Outline

The Kronecker Product

- Let $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{P \times Q}$, where Namely,

$n = 1 \sim N$

$$\mathbf{A} \triangleq \begin{array}{c} m=1 \\ \\ \\ M \end{array} \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \dots & a_{M,N} \end{bmatrix}. \tag{23}$$

- The Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is defined as

$$\mathbf{A} \otimes \mathbf{B} \triangleq \begin{bmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \dots & a_{1,N}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \dots & a_{2,N}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1}\mathbf{B} & a_{M,2}\mathbf{B} & \dots & a_{M,N}\mathbf{B} \end{bmatrix} \in \mathbb{C}^{(MP) \times (NQ)}. \tag{24}$$

- Reference [GVL2013]

An Example of the Kronecker Product

- We consider the matrices $\mathbf{A}$ and $\mathbf{B}$

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2\times 3}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}_{2\times 2}. \tag{25}$$

- Then the matrix $\mathbf{A} \otimes \mathbf{B}$ becomes

scalar matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} (1)\mathbf{B} & (2)\mathbf{B} & (3)\mathbf{B} \\ (4)\mathbf{B} & \boxed{(5)\mathbf{B}} & (6)\mathbf{B} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 2 & 0 & 3 \\ -1 & 0 & -2 & 0 & -3 & 0 \\ 0 & 4 & 0 & 5 & 0 & 6 \\ -4 & 0 & -5 & 0 & -6 & 0 \end{bmatrix}_{4\times 6}, \tag{26}$$

$*$

$$\underline{\underline{B}} \otimes \underline{\underline{A}} = \begin{bmatrix} 0\,(\underline{\underline{A}}) & (1)(\underline{\underline{A}}) \\ (-1)(\underline{\underline{A}}) & (0)(\underline{\underline{A}}) \end{bmatrix}_{4\times 6} \checkmark$$

Some Properties [Seber2008, pp. 234]

- If $\alpha \in \mathbb{C}$ and $\mathbf{A} \in \mathbb{C}^{M\times N}$, then

1×1

$$\alpha \otimes \mathbf{A} = \boxed{\alpha \mathbf{A}} = \mathbf{A} \otimes \alpha. \qquad (27)$$

$$\underline{A} = \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}$$

$$[\alpha] \otimes \underline{A} = [\alpha \underline{A}] = \begin{bmatrix} \alpha & 3\alpha \\ 5\alpha & 7\alpha \end{bmatrix}$$

$$= \begin{bmatrix} (1)\alpha & (3)\alpha \\ (5)\alpha & (7)\alpha \end{bmatrix}$$

- If $\mathbf{u} \in \mathbb{C}^N$ and $\mathbf{v} \in \mathbb{C}^M$, then

N M

$$\mathbf{u}^\mathsf{T} \otimes \mathbf{v} = \boxed{\mathbf{v}\mathbf{u}^\mathsf{T}} = \mathbf{v} \otimes \mathbf{u}^\mathsf{T}. \qquad (28)$$

$1\times N$ $M\times 1$ u^T

tensor

- If $\mathbf{A} \in \mathbb{C}^{M\times N}$ and $\mathbf{B} \in \mathbb{C}^{P\times Q}$, then

transpose

$$(\mathbf{A} \otimes \mathbf{B})^\mathsf{T} = (\mathbf{A}^\mathsf{T}) \otimes (\mathbf{B}^\mathsf{T}). \qquad (29)$$

transpose

$$\underline{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \qquad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

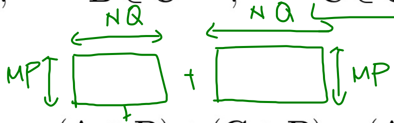
$$\underline{u}^T \otimes \underline{v} = \underset{1\times 2}{\begin{bmatrix} u_1 & u_2 \end{bmatrix}} \otimes \underset{3\times 1}{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}} = \underset{3\times 2}{\begin{bmatrix} u_1 v_1 & u_2 v_1 \\ u_1 v_2 & u_2 v_2 \\ u_1 v_3 & u_2 v_3 \end{bmatrix}}$$

The Addition Property

- We consider

$$\mathbf{A} \in \mathbb{C}^{M \times N}, \qquad \mathbf{B} \in \mathbb{C}^{P \times Q}, \qquad \mathbf{C} \in \mathbb{C}^{M \times N}, \qquad \mathbf{D} \in \mathbb{C}^{P \times Q}. \tag{30}$$

- Then



$$(\mathbf{A} \otimes \mathbf{B}) + (\mathbf{C} \otimes \mathbf{B}) = (\mathbf{A} + \mathbf{C}) \otimes \mathbf{B}, \tag{31}$$

$$(\mathbf{A} \otimes \mathbf{B}) + (\mathbf{A} \otimes \mathbf{D}) = \mathbf{A} \otimes (\mathbf{B} + \mathbf{D}). \tag{32}$$

The Product Property

- We consider

$$\mathbf{A} \in \mathbb{C}^{M_1 \times N_1}, \qquad \mathbf{B} \in \mathbb{C}^{M_2 \times N_2}, \qquad \mathbf{C} \in \mathbb{C}^{M_3 \times N_3}, \qquad \mathbf{D} \in \mathbb{C}^{M_4 \times N_4}. \tag{33}$$

- We assume that $N_1 = M_3$ and $N_2 = M_4$.
- Then

M3

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}). \tag{34}$$

N1 N2 M4

The Inverse Property

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

- Let $\mathbf{A} \in \mathbb{C}^{M \times M}$ and $\mathbf{B} \in \mathbb{C}^{N \times N}$.
- Assume that $\mathbf{A}$ and $\mathbf{B}$ are invertible.
- Then

A^{-1} & B^{-1} exist

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = (\mathbf{A}^{-1}) \otimes (\mathbf{B}^{-1}) . \tag{35}$$

Proof: sketch: $(M M^{-1} \stackrel{?}{=} I)$

$$\underbrace{(A \otimes B)}\underbrace{\left[(A^{-1}) \otimes (B^{-1}) \right]} \stackrel{(34)}{=} (A A^{-1}) \otimes (B B^{-1}) = I_M \otimes I_N = \begin{bmatrix} I_N & & 0 \\ & I_N & \\ 0 & & \ddots & I_N \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

The Eigenvector Property

Eigenvalues / eigenvectors for $(\underline{\underline{A}} \otimes \underline{\underline{B}})$?

- Let $\mathbf{A} \in \mathbb{C}^{M \times M}$ and $\mathbf{B} \in \mathbb{C}^{N \times N}$.
- We consider the eigenvalues and eigenvectors as follows:

$$\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \qquad \mathbf{B}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2. \tag{36}$$

MN × MN

- Then

$$\boxed{(\mathbf{A} \otimes \mathbf{B})}\boxed{(\mathbf{v}_1 \otimes \mathbf{v}_2)} = \underbrace{\boxed{(\lambda_1 \lambda_2)}}_{\text{eigenvalue}}\underbrace{\boxed{(\mathbf{v}_1 \otimes \mathbf{v}_2)}}_{\text{eigenvector}}. \tag{37}$$

$= (\underline{A}\,\underline{v}_1) \otimes (\underline{B}\,\underline{v}_2) = (\lambda_1 \underline{v}_1) \otimes (\lambda_2 \underline{v}_2)$

- Interpretations:
  - $(\mathbf{v}_1 \otimes \mathbf{v}_2)$ is an eigenvector of $(\mathbf{A} \otimes \mathbf{B})$.
  - $(\lambda_1 \lambda_2)$ is the corresponding eigenvalue of $(\mathbf{A} \otimes \mathbf{B})$.

The Orthogonal Property

$A = [1 \quad 3+j]$

- Let $\mathbf{A} \in \mathbb{C}^{M\times M}$ and $\mathbf{B} \in \mathbb{C}^{N\times N}$.
- We assume that $\mathbf{A}$ and $\mathbf{B}$ are unitary matrices, i.e.,

$$\mathbf{A}\mathbf{A}^{\mathsf{H}} = \mathbf{I}, \qquad \mathbf{B}\mathbf{B}^{\mathsf{H}} = \mathbf{I}. \tag{38}$$

- Let $\boxed{\mathbf{C} \triangleq \mathbf{A} \otimes \mathbf{B}.}$
- Then $\mathbf{C}$ is also an unitary matrix,

$$\mathbf{C}\mathbf{C}^{\mathsf{H}} = \mathbf{I}. \tag{39}$$

$(\underline{\underline{A}} \otimes \underline{\underline{B}})(\underline{\underline{A}} \otimes \underline{\underline{B}})^H$

$(\underline{\underline{A}} \otimes \underline{\underline{B}})^*$

$= \begin{bmatrix} a_{11}B & a_{12}B & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}^*$

$= \begin{bmatrix} a_{11}^* B^* & a_{12}^* \underline{B}^* & \vdots \end{bmatrix}$

$= (A^*) \otimes (\underline{\underline{B}}^*)$

The Rank Property

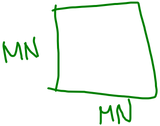
- Let $\mathbf{A} \in \mathbb{C}^{M\times N}$ and $\mathbf{B} \in \mathbb{C}^{P\times Q}$.
- Then

$$\operatorname{rank}(\mathbf{A} \otimes \mathbf{B}) = \operatorname{rank}(\mathbf{A}) \operatorname{rank}(\mathbf{B}). \tag{40}$$

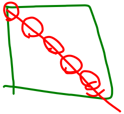
The Determinant Property

- Let $\mathbf{A} \in \mathbb{C}^{M \times M}$ and $\mathbf{B} \in \mathbb{C}^{N \times N}$.
- Then

$$\det(\mathbf{A} \otimes \mathbf{B}) = (\det(\mathbf{A}))^{N} (\det(\mathbf{B}))^{M}. \tag{41}$$

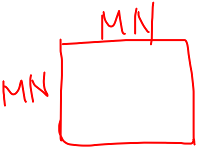


The Trace Property



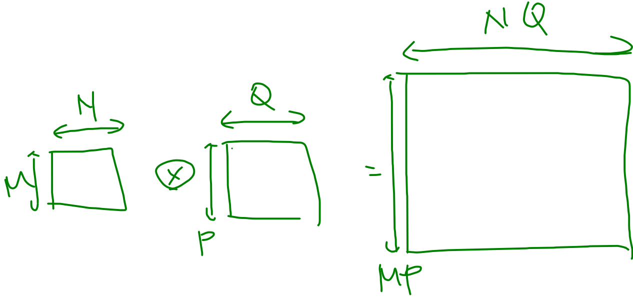
- Let $\mathbf{A} \in \mathbb{C}^{M\times M}$ and $\mathbf{B} \in \mathbb{C}^{N\times N}$.
- Then

$$\operatorname{tr}(\mathbf{A}\otimes\mathbf{B}) = \operatorname{tr}(\mathbf{A})\operatorname{tr}(\mathbf{B}). \tag{42}$$



Outline

1. Review of Linear Algebra
   - Matrix Operations
   - Eigenvalues and Eigenvectors
2. The Kronecker Product
3. The Hadamard (Element-Wise) Product
4. The Vectorization Operator
5. Generalized Norms
   - Vector Norms
   - The Entry-Wise Matrix Norms



The Hadamard (Element-Wise) Product

- $\mathbf{A} \in \mathbb{C}^{M \times N}$ (with entries $a_{m,n}$)
- $\mathbf{B} \in \mathbb{C}^{M \times N}$ (with entries $b_{m,n}$)
- The Hadamard product: $\mathbf{C} = \mathbf{A} \circ \mathbf{B} \in \mathbb{C}^{M \times N}$.

"$*$"

$$\underbrace{\begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,N} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M,1} & c_{M,2} & \cdots & c_{M,N} \end{bmatrix}}_{\mathbf{C}} = \underbrace{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \cdots & a_{M,N} \end{bmatrix}}_{\mathbf{A}} \circ \underbrace{\begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{M,1} & b_{M,2} & \cdots & b_{M,N} \end{bmatrix}}_{\mathbf{B}},$$

$$\boxed{c_{m,n} = a_{m,n} \times b_{m,n}.}$$

Examples

- Let the matrices $\mathbf{A}$ and $\mathbf{B}$ be

$$\mathbf{A} = \begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}. \tag{43}$$

- Then the Hadamard product $\mathbf{A} \circ \mathbf{B}$ is

$$\mathbf{A} \circ \mathbf{B} = \begin{bmatrix} (5) \times (1) & (4) \times (-1) \\ (3) \times (2) & (2) \times (0) \end{bmatrix} = \begin{bmatrix} 5 & -4 \\ 6 & 0 \end{bmatrix}. \tag{44}$$

Properties (1/2)

- Let the matrices $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{B} \in \mathbb{C}^{M\times N}$, and $\mathbf{C} \in \mathbb{C}^{M\times N}$.
- The commutative property:

$$\underbrace{\mathbf{A}\circ\mathbf{B}}_{a_{m,n}\times b_{m,n}} = \underbrace{\mathbf{B}\circ\mathbf{A}}_{b_{m,n}\times a_{m,n}}. \tag{45}$$

- The associative property:

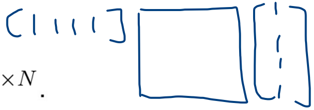
$$\underbrace{(\mathbf{A}\circ\mathbf{B})}_{(a_{m,n} b_{m,n})\, c_{m,n}}\circ\mathbf{C} = \mathbf{A}\circ\underbrace{(\mathbf{B}\circ\mathbf{C})}_{a_{n,n}(b_{m,n}\, c_{m,n})}. \tag{46}$$

- The distributive property

$$\mathbf{A}\circ(\mathbf{B}+\mathbf{C}) = (\mathbf{A}\circ\mathbf{B}) + (\mathbf{A}\circ\mathbf{C}), \tag{47}$$

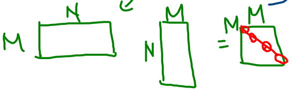
$$(\mathbf{A}+\mathbf{B})\circ\mathbf{C} = (\mathbf{A}\circ\mathbf{C}) + (\mathbf{B}\circ\mathbf{C}). \tag{48}$$

Properties (2/2)



- Let the matrices $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{B} \in \mathbb{C}^{M\times N}$, and $\mathbf{C} \in \mathbb{C}^{M\times N}$.
- Let $\mathbf{1}_M \triangleq \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^\mathsf{T}$ be a length-M vector of all ones.
- The trace function [Seber2008, pp. 252]:

$$\operatorname{tr}\left(\mathbf{A}\boxed{\mathbf{B}^\mathsf{T}}\right) = \mathbf{1}_M^\mathsf{T}(\mathbf{A} \circ \mathbf{B})\mathbf{1}_N, \tag{49}$$



$$\operatorname{tr}\left(\underbrace{(\mathbf{A} \circ \mathbf{B})\,\mathbf{C}^\mathsf{T}}\right) \overset{?}{=} \operatorname{tr}\left((\mathbf{A} \circ \mathbf{C})\,\mathbf{B}^\mathsf{T}\right). \tag{50}$$

$$\mathrm{tr}(\underline{\underline{A}}\,\underline{\underline{B}}^T)$$

$$\underline{\underline{A}} = \begin{bmatrix} \boxed{\underline{a}_1} & \boxed{\underline{a}_2} & \boxed{\underline{a}_N} \end{bmatrix}_{M\times N}$$

$M\times 1$ $M\times 1$ $M\times 1$

$$= \mathrm{tr}(\underline{\underline{B}}^T \underline{\underline{A}})$$

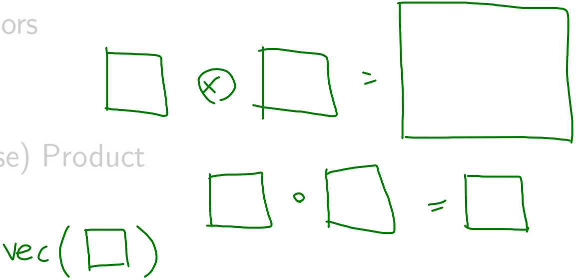
$$\underline{\underline{B}} = \begin{bmatrix} \boxed{\underline{b}_1} & \underline{b}_2 & \underline{b}_N \end{bmatrix}_{M\times N}$$

$$= \mathrm{tr}\left(\begin{bmatrix} \underline{b}_1^T \\ \underline{b}_2^T \\ \vdots \\ \underline{b}_N^T \end{bmatrix} \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_N \end{bmatrix} \right)$$

$$[\underline{a}_1]_1 [\underline{b}_1]_1 + [\underline{a}_1]_2 [\underline{b}_1]_2$$

$$= \mathrm{tr}\begin{bmatrix} \underline{b}_1^T \underline{a}_1 & \underline{b}_1^T \underline{a}_2 & & \\ & \underline{b}_2^T \underline{a}_2 & & \\ & & \ddots & \\ & & & \end{bmatrix} = \sum_{i=1}^{N} \boxed{\underline{b}_i^T \underline{a}_i}$$

Outline



Definition of the Vectorization Operator

- Let the matrix $\mathbf{A}$ be

$$\mathbf{A} = \begin{array}{cc} & \begin{array}{cccc} \mathbf{a}_1 & \mathbf{a}_2 & & \end{array} \\ & \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M,1} & a_{M,2} & \dots & a_{M,N} \end{bmatrix} \end{array} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_N \end{bmatrix}, \qquad (51)$$

$M \times N$

where $\mathbf{a}_n = \begin{bmatrix} a_{1,n} & a_{2,n} & \dots & a_{M,n} \end{bmatrix}^\mathsf{T} \in \mathbb{C}^M$.

- Then the vectorization of $\mathbf{A}$ is

$$\operatorname{vec}(\mathbf{A}) \triangleq \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_N \end{bmatrix}. \qquad (52)$$

M

M

MN

An Example of the Vectorization

$u^T \otimes v$ v

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

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}. \tag{53}$$

- The vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \qquad \mathbf{a}_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}. \tag{54}$$

- Then

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \\ 5 \\ 3 \\ 6 \end{bmatrix}. \tag{55}$$

The Kronecker Product [Seber2008, pp. 240]

- Let the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{B}$ satisfy

$$\mathbf{A} \in \mathbb{C}^{M\times N}, \qquad \mathbf{B} \in \mathbb{C}^{N\times P}, \qquad \mathbf{C} \in \mathbb{C}^{P\times Q}, \tag{56}$$

Two Matrices

$$\begin{aligned} &\operatorname{vec}(\mathbf{AB}) \\ &= (\mathbf{I}_P \otimes \mathbf{A}) \operatorname{vec}(\mathbf{B}) \\ &= (\mathbf{B}^\mathsf{T} \otimes \mathbf{A}) \operatorname{vec}(\mathbf{I}_N) \\ &= (\mathbf{B}^\mathsf{T} \otimes \mathbf{I}_M) \operatorname{vec}(\mathbf{A}). \end{aligned}$$

Three Matrices

$$\begin{aligned} &\operatorname{vec}(\mathbf{ABC}) \\ &= (\mathbf{C}^\mathsf{T} \otimes \mathbf{A}) \operatorname{vec}(\mathbf{B}) \\ &= (\mathbf{I}_Q \otimes (\mathbf{AB})) \operatorname{vec}(\mathbf{C}) \\ &= \left((\mathbf{BC})^\mathsf{T} \otimes \mathbf{I}_M\right) \operatorname{vec}(\mathbf{A}). \end{aligned}$$

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} \qquad \underline{\underline{B}} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$$

$$\underline{\underline{A}}\,\underline{\underline{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} = \begin{bmatrix} \alpha + 2\beta & \gamma + 2\delta \\ 3\alpha + 4\beta & 3\gamma + 4\delta \end{bmatrix}$$

$$\mathrm{vec}(\underline{\underline{A}}\,\underline{\underline{B}}) = \begin{bmatrix} \alpha + 2\beta \\ 3\alpha + 4\beta \\ \gamma + 2\delta \\ 3\gamma + 4\delta \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$$

$\mathrm{vec}(\underline{\underline{B}})$

$$\begin{bmatrix} \underline{\underline{A}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{A}} \end{bmatrix} = \underline{\underline{I}}_2 \otimes \underline{\underline{A}}$$

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

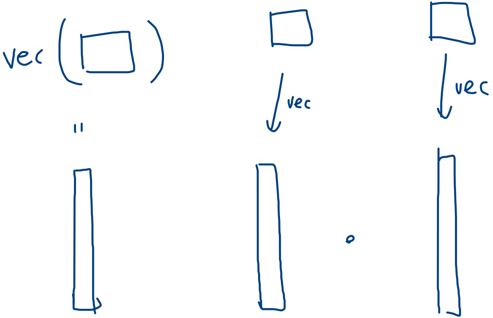
$$\operatorname{vec}(\underline{\underline{A}}\,\underline{\underline{B}}) = \begin{bmatrix} \alpha + 2\beta \\ 3\alpha + 4\beta \\ \gamma + 2\delta \\ 3\gamma + 4\delta \end{bmatrix} = \begin{bmatrix} \alpha & 0 & \beta & 0 \\ 0 & \alpha & 0 & \beta \\ & & & \\ & & & \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}}_{\operatorname{vec}(\underline{A})}$$

$$= \left(\quad \otimes \quad \right) \operatorname{vec}(\underline{\underline{A}})$$

The Hadamard Product

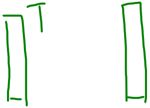
- Let $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{M \times N}$.
- Then

$$\operatorname{vec}(\mathbf{A} \circ \mathbf{B}) = (\operatorname{vec}(\mathbf{A})) \circ (\operatorname{vec}(\mathbf{B})) . \tag{57}$$

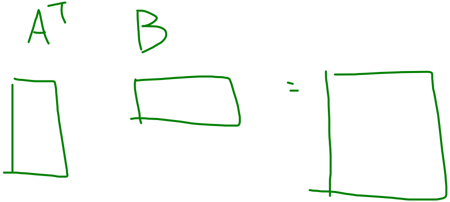


The Trace Function [Seber2008, pp. 240]

- Let $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{B} \in \mathbb{C}^{M \times \cancel{P}}$. H
- Then



$$\operatorname{tr}\left(\mathbf{A}^{\mathsf{T}}\mathbf{B}\right) = \left(\operatorname{vec}\left(\mathbf{A}\right)\right)^{\mathsf{T}} \operatorname{vec}\left(\mathbf{B}\right). \tag{58}$$



Outline

Norm : $\| \underline{v} \|$

$\underline{v} \in \mathbb{C}^N$, $\|\underline{v}\|$: norm.

Outline

The Vector p-Norm

Definition (The vector p-norm, the ℓ_p norm)

For a vector $\mathbf{x} \triangleq \begin{bmatrix} x_1 & x_2 & \dots & x_N \end{bmatrix}^\mathsf{T} \in \mathbb{C}^N$ and $p \geq 1$, the vector p-norm is defined as

$$\|\mathbf{x}\|_p \triangleq \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p}. \tag{59}$$

$\frac{1}{p}$

- $p = 2$: The Euclidean norm or the ℓ_2 norm
- $p = 1$: The ℓ_1 norm
- $p \to \infty$: The ℓ_∞ norm

The ℓ_2 Norm ($p = 2$)

$$\|\mathbf{x}\|_2 \triangleq \left(\sum_{i=1}^{N} |x_i|^2\right)^{1/2} \tag{60}$$

$$= \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_N|^2} \tag{61}$$

$$= \sqrt{\boxed{\mathbf{x}^{\mathsf{H}}\mathbf{x}}} = \sqrt{\mathrm{tr}\left(\mathbf{x}^{\mathsf{H}}\mathbf{x}\right)} = \sqrt{\mathrm{tr}\left(\mathbf{x}\mathbf{x}^{\mathsf{H}}\right)}. \tag{62}$$

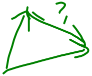
$\in \mathbb{R}$

$$\underline{x}^H \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}^H \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} x_1^* & x_2^* & \cdots & x_N^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

$$= |x_1|^2 + |x_2|^2 + \cdots + |x_N|^2$$

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$

- The Euclidean norm
- The ℓ_2 norm $\|\mathbf{u} - \mathbf{v}\|_2$ measures the distance between two vectors $\mathbf{u}$ and $\mathbf{v}$.
- Example: If $\mathbf{x} = \begin{bmatrix} 3 & -4 & 2 \end{bmatrix}^{\mathsf{T}}$, then



$$\|\mathbf{x}\|_2 = \sqrt{|3|^2 + |-4|^2 + |2|^2} = \sqrt{29}.$$

The ℓ_1 Norm $(p = 1)$

ℓ_p

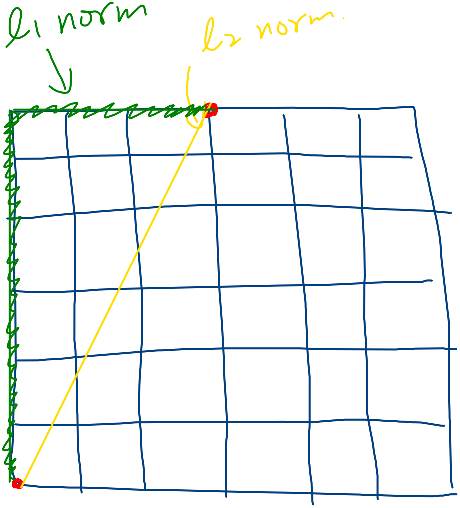
$p=1$

$$\|\mathbf{x}\|_1 \triangleq \left(\sum_{i=1}^{N} |x_i|^{1}\right)^{1/1} \tag{63}$$

$$= \underbrace{|x_1| + |x_2| + \cdots + |x_N|}. \tag{64}$$

- The sum of amplitudes
- Example: If $\mathbf{x} = \begin{bmatrix} 3 & -4 & 2 \end{bmatrix}^{\mathsf{T}}$, then

$$\|\mathbf{x}\|_1 = \underbrace{|3|}_{3} + \underbrace{|-4|}_{4} + \underbrace{|2|}_{2} = 9.$$


l1 norm
l2 norm.

The ℓ_∞ Norm ($p \to \infty$)

$$\|x\|_\infty \overset{(65)}{=} \lim_{p\to\infty} \left(3^p + 4^p + 2^p\right)^{1/p} = \lim_{p\to\infty} \left[4^p\left(\left(\frac{3}{4}\right)^p + 1 + \left(\frac{2}{4}\right)^p\right)\right]^{1/p}$$

$$\|\mathbf{x}\|_\infty \triangleq \lim_{p\to\infty} \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} \tag{65}$$

$$= \max_{i\in[N]} |x_i|. \tag{66}$$

$$= \lim_{p\to\infty} 4 \times \left[\boxed{\left(\frac{3}{4}\right)^p} + 1 + \boxed{\left(\frac{2}{4}\right)^p}\right]^{\frac{1}{p}}$$

0 if $p\to\infty$; 0 if $p\to\infty$

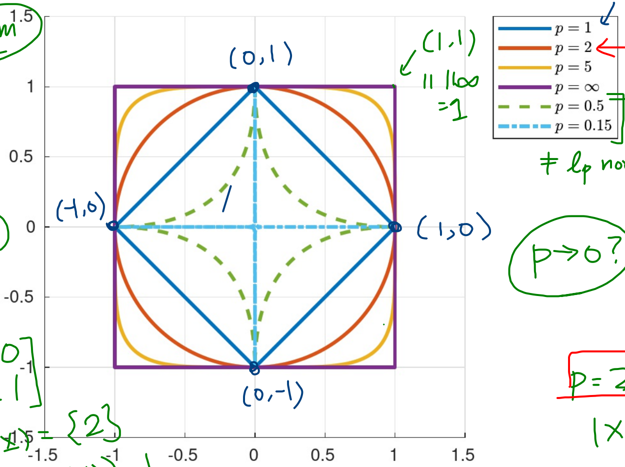
- $[N] \triangleq \{1, 2, \ldots, N\}$.
- $\|\mathbf{x}\|_\infty$ represents the maximal amplitudes of $\mathbf{x}$.
- Example: If $\mathbf{x} = \begin{bmatrix} 3 & -4 & 2 \end{bmatrix}^\mathsf{T}$, then

$$\|\mathbf{x}\|_\infty = \max\{|3|, |-4|, |2|\} = 4.$$

Examples: Contour Plots of $\|\mathbf{x}\|_p = 1$ (The ℓ_p Ball)

$p=2$
$p=1$
$p\to\infty$

norm



$\neq \ell_p$ norm

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2.$$

$p=1$

$$|x_1|^p + |x_2|^p = 1^p = 1.$$

$p \to 0$?

$\|\underline{x}\|_p^p = 1^p$

$p=1$: $|x_1| + |x_2| = 1$

$p=2$: $|x_1|^2 + |x_2|^2 = 1 \Rightarrow x_1^2 + x_2^2 = 1$

circle radius $=1$

$\underline{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\mathrm{supp}(\underline{x}) = \{2\}$

$\mathrm{card}(\mathrm{supp}(\underline{x})) = 1$

The $\boxed{\ell_0 \text{ Function}}$ (1/2)

$\boxed{\| \ \|_p \quad p \to 0?}$

- Notations
  - $[N] \triangleq \{1, 2, \ldots, N\}$.
  - $\text{card}(\mathbb{S})$: The cardinality of a set $\mathbb{S}$.
  - $\mathbf{x} \triangleq \begin{bmatrix} x_1 & x_2 & \ldots & x_N \end{bmatrix}^\mathsf{T}$.

$S = \{0, 1, 2\}$

$\text{card}(S) = 3$

Definition (Support)

The $\boxed{\text{support}}$ of a vector $\mathbf{x} \in \mathbb{C}^N$ is defined as $\text{supp}(\mathbf{x}) \triangleq \{i \in [N] : x_i \neq 0\}$

The support of a vector

If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$, then $\text{supp}(\mathbf{x}) = \{1, 2, 4\}$.

$2 \neq 0$, $3 \neq 0$, $0 = 0$, $1 \neq 0$

$\text{card}(\text{supp}(\underline{x}))$

$= 3 \leftarrow \ell_0$ (sparsity)

The ℓ_0 Function (2/2)

$p \to 0$

Definition (The ℓ_0 function)

For a vector $\mathbf{x} \in \mathbb{C}^N$, the ℓ_0 function is defined as

$$\|\mathbf{x}\|_0 \triangleq \mathrm{card}(\mathrm{supp}(\mathbf{x})). \tag{67}$$

- Precisely, $\|\mathbf{x}\|_0$ is not a norm.
- $\|\mathbf{x}\|_0$ is related to the following limit

$$\lim_{p\to 0} \sum_{i=1}^{N} |x_i|^p. \tag{68}$$

Remarks

- The $\boxed{\ell_p \text{ norm}}$ and the ℓ_0 function are often used in the objective function of optimization problems.
- The ℓ_p norm with $p \geq 1$ is a convex function.
- The ℓ_0 function promotes sparse solutions.
- The ℓ_0 function is non-convex.
- Convex relexation in compressive sensing
  - Replacing the ℓ_0 function with the ℓ_1 norm
  - Basis pursuit

Outline

ℓ_1

ℓ_2 $\to$ ℓ_∞

$\to$ (ℓ_0)

The Entry-Wise L_p Norm

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & & \ddots & \\ \vdots & & & \ddots \\ a_{M,1} & & & \end{bmatrix}_{M \times N}$$

- For $p \geq 1$, the entry-wise L_p norm is defined as

$$\|\mathbf{A}\|_p \triangleq \Bigg(\underbrace{\sum_{m=1}^{M}\sum_{n=1}^{N}}_{\text{sum}} |a_{m,n}|^p\Bigg)^{\frac{1}{p}}. \tag{69}$$

$a_{m,n}$

vec

- It can be shown that

$$\|\mathbf{A}\|_p = \|\mathrm{vec}(\mathbf{A})\|_p \tag{70}$$

$$\mathrm{vec}(\mathbf{A}) = \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{M,1} \\ a_{1,2} \\ a_{2,2} \\ \vdots \end{bmatrix}$$

where the latter $\|\cdot\|_p$ denotes the vector p-norm.

The Frobenius Norm (The L_2 Norm)

Definition

For a matrix $\mathbf{A} \in \mathbb{C}^{M\times N}$, the Frobenius norm $\|\mathbf{A}\|_F$ is defined as

(69), p=2

$$\|\mathbf{A}\|_F \triangleq \sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}|a_{m,n}|^2}. \tag{71}$$

The Frobenius Norm and the Trace Function

- The Frobenius norm of $\mathbf{A}$ can be expressed as

$$\|\mathbf{A}\|_F = \boxed{\sqrt{\operatorname{tr}\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right)}}. \tag{72}$$

- Equation (62), which is $\boxed{\|\mathbf{x}\|_2 = \sqrt{\operatorname{tr}\left(\mathbf{x}\mathbf{x}^{\mathsf{H}}\right)}}$, shares a similar form to (72).

Properties of the Frobenius Norm

- Let $\mathbf{A} \in \mathbb{C}^{M \times N}$.
- Let $\mathbf{U} \in \mathbb{C}^{M \times M}$ and $\mathbf{V} \in \mathbb{C}^{N \times N}$ be unitary matrices.
- Namely, $\mathbf{U}^{\mathsf{H}}\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^{\mathsf{H}}\mathbf{V} = \mathbf{I}$.
- The Frobenius norm is unitarily invariant,

$$\|\mathbf{UAV}\|_F = \|\mathbf{A}\|_F \,. \tag{73}$$

Other Entry-Wise L_p Norms

- The entry-wise L_1 norm for $\mathbf{A} \in \mathbb{C}^{M \times N}$
  - $p = 1$. ℓ_1
  - Definition:

$$\|\mathbf{A}\|_1 \triangleq \sum_{m=1}^{M} \sum_{n=1}^{N} |a_{m,n}|. \tag{74}$$

- The entry-wise L_∞ norm for $\mathbf{A} \in \mathbb{C}^{M \times N}$
  - $p \to \infty$.
  - The max norm
  - Definition:

A

$$\|\mathbf{A}\|_\infty \triangleq \max_{(m,n) \in [M] \times [N]} |a_{m,n}|. \tag{75}$$

Remarks on the Matrix Norms

- The $\boxed{\text{entry-wise}}$ L_p norm
  - The Frobenius norm (L_2)
  - The L_1 norm
  - The max norm (L_∞)

|| matrix ||

- Other matrix norms
  - Operator norms or induced norms
  - Nuclear norms
  - We will cover these matrix norms after $\underbrace{\text{the singular value decomposition}}_{\text{SVD}}$.

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

$A =$

$A =$

EVD

eigenvectors

eigenvalues

$A =$

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

1. Motivations

Jordan form

2. Jordan Canonical Form
   - Definition and Examples
   - The Integer Power of a Matrix

SVD

3. Singular Value Decomposition (SVD)
   - Definition and Properties
   - Matrix Norms and SVD
4. Principal Component Analysis (PCA)

The Eigen-Decomposition of Square Matrices

- Let $\mathbf{A} \in \mathbb{C}^{N \times N}$. 2
- 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. 1
- 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

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

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 \ddots & \\ 0 & & \lambda_N \end{bmatrix}$$

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

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

{ eigenvalues
{ eigenvectors

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

3x3

- First, we consider the characteristic equation $\boxed{\det(\mathbf{A} - \lambda \mathbf{I}) = 0.}$
- 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

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

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 $\mathbf{A}$:

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

$\times\ VDV^{-1}$

Question

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

Example: Generalized Eigenvectors (1/3)

$(\underline{\underline{A}} - \lambda \underline{\underline{I}}) \underline{v_1} = \underline{0}$

$(\underline{\underline{A}} - \lambda \underline{\underline{I}}) \boxed{\underline{v_2}} = \underline{v_1}$

- 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.

(8) $(A-\lambda I)v_2 = v_1$

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 $\mathbf{A}$.