LECTURE 5

SINGULAR VALUE DECOMPOSITION

Orthogonality:

A matrix $Q \in \mathbb{R}^{n\times n}$ ($U \in \mathbb{C}^{n\times n}$) is called orthogonal (unitary) if

$$Q^T Q = I \quad (U^* U = I)$$

e.g.

$$\tilde{Q} = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$ is orthogonal.

$$Q^T Q = I \iff q_i^T q_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$$\iff q_i \underset{i\neq j}{\perp} q_j \quad \text{and} \quad \|q_i\| = 1$$

$$\iff \{q_1, \dots, q_n\} \text{ is an orthonormal set}$$

Similarly for a unitary matrix $U \in \mathbb{C}^{n\times n}$ the set $\{u_1, \dots, u_n\}$ of columns is an orthonormal set.

g.

For the example above $\left\{ \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \right\}$ is an orthonormal set in $\mathbb{R}^2$.

THM (Unitary Invariance)

Let $U_1 \in \mathbb{C}^{m\times m}$ and $U_2 \in \mathbb{C}^{n\times n}$ be unitary matrices. For all $A \in \mathbb{C}^{m\times n}$ we have

(i) $\|A\|_2 = \|U_1 A\|_2 = \|A U_2\|_2$

(ii) $\|A\|_F = \|U_1 A\|_F = \|A U_2\|_F$

PROOF

(i) $\boxed{\|A\|_2 = \|U_1 A\|_2}$

$$\|U_1 A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_2 = 1}} \|U_1 A x\|_2$$

$$= \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_2 = 1}} \sqrt{(U_1 A x)^* (U_1 A x)}$$

$$= \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_2 = 1}} \sqrt{x^* A^* \underbrace{U_1^* U_1}_{I} A x}$$

$$= \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t. } \|x\|_2 = 1}} \|Ax\|_2 = \|A\|_2$$

(i) $\boxed{\|A\|_2 = \|AU_2\|_2}$

$$\|AU_2\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_2 = 1}} \|AU_2 x\|_2$$

Letting $y := U_2 x$ we have

(1) $\|y\|_2 = \sqrt{x^* U_2^* U_2 x} = \sqrt{x^* x} = \|x\|_2 = 1.$

(2) Each x is mapped to unique y, i.e., $x = U_2^* y$.

Consequently

$$\|AU_2\|_2 = \max_{\substack{y \in \mathbb{R}^n \\ \text{s.t.} \\ \|y\|_2 = 1}} \|Ay\|_2 = \|A\|_2$$

(ii) Exercise

~~Existence of SVD~~

Definition of Singular Values and Vectors

Let $A \in \mathbb{R}^{2\times 2}$ with SVD

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \end{bmatrix}$$

Then

$$(1)\quad u_1^* A = \sigma_1 v_1^* \quad \text{and} \quad A v_1 = \sigma_1 u_1$$

$$(2)\quad u_2^* A = \sigma_2 v_2^* \quad \text{and} \quad A v_2 = \sigma_2 u_2$$

<u>DEFN</u> (Singular Values)

Let $A \in \mathbb{C}^{m \times n}$. Suppose that the equations

$$(i)\quad Av = \sigma u, \quad \text{and}$$

$$(ii)\quad u^* A = \sigma v^*$$

hold for some $v \neq 0 \in \mathbb{C}^m$, $u \neq 0 \in \mathbb{C}^n$ and real scalar $\sigma \geq 0$. Then

(i) σ is called a singular value,

(ii) and u, v are called the left and right singular vectors associated with σ.

e.g.

SVD for

$$\begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

$\sigma_1 = \sqrt{10}$ is a singular value,

and $u_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, $v_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ are the associated left/right singular vectors.

...oof of Existence of SVD

Basic Idea

* Find unitary transformations U and V s.t.

$$U^* A V \text{ is diagonal}$$

EXAMPLE

Let

$$A = \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}.$$

For all x s.t $\|x\|_2 = 1$

$$\|Ax\|_2 = \left\| \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + x_2 \end{bmatrix} \right\|_2 = \sqrt{10x_1^2 + 10x_2^2} = \sqrt{10}$$

Consequently $\|A\|_2 = \sqrt{10}$. For instance for $v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ we have

$$A v_1 = \sqrt{10} \underbrace{\begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}}_{u_1 \ (\|u_1\|_2 = 1)} \qquad \left(\frac{\text{Note}}{u_1^T A = \sigma_1 v_1^T} \right)$$

Choose v_2 so that

(1) $\{v_1, v_2\}$ is orthonormal, equivalently

(2) $V = [v_1 \ v_2]$ is orthogonal

Similarly choose u_2 so that

$$U = [u_1 \quad u_2] \text{ is orthogonal}$$

Specifically

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \text{ and } U = \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Now consider

$$u_1^T A v_1 = \sqrt{10}\, u_1^T u_1 = \sqrt{10}$$
$$u_2^T A v_1 = \sqrt{10}\, u_2^T u_1 = 0 \quad (\text{Because } u_2 \perp u_1)$$

That is

$$U^T A v_1 = \begin{bmatrix} \sqrt{10} \\ 0 \end{bmatrix}$$

whereas

$$U^T A v_2 = \begin{bmatrix} u_1^T A v_2 \\ u_2^T A v_2 \end{bmatrix} = \begin{bmatrix} \sqrt{10}\, v_1^T v_2 \\ u_2^T A v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ u_2^T A v_2 \end{bmatrix}$$

In other words

$$U^T A V = [U^T A v_1 \quad U^T A v_2] \text{ is diagonal.}$$

Indeed

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

Consequently

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

<u>THM</u> (Existence of SVD)

Every matrix $A \in \mathbb{C}^{n\times n}$ has an SVD of the form

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

where $V, U \in \mathbb{C}^{n\times n}$ are unitary, and Σ is diagonal with non-negative entries.

<u>PROOF</u>

Let $\sigma_1 := \|A\|_2$. Since the maximization

$$\max_{\substack{x \in \mathbb{C}^n \\ \text{s.t.} \\ \|x\|_2 = 1}} \|Ax\|_2$$

involves a

* continuous objective function

$$x \longrightarrow \|Ax\|_2$$

* over a compact domain

$$\{x \in \mathbb{C}^n \mid \|x\|_2 = 1\},$$

it will be attained at some $v_1 \in \mathbb{C}^{n\times n}$, i.e.,

$$Av_1 = \sigma_1 u_1$$

for some $v_1, u_1 \in \mathbb{C}^n$, where $\|v_1\| = \|u_1\| = 1$.

Choose vectors $v_2, \cdots, v_n \in \mathbb{C}^n$ s.t.

$$\hat{V} = [v_1 \ v_2 \ \cdots \ v_n] \in \mathbb{C}^{n\times n}$$

is orthogonal. Similarly, choose $u_2, \cdots, u_n \in \mathbb{C}^n$ s.t.

$$\hat{U} = [u_1 \ u_2 \ \cdots \ u_n] \in \mathbb{C}^{n\times n}$$

is orthogonal.

Since $u_2^* u_1 = u_3^* u_1 = \cdots = u_n^* u_1 = 0$, we have

$$\hat{U}^* A \hat{V} = \begin{bmatrix} \begin{matrix} u_1^T \sigma_1 u_1 \\ u_2^T \sigma_1 u_1 \\ \vdots \\ u_n^T \sigma_1 u_1 \end{matrix} & \hat{U}^* A v_2 & \cdots & \hat{U}^* A v_n \end{bmatrix}$$

$$= \begin{bmatrix} \overbrace{\sigma_1}^{1\times 1} & \overbrace{w}^{1\times n} \\ \underbrace{0}_{n\times 1} & \underbrace{\tilde{A}}_{(n-1)\times(n-1)} \end{bmatrix}$$

Because of unitary invariance

$$\|\hat{U}^* A \hat{V}\|_2 = \|A\|_2 = \sigma_1 \geqslant \frac{\left\| \begin{bmatrix} \sigma_1 & w \\ 0 & \tilde{A} \end{bmatrix} \begin{bmatrix} \sigma_1 \\ w^* \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2}$$

$$\geqslant (\sigma_1^2 + w^T w^*) / \sqrt{\sigma_1^2 + w^T w^*}$$

Therefore $w=0$, and $U^T A V$ is of the form

$$(*)\quad \hat{U}^* A \hat{V} = \begin{bmatrix} \overbrace{\sigma_1}^{1\times 1} & 0 \\ 0 & \underbrace{\tilde{A}}_{(n-1)\times(n-1)} \end{bmatrix}.$$

Now the proof follows from induction. For a 1×1 scalar, there is obviously an SVD. Let us suppose for the $(n-1)\times(n-1)$ matrix $\tilde{A}$ there is an SVD of the form

$$\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^*$$

Then from $(*)$

$$\hat{U}^* A \hat{V} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \tilde{U}\tilde{\Sigma}\tilde{V}^* \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \tilde{\Sigma} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{V}^* \end{bmatrix}$$

$$\Longrightarrow$$

$$A = \underbrace{\hat{U}\begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \tilde{\Sigma} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \tilde{V}^* \end{bmatrix}\hat{V}^*}_{V^*}$$

where $U := \hat{U}\begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix}$ and $V := \hat{V}\begin{bmatrix} 1 & 0 \\ 0 & \tilde{V} \end{bmatrix}$ $\square$