Definition : Subspaces of $\mathbb{R}^n$

A set U of vectors in $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if it satisfies the following properties:

1. The zero vector $\mathbf{0} \in U$
2. If $\mathbf{x} \in U$ and $\mathbf{y} \in U$, then $\mathbf{x} + \mathbf{y} \in U$
3. If $\mathbf{x} \in U$, then $a\mathbf{x} \in U$ for all $a \in \mathbb{R}$

Definition : Null Space of a Matrix

The null space of an $m \times n$ matrix A is the set of all vectors for which multiplying by A gives 0. This is denoted:

$$\operatorname{null} A = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

Definition : Image Space of a Matrix

The image space of an $m \times n$ matrix A is the set of all vectors which can be obtained by multiplying A by a vector in $\mathbb{R}^n$. This is denoted:

$$\operatorname{im} A = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$$

Definition : Spanning Set

The set of all such linear combinations of a set of vectors $\mathbf{x}_i$ is called the span, and is denoted:

$$\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} = \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \mid t_i \in \mathbb{R}\}$$

If $V = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$, we say V is spanned by the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, or that the vectors span the space V.

Theorem : Span Theorem

Let $U = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ where each $\mathbf{x}_i \in \mathbb{R}^n$, then

1. U is a subspace of $\mathbb{R}^n$ containing each $\mathbf{x}_i$
2. If W is a subspace of $\mathbb{R}^n$ and each $\mathbf{x}_i \in W$, then $U \subseteq W$

Definition : Linear Independence in $\mathbb{R}^n$

We call a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k\}$ of vectors linearly independent if:

$$t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \ldots t_k\mathbf{x}_k = \mathbf{0} \Longrightarrow t_1 = t_2 = \cdots = t_k = 0$$

Theorem

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an independent set of vectors in $\mathbb{R}^n$, then every vector in $\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ has a unique representation as a linear combination of the $\mathbf{x}_i$.

Theorem : Independence Test

To verify that a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is independent, proceed as follows:

1. Set a linear combination equal to zero: $t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}$
2. Show that $t_i = 0$ for each i

If some nontrivial linear combination vanishes, the vectors are not independent.

Theorem : Inverse Theorem

The following conditions are equivalent for an $n \times n$ matrix A:

1. A is invertible
2. $A\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$ for $\mathbf{x} \in \mathbb{R}^n$
3. $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ for every $\mathbf{b} \in \mathbb{R}^n$

Theorem

Let A be an $m \times n$ matrix and let $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ denote the columns of A

1. $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ is independent in $\mathbb{R}^m$ iff $A\mathbf{x} = \mathbf{0}$ $(\mathbf{x} \in \mathbb{R}^n)$ implies $\mathbf{x} = \mathbf{0}$
2. $\mathbb{R}^m = \operatorname{span}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$ iff $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ for every $\mathbf{b} \in \mathbb{R}^m$

Theorem

The following are equivalent for an $n \times n$ matrix A:

1. A is invertible
2. The columns of A are linearly independent
3. The columns of A span $\mathbb{R}^n$
4. The rows of A are linearly independent
5. The rows of A span the set of all $1 \times n$ rows

Theorem : Fundamental Theorem

Let U be a subspace of $\mathbb{R}^n$. If U is spanned by m vectors, and if U contains k linearly independent vectors, then $k \leq m$.

Definition : Basis of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of vectors in U is called a basis of U if it satisfies the two following conditions:

1. $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent
2. $U = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$

Theorem : Invariance Theorem in $\mathbb{R}^n$

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ and $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$ are both bases of a subspace U of $\mathbb{R}^n$, then $m = k$.

Definition : Dimension of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$ and $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of U, the number m of vectors in the basis is called the dimension of U. This is denoted by:

$$\dim U = m$$

Note

We define:

$$\dim\{\mathbf{0}\} = 0$$

This means that $\{\mathbf{0}\}$ has a basis containing no vectors. This makes sense since $\mathbf{0}$ cannot belong to any independent set.

Theorem

Let $U \neq \{\mathbf{0}\}$ be a subspace of $\mathbb{R}^n$. Then,

1. U has a basis, and $\dim U \leq n$
2. Any independent set in U can be enlarged (by adding vectors from any fixed basis of U) to a basis of U
3. Any spanning set for U can be cut down (by deleting vectors) to a basis of U

Theorem

Let U be a subspace of $\mathbb{R}^n$ where $\dim U = m$ and let $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ be a set of m vectors in U. Then, B is independent iff B spans U.

Theorem

Let $U \subseteq W$ be subspaces of $\mathbb{R}^n$. Then,

1. $\dim U \leq \dim W$
2. If $\dim U = \dim W$, then $U = W$

Corollary

It follows from the previous theorem that if U is a subspace of $\mathbb{R}^n$, then $\dim U \in \{0, 1, 2, \ldots, n\}$ and

$$\dim U = 0 \iff U = \{\mathbf{0}\}$$

$$\dim U = n \iff U = \mathbb{R}^n$$

The other subspaces of $\mathbb{R}^n$ are called proper.

Definition : Length in $\mathbb{R}^n$

The length of a vector $\mathbf{x}$, denoted $\|\mathbf{x}\|$, is defined by:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Theorem

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, then:

1. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$
2. $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$
3. $(a\mathbf{x}) \cdot \mathbf{y} = a\,(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (a\mathbf{y}) \;\; \forall a \in \mathbb{R}$
4. $\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x}$
5. $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0$ iff $\mathbf{x} = \mathbf{0}$
6. $\|a\mathbf{x}\| = |a|\|\mathbf{x}\| \; \forall a \in \mathbb{R}$

Theorem : Cauchy Inequality

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

Moreover, $|\mathbf{x} \cdot \mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\|$ iff one of $\mathbf{x}$ and $\mathbf{y}$ is a multiple of the other.

The Cauchy inequality is equivalent to $(\mathbf{x} \cdot \mathbf{y})^2 \leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2$

Corollary : Triangle Inequality

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Definition : Distance in $\mathbb{R}^n$

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we define the distance function $d\,(\mathbf{x}, \mathbf{y})$ between $\mathbf{x}$ and $\mathbf{y}$ by

$$d\,(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Theorem

If $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, we have:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$
2. $d(\mathbf{x}, \mathbf{y}) = 0$ iff $\mathbf{x} = \mathbf{y}$
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
4. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

Definition : Orthogonal and Orthonormal Sets in $\mathbb{R}^n$

We say that two vectors $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ are orthogonal if $\mathbf{x} \cdot \mathbf{y} = 0$. More generally, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is called an orthogonal set if

$$\mathbf{x}_i \cdot \mathbf{x}_j = 0 \; \forall \; i \neq j \quad \text{and} \quad \mathbf{x}_i \neq \mathbf{0} \; \forall \; i$$

Note that $\{\mathbf{x}\}$ is an orthogonal set if $\mathbf{x} \neq \mathbf{0}$. A set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in $\mathbb{R}^n$ is called orthonormal if it is orthogonal and each $\mathbf{x}_i$ is a unit vector.

Theorem : Normalizing an Orthogonal Set

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set, then $\left\{\frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1, \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2, \ldots, \frac{1}{\|\mathbf{x}_k\|}\mathbf{x}_k\right\}$ is an orthonormal set, and we say that is the result of normalizing the orthogonal set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$.

Theorem : Pythagoras' Theorem in $\mathbb{R}^n$

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set in $\mathbb{R}^n$, then

$$\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 + \cdots + \|\mathbf{x}_k\|^2$$

Theorem : Expansion Theorem in $\mathbb{R}^n$

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal basis of a subspace U of $\mathbb{R}^n$. If $\mathbf{x}$ is any vector in U, we have

$$\mathbf{x} = \left(\frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\right)\mathbf{f}_1 + \left(\frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\right)\mathbf{f}_2 + \cdots + \left(\frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\right)\mathbf{f}_m$$

The expansion of $\mathbf{x}$ as a linear combination of the orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is called the Fourier expansion of $\mathbf{x}$ and the coefficients $t_i = \frac{\mathbf{x} \cdot \mathbf{f}_i}{\|\mathbf{f}_i\|^2}$ are called the Fourier coefficients.

Definition : Column and Row Space of a Matrix

The column space, denoted $\operatorname{col} A$, of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A.

The row space, denoted $\operatorname{row} A$, of A is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

Lemma

Let A and B denote $m \times n$ matrices.

1. If $A \to B$ by elementary row operations, then $\operatorname{row} A = \operatorname{row} B$
2. If $A \to B$ by elementary column operations, then $\operatorname{col} A = \operatorname{col} B$

Lemma

If R is a row-echelon matrix, then:

1. The nonzero rows of R are a basis of $\operatorname{row} R$
2. The columns of R containing leading ones are a basis of $\operatorname{col} R$

Theorem : Rank Theorem

Let A denote any $m \times n$ matrix of rank r. Then,

$$\dim(\operatorname{col} A) = \dim(\operatorname{row} A) = r$$

Moreover, if A is carried to a row-echelon matrix R by row operations, then

1. The r nonzero rows of R are a basis of $\operatorname{row} A$
2. If the leading 1s lie in columns $j_1, j_2, \dots, j_r$ of R, then columns $j_1, j_2, \dots, j_r$ of A are a basis of $\operatorname{col} A$

Corollary

If A is any matrix, then $\operatorname{rank} A = \operatorname{rank}\left(A^T\right)$.

Corollary

If A is an $m \times n$ matrix, then $\operatorname{rank} A \leq m$ and $\operatorname{rank} A \leq n$.

Corollary

If A is an $m \times n$ matrix, U is an $m \times m$ invertible matrix, and V is an $n \times n$ invertible matrix, then:

$$\operatorname{rank} A = \operatorname{rank}(UA) = \operatorname{rank}(AV)$$

Lemma

Let A, U, V be matrices of sizes $m \times n$, $p \times m$, and $n \times q$ respectively.

1. $\operatorname{col}(AV) \subseteq \operatorname{col} A$, with equality if $VV' = I_n$ for some V'
2. $\operatorname{row}(UA) \subseteq \operatorname{row} A$, with equality if $U'U = I_m$ for some U'

Corollary

If A is $m \times n$ and B is $n \times m$, then $\operatorname{rank}(AB) \leq \operatorname{rank} A$ and $\operatorname{rank}(AB) \leq \operatorname{rank} B$.

Theorem

Let A denote an $m \times n$ matrix of rank r. Then,

1. The $n - r$ basic solutions to the system $A\mathbf{x} = \mathbf{0}$ provided by the Gaussian algorithm are a basis of $\operatorname{null} A$, so $\dim(\operatorname{null} A) = n - r$
2. The rank theorem provides a basis of $\operatorname{im} A = \operatorname{col} A$, and $\dim(\operatorname{im} A) = r$

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\operatorname{rank} A = n$
2. The rows of A span $\mathbb{R}^n$
3. The columns of A are linearly independent in $\mathbb{R}^m$
4. The $n \times n$ matrix $A^T A$ is invertible
5. $CA = I_n$ for some $n \times m$ matrix C
6. If $A\mathbf{x} = \mathbf{0}$, $\mathbf{x} \in \mathbb{R}^n$, then $\mathbf{x} = \mathbf{0}$

Theorem

The following are equivalent for an $m \times n$ matrix A:

1. $\operatorname{rank} A = m$
2. The columns of A span $\mathbb{R}^m$
3. The rows of A are linearly independent in $\mathbb{R}^n$
4. The $m \times m$ matrix AA^T is invertible
5. $AC = I_m$ for some $n \times m$ matrix C
6. The system $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$

Definition : Similar Matrices

If A and B are $n \times n$ matrices, we say that A and B are similar, and write $A \sim B$, if $B = P^{-1}AP$ for some invertible matrix P.

If $A \sim B$, then necessarily $B \sim A$. To see why, suppose that $B = P^{-1}AP$. Then $A = PBP^{-1} = Q^{-1}BQ$ where $Q = P^{-1}$ is invertible. This proves the second of the following properties of similarity:

1. $A \sim A$ for all square matrices A
2. If $A \sim B$, then $B \sim A$
3. If $A \sim B$ and $B \sim C$, then $A \sim C$

These properties say that the similarity relation $\sim$ is an equivalence relation on the set of $n \times n$ matrices.

Note

Similarity is compatible with inverses, transposes and powers:

$$\text{If } A \sim B \text{ then } A^{-1} \sim B^{-1}, \quad A^T \sim B^T, \quad \text{and} \quad A^k \sim B^k \ \forall\, k \in \mathbb{Z}^+$$

Definition : Trace of a Matrix

The trace of an $n \times n$ matrix A, denoted $\operatorname{tr} A$, is defined to be the sum of the main diagonal elements of A. That is, if $A = [a_{ij}]$, then $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$.

Lemma

Let A and B be $n \times n$ matrices. Then $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Theorem

If A and B are similar $n \times n$ matrices, then A and B have the same determinant, rank, trace, characteristic polynomial, and eigenvalues.

Theorem

Let A be an $n \times n$ matrix.

1. The eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x)$ of A
2. The λ-eigenvectors $\mathbf{x}$ are the nonzero solutions to the homogeneous system

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

Theorem

Let A be an $n \times n$ matrix.

1. A is diagonalizable iff $\mathbb{R}^n$ has a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ consisting of eigenvectors of A
2. When this is the case, the matrix $P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$ is invertible and $P^{-1}AP = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where, for each i, λ_i is the eigenvalue of A corresponding to $\mathbf{x}_i$

Theorem

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ be eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ of an $n \times n$ matrix A. Then $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a linearly independent set.

Theorem

If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

Lemma

Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be a linearly independent set of eigenvectors of an $n \times n$ matrix A, extend it to a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$, and let

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$$

be the (invertible) $n \times n$ matrix with the $\mathbf{x}_i$ as its columns. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the (not necessarily distinct) eigenvalues of A corresponding to $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ respectively, then $P^{-1}AP$ has block form

$$P^{-1}AP = \begin{bmatrix} \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k) & B \\ 0 & A_1 \end{bmatrix}$$

where B has size $k \times (n-k)$ and A_1 has size $(n-k) \times (n-k)$.

Note that this lemma with $k = n$ shows that an $n \times n$ matrix A is diagonalizable if $\mathbb{R}^n$ has a basis of eigenvectors of A.

Definition : Eigenspace of a Matrix

If λ is an eigenvalue of an $n \times n$ matrix A, define the eigenspace of A corresponding to λ by:

$$E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\}$$

Recall the multiplicity of an eigenvalue λ of A is the number of times λ occurs as a root of the characteristic polynomial $c_A(x)$ of A. In other words, the multiplicity of λ is the largest integer $m \geq 1$ such that

$$c_A(x) = (x - \lambda)^m g(x)$$

for some polynomial $g(x)$.

Lemma

Let λ be an eigenvalue of multiplicity m of a square matrix A. Then $\dim(E_\lambda(A)) \leq m$.

Note

When does $\dim(E_\lambda(A)) = m$ for each eigenvalue λ? It turns out that this characterizes the diagonalizable $n \times n$ matrices A for which $c_A(x)$ factors completely over $\mathbb{R}$. By this we mean that $c_A(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$, where the λ_i are real numbers (not necessarily distinct). In other words, every eigenvalue of A is real.

Theorem

The following are equivalent for a square matrix A for which $c_A(x)$ factors completely.

1. A is diagonalizable
2. $\dim(E_\lambda(A))$ equals the multiplicity of λ for every eigenvalue λ of the matrix A

Definition : Vector Spaces

A vector space consists of a nonempty set V of objects (called vectors) that can be added, multiplied by a real number, and for which certain axioms hold. If $\mathbf{u}$ and $\mathbf{v}$ are two vectors in V, their sum is expressed as $\mathbf{u}+\mathbf{v}$, and the scalar product of $\mathbf{v}$ by a real number a is denoted as $a\mathbf{v}$. These axioms are called vector addition and scalar multiplication, respectively. The following axioms are assumed to hold.

Axioms for vector addition

(i) If $\mathbf{u}$ and $\mathbf{v}$ are in V, then $\mathbf{u}+\mathbf{v}$ is in V

(ii) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ for all $\mathbf{u}$ and $\mathbf{v}$ in V

(iii) $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$ for all $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in V

(iv) An element $\mathbf{0}$ in V exists such that $\mathbf{v}+\mathbf{0}=\mathbf{0}+\mathbf{v}=\mathbf{v}$ for every $\mathbf{v}$ in V

(v) For each $\mathbf{v}$ in V, an element $-\mathbf{v}$ in V exists such that $-\mathbf{v}+\mathbf{v}=\mathbf{0}$ and $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$

Axioms for scalar multiplication

(i) If $\mathbf{v}$ is in V, then $a\mathbf{v}$ is in V for all a in $\mathbb{R}$

(ii) $a\left(\mathbf{v}+\mathbf{w}\right)=a\mathbf{v}+a\mathbf{w}$ for all $\mathbf{v}$ and $\mathbf{w}$ in V and all a in $\mathbb{R}$

(iii) $(a+b)\,\mathbf{v}=a\mathbf{v}+b\mathbf{v}$ for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$

(iv) $a\left(b\mathbf{v}\right)=\left(ab\right)\mathbf{v}$ for all $\mathbf{v}$ in V and all a and b in $\mathbb{R}$

(v) $1\mathbf{v}=\mathbf{v}$ for all $\mathbf{v}$ in V

Definition : Polynomials

The set $\mathbf{P}$ of all polynomials is a vector space with addition and scalar multiplication defined as follows for $p\left(x\right)=a_0+a_1x+a_2x^2+\dots$ and $q\left(x\right)=b_0+b_1x+b_2x^2+\dots$:

$$p\left(x\right)+q\left(x\right)=\left(a_0+b_0\right)+\left(a_1+b_1\right)x+\left(a_2+b_2\right)x^2+\dots$$

$$ap(x)=aa_0+\left(aa_1\right)x+\left(aa_2\right)x^2+\dots$$

The zero vector is the zero polynomial, and the negative of a polynomial $p\left(x\right)=a_0+a_1x+a_2x^2+\dots$ is the polynomial $-p\left(x\right)=-a_0-a_1x-a_2x^2-\dots$ obtained by negating all the coefficients.

Definition : Polynomials of a Given Degree

Given $n \geq 1$, let $\mathbf{P}_n$ denote the set of all polynomials of degree at most n, together with the zero polynomial. That is

$$\mathbf{P}_n = \left\{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_0, a_1, a_2, \ldots, a_n \in \mathbb{R}\right\}$$

Then $\mathbf{P}_n$ is a vector space.

Definition : Functions over an Interval

The set $\mathbf{F}\left[a, b\right]$ of all functions on the interval $[a, b]$ is a vector space using pointwise addition $\left(\left(f+g\right)\left(x\right) = f\left(x\right) + g\left(x\right) \ \forall x \in [a, b]\right)$ and scalar multiplication $\left(\left(rf\right)\left(x\right) = rf\left(x\right) \ \forall x \in [a, b]\right)$. The zero function denoted 0 is the constant function defined by

$$0\left(x\right) = 0 \ \forall x \in [a, b]$$

The negative of a function f is denoted $-f$ and has action defined by

$$\left(-f\right)\left(x\right) = -f\left(x\right) \ \forall x \in [a, b]$$

Theorem : Cancellation

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in a vector space V. If $\mathbf{v} + \mathbf{u} = \mathbf{v} + \mathbf{w}$, then $\mathbf{u} = \mathbf{w}$.

Definition : Difference

We define the difference $\mathbf{u} - \mathbf{v}$ of two vectors in V as

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + \left(-\mathbf{v}\right)$$

Theorem

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a vector space V, the equation

$$\mathbf{x} + \mathbf{v} = \mathbf{u}$$

has one and only one solution $\mathbf{x}$ in V given by

$$\mathbf{x} = \mathbf{u} - \mathbf{v}$$

Theorem

Let $\mathbf{v}$ denote a vector in a vector space V and let a denote a real number

(i) $0\mathbf{v} = \mathbf{0}$

(ii) $a\mathbf{0} = \mathbf{0}$

(iii) If $a\mathbf{v} = \mathbf{0}$, then either $a = 0$ or $\mathbf{v} = \mathbf{0}$

(iv) $(-1)\,\mathbf{v} = -\mathbf{v}$

(v) $(-a)\,\mathbf{v} = -\,(a\mathbf{v}) = a\,(-\mathbf{v})$

Definition : Subspaces of a Vector Space

If V is a vector space, a nonempty subset $U \subseteq V$ is called a subspace of V if U is itself a vector space using the addition and scalar multiplication of V.

Theorem : Subspace Test

A subset U of a vector space is a subspace of V if and only if it satisfies the following three conditions:

(i) $\mathbf{0}$ lies in U where $\mathbf{0}$ is the zero vector of V

(ii) If $\mathbf{u}_1$ and $\mathbf{u}_2$ are in U, then $\mathbf{u}_1 + \mathbf{u}_2$ is also in U

(iii) If $\mathbf{u}$ is in U, then $a\mathbf{u}$ is in U for each scalar a

Note

If V is any vector space, both $\{\mathbf{0}\}$ and V are subspaces of V. The vector space $\{\mathbf{0}\}$ is called the zero subspace of V.

Definition : $m \times n$ Matrices

The set $\mathbf{M}_{mn}$ of all $m \times n$ matrices is a vector space where addition and scalar multiplication are the standard definitions for matrices.

Definition : Evaluation

Suppose $p\,(x)$ is a polynomial and a is a number. Then the number $p\,(a)$ obtained by replacing x by a in the expression for $p\,(x)$ is called the evaluation of $p\,(x)$ at a. If $p\,(a) = 0$, the number a is called a root of $p\,(x)$.

Theorem

$\mathbf{P}_n$ is a subspace of $\mathbf{P}$ for each $n \geq 0$.

Theorem

The subset $\mathbf{D}[a,b]$ of all differentiable functions on $[a,b]$ is a subspace of the vector space $\mathbf{F}[a,b]$.

Definition : Linear Combinations and Spanning

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a set of vectors in a vector space V. As in $\mathbb{R}^n$, a vector $\mathbf{v}$ is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ if it can be expressed in the form

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$$

where $a_1, a_2, \dots, a_n$ are scalars, called the coefficients of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. The set of all linear combinations of these vectors is called their span, and is denoted by

$$\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \mid a_i \in \mathbb{R}\}$$

If it happens that $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, these vectors are called a spanning set for V.

Theorem

Let $U = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ in a vector space V. Then:

(i) U is subspace of V containing each of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$

(ii) U is the "smallest" subspace containing these vectors in the sense that any subspace that contains each of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ must contain U

Theorem : Linear Independence and Dependence

As in $\mathbb{R}^n$, a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ in a vector space V is called linearly independent if it satisfies the following condition:

$$\text{If } s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n = \mathbf{0}, \quad \text{then} \quad s_1 = s_2 = \cdots = s_n = 0$$

Theorem

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a linearly independent set of vectors in a vector space V. If a vector $\mathbf{v}$ has two representations:

$$\mathbf{v} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 + \cdots + s_n\mathbf{v}_n$$

$$\mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_n\mathbf{v}_n$$

as a linear combination of these vectors, then $s_1 = t_1$, $s_2 = t_2$, $\dots$, $s_n = t_n$. In other words, every vector in V can be written in a unique way as a linear combination of the $\mathbf{v}_i$.

Theorem : Fundamental Theorem

Suppose a vector space V can be spanned by n vectors. If any set of m vectors in V is linearly independent, then $m \leq n$.

Definition : Basis of a Vector Space

As in $\mathbb{R}^n$, a set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of vectors in a vector space V is called a basis of V if it satisfies the two following conditions:

(i) $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is linearly independent

(ii) $V = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$

Theorem : Invariance Theorem

Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be two bases of a vector space V. Then $n = m$.

Definition : Dimension of a Vector Space

If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis of the nonzero vector space V, the number n of vectors in the basis is called the dimension of V, and we write

$$\dim V = n$$

The zero vector space $\{\mathbf{0}\}$ is defined to have dimension 0:

$$\dim\{\mathbf{0}\} = 0$$

Note

The space $\mathbf{M}_{mn}$ has dimension mn, and one basis consists of all $m \times n$ matrices with exactly one entry equal to 1 and all other entries equal to 0. We call this the standard basis of $\mathbf{M}_{mn}$.

Note

The space $\mathbf{P}_n$ has dimension $n+1$. The standard basis of this space is $\{1, x, x^2, \ldots, x^n\}$.

Lemma : Independent Lemma

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an independent set of vectors in a vector space V. If $\mathbf{u} \in V$ but $\mathbf{u} \notin \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, then $\{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is also independent.

Definition : Finite Dimensional and Infinite Dimensional Vector Spaces

A vector space V is called finite dimensional if it is spanned by a finite set of vectors. Otherwise, V is called infinite dimensional.

Note

The zero vector space $\{\mathbf{0}\}$ is finite dimensional because $\{\mathbf{0}\}$ is a spanning set.

Lemma

Let V be a finite dimensional vector space. If U is any subspace of V, then any independent subset of U can be enlarged to a finite basis of U.

Theorem

Let V be a finite dimensional vector space spanned by m vectors.

(i) V has a finite basis, and $\dim V \leq m$

(ii) Every independent set of vectors in V can be enlarged to a basis of V by adding vectors from any fixed basis of V

(iii) If U is a subspace of V, then

  (a) U is finite dimensional and $\dim U \leq \dim V$

  (b) If $\dim U = \dim V$ then $U = V$

Theorem

Let U and W be subspaces of the finite dimensional space V

(i) If $U \subseteq W$, then $\dim U \leq \dim W$

(ii) If $U \subseteq W$ and $\dim U = \dim W$, then $U = W$

Lemma : Independent Lemma

A set $D = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in a vector space V is dependent if and only if some vector in D is a linear combination of the others.

Theorem

Let V be a finite dimensional vector space. Any spanning set for V can be cut down (by deleting vectors) to a basis of V.

Theorem

Let V be a vector space with $\dim V = n$, and suppose S is a set of exactly n vectors in V. Then S is independent if and only if S spans V.

Definition : Sum and Intersection

If U and W are subspaces of V, their sum $U+W$ and their intersection $U \cap W$ are:

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U \text{ and } \mathbf{w} \in W\}$$

$$U \cap W = \{\mathbf{v} \mid \mathbf{v} \in U \text{ and } \mathbf{v} \in W\}$$

Theorem

Suppose that U and W are finite dimensional subspaces of a vector space V. Then $U+W$ is finite dimensional, and

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

Definition : Linear Transformations of Vector Spaces

If V and W are two vector spaces, a function $T : V \to W$ is called a linear transformation if it satisfies the following axioms.

(i) $T(\mathbf{v} + \mathbf{v}_1) = T(\mathbf{v}) + T(\mathbf{v}_1) \ \forall \mathbf{v}, \mathbf{v}_1 \in V$

(ii) $T(r\mathbf{v}) = rT(\mathbf{v}) \ \forall \mathbf{v} \in V, r \in \mathbb{R}$

Definition : Linear Operator

A linear transformation $T : V \to V$ is called a linear operator on V.

Note

If V and W are vector spaces, the following are linear transformations

(i) Identity Operator $V \to V$

$$1_V : V \to V \text{ where } 1_V(\mathbf{v}) = \mathbf{v} \ \forall \mathbf{v} \in V$$

(ii) Zero Transformation $V \to W$

$$0 : V \to W \text{ where } 0(\mathbf{v}) = \mathbf{0} \ \forall \mathbf{v} \in V$$

(iii) Scalar Operator $V \to V$

$$a : V \to V \text{ where } a(\mathbf{v}) = a\mathbf{v} \ \forall \mathbf{v} \in V, a \in \mathbb{R}$$

Theorem

Let $T : V \to W$ be a linear transformation

(i) $T(\mathbf{0}) = \mathbf{0}$

(ii) $T(-\mathbf{v}) = -T(\mathbf{v})\ \forall \mathbf{v} \in V$

(iii) $T(r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_k\mathbf{v}_k) = r_1T(\mathbf{v}_1) + r_2T(\mathbf{v}_2) + \cdots + r_kT(\mathbf{v}_k)\ \forall v_i \in V, r_i \in \mathbb{R}$

Definition : Equality of Linear Transformations

Two linear transformations $T : V \to W$ and $S : V \to W$ are called equal (written $T = S$) if they have the same action, that is, $T(\mathbf{v}) = S(\mathbf{v})\ \forall \mathbf{v} \in V$

Theorem

Let $T : V \to W$ and $S : V \to W$ be two linear transformations. Suppose that $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. If $T(\mathbf{v}_i) = S(\mathbf{v}_i)$ for each i, then $T = S$.

Theorem

Let V and W be vector spaces and let $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis of V. Given any vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ in W (not necessarily distinct), there exists a unique linear transformation $T : V \to W$ satisfying $T(\mathbf{b}_i) = \mathbf{w_i}$, for each $i = 1, 2, \ldots, n$. In fact, the action of T is as follows:

Given $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \in V$, $v_i \in \mathbb{R}$, then

$$T(\mathbf{v}) = T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdots + v_n\mathbf{w}_n$$

Definition : Kernel and Image of a Linear Transformation

The kernel of T (denoted $\ker T$) and the image of T (denoted $\operatorname{im} T$ or $T(V)$) are defined by

$$\ker T = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

$$\operatorname{im} T = \{T(\mathbf{v}) \mid \mathbf{v} \in V\} = T(V)$$

The kernel of T is often called the nullspace of T because it consists of all vectors $\mathbf{v} \in V$ satisfying the condition that $T(\mathbf{v}) = \mathbf{0}$. The image of T is often called the range of T and consists of all vectors $\mathbf{w} \in W$ of the form $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V$.

Theorem

Let $T : V \to W$ be a linear transformation.

(i) $\ker T$ is a subspace of V

(ii) $\operatorname{im} T$ is a subspace of W

Definition : Nullity and Rank

Given a linear transformation $T : V \to W$:

(i) $\dim(\ker T)$ is called the nullity of T and denoted as $\text{nullity}(T)$

(ii) $\dim(\operatorname{im} T)$ is called the rank of T and denoted as $\text{rank}(T)$

Definition : One-to-one and Onto Linear Transformations

Let $T : V \to W$ be a linear transformation.

(i) T is said to be onto if $\operatorname{im} T = W$

(ii) T is said to be one-to-one if $T(\mathbf{v}) = T(\mathbf{v}_1)$ implies $\mathbf{v} = \mathbf{v}_1$

Theorem

If $T : V \to W$ is a linear transformation, then T is one-to-one if and only if $\ker T = \{\mathbf{0}\}$.

Note

The identity transformation $1_V : V \to V$ is both one-to-one and onto for any vector space V.

Theorem

Let A be an $m \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation induced by A, that is $T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x} \in \mathbb{R}^n$.

(i) T_A is onto if and only if $\operatorname{rank} A = m$

(ii) T_A is one-to-one if and only if $\operatorname{rank} A = n$

Theorem : Dimension Theorem

Let $T : V \to W$ be any linear transformation and assume that $\ker T$ and $\operatorname{im} T$ are both finite dimensional. Then V is also finite dimensional and

$$\dim V = \dim(\ker T) + \dim(\operatorname{im} T)$$

In other words, $\dim V = \text{nullity}(T) + \text{rank}(T)$.

Theorem

Let $T : V \to W$ be a linear transformation, and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ be a basis of V such that $\{\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n\}$ is a basis of $\ker T$. Then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_r)\}$ is a basis of $\operatorname{im} T$, and hence $r = \operatorname{rank} T$.

Definition : Isomorphic Vector Spaces

A linear transformation $T : V \to W$ is called an isomorphism if it is both one-to-one and onto. The vector spaces V and W are said to be isomorphic if there exists an isomorphism $T : V \to W$, and we write $V \cong W$ when this is the case.

Note

An isomorphism $T : V \to W$ induces a pairing

$$\mathbf{v} \leftrightarrow T(\mathbf{v})$$

between vectors $\mathbf{v} \in V$ and vectors $T(\mathbf{v}) \in W$ that preserves vector addition and scalar multiplication. Hence, as far as their vector space properties are concerned, the spaces V and W are identical except for notation.

Theorem

If V and W are finite dimensional spaces, the following conditions are equivalent for a linear transformation $T : V \to W$.

(i) T is an isomorphism

(ii) If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is any basis of V, then $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.

(iii) There exists a basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of V such that $\{T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)\}$ is a basis of W.

Theorem

If V and W are finite dimensional vector spaces, then $V \cong W$ if and only if $\dim V = \dim W$

Corollary

Let U, V and W denote vector spaces. Then:

(i) $V \cong V$ for every vector space V

(ii) If $V \cong W$ then $W \cong V$

(iii) If $U \cong V$ and $V \cong W$, then $U \cong W$

Corollary

If V is a vector space and $\dim V = n$, then V is isomorphic to $\mathbb{R}^n$.

Definition : Coordinate Isomorphism

Let V be an n-dimensional vector space with basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$. The coordinate isomorphism corresponding to B is the linear transformation $C_B : V \to \mathbb{R}^n$ defined by

$$C_B\left(v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n\right) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

It satisfies $C_B\left(\mathbf{b}_i\right) = \mathbf{e}_i$ for each i and is an isomorphism.

Theorem

If V and W have the same dimension n, a linear transformation $T : V \to W$ is an isomorphism if it is either one-to-one or onto.

Definition : Composition of Linear Transformations

Given linear transformations $V \xrightarrow{T} W \xrightarrow{S} U$, the composite $ST : V \to U$ of T and S is defined by

$$ST\left(\mathbf{v}\right) = S\left[T\left(\mathbf{v}\right)\right] \;\; \forall \mathbf{v} \in V$$

The operation of forming the new function ST is called composition (sometimes denoted $S \circ T$).

Theorem

Let $V \xrightarrow{T} W \xrightarrow{S} U \xrightarrow{R} Z$ be linear transformations.

(i) The composite ST is again a linear transformation

(ii) $T1_V = T$ and $1_W T = T$

(iii) $\left(RS\right)T = R\left(ST\right)$

Theorem

Let V and W be finite dimensional vector spaces. The following conditions are equivalent for a linear transformation $T : V \to W$.

(i) T is an isomorphism

(ii) There exists a linear transformation $S : W \to V$ such that $ST = 1_V$ and $TS = 1_W$

Definition : Inverse of Linear Transformations

Given an isomorphism $T : V \to W$, there exists another isomorphism called the inverse of T, denoted by T^{-1}. $T : V \to W$ and $T^{-1} : W \to V$ are related by the fundamental identities:

$$T^{-1}\left[T\left(\mathbf{v}\right)\right] = \mathbf{v}\ \forall \mathbf{v} \in V \text{ and } T\left[T^{-1}\left(\mathbf{w}\right)\right] = \mathbf{w}\ \forall \mathbf{w} \in W$$

Lemma : Orthogonal Lemma in $\mathbb{R}^n$

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal set in $\mathbb{R}^n$. Given $\mathbf{x} \in \mathbb{R}^n$, write

$$\mathbf{f}_{m+1} = \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then:

(i) $\mathbf{f}_{m+1} \cdot \mathbf{f}_k = 0$ for $k = 1, 2, \ldots, m$

(ii) If $\mathbf{x} \notin \operatorname{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, then $\mathbf{f}_{m+1} \neq 0$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set.

Theorem

Let U be a subspace of $\mathbb{R}^n$

(i) Every orthogonal subset $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ in U is a subset of an orthogonal basis of U

(ii) U has an orthogonal basis

Theorem : Gram-Schmidt Orthogonalization Algorithm in $\mathbb{R}^n$

If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$ in U successively as follows:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \dots m$. Then:

(i) $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is an orthogonal basis of U

(ii) $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ for each $k = 1, 2, \dots, m$

Remark

Observe that the vector $\frac{\mathbf{x} \cdot \mathbf{f}_i}{\|\mathbf{f}_i\|^2}\mathbf{f}_i$ is unchanged if a nonzero scalar multiple of $\mathbf{f}_i$ is used in place of $\mathbf{f}_i$. Hence, if a newly constructed $\mathbf{f}_i$ is multiplied by a nonzero scalar at some stage of the Gram-Schmidt algorithm, the subsequent $\mathbf{f}$'s will be unchanged. This is useful in actual calculations.

Definition : Orthogonal Complement of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}$, define the orthogonal complement $U^\perp$ of U by:

$$U^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \ \forall \mathbf{y} \in U\}$$

Lemma

Let U be a subspace of $\mathbb{R}^n$

(i) $U^\perp$ is a subspace of $\mathbb{R}^n$

(ii) $\{\mathbf{0}\}^\perp = \mathbb{R}^n$ and $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$

(iii) If $U = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, then $U^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{x}_i = 0 \text{ for } i = 1, 2, \dots, k\}$

Definition : Projection onto a Subspace of $\mathbb{R}^n$

Let U be a subspace of $\mathbb{R}^n$ with orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. If $\mathbf{x} \in \mathbb{R}^n$, the vector

$$\operatorname{proj}_U \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

is called the orthogonal projection of $\mathbf{x}$ on U. For the zero subspace $U = \{\mathbf{0}\}$, we define

$$\operatorname{proj}_{\{\mathbf{0}\}} \mathbf{x} = \mathbf{0}$$

Theorem : Projection Theorem in $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$, write $\mathbf{p} = \operatorname{proj}_U \mathbf{x}$. Then:

(i) $\mathbf{p} \in U$ and $\mathbf{x} - \mathbf{p} \in U^\perp$

(ii) $\mathbf{p}$ is the vector in U closest to $\mathbf{x}$ in the sense that

$$\|\mathbf{x} - \mathbf{p}\| < \|\mathbf{x} - \mathbf{y}\| \ \forall \mathbf{y} \in U, \mathbf{y} \neq \mathbf{p}$$

Theorem

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \operatorname{proj}_U \mathbf{x} \ \forall \mathbf{x} \in \mathbb{R}^n$$

(i) T is a linear operator

(ii) $\operatorname{im} T = U$ and $\ker T = U^\perp$

(iii) $\dim U + \dim U^\perp = n$

Theorem : Orthogonal Matrix Criteria

The following conditions are equivalent for an $n \times n$ matrix P.

(i) P is invertible and $P^{-1} = P^T$

(ii) The rows of P are orthonormal

(iii) The columns of P are orthonormal

Definition : Orthogonal Matrices

An $n \times n$ matrix P is called an orthogonal matrix if it satisfies one (and hence all) of the conditions in the "Orthogonal Matrix Criteria" theorem.

Definition : Orthogonal Diagonalizable Matrices

An $n \times n$ matrix A is said to be orthogonally diagonalizable when an orthogonal matrix P can be found such that $P^{-1}AP = P^TAP$ is diagonal.

Theorem : Principal Axes Theorem for Matrices

The following conditions are equivalent for an $n \times n$ matrix A.

(i) A has an orthonormal set of n eigenvectors

(ii) A is orthogonally diagonalizable

(iii) A is symmetric

Since the eigenvalues of a (real) symmetric matrix are real, this theorem is sometimes called the real spectral theorem, and the set of distinct eigenvalues is called the spectrum of the matrix.

Definition : Principal Axes

A set of orthonormal eigenvectors of a symmetric matrix A is called a set of principal axes for A.

Theorem

If A is an $n \times n$ symmetric matrix, then

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$$

for all columns $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Theorem

If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Definition : Quadratic Forms & Principal Axes

If A is symmetric and a set of orthogonal eigenvectors of A is given, the eigenvectors are called principal axes of A. An expression $q = ax_1^2 + bx_1x_2 + cx_2^2$ is called a quadratic form in the variables x_1 and x_2, and the graph of the equation $q = 1$ is called a conic in these variables.

If we introduce new variables y_1 and y_2 by setting $x_1 = y_1 + y_2$ and $x_2 = y_1 - y_2$, then $q = y_1^2 - y_2^2$, a diagonal form with no cross term y_1y_2. Then, the y_1 and y_2 axes are called the principal axes for the conic.

Theorem : Triangulation Theorem

If A is an $n \times n$ matrix with n real eigenvalues, an orthogonal matrix P exists such that P^TAP is upper triangular.

Corollary

If A is an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (not necessarily distinct), then $\det A = \lambda_1 \lambda_2 \ldots \lambda_n$ and $\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Definition : Positive Definite Matrices

A square matrix is called positive definite if it is symmetric and all its eigenvalues λ are positive.

Since these matrices are symmetric, the principal axes theorem plays a central role in the theory.

Theorem

If A is positive definite, then it is invertible and $\det A > 0$.

Theorem

A symmetric matrix A is positive definite if and only if $\mathbf{x}^T A \mathbf{x} > 0$ for every column $\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n$

Definition : Principal Submatrices

If A is any $n \times n$ matrix, let ${}^{(r)}A$ denote the $r \times r$ submatrix in the upper left corner of A; that is, ${}^{(r)}A$ is the matrix obtained from A by deleting the last $n - r$ rows and columns. The matrices ${}^{(1)}A, {}^{(1)}A, {}^{(2)}A, \ldots, {}^{(n)}A = A$ are called the principal submatrices of A.

Lemma

If A is positive definite, so is each principal submatrix ${}^{(r)}A$ for $r = 1, 2, \ldots, n$.

Theorem

The following conditions are equivalent for a symmetric $n \times n$ matrix A:

(i) A is positive definite

(ii) $\det\left({}^{(r)}A\right) > 0$ for each $r = 1, 2, \ldots, n$

(iii) $A = U^T U$ where U is an upper triangular matrix with positive entries on the main diagonal

The factorization in (iii) is unique, and called the Cholesky factorization of A.

Theorem : Algorithm for the Cholesky Factorization

If A is a positive definite matrix, the Cholesky factorization $A = U^TU$ can be obtained as follows:

1. Carry A to an upper triangular matrix U_1 with positive diagonal entries using row operations each of which adds a multiple of a row to a lower row
2. Obtain U from U_1 by dividing each row of U_1 by the square root of the diagonal entry in that row

Definition : Complex Matrix

A matrix $A = [a_{ij}]$ is called a complex matrix if every entry a_{ij} is a complex number.

Definition : Complex Conjugate

If $z = a + bi$ is a complex number, the conjugate of z is also a complex number, denoted by $\overline{z}$, given by:

$$\overline{z} = a - bi$$

Definition : Conjugate Matrix

If A is a complex matrix, the conjugate of $A = [a_{ij}]$ is the matrix:

$$\overline{A} = [\overline{a}_{ij}]$$

obtained from A by conjugating every entry.

The properties

$$\overline{A + B} = \overline{A} + \overline{B} \quad \text{and} \quad \overline{AB} = \overline{A}\,\overline{B}$$

hold for all complex matrices of appropriate sizes.

Definition : Standard Inner Product in $\mathbb{C}^n$

Given $\mathbf{z} = (z_1, z_2, \dots, z_n)$ and $\mathbf{w} = (w_1, w_2, \dots, w_n)$ in $\mathbb{C}^n$, define their standard inner product by:

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n = \mathbf{z} \cdot \overline{\mathbf{w}}$$

Theorem

Let $\mathbf{z}, \mathbf{z}_1, \mathbf{w}, \mathbf{w}_1$ denote vectors in $\mathbb{C}^n$, and let λ denote a complex number.

(i) $\langle \mathbf{z} + \mathbf{z}_1, \mathbf{w} \rangle = \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{z}_1, \mathbf{w} \rangle$ and $\langle \mathbf{z}, \mathbf{w} + \mathbf{w_1} \rangle = \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{z}, \mathbf{w}_1 \rangle$

(ii) $\langle \lambda\mathbf{z}, \mathbf{w} \rangle = \lambda\langle \mathbf{z}, \mathbf{w} \rangle$ and $\langle \mathbf{z}, \lambda\mathbf{w} \rangle = \overline{\lambda}\langle \mathbf{z}, \mathbf{w} \rangle$

(iii) $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$

(iv) $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$, and $\langle \mathbf{z}, \mathbf{z} \rangle = 0$ iff $\mathbf{z} = \mathbf{0}$

Definition : Norm and Length in $\mathbb{C}^n$

We define the norm or length $\|\mathbf{z}\|$ of a vector $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$ by:

$$\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

Theorem

If $\mathbf{z} \in \mathbb{C}^n$, then:

(i) $\|\mathbf{z}\| \geq 0$ and $\|\mathbf{z}\| = 0$ iff $\mathbf{z} = \mathbf{0}$

(ii) $\|\lambda \mathbf{z}\| = |\lambda| \|\mathbf{z}\| \; \forall \lambda \in \mathbb{C}$

Definition : Conjugate Transpose in $\mathbb{C}^n$

The conjugate transpose A^H of a complex matrix A is defined by:

$$A^H = \left(\overline{A}\right)^T = \overline{(A^T)}$$

Theorem

Let A and B denote complex matrices, and let λ be a complex number.

(i) $\left(A^H\right)^H = A$

(ii) $(A + B)^H = A^H + B^H$

(iii) $(\lambda A)^H = \overline{\lambda} A^H$

(iv) $(AB)^H = B^H A^H$

Definition : Hermitian Matrices

A square complex matrix A is called hermitian if $A^H = A$, or equivalently if $\overline{A} = A^T$.

Theorem

An $n \times n$ complex matrix A is hermitian if and only if:

$$\langle A\mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{z}, A\mathbf{w} \rangle$$

for all n-tuples $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$.

Theorem

Let A denote a hermitian matrix.

(i) The eigenvalues of A are real

(ii) Eigenvectors of A corresponding to distinct eigenvalues are orthogonal

Definition : Orthogonal and Orthonormal Vectors in $\mathbb{C}^n$

As in the real case, a set of nonzero vectors $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m\}$ in $\mathbb{C}^n$ is called orthogonal if $\langle \mathbf{z}_i, \mathbf{z}_j \rangle = 0$ whenever $i \neq j$, and it is orthonormal if, in addition, $\|\mathbf{z}_i\| = 1$ for each i.

Theorem

The following are equivalent for an $n \times n$ complex matrix A.

(i) A is invertible and $A^{-1} = A^H$

(ii) The rows of A are an orthonormal set in $\mathbb{C}^n$

(iii) The columns of A are an orthonormal set in $\mathbb{C}^n$

Definition : Unitary Matrices

A square complex matrix U is unitary if $U^{-1} = U^H$. Thus, a real matrix is unitary iff it is orthogonal.

Definition : Unitary Diagonalizable

An $n \times n$ complex matrix A is called unitarily diagonalizable if $U^H AU$ is diagonal for some unitary matrix U.

Theorem : Schur's Theorem

If A is any $n \times n$ complex matrix, there exists a unitary matrix U such that

$$U^H AU = T$$

is upper triangular. Moreover, the entries on the main diagonal of T are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A (including multiplicities).

Corollary

Let A be an $n \times n$ complex matrix, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of A, including multiplicities. Then,

$$\det A = \lambda_1 \lambda_2 \ldots \lambda_n \quad \text{and} \quad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Theorem : Spectral Theorem

If A is hermitian, there is a unitary matrix U such that $U^H AU$ is diagonal.

Definition : Normal Matrix

An $n \times n$ complex matrix N is called normal if $NN^H = N^H N$.

Theorem

An $n \times n$ complex matrix A is unitarily diagonalizable if and only if A is normal.

Theorem : Cayley-Hamilton Theorem

If A is an $n \times n$ complex matrix, then $c_A(A) = 0$; that is, A is a root of its characteristic polynomial.

Definition : Quadratic Form

A quadratic form q in the n variables $x_1, x_2, \ldots, x_n$ is a linear combination of terms $x_1^2, x_2^2, \ldots x_n^2$ and cross terms $x_1x_2, x_1x_3, x_2x_3, \ldots$

Theorem : Diagonalization Theorem

Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in the variables $x_1, x_2, \ldots x_n$, where $\mathbf{x} = (x_1, x_2, \ldots x_n)^T$ and A is a symmetric $n \times n$ matrix. Let P be an orthogonal matrix such that $P^T AP$ is diagonal, and define new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ by:

$$\mathbf{x} = P\mathbf{y} \quad \Leftrightarrow \quad \mathbf{y} = P^T\mathbf{x}$$

If q is expressed in terms of these new variables $y_1, y_2, \ldots, y_n$, the result is:

$$q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A repeated according to their multiplicities.

Theorem

If $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form given by a symmetric matrix A, then A is uniquely determined by q.

Definition : Congruent Matrices

Two $n \times n$ matrices A and B are called congruent, written $A \overset{c}{\sim} B$ if $B = U^T AU$ for some invertible matrix U. Some properties of congruence are:

(i) $A \overset{c}{\sim} A$ for all A

(ii) If $A \overset{c}{\sim} B$, then $B \overset{c}{\sim} A$

(iii) If $A \overset{c}{\sim} B$ and $B \overset{c}{\sim} C$, then $A \overset{c}{\sim} C$

(iv) If $A \overset{c}{\sim} B$, then A is symmetric iff B is symmetric

(v) If $A \overset{c}{\sim} B$, then $\operatorname{rank} A = \operatorname{rank} B$

Theorem : Sylvester's Law of Inertia

If $A \overset{c}{\sim} B$, then A and B have the same number of positive eigenvalues, counting multiplicities.

Definition : Index

The index of a symmetric matrix A is the number of positive eigenvalues of A. If $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form, the index and rank of q are defined to be the index and rank of the matrix A respectively.

Note : Complete Diagonalization

Let $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be any quadratic form in n variables, of index k and rank r where A is symmetric. We claim there are new variables $\mathbf{z}$ can be found so that q is completely diagonalized, that is,

$$q(\mathbf{z}) = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

If $k \leq r \leq n$, let $D_n(k, r)$ denote the $n \times n$ diagonal matrix whose main diagonal consists of k 1's, followed by $r-k$ -1's, followed by $n-r$ 0's. Then we seek variables $\mathbf{z}$ such that:

$$q(\mathbf{z}) = \mathbf{z}^T D_n(k, r)\,\mathbf{z}$$

To determine $\mathbf{z}$, first diagonalize A as follows: Find an orthogonal matrix P_0 such that:

$$P_0^T A P_0 = D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$$

is diagonal with the nonzero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ of A on the main diagonal (followed by $n-r$ zeros). By reordering the columns of P_0, if necessary, we may assume that $\lambda_1, \ldots, \lambda_k$ are positive and $\lambda_{k+1}, \ldots, \lambda_r$ are negative. This being the case, let D_0 be the $n \times n$ diagonal matrix:

$$D_0 = \operatorname{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_k}}, \frac{1}{\sqrt{-\lambda_{k+1}}}, \ldots, \frac{1}{\sqrt{\lambda_r}}, 1, \ldots, 1\right)$$

Then $D_0^T D D_0 = D_n(k, r)$, so if new variables $\mathbf{z}$ are given by $\mathbf{x} = (P_0 D_0)\,\mathbf{z}$, we obtain:

$$q(\mathbf{z}) = \mathbf{z}^T D_n(k, r)\,\mathbf{z} = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

as required. Note that the change-of-variables matrix $P_0 D_0$ from $\mathbf{z}$ to $\mathbf{x}$ has orthogonal columns (in fact, scalar multiples of the columns of P_0).

Theorem

Let A and B be symmetric $n \times n$ matrices, and let $0 \leq k \leq r \leq n$.

(i) A has index k and rank r if and only if $A \overset{c}{\sim} D_n(k, r)$

(ii) $A \overset{c}{\sim} B$ if and only if they have the same rank and index

Definition : Ordered Basis

An ordered basis is a basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ where the order of the vectors is taken into account.

Definition : Coordinate Vector $C_B(\mathbf{v})$

Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an ordered basis of a vector space V, and let $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n$. Then, the coordinate vector of $\mathbf{v}$ with respect to B is defined to be

$$C_B(\mathbf{v}) = (v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Note that $C_B(\mathbf{b}_i) = \mathbf{e}_i$ is column i of I_n.

Theorem

If V has dimension n and $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is any ordered basis of V, the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism. In fact, $C_B^{-1} : \mathbb{R}^n \to V$ is given by:

$$C_B^{-1}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \quad \forall \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

Definition : Matrix of T corresponding to the ordered bases B and D

$M_{DB}(T)$ of $T : V \to W$ for bases B and D is called the matrix of T corresponding to the ordered bases B and D, we use the following notation:

$$M_{DB}(T) = \begin{bmatrix} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{bmatrix}$$

Theorem

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$, and let $B = \{\mathbf{b}_1 \dots, \mathbf{b}_n\}$ and D be ordered bases of V and W, respectively. Then the matrix $M_{DB}(T)$ is the unique $m \times n$ matrix A that satisfies:

$$C_D T = T_A C_B$$

Hence the defining property of $M_{DB}(T)$ is:

$$C_D [T(\mathbf{v})] = M_{DB}(T) C_B(\mathbf{v}) \;\; \forall \mathbf{v} \in V$$

The matrix $M_{DB}(T)$ is given in terms of its columns by:

$$M_{DB}(T) = \begin{bmatrix} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{bmatrix}$$

Theorem

Let $V \xrightarrow{T} W \xrightarrow{S} U$ be linear transformations and let B, D, and E be finite ordered bases of V, W, and U, respectively. Then:

$$M_{EB}(ST) = M_{ED}(S) \cdot M_{DB}(T)$$

Theorem

Let $T : V \to W$ be a linear transformation, where $\dim V = \dim W = n$. The following are equivalent:

(i) $M_{DB}(T)$ is invertible for all ordered bases B and D of V and W

(ii) $M_{DB}(T)$ is invertible for some pair of ordered bases B and D of V and W

When this is the case, $[M_{DB}(T)]^{-1} = M_{BD}(T^{-1})$.

Definition : Rank of a Linear Transformation

The rank of a linear transformation $T : V \to W$ is $\operatorname{rank} T = \dim(\operatorname{im} T)$. Moreover, if A is any $m \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is the matrix transformation, then $\operatorname{rank}(T_A) = \operatorname{rank} A$.

Theorem

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$. If B and D are any ordered bases of V and W, then $\operatorname{rank} T = \operatorname{rank}[M_{DB}(T)]$.

Note

If $T : V \to V$ is a linear operator where $\dim V = n$, it is possible to choose bases B and D of V such that the matrix $M_{DB}(T)$ has a very simple form:

$$M_{DB}(T) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where $r = \operatorname{rank} T$.

Definition : Matrix $M_B(T)$ of $T : V \to V$ for basis B

If $T : V \to V$ is an operator on a vector space V, and if B is an ordered basis of V, define $M_B(T) = M_{BB}(T)$ and call this the B-matrix of T.

Recall that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, then $C_B(\mathbf{x}) = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, so $M_E(T) = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$ is the standard matrix of the operator T.

Theorem

Let $T : V \to V$ be an operator where $\dim V = n$, and let B be an ordered basis of V.

(i) $C_B(T(\mathbf{v})) = M_B(T)\, C_B(\mathbf{v})\ \ \forall \mathbf{v} \in V$

(ii) If $S : V \to V$ is another operator on V, then $M_B(ST) = M_B(S)\, M_B(T)$

(iii) T is an isomorphism if and only if $M_B(T)$ is invertible. In this case $M_D(T)$ is invertible for every ordered basis D of V

(iv) If T is an isomorphism, then $M_B(T^{-1}) = [M_B(T)]^{-1}$

(v) If $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, then $M_B(T) = \begin{bmatrix} C_B[T(\mathbf{b}_1)] & C_B[T(\mathbf{b}_2)] & \cdots & C_B[T(\mathbf{b}_n)] \end{bmatrix}$

Definition : Change Matrix $P_{D \leftarrow B}$ for bases B and D

Define the change matrix $P_{D \leftarrow B}$ by:

$$P_{D \leftarrow B} = M_{DB}(1_V) \quad \text{for any ordered bases } B \text{ and } D \text{ of } V$$

Theorem

Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and D denote ordered bases of a vector space V. Then the change matrix $P_{D\leftarrow B}$ is given in terms of its columns by:

$$P_{D\leftarrow B} = \left[\begin{array}{cccc} C_D(\mathbf{b}_1) & C_D(\mathbf{b}_2) & \cdots & C_D(\mathbf{b}_n) \end{array}\right]$$

and has the property that:

$$C_D(\mathbf{v}) = P_{D\leftarrow B} C_B(\mathbf{v}) \ \forall \mathbf{v} \in V$$

Moreover, if E is another ordered basis of V, we have:

(i) $P_{B\leftarrow B} = I_n$

(ii) $P_{D\leftarrow B}$ is invertible and $(P_{D\leftarrow B})^{-1} = P_{B\leftarrow D}$

(iii) $P_{E\leftarrow D} P_{D\leftarrow B} = P_{E\leftarrow B}$

Theorem : Similarity Theorem

Let B_0 and B be two ordered bases of a finite dimensional vector space V. If $T : V \to V$ is any linear operator, the matrices $M_B(T)$ and $M_{B_0}(T)$ of T with respect to these bases are similar. More precisely,

$$M_B(T) = P^{-1} M_{B_0}(T) P$$

where $P = P_{B_0\leftarrow B}$ is the change matrix from B to B_0.

Theorem

Let A be an $n \times n$ matrix and let E be the standard basis of $\mathbb{R}^n$.

(i) Let A' be similar to A, say $A' = P^{-1}AP$, and let B be the ordered basis of $\mathbb{R}^n$ consisting of the columns of P in order. Then $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is linear and:

$$M_E(T_A) = A \quad \text{and} \quad M_B(T_A) = A'$$

(ii) If B is any ordered basis of $\mathbb{R}^n$, let P be the (invertible) matrix whose columns are the vectors in B in order. Then:

$$M_B(T_A) = P^{-1}AP$$

Definition : Similarity Invariant

A property of $n \times n$ matrices is similarity invariant if, whenever a given $n \times n$ matrix A which has the property, every matrix similar to A also has the property.

Definition : Determinant

If $T : V \to V$ is a linear operator on a finite dimensional space V, define the determinant of T (denoted $\det T$) by:

$$\det T = \det M_B(T), \quad B \text{ any basis of } V$$

Definition : Trace

If $T : V \to V$ is a linear operator on a finite dimensional space V, define the trace of T (denoted $\operatorname{tr} T$) by:

$$\operatorname{tr} T = \operatorname{tr} M_B(T), \quad B \text{ any basis of } V$$

Definition : Characteristic Polynomial

If $T : V \to V$ is a linear operator on a finite dimensional space V, define the characteristic polynomial of T by:

$$c_T(x) = c_A(x) \text{ where } A = M_B(T), \quad B \text{ any basis of } V$$

In other words, the characteristic polynomial of an operator T is the characteristic polynomial of any matrix representing T.

Definition : Image under T

If U is a subspace of V, write its image under T as:

$$T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \in U\}$$

Definition : T-invariant Subspace

Let $T : V \to V$ be an operator. A subspace $U \subseteq V$ is called T-invariant if $T(U) \subseteq U$, that is $T(\mathbf{u}) \in U$ for every vector $\mathbf{u} \in U$. Hence, T is a linear operator on the vector space U.

Theorem

Let $T : V \to V$ be any linear operator. Then:

(i) $\{\mathbf{0}\}$ and V are T-invariant subspaces

(ii) Both $\ker T$ and $\operatorname{im} T = T(V)$ are T-invariant subspaces

(iii) If U and W are T-invariant subspaces, so are $T(U)$, $U \cap W$, and $U + W$

Definition : Restriction of an Operator

Let $T : V \to V$ be a linear operator. If U is any T-invariant subspaces of V, then:

$$T : U \to U$$

is a linear operator on the subspace U, called the restriction of T to U

Theorem

Let $T : V \to V$ be a linear operator where V has dimension n and suppose that U is any T-invariant subspace of V. Let $B_1 = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ be any basis of U and extend it to a basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ of V is any way. Then $M_B(T)$ has the block triangular form:

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & Y \\ 0 & Z \end{bmatrix}$$

where Z is $(n-k) \times (n-k)$ and $M_{B_1}(T)$ is the matrix of the restriction of T to U.

Theorem

Let A be a block upper triangular matrix, say

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & A_{22} & A_{23} & \cdots & A_{2n} \\ 0 & 0 & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & A_{nn} \end{bmatrix}$$

where the diagonal blocks are square. Then:

(i) $\det A = (\det A_{11})(\det A_{22})(\det A_{33}) \cdots (\det A_{nn})$

(ii) $c_A(x) = c_{A_{11}}(x)\, c_{A_{22}}(x)\, c_{A_{33}}(x) \cdots c_{A_{nn}}(x)$

Definition : One Dimensional T-Invariant Subspace

Let $T : V \to V$ be a linear operator. A one dimensional subspace $\mathbb{R}\mathbf{v}$, $\mathbf{v} \neq 0$, is T-invariant if and only if $T(r\mathbf{v}) = rT(\mathbf{v})$ lies in $\mathbb{R}\mathbf{v}$ for all $r \in \mathbb{R}$. This holds if and only if $T(\mathbf{v})$ lies in $\mathbb{R}\mathbf{v}$, that is, $T(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}$.

Definition : Eigenvalue of an Operator

A real number λ is called an eigenvalue of an operator $T : V \to V$ if:

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

holds for some nonzero vector $\mathbf{v} \in V$. In this case, $\mathbf{v}$ is called an eigenvector of T corresponding to λ.

Definition : Eigenspace of a Linear Operator

Let $T : V \to V$. The subspace:

$$E_\lambda(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \lambda \mathbf{v}\}$$

is called the eigenspace of T corresponding to λ.

Theorem

If A is an $n \times n$ matrix, a real number λ is an eigenvalue of the matrix operator $T_A : \mathbb{R}^n \to \mathbb{R}^n$ if and only if λ is an eigenvalue of the matrix A. Moreover, the eigenspaces agree:

$$E_\lambda(T_A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda \mathbf{x}\} = E_\lambda(A)$$

Theorem

Let $T : V \to V$ be a linear operator where $\dim V = n$ and let B denote any ordered basis of V, and let $C_B : V \to \mathbb{R}^n$ denote the coordinate isomorphism. Then:

(i) The eigenvalues λ of T are precisely the eigenvalues of the matrix $M_B(T)$ and thus are the roots of the characteristic polynomial $c_T(x)$

(ii) In this case, the eigenspaces $E_\lambda(T)$ and $E_\lambda[M_B(T)]$ are isomorphic via the restriction $C_B : E_\lambda(T) \to E_\lambda[M_B(T)]$

Theorem

Each eigenspace of the linear operator $T : V \to V$ is a T-invariant subspace of V.

Definition : Direct Sum of Subspaces

A vector space V is said to be the direct sum of subspaces U and W if:

$$U \cap W = \{\mathbf{0}\} \quad \text{and} \quad U + W = V$$

In this case, we write $V = U \oplus W$. Given a subspace U, any subspace W such that $U \oplus W$ is called a complement of U in V.

Theorem

Let U and W be subspaces of a finite dimensional vector space V. The following three conditions are equivalent:

(i) $V = U \oplus W$

(ii) Each vector $\mathbf{v} \in V$ can be written uniquely in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{w} \quad \text{where } \mathbf{u} \in U, \mathbf{w} \in W$$

(iii) If $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases of U and W respectively, then $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is a basis of V.

Theorem

Let $T : V \to V$ be a linear operator where V has dimension n. Suppose $V = U_1 \oplus U_2$ where both U_1 and U_2 are T-invariant. If $B_1 = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ and $B_2 = \{\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ are bases of U_1 and U_2 respectively, then:

$$B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$$

is a basis of V, and $M_B(T)$ has the block diagonal form:

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix}$$

where $M_{B_1}(T)$ and $M_{B_2}(T)$ are the matrices of the restrictions of T to U_1 and to U_2 respectively.

Definition : Reducible Linear Operator

A linear operator $T : V \to V$ is said to be reducible if nonzero T-invariant subspaces U_1 and U_2 can be found such that $V = U_1 \oplus U_2$.

Definition : Involution

Let $T : V \to V$ be a linear operator such that $T^2 = 1_V$. Then, T is an involution.

Definition : Inner Product Spaces

An inner product on a real vector space V is a function that assigns a real number $\langle \mathbf{v}, \mathbf{w} \rangle$ to every pair $\mathbf{v}, \mathbf{w}$ of vectors in V in such a way that the following axioms are satisfied:

(i) $\langle \mathbf{v}, \mathbf{w} \rangle$ is a real number for all $\mathbf{v}, \mathbf{w} \in V$

(ii) $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$

(iii) $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$

(iv) $\langle r\mathbf{v}, \mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{v}, \mathbf{w} \in V$ and all $r \in \mathbb{R}$

(v) $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ for all $\mathbf{v} \neq \mathbf{0}$ in V

Definition : Inner Product Space

A real vector space V with an inner product $\langle \ , \ \rangle$ is called an inner product space.

Note

Every subspace of an inner product space is again an inner product space using the same inner product.

Theorem : L

t V be an inner product space with inner product $\langle \ , \ \rangle$; let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ denote vectors in V, and let r denote a real number.

(i) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(ii) $\langle \mathbf{v}, r\mathbf{w} \rangle = r\langle \mathbf{v}, \mathbf{w} \rangle = \langle r\mathbf{v}, \mathbf{w} \rangle$

(iii) $\langle \mathbf{v}, \mathbf{0} \rangle = 0 = \langle \mathbf{0}, \mathbf{v} \rangle$

(iv) $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$

Theorem

If A is any $n \times n$ positive definite matrix, then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y} \quad \text{for all columns } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

defines an inner product on $\mathbb{R}^n$, and every inner product on $\mathbb{R}^n$ arises in this way.

Note

Just as every linear operator $\mathbb{R}^n \to \mathbb{R}^n$ corresponds to an $n \times n$ matrix, every inner product on $\mathbb{R}^n$ corresponds to a positive definite $n \times n$ matrix. In particular, the dot product corresponds to the identity matrix I_n.

Remark

If we refer to the inner product space $\mathbb{R}^n$ without specifying the inner product, we mean that the dot product is to be used.

Definition : Norm and Distance

As in $\mathbb{R}^n$, if $\langle\ ,\ \rangle$ is an inner product on a space V, the norm $\|\mathbf{v}\|$ of a vector $\mathbf{v} \in V$ is defined by:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

We define the distance between vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space to be:

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

Definition : Unit Vector

A vector $\mathbf{v}$ in an inner product space V is called a unit vector if $\|\mathbf{v}\| = 1$. The set of all unit vectors in V is called the unit ball in V.

Theorem

If $\mathbf{v} \neq \mathbf{0}$ is any vector in an inner product space V, then $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ is the unique unit vector that is a positive multiple of $\mathbf{v}$.

Theorem : Cauchy-Schwarz Inequality

If $\mathbf{v}$ and $\mathbf{w}$ are two vectors in an inner product space V, then

$$\langle \mathbf{v}, \mathbf{w} \rangle^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2$$

Moreover, equality occurs if and only if one of $\mathbf{v}$ and $\mathbf{w}$ is a scalar multiple of the other.

Theorem

If V is an inner product, the norm $\|\ \|$ has the following properties:

(i) $\|\mathbf{v}\| \geq 0$ for every vector $\mathbf{v} \in V$

(ii) $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$

(iii) $\|r\mathbf{v}\| = |r|\|\mathbf{v}\|$ for every $\mathbf{v} \in V$ and every $r \in \mathbb{R}$

(iv) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ for all $\mathbf{v}, \mathbf{w} \in V$ (triangle inequality)

Theorem

Let V be an inner product space.

(i) $d(\mathbf{v}, \mathbf{w}) \geq 0$ for all $\mathbf{v}, \mathbf{w} \in V$

(ii) $d(\mathbf{v}, \mathbf{w}) = 0$ if and only if $\mathbf{v} = \mathbf{w}$

(iii) $d(\mathbf{v}, \mathbf{w}) = d(\mathbf{w}, \mathbf{v})$ for all $\mathbf{v}, \mathbf{w} \in V$

(iv) $d(\mathbf{v}, \mathbf{w}) \leq d(\mathbf{v}, \mathbf{u}) + d(\mathbf{u}, \mathbf{w})$ for all $\mathbf{v}, \mathbf{u}, \mathbf{w} \in V$

Definition : Orthogonality

Two vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V are said to be orthogonal if

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0$$

Definition : Orthogonal and Orthonormal Sets

A set $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of vectors is called an orthogonal set of vectors if:

(i) Each $\mathbf{f}_i \neq \mathbf{0}$

(ii) $\langle \mathbf{f}_i, \mathbf{f}_j \rangle$ for all $i \neq j$

If, in addition, $\|\mathbf{f}_i\| = 1$ for each i, the set $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is called an orthonormal set.

Theorem : Pythagoras' Theorem

If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is an orthogonal set of vectors, then:

$$\|\mathbf{f}_1 + \mathbf{f}_2 + \cdots + \mathbf{f}_n\|^2 = \|\mathbf{f}_1\|^2 + \|\mathbf{f}_2\|^2 + \cdots + \|\mathbf{f}_n\|^2$$

Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ be an orthogonal set of vectors.

(i) $\{r_1\mathbf{f}_1, r_2\mathbf{f}_2, \ldots, r_n\mathbf{f}_n\}$ is also orthogonal for any $r_i \neq 0 \in \mathbb{R}$

(ii) $\left\{\frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1, \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2, \ldots, \frac{1}{\|\mathbf{f}_n\|}\mathbf{f}_n\right\}$ is an orthonormal set

The process of passing from an orthogonal set to an orthonormal one is called normalizing the orthogonal set.

Theorem

Every orthogonal set of vectors is linearly independent.

Theorem : Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ be an orthogonal basis of an inner product space V. If $\mathbf{v}$ is any vector in V, then:

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}\mathbf{f}_n$$

is the expansion of $\mathbf{v}$ as a linear combination of the basis vectors.

The coefficients $\frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1, \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2, \ldots, \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}\mathbf{f}_n$ in the expansion theorem are sometimes called the Fourier coefficients of $\mathbf{v}$ with respect to the orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$.

Definition : Lagrange Polynomials

Let $a_0, a_1, \ldots, a_n$ be distinct numbers. The Lagrange polynomial corresponding to a_k is

$$\delta_k(x) = \frac{\prod_{i \neq k}(x - a_i)}{\prod_{i \neq k}(a_k - a_i)} \quad k = 0, 1, 2, \ldots, n$$

It satisfies the interpolation property

$$\delta_k(a_i) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}$$

with the inner product

$$\langle p(x), q(x) \rangle = \sum_{i=0}^{n} p(a_i) q(a_i)$$

the set $\{\delta_0, \delta_1, \ldots, \delta_n\}$ forms an orthonormal basis for $\mathbf{P}_n$.

Definition : Lagrange Interpolation Expansion

Every polynomial $p(x) \in \mathbf{P}_n$ can be written as

$$p(x) = \sum_{k=0}^{n} p(a_k) \delta_k(x)$$

This is called the Lagrange interpolation expansion (or Lagrange interpolation formula). It reconstructs a polynomial from its values at the distinct points $a_0, a_1, \ldots, a_n$.

Lemma : Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal set of vectors in an inner product space V, and let $\mathbf{v}$ be any vector not in span $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. Define

$$\mathbf{f}_{m+1} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set of vectors.

Theorem : Gram-Schmidt Orthogonalization Algorithm

Let V be an inner product space and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis of V. Define vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n \in V$ successively as follows:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{v}_1 \\
\mathbf{f}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{v}_k - \frac{\langle \mathbf{v}_k, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_k, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}_k, \mathbf{f}_{k-1} \rangle}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \ldots n$. Then:

(i) $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is an orthogonal basis of V

(ii) span $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ = span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ for each $k = 1, 2, \ldots, n$

The purpose of the Gram-Schmidt algorithm is to convert a basis of an inner product space into an orthogonal basis. In particular, it shows that every finite dimensional inner product space has an orthogonal basis.

Corollary

If V is any n-dimensional inner product space, then V is isomorphic to $\mathbb{R}^n$ as inner product spaces. More precisely, if E is any orthonormal basis of V, the coordinate isomorphism

$$C_E : V \to \mathbb{R}^n \text{ satisfies } \langle \mathbf{v}, \mathbf{w} \rangle = C_E(\mathbf{v}) \cdot C_E(\mathbf{w})$$

for all $\mathbf{v}, \mathbf{w} \in V$.

Definition : Orthogonal Complement

Let U be a subspace of an inner product space V. As in $\mathbb{R}^n$, the orthogonal complement $U^\perp$ of U in V is defined by:

$$U^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \ \forall \mathbf{u} \in U\}$$

Theorem

Let U be a finite dimensional subspace of an inner product space V

(i) $U^\perp$ is a subspace of V and $V = U \oplus U^\perp$

(ii) If $\dim V = n$, then $\dim U + \dim U^\perp = n$

(iii) If $\dim V = n$, then $U^{\perp\perp} = U$

Definition : Orthogonal Projection on a Subspace

The projection on U with kernel $U^\perp$ is called the orthogonal projection on U (or simply the projection on U) and is denoted $\mathrm{proj}_U : V \to V$

Theorem : Projection Theorem

Let U be a finite dimensional subspace of an inner product space V and let $\mathbf{v} \in V$

(i) $\mathrm{proj}_U : V \to V$ is a linear operator with image U and kernel $U^\perp$

(ii) $\mathrm{proj}_U \mathbf{v} \in U$ and $(\mathbf{v} - \mathrm{proj}_U \mathbf{v}) \in U^\perp$

(iii) If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is any orthogonal basis of U, then

$$\mathrm{proj}_U \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Note that there is no requirement that V is finite dimensional.

Theorem : Approximation Theorem

Let U be a finite dimensional subspace of an inner product space V. If $\mathbf{v}$ is any vector in V, then $\mathrm{proj}_U \mathbf{v}$ is the vector in U that is closest to $\mathbf{v}$. Here closest means that

$$\|\mathbf{v} - \mathrm{proj}_U \mathbf{v}\| < \|\mathbf{v} - \mathbf{u}\|$$

for all $\mathbf{u} \in U, u \neq \mathrm{proj}_U \mathbf{v}$.

Theorem

Let $T : V \to V$ be a linear operator on a finite dimensional space V. Then the following conditions are equivalent:

(i) V has a basis consisting of eigenvectors of T

(ii) There exists a basis B of V such that $M_B(T)$ is diagonal

Definition : Diagonalizable Linear Operators

A linear operator T on a finite dimensional space V is called diagonalizable if V has a basis consisting of eigenvectors of T.

Theorem

Let $T : V \to V$ be a linear operator on an inner product space V. If $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthogonal basis of V, then

$$M_B(T) = \left[\frac{\langle \mathbf{b}_i, T(\mathbf{b}_j)\rangle}{\|\mathbf{b}_i\|^2}\right]$$

Theorem

Let V be a finite dimensional inner product space. The following conditions are equivalent for a linear operator $T : V \to V$.

(i) $\langle \mathbf{v}, T(\mathbf{v})\rangle = \langle T(\mathbf{v}), \mathbf{w}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$

(ii) The matrix of T is symmetric with respect to every orthonormal basis of V

(iii) The matrix of T is symmetric with respect to some orthonormal basis of V

(iv) There is an orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of V such that $\langle \mathbf{f}_i, T(\mathbf{f}_j)\rangle = \langle T(\mathbf{f}_i), \mathbf{f}_j\rangle$ holds for all i and j.

Definition : Symmetric Linear Operator

A linear operator T on an inner product space V is called symmetric if $\langle \mathbf{v}, T(\mathbf{w})\rangle = \langle T(\mathbf{v}), \mathbf{w}\rangle$ holds for all $\mathbf{v}, \mathbf{w} \in V$.

Theorem

A symmetric linear operator on a finite dimensional inner product space has real eigenvalues.

Theorem

Let $T : V \to V$ be a symmetric linear operator on an inner product space V, and let U be a T-invariant subspace of V. Then:

(i) The restriction of T to U is a symmetric linear operator on U

(ii) $U^{\perp}$ is also T-invariant

Theorem : Principal Axes Theorem

The following conditions are equivalent for a linear operator T on a finite dimensional inner product space V.

(i) T is symmetric

(ii) V has an orthogonal basis consisting of eigenvectors of T

Definition : Distance Preserving

If V is an inner product space, a transformation $S : V \to V$ (not necessarily linear) is said to be distancing preserving if:

$$\|S(\mathbf{v}) - S(\mathbf{w})\| = \|\mathbf{v} - \mathbf{w}\| \quad \text{for all } \mathbf{v}, \mathbf{w} \in V$$

Lemma

Let V be an inner product space of dimension n, and consider a distance preserving transformation $S : V \to V$. If $S(\mathbf{0}) = \mathbf{0}$, then S is linear.

Definition : Isometries

Distance preserving linear operators are called isometries.

Theorem

If V is a finite dimensional inner product space, then every distance preserving transformation $S : V \to V$ is the composite of a translation and an isometry.

Theorem

Let $T : V \to V$ be a linear operator on a finite dimensional inner product space V. The following conditions are equivalent:

(i) T is an isometry (T preserves distance)

(ii) $\|T(\mathbf{v})\| = \|\mathbf{v}\| \ \forall \mathbf{v} \in V$ (T preserves norms)

(iii) $\langle T(\mathbf{v}), T(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle \ \forall \mathbf{v}, \mathbf{w} \in V$ (T preserves inner products)

(iv) If $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ is an orthonormal basis of V, then $\{T(\mathbf{f}_1), T(\mathbf{f}_2), \ldots, T(\mathbf{f}_n)\}$ is also an orthonormal basis (T preserves orthonormal bases)

(v) T carries some orthonormal basis to an orthonormal basis

Corollary

Let V be a finite dimensional inner product space.

(i) Every isometry of a finite dimensional V is an isomorphism

  (a) $1_V : V \to V$ is an isometry
  (b) The composite of two isometries of V is an isometry
  (c) The inverse of an isometry of V is an isometry

The conditions in (ii) assert that the set of isometries of a finite dimensional inner product space form a group.

Theorem

Let $T : V \to V$ be an operator where V is a finite dimensional inner product space. The following conditions are equivalent.

(i) T is an isometry

(ii) $M_B(T)$ is an orthogonal matrix for every orthonormal basis B

(iii) $M_B(T)$ is an orthogonal matrix for some orthonormal basis B

Corollary

If $T : V \to V$ is an isometry where V is a finite dimensional inner product space, then $\det T = \pm 1$.

Theorem

Let $T : V \to V$ be an isometry on the two-dimensional inner product space V. Then there are two possibilities:

(i) There is an orthonormal basis B of V such that

$$M_B(T) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, 0 \leq \theta < 2\pi$$

(ii) There is an orthonormal basis B of V such that

$$M_B(T) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Furthermore, type (i) occurs if and only if $\det T = 1$, and type (ii) occurs if and only if $\det T = -1$.

Corollary

An operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ is an isometry if and only if T is a rotation or a reflection.

Lemma

Let $T : V \to V$ be an isometry of a finite dimensional inner product space V. If U is a T-invariant subspace of V, then $U^\perp$ is also T-invariant.

Lemma

Let $T : V \to V$ be an isometry of the finite dimensional inner product space V. If λ is a complex eigenvalue of T, then $|\lambda| = 1$.

Lemma

Let $T : V \to V$ be an isometry of the n-dimensional inner product space V. If T has a non-real eigenvalue, then V has a two-dimensional T-invariant subspace.

Theorem

Let $T : V \to V$ be an isometry of the n-dimensional inner product space V. Given an angle θ, write $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Then there exists an orthonormal basis B of V such that $M_B(T)$ has one of the following block diagonal forms, classified for convenience by whether n is odd or even:

$$n = 2k+1 : \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix}$$

$$n = 2k : \begin{bmatrix} R(\theta_1) & 0 & \cdots & 0 \\ 0 & R(\theta_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R(\theta_{k-1}) \end{bmatrix}$$

Theorem

If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is an isometry, there are three possibilities.

(i) T is a rotation, and $M_B(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ for some orthonormal basis B.

(ii) T is a reflection, and $M_B(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ for some orthonormal basis B.

(iii) $T = QR = RQ$ where Q is a reflection, R is a rotation about an axis perpendicular to the fixed plane of Q and $M_B(T) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ for some orthonormal basis B.

Hence, T is a rotation if and only if $\det T = 1$.

Definition : Hyperplane

Let V be an n-dimensional inner product space. A subspace of V of dimension $n-1$ is called a hyperplane in V.

Definition : Fixed Hyperplanes & Reflection

Let V be an n-dimensional inner product space, and let $Q : V \to V$ be an isometry with matrix

$$M_B(Q) = \begin{bmatrix} -1 & 0 \\ 0 & I_{n-1} \end{bmatrix}$$

for some orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$. Then $Q(\mathbf{f}_1) = -\mathbf{f}_1$ whereas $Q(\mathbf{u}) = \mathbf{u}$ for each $\mathbf{u} \in U = \operatorname{span}\{\mathbf{f}_2, \ldots, \mathbf{f}_n\}$. Hence U is called the fixed hyperplane of Q and Q is called reflection in U. Note that each hyperplane in V is the fixed hyperplane of a (unique) reflection of V.

Theorem

Let $T : V \to V$ be an isometry of a finite dimensional inner product space V. Then there exist isometries $T_1, \ldots, T_k$ such that

$$T = T_k T_{k-1} \cdots T_2 T_1$$

where each T_i is either a rotation or a reflection, at most one is a reflection, and $T_i T_j = T_j T_i$ holds for all i and j. Furthermore, T is a composite of rotations if and only if $\det T = 1$.