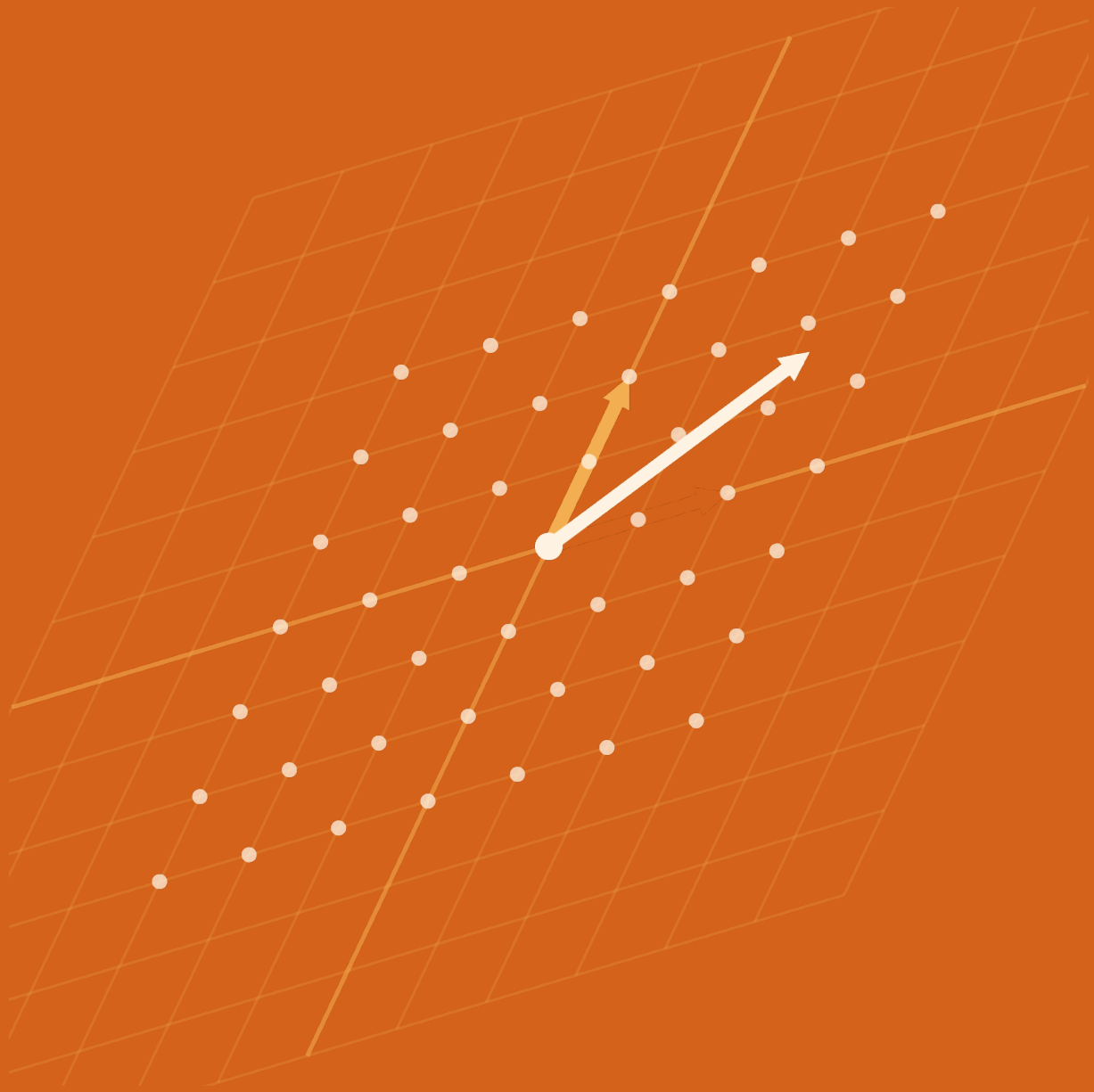
LINEAR ALGEBRA FOR AIML AND 6G

Pseudoinverse and Jordan Normal Form



Lecture 19

Prof. Aditya K. Jagannatham
IIT Kanpur

Outline

1. **Pseudoinverse**
   Pseudoinverse from the SVD; left and right inverses
2. **Incomplete eigenspace**
   A repeated eigenvalue with too few eigenvectors
3. **Jordan normal form**
   Jordan blocks and the Jordan canonical form

01

Pseudoinverse

Pseudoinverse from the SVD; left and right inverses



Pseudoinverse

- Consider the $m \times n$ matrix Σ with **non-zero** elements σ_i only along the **principal diagonal**
- Note that such a matrix Σ does not have an inverse if $m \neq n$

Pseudoinverse

- Its **pseudoinverse** is denoted by the $n \times m$ matrix $\Sigma^{\dagger}$ and is defined as follows
- $\Sigma^{\dagger}$ is a diagonal matrix and comprises as its **principal diagonal** entries the reciprocals of the non-zero **singular values** i.e. $1/\sigma_1$, …, $1/\sigma_r$

Pseudoinverse

- $\Sigma^\dagger$ is given below

$$\Sigma^\dagger = \begin{bmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ & & & \mathbf{0} \end{bmatrix}$$

Pseudoinverse – example

- Consider the 3×4 matrix Σ

$$\Sigma = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Pseudoinverse – example

- Its pseudoinverse $\Sigma^\dagger$ is given as the 4×3 matrix

$$\Sigma^\dagger = \begin{bmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Pseudoinverse – property

- Note that if $m \geq n = r$

$$\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma} = \mathbf{I}$$

- where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

Pseudoinverse – property

- This can be seen as follows

$$\boldsymbol{\Sigma}^{\dagger}\boldsymbol{\Sigma} = \begin{bmatrix} \frac{1}{\sigma_1} & & \mathbf{0} \\ & \ddots & \vdots \\ & \frac{1}{\sigma_n} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} = \mathbf{I}$$

Pseudoinverse – property

- Similarly if $n \geq m = r$

$$\Sigma\Sigma^{\dagger} = \mathbf{I}$$

Pseudoinverse

- The **pseudoinverse** of the $m \times n$ matrix $\mathbf{A}$ with SVD $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is given as

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

Pseudoinverse – tall matrix

- Consider a **tall matrix** $\mathbf{A}$ of size $m \times n$ with $m > n = \text{rank}(\mathbf{A})$

$$\begin{aligned}\mathbf{A}^\dagger\mathbf{A} &= \mathbf{V}\boldsymbol{\Sigma}^\dagger\mathbf{U}^T\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \\ &= \mathbf{V}\boldsymbol{\Sigma}^\dagger\boldsymbol{\Sigma}\mathbf{V}^T = \mathbf{V}\mathbf{I}\mathbf{V}^T \\ &= \mathbf{V}\mathbf{V}^T = \mathbf{I}\end{aligned}$$

  - Thus, $\mathbf{A}^\dagger$ is a **left-inverse** of $\mathbf{A}$.

Pseudoinverse – tall matrix

- Furthermore, for this case of a tall matrix with full column rank n, it can be seen that

$$\mathbf{A}^{\dagger} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$$

- This can be verified as follows

$$\begin{aligned}\mathbf{A}^T\mathbf{A} &= \mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \\ &= \mathbf{V}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\mathbf{V}^T = \mathbf{V}\tilde{\boldsymbol{\Sigma}}^2\mathbf{V}^T\end{aligned}$$

Pseudoinverse – tall matrix

- Observe that $\mathbf{\Sigma}^2 = \mathbf{\Sigma}^T\mathbf{\Sigma}$ for this scenario is an $n \times n$ diagonal matrix with elements, σ_i^2, $i = 1,2,\ldots,n$
- Simplifying further, and noting $\mathbf{\Sigma}^{-2}\mathbf{\Sigma}^T = \mathbf{\Sigma}^\dagger$, it follows that

$$\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \mathbf{V}\tilde{\mathbf{\Sigma}}^{-2}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T$$

$$= \mathbf{V}\tilde{\mathbf{\Sigma}}^{-2}\mathbf{\Sigma}^T\mathbf{U}^T = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^T = \mathbf{A}^\dagger$$

Pseudoinverse – wide matrix

- On the other hand, for a **wide matrix** $\mathbf{A}$ of size $m \times n$ with $n > m = \text{rank}(\mathbf{A})$

$$\begin{aligned}\mathbf{A}\mathbf{A}^\dagger &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^T \\ &= \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^\dagger\mathbf{U}^T = \mathbf{U}\mathbf{I}\mathbf{U}^T \\ &= \mathbf{U}\mathbf{U}^T = \mathbf{I}\end{aligned}$$

  - Thus, $\mathbf{A}^\dagger$ is a **right-inverse** of $\mathbf{A}$.

Pseudoinverse – wide matrix

- For this case of a wide matrix with full row rank m, it can be seen that

$$\mathbf{A}^{\dagger} = \mathbf{A}^{T}\left(\mathbf{A}\mathbf{A}^{T}\right)^{-1}$$

- This can be verified as follows

$$\mathbf{A}^{T}\left(\mathbf{A}\mathbf{A}^{T}\right)^{-1} = \mathbf{V}\boldsymbol{\Sigma}^{T}\mathbf{U}^{T}\left(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{T}\mathbf{V}\boldsymbol{\Sigma}^{T}\mathbf{U}^{T}\right)^{-1}$$

$$= \mathbf{V}\boldsymbol{\Sigma}^{T}\mathbf{U}^{T}\left(\mathbf{U}\tilde{\boldsymbol{\Sigma}}^{2}\mathbf{U}^{T}\right)^{-1} = \mathbf{V}\boldsymbol{\Sigma}^{T}\mathbf{U}^{T}\mathbf{U}\tilde{\boldsymbol{\Sigma}}^{-2}\mathbf{U}^{T}$$

$$= \mathbf{V}\boldsymbol{\Sigma}^{T}\tilde{\boldsymbol{\Sigma}}^{-2}\mathbf{U}^{T} = \mathbf{V}\boldsymbol{\Sigma}^{\dagger}\mathbf{U}^{T} = \mathbf{A}^{\dagger}$$

02

Incomplete eigenspace

A repeated eigenvalue with too few eigenvectors



Incomplete eigenspace

- Consider now the matrix below

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

Incomplete eigenspace

- Its characteristic equation, expanding along the first row, is

$$|\mathbf{A}-\lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & -1 & 3 \\ 2 & -\lambda & 2 \\ 2 & -1 & 2-\lambda \end{vmatrix}$$

$$= (1-\lambda)\left((-\lambda)(2-\lambda)+2\right)+(4-2\lambda-4)+3(-2+2\lambda)$$

$$= (1-\lambda)\left(\lambda^2-2\lambda+2\right)+4\lambda-6$$

$$= -\lambda^3+3\lambda^2-4=0$$

Incomplete eigenspace

- This can be factored as

$$-\lambda^3 + 3\lambda^2 - 4 = 0$$
$$\Rightarrow -(\lambda - 2)^2 (\lambda + 1) = 0$$

- It has a **double eigenvalue** at 2 and a simple eigenvalue at -1

Incomplete eigenspace

- Consider

$$\mathbf{A} - \lambda_1 \mathbf{I} = \mathbf{A} - 2\mathbf{I}$$
$$= \begin{bmatrix} -1 & -1 & 3 \\ 2 & -2 & 2 \\ 2 & -1 & 0 \end{bmatrix}$$

Incomplete eigenspace

- To find the eigenvector, set

$$(\mathbf{A} - \lambda_1 \mathbf{I})\,\bar{\mathbf{v}} = (\mathbf{A} - 2\mathbf{I})\,\bar{\mathbf{v}} = \mathbf{0}$$

$$\Rightarrow \begin{bmatrix} -1 & -1 & 3 \\ 2 & -2 & 2 \\ 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \mathbf{0}$$

Incomplete eigenspace

- Which can be simplified as

$$2v_1 - v_2 = 0 \Rightarrow v_2 = 2v_1$$
$$-v_1 - v_2 + 3v_3 = 0$$
$$\Rightarrow 3v_3 = 3v_1 \Rightarrow v_3 = v_1$$

Incomplete eigenspace

- Therefore, the **eigenvector** is

$$\begin{bmatrix} v_1 \\ 2v_1 \\ v_1 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

- Thus corresponding to $\lambda_1 = 2$ it only has a **one-dimensional eigenspace**!

Incomplete eigenspace

- Therefore, the **eigenspace is incomplete**
  - as the eigenvectors do not span $\mathbb{R}^3$

03

Jordan normal form

Jordan blocks and the Jordan canonical form



Jordan block matrix

- Consider the **bidiagonal Jordan block**

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

Jordan block matrix

- Observe that $\mathbf{J}_\lambda - \mu\mathbf{I}$ is

$$\mathbf{J}_\lambda - \mu\mathbf{I} = \begin{bmatrix} \lambda-\mu & 1 & & \\ & \lambda-\mu & \ddots & \\ & & \ddots & 1 \\ & & & \lambda-\mu \end{bmatrix}$$

Jordan block matrix

- To determine the eigenvalues, we set the characteristic equation to 0

$$|\mathbf{J}_\lambda - \mu\mathbf{I}| = (\lambda - \mu)^n = 0$$
$$\Rightarrow \mu = \lambda$$

Jordan block matrix

- For $\mu = \lambda$, the equation $(\mathbf{J}_\lambda - \lambda\mathbf{I})\bar{\mathbf{v}} = \mathbf{0}$ gives

$$v_2 = v_3 = \ldots = v_n = 0$$

Jordan block matrix

- Therefore, the eigenvector is given as

$$\bar{\mathbf{v}} = k \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

  - Thus it has only a **single eigenvector**!

Jordan normal form

- A matrix $\mathbf{A}$ is diagonalizable if and only if it has n **linearly independent eigenvectors** that span $\mathbb{C}^n$
- However, any matrix $\mathbf{A}$ can be reduced to a **Jordan normal form**, i.e.

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$$
$$\Rightarrow \mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$$

Jordan normal form

- The matrix $\mathbf{J}$ is a **block diagonal matrix**

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & \\ & \ddots & \\ & & \mathbf{J}_p \end{bmatrix}$$

Jordan normal form

- The matrices $\mathbf{J}_i$ are **Jordan blocks** of the form

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Jordan form – example

- Consider the matrix below

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

- We determine the **Jordan normal form** for this matrix

Jordan form – example

- The **characteristic polynomial** of the matrix is

$$\lambda^4 - 8\lambda^3 + 23\lambda^2 - 28\lambda + 12$$
$$= (\lambda - 1)(\lambda - 3)(\lambda - 2)^2 = 0$$
$$\Rightarrow \lambda = 1, 3, 2$$

Jordan form – example

- Eigenvectors corresponding to eigenvalues $1, 3$ are

$$\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Jordan form – example

- The eigenvector corresponding to eigenvalue 2, which has multiplicity 2, is

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

- Therefore, the **eigenspace is incomplete**

Jordan form – example

- For eigenvalue 2, we compute the kernel or nullspace of

$$\mathcal{N}\left((\mathbf{A}-2\mathbf{I})^2\right) = \text{span}\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}\right\}$$

- These are termed **generalized eigenvectors**.

Jordan form – example

- The **Jordan canonical form** can be obtained as

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Jordan form – example

- The matrix $\mathbf{P}$ can be determined as

$$\mathbf{P} = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

- Its last column $\bar{\mathbf{p}}_4$ satisfies $(\mathbf{A} - 2\mathbf{I})\bar{\mathbf{p}}_4 = \bar{\mathbf{p}}_3$

Summary

- The **pseudoinverse** $\boldsymbol{\Sigma}^{\dagger}$ inverts the non-zero singular values; $\mathbf{A}^{\dagger} = \mathbf{V}\boldsymbol{\Sigma}^{\dagger}\mathbf{U}^{T}$
- Tall, full column rank: $\mathbf{A}^{\dagger} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T}$ is a left inverse
- Wide, full row rank: $\mathbf{A}^{\dagger} = \mathbf{A}^{T}(\mathbf{A}\mathbf{A}^{T})^{-1}$ is a right inverse
- A repeated eigenvalue can have too few eigenvectors: the eigenspace is **incomplete**
- Every matrix has a **Jordan normal form** $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$ built from Jordan blocks



Thank You!

Questions?