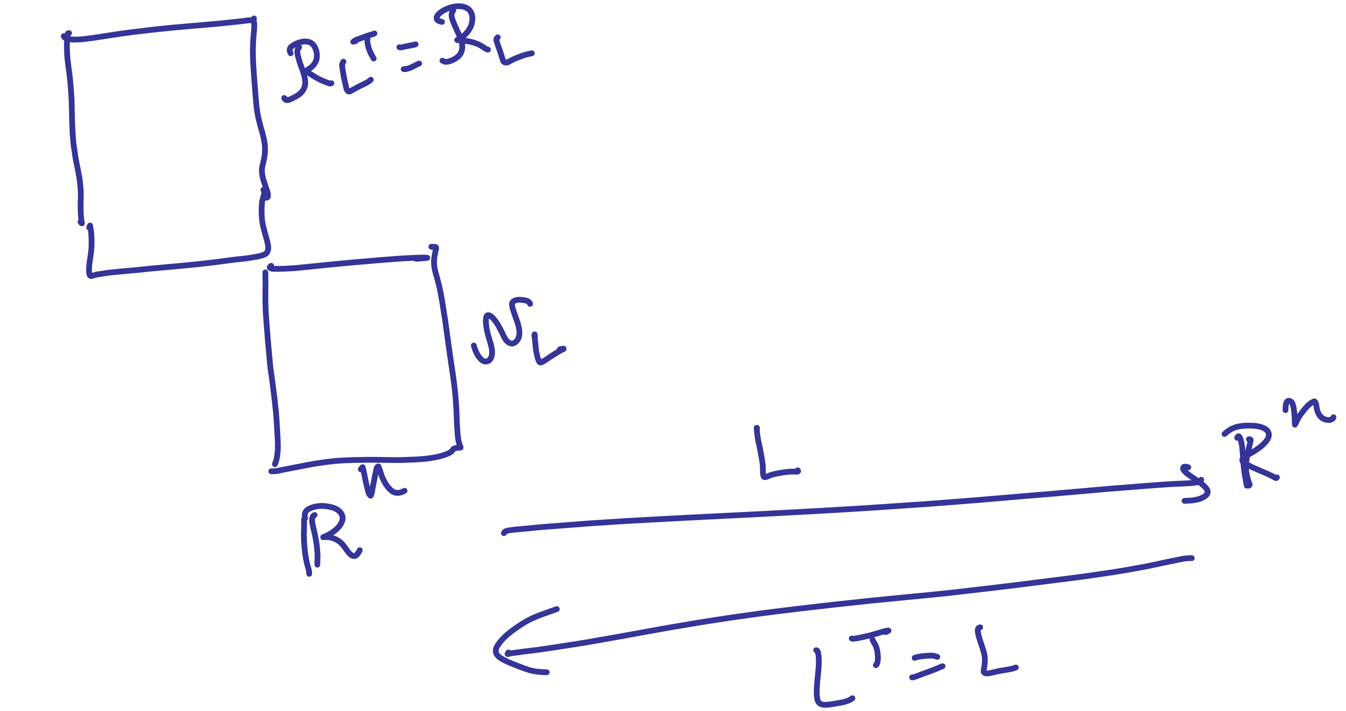
$\Rightarrow$ L is a Positive semi definite matrix $\in \mathbb{R}^{n \times n}$

3) Analogously we can define

$$M = AA^T$$

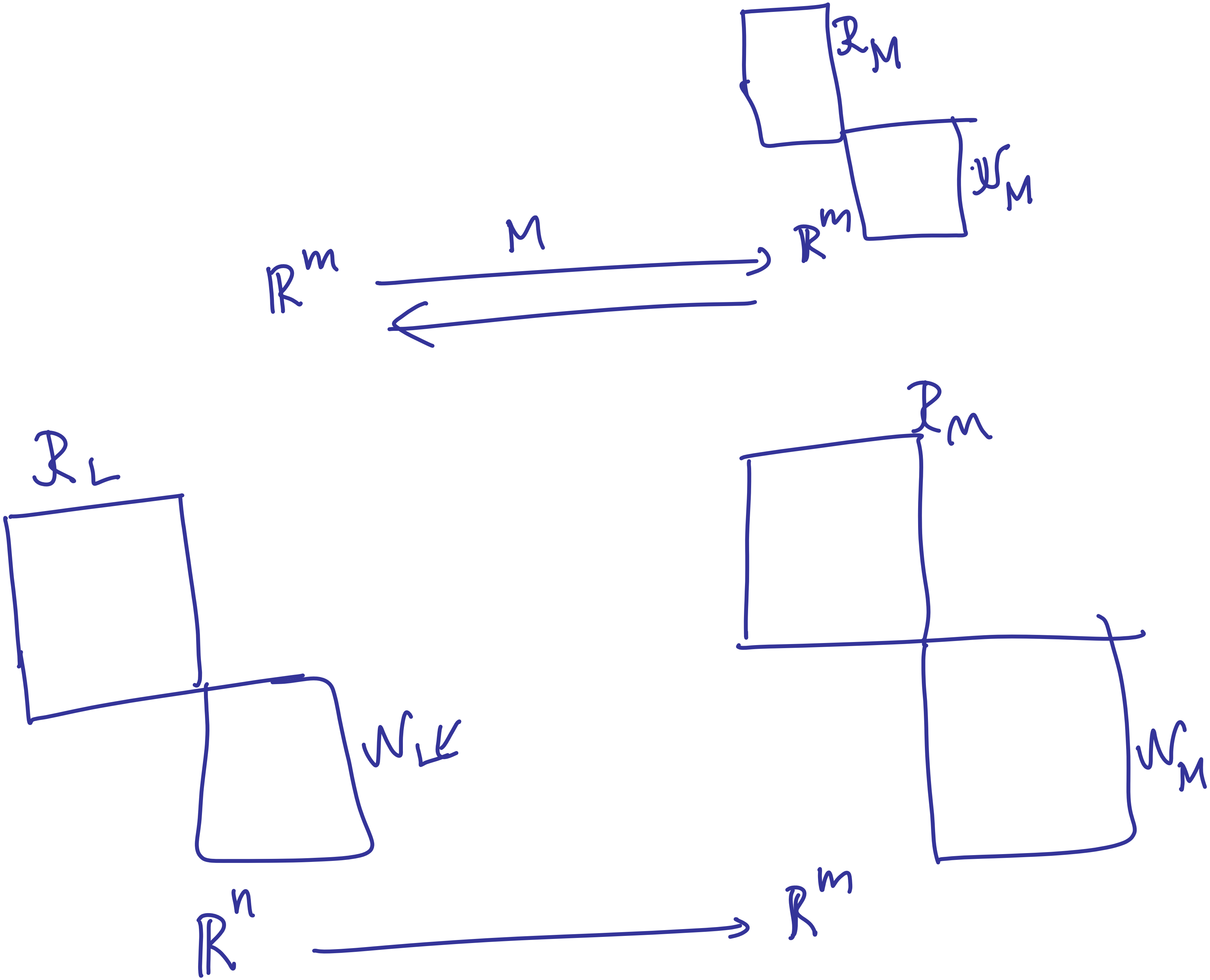
This is a PSD matrix $\in \mathbb{R}^{m \times m}$

Given $L \in \mathbb{R}^{n \times n}$

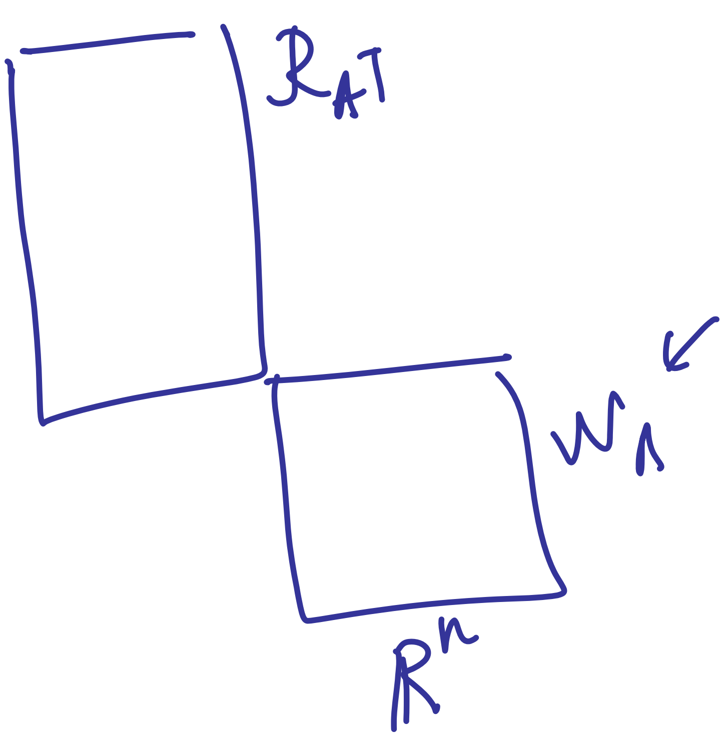


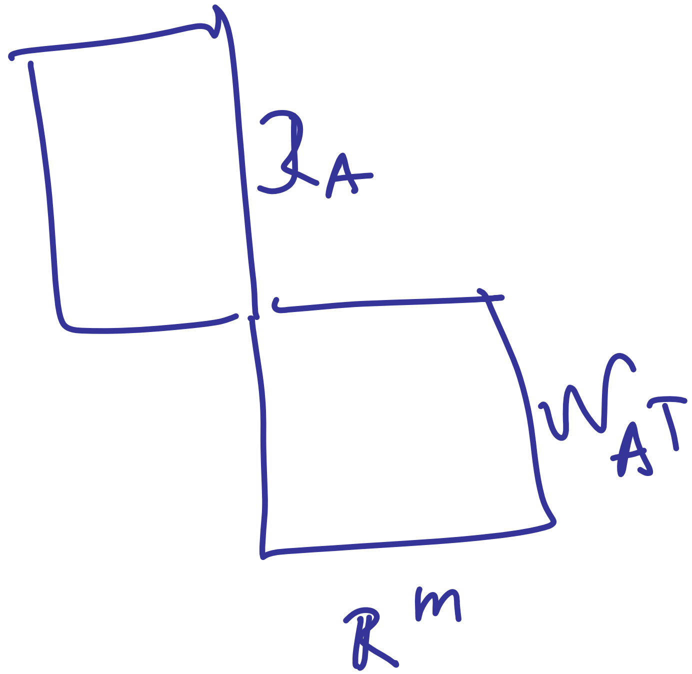
Subspace of L

$$\begin{array}{ccc} R_L & , & N_L \\ \| & & \| \\ R_{L^T} & & N_{L^T} \end{array}$$



$\mathcal{R}_M$
$\mathcal{W}_M$
$\mathbb{R}^m$
M
$\mathbb{R}^m$
$\mathcal{R}_L$
$\mathcal{W}_L$
$\mathcal{R}_M$
$\mathcal{W}_M$
$\mathbb{R}^n$
$\mathbb{R}^m$





For $\mathbb{R}^n$ we have two orthogonal complement decomposition

1) R_L and N_L given by L
2) R_{A^T} and N_A given by A

Is there any connection between these two decompositions?

To look at this we shall first look at some properties of $\mathcal{N}_L$ & $\mathcal{R}_L$

$$\text{(I)}\quad \left.\begin{array}{l} x \in \mathbb{R}^n \\ x \in \mathcal{N}_L \end{array}\right\} \Longrightarrow \underset{\in \mathbb{R}^n}{Lx} = \theta_n$$

$$\Longrightarrow (Lx, x) = (\theta_n, x) = 0$$

$$\Longrightarrow (A^T A x, x) = 0 \qquad (\because L = A^T A)$$

$$\Longrightarrow (Ax, Ax) = 0$$

$$\Longrightarrow \|Ax\|^2 = 0$$

$$\Longrightarrow Ax = \theta_m$$

$$\Rightarrow x \in N_A$$

$$\therefore \boxed{N_L \subseteq N_A} \quad --- \quad (1)$$

On the other hand

$$x \in N_A \Rightarrow Ax = \theta_m$$

$$\Rightarrow \underset{n\times m}{A^T} A x = \underset{m\times n}{A^T \theta_m} = \theta_n$$

$$\Rightarrow Lx = \theta_n$$

$$\Rightarrow x \in N_L$$

$$\Rightarrow \boxed{N_A \subseteq N_L} \quad --- \quad (2)$$

By (1) & (2) we get

$$N_L = N_A$$

(II) Consequently we have

$$R_A = R_{L^T} = N_{L^T}^{\perp} = N_L^{\perp}$$

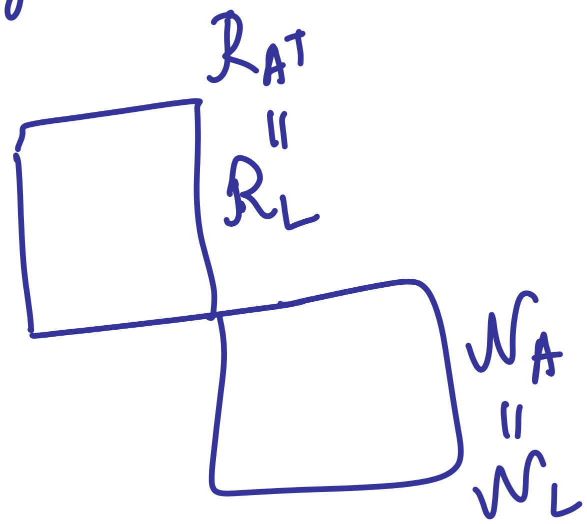
$\because L^T = L$ $\qquad$ $\because L^T = L$

$$= N_A^{\perp} \quad (\because N_L = N_A)$$

$$= R_{A^T}$$

Hence we get $R_L = R_{A^T}$

Hence the decomp of $\mathbb{R}^n$ into orthogonal complements given by L & A are the same



$\mathbb{R}^n$

We get $\dim N_L = \dim N_A$

$\Rightarrow \quad \nu_L = \nu_A$

$$\dim \mathcal{R}_{A^T} = \dim \mathcal{R}_L$$

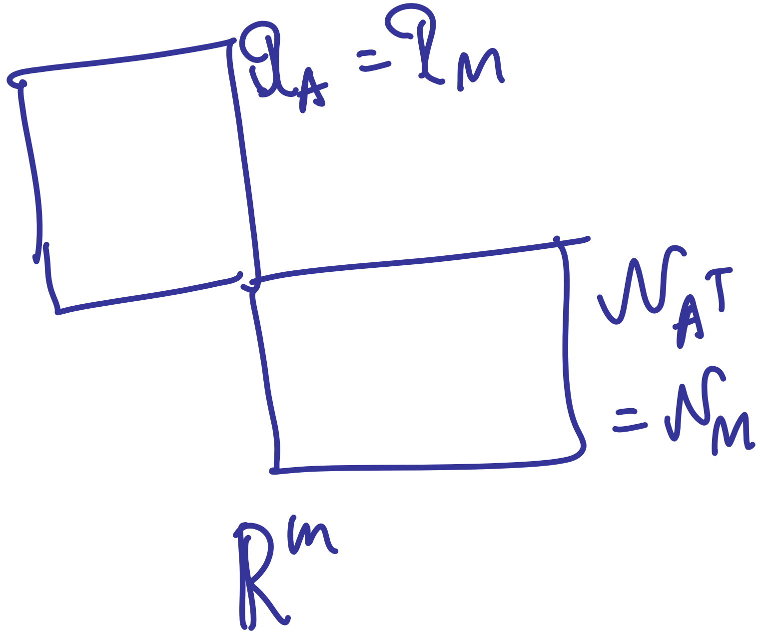
$$\Rightarrow \text{Rank } A^T = \text{Rank } L$$

$$\Rightarrow \rho_{A^T} = \rho_L$$

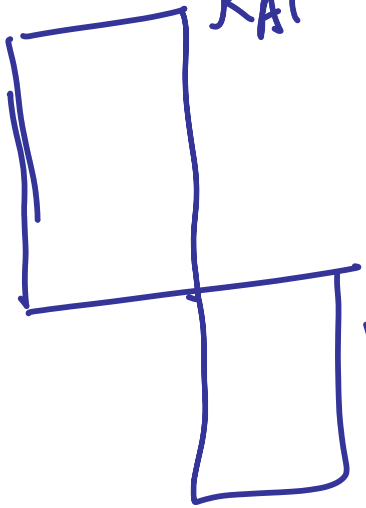
$$\Rightarrow \rho_A = \rho_L$$

$$\left\| \begin{array}{l} \nu_A = \nu_L \\ \rho_A = \rho_L \quad (= \rho_{A^T}) \end{array} \right\|$$

Analogously on the decomp of $\mathbb{R}^m$ we have,

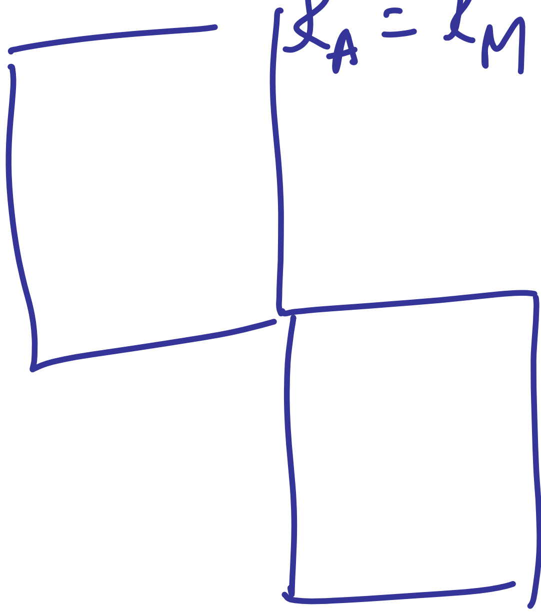


We have



$\mathcal{R}_{A^T} = \mathcal{R}_L$

$\mathcal{N}_A = \mathcal{N}_L$



$\mathcal{R}_A = \mathcal{R}_M$

$\mathcal{N}_{A^T} = \mathcal{N}_M$

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

$$\dim \mathcal{N}_{A^T} = \dim \mathcal{N}_M$$

$$\nu_{A^T} = \nu_M$$

$$\dim R_A = \dim R_M$$

$$\rho_A = \rho_M$$

$$\|$$

$$\rho_L = \rho_{A^T}$$

$$\boxed{\begin{array}{l} \nu_A = \nu_L \quad , \quad \nu_{A^T} = \nu_M \\ \rho_A = \rho_{A^T} = \rho_L = \rho_M \end{array}}$$

How do we use then to find o.n bases for these four subspaces

1. $\mathcal{N}_A$ Same as $\mathcal{N}_L$

But L is PSD

$\therefore$ $\mathcal{N}_L$ basis (o.n) is given by the o.n. eigenvectors

$$\varphi_1, \varphi_2, \cdots, \varphi_{\nu_A}$$

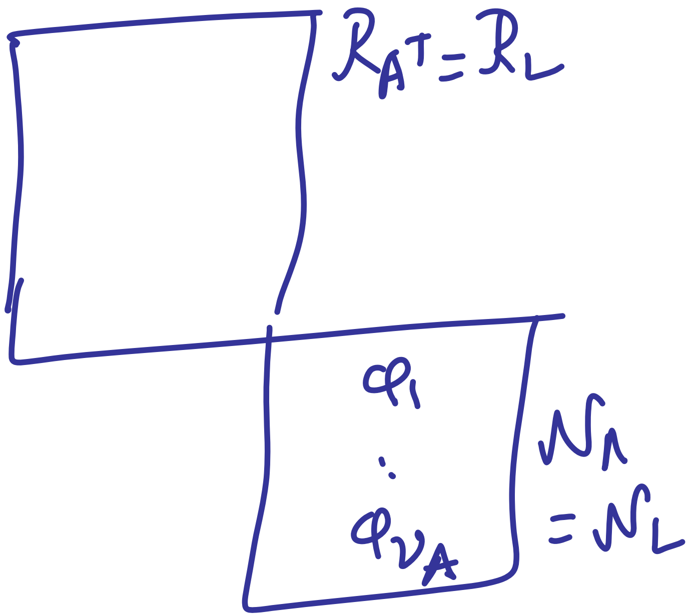
corresp to eigenvalue 0 of the matrix L

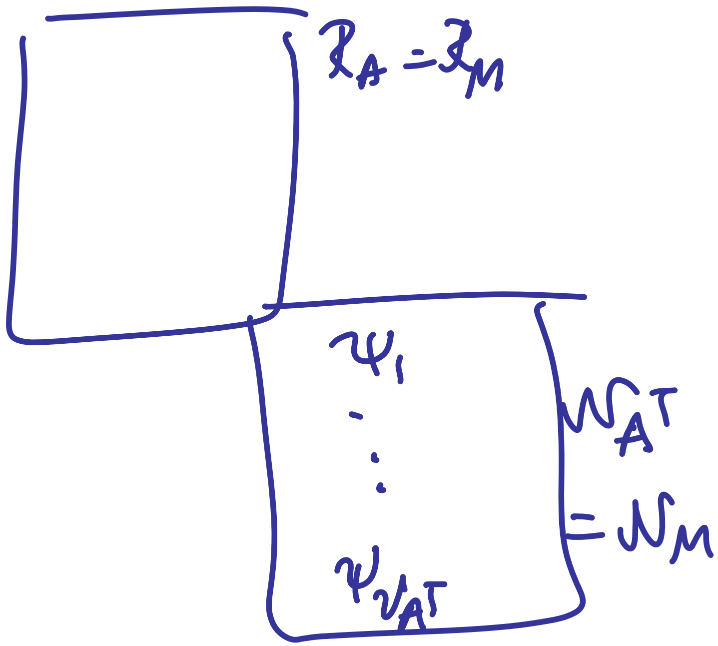
(Note $\nu_A = \nu_L$)

2. $\mathcal{N}_{A^T}$: o.n. basis corresponds to the o.n. eigenvectors corresp to the

eigenvalue 0 of the matrix M

$\psi_1, \psi_2, \ldots, \psi_{\nu_{A^T}}$ $\quad (\because \nu_{A^T} = \nu_M)$





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

3) $\underline{\mathcal{R}_{A^T} \,(= \mathcal{R}_L)}$

L is PSD. Its positive eigenvalues can be arranged as

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

We get o.n. eigenvectors

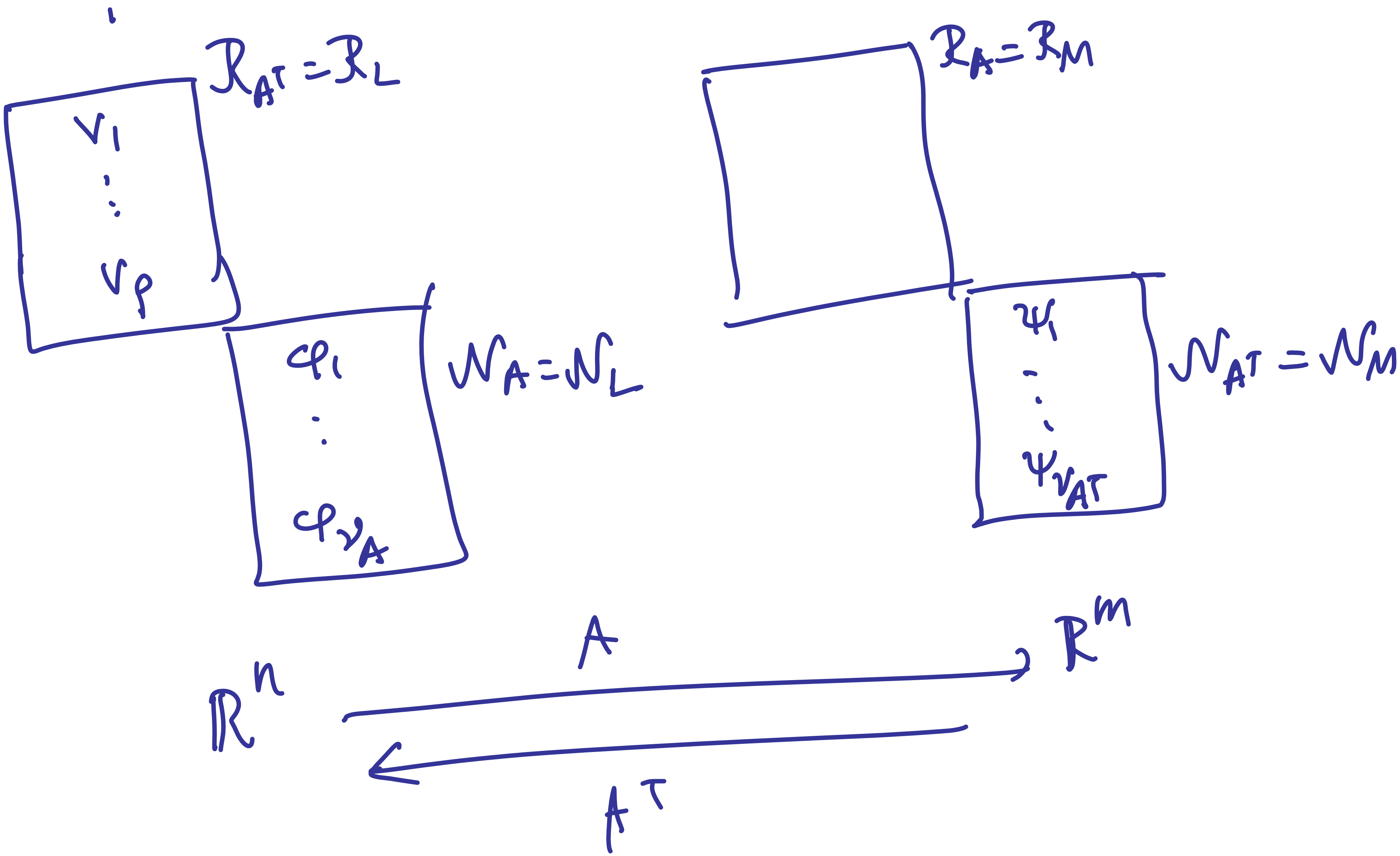
$$v_1, v_2, \ldots, v_p$$

We found that

$$v_1, \ldots, v_p$$

provide an o.n.b. for $\mathcal{R}_L$ ($\because L$ is PSD)

& an o.n.b for $\mathcal{R}_{A^T}$


$\mathcal{R}_{A^T} = \mathcal{R}_L$
v_1
$\vdots$
v_ρ
$\mathcal{N}_A = \mathcal{N}_L$
φ_1
$\vdots$
φ_{ν_A}
$\mathcal{R}_A = \mathcal{R}_M$
$\mathcal{N}_{A^T} = \mathcal{N}_M$
ψ_1
$\vdots$
$\psi_{\nu_{A^T}}$
$\mathbb{R}^n$
A
$\mathbb{R}^m$
A^T

To Choose an onb for $R_A (= R_M)$

We have

$v_1, v_2, \dots, v_p$ o.n. basis for $R_{A^T} (= R_L)$

& there are o.n. eigenvectors corresponding to the positive eigenvalues

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ of L

Hence $Lv_j = \lambda_j v_j$, $j = 1, 2, \dots, p$ -- (1)

We look at

$$\underset{m \times n}{A} \; \underset{n \times 1}{v_j} \in R^m$$

$$A v_j \in R_A \quad , \quad j = 1, 2, \cdots, \rho$$

Define

$$\longrightarrow \quad w_j = A v_j \longleftarrow, \quad j = 1, 2, \cdots, \rho$$

Then $\underline{\underline{w_j \in R_A}}$

Will w_j form a basis for R_A?

" " " an onb for R_A?

<u>We look at then w_j</u>

We have

$$(w_j, w_r) = (A v_j, A v_r)$$

$$= (v_j, A^T A v_r)$$

$$= (v_j, L v_r)$$

$$= (v_j, \lambda_r v_r)$$

$$= \lambda_r (v_j, v_r)$$

$$= \begin{cases} 0 & \text{if } j \neq r \\ \lambda_j & \text{if } j = r \end{cases}$$

What this says is that

$w_1, w_2, \ldots, w_p$ are orthogonal to each other

and $\|w_j\|^2 = \lambda_j \implies \|w_j\| = \sqrt{\lambda_j}$

Define
$$u_j = \frac{w_j}{\|w_j\|} = \frac{w_j}{\sqrt{\lambda_j}}$$ when we take positive sq. root of λ_j

We denote by
$$s_j = \sqrt{\lambda_j}$$ Singular Values of A

$$u_j = \frac{w_j}{s_j}$$

$u_1, u_2, \ldots, u_p$ are o.n. vectors in R_A

But R_A has dimension p

$\Rightarrow$ $u_1, u_2, \ldots, u_p$ form an o.n. basis for R_A.

Note:

$$u_j = \frac{w_j}{s_j} = \frac{Av_j}{s_j} \longleftarrow$$

$$\Longrightarrow \boxed{Av_j = s_j u_j} \qquad j = 1, 2, \ldots, p$$

We have

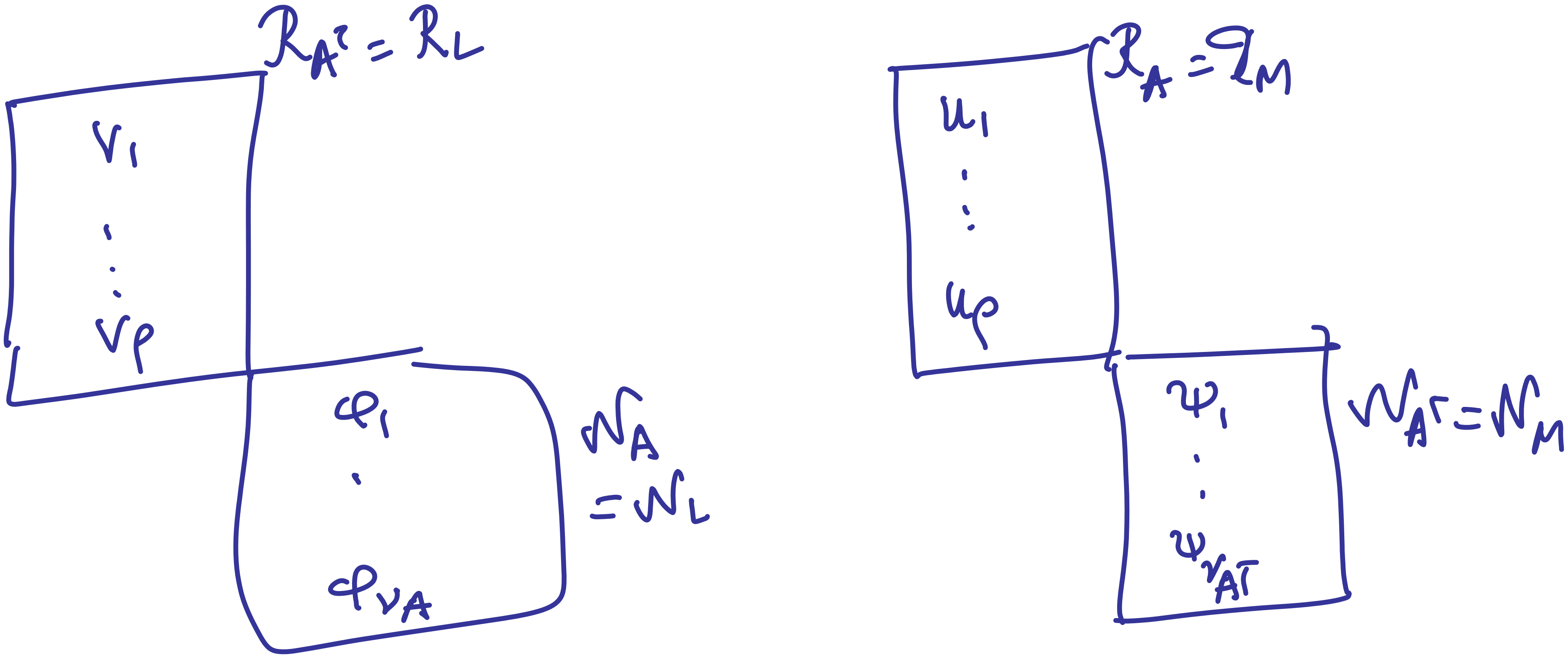
$$u_j \in \mathbb{R}^m$$

$$A^T u_j = \frac{A^T A v_j}{s_j} = \frac{L v_j}{s_j} = \frac{\lambda_j v_j}{s_j}$$

$$= \sqrt{\lambda_j}\, v_j \qquad (\because s_j = \sqrt{\lambda_j})$$

$$= s_j v_j$$

$$\boxed{A^T u_j = s_j v_j}$$



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

$$\boxed{A v_j = s_j u_j \quad ; \quad A^T u_j = s_j v_j}$$

$j = 1, 2, \cdots, \rho$

We have

$v_1, \ldots, v_p, \varphi_1, \varphi_2, \ldots, \varphi_{\nu_A}$ is an o.n. basis for $\mathbb{R}^n$

$u_1, \ldots, u_p, \psi_1, \psi_2, \ldots, \psi_{\nu_{A^T}}$ is an o.n.b for $\mathbb{R}^m$

$$\left.\begin{aligned} A v_j &= s_j u_j \\ A^T u_j &= s_j v_j \end{aligned}\right\} \quad 1 \le j \le p$$

$s_j = \sqrt{\lambda_j}$, ($\lambda_1, \lambda_2, \ldots, \lambda_p$ are the positive eigvalues of L)