NYU

Lecture 20

Singular Value Decomposition

Dr. Ralph Chikhany

Quiz 9

Q7. $A = \begin{bmatrix} 2 & ! & 4 \\ * & 5 & @ \\ \% & \# & 2 \end{bmatrix}$ $\lambda_1 = \lambda_2 = 6$ what is λ_3

$\underbrace{trace(A)}_{2+5+2} = \lambda_1 + \lambda_2 + \lambda_3 = 6+6+\lambda_3$

$\det(A) = \prod \lambda_i$ reason. $A = X^{-1} \Lambda X$, $\Lambda = diag([\lambda_1\ \lambda_2\ \lambda_3])$

$\det(\Lambda) = \lambda_1 \lambda_2 \cdots \lambda_n$

$\det(A) = \det(X^{-1}) \det(\Lambda) \det(X)$

$= \det(\Lambda)$

$= \lambda_1 \cdots \lambda_n$

$\det(X^{-1})\det(X) = 1$
$\parallel$
$\det(X^{-1}X)$
$\parallel$
$\det(I)$

B is a 3x3 matrix.

v_1, v_2 are eigenvalue -2 $\rightarrow$ $c_1 v_1 + c_2 v_2$ is still eigenvector of B — vector space.

w are eigenvalue -3

$4w$. $B(4w) = 4Bw$
$= 4(3w)$
$= 12w = 3(4w)$

Remark. for eigenvalue λ

all eigenvectors is $Nul(A - \lambda I)$ which is a vector space.

---

1. all nxn matrix have n eigenvalues.

- This is because. $p(\lambda) = \det(A - \lambda I)$ is a n-th order polynomial.

But it doesn't mean we have "n" - eigenvectors, (counting the basis)

n- linear independent eigenvectors.

- If we have "n" eigen-vectors, $\vec{x}_1 \cdots \vec{x}_n$. $X = [\vec{x}_1 \cdots \vec{x}_n]$

$A = X^{-1} \Lambda X$ $\Downarrow$ matrix A can be diagonalized.

- If $\lambda_1, \lambda_2 \cdots \lambda_n$ they are different from each other.

$\downarrow$ $x_1 \cdots x_2 \cdots x_n$ are linear independent!

- If A can't be diagonalized. means A have two repeated eigenvalues.

Example $B=\begin{bmatrix}1 & 1\\ & 1\end{bmatrix}$

$p(\lambda)=\det(B-\lambda I)$

$=\det\begin{bmatrix}1-\lambda & 1\\ & 1-\lambda\end{bmatrix}$ (This upper diag)

$=(1-\lambda)^2 \rightarrow \underbrace{\lambda_1=1,\ \lambda_2=1}_{\text{repeated eigenvalue}}$

$C=\begin{bmatrix}1 & 1 & \\ & 1 & \\ & & 2\end{bmatrix} \rightarrow \underbrace{\lambda_1=1,\ \lambda_2=1}_{\text{repeated eigenvalue}},\ \lambda_3=2$

Why we can't diagonalize B and C

$B-I=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} \rightarrow$ rank is 1, $\dim(\mathrm{Nul}(B-I))=1$

$\Rightarrow$ only "1" eigen vector $\rightarrow$ counting the basis but not 2

$C-I=\begin{bmatrix}0 & 1 & \\ & 0 & \\ & & 1\end{bmatrix} \rightarrow$ rank is 2 $\rightarrow \dim(\mathrm{Nul}(C-I))=1$

$C-2I=\begin{bmatrix}-1 & 1 & \\ & -1 & \\ & & 0\end{bmatrix} \rightarrow$ rank is 2 $\rightarrow \dim(\mathrm{Nul}(C-I))=1$

$\rightarrow$ "2" eigen vector. but not 3

Matrix C $\lambda_1=\lambda_2=1$, $\lambda_3=2$

) I want eigenvalue repeated 2-times can provide 2-eigen vectors!

$\begin{bmatrix}1 & & \\ & 1 & \\ & & 2\end{bmatrix} \rightarrow$ 2 eigenvectors of eigenvalue 1

$\hookrightarrow$ 1 eigenvector of eigenvalue 2.

"Diagonalize" means repeat times = "number" of eigen vectors $\rightarrow$ count the basis.

but for C matrix

repeat times > number of eigen vectors

$\lambda_1=\lambda_2=1$ one eigen vectors.

Symmetric matrix $A = A^T \rightarrow \vec{x}^T A \vec{x}$ is qudratic function respect to $x_1 \cdots x_n$

$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $x_1 \rightarrow$ real number, $x_n \rightarrow$ real number.

- A can be diagnolized !
- $x_1 \cdots x_n$ (eigenvector) orthogonal) $\Rightarrow A = X \Lambda X^{-1} = X \Lambda X^T$

X is orthogonal.

$$A = \lambda_1 \underbrace{x_1 x_1^T}_{\text{rank } 1} + \lambda_2 \underbrace{x_2 x_2^T}_{\text{rank } 1} + \cdots + \lambda_n \underbrace{x_n x_n^T}_{\text{rank } 1}$$

ex. $x_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \rightarrow x_1 x_1^T = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \leftarrow$ rank is 1

$\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

$$\vec{a}^T A \vec{a} = \lambda_1 (x_1^T \vec{a})^2 + \lambda_2 (x_2^T \vec{a})^2 + \cdots + \lambda_n (x_n^T \vec{a})^2$$

Qudratic function respect to $a_1 \cdots a_n$

ex. $x_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ Then $(x_1^T a)^2 = (\underbrace{1/\sqrt{2}\, a_1 + 1/\sqrt{2}\, a_2}_{\text{Simple}})^2$. square function!

Positive. Definite Matrix.

For matrix $A = A^T$, $\lambda_1 \cdots \lambda_n > 0$

$$\Leftrightarrow a^T A a = \Sigma\, \lambda_i\, (x_i^T a)^2 \rightarrow a^T A a \geq 0$$

λ_i: positive coef; $(x_i^T a)^2$: square function ≥ 0

Remark. A is positive definite.

$$\vec{a}^T A \vec{a} = 0 \Leftrightarrow \vec{a} = \vec{0}$$

Strang Chapter 7 – The Singular Value Decomposition

Course notes adapted from N. Hammoud's NYU lecture notes.

SVD

SVD → holds for all matrix

↳ holds for rectgular Matrix

$$\underset{m\times n}{A} = \underset{m\times m}{U}\underset{m\times n}{\Sigma}V^{T} \rightarrow n\times n$$

orthogonal matrix

$U^{-1} = U^{T}$

"diagonal" matrix

orthogonal matrix

$V^{-1} = V^{T}$

$$A^{T}A = (U\Sigma V^{T})^{T}(U\Sigma V^{T})$$
$$= V\Sigma^{T}U^{T}U\Sigma V^{T}$$
$$= V\Sigma^{T}\Sigma V^{T}$$

eigenvectors of $A^{T}A$

eigenvalues of $A^{T}A$

$$AA^{T} = (U\Sigma V^{T})(U\Sigma V^{T})^{T}$$
$$= U\Sigma V^{T}V\Sigma^{T}U^{T}$$
$$= U\Sigma\Sigma^{T}U^{T}$$

eigenvectors of AA^{T}

eigenvalues of AA^{T}

4x4

should be zero

size!!

2x2

2x4

Example

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \underset{2\times 2}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \underset{2\times 4}{\begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix}} \underset{4\times 4}{\begin{bmatrix} \sqrt{2}/2 & 0 & \sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ -\sqrt{2}/2 & 0 & \sqrt{2}/2 & 0 \\ 0 & -\sqrt{2}/2 & 0 & \sqrt{2}/2 \end{bmatrix}}$$

SVD.

$$A = U \Sigma V^T.$$ ↙ don't forget the Transpose here.

Key idea. Even if $\overset{m\times n}{A}$ is rectangular.

$\underbrace{\overset{n\times m}{A^T}\,\overset{m\times n}{A}}_{n\times n}$ and $\underbrace{\overset{m\times n}{A}\,\overset{n\times m}{A^T}}_{m\times m}$ are square and symmetric.

$$(A^T A)^T = A^T (A^T)^T = A^T A. \Rightarrow A^T A \text{ is symmetric!!}$$

$(AB)^T = B^T A^T$

Assume we have $A = U\Sigma V^T$ (SVD of my A).

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

This indentity matrix for U is orthogonal.

$$= ((V^T)^T \Sigma^T U^T)(U \Sigma V^T)$$

$(ABC)^T = C^T B^T A^T$

$$= \underbrace{(V)}_{\text{orthogonal}} \underbrace{\Sigma^T\Sigma}_{\text{diag}} \underbrace{(V^T)}_{\text{orthonogonal matrix Transpose.}} \rightarrow A^T A!!$$

diagonalization of symmetric matrix

V is $[v_1 \cdots v_n]$, $v_1 \cdots v_n$ is eigenvector of $A^T A$

$$A A^T = (U) \Sigma \Sigma^T (U^T)$$

↳ diagonalization of symmetric matrix AA^T.

$U = [u_1 \cdots u_n]$. $u_1 \cdots u_n$ is eigenvector of AA^T.

Size!!

Example $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{2\times 2} \overset{2\times 4}{\begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix}} \overset{4\times 4}{\begin{bmatrix} \sqrt{2}/2 & 0 & \sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ -\sqrt{2}/2 & 0 & \sqrt{2}/2 & 0 \\ 0 & -\sqrt{2}/2 & 0 & \sqrt{2}/2 \end{bmatrix}}$

should be zero

Size!!

Example $A = \begin{bmatrix}1&0&1&0\\0&1&0&1\end{bmatrix} = \underbrace{\overset{2\times2}{\begin{bmatrix}1&0\\0&1\end{bmatrix}}}_{U}\underbrace{\overset{2\times4}{\begin{bmatrix}\sqrt2&0&0&0\\0&\sqrt2&0&0\end{bmatrix}}}_{\Sigma}\underbrace{\overset{4\times4}{\begin{bmatrix}\sqrt2/2&0&\sqrt2/2&0\\0&\sqrt2/2&0&\sqrt2/2\\-\sqrt2/2&0&\sqrt2/2&0\\0&-\sqrt2/2&0&\sqrt2/2\end{bmatrix}}}_{V^T}$

should be zero

① What is the size of $\overset{4\times2}{A^T}\,\overset{2\times4}{A}$, $\overset{4\times2}{\Sigma^T}\,\overset{2\times4}{\Sigma} \to 4\times4$

$\overset{2\times4}{A}\,\overset{4\times2}{A^T}$, $\overset{2\times4}{\Sigma}\,\overset{4\times2}{\Sigma^T} \to 2\times2$

A^TA, $\Sigma^T\Sigma$ have same eigenvalues because they are similar

$$\Sigma^T\Sigma = \begin{bmatrix}\sqrt2&0\\0&\sqrt2\\0&0\\0&0\end{bmatrix}\begin{bmatrix}\sqrt2&0&0&0\\0&\sqrt2&0&0\end{bmatrix} = \begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

A^TA have eigenvalue $\lambda_1^2, \lambda_2^2, 0, 0$

$$\Sigma\Sigma^T = \begin{bmatrix}\sqrt2&0&0&0\\0&\sqrt2&0&0\end{bmatrix}\begin{bmatrix}\sqrt2&0\\0&\sqrt2\\0&0\\0&0\end{bmatrix} = \begin{bmatrix}2&0\\0&2\end{bmatrix}$$

AA^T have eigenvalue λ_1^2, λ_2^2

If my $\Sigma = \begin{bmatrix}\lambda_1&0&0&0\\0&\lambda_2&0&0\end{bmatrix}$

$\Rightarrow$ Eigenvalue of $\Sigma^T\Sigma$: $\lambda_1^2, \lambda_2^2, 0, 0$

$\Sigma\Sigma^T$: λ_1^2, λ_2^2

the same

$4-2=2$ 0 eigenvalues

4 eigen values ↑ 4×4 matrix

2×2 matrix ↓ 2 eigen values

---

a new understanding of $\text{rank}(A^TA) = \text{rank}(AA^T) = \text{rank}(\Sigma) = \text{rank}(A)$

① Recall if X, Y are invertible matrix

$$\text{rank}(AX) = \text{rank}(YA) = \text{rank}(A)$$

We know $\text{rank}(AB) \le \text{rank}(A)$ $\quad \text{rank}(AB) \le \text{rank}(B)$

$\Rightarrow \text{rank}(AX) \le \text{rank}(A)$

$\text{rank}(A) = \text{rank}(AXX^{-1}) \le \text{rank}(AX)$ $\Big\} \Rightarrow \text{rank}(A) = \text{rank}(AX)$

$A = U \Sigma V^T$

mxm orthogonal

nxn orthogonal.

U. V^T are both invertible.

$\Rightarrow \text{rank}(A) = \text{rank}(\Sigma) = \text{rank}(\Sigma^T \Sigma) = \text{rank}(A^T A)$

$= \text{rank}(\Sigma^T \Sigma) = \text{rank}(A A^T)$

number of non-zero in the digonal of Σ

Example

Suppose: $A \in M_{5\times 10}$; $rk(A) = 2$

Then:

$rk(\underbrace{AA^T}_{5\times 5}) = 2$ $\qquad$ $rk(\underbrace{A^TA}_{10\times 10}) = 2$

2 nonzero e-values of AA^T $\qquad$ 2 nonzero e-values of A^TA

$\Rightarrow$ 3 zero e-values (counting multiplicities) $\qquad$ $\Rightarrow$ 8 zero e-values (counting multiplicities)

Ignoring multiplicities, the e-values of AA^T and A^TA are the same

The Eigenvalues of A^TA and AA^T

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

Therefore, the matrices A^TA and AA^T have the same nonzero eigenvalues.

$\sigma_1 = \sqrt{\lambda_1}$, where λ_1 is the largest eigenvalue of A^TA (or AA^T)

$\vdots$

$\sigma_r = \sqrt{\lambda_r}$, where λ_r is the smallest (nonzero) eigenvalue of A^TA (or AA^T)

Example

The Connection between *U* and *V*

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

$$A\vec{v}_1 = \sigma_1 \vec{u}_1 \implies \vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1$$

$$A\vec{v}_2 = \sigma_1 \vec{u}_2 \implies \vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2$$

$$\vdots$$

$$A\vec{v}_r = \sigma_1 \vec{u}_r \implies \vec{u}_r = \frac{1}{\sigma_r} A\vec{v}_r$$

The remaining $\vec{u}_{r+1}, \ldots, \vec{u}_m$ are determined as eigenvectors of AA^T corresponding to zero eigenvalue.

How to find the SVD Decomposition?

1. Σ is the same size as A and has the singular values σ on its diagonal. The singular values are the square roots of the nonzero eigenvalues of $A^T A$ or AA^T.

2. V contains the eigenvectors of $A^T A$.

3. U contains the eigenvectors of AA^T. But we will find the vectors $\vec{u}$ corresponding to nonzero eigenvalues using the SVD.

Example

Bases of the Four Fundamental Subspaces using the SVD

The Matrix V

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

$$\left.\begin{aligned} A^T A\vec{v}_1 &= \sigma_1^2 \vec{v}_1 \\ A^T A\vec{v}_2 &= \sigma_2^2 \vec{v}_2 \\ &\vdots \\ A^T A\vec{v}_r &= \sigma_r^2 \vec{v}_r \end{aligned}\right\} \begin{array}{l} \{\vec{v}_1, \ldots, \vec{v}_r\} \perp \{\vec{v}_{r+1}, \ldots, \vec{v}_n\} \\ \{\vec{v}_1, \ldots, \vec{v}_r\} \text{ orthonormal basis for Row} A \end{array}$$

$$\left.\begin{aligned} A^T A\vec{v}_{r+1} &= \sigma_{r+1}^2 \vec{v}_{r+1} = \vec{0} \\ &\vdots \\ A^T A\vec{v}_n &= \sigma_n^2 \vec{v}_n = \vec{0} \end{aligned}\right\} \begin{array}{l} \vec{v}_{r+1}, \ldots, \vec{v}_n \text{ span } \text{Nul}(A^T A) = \text{Nul} A \\ \{\vec{v}_{r+1}, \ldots, \vec{v}_n\} \text{ orthonormal basis for Nul} A \end{array}$$

The Matrix U

$$AA^T = U\Sigma\Sigma^T U^T$$

$$AA^T\vec{u}_1 = \sigma_1^2\vec{u}_1$$

$$AA^T\vec{u}_2 = \sigma_2^2\vec{u}_2$$

$$\vdots$$

$$AA^T\vec{u}_r = \sigma_r^2\vec{u}_r$$

$\{\vec{u}_1, \ldots, \vec{u}_r\} \perp \{\vec{u}_{r+1}, \ldots, \vec{u}_m\}$

$\{\vec{u}_1, \ldots, \vec{u}_r\}$ orthonormal basis for Col$A$

$$AA^T\vec{u}_{r+1} = \sigma_{r+1}^2\vec{u}_{r+1} = \vec{0}$$

$$\vdots$$

$$AA^T\vec{u}_m = \sigma_m^2\vec{u}_m = \vec{0}$$

$\vec{u}_{r+1}, \ldots, \vec{u}_m$ span $\text{Nul}(AA^T) = \text{Nul}A^T$

$\{\vec{u}_{r+1}, \ldots, \vec{u}_m\}$ orthonormal basis for NulA^T

Summary and Examples

$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ is an orthonormal basis for the **column space**
$\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m$ is an orthonormal basis for the **left nullspace** $\boldsymbol{N}(A^{\mathrm{T}})$
$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ is an orthonormal basis for the **row space**
$\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$ is an orthonormal basis for the **nullspace** $\boldsymbol{N}(A)$.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\frac{\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}{\sqrt{5}}$$