<from last time>

Thm (SVD): Let A be an $m \times n$ matrix of rank r. Then there exist $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and orthonormal sets $\{u_1, \ldots, u_r\}$ in $\mathbb{R}^m$ and $\{v_1, \ldots, v_r\}$ in $\mathbb{R}^n$ such that

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

$A^T = \sigma_1 v_1 u_1^T + \sigma_2 v_2 u_2^T + \cdots + \sigma_r v_r u_r^T$ is the SVD of A^T.

$$Ax = \sigma_1 (v_1 \cdot x) u_1 + \sigma_2 (v_2 \cdot x) u_2 + \cdots + \sigma_r (v_r \cdot x) u_r$$

$$\boxed{Av_i = \sigma_i u_i} \qquad \boxed{\|Av_i\| = \sigma_i}$$

Naive Schoolbook Procedure to Compute SVD:

Let A be an $m \times n$ matrix of rank r.

(1) Compute the nonzero eigenvalues of A^TA:
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ counted w/multiplicity
(automatically r of them, & they're positive)

(2) Compute orthonormal bases of eigenspaces of A^TA $\rightsquigarrow$ orthonormal set $\{v_1, \ldots, v_r\}$ s.t. $A^TAv_i = \lambda_i v_i$ for all i.

(3) Let $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{1}{\sigma_i} Av_i$. Then $\{u_1, \ldots, u_r\}$ is orthononormal and

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

<\ from>

NB: If A is a <u>wide</u> matrix $(n > m)$ then A^TA is $n \times n$ but AA^T is $m \times m$. So compute the SVD of A^T instead!

Eg: $A = \begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix}$

$$A^TA = \begin{pmatrix} 200 & -50 & 200 & -50 \\ -50 & 125 & -50 & 125 \\ 200 & -50 & 200 & -50 \\ -50 & 125 & -50 & 125 \end{pmatrix}$$ yikes!

compute SVD of A^T

$$AA^T = \begin{pmatrix} 400 & -100 \\ -100 & 250 \end{pmatrix} \qquad p(\lambda) = (\lambda - 450)(\lambda - 200)$$

$$\lambda_1 = 450 \Rightarrow \sigma_1 = \sqrt{450} = 15\sqrt{2} \quad v_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad u_1 = \frac{1}{\sigma_1} A^T v_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 200 \Rightarrow \sigma_2 = \sqrt{200} = 10\sqrt{2} \quad v_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad u_2 = \frac{1}{\sigma_2} A^T v_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

$$\Rightarrow A^T = 15\sqrt{2}\, u_1 v_1^T + 10\sqrt{2}\, u_2 v_2^T$$
$$\Rightarrow A = 15\sqrt{2}\, v_1 u_1^T + 10\sqrt{2}\, v_2 u_2^T$$

Why does this work?

Nonzero eigenvalues of A^TA are positive & there are r of them:

- A^TA is positive-semidefinite: if $A^TAx = \lambda x$, $x \neq 0$
  $\Rightarrow \|Ax\|^2 = (Ax)^T(Ax) = x^T(A^TAx) = x^T\lambda x = \lambda\|x\|^2$
  $\Rightarrow \lambda \geq 0$

- $\underset{n\times n}{\operatorname{rank}(A^TA)} = n - \dim \operatorname{Nul}(A^TA) \overset{\operatorname{Nul}(A^TA)=\operatorname{Nul}(A)}{=} n - \dim \operatorname{Nul}(A) = \underset{m\times n}{\operatorname{rank}(A)}$

- A^TA symmetric $\overset{\text{spectral thm}}{\Longrightarrow}$ diagonalizable over $\mathbb{R}$ $\Rightarrow$ all AM=GM
  $\Rightarrow 0$ is an eigenvalue of mult. $n - r = \dim \operatorname{Nul}(A^TA)$
  $\Rightarrow p(\lambda) = \pm(\lambda - \lambda_1)\cdots(\lambda - \lambda_r)\lambda^{n-r} \qquad \lambda_1, \ldots, \lambda_r \neq 0$
  $\Rightarrow$ there are r nonzero eigenvalues (with multiplicity).

Get $\{v_1, \ldots, v_r\}$ = orthonormal eigenvectors s.t. $A^TAv_i = \lambda_i v_i$.
Set $\sigma_i = \sqrt{\lambda_i}$, $u_i = \frac{1}{\sigma_i} A v_i$.

Orthonormality of $\{u_1, \ldots, u_r\}$:

$$u_i \cdot u_j = \left(\frac{1}{\sigma_i} A v_i\right)^T \left(\frac{1}{\sigma_j} A v_j\right) = \frac{1}{\sigma_i \sigma_j} v_i^T (A^T A v_j)$$
$$= \frac{1}{\sigma_i \sigma_j} v_i^T (\sigma_j^2 v_j) = \frac{\sigma_j}{\sigma_i} \underbrace{v_i \cdot v_j}_{\text{orthonormality of } \{v_1, \ldots, v_r\}} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \checkmark$$

Outer Product Formula:

Let $A' = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$. We'll show $A' = A$ by showing $A'v_i = Av_i$ for $\{v_1, \ldots, v_n\}$ a basis of $\mathbb{R}^n$.
(This means $v_1, \ldots, v_n \in \mathrm{Nul}(A - A') \Rightarrow \mathrm{Nul}(A - A') = \mathbb{R}^n \Rightarrow A - A' = 0$)
Let $\{\underbrace{v_{r+1}, \ldots, v_n}_{n-r \text{ vectors}}\}$ be an o.n. basis for $\mathrm{Nul}(A) = \mathrm{Nul}(A^TA)$
= 0-eigenspc of A^TA
$\Rightarrow \{v_1, \ldots, v_n\}$ is an o.n. eigenbasis of A^TA.
$\Rightarrow \{v_1, \ldots, v_n\}$ is a basis of $\mathbb{R}^n$.

For $i \leq r$,
$$A'v_i = \sigma_1 (\cancelto{0}{v_1 \cdot v_i}) u_1 + \cdots + \sigma_i (\cancelto{1}{v_i \cdot v_i}) u_i + \cdots + \sigma_r (\cancelto{0}{v_r \cdot v_i}) u_r$$
$$= \sigma_i u_i = Av_i \quad \text{b/c } u_i = \frac{1}{\sigma_i} A v_i \text{ by definition.}$$

For $i > r$,
$A'v_i = 0$ because $v_j \cdot v_i = 0$ for $j \leq i$.
$Av_i = 0$ because $v_i \in \mathrm{Nul}(A)$. $\checkmark$

This justifies the SVD!

SVD: Matrix Form

NB: $\{u_1,\dots,u_r\}$ is an o.n. basis for $\mathrm{Col}(A)$:

- $u_i = \frac{1}{\sigma_i} A v_i \in \mathrm{Col}(A)$ • $\{u_1,\dots,u_r\}$ LI ; $\dim \mathrm{Col}(A) = r$

NB: $\{v_1,\dots,v_r\}$ is an o.n. basis for $\mathrm{Row}(A)$:

- replace A by A^T ↑

Let $\{u_{r+1},\dots,u_m\}$ be an o.n. basis for $\mathrm{Nul}(A^T) = \mathrm{Nul}(AA^T)$

$\implies \{u_1,\dots,u_m\}$ is an o.n. eigenbasis for AA^T

Let $\{v_{r+1},\dots,v_n\}$ be an o.n. basis for $\mathrm{Nul}(A) = \mathrm{Nul}(A^TA)$

$\implies \{v_1,\dots,v_n\}$ is an o.n. eigenbasis for A^TA.

Thm (SVD): Let A be an $m\times n$ matrix of rank r. Then

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

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix} \quad m\times m \text{ orthogonal} \qquad V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \quad n\times n \text{ orthogonal}$$

same size as A → $\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}$ $m\times n$ diagonal

Check: $U\Sigma V^T v_i = U\Sigma e_i = \begin{cases} 0 & \text{for } i > r \\ U\sigma_i e_i & \text{for } i \le r \end{cases} = \begin{cases} 0 \\ \sigma_i u_i \end{cases} = A v_i$

orthogonal diagonalizations

Compare: $A^TA = V\begin{pmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & 0 \end{pmatrix}V^T \qquad AA^T = U\begin{pmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & 0 \end{pmatrix}U^T$

Combine both: "diagonalize" $A = U\begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix}V^T$

SVD & the Four Subspaces:

$$\begin{array}{cc} \text{o.n. basis for } Col(A) & \text{o.n. basis for } Nul(A^T) \end{array}$$

$$U = \begin{pmatrix} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{pmatrix}$$

$$A^T u_i = \sigma_i v_i \quad \downarrow \uparrow \quad A v_i = \sigma_i u_i$$

$$\underbrace{u_{r+1}, \dots, u_m}_{} \xrightarrow{A^T} 0 \qquad \underbrace{v_{r+1}, \dots, v_n}_{} \xrightarrow{A} 0$$

$$V = \begin{pmatrix} | & & | & | & & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & & | & | & & | \end{pmatrix}$$

$$\begin{array}{cc} \text{o.n. basis for } Row(A) & \text{o.n. basis for } Nul(A) \end{array}$$

NB: SVD of A^T is $A^T = V\Sigma^T U^T$

Procedure for $A = U\Sigma V^T$:

(1) Compute σ_i, u_i, v_i as for outer product version

(2) Compute $\{u_{r+1}, \dots, u_m\}$ o.n. basis for $Nul(A^T)$

$\{v_{r+1}, \dots, v_n\}$ o.n. basis for $Nul(A)$

(do elimination & run Gram-Schmidt)

(3) $U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix}$ $V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ $\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}$

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