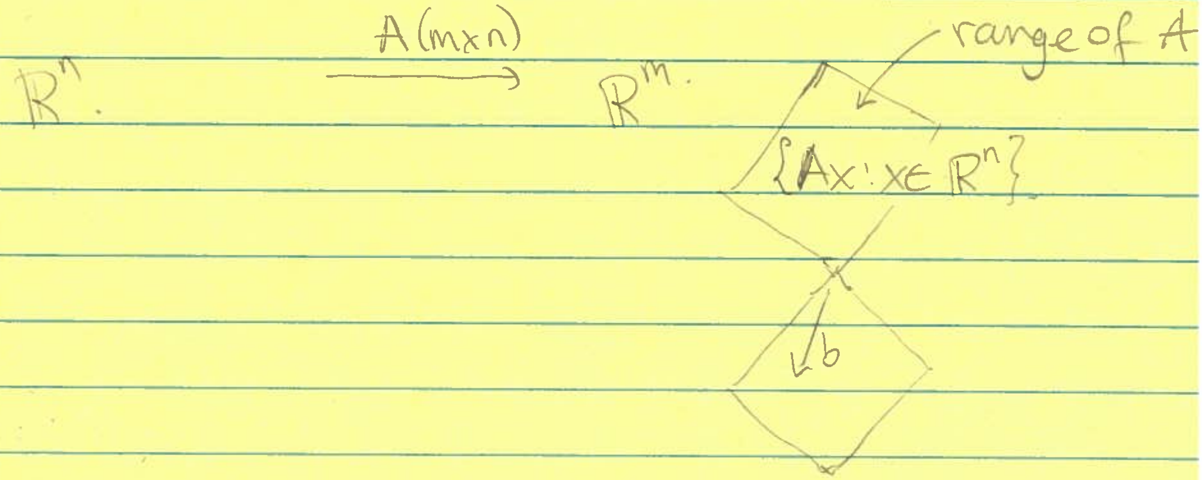
Lec 25 — **Singular Value Decomposition**

SYSTEMS OF EQNS — Consider the system of equations

$$Ax = b \qquad A = m \times n.$$

When is there no unique solution?

CASE 1: $Ax = b$ has no solution, ie. there is no x that maps to b under the linear transformation A, ie. $b \notin$ range of A:



Lemma — Range of A is the vector subspace spanned by its column vectors, called the column space.

Pf — Let $A = \begin{bmatrix} | & | & & | \\ a^1 & a^2 & \cdots & a^n \\ | & | & & | \end{bmatrix}$ and $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Then

$$[Ax]_i = a^1_i x_1 + \cdots a^n_i x_n$$

$$= x_1 a_i^1 + \cdots + x_n a_i^n$$

$$= x_1 (a^1)_i + \cdots + x_n (a^n)_i$$

$$= [x_1 a^1 + \cdots + x_n a^n]_i$$

Thus

$$Ax = x_1 a^1 + \cdots x_n a^n.$$

ie. Every vector in the range of A is a linear combination of the columns of A.

CASE 2: $Ax = b$ has multiple solutions. To see how this can come about, consider

$$Ax = 0$$

and suppose that the vector space of its solutions – called the NULLSPACE of A – contains more than just the zero vector. Then you can add any of these vectors to a "particular" sol$^{\underline{n}}$ of $Ax = b$ to yield another sol$^{\underline{n}}$ of $Ax = b$:

$$Ax_n = 0 \qquad x_n \neq 0$$

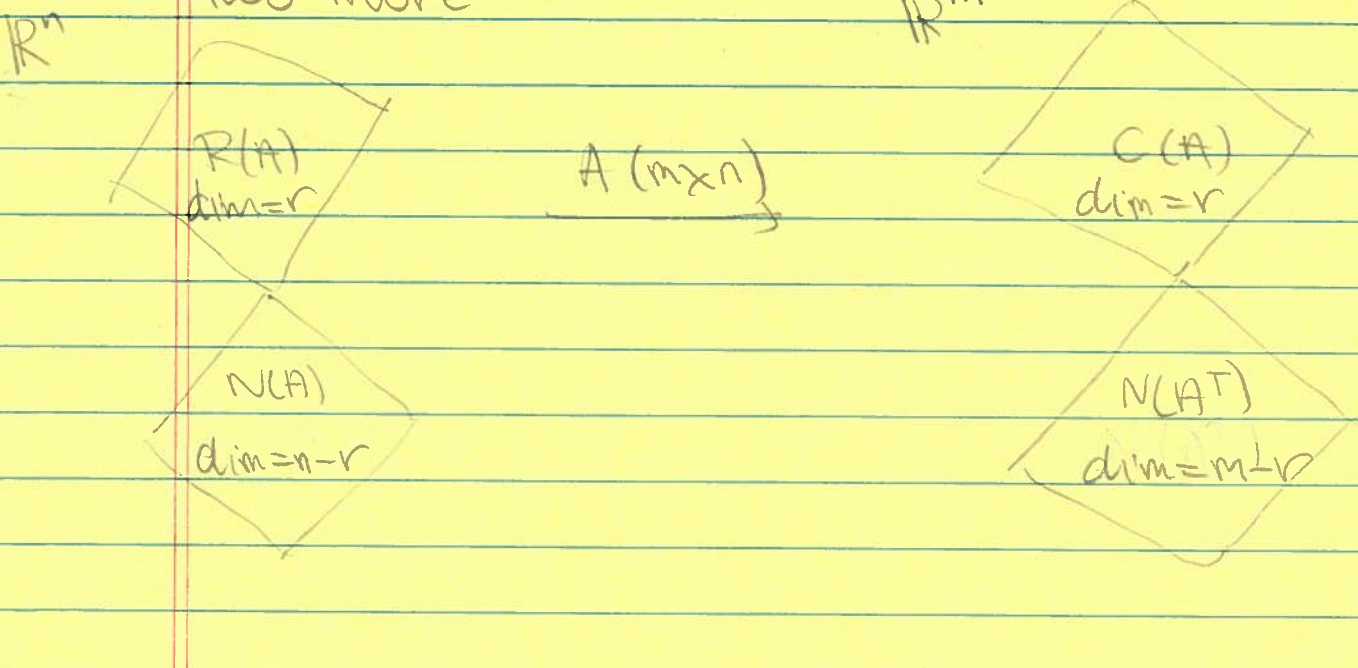
$$Ax_p = b$$

$$\Rightarrow \quad A(x_p + x_n) = b + 0 = b$$

$\Rightarrow Ax=b$ has multiple sol^{n}s.

FUNDAMENTAL SUBSPACES

So far we have described one subspace in $\mathbb{R}^m$, the column space of A, denoted $C(A)$, and one subspace in $\mathbb{R}^n$, the null space of A, denoted $N(A)$. In fact, there are two more:



$R(A)$ = row space of A
= all combos of rows of A.
(= all combos of cols of A^T)

$N(A^T)$ = all x st $A^T x = 0$
= all x st $x^T A = 0$. ("left nullspace of A")

FACT 1: $\dim R(A) = \dim C(A) = \text{rank of } A$

FACT 2: $N(A) \perp R(A)$

Pf (of FACT 2)

Suppose $x \in N(A)$

$\Rightarrow Ax = 0.$ $(*)$

Suppose:

$$A = \begin{bmatrix} - r_1^T - \\ \vdots \\ - r_n^T - \end{bmatrix}$$

Then $(*) \Rightarrow$

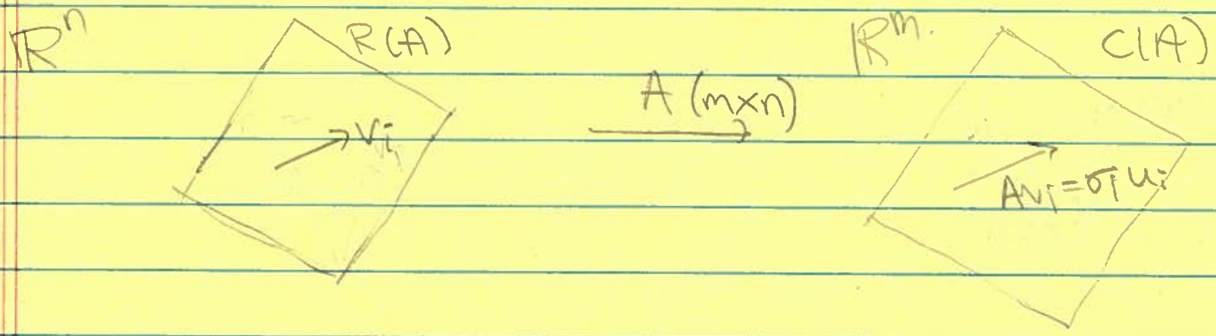
$$\begin{bmatrix} r_1^T x \\ r_2^T x \\ \vdots \\ r_n^T x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\Rightarrow$ $x \perp r_1, r_2, \ldots, r_n$

$\Rightarrow$ $N(A) \perp R(A)$

$\blacksquare$

SVD Singular Value Decomposition of A finds a orthonormal basis $\{v_1, \cdots, v_r\}$ in $R(A)$ that maps into an orthonormal basis $\{u_1, \cdots, u_r\}$ in $C(A)$:



In matrix form, this picture is:

$$\overset{m\times n}{A}\ \underbrace{\overset{n\times r}{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_r \\ | & & | \end{bmatrix}}}_{\substack{\text{orthogonal since} \\ v_i^T v_j = \delta_{ij}}} = \overset{m\times r}{\begin{bmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{bmatrix}} \overset{r\times r}{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}}$$

Next, extend

$$\begin{bmatrix} | & & | \\ v_1 & \cdots & v_r \\ | & & | \end{bmatrix}$$

to include an orthonormal basis of $N(A)$, denoted $\{v_{r+1}, \ldots, v_n\}$:

$$\overset{m\times n}{A}\;\overset{V:\,n\times n}{\begin{bmatrix} | & & | & | & & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & & | & | & & | \end{bmatrix}}$$

$$\qquad\quad \underbrace{\qquad\quad}_{R(A)} \quad \underbrace{\qquad\quad}_{N(A)}$$

arbitrary since $Av_i =$
$i = r+1, \ldots, n$, but must b
orthonormal, obtainable
e.g. via Gram Schmidt

$$= \underset{U:\,m\times m}{\begin{bmatrix} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{bmatrix}} \underset{\Sigma:\,m\times n}{\begin{bmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix}}$$

$$\qquad\quad \underbrace{\qquad\quad}_{C(A)} \quad \underbrace{\qquad\quad}_{N(A^T)}$$

Fortunately V is orthogonal, $v_i^T v_j = \delta_{ij}$, $i,j = 1, \ldots, n$, because $N(A) \perp R(A)$, so we may write:

$$AV = U\Sigma \Rightarrow A = U\Sigma V^{-1}$$

$$\Rightarrow \boxed{A = U\Sigma V^T}$$

This is called the SVD of A.

COMPUTING THE SVD

Consider

$$\overbrace{A^T A}^{n\times n} = V\Sigma^T \underbrace{U^T U}_{I} \Sigma V^T$$

$$= V\Sigma^T \Sigma V^T$$

$$= \overset{n\times n}{V} \overset{n\times n}{\begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}} \overset{n\times n.}{V^T}$$

ie. A^TA is diagonalized by an orthogonal matrix and it has positive eigenvalues. In fact, this is a property of all symmetric,

$$(A^TA)^T = A^TA,$$

and positive definite matrices,

$$x^T(A^TA)x = (Ax)^T(Ax) = \|Ax\|^2 \geq 0.$$

(see Thm 9.16 and 9.18 on p. 572, 573 of the TEXT).

In summary, we have decomposed a symmetric, positive definite matrix into its eigenvectors (the columns of V) and its eigenvalues ($\sigma_1^2, \ldots \sigma_r^2 \geq 0$) In other words, to compute V and Σ we should compute eigenvectors and eigenvalues of A^TA.

Similarly we compute U by finding eigenvectors of AA^T since

$$AA^T = U \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \end{bmatrix} U^T$$

NOTE 1: $\|u_i\| = 1$. To see this, consider

$$A^T A v_i = \sigma_i^2 v_i$$

$$\Rightarrow v_i^T A^T A v_i = \sigma_i^2 v_i^T v_i$$

$$\Rightarrow \|\underbrace{Av_i}_{\sigma_i u_i}\|^2 = \sigma_i^2 \underbrace{\|v_i\|^2}_{1}$$

$$\Rightarrow \sigma_i^2 \|u_i\|^2 = \sigma_i^2$$

$$\Rightarrow \|u_i\| = 1. \qquad \boxtimes$$

NOTE 2: $A^T A$ and AA^T have same eigenvalues. This is no accident. In general: AB and BA have the same eigenvalues:

$$A\underbrace{Bv_i}_{w_i} = \lambda_i v_i$$

$$\Rightarrow Aw_i = \lambda_i v_i$$

$$\Rightarrow BAw_i = \lambda_i \underbrace{Bv_i}_{w_i}$$

$$\Rightarrow BAw_i = \lambda_i w_i.$$

Thus evalues are same and evectors are related by

$$Bv_i = w_i. \qquad \boxtimes$$