

Matrices

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

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

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 2.5.15. Let \mathbf{A} be a square matrix.

1. If \mathbf{B} is obtained from \mathbf{A} by multiplying one row of \mathbf{A} by a constant k , then $\det(\mathbf{B}) = k \det(\mathbf{A})$.
2. If \mathbf{B} is obtained from \mathbf{A} by interchanging two rows, then $\det(\mathbf{B}) = -\det(\mathbf{A})$.
3. If \mathbf{B} is obtained from \mathbf{A} by adding a multiple of one row of \mathbf{A} to another row, then $\det(\mathbf{B}) = \det(\mathbf{A})$.
4. Let \mathbf{E} be an elementary matrix of the same size as \mathbf{A} . Then $\det(\mathbf{EA}) = \det(\mathbf{E}) \det(\mathbf{A})$.

Theorem 2.5.25. If \mathbf{A} is invertible, then $\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \text{adj}(\mathbf{A})$.

Theorem 2.5.27. 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

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

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} .

Euclidean Spaces

Definition 3.2.3. Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n . Then the set of all linear combinations of $\mathbf{u}_1, \dots, \mathbf{u}_k$,

$$\{c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k \mid c_1, \dots, c_k \in \mathbb{R}\}$$

is called the *linear span* of S (or the *linear span* of $\mathbf{u}_1, \dots, \mathbf{u}_k$) and is denoted by $\text{span}(S)$ (or $\text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$).

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 $\text{span}(S_1) \subseteq \text{span}(S_2)$ if and only if each \mathbf{u}_i is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_m$.

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 .

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

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

Definition 3.4.2. Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a set of vectors in \mathbb{R}^n . Consider the equation

$$c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$$

where c_1, \dots, c_k are variables.

1. S is called a *linearly dependent set* and $\mathbf{u}_1, \dots, \mathbf{u}_k$ are said to be *linearly dependent* if the equation has only the trivial solution $c_1 = \cdots = c_k = 0$.
2. S is called a *linearly independent set* and $\mathbf{u}_1, \dots, \mathbf{u}_k$ are said to be *linearly independent* if the equation has non-trivial solutions.

Definition 3.5.4. Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a subset of a vector space V . Then S is called a *basis* for V if S is linearly independent and S spans V .

Definition 3.5.8. Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis for a vector space V and \mathbf{v} a vector in V . By Theorem 3.5.7, \mathbf{v} is expressed uniquely as a linear combination

$$\mathbf{v} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$$

The coefficients c_1, \dots, c_k are called the *coordinates* of \mathbf{v} relative to the basis S .

The vector $(\mathbf{v})_S = (c_1, \dots, c_k) \in \mathbb{R}^k$ is called the *coordinate vector* of \mathbf{v} relative to the basis S .

Theorem 3.6.1. Let V be a vector space which has a basis with k vectors. Then

1. any subset of V with more than k vectors is always linearly dependent;
2. any subset of V with less than k vectors cannot span V .

Definition 3.6.3. The *dimension* of a vector space V , denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V . In addition, we define the dimension of the zero space to be zero.

Theorem 3.6.7. Let V be a vector space of dimension k and S a subset of V . The following are equivalent:

1. S is a basis for V .
2. S is linearly independent and $|S| = k$.
3. S spans V and $|S| = k$.

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

Vector Space of Matrices

Definition 4.1.2. Let $\mathbf{A} = (a_{ij})$ be an $m \times n$ matrix. The *row space* of \mathbf{A} is the subspace of \mathbb{R}^n spanned by the rows of \mathbf{A} . The *column space* of \mathbf{A} is the subspace of \mathbb{R}^m spanned by the columns of \mathbf{A} .

Theorem 4.1.7. Let \mathbf{A} and \mathbf{B} be row equivalent matrices. Then the row space of \mathbf{A} and the row space of \mathbf{B} are identical, i.e. elementary row operations preserve the row space of a matrix.

Theorem 4.1.11. Let \mathbf{A} and \mathbf{B} be row equivalent matrices. Then the following statements hold:

1. A given set of columns of \mathbf{A} is linearly independent if and only if the set of corresponding columns of \mathbf{B} is linearly independent.
2. A given set of columns of \mathbf{A} forms a basis for the column space of \mathbf{A} if and only if the set of corresponding columns of \mathbf{B} forms a basis for the column space of \mathbf{B} .

Theorem 4.2.1. The row space and column space of a matrix have the same dimension.

Definition 4.2.3. The *rank* of a matrix is the dimension of its row space (or column space). We denote the rank of a matrix \mathbf{A} by $\text{rank}(\mathbf{A})$. Note that $\text{rank}(\mathbf{A})$ is equal to the number of nonzero rows as well as the number of pivot columns in a row-echelon form of \mathbf{A} .

Theorem 4.2.8. Let \mathbf{A} and \mathbf{B} be $m \times n$ and $n \times p$ matrices respectively. Then

$$\text{rank}(\mathbf{AB}) \leq \min\{\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})\}$$

Definition 4.3.1. Let \mathbf{A} be an $m \times n$ matrix. The solution space of the homogeneous system of linear equations $\mathbf{Ax} = \mathbf{0}$ is known as the *nullspace* of \mathbf{A} .

The dimension of the null space of a matrix \mathbf{A} is known as the *nullity* of \mathbf{A} and is denoted by $\text{nullity}(\mathbf{A})$. If \mathbf{A} is an $m \times n$ matrix, it is clear that $\text{nullity}(\mathbf{A}) \leq n$ since the nullspace is a subspace of \mathbb{R}^n .

Theorem 4.3.4. Let \mathbf{A} be a matrix with n columns. Then

$$\text{rank}(\mathbf{A}) + \text{nullity}(\mathbf{A}) = n$$

Theorem 4.3.6. Suppose the linear equations $\mathbf{Ax} = \mathbf{b}$ has a solution \mathbf{v} . Then the solution set of the system is given by

$$\mathbf{M} = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \text{ is an element of the nullspace of } \mathbf{A}\}$$

Orthogonality

Definition 5.2.1.

- Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are called *orthogonal* if $\mathbf{u} \cdot \mathbf{v} = 0$.
- A set S of vectors in \mathbb{R}^n is called *orthogonal* if every pair of distinct vectors in S are orthogonal.
- A set S of vectors in \mathbb{R}^n is called *orthonormal* if S is orthogonal and every vector in S is a unit vector.

Definition 5.2.4.

- A basis S for a vector space is called an *orthogonal basis* if S is orthogonal.
- A basis S for a vector space is called an *orthonormal basis* if S is orthonormal.

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

$$\mathbf{w} = \frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k$$

i.e. $(\mathbf{w})_S = \left(\frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}, \dots, \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \right)$

Definition 5.2.10. Let V be a subspace of \mathbb{R}^n . A vector $\mathbf{u} \in \mathbb{R}^n$ is said to be *orthogonal* (or *perpendicular*) to V if \mathbf{u} is orthogonal to all vectors in V .

Definition 5.2.13. Let V be a subspace of \mathbb{R}^n . Every vector $\mathbf{u} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{u} = \mathbf{n} + \mathbf{p}$$

such that \mathbf{n} is a vector orthogonal to V and \mathbf{p} is a vector in V . The vector \mathbf{p} is called the (*orthogonal*) *projection* of \mathbf{u} onto V .

Theorem 5.2.15. Let V be a subspace of \mathbb{R}^n and \mathbf{w} a vector in \mathbb{R}^n . If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is an orthogonal basis for V , then

$$\frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k$$

is the projection of \mathbf{w} onto V .

Theorem 5.2.19. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis for a vector space V . Let

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1 \\ \mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 \\ &\vdots \\ \mathbf{v}_k &= \mathbf{u}_k - \frac{\mathbf{u}_k \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 - \dots - \frac{\mathbf{u}_k \cdot \mathbf{v}_{k-1}}{\mathbf{v}_{k-1} \cdot \mathbf{v}_{k-1}} \mathbf{v}_{k-1} \end{aligned}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is an orthogonal basis for V . Furthermore, let

$$\mathbf{w}_1 = \frac{1}{\|\mathbf{v}_1\|} \mathbf{v}_1, \dots, \mathbf{w}_k = \frac{1}{\|\mathbf{v}_k\|} \mathbf{v}_k$$

Then $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ is an orthonormal basis for V .

Theorem 5.3.10. Let $\mathbf{Ax} = \mathbf{b}$ be a linear system. Then \mathbf{u} is a least squares solution to $\mathbf{Ax} = \mathbf{b}$ if and only if \mathbf{u} is a solution to $\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$.

Definition 5.4.3. A square matrix \mathbf{A} is called *orthogonal* if $\mathbf{A}^{-1} = \mathbf{A}^T$.

Eigens and Diagonalization

Definition 6.1.3. Let \mathbf{A} be a square matrix of order n . A nonzero column vector \mathbf{u} in \mathbb{R}^n is called an *eigenvector* of \mathbf{A} if

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

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

Definition 6.1.6. Let \mathbf{A} be a square matrix of order n . The equation

$$\det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

is called the *characteristic equation* of \mathbf{A} and the polynomial

$$\det(\lambda \mathbf{I} - \mathbf{A})$$

is called the *characteristic polynomial* of \mathbf{A} .

Theorem 6.1.8. Let \mathbf{A} be an $n \times n$ matrix. The following statements are equivalent:

- \mathbf{A} is invertible.
- The linear system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.
- The reduced row-echelon form of \mathbf{A} is an identity matrix.
- \mathbf{A} can be expressed as a product of elementary matrices.
- $\det(\mathbf{A}) \neq 0$
- The rows of \mathbf{A} form a basis for \mathbb{R}^n .
- The columns of \mathbf{A} form a basis for \mathbb{R}^n .
- $\text{rank}(\mathbf{A}) = n$
- 0 is not an eigenvalue of \mathbf{A} .

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

Definition 6.1.11. Let \mathbf{A} be a square matrix of order n and λ an eigenvalue of \mathbf{A} . Then the solution space of the linear system $(\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ is called the *eigenspace* of \mathbf{A} *associated* with the eigenvalue λ and is denoted by E_λ .

Note that if \mathbf{u} is a nonzero vector in E_λ , then \mathbf{u} is an eigenvector of \mathbf{A} associated with the eigenvalue λ .

Definition 6.2.1. A square matrix \mathbf{A} is called *diagonalizable* if there exists an invertible matrix \mathbf{P} such that $\mathbf{P}^{-1} \mathbf{AP}$ is a diagonal matrix. Here the matrix \mathbf{P} is said to *diagonalize* \mathbf{A} .

Theorem 6.2.3. Let \mathbf{A} be a square matrix of order n . Then \mathbf{A} is diagonalizable if and only if \mathbf{A} has n linearly independent eigenvectors.

Remark 6.2.5.2. The dimension of an eigenspace E_λ of a square matrix \mathbf{A} associated with the eigenvalue λ is at most the multiplicity of λ in the characteristic polynomial of \mathbf{A} .

Furthermore, \mathbf{A} is diagonalizable if and only if the dimension of each eigenspace of \mathbf{A} is equal to the multiplicity of its associated eigenvalue.

Theorem 6.2.7. Let \mathbf{A} be a square matrix of order n . If \mathbf{A} has n distinct eigenvalues, then \mathbf{A} is diagonalizable.

Definition 6.3.2. A square matrix \mathbf{A} is called *orthogonally diagonalizable* if there exists an orthogonal matrix \mathbf{P} such that $\mathbf{P}^T \mathbf{AP}$ is a diagonal matrix. Here the matrix \mathbf{P} is said to *orthogonally diagonalize* \mathbf{A} .

Theorem 6.3.4. A square matrix is orthogonally diagonalizable if and only if it is symmetric.

Linear Transformations

Definition 7.1.1. A *linear transformation* is a mapping $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ of the form

$$T \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} \text{ for } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

where a_{11}, \dots, a_{mn} are real numbers. In particular, if $n = m$, T is also called a *linear operator* on \mathbb{R}^n . We can rewrite the formula of T as

$$T \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

The matrix $(a_{ij})_{m \times n}$ above is called the *standard matrix* for T .

Definition 7.1.10. Let $S : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $T : \mathbb{R}^m \rightarrow \mathbb{R}^k$ be linear transformations. The *composition* of T with S , denoted by $T \circ S$, is a mapping from $\mathbb{R}^n \rightarrow \mathbb{R}^k$ such that

$$(T \circ S)(\mathbf{u}) = T(S(\mathbf{u})) \text{ for } \mathbf{u} \in \mathbb{R}^n$$

Definition 7.2.1. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. The *range* of T , denoted by $R(T)$, is the set of images of T , i.e.

$$R(T) = \{T(\mathbf{u}) \mid \mathbf{u} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Theorem 7.2.4. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation and \mathbf{A} the standard matrix for T . Then

$$R(T) = \text{the column space of } \mathbf{A}$$

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

Definition 7.2.5. Let T be a linear transformation. The dimension of $R(T)$ is called the *rank* of T and is denoted by $\text{rank}(T)$.

By Theorem 7.2.4, if \mathbf{A} is the standard matrix for T , then $\text{rank}(T) = \text{rank}(\mathbf{A})$.

Definition 7.2.7. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation. The *kernel* of T , denoted by $\text{Ker}(T)$, is the set of vectors in \mathbb{R}^n whose image is the zero vector in \mathbb{R}^m , i.e.

$$\text{Ker}(T) = \{\mathbf{u} \mid T(\mathbf{u}) = \mathbf{0}\} \subseteq \mathbb{R}^n$$

Theorem 7.2.9. Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation and \mathbf{A} the standard matrix for T . Then

$$\text{Ker}(T) = \text{the nullspace of } \mathbf{A}$$

Definition 7.2.10. Let T be a linear transformation. The dimension of $\text{Ker}(T)$ is called the *nullity* of T and is denoted by $\text{nullity}(T)$.

By Theorem 7.2.9, if \mathbf{A} is the standard matrix for T , then $\text{nullity}(T) = \text{nullity}(\mathbf{A})$.

Theorem 7.2.13. If $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation, then

$$\text{rank}(T) + \text{nullity}(T) = n$$