Lecture 5

Math Foundations Team

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

find e.val & e vec of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$A - \lambda I = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} \Rightarrow (1-\lambda)^2 = 0 \Rightarrow \lambda = 1, 1$

$\therefore$ Algebraic multiplicity of $\lambda = 1$ is 2.

for $\lambda = 1$:

e vec forms $N(A - \lambda I)$

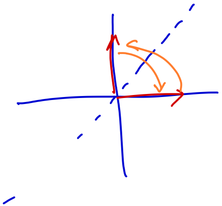
$$\left[\begin{array}{cc|c} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \end{array}\right] = \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \uparrow & \uparrow & \\ t & s & \end{array}\right]$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} s$$

$\Rightarrow$ any vector in x-y plane can be e.vec (use suitable value of t & s)

$\therefore$ $N(A - \lambda I)$ is entire x-y plane

$\therefore$ geometric multiplicity of $\lambda = 1$ is 2 find the transformation mat. for reflection about the line $x = y$



$$A = \begin{bmatrix} \text{Refl}\begin{bmatrix}1\\0\end{bmatrix} & \text{Refl}\begin{bmatrix}0\\1\end{bmatrix} \end{bmatrix}$$

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

Char eqn: $P(\lambda) = |A - I\lambda| = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 \Rightarrow \lambda = \pm 1$

AM of $+1 = 1$

AM of $-1 = 1$

for $\lambda_1 = 1$

$$N(A - \lambda I) = \left[\begin{array}{cc|c} 0-\lambda & 1 & 0 \\ 1 & 0-\lambda & 0 \end{array}\right] = \left[\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} \enclose{circle}{1} & -1 & 0 \\ 0 & \underset{t}{0} & 0 \end{array}\right]$$

$x - t = 0$

$y = t$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} t$$

GM of $\lambda = 1$ is 1

$\lambda_1 = -1$

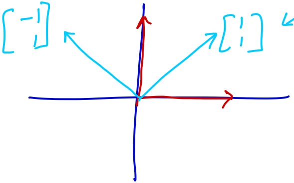
$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} \enclose{circle}{1} & 1 & 0 \\ 0 & \underset{t}{0} & 0 \end{array}\right]$$

$x + t = 0$

$y = t$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} t$$

GM of $\lambda = -1$ is 1



dir. along which Refl about $x=y$ results only in scaling

What is inverse of Refl about $x=y$

$$\underbrace{R_{x=y}^{-1}}_{\hookrightarrow \text{ undoes what } R_{x=y} \text{ does}} R_{x=y} = I$$

$$\Rightarrow R_{x=y}^{-1} = \left[R_{x=y}^{-1}\begin{bmatrix}1\\0\end{bmatrix} \quad R_{x=y}^{-1}\begin{bmatrix}0\\1\end{bmatrix} \right]$$

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

in $R_{x=y}$, x-axis is moved to y-axis. ∴ to undo y-axis is moved to x

$$\begin{bmatrix}a & b\\c & d\end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}$$

using formula:

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}^{-1} = \frac{1}{0-1}\begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = R_{x=y}^{-1}$$

Check

$$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = I$$

Introduction

- In the previous lecture, we discussed eigenvalues and eigenvectors of matrices
- In this lecture, we will look at two related methods for factorizing matrices into canonical forms.
- The first one is known as Eigenvalue decompostion. It uses the concepts of eigenvalues and eigenvectors to generate the decomposition
- The second method known as singular value decomposition or SVD is applicable to all matrices

Diagonal Matrices

- A diagonal matrix is a matrix that has value zero on all off diagonal elements.

$$\mathcal{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

- For a diagonal matrix $\mathcal{D}$, the determinant is the product of its diagonal entries.
- A matrix power $\mathcal{D}^k$ is given by each diagonal element raised to the power k.

$$D.D = \begin{bmatrix} d^2 & & \\ & d^2 & \\ & & \ddots \end{bmatrix}$$

- Inverse of a diagonal matrix is obtained by taking inverse of non-zero diagonal entry.

$$D^{-1}.D = I$$

Diagonalizable Matrices

- A matrix $A \in \mathbb{R}^{n \times n}$ is diagonalizable if there exists an invertible matrix $P \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\mathcal{D}$ such that $\mathcal{D} = \mathbf{P}^{-1}\mathbf{AP}$ or $A = P\Lambda P^{-1}$ $M = PNP^{-1}$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

- In the definition of diagonalization, it is required that $\mathbf{P}$ is an invertible matrix. Assume $p_1, p_2, ..., p_n$ are the n columns of $\mathbf{P}$
- Rewriting we get $\mathbf{AP} = \mathbf{PD}$. By observing that $\mathcal{D}$ is a diagonal matrix, we can simplify as

$$\mathcal{A}p_i = \lambda_i p_i$$

where λ_i is the i^{th} diagonal entry in $\mathcal{D}$.

Recall:

unique eval $\Rightarrow$ L.I. e.vec

$\Rightarrow$ if there are n. L.I. e.vec

$\rightarrow$ invertible P

Diagonalizable Matrix

- Consider a square matrix

$$\mathcal{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

- Consider the invertible matrix

$$\mathbf{P} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$$

- Now consider the product $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ as follows

$$\begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

$$\det(A - \lambda I) = 0 \Rightarrow \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)(3-\lambda) - 8 = 0$$

$$\Rightarrow \lambda^2 - 4\lambda - 5 = 0 \Rightarrow (\lambda+1)(\lambda-5) = 0$$

$$\lambda = -1, 5$$

$\lambda = -1$

$$N(A - \lambda I) \Rightarrow \left[\begin{array}{cc|c} 1-(-1) & 4 & 0 \\ 2 & 3-(-1) & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 2 & 4 & 0 \\ 2 & 4 & 0 \end{array}\right] \rightarrow \begin{array}{c} \left[\begin{array}{cc|c} \textcircled{1} & 2 & 0 \\ 0 & 0 & 0 \end{array}\right] \\ \begin{array}{ccc} x & t & \end{array} \end{array}$$

$$x + 2t = 0 \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \underline{\begin{bmatrix} -2 \\ 1 \end{bmatrix} t}$$

$\lambda = 5$

$$N(A - \lambda I) \Rightarrow \left[\begin{array}{cc|c} 1-5 & 4 & 0 \\ 2 & 3-5 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} -4 & 4 & 0 \\ 2 & -2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 1 & -1 & 0 \end{array}\right]$$

$$\rightarrow \begin{array}{c} \left[\begin{array}{cc|c} \textcircled{1} & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \\ \begin{array}{ccc} x & t & \end{array} \end{array} \qquad \begin{bmatrix} x \\ y \end{bmatrix} = \underline{\begin{bmatrix} 1 \\ 1 \end{bmatrix} t}$$

$$x - t = 0$$

Eigendecomposition of a matrix

- Recall the existence of eigenvalues and eigenvectors for square matrices
- Eigenvalues can be used to create a matrix decomposition known as Eigenvalue Decomposition
- A square matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ can be factored into

$$\mathbf{A} = \mathbf{PDP}^{-1}$$

- where $\mathcal{P}$ is an invertible matrix of eigenvectors of A assuming we can find n eigenvectors that form a basis of $\mathbb{R}^n$
- and $\mathcal{D}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathcal{A}$

Example of Eigendecomposition

Let us compute the eigendecomposition of the matrix A

$$\mathbf{A} = \begin{bmatrix} 2.5 & -1 \\ -1 & 2.5 \end{bmatrix}$$

- Step 1: Find the eigenvalues and eigenvectors

$$\mathbf{A} - \lambda \mathbf{I} = \begin{bmatrix} 2.5 - \lambda & -1 \\ -1 & 2.5 - \lambda \end{bmatrix}$$

- The characteristic equation is given by $\det(\mathbf{A} - \lambda \mathbf{I}) = \mathbf{0}$
- This leads to the equation $\lambda^2 - 5\lambda + \frac{21}{4} = 0$
- Solving the quadratic equation gives us $\lambda_1 = 3.5$ and $\lambda_2 = 1.5$

Example of Eigendecomposition

- The eigenvector corresponing to $\lambda_1 = 3.5$ is derived as

HW

$$p_1 = \begin{bmatrix} -\frac{1}{\sqrt{2})} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

- The eigenvector corresponing to $\lambda_1 = 1.5$ is derived as

$$p_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

- Step 2 : Construct the matrix **P** to diagonalize **A**

$$\mathbf{P} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example of Eigendecomposition

- The inverse of matrix P is given by

$$\mathbf{P}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

- The eigendecompostion of the matrix A is given by

✓

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3.5 & 0 \\ 0 & 1.5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

- In summary we have obtained the required matrix factorization using eigenvalues and eigenvectors.

$$Ax = \underbrace{P \underbrace{D \underbrace{P^{-1} x}}}$$

Symmetric Matrices and Diagonalizability

- Recall that a matrix A is called symmetric matrix if $\mathbf{A} = \mathbf{A^T}$

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

- A Symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ can always be diagonalized.
- This follows directly from the spectral theorem discussed in previous lecture
- Moreover the spectral theorem states that we can find an orthogonal matrix $\mathbf{P}$ of eigenvectors of $\mathcal{A}$.

if A is diagonizable : $A = P \Lambda P^{-1}$

if A is symm: $A = Q \Lambda Q^T$

Why Diagonalize?

$$A = Q \Lambda Q^T$$

$$A^N = (Q \Lambda Q^T)^N$$

$$= (Q \Lambda \underbrace{Q^T)(Q}_{I} \Lambda \underbrace{Q^T)(Q}_{I} \Lambda Q^T) \ldots$$

$$A^N = Q \underbrace{\Lambda \Lambda \ldots}_{\Lambda^N} Q^T = Q \Lambda^N Q^T$$

<u>HW</u>

- N^{th} term in fibonacci series

- # of ways to get from vertex i to vertex j in a graph in N steps

  = $(i,j)^{th}$ element of A^N

$$A = \begin{array}{c|cccc} & v_1 & v_2 & \cdots & \\ \hline v_1 & 1 & 1 & & \\ v_2 & & & & 1 \\ \vdots & & & & \\ \vdots & & & & \end{array}$$

Motivation for Singular Value Decomposition

- The singular value decomposition or (SVD) of a matrix is a central matrix decomposition method in linear algebra.
- The eigenvalue decomposition is applicable to square matrices only.
- The singular value decomposition exists for all rectangular matrices
- SVD involves writing a matrix as a product of three matrices $\mathbf{U}, \mathbf{\Sigma}$ and $\mathbf{V^T}$.
- The three component matrices are derived by applying eigenvalue decomposition discussed previously

Eigen value is defined only for $A \in R^{n\times n}$

What about $A \in R^{N\times D}$ $N \neq D$?

$\det(A - \lambda I)$

SVD Decomposes A as:

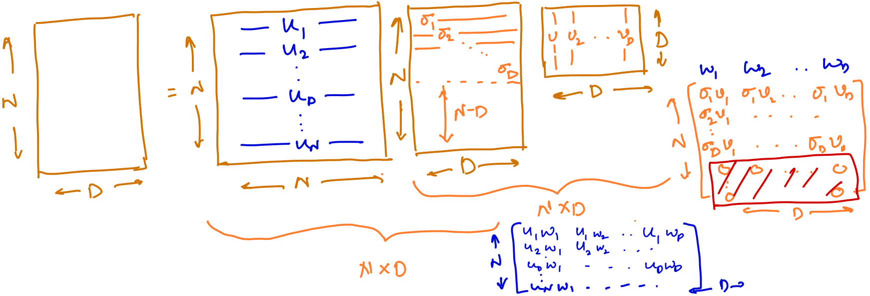
$$\underset{N\times D}{A} = \underset{N\times N}{U}\;\underset{N\times D}{\Sigma}\;\underset{D\times D}{V^T}$$

orthogonal

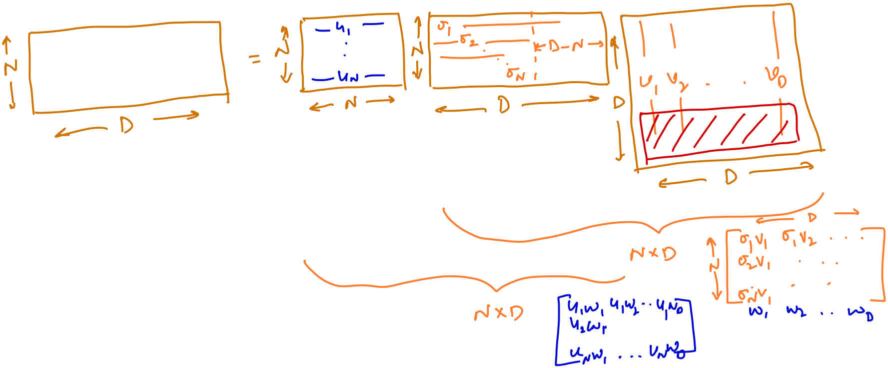
Diagonal

orthogonal

Consider $N > D$



Consider $N < D$


N
D
=
N
u₁
uₙ
N
N
σ₁
σ₂
σₙ
D−N
D
D
v₁ v₂ . . . v_D
D
N×D
N×D
D
N
σ₁v₁ σ₁v₂ . . .
σ₂v₁
σₙv₁
w₁ w₂ . . w_D
u₁w₁ u₁w₂ . . u₁w_D
u₂w₁
uₙw₁ . . . uₙw_D

Singular Value Decomposition Theorem

- Let $\mathbf{A} \in \mathbb{R}^{\mathbf{m \times n}}$ be a rectangular matrix. Assume that $\mathbf{A}$ has rank r.
- The Singular value decomposition of $\mathcal{A}$ is defined as

$$\mathbf{A} = \mathbf{U \Sigma V^T}$$

- $\mathbf{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix with column vectors u_i where $i = 1, ..., m$
- $\mathbf{V} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix with column vectors v_j where $j = 1, ..., n$
- $\mathbf{\Sigma}$ is an m × n matrix with $\mathbf{\Sigma_{ii}} = \sigma_\mathbf{i} > \mathbf{0}$
- The diagonal entries $\sigma_i, i = 1, ..., r$ of $\mathbf{\Sigma}$ are called the singular values.
- By convention, the singular values are ordered i.e $\mathbf{\Sigma}_{ii} > \mathbf{\Sigma}_{jj}$ where $i < j$.

Finding V

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

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = \underset{(V^T)^T = V}{V} \underset{\Sigma^T = \Sigma}{\Sigma} \underbrace{U^T U}_{I} \Sigma V^T = V \Sigma^2 V^T$$

$$\underbrace{A^T A}_{\text{symm matrix}} = \underbrace{V}_{\text{orthogonal, e.vec. of } A^TA} \underbrace{\Sigma^2}_{\text{diag eval of } A^TA} V^T$$

eigen decomposition of symm. mat:

$$A = \underset{\text{cols are ortho. e.vec}}{P} \underset{\text{diag. e.val}}{\Lambda} P^T$$

$$\boxed{\begin{array}{l} \therefore \text{ cols of } V \iff \text{e.vec of } A^TA \\ \text{entries of } \Sigma^2 \iff \text{eval of } A^TA \\ \\ \underline{\text{Similarly}} \\ \quad \text{cols of } U \iff \text{evec of } AA^T \\ \text{entries of } \Sigma^2 \iff \text{eval of } AA^T \end{array}}$$

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$
$$= U \Sigma V^T V \Sigma U^T$$
$$= U \Sigma^2 U^T$$

$$\underbrace{A^T A} x = \lambda x \Rightarrow A A^T A x = A \lambda x$$
$$\Rightarrow (AA^T)(\underset{v}{Ax}) = \lambda(\underset{v}{Ax}) \Rightarrow$$

$(\lambda \neq 0)$
if λ is e.val of A^TA, it is eval of AA^T

Properties of Σ

- The singular value matrix $\boldsymbol{\Sigma}$ is unique.
- Observe that the $\boldsymbol{\Sigma} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$matrix is rectangular. In particular, $\boldsymbol{\Sigma}$ is of the same size as $\mathcal{A}$.
- This means that $\boldsymbol{\Sigma}$ has a diagonal submatrix that contains the singular values and needs additional zero padding.
- ✓ Specifically, if $m > n$, then the matrix $\boldsymbol{\Sigma}$ has diagonal structure up to row n and then consists of zero rows.
- ✓ If $m < n$, the matrix Σ has a diagonal structure up to column m and columns that consist of 0 from $m + 1$ to n.

Construction of V

- It can be observed that

$$\mathbf{A}^{\mathbf{T}}\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{V}^{\mathbf{T}}$$

- Since A^TA has the following eigendecomposition

$$\mathbf{A}^{\mathbf{T}}\mathbf{A} = \mathbf{P}\mathcal{D}\mathbf{P}^{\mathbf{T}}$$

- ✓ Therefore, the eigenvectors of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ that compose $\mathbf{P}$ are the right-singular vectors $\mathbf{V}$ of $\mathcal{A}$.
- The eigenvalues of $\mathcal{A}^{\mathbf{T}}\mathbf{A}$ are the squared singular values of $\boldsymbol{\Sigma}$

Construction of U

- It can be observed that

$$\mathbf{AA^T} = \mathbf{U\Sigma V^T V \Sigma^T U^T}$$

- Since $\mathbf{AA^T}$ has the following eigendecomposition

$$\mathbf{AA^T} = \mathbf{SDS^T}$$

- Therefore, the eigenvectors of $\mathbf{AA^T}$ that compose $\mathbf{S}$ are the left-singular vectors $\mathbf{U}$ of $\mathcal{A}$

Construction of U continued

- $\mathcal{A} = \mathbf{U\Sigma V^T}$ can be rearranged to obtain a simple formulation for u_i
- By postmultiplying by $\mathbf{V}$ we get $\mathbf{AV} = \mathbf{U\Sigma V^T V}$
- By observing that $\mathbf{V}$ is orthogonal we obtain a simple form

$$\mathbf{AV} = \mathbf{U\Sigma}$$

$$\overset{n\times D}{A} = \overset{n\times n}{U}\;\overset{n\times D}{\Sigma}\;\overset{D\times D}{V^T}$$

$$\underset{n\times D}{A}\;\underset{D\times D}{V} = \underset{n\times n}{U}\;\underset{n\times D}{\Sigma}$$

- This is equivalent to the following

$$u_i = \frac{1}{\sigma_i}\mathcal{A}v_i \quad \forall i = 1, 2, ..., r$$

$$\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n\end{bmatrix}\underbrace{\begin{bmatrix}| & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & |\end{bmatrix}}_{U} = \underbrace{\begin{bmatrix}| & | & & \\ a_1 & a_2 & \cdots & \cdot \\ | & | & & \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} v_{11} & | & | \\ v_{12} & v_2 & v_D \\ \vdots & | & | \end{bmatrix}}_{V}$$

Computing Singular Value Decomposition 1

- We want to find SVD of the following rectangular matrix $\mathcal{A}$

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

- Let us consider the matrix $\mathcal{A}^T\mathcal{A}$ derived from $\mathcal{A}$ given by

✓

$$\mathcal{A}^T\mathcal{A} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- It is a symmetric matrix

Computing Singular Value Decomposition 2

- Derive the eigendecomposition of $\mathcal{A}^T\mathcal{A}$ in the form PDP^T

HW

- The matrix **P** is given by

$$\begin{array}{c} \mathbf{P} = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix} \\ \qquad\quad v_1 \qquad v_2 \qquad v_3 \end{array}$$

- The matrix $\mathcal{D}$ is given by

$$\mathcal{D} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Computing Singular Value Decomposition 3

Now we construct the singular value matrix $\boldsymbol{\Sigma}$

- The matrix $\boldsymbol{\Sigma}$ has the dimension same as $\mathcal{A}$. In this case $\boldsymbol{\Sigma}$ is hence a 2×3 matrix.
- The diagonal entries of submatrix is obtained by taking square root of 6 and 1 respectively
- Singular-value matrix $\boldsymbol{\Sigma}$ is given by:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

e.val are Squares of Singular value Σ

- The last column is a column of zeros only

$$\underset{2\times 3}{A} = \underset{2\times 2}{U}\ \underset{2\times 3}{\Sigma}\ \underset{3\times 3}{V^T}$$

↳ $\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$

Computing Singular Value Decomposition 4

Left singular vectors as the normalized image of the right singular vectors. Recall that $u_i = \frac{1}{\sigma_i}\mathbf{A}v_i$

$$\begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} \frac{1}{\sqrt{30}}$$

- The first vector

$$u_1 = \frac{1}{\sigma_1}\mathcal{A}v_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix}$$

- The second vector

$$u_2 = \frac{1}{\sigma_2}\mathcal{A}v_2 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

Final Step : Combining U, Σ and V

We compile all the three matrices together to generate the SVD

- $$\mathcal{A} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}^T$$
- The matrix **U** is an 2×2 matrix satifying orthogonality property.
- The matrix **V** is an 3×3 matrix satifying orthogonality property.

Comparing SVD and EVD

- The left-singular vectors of $\mathcal{A}$ are eigenvectors of $\mathbf{AA^T}$
- The right-singular vectors of $\mathcal{A}$ are eigenvectors of $\mathcal{A}^T\mathcal{A}$
- The non-zero singular values of $\mathcal{A}$ are the square roots of the nonzero eigenvalues of $\mathcal{A}^T\mathcal{A}$.
- The SVD always exists for any matrix in $\mathbb{R}^{m\times n}$
- The eigendecomposition is only defined for square matrices in $\mathbb{R}^{n\times n}$ and only exists if we can find a basis of eigenvectors of $\mathbb{R}^n$

Comparing SVD and EVD

- The vectors in the eigendecomposition matrix **P** are not necessarily orthogonal.
- On the other hand, the vectors in the matrices **U** and **V** in the SVD are orthonormal. $A^T A$ is symm
- Both the eigendecomposition and the SVD are compositions of three linear mappings:
- A key difference between the eigendecomposition and the SVD is that in the SVD, domain and codomain can be of different dimensions
- In the SVD, the left and right singular vector matrices **P** and **P** are generally not inverse of each other.

Comparing SVD and EVD 3

- In the eigendecomposition, the matrices in decomposition are inverse of each other
- In the SVD, the entries in the diagonal matrix $\mathbf{\Sigma}$ are all real and nonnegative,

  ↳ symm

  → pos. def
- In eigendecomposition diagonal matrix entries need not be real always.
- The leftsingular vectors of $\mathcal{A}$ are eigenvectors of $\mathcal{A}\mathcal{A}^T$
- The rightsingular vectors of $\mathcal{A}$ are eigenvectors of $\mathcal{A}^T\mathcal{A}$.

→ Consider $x^TA^TAx = (Ax)^TAx \geqslant 0$

Matrix Approximation

- We considered the SVD as a way to factorize $\mathcal{A} = \mathbf{U\Sigma V^T}$ into the product of three matrices, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal and $\mathbf{\Sigma}$ contains the singular values on its main diagonal.
- Instead of doing the full SVD factorization, we will now investigate how the SVD allows us to represent a matrix $\mathcal{A}$ as a sum of simpler matrices $\mathcal{A}_i$
- This representation which lends itself to a matrix approximation scheme that is cheaper to compute than the full SVD.

$$\mathcal{A} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}^T$$

$$= \sqrt{6} \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{30}} \end{bmatrix} + 1 \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$= \sqrt{6} \begin{bmatrix} \frac{5}{\sqrt{5}\sqrt{30}} & -\frac{2}{\sqrt{5}\sqrt{30}} & \frac{1}{\sqrt{5}\sqrt{30}} \\ \frac{-10}{\sqrt{5}\sqrt{30}} & \frac{+4}{\sqrt{5}\sqrt{30}} & -\frac{2}{\sqrt{5}\sqrt{30}} \end{bmatrix} + \begin{bmatrix} 0 & \frac{2}{5} & \frac{4}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \end{bmatrix}$$

$$= \frac{\sqrt{6}}{\sqrt{5}\sqrt{6}\sqrt{5}} \begin{bmatrix} 5 & -2 & 1 \\ -10 & 4 & -2 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 0 & 2 & 4 \\ 0 & 1 & 2 \end{bmatrix}$$

$$= \underbrace{\underbrace{\frac{1}{5} \begin{bmatrix} 5 & -2 & 1 \\ -10 & 4 & -2 \end{bmatrix}}_{\text{approx 1}} + \frac{1}{5} \begin{bmatrix} 0 & 2 & 4 \\ 0 & 1 & 2 \end{bmatrix}}_{\text{approx 2}}$$

$$= \frac{1}{5} \begin{bmatrix} 5 & 0 & 5 \\ -10 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = A$$

Matrix Approximation continued

- A matrix $\mathcal{A} \in \mathbb{R}^{m \times n}$ of rank r can be written as a sum of rank-1 matrices so that $\mathcal{A} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$
- ✓ The diagonal structure of the singular value matrix $\mathbf{\Sigma}$ multiplies only matching left and right singular vectors $u_i v_i^T$ and scales them by the corresponding singular value σ_i.
- All terms $\sigma_i u_i v_i^T$ vanish for $i \neq j$ because $\mathbf{\Sigma}$ is a diagonal matrix.
- Any term for $i > r$ would vanish because the corresponding singular value is 0.

Rank k Approximation

- We summed up the r individual rank-1 matrices to obtain a rank r matrix $\mathcal{A}$.
- If the sum does not run over all matrices $A_i \;\; i = 1, ..., r$ but only up to an intermediate value k we obtain a rank-k approximation
- The approximation represented by $\hat{\mathcal{A}}(k)$ is defined as follows

$$\hat{\mathcal{A}}(k) = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

- To measure the difference between $\mathcal{A}$ and its rank-k approximation we need the notion of a norm which is introduced next

Spectral Norm of a matrix

- We introduce the notation of a subscript in the matrix norm
- Spectral Norm of a Matrix. For $x \in \mathbb{R}^n$, $x \neq \mathbf{0}$, the spectral norm norm of a matrix $\mathcal{A} \in \mathbb{R}^{m \times n}$ is defined as

  $$\|\mathcal{A}\|_2 = \max_x \frac{\|\mathcal{A}x\|_2}{\|x\|_2}$$

  where $\|y\|_2$ is the euclidean norm of y
- Theorem : The spectral norm of a matrix $\mathcal{A}$ is its largest singular value

Example : Spectral Norm of a matrix

- Example : Consider the following matrix A

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

- Singular value decomposition of this matrix will provide the matrix Σ as follows

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5.465 & 0 \\ 0 & 0.366 \end{bmatrix}$$

- The 2 singular values are 5.4650 and 0.366.
- By definition the spectral norm is the largest singular value.
- Hence, the spectral norm is 5.4650

SVD on Image

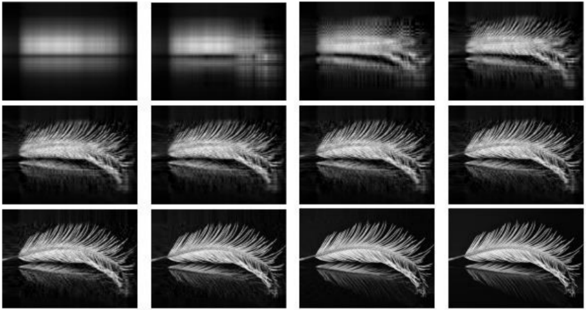


Figure 2: Number of Singular Values: $\{1, 2, 5, 10\}\{15, 18, 24, 30\}\{35, 60, 120, 680\}$

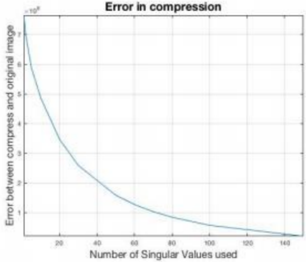
SVD → Eigenvalus

↓

vec. spaces ← Null space

(orthogonal & complimentary)

↳ inner prod



error → MSE → distance metric

↓

inner prod space ← norm