Midterm exam (2 weeks from today)

- closed book + notes
- no calculators / computer simulations
- 50 mins (during class time)
- exam pdf available starting exam time.
- upload (to canvas) within 20 mins after exam time finishes. (pencil + paper problems)

☆ - Heavily based on HW problems

The singular value decomposition

MATH 6610 Lecture 09

September 23, 2020

Trefethen & Bau: Lectures 4, 5

Diagonalizability → transforming a matrix into a diagonal matrix via a similarity transform

Recall:

- All non-defective square matrices are diagonalizable (eigenvalue decomposition) $A = V \Lambda V^{-1}$
- All square matrices are bidiagonalizable (Jordan normal form) $A = V J V^{-1}$
- All square matrices are unitarily triangularizable (Schur decomposition)
- All normal matrices are unitarily diagonalizable (spectral theorem)

$A = UTU^*$

$A = U\Lambda U^*$

Diagonalizability

Recall:

- All non-defective square matrices are diagonalizable (eigenvalue decomposition)
- All square matrices are bidiagonalizable (Jordan normal form)
- All square matrices are unitarily triangularizable (Schur decomposition)
- All normal matrices are unitarily diagonalizable (spectral theorem)

What about rectangular matrices?

Ans: All matrices are diagonal, upon appropriate unitary transforms of the domain and range.

The singular value decomposition

Theorem (SVD) (Arbitrary m, n)

Any matrix $A \in \mathbb{C}^{m \times n}$ can be written as the product,

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

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

$\Sigma = \text{diag}(\sigma_1, \sigma_2 \ldots \sigma_{\min\{m,n\}}) \longrightarrow \{\sigma_j\}$ are "singular values"

The singular value decomposition

Theorem (SVD)

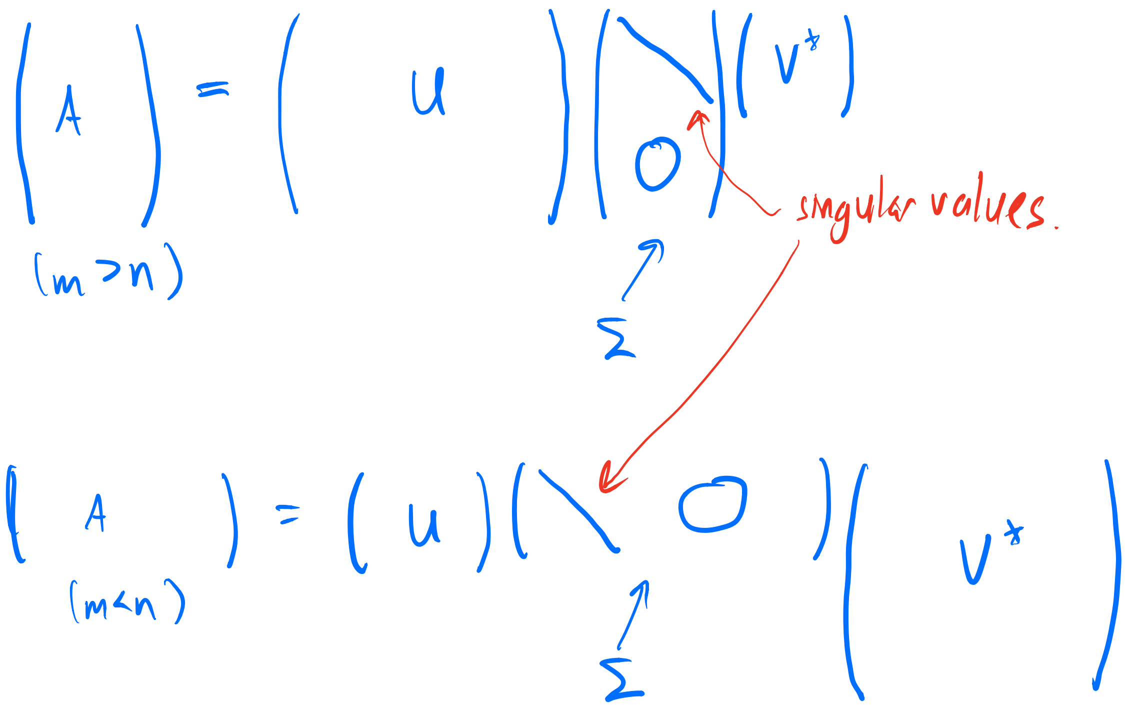
Any matrix $A \in \mathbb{C}^{m \times n}$ can be written as the product,

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

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

With $p = \min\{m, n\}$, notational convention:

- $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)$ ☆
- $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_p \geqslant 0$ (ordering is assumed)
- $U = [u_1,\ u_2,\ \ldots,\ u_m]$ (columns)
- $V = [v_1,\ v_2,\ \ldots,\ v_n]$ (columns)



Proof: Idea: induction on (m,n)

Base cases: (i) $m \geq 1$, $n = 1$ (A is a column vector)

$$A = \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} = \frac{a_1}{\|a_1\|} \cdot \|a_1\| \quad (\text{if } \|a_1\| > 0)$$

Define $U = \begin{pmatrix} | & | & & | \\ \frac{a_1}{\|a_1\|} & u_2 & \cdots & u_m \\ | & | & & | \end{pmatrix}$ where $\{u_j\}_{j=2}^{m}$ is an orthonormal completion of $\mathbb{C}^m$!

$$V = 1, \quad \sigma_1 = \|a_1\|$$

$$A = \frac{a_1}{\|a_1\|} \cdot \|a_1\| \cdot 1 = U \Sigma V^*$$

(ii) $m=1$, $n\geq 1$ (A is a row vector)

Just use SVD for A^* (column vector)

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

$$\Rightarrow A = V\Sigma^* U^* \quad \text{(this is an SVD of } A\text{)}.$$

(iii) general m, n, $A=0$.

$$\underset{(U)}{A = I_{m\times m}} \; \underset{(\Sigma)}{0_{m\times n}} \; \underset{(V^*)}{I_{n\times n}}$$

Inductive step: assume $m,n \geq 2$

consider $A^*A \in \mathbb{C}^{n\times n}$

~~assume~~ A^*A is ~~spd~~ Hermitian semi-positive define.

(because A^*A is Hermitian and $R_{A^*A}(x) \geq 0$).

$$A^*A = V\Lambda V^*$$

$\max_j |\lambda_j|$

$\parallel$

Choose eigenvalue λ s.t. $\rho(A^*A) = \lambda$

Note: ~~$\lambda > 0$, becau~~

We can assume $\lambda > 0$ because otherwise $A=0$

Let v_1 be an(y) normalized eigenvector for λ.

$$A^*Av_1 = \lambda v_1 \qquad (v_1 \in \mathbb{C}^n)$$

Define $\sigma_1 = \sqrt{\lambda} > 0$.

Define $u_1 = \frac{Av_1}{\sigma_1}$

First: note that $\sigma_1 = \|A\|_2$ $(\star)$

$$R_{A^*A}(x) = \frac{\langle A^*Ax, x\rangle}{\langle x, x\rangle} = \frac{\|Ax\|_2^2}{\|x\|_2^2}$$

$$\Rightarrow \|A\|_2^2 = \sup_{x \neq 0} R_{A^*A}(x)$$

$$= \sup_{x\neq 0} \frac{\langle \Lambda V^*x, V^*x\rangle}{\langle V^*x, V^*x\rangle}$$

$$= \sup_{y \neq 0} \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} = \lambda = \sigma_1^2 \; (\star)$$

Now: "reduce" A:

$$U_1 = \begin{bmatrix} | & | & & | \\ u_1 & r_2 & \cdots & r_m \\ | & | & & | \end{bmatrix}$$

$\underbrace{r_2 \cdots r_m}_{\text{any ON completion of } \mathbb{C}^m}$

$$V_1 = \Big[\, v_1 \;\; \underbrace{s_2 \cdots s_n}_{\text{any}} \,\Big] \quad \text{ON completion of } \mathbb{C}^n.$$

U_1 and V_1 are unitary.

$$U_1^* A V_1 = U_1^* \begin{pmatrix} | & | & & | \\ Av_1 & As_2 & \cdots & As_n \\ | & | & & | \end{pmatrix}$$

$$= \begin{array}{c} \\ r_2^* \\ \\ r_m^* \end{array} \begin{pmatrix} u_1^* A v_1 & \text{—— } x \text{ ——} \\ \cancel{u_2}^* A v_1 & \\ \vdots & \boxed{\;\tilde{A}\;} \\ \cancel{u_m}^* A v_1 & \end{pmatrix}$$

$$j \neq 1: \quad \underset{r_j^*}{\cancel{u_j}^*} A v_1 = \underset{r_j^*}{\cancel{u_j}^*} u_1 \cdot \sigma_1 = 0$$

$$u_1^* A v_1 = u_1^* u_1 \cdot \sigma_1 = \sigma_1$$

Next: $x = 0$

Define $A_1 = U_1^* A V_1$, Let $w = \dfrac{1}{\sqrt{\sigma_1^2 + \|x\|_2^2}} \begin{pmatrix} \sigma_1 \\ | \\ x^* \\ | \end{pmatrix}$

$$\sigma_1 = \|A\|_2 = \|A_1\|_2 \geq \|A_1 w\|_2 = \sqrt{\left(\frac{\sigma_1^2 + \|x\|_2^2}{\sqrt{\sigma_1^2 + \|x\|_2^2}}\right)^2 + \cdots}$$

def'n of 2-norm

non-negative contributions.

$$\geq \sqrt{\sigma_1^2 + \|x\|_2^2}$$

$$\Rightarrow \sigma_1 \geq \sqrt{\sigma_1^2 + \|x\|_2^2} \Rightarrow \|x\| = 0 \Rightarrow x = 0$$

I.e., $A_1 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{A} \end{pmatrix}$

By inductive hypothesis, $\tilde{A}$ has an SVD, and we can use this to get SVD for A... ∎

The SVD

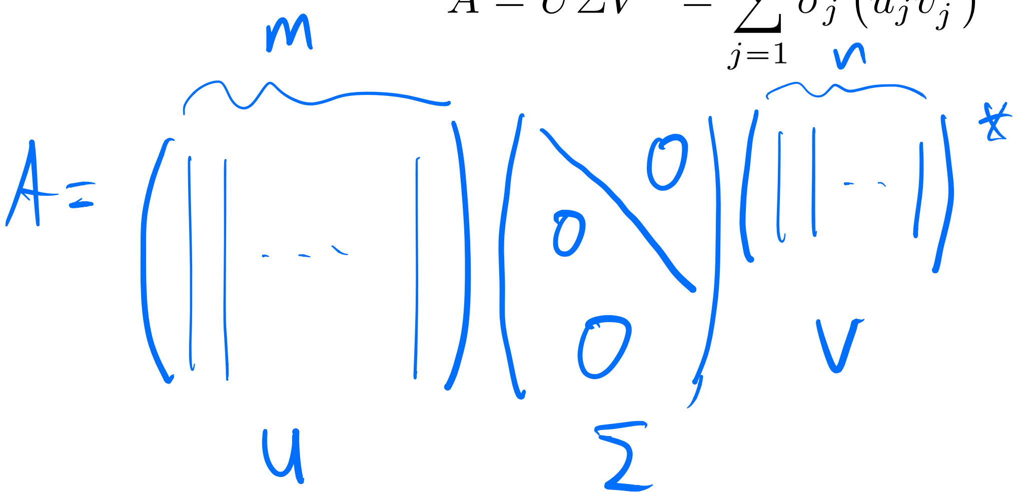
The SVD is arguably the most powerful matrix decomposition.

The SVD

min $\{m, n\}$

The SVD is arguably the most powerful matrix decomposition.

$$A = U\Sigma V^* = \sum_{j=1}^{p} \sigma_j \left(u_j v_j^*\right)$$



U, V unitary, elements of Σ are non-negative.

Define $r := \operatorname{rank}(A) \leq p$

The SVD : singular values are ordered: $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_p \geq 0$.

The SVD is arguably the most powerful matrix decomposition.

$$A = U\Sigma V^* = \sum_{j=1}^{p} \sigma_j \left(u_j v_j^*\right)$$

If $r = \mathrm{rank}(A)$:

$$\sigma_j = 0, \qquad\qquad j > r.$$

(if not: dim. range$(A) > r$)

The SVD

The SVD is arguably the most powerful matrix decomposition.

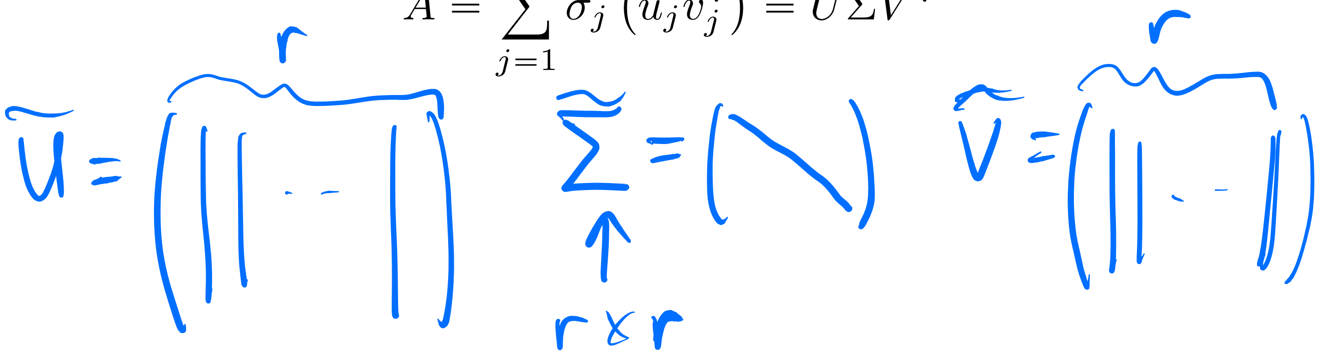
$$A = U\Sigma V^* = \sum_{j=1}^{p} \sigma_j \left(u_j v_j^*\right)$$

If $r = \operatorname{rank}(A)$:

$$\sigma_j = 0, \qquad\qquad j > r.$$

A has a *reduced* SVD:

$$A = \sum_{j=1}^{r} \sigma_j \left(u_j v_j^*\right) = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^*$$



The SVD

The SVD is arguably the most powerful matrix decomposition.

$$A = U\Sigma V^* = \sum_{j=1}^{p} \sigma_j \left(u_j v_j^*\right)$$

If $r = \mathrm{rank}(A)$:

$$\sigma_j = 0, \qquad\qquad j > r.$$

A has a *reduced* SVD:

$$A = \sum_{j=1}^{r} \sigma_j \left(u_j v_j^*\right) = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^*$$

$$\|A\|_2 = \sigma_1$$

$$\|A\|_2 = \sup_{\|x\|=1} \|Ax\|_2 = \sup_{\|x\|=1} \|U\Sigma V^* x\|_2$$

$$= \sup_{\|x\|=1} \|\Sigma \underbrace{V^* x}_{y}\|$$

$$= \sup_{\|y\|=1} \|\Sigma y\|_2 \longrightarrow \text{we did this before}$$

$$= \sigma_1$$

The SVD

The SVD is arguably the most powerful matrix decomposition.

$$A = U\Sigma V^* = \sum_{j=1}^{p} \sigma_j \left(u_j v_j^*\right)$$

If $r = \mathrm{rank}(A)$:

$$\sigma_j = 0, \qquad\qquad j > r.$$

A has a *reduced* SVD:

$$A = \sum_{j=1}^{r} \sigma_j \left(u_j v_j^*\right) = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^*$$

$$\|A\|_2 = \sigma_1$$

If A is square and invertible:

$$\|A^{-1}\|\|A\| =: \kappa(A) = \frac{\sigma_1}{\sigma_n}$$

If A is invertible: $\text{rank}(A) = n = m$

$\Downarrow$

Σ is invertible

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

is an SVD of A^{-1}

$$\Rightarrow \|A^{-1}\|_2 = \max_j \frac{1}{\sigma_j} = \frac{1}{\sigma_n}$$

U, Σ, V^* are also matrices related to A^*A, and AA^*

$A^*A \to$ Hermitian positive-definite.

$A^*A = V\Lambda V^*$ (☆)

$A = U\Sigma V^* \Rightarrow A^*A = V\underbrace{\Sigma^*\Sigma}_{n\times n,\text{ diagonal matrix}}V^*$ (☆)

$$\Sigma^*\Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}$$

$$\Longrightarrow \lambda_1(A^*A) = \sigma_1^2(A) \qquad (\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n)$$

$$\vdots$$

$$\lambda_n(A^*A) = \sigma_n^2(A)$$

eigenvectors V of $A^*A \equiv$ right-singular vectors V of A.

(Similar ~~a~~characterization for U, AA^*).