

Matrices

Theorem 1.2.7. If **augmented matrices** of two systems of linear equations are row equivalent, then the two systems have the same set of solutions. (* Even for two homogeneous linear systems, we still need to say that $(\mathbf{A} \mid \mathbf{0})$ is row equivalent to $(\mathbf{B} \mid \mathbf{0})$, not that \mathbf{A} is row equivalent to \mathbf{B} .)

Example 1.4.10. Suppose augmented matrix \mathbf{R} is in (R)REF:

1. LS has no solution
 \Leftrightarrow Last column of \mathbf{R} is pivot.
2. LS has one unique solution
 \Leftrightarrow **Only** last column of \mathbf{R} is non-pivot.
3. LS has infinite number of solution
 \Leftrightarrow At least one column other than the last one is non-pivot
 \Leftrightarrow Number of variables $>$ Number of non-zero rows in \mathbf{R}
(* # non-pivot columns in (R)REF $-1 = \#$ unique solutions)

Theorem 6.1.8. \mathbf{A} is invertible when:

1. $\exists \mathbf{B}$ s.t. $\mathbf{AB} = \mathbf{I} \vee \mathbf{BA} = \mathbf{I}$
2. Refer to **Theorem 2.4.7.2** below
3. $\text{rref}(\mathbf{A}) = \mathbf{I}$
4. \mathbf{A} is a product of elementary matrices
5. $\det(\mathbf{A}) \neq 0$
6. Rows of \mathbf{A} is a basis of \mathbb{R}^n
7. Columns of \mathbf{A} is a basis of \mathbb{R}^n
8. 0 is not an eigenvalue of \mathbf{A}

Remark 2.3.4 (Cancellation Laws for Matrices). Let \mathbf{A} be an invertible $m \times m$ matrix,

- (a) If \mathbf{B}_1 and \mathbf{B}_2 are $m \times n$ matrices with $\mathbf{AB}_1 = \mathbf{AB}_2$, then $\mathbf{B}_1 = \mathbf{B}_2$
- (b) If \mathbf{C}_1 and \mathbf{C}_2 are $n \times m$ matrices with $\mathbf{C}_1\mathbf{A} = \mathbf{C}_2\mathbf{A}$, then $\mathbf{C}_1 = \mathbf{C}_2$

Theorem 2.4.7.2 (generalised). Relationship between singularity of \mathbf{A} and the number of solutions of a linear system $\mathbf{Ax} = \mathbf{b}$:

1. \mathbf{A} is singular $\Leftrightarrow \mathbf{Ax} = \mathbf{b}$: has ∞ solutions (only case for homogeneous LS) or no solutions
2. \mathbf{A} is invertible $\Leftrightarrow \mathbf{Ax} = \mathbf{b}$: has one unique solution (trivial solution for homogeneous LS)

Definition 2.5.2. Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix. Let \mathbf{M}_{ij} be an $(n-1) \times (n-1)$ matrix obtained from \mathbf{A} by deleting the i th row and the j th column. Then the *determinant* of \mathbf{A} is defined as

$$\det(\mathbf{A}) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + \dots + a_{1n}A_{1n} & \text{if } n > 1 \end{cases}$$

where

$$A_{ij} = (-1)^{i+j} \det(\mathbf{M}_{ij})$$

The number A_{ij} is called the (i, j) -*cofactor* of \mathbf{A} .

Theorem 2.5.8. The determinant of a triangular matrix is equal to the product of its diagonal entries.

Theorem 2.5.12 (added-on). The determinant of a square matrix is 0 when:

1. it has two identical rows, or
2. it has two identical columns
3. any row/column of its (R)REF is zero

Theorem 2.5.15. Let \mathbf{A} be a square matrix. k is a non-zero constant.

1. $\mathbf{A} \xrightarrow{k\mathbf{R}_i} \mathbf{B} \Rightarrow \det(\mathbf{B}) = k \det(\mathbf{A})$
2. $\mathbf{A} \xrightarrow{\mathbf{R}_i \leftrightarrow \mathbf{R}_j} \mathbf{B} \Rightarrow \det(\mathbf{B}) = -\det(\mathbf{A})$
3. $\mathbf{A} \xrightarrow{\mathbf{R}_i + k\mathbf{R}_j} \mathbf{B} \Rightarrow \det(\mathbf{B}) = \det(\mathbf{A})$
4. Let \mathbf{E} be an elementary matrix of the same size as \mathbf{A} . Then $\det(\mathbf{EA}) = \det(\mathbf{E}) \det(\mathbf{A})$.

Remark 2.5.18. Since $\det(\mathbf{A}) = \det(\mathbf{A}^T)$, theorem 2.5.15 holds if “rows” are changed to “columns”.

Theorem 2.5.22. Let \mathbf{A} and \mathbf{B} are two square matrices of order n and c is a scalar. Then

1. $\det(c\mathbf{A}) = c^n \det(\mathbf{A})$
2. $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$
3. if \mathbf{A} is invertible, $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$

Definition 2.5.24. Let \mathbf{A} be a square matrix of order n . Then the (*classical*) *adjoint* of \mathbf{A} is the $n \times n$ matrix

$$\text{adj}(\mathbf{A}) = (A_{ij})_{n \times n}^T$$

where A_{ij} is the (i, j) -cofactor of \mathbf{A} .

Theorem 2.5.27 (Cramer's Rule). Suppose $\mathbf{Ax} = \mathbf{b}$ is a linear system where \mathbf{A} is an $n \times n$ matrix. Let \mathbf{A}_i be the matrix obtained from \mathbf{A} by replacing the i th column of \mathbf{A} by \mathbf{b} . If \mathbf{A} is invertible, then the system has only one solution

$$\mathbf{x} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} \det(\mathbf{A}_1) \\ \vdots \\ \det(\mathbf{A}_n) \end{pmatrix}$$

Mixed Notes 1. \mathbf{A}^{-1} is able to be computed by:

1. Find \mathbf{B} s.t. $\mathbf{AB} = \mathbf{I} \vee \mathbf{BA} = \mathbf{I}$
2. Find using **Theorem 2.5.25**: $\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A})$
3. Find using: $(\mathbf{A} \mid \mathbf{I}) \xrightarrow{GJE} (\mathbf{I} \mid \mathbf{A}^{-1})$

Mixed Notes 2. $\det(\mathbf{A})$ is able to be computed by:

1. Using **Theorem 2.5.2**
2. Using cross multiplication (for 2×2 and 3×3 matrices only)
3. Doing some ERO (e.g. GE, consider **Theorem 2.5.15**) and making it triangular then using **Theorem 2.5.8** or making it have properties in **Theorem 2.5.12**
4. Using **Theorem 2.5.22**

Mixed Notes 3. Some random notes:

1. In \mathbb{R}^n where $n \geq 2$, a set with 1 parameter is a line and that with 2 parameters is a space.
2. $\mathbf{M}^2 + \mathbf{M} = \mathbf{I} \Rightarrow \mathbf{M}(\mathbf{M} + \mathbf{I}) = \mathbf{I}$ (Don't put that \mathbf{I} to be scalar 1!)
3. Two matrices have same RREF \Leftrightarrow They are row equivalent
4. In exam, express a matrix in the form $\mathbf{A} = (a_{ij})_{m \times n}$. **DO NOT** use dots form
5. When using ERO $\mathbf{R}_i = \frac{1}{k}\mathbf{R}_j$, discuss whether k is 0 when necessary

Mixed Notes 4. Generally, for (square) matrices \mathbf{A} and \mathbf{B} ,

1. $\mathbf{AB} \neq \mathbf{BA}$
2. $(\mathbf{AB})^2 \neq \mathbf{A}^2\mathbf{B}^2$
3. $\mathbf{AB} = \mathbf{0} \nRightarrow \mathbf{A} = \mathbf{0} \vee \mathbf{B} = \mathbf{0}$
4. $\mathbf{A}^2 = \mathbf{I} \nRightarrow \mathbf{A} = \pm \mathbf{I}$ (For example: 2 EMs of 2nd type ERO)

Mixed Notes 5. When expanding a row/column with cofactors of the other row/column, 0 will be yielded:

$$\sum_{m=1}^n a_{im}A_{jm} = \sum_{m=1}^n a_{mi}A_{mj} = 0, \text{ for some } i \neq j$$

Euclidean Spaces

Discussion 3.2.5. Given $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$, show $\text{span}(S) = \mathbb{R}^n$:

Consider $\mathbf{v}_i = (v_{i1}, \dots, v_{in})$,

$$\begin{pmatrix} \mathbf{v}_{11} & \dots & \mathbf{v}_{m1} \\ \vdots & \ddots & \vdots \\ \mathbf{v}_{1n} & \dots & \mathbf{v}_{mn} \end{pmatrix} \xrightarrow{GE} \mathbf{R}$$

$\text{span}(S) = \mathbb{R}^n \Leftrightarrow \mathbf{R}$ has no zero rows

Theorem 3.2.7. If $|S| < n$, $\text{span}(S) \neq \mathbb{R}^n$.

Theorem 3.2.10. Let $S_1 = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $S_2 = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ be subsets of \mathbb{R}^n . Then, $\text{span}(S_1) \subseteq \text{span}(S_2) \Leftrightarrow \forall i = 1, 2, \dots, k, \mathbf{u}_i \in \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$.

Definition 3.3.2. Let V be a subset of \mathbb{R}^n . Then V is called a *subspace* of \mathbb{R}^n if $V = \text{span}(S)$ where $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ for some vectors $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{R}^n$.

More precisely, V is called the *subspace spanned* by S (or the *subspace spanned* by $\mathbf{u}_1, \dots, \mathbf{u}_k$). We also say that S *spans* (or $\mathbf{u}_1, \dots, \mathbf{u}_k$ *span*) the subspace V .

By contraposition, $V = \text{span}(S) \Rightarrow \mathbf{0} \in V \equiv \mathbf{0} \notin V \Rightarrow V \neq \text{span}(S)$. (* i.e., If $\mathbf{0}$ is not in V , V is not a subspace of \mathbb{R}^n)

Theorem 3.3.6. If $V = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{0}\}$, V is a subspace of \mathbb{R}^n .

Remark 3.3.8. Let V be a non-empty subset of \mathbb{R}^n . Then V is a subspace of \mathbb{R}^n if and only if

for all $\mathbf{u}, \mathbf{v} \in V$ and $c, d \in \mathbb{R}$, $c\mathbf{u} + d\mathbf{v} \in V$

(* This checks whether V is **closed** under addition and scalar multiplication)

Definition 3.4.2/4. Consider $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ which are column vectors, set $S = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ is **Linear Independent** iff. any of:

- 1. $(\mathbf{u}_1\mathbf{u}_2\dots\mathbf{u}_k)\mathbf{x} = \mathbf{0}$ has only trivial solution.
- 2. No vectors in S can be written as a linear combination of other vectors in S .
- 3. S is a subset of a **Linear Independent** set.

Definition 3.5.4/Theorem 3.6.7. A set S is a basis of a vector space if:

- 1. $S \subseteq V$
- 2. Any 2 of the 3 below:
 - 2.1. S is Linear Independent
 - 2.2. S spans V
 - 2.3. $|S| = \dim(V)$

Definition 3.5.8. Let $S = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ be a basis for a vector space V and \mathbf{v} is a vector in V . By T3.5.7, \mathbf{v} is expressed uniquely as a LC:

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \dots + c_k\mathbf{u}_k$$

Then we shall have the **coordinate vector** of \mathbf{v} relative to the basis S : $(\mathbf{v})_S = (c_1, c_2, \dots, c_k) \in \mathbb{R}^k$ (assuming vectors in S are in fixed order).

Remark 3.5.10/Theorem 3.5.11. Let S be a basis for a vector space V ,

- 1. $\forall \mathbf{u}, \mathbf{v} \in V, \mathbf{u} = \mathbf{v} \Leftrightarrow (\mathbf{u})_S = (\mathbf{v})_S$
- 2. Coordinate vectors are closed under scalar multiplication and addition
- 3. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r \in V$, they are LI iff. $(\mathbf{v}_1)_S, (\mathbf{v}_2)_S, \dots, (\mathbf{v}_k)_S$ are LI
- 4. $\text{span } \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r = V \Leftrightarrow \text{span } (\mathbf{v}_1)_S, (\mathbf{v}_2)_S, \dots, (\mathbf{v}_k)_S = \mathbb{R}^{|S|}$

Theorem 3.6.9. Let U be a subspace of V , then $\dim(U) \leq \dim(V)$. Furthermore, if $U \neq V$, then $\dim(U) < \dim(V)$.

Definition 3.7.3. Let $S = \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ and T be two bases for a vector space. The square matrix $\mathbf{P} = ([\mathbf{u}_1]_T \quad [\mathbf{u}_2]_T \quad \dots \quad [\mathbf{u}_k]_T)$ is called the **transition matrix** from S to T .

Mixed Theorem 6. Consider S and T are two bases for vector space V and \mathbf{P} is the transition matrix from S to T . If \mathbf{A} and \mathbf{B} are matrices with elements of S and T respectively as columns, we have $\mathbf{BP} = \mathbf{A}$.

Mixed Theorem 7. ERO preserves row space (T4.1.7), and we have:

- (R4.1.9) \mathbf{R} is RREF of \mathbf{A} . Non-empty rows in \mathbf{R} forms the basis of row space of \mathbf{A} .
- (T4.2.1) Row space and column space of a matrix have the same dimension.

Remark 4.2.5. Regarding $\text{rank}(\mathbf{A})$:

- 1. For $m * n$ matrix \mathbf{A} , $\text{rank}(\mathbf{A}) \leq \min m, n$. If $\text{rank}(\mathbf{A}) = \min m, n$, \mathbf{A} is said to have **full rank**.
- 2. A square matrix \mathbf{A} have full rank iff. it is invertible.
- 3. $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$.

Theorem 4.3.6. Suppose linear system $\mathbf{Ax} = \mathbf{b}$ has solution \mathbf{v} , then the solution set of this system is given by:

$$M = \{\mathbf{u} + \mathbf{v} | \mathbf{u} \in \text{nullspace}(\mathbf{A})\}$$

Orthogonality

Definition 5.1.2.3/4. For two vectors \mathbf{u} and \mathbf{v} :

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Angle between \mathbf{u} and \mathbf{v} is:

$$\cos^{-1}(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|})$$

Theorem 5.2.4. If S is an orthogonal set of non-zero vectors in a vector space, S is **LI**.

Theorem 5.2.8. Consider $S = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is a basis for a vector space V , then for any vector \mathbf{w} in V :

- 1. If S is orthogonal, we have
$$(\mathbf{w})_S = (\frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1, \frac{\mathbf{w} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2, \dots, \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k)$$
- 2. If S is orthonomal, we have
$$(\mathbf{w})_S = (\mathbf{w} \cdot \mathbf{u}_1, \mathbf{w} \cdot \mathbf{u}_2, \dots, \mathbf{w} \cdot \mathbf{u}_k)$$

T5.2.15: $(\mathbf{w})_S$ is the projection of \mathbf{w} onto V if $\mathbf{w} \in \mathbb{R}^n \wedge V$ is a subspace of \mathbb{R}^n (condition of \mathbf{w} changed but same formula applies).

Theorem 5.2.19 (Gram-Schmidt Process). Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ be a basis for a vector space V . Let

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1, \\ \mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1, \\ \mathbf{u}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \mathbf{v}_2, \\ &\vdots \end{aligned}$$

Then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is an orthogonal basis for V . Normalize all vectors in it then we have a orthonormal basis for V .

Definition 5.3.6. Let $\mathbf{Ax} = \mathbf{b}$ be a linear system where \mathbf{A} is an $m * n$ matrix. A vector $\mathbf{u} \in \mathbb{R}^n$ is called a **least squares solution** to the linear system if $\forall \mathbf{u} \in \mathbb{R}^n, \|\mathbf{b} - \mathbf{Au}\| \leq \|\mathbf{b} - \mathbf{Av}\|$.

Theorem 5.3.8. Continuing D5.3.6, let \mathbf{p} be the projection of \mathbf{b} onto the column space of \mathbf{A} . \mathbf{u} is the least squares solution iff. $\mathbf{Au} = \mathbf{p}$.

Theorem 5.3.10. Continuing D5.3.6, \mathbf{u} is the least squares solution iff. \mathbf{u} is a solution to $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$.

D5.4.3/R5.4.4/T5.4.6. \mathbf{A} is a square matrix of order n . The following are equivalent:

- 1. \mathbf{A} is orthogonal
- 2. $\mathbf{A}^{-1} = \mathbf{A}^T$
- 3. $\mathbf{AA}^T = \mathbf{A}^T \mathbf{A} = \mathbf{I}$
- 4. The rows of \mathbf{A} form an **orthonormal** basis for \mathbb{R}^n
- 5. The columns of \mathbf{A} form an **orthonormal** basis for \mathbb{R}^n

Theorem 5.4.7. Let S and T be two **orthonormal** bases for a vector space and let \mathbf{P} be the transition matrix from S to T . Then \mathbf{P} is orthogonal and \mathbf{P}^T is the transition matrix from T to S .

Diagonalization

Definition 6.1.3. \mathbf{A} is a square matrix of order n . $\mathbf{u} \in \mathbb{R}^n$ is a non-zero column vector that satisfies:

$$\mathbf{Au} = \lambda \mathbf{u}$$

for some scalar λ . λ is called an **eigenvalue** of \mathbf{A} . \mathbf{u} is said to be an **eigenvector** of \mathbf{A} associated with the eigenvalue λ .

Theorem 6.1.9. If \mathbf{A} is triangular, the eigenvalues of \mathbf{A} are the diagonal entries of \mathbf{A} .

Remark 6.2.5. Suppose the characteristic polynomial of the matrix \mathbf{A} can be factorized as

$$\det(\lambda \mathbf{I} - \mathbf{A}) = (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \dots (\lambda - \lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct eigenvalues of \mathbf{A} . Then for each eigenvalue λ_i ,

$$\dim(E_{\lambda_i}) \leq r_i$$

Furthermore, \mathbf{A} is diagonalizable iff. $\forall 1 \leq i \leq k, \dim(E_{\lambda_i}) = r_i$.

Definition 6.3.2/T*.4. A square matrix \mathbf{A} is said to be orthogonally diagonalizable iff. there exists an orthogonal matrix \mathbf{P} such that $\mathbf{P}^T \mathbf{AP}$ is diagonal.

A square matrix is orthogonally diagonalizable iff. it is **symmetric**.

Algorithm 6.3.5. Similar to the process for the normal matrix, orthogonal matrix \mathbf{P} can be found by using vectors of T as **its columns** where $T = T_{\lambda_1} \cup T_{\lambda_2} \cup \dots \cup T_{\lambda_k}$ and T_{λ_i} is transformed from S_{λ_1} using Gram-Schmidt Process.

Linear Transformation

Theorem 7.1.4. Let T be a linear transformation, we have:

- 1. $T(\mathbf{0}) = \mathbf{0}$
- 2. T is closed under scalar multiplication and addition

Discussion 7.1.8. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation with the standard matrix \mathbf{A} . Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be the standard basis for \mathbb{R}^n . We then have:

$$\mathbf{A} = (T(\mathbf{e}_1) \quad T(\mathbf{e}_2) \quad \dots \quad T(\mathbf{e}_n))$$

Theorem 7.2.4. Continuing D7.1.8. We have:

$$R(T) = \text{span}\{T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)\} = \text{the column space of } \mathbf{A}$$

which is a subspace of \mathbb{R}^m

D7.2.5/T7.2.9/D7.2.10/T7.2.12. Continuing T7.2.4. We have:

- $\text{rank}(T) = \dim(R(T)) = \text{rank}(\mathbf{A})$
- $\text{nullity}(T) = \text{nullity}(\mathbf{A})$
- $\text{rank}(T) + \text{nullity}(T) = n$
- $\ker(T) = \text{the nullspace of } \mathbf{A}$