Section 5.11 Orthogonal Complement and Orthogonal Decomposition

Let M be a set of V, a vector space with an inner product

$$M^{\perp} = \{x \in V \,/\, (m, x) = 0 \;\; \forall m \in M\}$$

We already saw that if M is a subspace, $M^{\perp}$ is a subspace.

We also saw last class

$$H = \{x \in V \,/\, (u, x) = 0\}$$

for some vector $u \in V$. Hyperplane

$$H = L^{\perp} \quad \text{where } L = \{x = \alpha u, \alpha \in \mathbb{R}\}$$

When M is a subspace

$$M \oplus M^{\perp} = V$$

complementary subspaces

M "perp" is the orthogonal complement of M

<u>Lemma</u> M and $N = M^\perp$ are complementary subspaces of V.

Proof First. $M \cap M^\perp = \{0\}$

take $x \in M \cap M^\perp$, then $\langle x, x\rangle = 0$

$\Rightarrow x = 0$.

Now consider B_M a basis of M with orthonormal vectors.
and B_N an orthonormal basis of $M^\perp$

Since $M \cap M^\perp = \{0\}$

$B_M \cup B_N$ is an orthonormal basis of

a subspace $S \subseteq V$

If $\exists\ v \in V$, $v \neq 0$, $v \notin M + M^\perp$ (ie $v \notin M$, $v \notin M^\perp$)

then we can find $u \in V$ (by Gram-Schmidt)

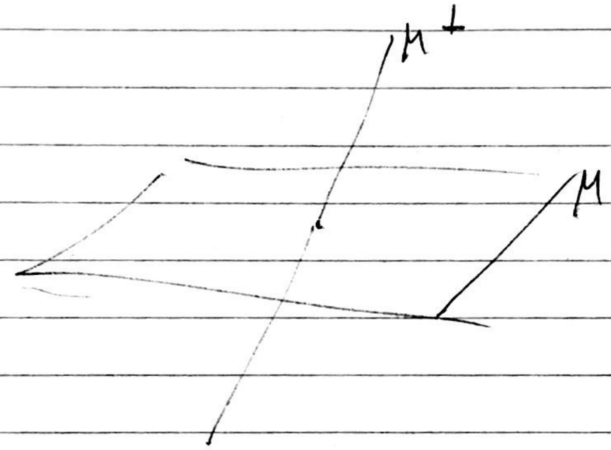
so that $u \perp M$ ~~[illegible]~~

but then $u \in M^\perp$, or $u = 0$.

$\Rightarrow\ M + M^\perp = V$.

q.e.d

Note that $(M^{\perp})^{\perp} = M$



A $m \times n$

<u>Lemma</u> $R(A) \perp N(A^T)$, $\left(R(A) = N(A^T)^{\perp}\right)$

and similarly $R(A^T) \perp N(A)$

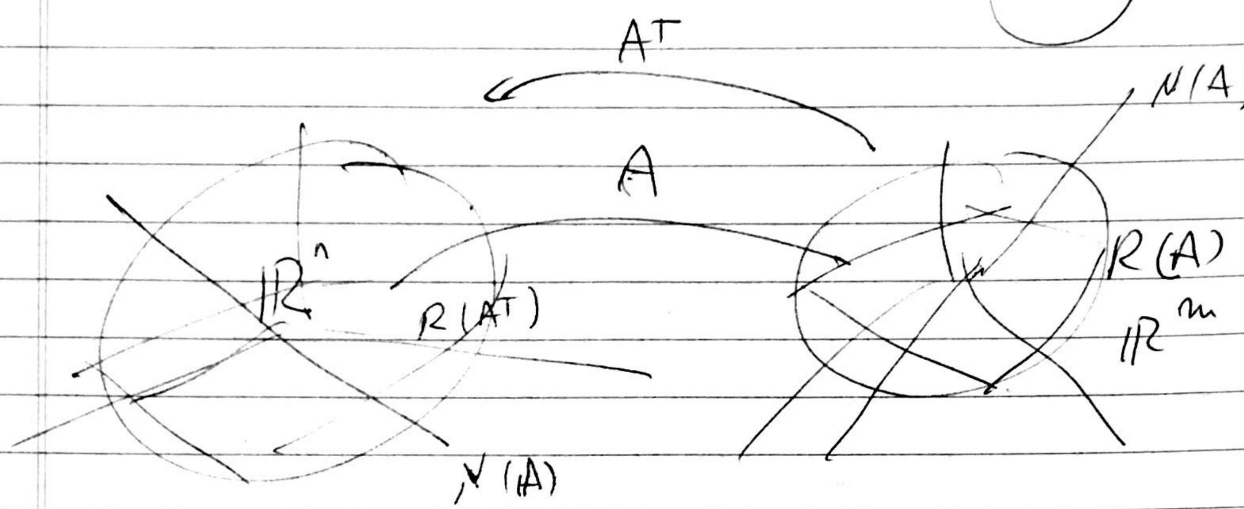
We already saw

$$\dim R(A) + \dim N(A^T) = n$$

$$\dim R(A^T) + \dim N(A) = m$$

so that $R(A) \oplus N(A^T)$.

$$R(A^T) \oplus N(A)$$



$R(A) \perp N(A^T)$

if $x \in R(A)$ $\quad x = Aw$ for some w

if $y \in N(A^T)$ $\quad A^T y = 0$

$$(x,y) = x^T y = \underline{(Aw)^T} y = w^T A y = 0$$

$$R(A) = N(A^T)^{\perp}$$

Section 5.12 Singular Value Decomposition

We saw $A = LU$ (if no pivoting)

$A = QR \qquad Q^T Q = I$

QR useful for least squares.
More stable for computations

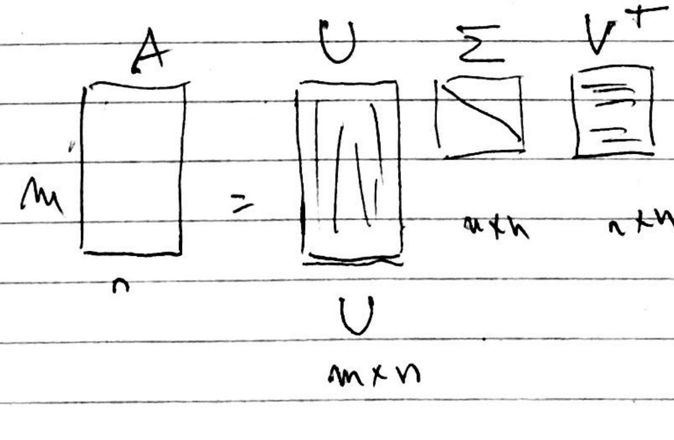
New Decomposition (factorization)

$$A = U \Sigma V^T$$

U, V orthogonal, Σ diagonal

$$\Sigma = \operatorname{diag}(\sigma_1\ \sigma_2\ \sigma_3 \cdots \sigma_n)$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$



$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T \qquad \text{if } \sigma_n > 0$$

rank $A = n$

if $\sigma_r > 0$

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T \qquad \text{rank } A = r$$

~~Rank A~~ $\mathcal{R}(A) = \mathcal{R}(U)$

$$Ax = U\Sigma V^T x = U(\Sigma V^T x) = Uw$$

$\{v_1 v_2 \dots v_r\}$ basis of $N(A)^\perp = R(A^T)$

For $x \in N(A)$ i.e. $v_i^T x = 0$ $i = 1 - r$ $V^T x = 0$

$$Ax = U\Sigma V^T x = U\Sigma 0 = 0$$

Recall $A = uv^T$ rank 1

$A = u_1 v_1^T + u_2 v_2^T$ rank 2

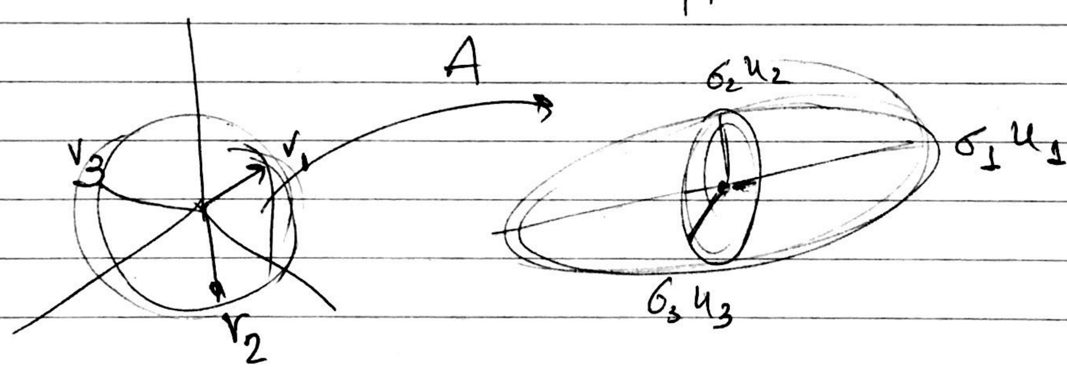
etc.

v_i orthogonal basis of $\mathbb{R}^n$
(lie in unit sphere)

$$A v_j = \sum \sigma_i u_i v_i^T v_j$$
$$= \sigma_j u_j v_j^T v_j = \sigma_j u_j$$

u_j length 1 (orthonormal vectors)

σ_j, length of image of v_j, when A is applied



remember $\sigma_1 \geq \sigma_2 \geq \sigma_3$

so σ_1 is maximum streach $= \|A\|$

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \sigma_1$$

Recall U V orthogonal so that

$$\|Ux\| = \|x\|$$
$$\|V^T y\| = \|y\|$$

$$\|Ax\| = \|U\Sigma V^T x\|$$

$$\max_{\|x\|=1} \|Ax\| = \cancel{\|\Sigma\|} \max_{\|V^T x\|=1} \|U\Sigma V^T x\|$$

$$= \max_{\|y\|=1} \|\Sigma y\| = \sigma_1$$

$$\Sigma = \begin{vmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{vmatrix}$$

Note if A is square and nonsingular, U, V square

$$A^{-1} = (U\Sigma V^T)^{-1} =$$
$$= V^{-T}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$$

the singular value decomposition of A^{-1}

So that $\|A^{-1}\|_2 = \frac{1}{\sigma_n}$

(the maxim of $\frac{1}{\sigma_1}, \frac{1}{\sigma_2} \cdots \frac{1}{\sigma_n}$)

SVD: $A = U \Sigma V^T$

$$A V = U \Sigma$$

$$A \begin{vmatrix} v_1 & v_2 & \cdots & v_n \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{vmatrix} \begin{vmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots & \\ & & & \sigma_n \end{vmatrix}$$

$$= \begin{vmatrix} \sigma_1 u_1 & \sigma_2 u_2 & \cdots & \sigma_n u_n \\ | & | & & | \end{vmatrix}$$

How to prove existence of SVD

take v_1 of norm 1 so that

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \|Av_1\| = \sigma_1$$

let $u_1 = \dfrac{Av_1}{\|Av\|}$

then $Av_1 = \sigma_1 u_1$

Note $u_1^T A v_1 = \sigma_1 \underbrace{u_1^T u_1}_{=1} = \sigma_1$

let $V_1 = \{v_1, \bar{v}_2, \bar{v}_3, \dots \bar{v}_n\}$ orthonormal basis of $\mathbb{R}^n$

$U_1 = \{u_1, \bar{u}_2, \bar{u}_3, \dots \bar{u}_m\}$ orthonormal basis of $\mathbb{R}^n$

then
$$T = U_1^T A V_1 = \begin{bmatrix} \sigma_1 & w^T \\ 0 & \\ \vdots & B \\ 0 & \end{bmatrix}$$

$\boxed{\|T\|_2 = \|A\|_2 = \sigma_1 \quad (U_1\ V_1 \text{ orthogonal})}$

We can show that $w \in \mathbb{R}^{n-1}$ is $= 0$

then we can repeat the

process with $\sigma_2 = \|B\|$

$$B \hat{v}_2 = \sigma_2 \hat{u}_2$$

$$\text{get } V_2 = \begin{vmatrix} 0 \\ \hat{v}_2 \end{vmatrix} \quad \tilde{u}_2 = \begin{vmatrix} 0 \\ u_2 \end{vmatrix}$$

$$\begin{vmatrix} u_1 & u_2 \end{vmatrix}^T A \begin{vmatrix} V_1 & V_2 \end{vmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

Repeating the process we get the SVD.

Proof that $w = 0$. Assume $w \neq 0$

consider the vector $\begin{vmatrix} \sigma_1 \\ w \end{vmatrix}$

$$T \begin{vmatrix} \sigma_1 \\ w \end{vmatrix} = \begin{vmatrix} \sigma_1 & w^T \\ 0 & B \\ \vdots & \\ 0 & \end{vmatrix} \begin{vmatrix} \sigma_1 \\ w \end{vmatrix} = \begin{vmatrix} \sigma_1^2 + w^T w \\ B w \end{vmatrix}$$

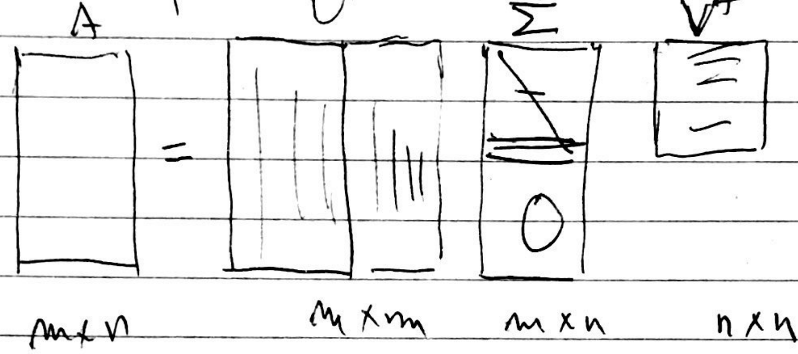
$$\left\| T \begin{vmatrix} \sigma_1 \\ w \end{vmatrix} \right\| \geq |\sigma_1^2 + w^T w| = (\sigma_1^2 + w^T w)^{1/2} \left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2$$

$$\Rightarrow \underset{\sigma_1}{\overset{\|}{\|T\|_2}} \geq (\sigma_1^2 + w^T w)^{1/2} \Rightarrow w = 0.$$

QED.

Now, let us repeat the picture in p. 154, and enlarge U

complete the orthogonal basis to $\mathbb{R}^m$



so now both U and V are square

this is the Full SVD

Note if Q is orthogonal $Q^T Q = I$

$$Q = Q \, I \, I$$
$$\quad\;\; U \, \Sigma \, V^T$$

__Every__ matrix has an SVD. We showed existence. Σ unique

U, V unique up to sign say $u_1 \to -u_1$, $v_1 \to -v_1$

Super important ~~application~~ Result
(with many applications)

if A has rank n

$$A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

Find the matrix B of rank k
such that
$\|A - B\|$ is smallest.

Well $B = \sum_{i=1}^{k} \sigma_i u_i v_i^T$

$$A - B = \sum_{i=k+1}^{n} \sigma_i u_i v_i^T$$

$$\|A - B\| = \sigma_{k+1} = \min_{\operatorname{rank} E = k} \|A - E\|$$

Recall $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_k \geq \sigma_{k+1} \geq \cdots$

One application <u>image compression</u>

Example

$$A = U\Sigma V^T = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{vmatrix} \begin{vmatrix} 3 & & \\ & 2 & \\ & & 1 \end{vmatrix} \begin{vmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{vmatrix}$$

$$= \frac{1}{2}\begin{vmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{vmatrix} \begin{vmatrix} 3 & 3 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & \sqrt{2} \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 5 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 5 & 0 \end{vmatrix}$$

B of rank 2

$$B = 3u_1 v_1^T + 2 u_2 v_2^T$$

$$= \frac{1}{2}\begin{vmatrix} 1 & 1 \\ 0 & 0 \\ 1 & -1 \end{vmatrix} \begin{bmatrix} 3 & \\ & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} =$$

$$= \frac{1}{2}\begin{vmatrix} 1 & 1 \\ 0 & 0 \\ 1 & -1 \end{vmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 2 & -2 & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 5 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 5 & 0 \end{bmatrix}$$

rank 2

$$A - B = \frac{1}{2}\begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} \qquad \|A - B\| = 1 = \sigma_3$$