MA1522 CheatSheet AY23/24 —— @Jin Hang

Table of Contents

Chapter 7 Linear Transformation

1. **Linear Transformationsfrom $\mathbb{R}^n$ to $\mathbb{R}^m$**
   - Definition
   - Examples
   - Linearity
   - Representation
   - Change of Bases
   - Composition
2. **Ranges and Kernels**
   - Range of Function
   - Range of Linear Transformation
   - Representation of Range
   - Kernel of Linear Transformation.
   - Representation of Kernel
   - Properties

Chapter 2 Matrix

Definition: A matrix (plural matrices) is a rectangular array of numbers. m is the number of **rows** in the matrix., n is the number of **columns** in the matrix, the size of the matrix is given by $m \times n$

Special Matrices

- **Square matrix** same number of rows and columns i.e. $n \times n$ matrix.
- A square matrix is a **diagonal matrix** if all its non-diagonal entries are zero.
$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$
$\boldsymbol{A} = (a_{ij})_{n\times n}$ is diagonal $\Leftrightarrow a_{ij} = 0$ for all $i \neq j$.
- A diagonal matrix is a **scalar matrix**(标量矩阵) if all its diagonal entries are the same.
$$\boldsymbol{A} = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}$$
where c is a constant.
$\boldsymbol{A} = (a_{ij})_{n\times n}$ is scalar $\Leftrightarrow a_{ij} = \begin{cases} c, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$
- **identity matrix** $\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$
- A square matrix is **symmetric** if it is symmetric with respect to the diagonal.
$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots & a_{4n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & a_{nn} \end{pmatrix}$$
$\boldsymbol{A} = (a_{ij})_{n\times n}$ is symmetric $\Leftrightarrow a_{ij} = a_{ji}$ for all i, j.
- A square matrix is **upper triangular** if all entries below the diagonal are zero.
$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & a_{34} & \cdots & a_{3n} \\ 0 & 0 & 0 & a_{44} & \cdots & a_{4n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$
$\boldsymbol{A} = (a_{ij})_{n\times n}$ is upper triangular $\Leftrightarrow a_{ij} = 0$ if $i > j$.
- A square matrix is lower triangular if all entries above the diagonal are zero
$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & 0 & \cdots & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & \cdots & a_{nn} \end{pmatrix}.$$
$\boldsymbol{A} = (a_{ij})_{n\times n}$ is lower triangular $\Leftrightarrow a_{ij} = 0$ if $i < j$
- Both upper triangular matrices and lower triangular matrices are called **triangular matrices**.
- A matrix is both upper and lower triangular $\Leftrightarrow$ it is a **diagonal matrix**.

Matrix Multiplication

Let $\boldsymbol{A} = (a_{ij})_{m\times p}$ and $\boldsymbol{B} = (b_{ij})_{p\times n}$. $\boldsymbol{AB}$ is the $m \times n$ matrix such that its (i, j)-entry is
$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

Properties

Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ be $m \times p, p \times q, q \times n$ matrices, respectively.

- Associative Law: $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$.
- Distributive Law: $\boldsymbol{A}(\boldsymbol{B}_1 + \boldsymbol{B}_2) = \boldsymbol{AB}_1 + \boldsymbol{AB}_2$.
- Distributive Law: $(A_1 + A_2)B = A_1B + A_2B$.
- $c(\boldsymbol{AB}) = (c\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(c\boldsymbol{B})$.
- **matrix multiplication is NOT commutative**

Let $\boldsymbol{A}$ be an $m \times n$ matrix.

- $\boldsymbol{A0}_{n\times p} = \boldsymbol{0}_{m\times p}$; $\quad \boldsymbol{0}_{p\times m}\boldsymbol{A} = \boldsymbol{0}_{p\times n}$.
- $\boldsymbol{AI}_n = \boldsymbol{A}$; $\quad \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}$.

Powers of Square Matrices: Let $\boldsymbol{A}$ be an $m \times n$ matrix. $\boldsymbol{AA}$ is **well-defined** $\Leftrightarrow m = n \Leftrightarrow \boldsymbol{A}$ is square.

Let $\boldsymbol{A}$ be a square matrix of order n. For nonnegative integers k, the powers of $\boldsymbol{A}$ are defined as
$$\boldsymbol{A}^k = \begin{cases} \boldsymbol{I}_n & \text{if } k = 0 \\ \underbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}_{k \text{ times}} & \text{if } k \geq 1 \end{cases}$$

Properties

Let $\boldsymbol{A}$ be a square matrix, and m, n nonnegative integers.

- $\boldsymbol{A}^m\boldsymbol{A}^n = \boldsymbol{A}^{m+n}$
- $(\boldsymbol{A}^m)^n = \boldsymbol{A}^{mn}$.
- In general, $(\boldsymbol{AB})^n \neq \boldsymbol{A}^n\boldsymbol{B}^n$ for $n = 2, 3, \ldots$
- Only if $\boldsymbol{AB} = \boldsymbol{BA}$ then $(\boldsymbol{AB})^n = \boldsymbol{A}^n\boldsymbol{B}^n$

Representation of Linear System
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

- **coefficient matrix** $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$
- **variable matrix** $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- **constant matrix** $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

Then $\boldsymbol{Ax} = \boldsymbol{b}$

Transpose(转置矩阵): Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ be a matrix. The **transpose** of $\boldsymbol{A}$ is the $n \times m$ matrix $\boldsymbol{A}^{\mathrm{T}}$ (or $\boldsymbol{A}^{\mathrm{t}}$) whose (i, j)-entry is a_{ji}.

Properties Let $\boldsymbol{A}$ be an $m \times n$ matrix.

- $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.
- $\boldsymbol{A}$ is symmetric $\Leftrightarrow A = A^{\mathrm{T}}$.
- $(c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}$.
- $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$
- $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

Inverses of Square Matrices: Let $\boldsymbol{A}$ be a square matrix of order n. If there exists a square matrix $\boldsymbol{B}$ of order n so that
$$\mathbf{AB} = I_n \text{ and } \mathbf{B}\boldsymbol{A} = \boldsymbol{I}_n$$
then $\boldsymbol{A}$ is called invertible, and $\boldsymbol{B}$ is an **inverse** of $\boldsymbol{A}$.

If $\boldsymbol{A}$ is not invertible, $\boldsymbol{A}$ is called singular.(奇矩阵)

Note: Non-square matrix is neither invertible nor singular.

Property

- Let $\boldsymbol{A}$ be an invertible matrix, then its inverse is unique.
- **Cancellation Law** Let $\boldsymbol{A}$ be an invertible matrix.
$\boldsymbol{AB}_= \boldsymbol{AB}_2 \Rightarrow \boldsymbol{B}_1 = \boldsymbol{B}_2$.
$\boldsymbol{C}_1\boldsymbol{A} = \boldsymbol{C}_2\boldsymbol{A} \Rightarrow \boldsymbol{C}_1 = \boldsymbol{C}_2$.
In particular, if $\boldsymbol{A}$ is **invertible**, $\boldsymbol{AB} = \boldsymbol{0} \Rightarrow B = \boldsymbol{0}$ and $\boldsymbol{CA} = \boldsymbol{0} \Rightarrow \boldsymbol{C} = \boldsymbol{0}$.
- $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.
- $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$.
- $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$
- $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
- $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)^{-1} = \boldsymbol{A}_k^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$.
- $(\underbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}_{k \text{ times}})^{-1} = \underbrace{\boldsymbol{A}^{-1}\cdots\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}}_{k \text{ times}} \Leftrightarrow (\boldsymbol{A}^k)^{-1} = (\boldsymbol{A}^{-1})^k$.
- If $\boldsymbol{A}$ is singular, then $\boldsymbol{A}^{-1}$ is undefined.
- $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k$ is invertible $\Leftrightarrow$ **all** $\boldsymbol{A}_i$ are invertible.
- $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k$ is singular $\Leftrightarrow$ **some** $\boldsymbol{A}_i$ are singular.
- 矩阵方程的等价变形: $\boldsymbol{A}^2 - 3\boldsymbol{A} - 4\boldsymbol{I} = \boldsymbol{0}$
$\Rightarrow 4\boldsymbol{I} = \boldsymbol{A}^2 - 3\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{A} - 3\boldsymbol{I})$
$\Rightarrow \boldsymbol{I} = \frac{1}{4}\boldsymbol{A}(\boldsymbol{A} - 3\boldsymbol{I}) = \boldsymbol{A}\left[\frac{1}{4}(\boldsymbol{A} - 3\boldsymbol{I})\right]$
$\Rightarrow \boldsymbol{A}$ is invertible with $\boldsymbol{A}^{-1} = \frac{1}{4}(\boldsymbol{A} - 3\boldsymbol{I})$.

Theorem 1: Let $\boldsymbol{A}$ be a square matrix. Then the following are equivalent

1. $\boldsymbol{A}$ is an invertible matrix.
2. Linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a **unique** solution.
3. Linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the **trivial solution**.
4. The **reduced row-echelon form** of $\boldsymbol{A}$ is $\boldsymbol{I}$.
5. $\boldsymbol{A}$ is the product of **elementary matrices**.

$\Rightarrow$ Let $\boldsymbol{A}$ be an invertible matrix, the **reduced row-echelon form** of $(\boldsymbol{A} \mid \boldsymbol{I})$ is $\left(\boldsymbol{I} \mid \boldsymbol{A}^{-1}\right)$

Theorem 2: (矩阵可逆的等价条件)

1. A square matrix is invertible
   $\Leftrightarrow$ Its reduced row-echelon form is $\boldsymbol{I}$
   $\Leftrightarrow$ All the columns in its row-echelon form are pivot
   $\Leftrightarrow$ All the rows in its row-echelon form are nonzero.
2. A square matrix is singular
   $\Leftrightarrow$ Its reduced row-echelon form is not $\boldsymbol{I}$
   $\Leftrightarrow$ Some columns in its row-echelon form are non-pivot
   $\Leftrightarrow$ Some rows in its row-echelon form are zero

Find inverse Apply the **elementary row operations** corresponding to $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_k$

$$\begin{aligned}(\boldsymbol{A} \mid \boldsymbol{I}_n) &\xrightarrow{\boldsymbol{E}_1} (\boldsymbol{E}_1\boldsymbol{A} \mid \boldsymbol{E}_1)\\ &\xrightarrow{\boldsymbol{E}_2} (\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} \mid \boldsymbol{E}_2\boldsymbol{E}_1)\\ &\rightarrow \cdots \rightarrow \cdots\\ &\xrightarrow{\boldsymbol{E}_k} (\boldsymbol{E}_k\cdots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} \mid \boldsymbol{E}_k\cdots\boldsymbol{E}_2\boldsymbol{E}_1)\\ &= \left(\boldsymbol{I}_n \mid \boldsymbol{A}^{-1}\right).\end{aligned}$$

Elementary Matrices A square matrix is called an elementary matrix if it can be obtained from the identity matrix by performing a single elementary row operation.

Row operation （注意是左乘）Let $\boldsymbol{E}$ be the elementary matrix obtained by performing an **elementary row operation** to $\boldsymbol{I}_m$.

- $\boldsymbol{I}_m \xrightarrow{cR_i} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{EA}$.
- $\boldsymbol{I}_m \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{EA}$.
- $\boldsymbol{I}_m \xrightarrow{R_i + cR_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{EA}$.

Column Operations (注意是右乘)

- $\boldsymbol{I} \xrightarrow{kC_i} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{kC_i} \boldsymbol{AE}$
- $\boldsymbol{I} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{AE}$
- $\boldsymbol{I} \xrightarrow{C_i + kC_j} \boldsymbol{E} \Rightarrow \boldsymbol{A} \xrightarrow{C_i + kC_j} \boldsymbol{AE}$

Theorem: Every elementary matrix is **invertible**. Its inverse is also an elementary matrix of the same type.

Invertibility and Elementary Row Operation

- $\boldsymbol{I} \xrightarrow{cR_i} \boldsymbol{E} \Rightarrow \boldsymbol{I} \xrightarrow{\frac{1}{c}R_i} \boldsymbol{E}^{-1}$
- $\boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \Rightarrow \boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E}^{-1} \Rightarrow \boldsymbol{E} = \boldsymbol{E}^{-1}$
- $\boldsymbol{I} \xrightarrow{R_i + cR_j} \boldsymbol{E} \Rightarrow \boldsymbol{I} \xrightarrow{R_i - cR_j} \boldsymbol{E}^{-1}$

Suppose that
$\quad \boldsymbol{A} = \boldsymbol{A}_0 \xrightarrow{\text{ero1}} \boldsymbol{A}_1 \xrightarrow{\text{ero2}} \boldsymbol{A}_2 \rightarrow \cdots \rightarrow \boldsymbol{A}_{k-1} \xrightarrow{\text{erok}} \boldsymbol{A}_k = \boldsymbol{B}$

Note that $\boldsymbol{I} \xrightarrow{\text{eroi}} \boldsymbol{E}_i$
$\quad \boldsymbol{A} = \boldsymbol{A}_0 \xrightarrow[\boldsymbol{E}_1]{\text{ero1}} \boldsymbol{A}_1 \xrightarrow[\boldsymbol{E}_2]{\text{ero2}} \boldsymbol{A}_2 \rightarrow \cdots \rightarrow \boldsymbol{A}_{k-1} \xrightarrow[\boldsymbol{E}_k]{\text{erok}} \boldsymbol{A}_k = \boldsymbol{B}$

Then $\boldsymbol{B} = \boldsymbol{E}_k\boldsymbol{E}_{k-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}$.

Similarly $\boldsymbol{A} \xleftarrow{\boldsymbol{E}_1^{-1}} \bullet \xleftarrow{\boldsymbol{E}_2^{-1}} \bullet \leftarrow \cdots \leftarrow \bullet \xleftarrow{\boldsymbol{E}_{k-1}^{-1}} \bullet \xleftarrow{\boldsymbol{E}_k^{-1}} \boldsymbol{B}$

Hence, $\boldsymbol{A} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1}\cdots\boldsymbol{E}_{k-1}^{-1}\boldsymbol{E}_k^{-1}\boldsymbol{B}$

LU Decomposition

1. Let $\boldsymbol{A}$ be an $m \times n$ matrix. $\boldsymbol{A} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \boldsymbol{U}$.
   (Only **type III operations** $(R_i + cR_j, i > j)$ are used)
2. Let $\boldsymbol{U} \xrightarrow[\text{reversed}]{\text{operations}} \boldsymbol{A}$, similarly $\boldsymbol{I}_m \xrightarrow[\text{reversed}]{\text{operations}} \boldsymbol{L}$.
3. Then $\boldsymbol{A} = \boldsymbol{LU}$.
4. Then $\boldsymbol{L}$ is an **unique** lower triangular matrix with 1 along the diagonal
5. If $\boldsymbol{A}$ is a square matrix, $\boldsymbol{U}$ is an **upper triangular matrix**

Application of LU Decomposition

1. Suppose $\boldsymbol{A} = \boldsymbol{LU}$, Consider $\boldsymbol{Ax} = \boldsymbol{b} \Rightarrow \boldsymbol{LUx} = \boldsymbol{b}$
2. Let $\boldsymbol{y} = \boldsymbol{Ux}$.
3. $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{Ly} = \boldsymbol{b}$, get $\boldsymbol{y}$
4. Solve $\boldsymbol{Ux} = \boldsymbol{y}$

Partial Pivoting: Suppose type II $(R_i \leftrightarrow R_j)$ operation **must** be used in Gaussian elimination.

1. $\boldsymbol{A} \xrightarrow{\boldsymbol{E}_1} \bullet \xrightarrow{\boldsymbol{E}_2} \bullet \xrightarrow{R_i \leftrightarrow R_j} \bullet \xrightarrow{\boldsymbol{E}_4} \bullet \xrightarrow{\boldsymbol{E}_5} \boldsymbol{R}$
2. $\boldsymbol{A} \xleftarrow{\boldsymbol{E}_1^{-1}} \bullet \xleftarrow{\boldsymbol{E}_2^{-1}} \bullet \xleftarrow[R_i \leftrightarrow R_j]{\boldsymbol{E}_3} \bullet \xleftarrow{\boldsymbol{E}_4^{-1}} \bullet \xleftarrow{\boldsymbol{E}_5^{-1}} \boldsymbol{R}$
3. $\boldsymbol{A} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1}\boldsymbol{E}_3\boldsymbol{E}_4^{-1}\boldsymbol{E}_5^{-1}\boldsymbol{R}$
4. $\boldsymbol{E}_3\boldsymbol{A} = \left(\boldsymbol{E}_3\boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1}\boldsymbol{E}_3\right)\boldsymbol{E}_4^{-1}\boldsymbol{E}_5^{-1}\boldsymbol{R}$
5. Let permutation matrix $\boldsymbol{P} = \boldsymbol{E}_3$
   $\boldsymbol{L} = \left(\boldsymbol{E}_3\boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1}\boldsymbol{E}_3\right)\boldsymbol{E}_4^{-1}\boldsymbol{E}_5^{-1}$ and $\boldsymbol{R} = \boldsymbol{U}$
6. $\boldsymbol{PA}$ has an LU decomposition, $\boldsymbol{PA} = \boldsymbol{LU}$

Determinant (行列式)

Properties:

- $\det(\boldsymbol{I}_n) = 1$
- $\boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = c\det(\boldsymbol{A})$
- $\boldsymbol{A} \xrightarrow{R_1 \leftrightarrow R_2} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = -\det(\boldsymbol{A})$
- $\boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{B} \Rightarrow \det(\boldsymbol{B}) = \det(\boldsymbol{A})$, where $i \neq j$
- $\boldsymbol{A}$ is invertible $\Leftrightarrow \det(\boldsymbol{A}) \neq 0$, $\boldsymbol{A}$ is singular $\Leftrightarrow \det(\boldsymbol{A}) = 0$
- Suppose
$$\boldsymbol{A} \xrightarrow{\boldsymbol{E}_1} \bullet \xrightarrow{\boldsymbol{E}_2} \bullet \ldots \bullet \xrightarrow{\boldsymbol{E}_4} \bullet \xrightarrow{\boldsymbol{E}_5} \boldsymbol{R}$$
$$\boldsymbol{A} \xleftarrow{\boldsymbol{E}_1^{-1}} \bullet \xleftarrow{\boldsymbol{E}_2^{-1}} \bullet \ldots \bullet \xleftarrow{\boldsymbol{E}_4^{-1}} \bullet \xleftarrow{\boldsymbol{E}_5^{-1}} \boldsymbol{R}$$
  $\det(\boldsymbol{A}) = \det\left(\boldsymbol{E}_1^{-1}\right)\det\left(\boldsymbol{E}_2^{-1}\right)\cdots\det\left(\boldsymbol{E}_k^{-1}\right)$
- $\det(\boldsymbol{A}) = \det\left(\boldsymbol{A}^{\mathrm{T}}\right)$
- $\det(c\boldsymbol{A}) = c^n\det(\boldsymbol{A})$
- $\boldsymbol{A}$ is a **triangular matrix** $\Rightarrow \det(\boldsymbol{A})$ is the product of the **diagonal entries** of $\boldsymbol{A}$
- $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$
- $\det\left(\boldsymbol{A}^{-1}\right) = [\det(\boldsymbol{A})]^{-1}$

Adjoint Matrix (伴随矩阵): Let $\boldsymbol{A}$ be a square matrix of order n. The (classical) adjoint (or adjugate, or adjunct) of $\boldsymbol{A}$, $\operatorname{adj}(\boldsymbol{A}) = (A_{ji})_{n\times n} = (A_{ij})_{n\times n}^{\mathrm{T}}$, where A_{ij} is the (i,j)-cofactor of $\boldsymbol{A}$.

Property: Let $\boldsymbol{A}$ be a square matrix.
$$\boldsymbol{A}[\operatorname{adj}(\boldsymbol{A})] = \det(\boldsymbol{A})\boldsymbol{I}$$
$$[\operatorname{adj}(\boldsymbol{A})]\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}$$
If $\boldsymbol{A}$ is **invertible**, then $A^{-1} = \frac{1}{\det(\boldsymbol{A})}\operatorname{adj}(\boldsymbol{A})$.

Cramer's Rule

For invertible matrix $\boldsymbol{A} = (a_{ij})_{n\times n}$ and $\boldsymbol{b} = (b_i)_{n\times 1}$.

1. The linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$
2. Then $\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}[\operatorname{adj}(\boldsymbol{A})]\boldsymbol{b}$.
   $x_j = \frac{1}{\det(\boldsymbol{A})}(A_{1j}b_1 + A_{2j}b_2 + \cdots + A_{nj}b_n)$.
3. Let $\boldsymbol{A}_j$ be the matrix obtained by replacing the j th column of $\boldsymbol{A}$ by $\boldsymbol{b}$. where b_i is the (i,j)-entry of $\boldsymbol{A}_j$ and A_{ij} is the (i,j)-cofactor of $\boldsymbol{A}_j$. (将矩阵$\boldsymbol{A}$的第j行替换为$\boldsymbol{b}$得到矩阵$\boldsymbol{A}_j$)
4. Therefore, $x_j = \frac{\det(\boldsymbol{A}_j)}{\det(\boldsymbol{A})}, j = 1, \ldots, n$.
5. Generally, $\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{pmatrix}\det(\boldsymbol{A}_1)\\ \vdots \\ \det(\boldsymbol{A}_n)\end{pmatrix}$

Chapter 3 Vector Space

Scalar Multiplication: Let $\boldsymbol{v}=(v_1,v_2)$ and $c\in\mathbb{R}$ Then $c\boldsymbol{v}=(cv_1,cv_2)$

Euclidean Spaces: An n-vector or ordered n-tuple of real numbers is $\boldsymbol{v}=(v_1,v_2,...,v_i,...,v_n)$. The Euclidean n-space (or n-space) is the set of all n-vectors of real numbers. $\mathbb{R}_n=\{(v_1,v_2,...,v_n)|v_1,v_2,...,v_n\in\mathbb{R}\}$. In particular,

- If $n=1$, then $\mathbb{R}=\mathbb{R}^1$ is the real line.
- If $n=2$, then $\mathbb{R}^2$ is the xy-plane.
- If $n=3$, then $\mathbb{R}^3$ is the xyz-space.

Linear system $Ax=b$ of m equations and n variables. Then the solution set of $Ax=b$ is a subset of $\mathbb{R}^n$.

Implicit and Explicit Forms

Implicit form: $\{(x,y,z)|ax+by+cz=d\}$

Explicit form: $\{(x_0,y_0,z_0)+s(a_1,b_1,c_1)+t(a_2,b_2,c_2)|s,t\in\mathbb{R}\}$

Lines in $\mathbb{R}^3$

$\{(x,y,z)|a_1x+b_1y+c_1z=d_1 \text{ and } a_2x+b_2y+c_2z=d_2\}$

Exp: A line is given explicitly by $\{(t-2,-2t+3,t+1)|t\in\mathbb{R}\}$

1. Express t in terms of x,y $\begin{cases} t=x+2\\ -2t=y-3\\ t=z-1\end{cases}$
2. Augmented Matric
$$\left(\begin{array}{c|c}1 & x+2\\ -2 & y-3\\ 1 & z-1\end{array}\right)\xrightarrow[R_3-R_1]{R_2+2R_1}\left(\begin{array}{c|c}1 & x+2\\ 0 & 2x+y+1\\ 0 & -x+z-3\end{array}\right)\tag{1}$$
3. Hence, implicit form:
$$\{(x,y,z)|2x+y+1=0 \text{ and } -x+z-3=0\}$$

Linear Combination Let $\boldsymbol{v}_1=(2,1,3),\boldsymbol{v}_2=(1,-1,2)$ and $\boldsymbol{v}_3=(3,0,5)$. Is $\boldsymbol{v}=(3,3,4)$ a linear combination of $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$?

1. Suppose that $\boldsymbol{v}=a\boldsymbol{v}_1+b\boldsymbol{v}_2+c\boldsymbol{v}_3$,
i.e.,$(3,3,4)=a(2,1,3)+b(1,-1,2)+c(3,0,5)$
2. Solve the linear system $\begin{cases}2a+b+3c=3\\ a-b=3\\ 3a+2b+5c=4.\end{cases}$
$$\left(\begin{array}{ccc|c}2&1&3&3\\1&-1&0&3\\3&2&5&4\end{array}\right)\xrightarrow[\text{elimination}]{\text{Gaussian}}\left(\begin{array}{ccc|c}2&1&3&3\\0&-\frac{3}{2}&-\frac{3}{2}&\frac{3}{2}\\0&0&0&0\end{array}\right)$$
3. The system is consistent $\Rightarrow$ linear combination.
The system is inconsistent $\Rightarrow$ **not** linear combination.

Linear Span: Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$ be a subset of $\mathbb{R}^n$. The set of **all linear combinations** of $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k$: $\{c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k \mid c_1,c_2,\dots,c_k\in\mathbb{R}\}$ is called the **linear span** (or simply span) of S (or $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_n$). It is denoted by $\operatorname{span}(S)$ or $\operatorname{span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$.

- $\boldsymbol{v}$ is a linear combination $\Leftrightarrow \boldsymbol{v}\in\operatorname{span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$.

Criterion for $\operatorname{span}(S)=\mathbb{R}^n$

1. Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}\subseteq\mathbb{R}^n$. Determine whether $\operatorname{span}(S)=\mathbb{R}^n$.
2. For an arbitrary $\boldsymbol{v}\in\mathbb{R}^n$, check the consistency of the equation $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k=\boldsymbol{v}$.
3. Let $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_k\end{pmatrix}$.
$\Rightarrow \boldsymbol{A}\boldsymbol{x}=\boldsymbol{v}\Rightarrow(\boldsymbol{A}\mid\boldsymbol{v})\xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}}(\boldsymbol{R}\mid\boldsymbol{v}')$.
   - If $\boldsymbol{R}$ has a zero row, then $\operatorname{span}(S)\neq\mathbb{R}^n$.
   - If $\boldsymbol{R}$ has no zero row, then $\operatorname{span}(S)=\mathbb{R}^n$.
4. $\operatorname{span}(S)=\mathbb{R}^n\Leftrightarrow \boldsymbol{A}\boldsymbol{x}=\boldsymbol{v}$ is consistent for every $\boldsymbol{v}\in\mathbb{R}^n$
$\Leftrightarrow \boldsymbol{R}\boldsymbol{x}=\boldsymbol{v}'$ is consistent for every $\boldsymbol{v}'\in\mathbb{R}^n$
$\Leftrightarrow$ rightmost column of $(\boldsymbol{R}\mid\boldsymbol{v}')$ is non-pivot
$\Leftrightarrow$ all rows of $\boldsymbol{R}$ are nonzero

Theorem: Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$ be a subset of $\mathbb{R}^n$. If $k<n$, then $\operatorname{span}(S)\neq\mathbb{R}^n$.

Properties

- Let $S=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_k\}$ be a subset of $\mathbb{R}^n\Rightarrow\boldsymbol{0}\in\operatorname{span}(S)$
- Let $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_r\in\operatorname{span}(S),c_1,c_2,\dots,c_r\in\mathbb{R}$
$\Rightarrow c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_r\boldsymbol{v}_r\in\operatorname{span}(S)$. (加法/乘法封闭性)
- For $S_1=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_k\},S_2=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_m\}\subseteq\mathbb{R}^n$, $\operatorname{span}(S_1)\subseteq\operatorname{span}(S_2)\Leftrightarrow$ Every $\boldsymbol{u}_i$ is a linear combination of $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_m$.
- Let $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_k\end{pmatrix}$.
If $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{u}$ is **consistent**, then $\boldsymbol{u}\in\operatorname{span}(S)$.
If $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{u}$ is **inconsistent**, then $\boldsymbol{u}\notin\operatorname{span}(S)$.
Example: Prove $\operatorname{span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}\subseteq\operatorname{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$
$$\left(\begin{array}{ccc|c|c}1&1&-1&1&2\\0&1&2&2&-1\\1&2&1&3&1\end{array}\right)\xrightarrow[\text{Elimination}]{\text{Gaussian}}\left(\begin{array}{ccc|c|c}1&1&-1&1&2\\0&1&2&2&-1\\0&0&0&0&0\end{array}\right)$$
The systems $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{v}_j,j=1,2$, are all **consistent** $\Rightarrow\boldsymbol{v}_1,\boldsymbol{v}_2\in\operatorname{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}\Rightarrow\operatorname{span}\{\boldsymbol{v}_1,\boldsymbol{v}_2\}\subseteq\operatorname{span}\{\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3\}$
- $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_{k-1},\boldsymbol{v}_k\in\mathbb{R}^n$. If $\boldsymbol{v}_k$ is a linear combination of $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_{k-1}\Rightarrow\operatorname{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_{k-1}\}=\operatorname{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_{k-1},\boldsymbol{v}_k\}$.
$\Rightarrow$ Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_{k-1},\boldsymbol{v}_k\}$ and $V=\operatorname{span}(S)$. Remove $\boldsymbol{v}_i$ from $S\to S'\subseteq S\Rightarrow V=\operatorname{span}(S')$.

Subspaces: Let V be a subset of $\mathbb{R}^n$. Then V is called a **subspace** of $\mathbb{R}^n$ if $\exists\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\in\mathbb{R}^n$ s.t. $V=\operatorname{span}\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$.

- $\{\boldsymbol{0}\}=\operatorname{span}\{\boldsymbol{0}\}$ is the **zero space**.
- Let $\boldsymbol{e}_i$ denote the n-vector whose i th coordinate is 1 and elsewhere 0 . Then for every $\boldsymbol{v}=(c_1,c_2,\dots,c_n)\in\mathbb{R}^n$ $\boldsymbol{v}=c_1\boldsymbol{e}_1+c_2\boldsymbol{e}_2+\cdots+c_n\boldsymbol{e}_n$. $\mathbb{R}^n=\operatorname{span}\{\boldsymbol{e}_1,\boldsymbol{e}_2,\dots,\boldsymbol{e}_n\}$ is a subspace of $\mathbb{R}^n$.

If any of these fails, then V is not a subspace (of $\mathbb{R}^n$).

- $\boldsymbol{0}\in V$
- $c\in\mathbb{R}$ and $\boldsymbol{v}\in V\Rightarrow c\boldsymbol{v}\in V$,
- $\boldsymbol{u}\in V$ and $\boldsymbol{v}\in V\Rightarrow\boldsymbol{u}+\boldsymbol{v}\in V$.

Subspaces of $\mathbb{R}^1,\mathbb{R}^2,\mathbb{R}^3$

1. Subspaces of $\mathbb{R}^1$: $\{\boldsymbol{0}\}=\{0\}$, $\mathbb{R}$.
2. Subspaces of $\mathbb{R}^2$:
   - $\{\boldsymbol{0}\}=\{(0,0)\}$
   - A straight line passing through the origin $(0,0)$
   - $\mathbb{R}^2$
3. Subspaces of $\mathbb{R}^3$:
   - $\{\boldsymbol{0}\}=\{(0,0,0)\}$
   - A straight line passing through the origin $(0,0,0)$
   - A plane containing the origin $(0,0,0)$
   - $\mathbb{R}^3$

A subspace of $\mathbb{R}^i,i=1,2,3$, is always the **solution set** of a **homogeneous linear system Linear Independence**

Definition: Let $S=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ be a subset of $\mathbb{R}^n$.The equation $c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k=\boldsymbol{0}$ has a trivial solution $c_1=c_2=\cdots=c_k=0$.

- If the equation has a **non-trivial solution**($\exists c_i\neq0$), S is **linearly dependent**
- If the equation has only the **trivial solution**, then S is **linearly independent**

Property

- Let S_1,S_2 be finite subsets of $\mathbb{R}^n$ such that $S_1\subseteq S_2$.
(1) S_1 is linearly dependent $\Leftrightarrow$ some $\boldsymbol{v}_i$ is a linear combination of other vectors in $S\Rightarrow S_2$ is linearly dependent.
(2) S_2 is linearly independent $\Leftrightarrow$ no vector in S can be written as a linear combination of other vectors $\Rightarrow S_1$ is linearly independent.
- $c\boldsymbol{0}=\boldsymbol{0}$ has infinitely many solutions $c\in\mathbb{R}$.
- $\{\boldsymbol{0}\}$ is **linearly dependent**.
- If $\boldsymbol{0}\in S\,(\subseteq\mathbb{R}^n)$ then S is linearly dependent.
- Let $\boldsymbol{v}\in\mathbb{R}^n$. Then $c\boldsymbol{v}=\boldsymbol{0}\Leftrightarrow c=0$ or $\boldsymbol{v}=\boldsymbol{0}$.
- $\{\boldsymbol{v}\}$ is linearly independent $\Leftrightarrow\boldsymbol{v}\neq\boldsymbol{0}$.
- Let $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^n$. Then $\{\boldsymbol{u},\boldsymbol{v}\}$ is linearly dependent $\Leftrightarrow\boldsymbol{u}=a\boldsymbol{v}$ or $\boldsymbol{v}=a\boldsymbol{u}$ for some $a\in\mathbb{R}$
- Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}\subseteq\mathbb{R}^n$. If $k>n$, then S is linearly **dependent**.

Vector Space: A set V is called a vector space if V is a subspace of $\mathbb{R}^n$ for some positive integer n. If W and V are vector spaces such that $W\subseteq V$, then W is a **subspace** of V.

Bases: Let $S=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ be a subset of a vector space V. Then S is called a basis (plural bases) for V if S is linearly **independent**, and $\operatorname{span}(S)=V$.

Property

1. A basis for a vector space V contains
   - smallest possible number of vectors that spans V
   - largest possible number of vectors that is linearly independent
2. For convenience, $\varnothing$ is said to be the basis for $\{\boldsymbol{0}\}$
3. Other than $\{\boldsymbol{0}\}$, any vector space has **infinitely many** different bases

Coordinate Vector Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$ be a subset of a vector space V. S is a basis for $V\Leftrightarrow$ every vector $\boldsymbol{v}\in V$ can be uniquely written as $\boldsymbol{v}=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k$

Definition: Let $S=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$ be a basis for a vector space V. $\forall\boldsymbol{v}\in V$, $\exists!c_1,\dots,c_k\in\mathbb{R}$ s.t. $\boldsymbol{v}=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_k\boldsymbol{v}_k$. $(\boldsymbol{v})_S=(c_1,c_2,\dots,c_k)$ is the coordinate vector of $\boldsymbol{v}$ relative to S.

Remark: The order of $\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k$ is relevant.

- $S_1=\{(1,0),(0,1)\}\Rightarrow\boldsymbol{v}==x(1,0)+y(0,1)\Rightarrow(\boldsymbol{v})_{S_1}=(x,y)$.
- $S_2=\{(0,1),(1,0)\}\Rightarrow(\boldsymbol{v})_{S_2}=(y,x)$

- Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a basis for a vector space V and $\boldsymbol{v} \in V$
  View each vector as a column vector and let
  $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$
  Let $[\boldsymbol{v}]_S = \{(\boldsymbol{v})_S\}^{\mathrm{T}}$ be the column form of coordinate vector of $\boldsymbol{v} \Rightarrow [\boldsymbol{v}]_S$ is the **(unique) solution** to $\boldsymbol{Ax} = \boldsymbol{v} \Leftrightarrow \boldsymbol{A}[\boldsymbol{v}]_S = \boldsymbol{v}$

Standard Basis Let $E = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. It is a subset of $\mathbb{R}^n$, $\boldsymbol{E}$ is the standard basis for $\mathbb{R}^n$. For any $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, $(\boldsymbol{v})_E = (v_1, v_2, \ldots, v_n) = \boldsymbol{v}$

Property: Let S be a basis for a vector space V

- $(\boldsymbol{v})_S = \mathbf{0} \Leftrightarrow \boldsymbol{v} = \mathbf{0}$.
- $\forall c \in \mathbb{R}$ and $\boldsymbol{v} \in \mathbb{R}, (c\boldsymbol{v})_S = c(\boldsymbol{v})_S$.
- $\forall \boldsymbol{u}, \boldsymbol{v} \in V, (\boldsymbol{u} + \boldsymbol{v})_S = (\boldsymbol{u})_S + (\boldsymbol{v})_S$.
- $(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k)_S = c_1(\boldsymbol{v}_1)_S + \cdots + c_k(\boldsymbol{v}_k)_S$.
- For $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in V$, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are **linearly independent** $\Leftrightarrow (\boldsymbol{v}_1)_S, \ldots, (\boldsymbol{v}_k)_S$ are linearly independent.
- $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = V$
  $\Leftrightarrow \operatorname{span}\left\{(\boldsymbol{v}_1)_S, \ldots, (\boldsymbol{v}_k)_S\right\} = \mathbb{R}^k$, where $k = |S|$
- The function $V \to \mathbb{R}^k$ by $\boldsymbol{v} \mapsto (\boldsymbol{v})_S$ identifies V and $\mathbb{R}^k$.
  $\Rightarrow V$ and $\mathbb{R}^k$ are **isomorphic** as vector spaces.
- If S and T are bases for a vector space V, then $|S| = |T|$.

Dimension: Let V be a vector space and S a basis for V. The **dimension** of V is $\dim(V) = |S|$. Let $\boldsymbol{Ax} = \mathbf{0}$ be a homogeneous linear system. The solution set of $\boldsymbol{Ax} = \mathbf{0}$ is a vector space V. Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$.
number of non-pivot columns of $\boldsymbol{R}$
= number of arbitrary parameters in general solution
= the dimension of V.

Properties

- Let S be a subset of a vector space V. The following are equivalent:
  1. S is a basis for V.
  2. S is linearly independent, and $|S| = \dim(V)$.
  3. S spans V, and $|S| = \dim(V)$.

  Remark: To show that S is a basis for V, it suffices to check any two of the following conditions:
  1. S is linearly independent,
  2. $\operatorname{span}(S) = V$,
  3. $|S| = \dim(V)$.
- Suppose $U \subseteq V$, S is a basis for U
  $\Rightarrow S$ is linearly independent in V
  $\Rightarrow \dim(V) \geq |S| = \dim(U)$.
  If $\dim(U) = \dim(V)$, then $|S| = \dim(V)$
  $\Rightarrow S$ is a basis for V and $V = \operatorname{span}(S) = U$.
- Let U be a subspace of V. Then $\dim(U) \leq \dim(V)$.
$$U = V \Leftrightarrow \dim(U) = \dim(V)$$
$$U \neq V \Leftrightarrow \dim(U) < \dim(V)$$
- Let $\boldsymbol{A}$ be a square matrix of order n, the following are equivalent:
  1. $\boldsymbol{A}$ is invertible.
  2. $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.
  3. $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution.
  4. The reduced row-echelon form of $\boldsymbol{A}$ is $\boldsymbol{I}_n$.
  5. A is a product of elementary matrices.
  6. $\det(\boldsymbol{A}) \neq 0$.
  7. The rows of A form a basis for $\mathbb{R}^n$.
  8. The columns of A form a basis for $\mathbb{R}^n$.

Transition Matrix: Let V be a vector space, and $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ and T be bases for V. $\begin{pmatrix} [\boldsymbol{u}_1]_T & \cdots & [\boldsymbol{u}_k]_T \end{pmatrix}$ is the **transition matrix** from S to T. Denoted by $\boldsymbol{P}$. Then $\boldsymbol{P}[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$ for all $\boldsymbol{w} \in V$.
Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ View all vectors as column vectors.
$$\left(\begin{array}{ccc|c|c|c} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} (\boldsymbol{I} \mid \boldsymbol{P})$$

Properties

- $[\boldsymbol{v}]_{S_1} \xrightarrow{\boldsymbol{P}} [\boldsymbol{v}]_{S_2} \xrightarrow{\boldsymbol{Q}} [\boldsymbol{v}]_{S_3}$
  $\Leftrightarrow [\boldsymbol{v}]_{S_3} = \boldsymbol{Q}[\boldsymbol{v}]_{S_2} = \boldsymbol{QP}[\boldsymbol{v}]_{S_1}$
  $\Leftrightarrow \boldsymbol{QP}$ is the transition matrix from S_1 to S_3.
- $[\boldsymbol{v}]_S \xrightarrow{\boldsymbol{P}} [\boldsymbol{v}]_T$
  $\Leftrightarrow [\boldsymbol{v}]_T \xrightarrow{\boldsymbol{P^{-1}}} [\boldsymbol{v}]_S$
  $\Leftrightarrow \boldsymbol{P}[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$ and $\boldsymbol{P}^{-1}[\boldsymbol{w}]_T = [\boldsymbol{w}]_S$

Chapter 4 Vector Space and Matrix

Let $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}$.

Row space: Let $\boldsymbol{r}_i=\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in}\end{pmatrix}$ denote the i th row of $\boldsymbol{A}$. Then $\boldsymbol{r}_i \in \mathbb{R}^n$ and $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m\end{pmatrix}$. The row space of $\mathbf{A}$ is vector space spanned by the rows of $\mathbf{A}$: span $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_m\}$.

Column space: Let $\boldsymbol{c}_j=\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj}\end{pmatrix}$ denote the j th column of $\boldsymbol{A}$. Then $\boldsymbol{c}_j \in \mathbb{R}^m$ and $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n\end{pmatrix}$. The column space of $\boldsymbol{A}$ is the vector space spanned by the columns of $\boldsymbol{A}$: span $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n\}$.

Properties

1. Both row space and column space are subspace of $\mathbb{R}^m$.
2. The row space of $\boldsymbol{A}=$ the column space of $\boldsymbol{A}^{\mathrm{T}}$
3. The column space of $\boldsymbol{A}=$ the row space of $\boldsymbol{A}^{\mathrm{T}}$
4. dim(Row Space) = dim(Column Space)
5. **Consistency:** Let $\boldsymbol{A}$ be an $m \times n$ matrix. The column space of $\boldsymbol{A}$ is $\{\boldsymbol{A}\boldsymbol{v} \mid \boldsymbol{v} \in \mathbb{R}^n\}$ The linear system $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ is consistent $\Leftrightarrow \boldsymbol{b}$ lies in the column space of $\boldsymbol{A}$.

Row Equivalence Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be matrices of the same size. $\boldsymbol{A}$ and $\boldsymbol{B}$ are **row equivalent** if one can be obtained from another by a series of **elementary row operations**.

$$A=A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_k \rightarrow A_{k-1} \rightarrow A_k=B$$

Properties

- $\boldsymbol{A}$ and $\boldsymbol{B}$ are row equivalent $\Leftrightarrow \boldsymbol{A}$ and $\boldsymbol{B}$ have same row spaces

Let $\boldsymbol{R}$ be a **row-echelon form** of $\boldsymbol{A}$.

Row Relation

- The row space of $\boldsymbol{A}=$ the row space of $\boldsymbol{R}$
- The nonzero rows of $\boldsymbol{R}$ are linearly independent.
- The nonzero rows of $\boldsymbol{R}$ form a **basis** for the row space of $\boldsymbol{A}$
- The number of nonzero rows of $\boldsymbol{R}=$ dim(row space of $\boldsymbol{A}$)
  $\Rightarrow$ Form the basis of $\boldsymbol{A}$

Column Relation

- The pivot columns of $\boldsymbol{R}$ form a basis for the column space of $\boldsymbol{R}$
  $\Rightarrow$ form a basis for the column space of $\boldsymbol{A}$
- The number of pivot columns of $\boldsymbol{R}=$ dim(column space of $\boldsymbol{A}$)
  $\Rightarrow$ Form the basis of $\boldsymbol{A}$

注意: 在求解行空间与列空间时请**不要直接使用**简化后的行或列，而是要去**原矩阵**找对应的行和列！

Extend S to a basis for $\mathbb{R}^n$:

$S=\{(1,4,-2,5,1),(2,9,-1,8,2),(2,9,-1,9,3)\} \rightarrow \boldsymbol{R}$

$$\begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 2 & 9 & -1 & 8 & 2 \\ 2 & 9 & -1 & 9 & 3\end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 & 1\end{pmatrix}$$

- S is linearly independent
- The 1st, 2nd, 4th columns of $\boldsymbol{R}$ are pivot.
- Add rows to the non pivot column of $\boldsymbol{R}$

$$\begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}\end{pmatrix}$$

$S \cup\{(0,0,1,0,0),(0,0,0,0,1)\}$ is a basis for $\mathbb{R}^5$

Rank: Let $\boldsymbol{A}$ be a matrix, dim(row space of $\boldsymbol{A}$) = dim(column space of $\boldsymbol{A}$) is called the rank of $\boldsymbol{A}$, denoted by rank($\boldsymbol{A}$).

Properties For matrix $\boldsymbol{A}_{m \times n}$

- $\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}\left(\boldsymbol{A}^{\mathrm{T}}\right)$.
- $\operatorname{rank}(\boldsymbol{A})=0 \Leftrightarrow \boldsymbol{A}=\mathbf{0}$.
- $\operatorname{rank}(\boldsymbol{A}) \leq m$ and $\operatorname{rank}(\boldsymbol{A}) \leq n$.
- $\operatorname{rank}(\boldsymbol{A}) \leq \min\{m, n\}$.
- **Full rank:** $\operatorname{rank}(\boldsymbol{A})=\min\{m, n\}$.
- A square matrix $\boldsymbol{A}$ is of full rank $\Leftrightarrow \boldsymbol{A}$ is invertible.

Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$, a row-echelon form of $(\boldsymbol{A} \mid \boldsymbol{b})$ is of the form $(\boldsymbol{R} \mid \boldsymbol{b}')$.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \text{ is consistent } & \Leftrightarrow \boldsymbol{b}' \text{ is non-pivot in } (\boldsymbol{R} \mid \boldsymbol{b}') \\ & \Leftrightarrow \operatorname{rank}(\boldsymbol{R})=\operatorname{rank}(\boldsymbol{R} \mid \boldsymbol{b}') \\ & \Leftrightarrow \operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b}) \end{aligned}$$

Generally, $\operatorname{rank}(\boldsymbol{A}) \leq \operatorname{rank}(\boldsymbol{A} \mid \boldsymbol{b}) \leq \operatorname{rank}(\boldsymbol{A})+1$

Theorem For matrice $\boldsymbol{A}_{m \times n}$ and $\boldsymbol{B}_{n \times p}$

- col. space of $\boldsymbol{AB} \subseteq$ col. space of $\boldsymbol{A} \Leftrightarrow \operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{A})$
- row space of $\boldsymbol{AB} \subseteq$ row space of $\boldsymbol{B} \Leftrightarrow \operatorname{rank}(\boldsymbol{AB}) \leq \operatorname{rank}(\boldsymbol{B})$
- $\operatorname{rank}(\boldsymbol{AB}) \leq \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}$

Nullspace: For matrix $\boldsymbol{A}_{m \times n}$, the **nullspace** of $\boldsymbol{A}$ is the solution space of $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$: $\{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{v}=\mathbf{0}\}$. The **dimension** of the nullspace is the **nullity** of $\boldsymbol{A}$, denoted by nullity ($\boldsymbol{A}$). **vectors in nullspace are viewed as column vectors**.

Remarks: Let $\boldsymbol{R}$ be a row-echelon form of $\boldsymbol{A}$
$\operatorname{nullity}(\boldsymbol{A})=\operatorname{nullity}(\boldsymbol{R})=$ the number of non-pivot columns of $\boldsymbol{R}$

Dimension Theorem For matrix $\boldsymbol{A}_{m \times n}$

$$\begin{aligned} & \operatorname{rank}(\boldsymbol{A})+\operatorname{nullity}(\boldsymbol{A}) \\ = & \operatorname{rank}(\boldsymbol{R})+\operatorname{nullity}(\boldsymbol{R}) \\ = & (\text{ number of pivot columns of } \boldsymbol{R})+(\text{ number of non-pivot columns of } \boldsymbol{R}) \\ = & \text{ number of columns of } \boldsymbol{R}=n . \end{aligned}$$

Solution Theorem Let $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ be consistent. Fix a solution $\boldsymbol{v}$.

$$\begin{aligned} \boldsymbol{u} \text{ is a solution to } \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} & \Leftrightarrow \boldsymbol{A}\boldsymbol{u}=\boldsymbol{b} \\ & \Leftrightarrow \boldsymbol{A}\boldsymbol{u}-\boldsymbol{b}=\mathbf{0} \\ & \Leftrightarrow \boldsymbol{A}(\boldsymbol{u}-\boldsymbol{v})=\mathbf{0} \\ & \Leftrightarrow \boldsymbol{u}-\boldsymbol{v} \in \text{ nullspace of } \boldsymbol{A} \\ & \Leftrightarrow \boldsymbol{u}=\boldsymbol{v}+\boldsymbol{w}, \boldsymbol{w} \in \text{ nullspace of } \boldsymbol{A} \end{aligned}$$

Suppose the **nullspace** of $\boldsymbol{A}$ is spanned by $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$. The solution set of $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ is $\{\boldsymbol{v}+t_1\boldsymbol{u}_1+\cdots+t_k\boldsymbol{u}_k \mid t_1, \ldots, t_k \in \mathbb{R}\}$.

Remark $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ has only one solution $\boldsymbol{v}$
$\Leftrightarrow \boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ has only one solution $\mathbf{0} \Leftrightarrow$ nullspace of $\boldsymbol{A}=\{\mathbf{0}\}$
$\Leftrightarrow \operatorname{nullity}(\boldsymbol{A})=0$
$\Leftrightarrow \operatorname{rank}(\boldsymbol{A})=$ number of columns of $\boldsymbol{A}$
$\Leftrightarrow$ columns of $\boldsymbol{A}$ are linearly independent.

Chapter 5 Orthogonality

Properties of Vector

1. $-1 \le \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} \le 1 \Rightarrow |\boldsymbol{u}\cdot\boldsymbol{v}| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$
2. $\boldsymbol{u}\cdot\boldsymbol{v} = u_1v_1 + u_2v_2 = \begin{pmatrix} u_1 & u_2 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \boldsymbol{u}\boldsymbol{v}^{\mathbf{T}} = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$
3. $d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{\sum_{i=1}^{n}(u_i-v_i)^2}$
4. **Inequality**
   - $|\boldsymbol{u}\cdot\boldsymbol{v}| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$. (Cauchy-Schwarz inequality)
   - $\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. (Triangle inequality)
   - $d(\boldsymbol{u},\boldsymbol{w}) \le d(\boldsymbol{u},\boldsymbol{v}) + d(\boldsymbol{v},\boldsymbol{w})$. (Triangle inequality)
5. **Transpose Matrix**

$$\begin{aligned}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \mathbf{0} &\Leftrightarrow \boldsymbol{c}_i\cdot\boldsymbol{c}_j = 0 \text{ for all } i,j = 1,\dots,m \\ &\Rightarrow \boldsymbol{c}_i\cdot\boldsymbol{c}_i = 0 \text{ for all } i = 1,\dots,m \\ &\Leftrightarrow \|\boldsymbol{c}_i\|^2 = 0 \text{ for all } i = 1,\dots,m \\ &\Leftrightarrow \boldsymbol{c}_i = \mathbf{0} \text{ for all } i = 1,\dots,m \\ &\Leftrightarrow \boldsymbol{A} = \mathbf{0}\end{aligned}$$

Orthogonal: $\boldsymbol{u}\cdot\boldsymbol{v} = 0$ denoted by $\boldsymbol{u}\perp\boldsymbol{v}$

Orthonormal: orthogonal + **unit vector**

Relationship

1. S is orthognal $\Rightarrow \begin{cases} \text{subset of } S \text{ is orthogonal} \\ S\cup\{\mathbf{0}\} \text{ is also orthogonal}\end{cases}$
2. S is orthonormal $\Rightarrow \begin{cases} \text{subset of } S \text{ is orthonormal} \\ S \text{ is also orthogonal} \\ \mathbf{0}\notin S\end{cases}$

Normalize: $\boldsymbol{u}_i \mapsto \boldsymbol{v}_i = \frac{\boldsymbol{u}_i}{\|\boldsymbol{u}_i\|}$

Properties Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$

- $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is a diagonal matrix $\Leftrightarrow \boldsymbol{A}$ is orthogonal
- $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}_k \Leftrightarrow \boldsymbol{A}$ is orthonormal $\Rightarrow \boldsymbol{A}$ is invertible $\Rightarrow \{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_n\}$ is a basis for $\mathbb{R}^n$
- Let $\boldsymbol{A}$ be an orthogonal matrix. $\boldsymbol{I} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}}$. $\Rightarrow \boldsymbol{A}^{\mathrm{T}}\left(=\boldsymbol{A}^{-1}\right)$ is an orthogonal matrix.
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be orthogonal matrices of the same order $\Rightarrow (\boldsymbol{AB})^{\mathrm{T}}(\boldsymbol{AB}) = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AB} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{I} \Rightarrow$ orthogonal
- $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}_n \Leftrightarrow$ columns of $\boldsymbol{A}$ form an orthonormal set in $\mathbb{R}^m$ $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}_m \Leftrightarrow$ rows of $\boldsymbol{A}$ form an orthonormal set in $\mathbb{R}^n$
- A square matrix $\boldsymbol{A}$ of order n is an orthogonal matrix $\Leftrightarrow$ the columns of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$ $\Leftrightarrow$ the rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$.
- Let $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}\subseteq\mathbb{R}^n$ be an orthonormal set of vectors, $\boldsymbol{P}$ be an orthogonal matrix of order n, then $\{\boldsymbol{P}\boldsymbol{u}_1,\dots,\boldsymbol{P}\boldsymbol{u}_k\}$ is an orthonormal set of vectors.
$$(\boldsymbol{PA})^{\mathrm{T}}(\boldsymbol{PA}) = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{PA} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$$
- An orthogonal set of nonzero vectors is linearly independent
$$\boldsymbol{v}_i\cdot(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k) = \boldsymbol{v}_i\cdot\mathbf{0} = 0$$
$$\begin{aligned}0 &= \boldsymbol{v}_i\cdot(c_1\boldsymbol{v}_1 + \boldsymbol{c_2}\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k) \\ &= \boldsymbol{v}_i\cdot(c_1\boldsymbol{v}_1) + \boldsymbol{v}_i\cdot(c_2\boldsymbol{v}_2) + \cdots + \boldsymbol{v}_i\cdot(c_k\boldsymbol{v}_k) \\ &= c_1(\boldsymbol{v}_i\cdot\boldsymbol{v}_1) + c_2(\boldsymbol{v}_i\cdot\boldsymbol{v}_2) + \cdots c_i(\boldsymbol{v}_i\cdot\boldsymbol{v}_i) + \cdots + c_k(\boldsymbol{v}_i\cdot\boldsymbol{v}_k)\end{aligned}$$
Recall that $\boldsymbol{v}_i\cdot\boldsymbol{v}_j = 0$ if $i\ne j$.
The above equation is reduced to $c_i(\boldsymbol{v}_i\cdot\boldsymbol{v}_i) = 0$.
$$\boldsymbol{v}_i \ne \mathbf{0} \Rightarrow \boldsymbol{v}_i\cdot\boldsymbol{v}_i > 0 \Rightarrow c_i = 0.$$
Therefore, S is linearly independent.
- Let $S = \{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}$ be a basis for V. $\exists! c_1,\dots,c_k\in\mathbb{R}$ s.t. $\boldsymbol{w} = c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k$, coordinate vector $(\boldsymbol{w})_S = (c_1,\dots,c_k)$.
$$\boldsymbol{w} = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \cdots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$
Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k\end{pmatrix}$. For $\boldsymbol{A}[\boldsymbol{w}]_S = \boldsymbol{w}$, if S is orthogonal.
$$\begin{aligned}\boldsymbol{w}\cdot\boldsymbol{u}_i &= (c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k)\cdot\boldsymbol{u}_i \\ &= c_1(\boldsymbol{u}_1\cdot\boldsymbol{u}_i) + \cdots + c_k(\boldsymbol{u}_k\cdot\boldsymbol{u}_i) \\ &= c_i(\boldsymbol{u}_i\cdot\boldsymbol{u}_i) \\ c_i &= \frac{\boldsymbol{w}\cdot\boldsymbol{u}_i}{\boldsymbol{u}_i\cdot\boldsymbol{u}_i} = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_i}{\|\boldsymbol{u}_i\|^2}\end{aligned}$$
- Let $V = \operatorname{span}\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ be a vector space. $\boldsymbol{w}$ is orthogonal to $V \Leftrightarrow \boldsymbol{w}\cdot\boldsymbol{v}_i = 0$ for all i. $\Rightarrow$ trivial since $\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\in V$.

Projection (u on v): $\boldsymbol{p} = \|\boldsymbol{u}\|\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{v}\|^2}\right)\boldsymbol{v}$

Property

- Let $\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}$ be orthogonal basis for V, the projection of $\boldsymbol{w}$ on V is $\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \cdots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$ (sum of projections of $\boldsymbol{w}$ onto $\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_k$)
- Let $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_k\}$ be a basis for a vector space V. Let
$$\begin{aligned}\boldsymbol{v}_1 &= \boldsymbol{u}_1 \\ \boldsymbol{v}_2 &= \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2\cdot\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 \\ \boldsymbol{v}_3 &= \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3\cdot\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3\cdot\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 \\ \vdots\ & \vdots \\ \boldsymbol{v}_k &= \boldsymbol{u}_k - \frac{\boldsymbol{u}_k\cdot\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_k\cdot\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{u}_k\cdot\boldsymbol{v}_{k-1}}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}\end{aligned}$$
Then, $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots,\boldsymbol{v}_k\}$ is an orthogonal basis for V.
- Projection $\boldsymbol{p}$ has the shortest distance to $\boldsymbol{u}$.
$d(\boldsymbol{u},\boldsymbol{p}) \le d(\boldsymbol{u},\boldsymbol{v}) \quad$ for all $\boldsymbol{v}\in V$.
The projection $\boldsymbol{p}$ is called the **best approximation** of $\boldsymbol{u}$ in V

Least Squares Solutions: $\boldsymbol{Ax} = \boldsymbol{b}$ is inconsistent, find least squares solution, $\|\boldsymbol{Au}-\boldsymbol{b}\| \le \|\boldsymbol{Ax}-\boldsymbol{b}\|$ for any $\boldsymbol{x}\in\mathbb{R}^n$.

- Let $\boldsymbol{A}$ be an $m\times n$ matrix and $\boldsymbol{b}\in\mathbb{R}^n$, $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{Au} = \boldsymbol{p}$, the projection of $\boldsymbol{b}$ onto the **column space** of $\boldsymbol{A}$.
- Let The column space $V = \operatorname{span}(A) = \operatorname{span}\{\boldsymbol{c}_1,\dots,\boldsymbol{c}_n\}$

$\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$
$\Leftrightarrow \boldsymbol{Au}$ is the projection of $\boldsymbol{b}$ onto V
$\Leftrightarrow \boldsymbol{b}-\boldsymbol{Au}$ is orthogonal to V
$\Leftrightarrow \boldsymbol{b}-\boldsymbol{Au}$ is orthogonal to $\boldsymbol{c}_i$ for all $i = 1,\dots,n$
$\Leftrightarrow \boldsymbol{c}_i\cdot(\boldsymbol{b}-\boldsymbol{Au}) = 0$ for all $i = 1,\dots,n$
$\Leftrightarrow \boldsymbol{c}_i^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{Au}) = 0$ for all $i = 1,\dots,n$
$\Leftrightarrow \begin{pmatrix} \boldsymbol{c}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{c}_n^{\mathrm{T}}\end{pmatrix}(\boldsymbol{b}-\boldsymbol{Au}) = \mathbf{0}$
$\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{Au}) = \mathbf{0}$
$\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{Au} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
$\Leftrightarrow \boldsymbol{p} = \boldsymbol{Au}$

- Find LSS $\Leftrightarrow$ Solve the system $\boldsymbol{Ax} = \boldsymbol{p}$ $\Leftrightarrow$ Solve the system $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
- If $\boldsymbol{Ax} = \boldsymbol{b}$ is consistent, its **least squares solutions** are **precisely** all its solutions.

QR Decomposition: Let $\boldsymbol{A}$ be an $m\times n$ matrix whose columns are linearly independent. Then there exist an $m\times n$ matrix $\boldsymbol{Q}$ whose columns form an **orthonormal** set, and an **invertible upper triangular matrix** $\boldsymbol{R}$ of order n such that $A = QR$.

Algorithm: Let $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\dots,\boldsymbol{u}_n\}$ be the columns of $\boldsymbol{A}$.

1. Use Gram-Schmidt process to obtain orthonormal basis $\{\boldsymbol{w}_1,\boldsymbol{w}_2,\dots,\boldsymbol{w}_n\}$ for the column space of $\boldsymbol{A}$.
2. $\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_n\end{pmatrix}$.
3. $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{w}_1\cdot\boldsymbol{u}_1 & \boldsymbol{w}_1\cdot\boldsymbol{u}_2 & \cdots & \boldsymbol{w}_1\cdot\boldsymbol{u}_n \\ 0 & \boldsymbol{w}_2\cdot\boldsymbol{u}_2 & \cdots & \boldsymbol{w}_2\cdot\boldsymbol{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{w}_n\cdot\boldsymbol{u}_n\end{pmatrix} = (\boldsymbol{w}_i\cdot\boldsymbol{u}_j)_{n\times n}$

Property

- $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{QRx} = \boldsymbol{b} \Rightarrow \boldsymbol{Rx} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$
- $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$
$\Leftrightarrow \boldsymbol{u}$ is a solution to $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$
$\Leftrightarrow \boldsymbol{u}$ is a solution to $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{Rx} = \boldsymbol{R}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$
$\Leftrightarrow \boldsymbol{u}$ is a solution to $\boldsymbol{Rx} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$.
- $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique least squares solution $\Leftrightarrow$ unique solution to $\boldsymbol{Rx} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{b}$.

Transition Matrix: Let $S = \{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\}$ be orthonormal bases for V.
Let $\boldsymbol{A} = \begin{pmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k\end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k\end{pmatrix}$, $\boldsymbol{w}\in V$
$$\boldsymbol{w} = \boldsymbol{A}[\boldsymbol{w}]_S = \boldsymbol{B}[\boldsymbol{w}]_T$$

- $[\boldsymbol{w}]_T = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}[\boldsymbol{w}]_T = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}[\boldsymbol{w}]_S$, $\boldsymbol{P} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ is the transition matrix from S to T
- $[\boldsymbol{w}]_S = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}[\boldsymbol{w}]_S = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}[\boldsymbol{w}]_T$, $\boldsymbol{Q} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}$ is the transition matrix from T to S.

$\boldsymbol{P}^{\mathrm{T}} = \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}\right)^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B} = \boldsymbol{Q}$. Since $\boldsymbol{P}^{-1} = \boldsymbol{Q} \Rightarrow \boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}^{-1}$
$\Rightarrow \boldsymbol{P}$ (and hence $\boldsymbol{Q}$) is an orthogonal matrix.

Geometric Representation Let $\boldsymbol{P}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

$\forall \boldsymbol{u} \in \mathbb{R}^2$, $\boldsymbol{P}_\theta \boldsymbol{u} =$ rotation of $\boldsymbol{u}$ about O by θ anticlockwise.

Fix angles α and β. Then for any $\boldsymbol{u} \in \mathbb{R}^2$,

$$\begin{aligned}
\boldsymbol{P}_\alpha \boldsymbol{u} &= \text{ rotation of } \boldsymbol{u} \text{ about } O \text{ by } \alpha \text{ anticlockwise} \\
\boldsymbol{P}_\beta \left(\boldsymbol{P}_\alpha \boldsymbol{u} \right) &= \text{ rotation of } \boldsymbol{P}_\alpha \boldsymbol{u} \text{ about } O \text{ by } \beta \text{ anticlockwise} \\
&= \text{ rotation of } \boldsymbol{u} \text{ about } O \text{ by } \alpha + \beta \\
&= \boldsymbol{P}_{\alpha+\beta} \boldsymbol{u}.
\end{aligned}$$

Therefore, $\boldsymbol{P}_{\alpha+\beta} = \boldsymbol{P}_\beta \boldsymbol{P}_\alpha$.

Chapter 6 Diagonalization

Theorem

Suppose $\exists$ invertible matrix $\boldsymbol{P}$ so that $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is diagonal.

$$\begin{aligned}\boldsymbol{D}^m &= \left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)^m\\ &= \underbrace{\left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)\left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)\cdots\left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)\left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)}_{m\text{ copies}}\\ &= \boldsymbol{P}^{-1}\boldsymbol{A}\left(\boldsymbol{P}\boldsymbol{P}^{-1}\right)\boldsymbol{A}\left(\boldsymbol{P}\boldsymbol{P}^{-1}\right)\boldsymbol{A}\cdots\boldsymbol{A}\left(\boldsymbol{P}\boldsymbol{P}^{-1}\right)\boldsymbol{A}\boldsymbol{P}\\ &= \boldsymbol{P}^{-1}\underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}\boldsymbol{A}}_{m\text{ copies}}\boldsymbol{P}\\ &= \boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}.\end{aligned}$$

Hence, $\boldsymbol{A}^m=\boldsymbol{P}\boldsymbol{D}^m\boldsymbol{P}$.

Diagonalizable Matrices Let A be a square matrix. It is called **diagonalizable** if there is an **invertible matrix** $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.

$$\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}$$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}\Leftrightarrow\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{D}\Leftrightarrow$$

$$\begin{pmatrix}\boldsymbol{A}\boldsymbol{v}_1 & \cdots & \boldsymbol{A}\boldsymbol{v}_n\end{pmatrix}=\begin{pmatrix}\lambda_1\boldsymbol{v}_1 & \cdots & \lambda_n\boldsymbol{v}_n\end{pmatrix}$$

<u>**Property**</u>

- $\boldsymbol{A}$ is diagonalizable $\Leftrightarrow$ $\boldsymbol{A}$ has n linearly independent eigenvectors.
- Solve $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\,\boldsymbol{x}=\boldsymbol{0}$ to find a basis S_i for the eigenspace E_{λ_i}.

$$\begin{aligned}\boldsymbol{A}\text{ is diagonalizable } &\Leftrightarrow |S_1|+\cdots+|S_k|=n,\\ \boldsymbol{A}\text{ is not diagonalizable } &\Leftrightarrow |S_1|+\cdots+|S_k|<n.\end{aligned}$$

- $\boldsymbol{D}$ is not unique unless $\boldsymbol{A}$ has only one eigenvalue.

Definition Let $\boldsymbol{A}$ be a square matrix of order n. Suppose that for some $\lambda\in\mathbb{R}$ and nonzero $v\in\mathbb{R}^n$ such that $\boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v}$

- λ is called an **eigenvalue**(特征值) of $\boldsymbol{A}$.
- $\boldsymbol{v}$ is called an **eigenvector**(特征向量) of $\boldsymbol{A}$ associated to the eigenvalue λ.

<u>**Characteristic Equation**</u> Let $\boldsymbol{A}$ be a square matrix of order n.

$$\begin{aligned}\lambda\in\mathbb{R}\text{ is an eigenvalue of }\boldsymbol{A} &\Leftrightarrow \boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v}\text{ for some }\boldsymbol{0}\neq\boldsymbol{v}\in\mathbb{R}^n\\ &\Leftrightarrow \lambda\boldsymbol{I}\boldsymbol{v}-\boldsymbol{A}\boldsymbol{v}=\boldsymbol{0}\text{ for some }\boldsymbol{0}\neq\boldsymbol{v}\in\mathbb{R}^n\\ &\Leftrightarrow (\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v}=\boldsymbol{0}\text{ for some }\boldsymbol{0}\neq\boldsymbol{v}\in\mathbb{R}^n\\ &\Leftrightarrow \lambda\boldsymbol{I}-\boldsymbol{A}\text{ is a singular matrix}\\ &\Leftrightarrow \det(\lambda\boldsymbol{I}-\boldsymbol{A})=0.\end{aligned}$$

- $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ is the **characteristic polynomial** of $\boldsymbol{A}$
- $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ is the **characteristic equation** of $\boldsymbol{A}$.

Remark: Suppose that $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ can be completely factorized: $(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$, $\lambda_1,\lambda_2,\ldots,\lambda_k$ are all distinct.

- r_i is the **algebraic multiplicity** $a(\lambda_i)$ of λ_i.
- $\dim(E_i)$ is the **geometric multiplicity** $g(\lambda_i)$ of λ_i.
- $g(\lambda_i)\leq a(\lambda_i)$
- $a(\lambda_1)+a(\lambda_2)+\cdots+a(\lambda_k)=n$
- $g(\lambda_i)<a(\lambda_i)$ for some $i=1,\ldots,k\Leftrightarrow\boldsymbol{A}$ is not diagonalizable.
- If $\boldsymbol{A}_{n\times n}$ has n distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable.

Special Case

$$\begin{aligned}0\text{ is an eigenvalue of }\boldsymbol{A} &\Leftrightarrow 0\text{ is a root to }\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0\\ &\Leftrightarrow \det(0\boldsymbol{I}-\boldsymbol{A})=0\\ &\Leftrightarrow (-1)^n\det(\boldsymbol{A})=0\\ &\Leftrightarrow \boldsymbol{A}\text{ is a singular matrix.}\end{aligned}$$

- 0 is not an eigenvalue of $\boldsymbol{A}\Leftrightarrow\boldsymbol{A}$ is an invertible matrix
- The **eigenspace** E_0 is the nullspace of $\boldsymbol{B}$.

<u>**Eigenspace**</u>: The eigenspace of $\boldsymbol{A}$ associated to λ is the nullspace of $(\lambda\boldsymbol{I}-\boldsymbol{A})$, denoted by $E_{\boldsymbol{A},\lambda}$ or simply E_λ.

For any nonzero vector $\boldsymbol{v}\in\mathbb{R}^n$

$$\begin{aligned}\boldsymbol{v}\text{ is an eigenvector of }\boldsymbol{A}\text{ associated to }\lambda &\Leftrightarrow \boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v}\\ &\Leftrightarrow (\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{v}=\boldsymbol{0}\\ &\Leftrightarrow \boldsymbol{v}\in\text{nullspace of }(\lambda\boldsymbol{I}-\boldsymbol{A})\end{aligned}$$

Note that $(\lambda\boldsymbol{I}-\boldsymbol{A})$ is singular, so $\dim\left(E_{\boldsymbol{A},\lambda}\right)\geq 1$.

Theorem: Let λ be an eigenvalue of a square matrix $\boldsymbol{A}$. The eigenvectors of $\boldsymbol{A}$ associated to the eigenvalue λ are precisely all nonzero vectors in the eigenspace $E_{\boldsymbol{A},\lambda}$.

Example The Fibonacci numbers a_n are defined by $a_0=0, a_1=1$ and $a_n=a_{n-1}+a_{n-2}$ for $n\geq 2$.

Note that $a_{n+1}=a_{n-1}+a_n$ for $n\geq 1$.

$$\begin{pmatrix}a_n\\ a_{n+1}\end{pmatrix}=\begin{pmatrix}a_n\\ a_{n-1}+a_n\end{pmatrix}=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}a_{n-1}\\ a_n\end{pmatrix}$$

Let $\boldsymbol{A}=\begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$ and $\boldsymbol{x}_n=\begin{pmatrix}a_n\\ a_{n+1}\end{pmatrix}$.

Then $\boldsymbol{x}_0=\begin{pmatrix}0\\ 1\end{pmatrix}$ and $\boldsymbol{x}_n=\boldsymbol{A}^n\boldsymbol{x}_0$

$$\boldsymbol{A}^n=\frac{1}{\sqrt{5}}\begin{pmatrix}\left(\frac{1+\sqrt{5}}{2}\right)^{n-1}-\left(\frac{1-\sqrt{5}}{2}\right)^{n-1} & \left(\frac{1+\sqrt{5}}{2}\right)^{n}-\left(\frac{1-\sqrt{5}}{2}\right)^{n}\\ \left(\frac{1+\sqrt{5}}{2}\right)^{n}-\left(\frac{1-\sqrt{5}}{2}\right)^{n} & \left(\frac{1+\sqrt{5}}{2}\right)^{n+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\end{pmatrix}$$

$$\boldsymbol{x}_n=\boldsymbol{A}^n\boldsymbol{x}_0=\frac{1}{\sqrt{5}}\begin{pmatrix}\left(\frac{1+\sqrt{5}}{2}\right)^{n}-\left(\frac{1-\sqrt{5}}{2}\right)^{n}\\ \left(\frac{1+\sqrt{5}}{2}\right)^{n+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\end{pmatrix}$$

Therefore, $a_n=\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n-\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$

Orthogonal Diagonalization A square matrix matrix $\boldsymbol{A}$ of order n is diagonalizable $\Leftrightarrow\boldsymbol{A}$ has n linearly independent eigenvectors $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n$ (associated to eigenvalues $\lambda_1,\ldots,\lambda_n$ respectively).

- $\boldsymbol{P}$ forms an orthonormal basis for $\mathbb{R}^n$
- If $\boldsymbol{P}$ is orthogonal, then $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$.
- $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}=\boldsymbol{D}^{\mathrm{T}}=\left(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\right)^{\mathrm{T}}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}$
- $\boldsymbol{A}$ is orthogonally diagonalizable $\Leftrightarrow\boldsymbol{A}$ is a symmetric matrix.

<u>**Singular Value Decomposition**</u>: Let $\boldsymbol{A}$ be an $m\times n$ matrix. Then $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is a symmetric matrix of order n.

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v}\Rightarrow\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}=\boldsymbol{v}^{\mathrm{T}}\lambda\boldsymbol{v}\Rightarrow\|\boldsymbol{A}\boldsymbol{v}\|^2=\lambda\|\boldsymbol{v}\|^2\geq 0$$

Theorem Let $\boldsymbol{A}$ be an $m\times n$ matrix. The eigenvalues of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ are **nonnegative**: $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n\geq 0$.

Singular Values: Let $\sigma_1=\sqrt{\lambda_1},\ldots,\sigma_n=\sqrt{\lambda_n}$. They are called the **singular values** of $\boldsymbol{A}$.

Suppose $\boldsymbol{V}=\begin{pmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{pmatrix}$ orthogonally diagonalizes $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$, such that

- if $i\leq r, \boldsymbol{v}_i$ is an eigenvector associated to λ_i
- if $i>r, \boldsymbol{v}_i$ is an eigenvector associated to 0

Let $\sigma_i=\sqrt{\lambda_i}$, and $\boldsymbol{u}_i=\frac{1}{\sigma_i}\boldsymbol{A}\boldsymbol{v}_i$ for $i\leq r$.

- $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r\}$ is an **orthonormal basis** for **column space** of $\boldsymbol{A}$

  For $\boldsymbol{v}\in\mathbb{R}^n$, $\boldsymbol{v}=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n$ for unique $c_1,\ldots,c_n\in\mathbb{R}$.

$$\begin{aligned}\boldsymbol{A}\boldsymbol{v}&=c_1\boldsymbol{A}\boldsymbol{v}_1+\cdots+c_n\boldsymbol{A}\boldsymbol{v}_n=c_1\boldsymbol{A}\boldsymbol{v}_1+\cdots+c_r\boldsymbol{A}\boldsymbol{v}_r\\ &=c_1\sigma_1\boldsymbol{u}_1+\cdots+c_r\sigma_r\boldsymbol{u}_r\in\operatorname{span}\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r\}\end{aligned}$$

- In particular, $\operatorname{rank}(\boldsymbol{A})=r$.

$$\boldsymbol{u}_i\cdot\boldsymbol{u}_j=\frac{1}{\sigma_i\sigma_j}\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_j=\frac{1}{\sigma_i\sigma_j}\boldsymbol{v}_i^{\mathrm{T}}\lambda_j\boldsymbol{v}_j=\frac{\lambda_j}{\sigma_i\sigma_j}\boldsymbol{v}_i\cdot\boldsymbol{v}_j=\begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j.\end{cases}$$

Algorithm Let $\boldsymbol{A}$ be an $m\times n$ matrix.

1. Find the eigenvalues $\lambda_1\geq\cdots\geq\lambda_r>\lambda_{r+1}=\cdots=\lambda_n=0$ of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.
2. Find corresponding orthonormal set $\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\}$ of eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.
3. Let $\sigma_i=\sqrt{\lambda_i}$ and $\boldsymbol{u}_i=\frac{1}{\sigma_i}\boldsymbol{A}\boldsymbol{v}_i$ for $i=1,\ldots,r$.
4. Extend $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r\}$ to an orthonormal basis $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m\}$ for $\mathbb{R}^m$.
5. Let $\boldsymbol{U}=\begin{pmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m\end{pmatrix}$ and $\boldsymbol{V}=\begin{pmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n\end{pmatrix}$. Then $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices.
6. Let $\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1 & \cdots & 0 & \cdots & 0\\ \vdots & \ddots & \vdots & & \vdots\\ 0 & \cdots & \sigma_r & \cdots & 0\\ \vdots & & \vdots & \ddots & \vdots\\ 0 & \cdots & 0 & \cdots & \end{pmatrix}$. It is a rectangular diagonal matrix.

- Then $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$ is called the singular value decomposition of $\boldsymbol{A}$.

Chapter 7 Linear Transformation

In this chapter, all vectors are viewed as column vectors.

Definition: We say the mapping $f:\mathbb{R}^n\to\mathbb{R}$ defined by
$$f(x_1,x_2,\dots,x_n)=a_1x_1+a_2x_2+\cdots+a_nx_n$$
a **linear transformation** from $\mathbb{R}^n$ to $\mathbb{R}$. It can be viewed in the **matrix form**:
$$f\left(\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\right)=\begin{pmatrix}a_1&a_2&\cdots&a_n\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$
Generally, the mapping $T:\mathbb{R}^n\to\mathbb{R}^m$ defined by
$$T\left(\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\right)=\begin{pmatrix}a_{11}x_1+a_{12}x_2+&\cdots&+a_{1n}x_n\\a_{21}x_1+a_{22}x_2+&\cdots&+a_{2n}x_n\\\vdots&\vdots&\vdots\\a_{m1}x_1+a_{m2}x_2+&\cdots&+a_{mn}x_n\end{pmatrix}$$
$$=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$
is a **linear transformation** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

- $T:\mathbb{R}^n\to\mathbb{R}^m$ such that $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$, for $\boldsymbol{x}\in\mathbb{R}^n$.
- $\boldsymbol{A}=(a_{ij})_{m\times n}$ is called the standard matrix for T.
- T is called a **linear operator** on $\mathbb{R}^n$ if $m=n$

Identity Operator: Let $\mathbb{R}^n\to\mathbb{R}^n$ be the linear operator s.t. $I(\boldsymbol{x})=\boldsymbol{x}$ for $\boldsymbol{x}\in\mathbb{R}^n$. It is called the **identity operator** on $\mathbb{R}^n$.

- $I(\boldsymbol{x})=\boldsymbol{x}=\boldsymbol{I}_n\boldsymbol{x}$; so $\boldsymbol{I}_n$ is the **standard matrix** for I.

Zero Transformation: Let $O:\mathbb{R}^n\to\mathbb{R}^m$ be the linear transformation s.t. $O(\boldsymbol{x})=\boldsymbol{0}$ for $\boldsymbol{x}\in\mathbb{R}^n$. It is called the **zero transformation** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

- $O(\boldsymbol{x})=\boldsymbol{0}=\boldsymbol{0}_{m\times n}\boldsymbol{0}$; so $\boldsymbol{0}_{m\times n}$ is the standard matrix for O.

Uniqueness: Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation, the standard matrix for a linear transformation is **unique**.

Remark: To show that a function $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, it suffices to find an $m\times n$ matrix $\boldsymbol{A}$ so that $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$ for all $\boldsymbol{x}\in\mathbb{R}^n$.

Linearity:

1. $T(\boldsymbol{0})=\mathbf{A0}=\boldsymbol{0}$.
2. $T(c\boldsymbol{v})=\boldsymbol{A}(c\boldsymbol{v})=c(\boldsymbol{A}\boldsymbol{v})=cT(\boldsymbol{v})$
3. $T(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{A}\boldsymbol{u}+\boldsymbol{A}\boldsymbol{v}=T(\boldsymbol{u})+T(\boldsymbol{v})$
4. For any $\boldsymbol{v}_1,\dots,\boldsymbol{v}_k\in\mathbb{R}^n$ and $c_1,\dots,c_k\in\mathbb{R}$,
$$\begin{aligned}T(c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k)&=\boldsymbol{A}(c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k)\\&=\boldsymbol{A}(c_1\boldsymbol{v}_1)+\cdots+\boldsymbol{A}(c_k\boldsymbol{v}_k)\\&=c_1(\boldsymbol{A}\boldsymbol{v}_1)+\cdots+c_k(\boldsymbol{A}\boldsymbol{v}_k)\\&=c_1T(\boldsymbol{v}_1)+\cdots+c_kT(\boldsymbol{v}_k)\end{aligned}$$
5. Let $E=\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_n\}$ be the standard basis for $\mathbb{R}^n$. Every $\boldsymbol{v}=c_1\boldsymbol{e}_1+\cdots+c_n\boldsymbol{e}_n=(c_1,\dots,c_n)\in\mathbb{R}^n$.
   Suppose $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation.
   $$T(\boldsymbol{v})=T(c_1\boldsymbol{e}_1+\cdots+c_n\boldsymbol{e}_n)=c_1T(\boldsymbol{e}_1)+\cdots+c_nT(\boldsymbol{e}_n)$$
   $T(\boldsymbol{v})$ is completely determined by $T(\boldsymbol{e}_1),\dots,T(\boldsymbol{e}_n)$.

Representation

For $T(\boldsymbol{e}_1)=\boldsymbol{A}\boldsymbol{e}_1,\dots,T(\boldsymbol{e}_n)=\boldsymbol{A}\boldsymbol{e}_n$, we have
$$\begin{aligned}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{I}&=\boldsymbol{A}\begin{pmatrix}\boldsymbol{e}_1&\cdots&\boldsymbol{e}_n\end{pmatrix}\\&=\begin{pmatrix}\boldsymbol{A}\boldsymbol{e}_1&\cdots&\boldsymbol{A}\boldsymbol{e}_n\end{pmatrix}\\&=\begin{pmatrix}T(\boldsymbol{e}_1)&\cdots&T(\boldsymbol{e}_n)\end{pmatrix}\end{aligned}$$
Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a function, T is a linear transformation

1. $\Leftrightarrow T(c_1\boldsymbol{v}_1+\cdots+c_k\boldsymbol{v}_k)=c_1T(\boldsymbol{v}_1)+\cdots+c_kT(\boldsymbol{v}_k)$
2. $\Leftrightarrow T(c\boldsymbol{v})=cT(\boldsymbol{v})$
3. $\Leftrightarrow T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$

To show that a mapping T is not a linear transformation.

1. Show that $T(\boldsymbol{0})\neq\boldsymbol{0}$;
2. Find $\boldsymbol{v}\in\mathbb{R}^n$ and $c\in\mathbb{R}$ such that $T(c\boldsymbol{v})\neq cT(\boldsymbol{v})$;
3. Find $\boldsymbol{u}\in\mathbb{R}^n$ and $\boldsymbol{v}\in\mathbb{R}^n$ such that $T(\boldsymbol{u}+\boldsymbol{v})\neq T(\boldsymbol{u})+T(\boldsymbol{v})$.

Transformation and Coordinate Vector:
$$\begin{aligned}T(\boldsymbol{v})&=T(c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n)\\&=c_1T(\boldsymbol{v}_1)+\cdots+c_nT(\boldsymbol{v}_n)\\&=\begin{pmatrix}T(\boldsymbol{v}_1)&\cdots&T(\boldsymbol{v}_n)\end{pmatrix}\begin{pmatrix}c_1\\\vdots\\c_n\end{pmatrix}\\&=\begin{pmatrix}T(\boldsymbol{v}_1)&\cdots&T(\boldsymbol{v}_n)\end{pmatrix}[\boldsymbol{v}]_S.\end{aligned}$$
Change of Bases: Let $S=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ be a basis for $\mathbb{R}^n$.

1. For $\boldsymbol{v}\in\mathbb{R}^n$, write $(\boldsymbol{v})_S=(c_1,\dots,c_n)$
   $$\boldsymbol{v}=c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n=\begin{pmatrix}\boldsymbol{v}_1&\cdots&\boldsymbol{v}_n\end{pmatrix}[\boldsymbol{v}]_S=\boldsymbol{P}[\boldsymbol{v}]_S$$
2. Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation
   $$\begin{aligned}T(\boldsymbol{v})&=c_1T(\boldsymbol{v}_1)+\cdots+c_nT(\boldsymbol{v}_n)\\&=\begin{pmatrix}T(\boldsymbol{v}_1)&\cdots&T(\boldsymbol{v}_n)\end{pmatrix}[\boldsymbol{v}]_S=\boldsymbol{B}[\boldsymbol{v}]_S\end{aligned}$$
3. Let $\boldsymbol{A}$ be the standard matrix for T
   $$\boldsymbol{B}[\boldsymbol{v}]_S=T(\boldsymbol{v})=\boldsymbol{A}\boldsymbol{v}=\boldsymbol{A}\boldsymbol{P}[\boldsymbol{v}]_S\Leftrightarrow\boldsymbol{A}=\boldsymbol{B}\boldsymbol{P}^{-1}$$

$S_1=\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_n\}$ is a basis for $\mathbb{R}^n$, $T(\boldsymbol{v})=\boldsymbol{B}[\boldsymbol{v}]_{S_1}$

$S_2=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ is a basis for $\mathbb{R}^n$, $T(\boldsymbol{v})=\boldsymbol{C}[\boldsymbol{v}]_{S_2}$

Relation between B **and** C: Let $\boldsymbol{P}$ be the transition matrix from S_1 to S_2.
$$\boldsymbol{P}[\boldsymbol{v}]_{S_1}=[\boldsymbol{v}]_{S_2}\Rightarrow\boldsymbol{C}\boldsymbol{P}[\boldsymbol{v}]_{S_1}=\boldsymbol{C}[\boldsymbol{v}]_{S_2}=T(\boldsymbol{v})=\boldsymbol{B}[\boldsymbol{v}]_{S_1}$$
Therefore, $\boldsymbol{B}=\boldsymbol{C}\boldsymbol{P}$.

Specifically, if $S_1=S$ is any basis and $S_2=E$ is the standard basis. $\boldsymbol{P}=\begin{pmatrix}\boldsymbol{u}_1&\cdots&\boldsymbol{u}_n\end{pmatrix}$ and $\boldsymbol{P}^{-1}$ is the transition matrix from E to S. $C=A$ is the standard matrix for T.

- $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{P}$
- $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{P}^{-1}$

Similarity: Let $T:\mathbb{R}^n\to\mathbb{R}^n$ be a linear operator on $\mathbb{R}^n$.

Let

- $\boldsymbol{A}$ be the standard matrix for T s.t. $T(\boldsymbol{v})=\boldsymbol{A}\boldsymbol{v}$
- $S=\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ be any basis for $\mathbb{R}^n$.
- $\boldsymbol{P}=\begin{pmatrix}\boldsymbol{v}_1&\cdots&\boldsymbol{v}_n\end{pmatrix}$. Then $\boldsymbol{P}$ is an invertible matrix.

Then

- $T(\boldsymbol{v})=\boldsymbol{P}[T(\boldsymbol{v})]_S$
- $\boldsymbol{A}\boldsymbol{v}=\boldsymbol{A}\boldsymbol{P}[\boldsymbol{v}]_S$
- $\boldsymbol{P}[T(\boldsymbol{v})]_S=\boldsymbol{A}\boldsymbol{P}[\boldsymbol{v}]_S\Rightarrow[T(\boldsymbol{v})]_S=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}[\boldsymbol{v}]_S$

T can be represented by $[\boldsymbol{v}]_S\mapsto\boldsymbol{B}[\boldsymbol{v}]_S$, where $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. We say $\boldsymbol{A}$ and $\boldsymbol{B}$ are **similar**.

- A square matrix is diagonalizable $\Leftrightarrow$ it is similar to a diagonal matrix.

Composition: $g\circ f(x)=g(f(x))$, $(T\circ S)(\boldsymbol{u})=T(S(\boldsymbol{u}))$

Properties

Let $S:\mathbb{R}^n\to\mathbb{R}^m$ and $T:\mathbb{R}^m\to\mathbb{R}^k$ be linear transformations.

- $\boldsymbol{A}$ be the standard matrix for S, s.t. $S(\boldsymbol{u})=\boldsymbol{A}\boldsymbol{u}$
- $\boldsymbol{B}$ be the standard matrix for T, s.t. $T(\boldsymbol{v})=\boldsymbol{B}\boldsymbol{v}$

For any $\boldsymbol{u}\in\mathbb{R}^n$,
$$(T\circ S)(\boldsymbol{u})=T(S(\boldsymbol{u}))=T(\boldsymbol{A}\boldsymbol{u})=\boldsymbol{B}(\boldsymbol{A}\boldsymbol{u})=(\boldsymbol{B}\boldsymbol{A})\boldsymbol{u}$$
$T\circ S:\mathbb{R}^n\to\mathbb{R}^k$ is a linear transformation with **standard matrix** $\boldsymbol{BA}$.

Range of Linear Transformation: Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation. The **range** of T is the set of all images of T: $\mathrm{R}(T)=\{T(\boldsymbol{v})\mid\boldsymbol{v}\in\mathbb{R}^n\}\subseteq\mathbb{R}^m$.

Representation of Range

$\because T(\boldsymbol{v})=c_1T(\boldsymbol{v}_1)+\cdots+c_nT(\boldsymbol{v}_n)\in\mathrm{span}\{T(\boldsymbol{v}_1),\dots,T(\boldsymbol{v}_n)\}$

$\therefore \mathrm{R}(T)=\{T(\boldsymbol{v})\mid\boldsymbol{v}\in\mathbb{R}^n\}\subseteq\mathrm{span}\{T(\boldsymbol{v}_1),\dots,T(\boldsymbol{v}_n)\}$.

Conversely,

$\because c_1T(\boldsymbol{v}_1)+\cdots+c_nT(\boldsymbol{v}_n)=T(c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n)\in\mathrm{R}(T)$

$\therefore \mathrm{span}\{T(\boldsymbol{v}_1),\dots,T(\boldsymbol{v}_n)\}\subseteq\mathrm{R}(T)$.

$\therefore \mathrm{R}(T)=\mathrm{span}\{T(\boldsymbol{v}_1),\dots,T(\boldsymbol{v}_n)\}$, where $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ is any basis for $\mathbb{R}^n$.

Note that T has standard matrix $\boldsymbol{A}=\begin{pmatrix}T(\boldsymbol{e}_1)&\cdots&T(\boldsymbol{e}_n)\end{pmatrix}$.

Then

1. $\mathrm{R}(T)=\mathrm{span}\{T(\boldsymbol{v}_1),\dots,T(\boldsymbol{v}_n)\}=$ (column space of $\boldsymbol{A}$)
2. $\mathrm{rank}(T)=\dim(\mathrm{R}(T))$
3. $\mathrm{rank}(T)=\mathrm{rank}(\boldsymbol{A})$

Kernel of Linear Transformation: The kernel of T is the set of all vectors in $\mathbb{R}^n$ whose image is $0\in\mathbb{R}^m$.
$$\mathrm{Ker}(T)=\{\boldsymbol{v}\in\mathbb{R}^n\mid T(\boldsymbol{v})=\boldsymbol{0}\}\subseteq\mathbb{R}^n$$
Recall that $T(\boldsymbol{0})=\boldsymbol{0}$. $\mathrm{Ker}(T)$ contains $\boldsymbol{0}\in\mathbb{R}^n$.

Representation of Kernel: Let $T_{n\to m}(\boldsymbol{v})=\boldsymbol{A}\boldsymbol{v}$ for all $\boldsymbol{v}\in\mathbb{R}^n$
$$\begin{aligned}\mathrm{Ker}(T)&=\{\boldsymbol{v}\in\mathbb{R}^n\mid T(\boldsymbol{v})=\boldsymbol{0}\}\\&=\{\boldsymbol{v}\in\mathbb{R}^n\mid\boldsymbol{A}\boldsymbol{v}=\boldsymbol{0}\}=(\text{ nullspace of }\boldsymbol{A}).\end{aligned}$$

- $\mathrm{Ker}(T)=($ nullspace of $\boldsymbol{A})$
- $\mathrm{nullity}(T)=\dim(\mathrm{Ker}(T))$
- $\mathrm{nullity}(T)=\mathrm{nullity}(\boldsymbol{A})$

Property: Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation.
$$\mathrm{rank}(T)+\mathrm{nullity}(T)=n$$