LECTURE 7

APPLICATIONS OF SVD (CONTINUES)

Rank of a matrix

Let V be a vector space with a basis

$$\mathcal{B} = \{b_1, \dots, b_\ell\}.$$

Then we say the dimension of V is ℓ and write $\dim V = \ell$.

e.g.

$\{1, x, x^2\}$ is a basis for

$$\mathbb{P}_2 = \{a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}$$

$$\dim \mathbb{P}_2 = 3$$

$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for

$$\mathcal{S}^{2\times 2} = \{A \in \mathbb{R}^{2\times 2} \mid A^T = A\}$$

$$\dim \mathcal{S}^{2\times 2} = 3$$

DEFN (Rank)

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

$$\operatorname{rank}(A) = \dim(\operatorname{Col}(A)).$$

e.g.

$$\operatorname{rank}\left(\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\right) = 1$$

$$\operatorname{rank}\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right) = 2$$

THM (Rank & Null Spaces)

Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$. Then

$$\operatorname{rank}(A) + \dim(\operatorname{Null}(A)) = n.$$

PROOF

Let $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$ and

$\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0.$

$$\operatorname{Col}(A) = \operatorname{span} \overbrace{\{u_1, \dots, u_r\}}^{\text{Basis}}$$

and

$$\operatorname{Null}(A) = \operatorname{span} \underbrace{\{v_{r+1}, \dots, v_n\}}_{\text{Basis}}$$

$$\Longrightarrow$$

$$\underbrace{\operatorname{rank}(A)}_{r} + \underbrace{\dim(\operatorname{Null}(A))}_{n-r} = n$$

e.g.

$$\operatorname{rank}\left(\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\right)=1 \quad \text{and} \quad \dim\left(\operatorname{Null}\left(\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\right)\right)=1$$

Indeed

$$\operatorname{Null}\left(\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}\right)=\operatorname{span}\left\{\begin{bmatrix}-2/\sqrt{5}\\1/\sqrt{5}\end{bmatrix}\right\}.$$

③ SVD and rank of a matrix

The following is deduced from the proof of previous thm.

THM

Let $A \in \mathbb{R}^{m\times n}$ with singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$$

and

$$\sigma_{r+1} = \cdots = \sigma_n = 0.$$

Then $\operatorname{rank}(A) = r$.

e.g.

$\boxed{\text{SVD}}$ $$\begin{bmatrix}1 & 2\\1 & 2\end{bmatrix}=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\0 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{5} & 2/\sqrt{5}\\2/\sqrt{5} & -1/\sqrt{5}\end{bmatrix}$$

$$\sigma_1=\sqrt{10}, \quad \sigma_2=0$$

$$\operatorname{rank}(A)=1$$

DEFN (Inverse of a matrix)
Let $A \in \mathbb{R}^{n \times n}$. The matrix $X = A^{-1}$ satisfying

$$AX = XA = I_n \ (\substack{n \times n \\ \text{identity}})$$

is called the inverse of A.

A is invertible ($\substack{\text{OR} \\ \text{non-singular}}$) — A^{-1} exists

$$AX = I_n \quad \exists X \in \mathbb{R}^{n \times n}$$

$$\Longleftrightarrow$$

$$Ax_1 = e_1 \ \exists x_1 \in \mathbb{R}^n \text{ and } \ \dots \ \text{ and } Ax_n = e_n \ \exists x_n \in \mathbb{R}^n$$

$$\Longleftrightarrow$$

$$Col(A) = \mathbb{R}^n$$

$$\Longleftrightarrow$$

$$rank(A) = n$$

$$\Longleftrightarrow$$

$$\dim(Null(A)) = 0$$

$$\Longleftrightarrow$$

$$Ax \neq 0 \quad \forall x \neq 0 \in \mathbb{R}^n$$

THM (Characterizations of Invertibility)
Let $A \in \mathbb{R}^{n \times n}$. The following are equivalent.

(1) A is invertible
(2) $rank(A) = n$
(3) $Ax \neq 0 \quad \forall x \neq 0 \in \mathbb{R}^n$
(4) $\det(A) \neq 0$

(5) $\{a_1, \cdots, a_n\}$ is linearly independent.

e.g.

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

rank$(A) = 1$

A is not invertible.

$Ax = 0$ for $x = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \neq 0$

$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$ is linearly dependent

COROLLARY

Let $A \in \mathbb{R}^{n \times n}$. Then

$$A \text{ is invertible} \iff \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$$

④ SVD and low rank approximations

THM (Optimal Rank)

Let $A \in \mathbb{R}^{m \times n}$ $(m \geq n)$ with SVD

$$A = [u_1 \cdots u_n] \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_n \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}.$$

Then

(i) $$\min_{\substack{\Delta A \in \mathbb{R}^{m \times n} \\ \text{s.t.} \\ \text{rank}(A + \Delta A) = r}} \|\Delta A\|_2 = \sigma_{r+1}.$$

(ii) $\Delta A_* = -\sum_{j=r+1}^{n} \sigma_j u_j v_j^T$ is such that $\|\Delta A_*\|_2 = \sigma_{r+1}$ and $\text{rank}(A + \Delta A_*) = r$.

e.g.

$$\boxed{\text{SVD}}\quad \underbrace{\begin{bmatrix} 5 & 5 & 0 \\ 1 & 7 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 3\sqrt{10} & 0 & 0 \\ 0 & \sqrt{10} & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{V^T}$$

* Distance to the nearest rank two matrix (Same as distance singularity)

$$\sigma_3 = 2$$

$$\left(\begin{matrix}\text{OPTIMAL}\\ \text{PERTURBATION}\end{matrix}\right) \Delta A_* = -\sigma_3 u_3 v_3^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$\left(\begin{matrix}\text{NEAREST}\\ \text{RANK TWO}\\ \text{OR SINGULAR}\\ \text{MATRIX}\end{matrix}\right) A + \Delta A_* = \begin{bmatrix} 5 & 5 & 0 \\ 1 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

* Distance to the nearest rank one matrix

$$\sigma_2 = \sqrt{10}$$

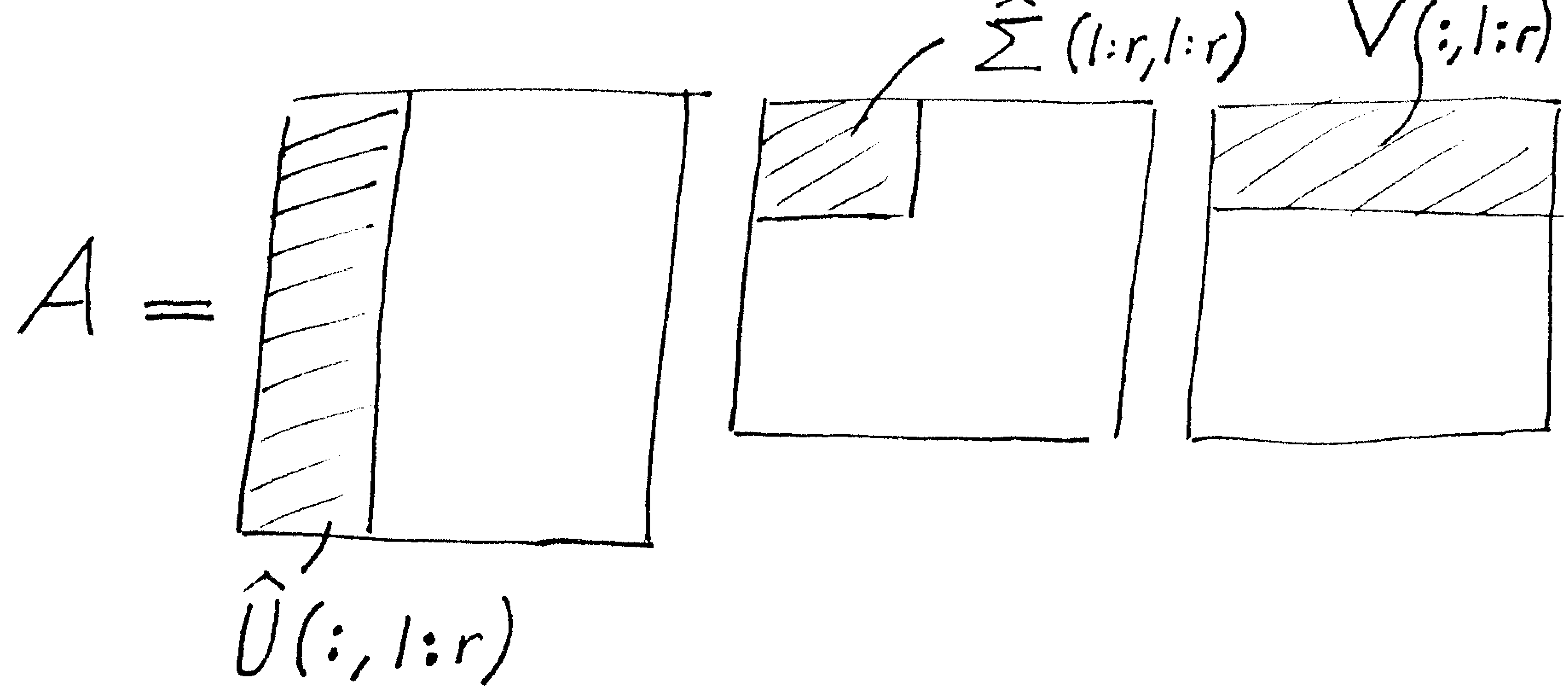
$$\left(\begin{matrix}\text{OPTIMAL}\\ \text{PERTURBATION}\end{matrix}\right) \Delta A_* = -\sigma_2 u_2 v_2^T - \sigma_3 u_3 v_3^T$$

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

$$\left(\begin{matrix}\text{NEAREST}\\ \text{RANK ONE}\\ \text{MATRIX}\end{matrix}\right) A + \Delta A_* = \begin{bmatrix} 3 & 6 & 0 \\ 3 & 6 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Application to Image Compression

An image can be stored in a matrix; the entries of the matrix keep the intensity values of the image.



$$A_r = \sum_{j=1}^{r} \sigma_j u_j v_j^T$$
$$= \hat{U}(:,1:r)\,\hat{\Sigma}(1:r,1:r)\,\hat{V}(:,1:r)^T$$

* $A \approx A_r$ when σ_{r+1} is small
* A_r is the best rank r representation of A.

PROOF OF THM (Optimal Rank)

(ii) Obvious

(i) We only need to show

$$(*) \quad \|\Delta A\|_2 \geq \sigma_{r+1}$$

$$\forall \Delta A \text{ s.t. } \operatorname{rank}(A + \Delta A) = r$$

Suppose otherwise, i.e., there exists a ΔA s.t.

$$\|\Delta A\|_2 < \sigma_{r+1} \text{ and } \operatorname{rank}(A + \Delta A) = r.$$

Consider following subspaces

(1) $S_1 = \operatorname{null}(A + \Delta A)$
of dimension $n-r$

(2) $S_2 = \operatorname{span}\{v_1, \ldots, v_{r+1}\}$
of dimension $r+1$.

Notice $S_1 \cap S_2 \neq \emptyset$ (since the dimensions sum up to $n+1$).

But for all $x \in S_1$

$$
\begin{aligned}
\|Ax\|_2 &= \|Ax - (A+\Delta A)x\|_2 \\
&= \|(-\Delta A)x\|_2 \\
&\leq \|\Delta A\|_2 \, \|x\|_2 \\
&< \sigma_{r+1} \, \|x\|_2 ,
\end{aligned}
$$

and for all $x \in S_2$ there exist $\alpha_1, \ldots, \alpha_{r+1}$ s.t.

$$
\begin{aligned}
\|Ax\|_2 &= \|A(\alpha_1 v_1 + \cdots + \alpha_{r+1} v_{r+1})\|_2 \\
&= \|\alpha_1 \sigma_1 u_1 + \cdots + \alpha_{r+1} \sigma_{r+1} u_{r+1}\|_2 \\
&\cancel{= \| \sum \sigma_i u_i} = \sqrt{\sum_{j=1}^{r+1} \alpha_j^2 \sigma_j^2} \\
&\geq \sqrt{\sum_{j=1}^{r+1} \alpha_j^2 \sigma_{r+1}^2} \\
&= \sigma_{r+1} \|x\|
\end{aligned}
$$

yielding the contradiction that $S_1 \cap S_2 = \phi$. Consequently we deduce $(*)$ $\square$

Note: $\|x\|_2 = \sqrt{(\alpha_1 v_1 + \cdots + \alpha_{r+1} v_{r+1})^T (\alpha_1 v_1 + \cdots + \alpha_{r+1} v_{r+1})}$
$= \sqrt{\alpha_1^2 + \cdots + \alpha_{r+1}^2}$