

Lesson 20

Read Ch 8

The matrix $A^T A$

SVD existence th and calculations

Recall: given $A \in \mathbb{R}^{m \times n}$ of rank r

$$\text{SVD of } A \text{ is } A = U \Sigma V^T$$

• orthonormal basis for $\text{Null}(A^T)$

$$U = \begin{bmatrix} \underbrace{u_1 \dots u_r}_{\text{orthonormal basis for } \text{Col}(A)} & \underbrace{u_{r+1} \dots u_m}_{\text{orthonormal basis for } \text{Null}(A^T)} \end{bmatrix}$$

orthonormal basis for $\text{Col}(A)$

left singular vectors U is orthogonal

$$\Sigma = \begin{bmatrix} \sigma_1 & \dots & 0 \\ & \sigma_r & \\ 0 & & 0_{..} \end{bmatrix}$$

$m \times n$

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

singular values

orthonormal basis for $\text{Null}(A)$

$$V = \begin{bmatrix} v_1 \dots v_r & \underbrace{v_{r+1} \dots v_n}_{\text{orthonormal basis for } \text{Null}(A)} \end{bmatrix}$$

orthonormal basis for $\text{row}(A)$

right singular vectors

Note:

$$\Sigma^T \Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r & & 0 \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix}_{n \times m} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r & & 0 \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix}_{m \times n} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_r^2 & \dots \\ 0 & \dots & \dots & 0 \end{bmatrix}$$

Proof:

$$\text{Let } \Sigma = (a_{lj}) \text{ , } a_{lj} = \begin{cases} \sigma_l & \text{if } l=j \leq r \\ 0 & \text{otherwise} \end{cases}$$

$$\Sigma^T = (c_{lj}) \quad c_{lj} = a_{jl}$$

$$\Sigma^T \Sigma = (b_{lj})$$

$$b_{ll} = \sum_{k=1}^n c_{lk} a_{kl} = \sum_{k=1}^n a_{kl} a_{kl} = a_{ll}^2$$

$$l \neq j \quad b_{lj} = \sum_{k=1}^n c_{lk} a_{kj} = \sum_{k=1}^n a_{kl} a_{kj} = 0 \quad \text{because}$$

$$a_{kl} = 0 \quad \text{if } k \neq l \quad a_{kj} = 0 \quad \text{if } k \neq j$$

$$\text{so } \Sigma^T \Sigma = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_r^2 & & 0 \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix}$$

How do we find the σ , U , V ?

Idea:

Suppose $A = U \Sigma V^T$ then

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

$n \times n$ $n \times n$ $n \times m$ $m \times n$ $n \times n$ $n \times n$

$$= V \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & \ddots & \\ & & & & & 0 \end{bmatrix} V^T \quad (*)$$

diagonalization of $A^T A$

σ_l^2 eigenvalues of $A^T A$

right singular vectors $v_l =$ eigenvectors of $A^T A$

To find left singular values use

$$A V = \begin{bmatrix} u_1 & \dots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

$$\frac{1}{\sigma_l} A v_l = u_l \quad \text{for } l=1 \dots r$$

Complete u_1, \dots, u_r with basis of $\text{Null}(A^T)$

Note that we have assumed here, to start with, what we were trying to prove i.e. that A has an SVD, and this is not OK in a proof. The argument in the previous page just gives us an idea of how a careful proof of the existence of SVD could go:

- 1) Consider $A^T A$ show it is orthogonally diagonalizable and has non negative eigenvalues $d_1 \geq d_2 \geq \dots \geq d_r > 0 \dots 0$. Let $\sigma_L = \sqrt{d_L}$
- 2) Let v_1, \dots, v_n be the eigenvectors for d_1, \dots, d_n then I can choose them to be orthonormal, v_1, \dots, v_r are a basis for $\text{row}(A)$ v_{r+1}, \dots, v_n are a basis for $\text{Null}(A)$
- 3) Let $u_L = \frac{1}{\sigma_L} A v_L$ for $L=1 \dots r$ then

u_1, \dots, u_r are an orthonormal basis for $\text{col}(A)$ which can be completed with a basis for $\text{Null}(A^T)$ to form an orthonormal basis for \mathbb{R}^m

4) with these choices $A = U \Sigma V^T$

Th: If $A \in \mathbb{R}^{m \times n}$ has rank r
 $A^T A \in \mathbb{R}^{n \times n}$, is PSD, of rank r
 $n \times n$

Proof 1) $(A^T A)^T = A^T A$ so $A^T A$ is symmetric

2) it is PSD because is the product
of the transpose of a matrix and
the matrix

3) $\text{rank}(A^T A) = r$:

We will show $\text{Null}(A) = \text{Null}(A^T A)$ ($\subseteq \mathbb{R}^n$)

so $\text{rank}(A^T A) = n - \text{nullity}(A^T A) = n - \text{nullity}(A)$

$= n - (n - r) = r$ by rank-nullity theorem

applied to $A^T A$ and A .

1) suppose $y \in \text{Null}(A)$ then $A^T A y = 0$ so $y \in \text{Null}(A^T A)$

2) suppose $y \in \text{Null}(A^T A)$ then $y^T A^T A y = 0$ so

$\|A y\|^2 = 0$ so $A y = 0$ so $y \in \text{Null}(A)$

This tells us that If we look at the eigenvalues

of $A^T A$ in descending order they are

$d_1 \dots d_r \ 0 \dots 0$

Given $A \in \mathbb{R}^{m \times n}$ of rank r to compute SVD:

1) Diagonalize $A^T A = V \begin{pmatrix} d_1 & & & \\ & \ddots & & \\ & & d_r & \\ & & & 0 \dots 0 \end{pmatrix} V^T \quad d_1 \geq d_2 \geq \dots \geq d_r$

2) $\sigma_L = \sqrt{d_L}$ are the singular values $L=1 \dots r$

3) The eigenvectors of $A^T A$ (columns of V) are the right singular vectors

4) $u_L = \frac{1}{\sigma_L} A v_L$ for $L=1 \dots r$

$u_{r+1} \dots u_m$ are an orthonormal basis for $\text{Null}(A^T)$

1

Ex What is the SVD of $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$? $r=1$

1) $A^T A = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix}$ has eigenvalues 15, 0

$$E_{15} = \text{span} \left(\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \right) \quad E_0 = \text{span} \left(\begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \right)$$

\parallel v_1 \parallel v_2

$$\sigma_1 = \sqrt{15}$$

$$U_1 = \frac{1}{\sqrt{15}} A v_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Now we need u_2, u_3 : orthonormal basis for $\text{Null}(A^T)$

$$A^T = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \quad \text{Null}(A^T) \text{ has dimension } 2$$

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \text{Null}(A^T) \quad \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} = u_2 \text{ has length } 1$$

I want to find another vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$$\begin{aligned} \text{s.t. } & x+y+z=0 \quad (\in \text{Null}(A^T)) \\ & \frac{x}{\sqrt{2}} - \frac{z}{\sqrt{2}} = 0 \quad (\perp \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}) \quad \text{So } 2x+y=0 \end{aligned}$$

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ would work} \quad u_3 = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \text{ has length } 1$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{15} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$$

svd (B)

Julia output

```
SVD{Float64,Float64,Array{Float64,2}}
U factor:
33 Array{Float64,2}:
(-0.30519  0.757315 -0.57735
 -0.503259 -0.64296 -0.57735
 -0.808449  0.114355  0.57735) = U
singular values:
3-element Array{Float64,1}:
3.904484344750072  σ₁
1.6598198702273692  σ₂
-0.0  σ₃
B has rank 2
Vt factor:
33 Array{Float64,2}:
(0.0507284 -0.285221 -0.957119
 0.84363  0.525159 -0.111784
 0.534522 -0.801784  0.267261) = Vᵀ
```

Truncating the above numbers at 5 decimal places, the singular value decomposition of B is

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} -0.30519 & 0.75731 & -0.57735 \\ -0.50325 & -0.64296 & -0.57735 \\ -0.80844 & 0.11435 & 0.57735 \end{bmatrix} \begin{bmatrix} 3.90448 & 0 \\ 0 & 1.65981 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.05072 & -0.28522 & -0.95711 \\ 0.84363 & 0.52515 & -0.11178 \\ 0.53452 & -0.80178 & 0.26726 \end{bmatrix}$$

Notice that B has only two singular values and $\text{rank}(B) = 2$.

Complexity of computing SVD of $A \in \mathbb{R}^{m \times n}$ $\Theta(m \cdot n \cdot \min(m, n))$

Ex: What is the SVD of $A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$?

$$A^T A = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} = \sigma_1 = 3 \quad \sigma_2 = 2$$

$$E_9 \text{ for } A^T A \text{ is span} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

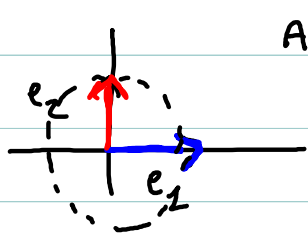
$$E_4 \text{ for } A^T A \text{ is span} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$u_1 = \frac{1}{3} A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

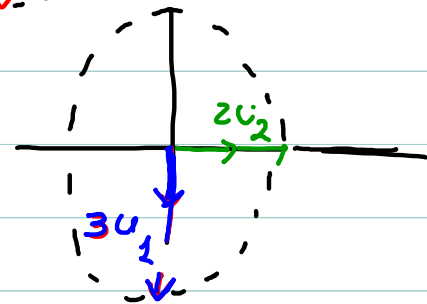
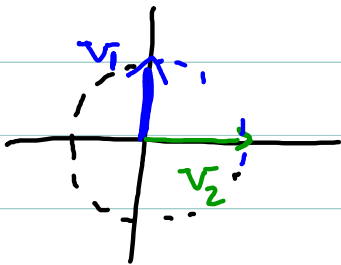
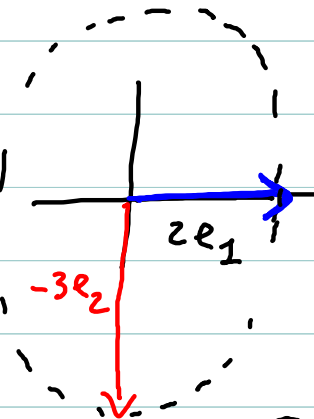
$$u_2 = \frac{1}{2} A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Geometric interpretation: $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ $T(w) = Aw$



$$A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$$



$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$Av_1 = 3u_1$$

$$Av_2 = 2u_2$$