Matrix Decomposition

Introduction

1145×1718 pixels image

➔ Require 1,967,110 numbers to store

An Approximation of the image needs only 50*(1718 +1145+1)

= 143,200 numbers

Matrix Approximation – Ex

image $\mathbf{A} \in \mathbb{R}^{1145\times 1718}$

$\hat{A}(5)$, error = 0.647

$\hat{A}(10)$, error = 0.592

$\hat{A}(50)$, error = 0.393

$\hat{A}(100)$, error = 0.280

$\hat{A}(200)$, error = 0.164

Contents

- Matrix Decomposition
  - Eigenvalue Decomposition **$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$**
    - Eigenvalues and Eigenvectors
  - Singular Value Decomposition (SVD) **$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$**
    - Left-singular vectors (columns of **U**)
    - Singular values (diagonal entries of singular matrix Σ)
    - Right-singular vectors (columns of **V**)
- Matrix Approximation
  - Rank-1 matrix $\mathbf{A}_i = \mathbf{u}_i\mathbf{v}_i^T$
  - $\mathbf{A} = \Sigma_{i=1..r}\, \sigma_i \mathbf{A}_i$
  - $\mathbf{A} \approx \widehat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{A}_i$

Computer project 1

- Chia nhóm: 6 nhóm x 5 bạn = 30 SV (Group = No. mod 6 + 1)
- Project 1
  - Sử dụng SVD xấp xỉ 1 ảnh (kích thước tuỳ chọn) với sai số < 20%
  - Yêu cầu: không sử dụng hàm SVD được dựng sẵn trong Python (nếu có)
- Tiêu chí đánh giá
  - Hoàn thành đúng yêu cầu: 5 đ
  - Giải thích được: 3đ
  - Nhóm khác cho điểm: 1, 2, 3, 4, 5 ➔ cao nhất & nhì: +2, ba & tư: +1, thứ năm, sáu: +0

Trace of a square matrix

Definition. The *trace* of a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ is defined as

$$tr\ A := \sum_{i=1}^{n} a_{ii}$$

Properties

- $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$, for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$
- $\mathrm{tr}(\alpha\mathbf{A}) = \alpha\,\mathrm{tr}(\mathbf{A})$, $\alpha\in\mathbb{R}$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$
- $\mathrm{tr}(\mathbf{I}_n) = n$
- $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB})$

for $\mathbf{A} \in \mathbb{R}^{n\times k}$, $\mathbf{B} \in \mathbb{R}^{k\times m}$, $\mathbf{C} \in \mathbb{R}^{m\times n}$

  - In particular, $\mathrm{tr}(\mathbf{x}\mathbf{y}^T) = \mathrm{tr}(\mathbf{x}^T\mathbf{y}) = \mathbf{x}^T\mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in\mathbb{R}^n$

$$\mathbf{A} = \begin{pmatrix} 1 & 5 & 2 \\ 2 & 8 & 4 \\ -3 & -12 & 7 \end{pmatrix}$$

$\Rightarrow$ $\mathrm{tr}(A) = 1 + 8 + 7 = 16$

Characteristic Polynomial

Definition. For $\lambda \in \mathbb{R}$ and a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$

$$p_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda\mathbf{I}) = c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n,$$

where $c_0, \ldots, c_{n-1} \in \mathbb{R}$,
is the *characteristic polynomial* of $\mathbf{A}$.

Note. $c_0 = \det(\mathbf{A})$, $c_{n-1} = (-1)^{n-1}\mathrm{trace}(\mathbf{A})$

Characteristic Polynomial – Ex

Find the *characteristic polynomial* of the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 0 \\ 1 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

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

$= (1 - \lambda)[(2- \lambda)(4 - \lambda) - 3] = 5 - 11\lambda + 7\lambda^2 - \lambda^3$

Note. $c_0 = \det(\mathbf{A}) = 5$, $c_3 = (-1)^3 = -1$, $c_2 = (-1)^{3-1}\mathrm{trace}(\mathbf{A}) = 7$

Eigenvalues and Eigenvectors

Definition. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then $\lambda \in \mathbb{R}$ is an *eigenvalue* of $\mathbf{A}$ and $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ is the corresponding eigenvector of $\mathbf{A}$ if *eigenvalue eigenvector*
$\mathbf{Ax} = \lambda\mathbf{x}$.
We call $\mathbf{Ax} = \lambda\mathbf{x}$ the *eigenvalue equation.*

Ex. If $\mathbf{A} = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ then $\mathbf{Ax} = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \end{pmatrix} = -3\begin{pmatrix} 2 \\ 0 \end{pmatrix} = -3\mathbf{x}$

➔ **x** is an *eigenvector* of A corresponding to *eigenvalue -3*

Theorem. $\lambda \in \mathbb{R}$ is an *eigenvalue* of $\mathbf{A} \in \mathbb{R}^{n\times n}$ if and only if λ is a *root* of the *characteristic polynomial* $p_{\mathbf{A}}(\lambda)$ of $\mathbf{A}$

Notes

The following statements are equivalent:

- λ is an *eigenvalue* of $\mathbf{A} \in \mathbb{R}^{n \times n}$
- There exists an $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ with $\mathbf{Ax} = \lambda\mathbf{x}$,
- $(A - \lambda\mathbf{I}_n)x = \mathbf{0}$ can be solved *non-trivially*, i.e., $\mathbf{x} \neq \mathbf{0}$
- $\mathrm{rk}(\mathbf{A} - \lambda\mathbf{I}_n) < n$
- $\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$

Eigenvalues and Eigenvectors – Ex

- Find *eigenvalues* and *eigenvectors*, eigenspaces of the matrix

$$\mathbf{A} = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix}.$$

- The *characteristic polynomial* $p_{\mathbf{A}}(\lambda)$ of **A**

$$p_{\mathbf{A}}(\lambda) = \lambda^2 + 2\lambda - 3 = 0$$

- $p_{\mathbf{A}}(\lambda) = 0 \Leftrightarrow \lambda_1 = 1, \lambda_2 = -3$ // *eigenvalues*

Eigenvalues and Eigenvectors – Ex

- $\lambda_1 = 1$: solve the system $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \mathbf{0}$

$\left(\begin{array}{cc|c} -4 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow \boldsymbol{x} = t\begin{pmatrix} 1 \\ 2 \end{pmatrix}, t \in \mathbb{R}$ and all $\boldsymbol{x} \neq \mathbf{0}$ are *1-eigenvectors*

- $\lambda_2 = -3$: solve the system $(\mathbf{A} - \lambda_2\mathbf{I})\mathbf{x} = \mathbf{0}$

$\left(\begin{array}{cc|c} 0 & 2 & 0 \\ 0 & -4 & 0 \end{array}\right) \Rightarrow \boldsymbol{x} = t\begin{pmatrix} 1 \\ 0 \end{pmatrix}, t \in \mathbb{R}$ and all $\boldsymbol{x} \neq \mathbf{0}$ are *-3-eigenvectors*

Theorems

Theorem. The *determinant* of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the *product* of its eigenvalues, i.e.,

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \cdots \lambda_n$$

where $\lambda_i \in \mathbb{C}$ are (possibly repeated) eigenvalues of $\mathbf{A}$

Theorem. The *trace* of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the *sum* of its eigenvalues, i.e.,

$$\operatorname{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Theorem

Theorem. The eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are *linearly independent*

Ex. If $\mathbf{A} = \begin{pmatrix} -3 & 2 \\ 0 & 1 \end{pmatrix}$, then $\mathbf{x}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ are two eigenvectors with *distinct* eigenvalues $\lambda_1 = -3$, $\lambda_2 = 1$, respectively

➔ x_1 and x_2 are *linearly independent*

Useful properties regarding eigenvalues and eigenvectors

For $\mathbf{A} \in \mathbb{R}^{n \times n}$, we define

- λ-eigenspace: $\mathbf{E}_\lambda = \{\mathbf{x} \in \mathbb{R}^n \mid (\mathbf{A} - \lambda \mathrm{I})\mathbf{x} = \mathbf{0}\}$
- *spectrum* of **A** is the set of *all eigenvalues* of **A**
- **A** and its transpose $\mathbf{A}^T$ possess the same eigenvalues
- $\mathbf{E}_\lambda$ is the *null space* of $\mathbf{A} - \lambda\mathbf{I}$ (the *kernel* of $\mathbf{A} - \lambda\mathbf{I}$)
- *Symmetric, positive definite* matrices always have *positive, real eigenvalues*

Algebraic multiplicity vs Geometric multiplicity

Let λ be an eigenvalue of a square matrix **A**

- *Algebraic multiplicity* of λ is the number of times λ appears
- *Geometric multiplicity* of λ is the number of linearly independent vectors associated with λ = dim($\mathbf{E}_\lambda$)
- Algebraic multiplicity of λ $\geqslant$ Geometric multiplicity of λ
- **Ex**. If A = $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, A has only one eigenvalue λ = 3 ($\lambda_1 = \lambda_2 = 3$)
  - *Algebraic multiplicity* of λ = 2,
  - *Algebraic multiplicity* of λ = dim($\mathbf{E}_2$) = 1

Spectral Theorem

Theorem (Spectral Theorem)**.** If $\mathbf{S} \in \mathbb{R}^{n \times n}$ is *symmetric*, there exists an *orthonormal basis* of the corresponding vector space **V** consisting of eigenvectors of S, and each eigenvalue is real.

Note.

If S is a symmetric matrix and λ_1, λ_2 are distinct eigenvalues of S, then eigenvectors v_1, v_2 corresponding to λ_1, λ_2 are orthogonal.

Spectral Theorem – Ex

Find an *orthonormal basis* of the corresponding vector space **V** consisting of eigenvectors of **A**, where

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$$

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{pmatrix} 3-\lambda & 2 & 2 \\ 2 & 3-\lambda & 2 \\ 2 & 2 & 3-\lambda \end{pmatrix} = -\lambda^3 + 9\lambda^2 - 15\lambda + 7 =$$

$$-(\lambda - 1)^2(\lambda - 7)$$

$$p_{\mathbf{A}}(\lambda) = 0 \Leftrightarrow \lambda_1 = 1, \lambda_2 = 7$$

Spectral Theorem – Ex

- For $\lambda_1 = 1$, solve the system $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = 0$, we have solutions $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -m-n \\ n \\ m \end{pmatrix} = m\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + n\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, m, n $\in \mathbb{R}$

So, *eigenspace* ***E₁*** = span$\{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\}$

Similarly, *eigenspace for* $\lambda_2 = 7$ is $\mathbf{E}_7$ = span$\{x_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\}$

Spectral Theorem – Ex

- Next, construct an *orthogonal basis* from

$$x_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\ x_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},\ \text{and } x_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

- Note that x_3 is *orthogonal* to both x_1, x_2, but x_1 and x_2 are not orthogonal to each other.
- Use *Gram-Schmidt* process:

$$x'_2 = x_2 - \mathrm{proj}_{x1}(x_2) = x_2 - \frac{x_1^T x_2}{\|x_1\|^2} x_1 = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \end{pmatrix}$$

Now, $\{x_1, x'_2, x_3\}$ is an *orthogonal basis* of $\mathbb{R}^3$

Definitions

- **Diagonal matrix.** A *diagonal matrix* is a matrix that has value zero on all off-diagonal elements
- **Similar matrices.** Two matrices **A**, **D** are *similar* if there exists an invertible matrix P, such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$
- **Diagonalizable.** A matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is *diagonalizable* if it is *similar* to a *diagonal matrix*, i.e., if there exists an invertible matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ such that $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

Example

- Two matrices **A** = $\begin{pmatrix} 2 & -3 \\ 0 & 1 \end{pmatrix}$ and D = $\begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix}$ are similar as

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

where **P** = $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{P}^{-1}$ = $\begin{pmatrix} 1 & -1/2 \\ 0 & 1/2 \end{pmatrix}$

- **A** is called *diagonalizable* as **D** is a diagonal matrix

How to find P

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

$\rightarrow$ **AP = PD**

We have,

$$\mathbf{AP} = \mathbf{A}[\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n] = [\mathbf{A}\mathbf{p}_1, \mathbf{A}\mathbf{p}_2, \ldots, \mathbf{A}\mathbf{p}_n]$$

$$\mathbf{PD} = [\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n]\begin{bmatrix} \lambda_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_n \end{bmatrix} = [\lambda_1\mathbf{p}_1, \ldots, \lambda_n\mathbf{p}_n]$$

$\rightarrow$ $\mathbf{A}\mathbf{p}_i = \lambda_i\mathbf{p}_i$, for $i = 1, \ldots, n$

Eigendecomposition

Theorem. A square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ can be factored into

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

where $\mathbf{P} \in \mathbb{R}^{n\times n}$ and $\mathbf{D}$ is a *diagonal matrix* whose diagonal entries are the *eigenvalues* of $\mathbf{A}$, if and only if the *eigenvectors* of $\mathbf{A}$ form a *basis* of $\mathbb{R}^n$.

Example. $\begin{pmatrix} 2 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix}$

Eigendecomposition – Ex

- Find the eigenvalue decomposition of $\mathbf{A} = \begin{pmatrix} 2 & -3 \\ 0 & 1 \end{pmatrix}$.
- $p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \det\begin{pmatrix} 2-\lambda & -3 \\ 0 & 1-\lambda \end{pmatrix} = (2-\lambda)(1-\lambda) = \lambda^2 - 3\lambda + 2$
- $p_{\mathbf{A}}(\lambda) = 0 \Leftrightarrow \lambda_1 = 1, \lambda_2 = 2$
- $\lambda_1 = 1$: solve the system $(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \mathbf{0}$

$\begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix}\mathbf{X} = \mathbf{0} \rightarrow \mathbf{X} = t\begin{pmatrix} 3 \\ 1 \end{pmatrix}$

- $\lambda_2 = 2$: solve the system $(\mathbf{A} - \lambda_2\mathbf{I})\mathbf{x} = \mathbf{0}$

$\begin{pmatrix} 0 & -3 \\ 0 & -1 \end{pmatrix}\mathbf{X} = \mathbf{0} \rightarrow \mathbf{X} = t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$A = PDP^{-1}$

- P has columns as eigenvectors of A
- $\mathbf{P} = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \mathbf{P}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix}$
- Then, $\mathbf{A} = \mathbf{PDP}^{-1}$

where $\mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ // 1, 2 are eigenvalues of A

Eigendecomposition of **A**

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

Theorems

Theorem. A *symmetric* matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$ can always be *diagonalized*.

Theorem. Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, we can always obtain a *symmetric, positive semidefinite* matrix $S \in \mathbb{R}^{n \times n}$ by defining

$$\mathbf{S} := \mathbf{A}^T\mathbf{A}$$

Morerover, if $rk(A) = n$, then A^TA is positive definite.

Singular Value Decomposition (SVD) - Introduction

- The *singular value decomposition* (SVD) of a matrix is a *central matrix decomposition* method in linear algebra.

- It has been referred to as the *"fundamental theorem of linear algebra"* because it can be applied to all matrices, not only to square matrices, and it always exists.

SVD - Theorem

Let $A^{m\times n}$ be a rectangular matrix of *rank* r. The **SVD** of **A** is a decomposition of the form

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

Moreover,

- **U**, **V**: *orthogonal matrices*, i.e., $\mathbf{U}^T\mathbf{U} = \mathbf{I}_m$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}_n$
- Σ: a *diagonal matrix* with $\Sigma_{kk} = \sigma_k \geqslant 0$ and $\Sigma_{ik} = 0$, $i \neq k$

σ_k are called the *singular values,* $\sigma_1 \geqslant \sigma_2 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r \geqslant 0$.

SVD - Ex

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

$$\mathbf{A} = \begin{bmatrix} 1 & -0.8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -0.79 & 0 & -0.62 \\ 0.38 & -0.78 & -0.49 \\ -0.48 & -0.62 & 0.62 \end{bmatrix} \begin{bmatrix} 1.62 & 0 \\ 0 & 1.0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0.78 & 0.62 \\ -0.62 & -0.78 \end{bmatrix}$$

How to find U, V^T and Σ

- First, find $\mathbf{P}$ such that $\mathbf{A}^T\mathbf{A} = \mathbf{PDP}^T$ (spectral theorem)

➔ $\mathbf{A}^T\mathbf{A} = \mathbf{P} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{bmatrix} \mathbf{P}^T, \lambda_i \geqslant 0$

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

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

Compare $\mathbf{A}^T\mathbf{A} = \mathbf{PDP}^T$

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

➔ Choose $\mathbf{V} = \mathbf{P}$, and $\mathbf{\Sigma}$ such that $\mathbf{\Sigma}^T\mathbf{\Sigma} = \mathbf{D}$, or $\sigma_k^2 = \lambda_k$

How to choose U

- If we know $\mathbf{A} = \mathbf{U\Sigma V}^T$

$$\Rightarrow \mathbf{AA^T} = (\mathbf{U\Sigma V}^T)(\mathbf{U\Sigma V}^T)^\mathbf{T} = \mathbf{U}\Sigma\Sigma^T\mathbf{U}^T = \mathbf{U}\begin{bmatrix} \sigma_1^2 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \sigma_m^2 \end{bmatrix}\mathbf{U}^T$$

$\Rightarrow$ The orthonormal eigenvectors of $\mathbf{AA}^T$ are the *left-singular vectors*

U can be chosen from *singular value equations* $\mathbf{Av}_k = \sigma_k\mathbf{u}_k$

Note. $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ is an *orthonormal* set

$$\|\mathbf{Av}_k\|^2 = (\mathbf{Av}_k)^T(\mathbf{Av}_k) = \lambda_k\mathbf{v}_k^T\mathbf{v}_k = \lambda_k$$

$$\Rightarrow \|\mathbf{u}_k\| = 1,\ k = 1, \dots, r$$

SVD – Ex2

Find the *singular value decomposition* of A = $\begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$

We follow three steps (suppose m < n)

- Step 1: *Right-singular vectors* as the eigenvectors of A^TA.
- Step 2: *Singular-value matrix* Σ with $\Sigma_{kk} = \sigma_k$ such that $\sigma_k^2 = \lambda_k$, eigenvalues of A^TA.
- Step 3: *Left-singular vectors* as the normalized image of the right-singular vectors, $\mathbf{u}_k = \frac{\mathbf{A}\mathbf{v}_k}{\sigma_k}$

SVD – Ex2

Step 1: *Right-singular vectors* as the eigenvectors of $\mathbf{A}^\mathsf{T}\mathbf{A}$.

- $\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$
- Apply *eigenvalue decomposition* to $\mathbf{A}^\mathsf{T}\mathbf{A}$, which is given as

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{5}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{30}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} = \mathbf{P}\mathbf{D}\mathbf{P}^\mathsf{T}$$

SVD – Ex2

- We obtain the *right-singular vectors* as the columns of **P**

$$\mathbf{V} = \mathbf{P} = \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

SVD – Ex2

Step 2: *Singular-value matrix* Σ with $\Sigma_{kk} = \sigma_k$, where $\sigma_k^2 = \lambda_k$

From the previous step, $\mathbf{D} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, with $\lambda_1 = 6, \lambda_2 = 1$

- Since $\text{rk}(\mathbf{A}) = 2$, there are only two *nonzero singular values*:

$$\sigma_1 = \sqrt{6} \text{ and } \sigma_2 = 1$$

- So, $\Sigma = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

SVD – Ex2

Step 3: *Left-singular vectors* can be computed by $\mathbf{u}_k = \dfrac{\mathbf{A}\mathbf{v}_k}{\sigma_k}$

$$\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}\begin{pmatrix} \frac{5}{\sqrt{30}} \\ \frac{-2}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{pmatrix},$$

$$\mathbf{u}_2 = \frac{\mathbf{A}\mathbf{v}_2}{\sigma_2} = \frac{1}{1}\begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$

We obtain $\mathbf{U} = \dfrac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$

SVD – Ex2

- The SVD of A = $\begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ is

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

SVD vs Eigenvalue Decomposition

SVD $A = U\Sigma V^T$	**Eigenvalue Decomposition** $A = PDP^{-1}$
Always exists for any matrix $\mathbb{R}^{m\times n}$	Only defined for square matrices $\mathbb{R}^{n\times n}$ and *NOT always exists*
The vectors in the matrices U and V in the SVD are *orthonormal*	The vectors in the eigendecomposition matrix P are *not necessarily orthogonal*
U and V are generally not inverse of each other	P and P^{-1} are inverses of each other
entries in the diagonal matrix Σ are all real and *nonnegative*	entries in the diagonal matrix Σ are real or complex
For *symmetric matrices* $A \in \mathbb{R}^{n\times n}$, the *eigenvalue decomposition* and the *SVD* are one and the same, which follows from the *spectral theorem*	

Different versions of SVD

- Our construction (full SVD) $\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\Sigma_{m\times n}\mathbf{V}^{T}{}_{n\times n}$
- Other versions
  - (reduced SVD) For $m \geqslant n$, $\mathbf{A}_{m\times n} = \mathbf{U}_{m\times n}\,\Sigma_{n\times n}\,\mathbf{V}^{T}{}_{n\times n}$
  - For matrix $\mathbf{A}_{m\times n}$ with $rk(A) = r$, $\mathbf{A}_{m\times n} = \mathbf{U}_{m\times r}\,\Sigma_{r\times r}\,\mathbf{V}^{T}{}_{r\times n}$

Matrix Approximation

- A: high rank matrix
- Use SVD to represent A as a sum of simpler (low-rank) matrices

➔ matrix *approximation*

image $\mathbf{A} \in \mathbb{R}^{1145\times1718}$

rank(A) = 1145

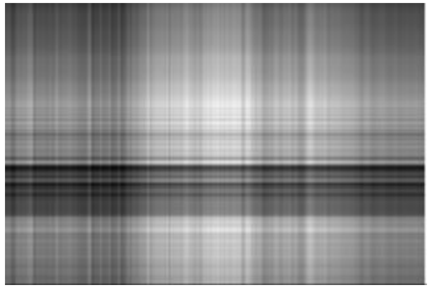
$\hat{\boldsymbol{A}}(50)$

Rank($\hat{\boldsymbol{A}}$) = 50

Matrix Approximation

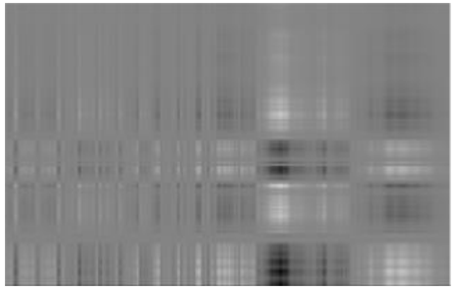
- Define *rank-1* matrices $\mathbf{A}_i := \mathbf{u}_i\mathbf{v}_i^T$

image $\mathbf{A} \in \mathbb{R}^{1145\times1718}$

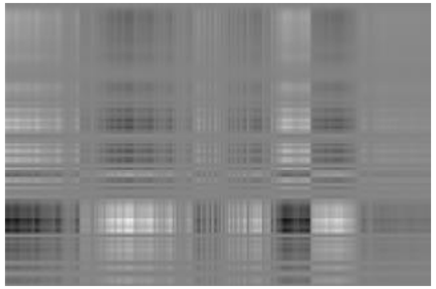


A_1, $\sigma_1 \approx 828$

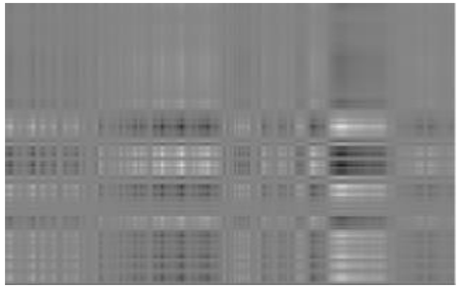
A_2, $\sigma_2 \approx 122.7$



A_3, $\sigma_3 \approx 59.4$



A_4, $\sigma_4 \approx 54$



A_5, $\sigma_5 \approx 48$

Matrix Approximation

- A matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ of *rank r* can be written as a sum of *rank-1* matrices so that

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sigma_1 A_1 + \sigma_2 A_2 + \dots + \sigma_r A_r$$

- **A** can be *approximated* by $\hat{A}(k)$, a *rank-k approximation* of A

$$\mathbf{A} \approx \hat{A}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{k} \sigma_i \mathbf{A}_i$$


Original image **A**


$\hat{A}(5)$


$\hat{A}(10)$


$\hat{A}(50)$

Error of an Approximation

Theorem (Eckart-Young, 1936). Consider a matrix $A \in \mathbb{R}^{m\times n}$ of *rank r* and let $B \in \mathbb{R}^{m\times n}$ be a matrix of *rank k*.
For any $k \leqslant r$, with $\widehat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ it holds that

$$\widehat{\boldsymbol{A}}(k) = \operatorname{argmin}_{rk(B)=k} \|A - B\|_2,$$

$$\|A - \widehat{\boldsymbol{A}}(k)\|_2 = \sigma_{k+1}$$

- The theorem states explicitly how much *error* we introduce by approximating A using a *rank-k approximation*.
- **Note.** $\|A\|_2 := \max_{x\neq 0} \frac{\|Ax\|}{\|x\|}$ (spectral norm of a matrix)

Errors of Approximations – Ex

image $\mathbf{A} \in \mathbb{R}^{1145\times1718}$

$\widehat{\boldsymbol{A}}(5)$, error = 0.647

$\widehat{\boldsymbol{A}}(10)$, error = 0.592

$\widehat{\boldsymbol{A}}(50)$, error = 0.393

$\widehat{\boldsymbol{A}}(100)$, error = 0.280

$\widehat{\boldsymbol{A}}(200)$, error = 0.164

Summary

- Matrix Decomposition
  - Eigenvalue Decomposition $\mathbf{A} = \mathbf{PDP}^{-1}$
    - Eigenvalues and Eigenvectors
  - Singular Value Decomposition (SVD) $\mathbf{A} = \mathbf{U\Sigma V}^T$
    - Left-singular vectors (columns of **U**)
    - Singular values (diagonal entries of singular matrix Σ)
    - Right-singular vectors (columns of **V**)
- Matrix Approximation
  - Rank-1 matrix $\mathbf{A}_i = \mathbf{u}_i\mathbf{v}_i^T$
  - $\mathbf{A} = \sum_{i=1..r} \sigma_i \mathbf{A}_i$
  - $\mathbf{A} \approx \widehat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i \mathbf{A}_i$

THANKS