Last time: orthogonal diagonalization

$$\boxed{B = UDU^{-1} = UDU^{T}}$$

where D is diagonal and U is orthogonal $(U^{-1} = U^{T})$

Orthogonal diagonalization of B is possible $\Longleftrightarrow$ B is symmetric

$\Updownarrow$

- all eigenvalues of B are real
- all geometric mult = algebraic mult
- distinct eigenspaces are orthogonal

$\Longleftrightarrow$ $\exists$ orthonormal basis of B consisting of eigenvectors

Today: singular value decomposition

	arbitrary vectors	orthogonal vectors
particular case	diagonalization $B = PDP^{-1}$	orthogonal diagonalization $B = UDU^{-1}$, $U^{T} = U^{-1}$

(B square)	(Lecture 17)	(Lecture 25)
general case (B rectangular)	$B = P \Sigma Q^{-1}$ (Lecture 16) $\Sigma = \begin{pmatrix} 1 & & 0 & 0 \\ & \ddots & & \\ 0 & & 1 & \\ 0 & & & 0 \end{pmatrix}$	SVD $B = U \Sigma V^{-1}$, $U^T = U^{-1}$, $V^T = V^{-1}$ $\Sigma = \begin{pmatrix} \sigma_1 & & 0 & 0 \\ & \ddots & & \\ 0 & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}$, $r = rank(B)$

THM 26.1 : any $B \in \mathbb{R}^{m \times n}$ has a SVD (singular value decomposition)

$$\boxed{B = U \Sigma V^{-1}} = U \Sigma V^T$$

U and V are orthogonal matrices, i.e. $U^{-1} = U^T$, $V^{-1} = V^T$

$$\Sigma = \text{a diagonal } (m \times n) \text{ matrix} = \begin{pmatrix} \sigma_1 & & & 0 & 0 \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ 0 & & & & 0 \end{pmatrix}$$

singular values (unique > 0)

$U = (u_1 \ldots\ldots u_m) \in \mathbb{R}^{m \times m}$ — singular vectors (not unique)

$V = (v_1 \ldots\ldots v_n) \in \mathbb{R}^{n \times n}$ — singular vectors (not unique)

Suppose you have an SVD: $B = U\Sigma V^{-1} = U\Sigma V^T$

So what?

$$\Downarrow$$

$$B^T = V\Sigma^T U^{-1} = V\Sigma^T U^T$$

$$\Sigma = \begin{pmatrix} \begin{matrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \ 0 \end{matrix} & 0 \end{pmatrix} \rightsquigarrow \Sigma^T = \begin{pmatrix} \begin{matrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{matrix} \\ 0 \end{pmatrix}$$

Upshot: singular values of B = singular values of B^T

$$B = U\Sigma V^{-1} = U\Sigma V^T$$

$$\Downarrow$$

$$B^T = V\Sigma^T U^{-1} = V\Sigma^T U^T$$

$$\Rightarrow B^T B = V\Sigma^T \underbrace{U^{-1} U}_{I_m} \Sigma \overset{V^T}{\overset{\parallel}{V^{-1}}}$$

$$\boxed{\begin{array}{c} \text{even if } B \text{ is rectangular,} \\ B^T B \text{ is symmetric} \\ (B^TB)^T = B^T B^{TT} = B^T B \end{array}}$$

$$B^T B = V \underbrace{\Sigma^T \Sigma}_{D} V^{-1}$$

$$B^T B = V \overbrace{\Sigma^T \Sigma} V^T$$

orthogonal diagonalization of $B^T B$

$$\begin{array}{ccc} n\times n & n\times m & m\times n \\ D = & \Sigma^T & \Sigma \end{array} = \begin{pmatrix} \begin{matrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{matrix} \\ 0 \end{pmatrix} \begin{pmatrix} \begin{matrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \ 0 \end{matrix} & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & & & & & 0 \\ & \sigma_2^2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r^2 & & \\ & & & & 0 & \\ 0 & & & & & \ddots \ 0 \end{pmatrix}$$

Singular values $= \sqrt{\text{diagonal entries of } D} = \sqrt{\text{eigenvalues of } B^TB}$

Singular vectors = eigenvectors of B^TB → symmetric

(columns of V)

$\left(BB^T = U\Sigma\Sigma^TU^{-1} \Rightarrow \text{columns of } U \text{ are eigenvectors of } BB^T \right)$

Note: BB^T and B^TB have the same multisets of non-zero eigenvalues

$B = \begin{pmatrix} 2 & 3 \end{pmatrix}$ $\quad B^T = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ $\rightsquigarrow$ $BB^T = \begin{pmatrix} 13 \end{pmatrix}$

$B^TB = \begin{pmatrix} 4 & 6 \\ 6 & 9 \end{pmatrix}$ $\qquad \chi_{B^TB}(t) = \det\begin{pmatrix} t-4 & -6 \\ -6 & t-9 \end{pmatrix}$

$= (t-4)(t-9) - 36$

$= t^2 - 13t + 36 - 36$

$= t(t-13)$ $\rightsquigarrow$ eigenvalues $0, 13$

PROP 26.2: all eigenvalues of B^TB are ≥ 0

$B^TB v = \lambda v \quad | \; v^T \cdot$

$$B^T B v = \lambda v$$

$$v^T B^T B v = \lambda \, v^T v$$

$$(Bv)^T B v = \lambda \, v^T v$$

$$\|Bv\|^2 = \lambda \, \|v\|^2 \quad \Rightarrow \lambda \geq 0$$

$$\left(\lambda \text{ can be } 0 \text{, i.e. } B^T B v = 0 \overset{*}{\Longrightarrow} Bv = 0 \right)$$

Upshot: singular values are well-defined.

Proof of THM 26.1:

$B \rightsquigarrow B^T B \rightsquigarrow$ eigenvalues $\underbrace{\lambda_1, \dots, \lambda_r}_{\text{rank } r} > 0, \underbrace{0, 0, \dots, 0}_{n-r}$

$\boxed{\sigma_i =: \sqrt{\lambda_i}}$

give you Σ

eigenvectors $v_1, \dots, v_r, v_{r+1}, \dots, v_n$

i.e. $B^T B \, v_i = \begin{cases} \lambda_i v_i & \text{if } i \leq r \\ 0 & \text{if } i > r \end{cases}$

Moreover, $v_1, \dots, v_n$ are orthonormal

give you V

($B^T B$ is symm $\Rightarrow$ orthogonally diagonalizable)

Lemma: $Bv_1, \dots, Bv_r$ are (1) mutually orthogonal

(2) lengths $\sigma_1, \dots, \sigma_r$

(2) lengths $\sigma_1, \dots, \sigma_r$

(hence non-zero)

(1) $Bv_i \cdot Bv_j = (Bv_i)^T B v_j = v_i^T B^T B v_j = v_i^T \lambda_j v_j = \lambda_j \; v_i \cdot v_j = 0 \quad \forall \, i \neq j$

(2) $Bv_i \cdot Bv_i = (Bv_i)^T B v_i = v_i^T B^T B v_i = v_i^T \lambda_i v_i = \lambda_i \; v_i \cdot v_i = \lambda_i$

$$\Downarrow$$

$$\|Bv_i\|^2 = \lambda_i$$

$$\Downarrow$$

$$\|Bv_i\| = \sigma_i$$

Lemma: $Bv_1, \dots, Bv_r$ form a basis of $\mathrm{Col}(B)$

$Bv_1, \dots, Bv_r \in \mathrm{Col}(B)$

↳ linearly independent (*)

$$\mathrm{Col}(B) = \mathrm{span}(Bv_1, Bv_2, \dots, Bv_r, \underbrace{Bv_{r+1}}_{=0}, \dots \underbrace{Bv_n}_{=0})$$

(because $B^T B v_{r+1} = \dots = B^T B v_n = 0$ and (*))

$$= \mathrm{span}(Bv_1, \dots Bv_r, 0, \dots, 0)$$

$$= \mathrm{span}(Bv_1, \dots Bv_r) \quad (**)$$

(*) and (**) imply that $Bv_1, \dots, Bv_r$ is a basis of $\mathrm{Col}(B)$

(*) and (**) imply that ...

$$\Downarrow$$

$$\left(\mathrm{rank}(BB^T) = \mathrm{rank}(B^T) = \right) \mathrm{rank}(B) = r = \mathrm{rank}(B^TB)$$

B^TB is symmetry

$\square$

Look at $\mathrm{Col}(B) \subset \mathbb{R}^m$

$\cup$

$Bv_1, \dots, Bv_r$ mutually orthogonal, lengths $\sigma_1, \dots, \sigma_r$

Define $\boxed{u_1 = \frac{Bv_1}{\sigma_1}}, \dots, \boxed{u_r = \frac{Bv_r}{\sigma_r}}$ an orthonormal basis of $\mathrm{Col}(B)$

Complete to an orthonormal basis $u_1, \dots, u_m$ of $\mathbb{R}^m$

$$U := (u_1 \dots u_m)^{m \times m} \quad \text{orthogonal}$$

$$\Sigma := \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}^{m \times n} \quad \text{"diagonal"}$$

$$V := (v_1 \dots v_n)^{n \times n} \quad \text{orthogonal}$$

Claim: $B = U \Sigma V^T = U \Sigma V^{-1}$

$$\Updownarrow$$

$$BV = U\Sigma \quad \mathbb{R}^n \to \mathbb{R}^m$$

$$BV = U\Sigma : \mathbb{R} \to \mathbb{R}$$

enough to prove $BVe_i = U\Sigma e_i \quad \forall\, 1 \leq i \leq n$

$$Ve_i = v_i$$

because $\boxed{u_i = \frac{Bv_i}{\sigma_i}}$

$$BVe_i = Bv_i = \begin{cases} \sigma_i u_i & i \leq r \quad \text{by definition} \\ 0 & i > r \quad \text{by } (*) \end{cases} \qquad (*)$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}$$

, check $\Sigma e_1 = \sigma_1 e_1, \ldots, \Sigma e_r = \sigma_r e_r$
$\Sigma e_{r+1} = \ldots = \Sigma e_n = 0$

$$\Sigma e_i = \begin{cases} \sigma_i e_i & i \leq r \\ 0 & i > r \end{cases}$$

$$U\Sigma e_i = \begin{cases} \sigma_i u_i & i \leq r \\ 0 & i > r \end{cases} \qquad (**)$$

Because $(*) = (**)$, then $BVe_i = U\Sigma e_i,\ \forall i$

$\square$

Ex: $B = \begin{pmatrix} -1 & 1 \\ 1 & 2 \\ -3 & 2 \end{pmatrix}$; find the SVD of B

Step 1: $B^T B = \begin{pmatrix} -1 & 1 & -3 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 2 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 11 & -5 \\ -5 & 9 \end{pmatrix}$

Step 2: $\chi_{B^TB}(t) = \det \begin{pmatrix} t-11 & 5 \\ 5 & t-9 \end{pmatrix} =$

(eigenvalues)

$$= (t-11)(t-9) - 25 = t^2 - 20t + 74$$

$$\left.\begin{aligned} \lambda_1 &= 10 + \sqrt{100-74} = 10 + \sqrt{26} \\ \lambda_2 &= 10 - \sqrt{100-74} = 10 - \sqrt{26} \end{aligned}\right\} > 0$$

Step 3: (orthonormal eigenvectors)

$$v_1 \in \operatorname{Ker}\left(\begin{pmatrix} 11 & -5 \\ -5 & 9 \end{pmatrix} - \lambda_1 I_2 \right)$$

$$\in \operatorname{Ker} \begin{pmatrix} 1-\sqrt{26} & -5 \\ -5 & -1-\sqrt{26} \end{pmatrix} = \operatorname{span} \begin{pmatrix} x \\ y \end{pmatrix}$$

s.t. $(1-\sqrt{26})x - 5y = 0$

Choose $v_1 = \frac{1}{\sqrt{1+\left(\frac{1-\sqrt{26}}{5}\right)^2}} \begin{pmatrix} 1 \\ \frac{1-\sqrt{26}}{5} \end{pmatrix} = \frac{5}{\sqrt{52-2\sqrt{26}}} \begin{pmatrix} 1 \\ \frac{1-\sqrt{26}}{5} \end{pmatrix}$

$$v_2 \in \mathrm{Ker}\left(\begin{pmatrix} 11 & -5 \\ -5 & 9 \end{pmatrix} - \lambda_2 I_2\right)$$

$$\in \mathrm{Ker}\begin{pmatrix} 1+\sqrt{26} & -5 \\ -5 & -1+\sqrt{26} \end{pmatrix} = \mathrm{span}\begin{pmatrix} x \\ y \end{pmatrix}$$

s.t. $(1+\sqrt{26})x - 5y = 0$;

Choose $v_2 = \frac{5}{\sqrt{52+2\sqrt{26}}} \begin{pmatrix} 1 \\ \frac{1+\sqrt{26}}{5} \end{pmatrix}$

Sanity check: $v_1 \cdot v_2 = 1 \cdot 1 + \frac{(1-\sqrt{26})(1+\sqrt{26})}{25} = 0$

$v_1 \cdot v_1 = 1$

$v_2 \cdot v_2 = 1$

$$V = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$$

Step 4:
(singular values)

$$\sigma_1 = \sqrt{10+\sqrt{26}} \qquad \sigma_2 = \sqrt{10-\sqrt{26}} \rightsquigarrow \Sigma = \begin{pmatrix} \sqrt{10+\sqrt{26}} & 0 \\ 0 & \sqrt{10-\sqrt{26}} \\ 0 & 0 \end{pmatrix}$$

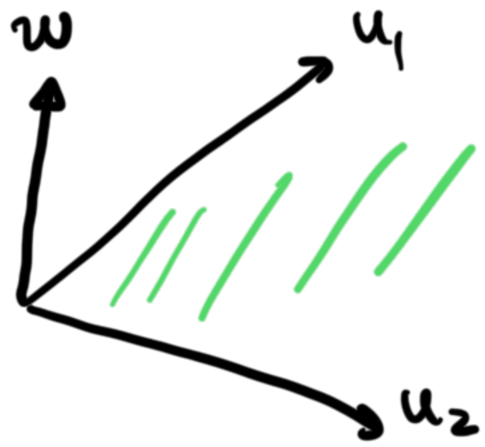
Step 5:
(find u_1, u_2)

$$u_1 = \frac{Bv_1}{\|Bv_1\|} = \frac{\begin{pmatrix} -1 & 1 \\ 1 & 2 \\ -3 & 2 \end{pmatrix} \frac{5}{\sqrt{52-2\sqrt{26}}} \begin{pmatrix} 1 \\ \frac{1-\sqrt{26}}{5} \end{pmatrix}}{\| \text{length of above vector in } \mathbb{R}^3 \|}$$

$$u_2 = \frac{Bv_2}{\|Bv_2\|} = \frac{\begin{pmatrix} -1 & 1 \\ 1 & 2 \\ -3 & 2 \end{pmatrix} \frac{5}{\sqrt{52+2\sqrt{26}}} \begin{pmatrix} 1 \\ \frac{1+\sqrt{26}}{5} \end{pmatrix}}{\| \text{length of above vector in } \mathbb{R}^3 \|}$$

Step 6:

(complete to an o.n.b. u_1, u_2, u_3 of $\mathbb{R}^3$)



pick a random vector w, e.g. $w = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$\rightarrow w \notin \operatorname{span}\{u_1, u_2\}$

Do G-S for $\{u_1, u_2, w\} \rightsquigarrow \{u_1, u_2, u_3\}$

$$U = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}$$

Upshot: we found U, Σ, V in $B = U \Sigma V^T$.