Selected Topics in Engineering Mathematics: Advanced Matrix Decompositions

Chun-Lin Liu (劉俊麟)

Department of Electrical Engineering
Graduate Institute of Communication Engineering
National Taiwan University

May 28, 2024

Reference

1. R. A. Horn and C. R. Johnson, Matrix Analysis, 2nd ed., New York: Cambridge University Press, 2013.
[HJ2013]
2. G. H. Golub and C. F. Van Loan, Matrix Computations, 4th ed., Baltimore: The Johns Hopkins University Press, 2013.
[GVL2013]
3. J.-J. Ding. (2023). Selected Topics in Engineering Mathematics [PowerPoint slides].

Outline

The Eigen-Decomposition of Square Matrices

- Let $\mathbf{A} \in \mathbb{C}^{N \times N}$.
- There exist N eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{N-1}, \lambda_N$.
- The eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1}, \mathbf{v}_N$ are assumed to be linearly independent.
- The eigen-equations are $\mathbf{A}\mathbf{v}_n = \lambda_n \mathbf{v}_n$ for $n = 1, 2, \ldots, N$.
- Then $\mathbf{A}$ can be decomposed into

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

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_{N-1} & \mathbf{v}_N \end{bmatrix}, \qquad \mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N-1}, \lambda_N). \tag{2}$$

Motivating Questions

1. What if N linearly independent eigenvectors do not exist? Jordan canonical forms.
2. What if the matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ is non-square? Singular value decomposition.

Outline

Outline

Example: Eigen-Decomposition of a Matrix (1/3)

- Find the eigen-decomposition of a matrix $\mathbf{A}$, which is

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 6 \\ 0 & 2 & 5 \\ 0 & 0 & 2 \end{bmatrix}. \tag{3}$$

- First, we consider the characteristic equation $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$.
- For the matrix $\mathbf{A}$ in (3), the characteristic equation becomes

$$\det\left(\begin{bmatrix} 2-\lambda & 1 & 6 \\ 0 & 2-\lambda & 5 \\ 0 & 0 & 2-\lambda \end{bmatrix}\right) = (2-\lambda)^3 = 0. \tag{4}$$

- Therefore, the eigenvalues of $\mathbf{A}$ are $2, 2, 2$.
- The eigenvalue 2 has an algebraic multiplicity of 3.

Example: Eigen-Decomposition of a Matrix (2/3)

- We assume that an eigenvector corresponding to the eigenvalue $\lambda = 2$ is $\mathbf{v}_1 = \begin{bmatrix}\alpha_1 & \beta_1 & \gamma_1\end{bmatrix}^\mathsf{T}$.
- The characteristic equation $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v}_1 = \mathbf{0}$ becomes

$$\begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{5}$$

- Equation (5) leads to $\beta_1 = \gamma_1 = 0$.
- For simplicity, we set $\alpha_1 = 1$.
- The eigenvector $\mathbf{v}_1$ becomes

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{6}$$

Example: Eigen-Decomposition of a Matrix (3/3)

- For simplicity, we set $\alpha_1 = 1$ in (6).
- There is only one independent solution to the eigenvector of $\mathbf{A}$.
- The eigenvalue 2 has a geometric multiplicity of 1.
- Also, there is only one eigen-equation for $\mathbf{A}$:

$$\mathbf{A}\mathbf{v}_1 = (2)\,\mathbf{v}_1. \tag{7}$$

Question

- Can we still decompose $\mathbf{A}$ into $\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}\boldsymbol{\mathcal{V}}^{-1}$?
- The matrix $\boldsymbol{\mathcal{V}}$ contains the (generalized) eigenvectors of $\mathbf{A}$.
- The matrix $\boldsymbol{\mathcal{J}}$ contains the eigenvalues of $\mathbf{A}$.

Example: Generalized Eigenvectors (1/3)

- Continued from the examples from pages 7 to 9
- We define a generalized eigenvector $\mathbf{v}_2 \in \mathbb{C}^3$ satisfying

$$(\mathbf{A} - \lambda \mathbf{I})\, \mathbf{v}_2 = \mathbf{v}_1. \tag{8}$$

$$\begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{9}$$

- (Exercise) It can be shown that

$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \tag{10}$$

  is a solution to (8).
- In addition, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Example: Generalized Eigenvectors (2/3)

- We define another generalized eigenvector $\mathbf{v}_3 \in \mathbb{C}^3$ satisfying

$$(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v}_3 = \mathbf{v}_2. \tag{11}$$

$$\begin{bmatrix} 0 & 1 & 6 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{12}$$

- We select

$$\mathbf{v}_3 = \begin{bmatrix} 0 \\ -\frac{6}{5} \\ \frac{1}{5} \end{bmatrix}, \tag{13}$$

such that (11) is satisfied and $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent.

Example: Generalized Eigenvectors (3/3)

- Equations (7), (8), and (11) can be rewritten as

$$\mathbf{A}\mathbf{v}_1 = \lambda\mathbf{v}_1, \qquad \mathbf{A}\mathbf{v}_2 = \lambda\mathbf{v}_2 + \mathbf{v}_1, \qquad \mathbf{A}\mathbf{v}_3 = \lambda\mathbf{v}_3 + \mathbf{v}_2, \tag{14}$$

- We obtain

$$\mathbf{A}\underbrace{\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix}}_{\boldsymbol{\mathcal{V}}} = \underbrace{\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3\end{bmatrix}}_{\boldsymbol{\mathcal{V}}}\underbrace{\begin{bmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{bmatrix}}_{\boldsymbol{\mathcal{J}}}. \tag{15}$$

- Since $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent, the matrix $\boldsymbol{\mathcal{V}}$ is invertible. We have

$$\mathbf{A} = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}\boldsymbol{\mathcal{V}}^{-1}. \tag{16}$$

- $\boldsymbol{\mathcal{J}}$ is the Jordan canonical form of $\mathbf{A}$.

The Jordan Canonical Form

- We decompose the matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ into $\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}\boldsymbol{\mathcal{V}}^{-1}$.
- The matrix $\boldsymbol{\mathcal{V}}$ contains the (generalized) eigenvectors.
- The Jordan canonical form $\boldsymbol{\mathcal{J}}$ of $\mathbf{A}$ is a block diagonal matrix of the form

$$\boldsymbol{\mathcal{J}} = \mathrm{blkdiag}\left(\boldsymbol{\mathcal{J}}_1, \boldsymbol{\mathcal{J}}_2, \ldots, \boldsymbol{\mathcal{J}}_K\right). \tag{17}$$

- For every $k \in [K]$, the Jordan block $\boldsymbol{\mathcal{J}}_k$ has the form of

$$\boldsymbol{\mathcal{J}}_k = \lambda_k \mathbf{I}_{L_k} + \mathbf{U}_{L_k}, \tag{18}$$

  for some $L_k \in [N]$.
- The matrix $\mathbf{I}_{L_k}$ denotes the identity matrix of size L_k by L_k.
- The matrix $\mathbf{U}_{L_k}$ is an upper shift matrix of size L_k by L_k.
- Let $(i, j) \in [L_k]^2$. The (i, j)th entry of $\mathbf{U}_{L_k}$ is

$$[\mathbf{U}_{L_k}]_{i,j} = \delta_{i+1,j}. \tag{19}$$

Examples: The Jordan Blocks

- If $k = 1$ and $L_k = 1$, then

$$\mathcal{J}_1 = \lambda_1 \mathbf{I}_1 + \mathbf{U}_1 = \lambda_1.$$

($\mathcal{J}_1$ becomes a scalar)

- If $k = 2$ and $L_k = 2$, then

$$\mathcal{J}_2 = \lambda_2 \mathbf{I}_2 + \mathbf{U}_2 = \begin{bmatrix} \lambda_2 & 1 \\ 0 & \lambda_2 \end{bmatrix}.$$

- If $k = 3$ and $L_k = 3$, then

$$\mathcal{J}_3 = \lambda_3 \mathbf{I}_3 + \mathbf{U}_3 = \begin{bmatrix} \lambda_3 & 1 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Example: The Jordan Canonical Form of a 4-by-4 Matrix (1/4)

- We consider the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 & -2 \\ 0 & 4 & -2 & 0 \\ -1 & -1 & 4 & 0 \\ -1 & -1 & 0 & 4 \end{bmatrix} \tag{20}$$

- From the characteristic equation, the eigenvalues of $\mathbf{A}$ are $\lambda = 2, 4, 4, 6$.
- For $\lambda = 2$, it can be shown that $\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^{\mathsf{T}}$ is an eigenvector.
- For $\lambda = 6$, it can be shown that $\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}^{\mathsf{T}}$ is an eigenvector.

Example: The Jordan Canonical Form of a 4-by-4 Matrix (2/4)

- For $\lambda = 4$, the eigenvector is assumed to be $\mathbf{v}_1 = \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 & \delta_1 \end{bmatrix}^{\mathsf{T}}$.
- The equation $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v}_1 = \mathbf{0}$ becomes

$$\begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \\ \delta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{21}$$

- For $\lambda = 4$, there is only one linearly independent eigenvector:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}. \tag{22}$$

Example: The Jordan Canonical Form of a 4-by-4 Matrix (3/4)

- As a result, we need to find the generalized eigenvector $\mathbf{v}_2 = \begin{bmatrix} \alpha_2 & \beta_2 & \gamma_2 & \delta_2 \end{bmatrix}^{\mathsf{T}}$.
- The equation $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v}_2 = \mathbf{v}_1$ can be expressed as

$$\begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \beta_2 \\ \gamma_2 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}. \tag{23}$$

- For $\lambda = 4$, the generalized eigenvector $\mathbf{v}_2$ is

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 1/2 \\ -1/2 \end{bmatrix}. \tag{24}$$

Example: The Jordan Canonical Form of a 4-by-4 Matrix (4/4)

- Based on the discussions on pages 15, 16, and 17, we obtain

$$\mathbf{A} = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}\boldsymbol{\mathcal{V}}^{-1}, \tag{25}$$

where

$$\boldsymbol{\mathcal{V}} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & 1/2 & -1 \\ 1 & 0 & -1/2 & -1 \end{bmatrix}, \qquad \boldsymbol{\mathcal{J}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}. \tag{26}$$

Example: The Jordan Canonical Form of a 5-by-5 Matrix (1/5)

- As an example, let the matrix $\mathbf{A}$ be

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 1 & 2 & 1 \\ 0 & 4 & 1 & -2 & 1 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 1 & 0 & 3 \end{bmatrix}. \tag{27}$$

- Solving the characteristic equation of $\mathbf{A}$ leads to the eigenvalues

$$\lambda = 2, \quad 4, \quad 4, \quad 4, \quad 4. \tag{28}$$

- For $\lambda = 2$, it can be shown that $\begin{bmatrix} 0 & 0 & 1 & 0 & -1 \end{bmatrix}^{\mathsf{T}}$ is an eigenvector.

Example: The Jordan Canonical Form of a 5-by-5 Matrix (2/5)

- For $\lambda = 4$, the eigenvector is assumed to be $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & v_5 \end{bmatrix}^{\mathsf{T}}$.
- From the equation $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{v} = \mathbf{0}$, we obtain

$$\begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{29}$$

- For $\lambda = 4$, there are only two linearly independent solutions, denoted by $\boldsymbol{\phi}_1$ and $\boldsymbol{\psi}_1$:

$$\boldsymbol{\phi}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathsf{T}}, \qquad \boldsymbol{\psi}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}^{\mathsf{T}}. \tag{30}$$

Example: The Jordan Canonical Form of a 5-by-5 Matrix (3/5)

- For $\lambda = 4$ and the eigenvector $\phi_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}^\mathsf{T}$, we solve the equation $(\mathbf{A} - \lambda \mathbf{I})\, \phi_2 = \phi_1$ for the generalized eigenvector.
- We obtain

$$\begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \phi_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (31)$$

- A solution to (31) is

$$\phi_2 = \begin{bmatrix} 0 & 0 & 1/4 & 1/4 & 1/4 \end{bmatrix}^\mathsf{T}, \qquad (32)$$

where the first and the second entries of ϕ_2 are set to zero for simplicity.

Example: The Jordan Canonical Form of a 5-by-5 Matrix (4/5)

- For $\lambda = 4$ and the eigenvector $\boldsymbol{\psi}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}^\mathsf{T}$, we solve the equation $(\mathbf{A} - \lambda \mathbf{I})\,\boldsymbol{\psi}_2 = \boldsymbol{\psi}_1$ for the generalized eigenvector.
- We obtain

$$\begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix} \boldsymbol{\psi}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{33}$$

- A solution to (33) is

$$\boldsymbol{\psi}_2 = \begin{bmatrix} 0 & 0 & 1/4 & -1/4 & 1/4 \end{bmatrix}^\mathsf{T}. \tag{34}$$

Example: The Jordan Canonical Form of a 5-by-5 Matrix (5/5)

- Therefore, we can decompose the matrix $\mathbf{A}$ into

$$\mathbf{A} = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}\boldsymbol{\mathcal{V}}^{-1}, \tag{35}$$

where

$$\boldsymbol{\mathcal{V}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 1 \\ 0 & 1/4 & 0 & -1/4 & 0 \\ 0 & 1/4 & 0 & 1/4 & -1 \end{bmatrix}, \qquad \boldsymbol{\mathcal{J}} = \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}. \tag{36}$$

Outline

The Integer Power of a Matrix

- We consider the Jordan canonical form of a matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$,

$$\mathbf{A} = \boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{V}}^{-1}. \tag{37}$$

- For a non-negative integer α, the matrix power $\mathbf{A}^{\alpha}$ becomes

$$\mathbf{A}^{\alpha} = \underbrace{\left(\boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{V}}^{-1}\right)\left(\boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{V}}^{-1}\right) \cdots \left(\boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{V}}^{-1}\right)}_{\alpha \text{ terms}} \tag{38}$$

$$= \boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}} \underbrace{\left(\boldsymbol{\mathcal{V}}^{-1} \boldsymbol{\mathcal{V}}\right)}_{\mathbf{I}} \boldsymbol{\mathcal{J}} \underbrace{\left(\boldsymbol{\mathcal{V}}^{-1} \boldsymbol{\mathcal{V}}\right)}_{\mathbf{I}} \boldsymbol{\mathcal{J}} \cdots \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{V}}^{-1} \tag{39}$$

$$= \boldsymbol{\mathcal{V}} \boldsymbol{\mathcal{J}}^{\alpha} \boldsymbol{\mathcal{V}}^{-1}. \tag{40}$$

- (Question) How do you determine $\boldsymbol{\mathcal{J}}^{\alpha}$?

The Power of $\mathcal{J}$

- From (17), we obtain

$$\boldsymbol{\mathcal{J}}^{\alpha} = \mathrm{blkdiag}\left(\boldsymbol{\mathcal{J}}_1^{\alpha}, \boldsymbol{\mathcal{J}}_2^{\alpha}, \ldots, \boldsymbol{\mathcal{J}}_K^{\alpha}\right). \tag{41}$$

- After dropping the subscript L_k in (18) for simplicity, we rewrite the matrix $\boldsymbol{\mathcal{J}}_k^{\alpha}$ as

$$\boldsymbol{\mathcal{J}}_k^{\alpha} = (\lambda_k \mathbf{I} + \mathbf{U})^{\alpha} = \sum_{\ell=0}^{\alpha} \binom{\alpha}{\ell} (\lambda_k \mathbf{I})^{\alpha-\ell}\, \mathbf{U}^{\ell} \tag{42}$$

$$= \sum_{\ell=0}^{\alpha} \binom{\alpha}{\ell} \lambda_k^{\alpha-\ell} \mathbf{U}^{\ell}. \tag{43}$$

- (Cross reference) The binomial expansion for scalars

$$(x+y)^N = \sum_{n=0}^{N} \binom{N}{n} x^{N-n} y^n, \qquad \binom{N}{n} = \frac{N!}{(N-n)!n!}. \tag{44}$$

Examples of the Powers of $\mathbf{U}$

- For instance, we assume that $\mathbf{U} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.
- The powers of $\mathbf{U}$ are

$$\mathbf{U}^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{U}^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{U}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

- It can be shown that $\mathbf{U}^\ell = \mathbf{0}$ for $\ell \geq 5$.

The General Form of $\mathbf{U}^\ell$

- If $\ell < L_k$, then $\mathbf{U}_{L_k}^\ell$ satisfies

$$\left[\mathbf{U}_{L_k}^\ell\right]_{m,n} = \delta_{n-m,\ell} = \begin{cases} 1, & \text{if } n - m = \ell, \\ 0, & \text{otherwise.} \end{cases} \tag{45}$$

- If $\ell \geq L_k$, then $\mathbf{U}_{L_k}^\ell = \mathbf{0}$.

The General Form of $\mathcal{J}_k^\alpha$

Powers of a Jordan block

The kth eigenvalue is denoted by λ_k. Let α be a non-negative integer. Let $\mathcal{J}_k$ be the kth Jordan block. Then

$$[\mathcal{J}_k^\alpha]_{m,n} = \begin{cases} \lambda_k^\alpha, & \text{if } m = n, \\ \binom{\alpha}{n-m}\lambda_k^{\alpha-n+m}, & \text{if } n > m \text{ and } \alpha \geq n-m, \\ 0, & \text{otherwise.} \end{cases} \tag{46}$$

An Example of $\mathcal{J}_k^\alpha$

- We assume that $k = 1$, $L_k = 5$, and $\alpha = 3$
- Then

$$\mathcal{J}_1^3 = \begin{bmatrix} \lambda_1^3 & \binom{\alpha}{1}\lambda_1^2 & \binom{\alpha}{2}\lambda_1^1 & \binom{\alpha}{3}\lambda_1^0 & 0 \\ 0 & \lambda_1^3 & \binom{\alpha}{1}\lambda_1^2 & \binom{\alpha}{2}\lambda_1^1 & \binom{\alpha}{3}\lambda_1^0 \\ 0 & 0 & \lambda_1^3 & \binom{\alpha}{1}\lambda_1^2 & \binom{\alpha}{2}\lambda_1^1 \\ 0 & 0 & 0 & \lambda_1^3 & \binom{\alpha}{1}\lambda_1^2 \\ 0 & 0 & 0 & 0 & \lambda_1^3 \end{bmatrix}.$$

Example: The Power of a 4-by-4 Matrix (1/2)

- Find the matrix power $\mathbf{A}^5$, where

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 & -2 \\ 0 & 4 & -2 & 0 \\ -1 & -1 & 4 & 0 \\ -1 & -1 & 0 & 4 \end{bmatrix}. \tag{47}$$

- According to the example on pages 15 to 18, the matrix power $\mathbf{A}^5$ becomes

$$\mathbf{A}^5 = \boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{J}}^5\boldsymbol{\mathcal{V}}^{-1} = \boldsymbol{\mathcal{V}}\mathrm{blkdiag}\left(\boldsymbol{\mathcal{J}}_1^5, \boldsymbol{\mathcal{J}}_2^5, \boldsymbol{\mathcal{J}}_3^5\right)\boldsymbol{\mathcal{V}}^{-1}. \tag{48}$$

- The Jordan blocks are

$$\boldsymbol{\mathcal{J}}_1 = 2, \qquad \boldsymbol{\mathcal{J}}_2 = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}, \qquad \boldsymbol{\mathcal{J}}_3 = 6, \tag{49}$$

Example: The Power of a 4-by-4 Matrix (2/2)

- The powers of the Jordan blocks can be expressed as

$$\mathcal{J}_1^5 = 2^5 = 32, \tag{50}$$

$$\mathcal{J}_2^5 = \begin{bmatrix} 4^5 & \binom{5}{1} \times 4^4 \\ 0 & 4^5 \end{bmatrix} = \begin{bmatrix} 1024 & 1280 \\ 0 & 1024 \end{bmatrix}, \tag{51}$$

$$\mathcal{J}_3^5 = 6^5 = 7776. \tag{52}$$

- Substituting (50), (50), and (50) into (48) yields

$$\mathbf{A}^5 = \begin{bmatrix} 2464 & 1440 & -656 & -3216 \\ 1440 & 2464 & -3216 & -656 \\ -1936 & -1936 & 2464 & 1440 \\ -1936 & -1936 & 1440 & 2464 \end{bmatrix}. \tag{53}$$

Outline

Outline

The Eigen-Decomposition of Hermitian Matrices

- Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ and $\mathbf{A}^{\mathsf{H}} = \mathbf{A}$ (Hermitian matrices).
- The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ are real numbers.
- After normalization, the set of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1}, \mathbf{v}_N\}$ is complete and orthonormal.
- The eigen-decomposition of a Hermitian matrix $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathsf{H}} = \sum_{n=1}^{N} \lambda_n \mathbf{v}_n \mathbf{v}_n^{\mathsf{H}}. \quad (54)$$

Motivating Questions

1. How do we extend the decomposition to M-by-N (non-square) matrices?

The Singular Value Decomposition [HJ2013, pp. 150], [GVL2013, pp. 76]

- We assume that $\mathbf{A} \in \mathbb{C}^{M \times N}$, $q = \min\{M, N\}$, and $\operatorname{rank}(\mathbf{A}) = r$.
- There are unitary matrices $\mathbf{U} \in \mathbb{C}^{M \times M}$ and $\mathbf{V} \in \mathbb{C}^{N \times N}$, and a square diagonal matrix

$$\boldsymbol{\Sigma}_q = \operatorname{diag}\left(\sigma_1, \sigma_2, \ldots, \sigma_q\right). \tag{55}$$

such that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_q, \tag{56}$$

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}}, \qquad \boldsymbol{\Sigma} = \begin{cases} \boldsymbol{\Sigma}_q \in \mathbb{R}^{M \times N} & \text{if } M = N, \\ \begin{bmatrix} \boldsymbol{\Sigma}_q & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{M \times N} & \text{if } M < N, \\ \begin{bmatrix} \boldsymbol{\Sigma}_q \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{M \times N} & \text{if } M > N, \end{cases} \tag{57}$$

Terminologies

- The scalars $\sigma_1, \sigma_2, \ldots, \sigma_q$ are the singular values of $\mathbf{A}$.
- The largest singular value of $\mathbf{A}$ is denoted by $\sigma_{\max}(\mathbf{A}) = \sigma_1$.
- Let

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_M \end{bmatrix} \in \mathbb{C}^{M \times M}. \quad (58)$$

  The column vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M$ are the left singular vectors of $\mathbf{A}$.
- Let

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_N \end{bmatrix} \in \mathbb{C}^{N \times N}. \quad (59)$$

  The column vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are the right singular vectors of $\mathbf{A}$.

An Example of the SVD

- It can be verified that

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ -1 & 1 \\ -1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{\mathsf{H}}}_{\mathbf{V}^{\mathsf{H}}}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{8} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}^{\mathsf{H}}.$$

- (Questions) How do we find the singular values and singular vectors for a matrix $\mathbf{A}$?

SVD and Eigen-Decompositions (1/2)

- Assume that $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}}$ is the SVD of $\mathbf{A} \in \mathbb{C}^{M \times N}$.
- The matrix $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ can be expressed as

$$\begin{aligned} \mathbf{A}\mathbf{A}^{\mathsf{H}} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}} \left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}}\right)^{\mathsf{H}} \\ &= \mathbf{U} \left(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathsf{H}}\right) \mathbf{U}^{\mathsf{H}}. \end{aligned} \tag{60}$$

- Remarks on (60):
  - The left singular vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_M$ are the eigenvectors of $\mathbf{A}\mathbf{A}^{\mathsf{H}}$.
  - The matrix $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathsf{H}}$ contains the eigenvalues of $\mathbf{A}\mathbf{A}^{\mathsf{H}}$.

SVD and Eigen-Decompositions (2/2)

- Similarly, the matrix $\mathbf{A}^{\mathsf{H}}\mathbf{A}$ can be expressed as

$$\begin{aligned}\mathbf{A}^{\mathsf{H}}\mathbf{A} &= \left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}}\right)^{\mathsf{H}}\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{H}} \\ &= \mathbf{V}\left(\boldsymbol{\Sigma}^{\mathsf{H}}\boldsymbol{\Sigma}\right)\mathbf{V}^{\mathsf{H}}. \end{aligned} \tag{61}$$

- Remarks on (61):
  - The right singular vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are the eigenvectors of $\mathbf{A}^{\mathsf{H}}\mathbf{A}$.
  - The matrix $\boldsymbol{\Sigma}^{\mathsf{H}}\boldsymbol{\Sigma}$ contains the eigenvalues of $\mathbf{A}\mathbf{A}^{\mathsf{H}}$.
- How do we find both the left and right singular vectors?

Relations among $\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}$ (1/2)

A Property rephrased from [GVL2013, Corollary 2.4.2]

If $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{H}}$ is the SVD of $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $M \geq N$, then for $i \in [N]$, we have

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i, \qquad \mathbf{A}^{\mathsf{H}}\mathbf{u}_i = \sigma_i \mathbf{v}_i. \tag{62}$$

- Proof sketch (1/2): We rewrite the SVD as $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$, which is

$$\mathbf{A}\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_N\end{bmatrix}$$
$$= \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_N & \mathbf{u}_{N+1} & \dots & \mathbf{u}_M\end{bmatrix}\begin{bmatrix}\sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0}\end{bmatrix}.$$

Relations among $\mathbf{U}$, $\boldsymbol{\Sigma}$, and $\mathbf{V}$ (2/2)

- Proof sketch (2/2): The SVD of $\mathbf{A}^{\mathsf{H}}$ can be expressed as

$$\mathbf{A}^{\mathsf{H}} = \left(\mathbf{U} \begin{bmatrix} \boldsymbol{\Sigma}_q \\ \mathbf{0} \end{bmatrix} \mathbf{V}^{\mathsf{H}} \right)^{\mathsf{H}} \tag{63}$$

$$= \mathbf{V} \begin{bmatrix} \boldsymbol{\Sigma}_q^{\mathsf{H}} & \mathbf{0}^{\mathsf{H}} \end{bmatrix} \mathbf{U}^{\mathsf{H}}. \tag{64}$$

$$= \mathbf{V} \begin{bmatrix} \boldsymbol{\Sigma}_q & \mathbf{0} \end{bmatrix} \mathbf{U}^{\mathsf{H}}. \tag{65}$$

  Comparing the columns of $\mathbf{A}^{\mathsf{H}}\mathbf{U} = \mathbf{V} \begin{bmatrix} \boldsymbol{\Sigma}_q & \mathbf{0} \end{bmatrix}$ shows the second equation in (62).

- Remarks on (65):
  - The matrices $\mathbf{A}$ and $\mathbf{A}^{\mathsf{H}}$ have the same singular values.
  - The left singular vectors of $\mathbf{A}$ become the right singular vectors of $\mathbf{A}^{\mathsf{H}}$.

Computation of the Singular Vectors

- If $\sigma_i \neq 0$, then (62) can be rewritten as

$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i}, \quad (66)$$

$$\mathbf{v}_i = \frac{\mathbf{A}^{\mathsf{H}}\mathbf{u}_i}{\sigma_i}. \quad (67)$$

- Implications of (66) and (67)
  - If the matrix $\mathbf{A}$, the non-zero singular values, and one set of singular vectors are provided, we can uniquely determine another set of singular vectors.

An Example of the SVD (1/3)

- Consider the matrix $\mathbf{A}$ on page 38. We obtain

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ -1 & 1 \\ -1 & 1 \end{bmatrix}, \qquad \mathbf{A}\mathbf{A}^{\mathsf{H}} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix}.$$

- The characteristic equation

$$\det\left(\mathbf{A}\mathbf{A}^{\mathsf{H}} - \lambda \mathbf{I}\right) = -\left(\lambda - 8\right)\left(\lambda - 4\right)\lambda = 0.$$

- The eigenvalues and eigenvectors are

$$\lambda_1(\mathbf{A}\mathbf{A}^{\mathsf{H}}) = 8, \qquad \lambda_2(\mathbf{A}\mathbf{A}^{\mathsf{H}}) = 4, \qquad \lambda_3(\mathbf{A}\mathbf{A}^{\mathsf{H}}) = 0, \tag{68}$$

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{u}_2 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \qquad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}. \tag{69}$$

An Example of the SVD (2/3)

- From the definition of SVD on page 36, we obtain

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}. \tag{70}$$

- According to (60), the matrix $\mathbf{\Sigma}\mathbf{\Sigma}^{\mathsf{H}}$ contains the eigenvalues of $\mathbf{A}\mathbf{A}^{\mathsf{H}}$.

$$\mathbf{\Sigma}\mathbf{\Sigma}^{\mathsf{H}} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{71}$$

- Since $\sigma_1, \sigma_2 \geq 0$, we obtain

$$\sigma_1 = \sqrt{8}, \qquad \sigma_2 = 2. \tag{72}$$

An Example of the SVD (3/3)

- Substituting (69) and (72) into (67) yields

$$\mathbf{v}_1 = \frac{\mathbf{A}^{\mathsf{H}}\mathbf{u}_1}{\sigma_1} = \frac{1}{\sqrt{8}} \begin{bmatrix} 2 & -1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad (73)$$

$$\mathbf{v}_2 = \frac{\mathbf{A}^{\mathsf{H}}\mathbf{u}_2}{\sigma_2} = \frac{1}{2} \begin{bmatrix} 2 & -1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}. \quad (74)$$

Outline

The Operator Norm [GVL2013, pp. 72]

- The operator norm $\|\mathbf{A}\|_{\alpha,\beta}$ is defined as

$$\|\mathbf{A}\|_{\alpha,\beta} \triangleq \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_\beta}{\|\mathbf{x}\|_\alpha}. \tag{75}$$

- $\|\cdot\|_{\alpha,\beta}$ is subordinate to the vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$.

The Matrix p-Norm

- By setting $\alpha = \beta = p$, the matrix p-norm is defined as

$$\|\mathbf{A}\|_p \triangleq \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}. \tag{76}$$

- According to (76), it can be shown that [HJ2013, pp. 344-345], [GVL2013, pp. 72]:

$$\|\mathbf{A}\|_1 = \max_{1 \leq j \leq N} \sum_{i=1}^{M} \left| [\mathbf{A}]_{i,j} \right|, \tag{77}$$

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq M} \sum_{j=1}^{N} \left| [\mathbf{A}]_{i,j} \right|. \tag{78}$$

- If $p = 2$, then $\|\mathbf{A}\|_2$ is the matrix 2-norm of $\mathbf{A}$.

The Matrix Norms and the Singular Values

- Assume that $\mathbf{A} \in \mathbb{C}^{M \times N}$ has singular values (c.f. page 36)

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_q. \tag{79}$$

- Then, the matrix 2-norm and the Frobenius norm of $\mathbf{A}$ satisfy [GVL2013, pp. 77]:

$$\|\mathbf{A}\|_2 = \sigma_1, \tag{80}$$

$$\|\mathbf{A}\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_q^2}. \tag{81}$$

The Interpretation of Matrix Norms

- The matrix $\mathbf{A}$ is mapped to a vector $\boldsymbol{\sigma}$

$$\boldsymbol{\sigma} \triangleq \begin{bmatrix} \sigma_1 & \sigma_2 & \dots & \sigma_r & \sigma_{r+1} & \dots & \sigma_q \end{bmatrix}^{\mathsf{T}}. \tag{82}$$

- Then, the matrix 2-norm and the Frobenius norm of $\mathbf{A}$ satisfy

$$\underbrace{\|\mathbf{A}\|_2}_{\text{matrix } 2\text{-norm}} = \underbrace{\|\boldsymbol{\sigma}\|_\infty}_{\text{vector } \infty\text{-norm}}, \tag{83}$$

$$\underbrace{\|\mathbf{A}\|_F}_{\text{Frobenius norm}} = \underbrace{\|\boldsymbol{\sigma}\|_2}_{\text{vector } 2\text{-norm}}. \tag{84}$$

The Rank of a Matrix

- Based on the vector $\boldsymbol{\sigma}$, the rank of a matrix $\mathbf{A}$ satisfies

$$\operatorname{rank}(\mathbf{A}) = \underbrace{\|\boldsymbol{\sigma}\|_0}_{\ell_0 \text{ function}} = \operatorname{card}(\operatorname{supp}(\boldsymbol{\sigma})). \tag{85}$$

- The rank of $\mathbf{A}$ is the number of non-zero singular values.
- Low-rank optimization in signal processing

The Nuclear Norm

- Based on the vector $\boldsymbol{\sigma}$, the nuclear norm of a matrix $\mathbf{A}$ is defined as

$$\|\mathbf{A}\|_* = \underbrace{\|\boldsymbol{\sigma}\|_1}_{\text{vector } 1\text{-norm}} = \sum_{i=1}^{q} \sigma_i. \tag{86}$$

- The nuclear norm is viewed as a convex surrogate of the rank function.

Outline

The Data Vectors

- Consider a set of data vectors (row vectors)

$$\mathbf{x}_m = \begin{bmatrix} x_{m,1} & x_{m,2} & x_{m,3} & \dots & x_{m,N} \end{bmatrix}, \tag{87}$$

  for $m = 1, 2, \dots M$.
- The number of data vectors: M
- The length of a data vector: N
- Usually $M \gg N$.
- Applications
  - Audio signals
  - Images
  - Communication signals
  - Array signal processing (linear arrays or planar arrays)

Mean Subtraction

- The mean vector $\overline{\mathbf{x}}$ (as a row vector) is

$$\overline{\mathbf{x}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}_m. \tag{88}$$

- The new data vector $\mathbf{a}_m$ after subtracting the mean vector from $\mathbf{x}_m$

$$\mathbf{a}_m \triangleq \mathbf{x}_m - \overline{\mathbf{x}}. \tag{89}$$

The Data Matrix

- The data matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_M \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 - \overline{\mathbf{x}} \\ \mathbf{x}_2 - \overline{\mathbf{x}} \\ \vdots \\ \mathbf{x}_M - \overline{\mathbf{x}} \end{bmatrix}. \tag{90}$$

- The data vector $\mathbf{x}_m$ can be expressed as

$$\mathbf{x}_m = \mathbf{e}_m^{\mathsf{T}} \mathbf{A} + \overline{\mathbf{x}}, \tag{91}$$

where $\mathbf{e}_m \in \mathbb{C}^M$ satisfies

$$[\mathbf{e}_m]_i = \begin{cases} 1 & \text{if } i = m, \\ 0 & \text{if } i \neq m. \end{cases} \tag{92}$$

SVD of $\mathbf{A}$

- According to Page 37, the SVD of $\mathbf{A}$ is

$$\begin{aligned} \mathbf{A} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{H}} && (93) \\ &= \sum_{i=1}^{N} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{H}} && (94) \\ &= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^{\mathsf{H}} + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^{\mathsf{H}} + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^{\mathsf{H}} + \cdots + \sigma_N \mathbf{u}_N \mathbf{v}_N^{\mathsf{H}}. && (95) \end{aligned}$$

- The singular values satisfy

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_N \geq 0. \tag{96}$$

- The ith component of $\mathbf{A}$ is $\sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{H}}$.

Dimensionality Reduction (1/2)

- We approximate the matrix $\mathbf{A}$ by L components:

$$\widehat{\mathbf{A}} \triangleq \sum_{i=1}^{L} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{H}} \tag{97}$$

$$= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^{\mathsf{H}} + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^{\mathsf{H}} + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^{\mathsf{H}} + \cdots + \sigma_L \mathbf{u}_L \mathbf{v}_L^{\mathsf{H}}. \tag{98}$$

- Dimensional reduction: $L \leq N$.

Dimensionality Reduction (2/2)

- According to (91) and (97), we define the approximated data vectors

$$\widehat{\mathbf{x}}_m \triangleq \mathbf{e}_m^{\mathsf{T}} \widehat{\mathbf{A}} + \overline{\mathbf{x}} = \left(\sum_{i=1}^{L} \sigma_i \left(\mathbf{e}_m^{\mathsf{T}} \mathbf{u}_i \right) \mathbf{v}_i^{\mathsf{H}} \right) + \overline{\mathbf{x}}. \tag{99}$$

  - $\mathbf{e}_m^{\mathsf{T}} \mathbf{u}_i$ is the mth entry of $\mathbf{u}_i$.
  - $\sigma_i \left(\mathbf{e}_m^{\mathsf{T}} \mathbf{u}_i \right)$ is the combination coefficient.
  - The set $\{\mathbf{v}_1^{\mathsf{H}}, \mathbf{v}_2^{\mathsf{H}}, \ldots, \mathbf{v}_L^{\mathsf{H}}\}$ contains the axes.
- A general form of the approximated data vectors is

$$\left(\sum_{i=1}^{L} c_i \mathbf{v}_i^{\mathsf{H}} \right) + \overline{\mathbf{x}}, \tag{100}$$

where $c_i \in \mathbb{C}$ for $i = 1, 2, \ldots L$.

An Example of the PCA (1/4)

Problem

Use the PCA with $L = 1$ to find a regression line that approximates the points in $\mathbb{R}^2$

$$\mathbf{x}_1 = \begin{bmatrix} 7 & 8 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 9 & 8 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 10 & 10 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} 11 & 12 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} 13 & 12 \end{bmatrix}.$$

We assume that the combination coefficients are real numbers.

- (Solution) The number of data $M = 5$.
- The length of the data vector $N = 2$.
- The mean vector

$$\overline{\mathbf{x}} = \begin{bmatrix} 10 & 10 \end{bmatrix}.$$

An Example of the PCA (2/4)

- The new data vectors

$$\mathbf{a}_1 = \begin{bmatrix} -3 & -2 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -1 & -2 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 & 0 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \mathbf{a}_5 = \begin{bmatrix} 3 & 2 \end{bmatrix}.$$

- The data matrix $\mathbf{A}$ and its SVD

$$\mathbf{A} = \begin{bmatrix} -3 & -2 \\ -1 & -2 \\ 0 & 0 \\ 1 & 2 \\ 3 & 2 \end{bmatrix}. \tag{101}$$

An Example of the PCA (3/4)

- The SVD of $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{H}}$, where

$$\mathbf{U} = \begin{bmatrix} -0.6116 & 0.3549 & 0 & 0.0393 & 0.7060 \\ -0.3549 & -0.6116 & 0 & 0.7060 & -0.0393 \\ 0 & 0 & 1.0000 & 0 & 0 \\ 0.3549 & 0.6116 & 0 & 0.7060 & -0.0393 \\ 0.6116 & -0.3549 & 0 & 0.0393 & 0.7060 \end{bmatrix}, \quad \mathbf{\Sigma} = \begin{bmatrix} 5.8416 & 0 \\ 0 & 1.3695 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{V} = \begin{bmatrix} 0.7497 & -0.6618 \\ 0.6618 & 0.7497 \end{bmatrix}.$$

An Example of the PCA (4/4)

- For $L = 1$ in (97), we obtain

$$\widehat{\mathbf{A}} = \underbrace{(5.8416)}_{\sigma_1} \underbrace{\begin{bmatrix} -0.6116 \\ -0.3549 \\ 0 \\ 0.3549 \\ 0.6116 \end{bmatrix}}_{\mathbf{u}_1} \underbrace{\begin{bmatrix} 0.7497 & -0.6618 \end{bmatrix}}_{\mathbf{v}_1^{\mathsf{H}}}.$$

- According to (100) and page 61, an approximation of the data points is

$$\begin{bmatrix} 10 & 10 \end{bmatrix} + c \begin{bmatrix} 0.7497 & -0.6618 \end{bmatrix},$$

where $c \in \mathbb{R}$.