Contents

Chapter 1

Some things to remember

1.1 Determinants

We extend the definition of the determinant to $n \times n$ matrices for $n \geq 3$. For this definition, it is convenient to introduce the following notation: Given $A \in \mathrm{M}_{n \times n}(F)$, for $n \geq 2$, denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j by $\tilde{A}_{ij}$. Thus for

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathrm{M}_{3 \times 3}(R),$$

we have

$$\tilde{A}_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, \quad \tilde{A}_{13} = \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}, \quad \text{and} \quad \tilde{A}_{32} = \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix},$$

and for

$$B = \begin{pmatrix} 1 & -1 & 2 & -1 \\ -3 & 4 & 1 & -1 \\ 2 & -5 & -3 & 8 \\ -2 & 6 & -4 & 1 \end{pmatrix} \in \mathrm{M}_{4 \times 4}(R),$$

we have

$$\tilde{B}_{23} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -5 & 8 \\ -2 & 6 & 1 \end{pmatrix} \quad \text{and} \quad \tilde{B}_{42} = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & -1 \\ 2 & -3 & 8 \end{pmatrix}.$$

Let $A \in M_{n \times n}(F)$. If $n = 1$, so that $A = (A_{11})$, we define $\det(A) = A_{11}$. For $n \geq 2$, we define $\det(A)$ recursively as
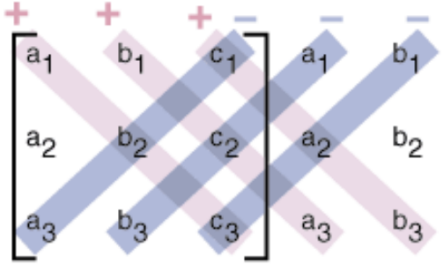
$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det\left(\tilde{A}_{1j}\right).$$

The scalar $\det(A)$ is called the determinant of A and is also denoted by $|A|$. The scalar

$$(-1)^{i+j} \det\left(\tilde{A}_{ij}\right)$$

is called the cofactor of the entry of A in row i, column j.

The determinant of a $2x2$ matrix A would be $det(A) = A_{11}A_{22} - A_{21}A_{12}$. Also, remember that the determinant of a $3x3$ matrix can be obtained from the following algorithm

$$\det A = (a_1 \, b_2 \, c_3 + b_1 \, c_2 \, a_3 + c_1 \, a_2 \, b_3) - (a_3 \, b_2 \, c_1 + b_3 \, c_2 \, a_1 + c_3 \, a_2 \, b_1)$$

Example 1.1.2 Let

$$A = \begin{pmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{pmatrix} \in M_{3\times 3}(R).$$

Using cofactor expansion along the first row of A, we obtain

$$\begin{aligned}
\det(A) =& (-1)^{1+1} A_{11} \cdot \det\left(\tilde{A}_{11}\right) + (-1)^{1+2} A_{12} \cdot \det\left(\tilde{A}_{12}\right) \\
& + (-1)^{1+3} A_{13} \cdot \det\left(\tilde{A}_{13}\right) \\
=& (-1)^2 (1) \cdot \det\begin{pmatrix} -5 & 2 \\ 4 & -6 \end{pmatrix} + (-1)^3 (3) \cdot \det\begin{pmatrix} -3 & 2 \\ -4 & -6 \end{pmatrix} \\
& + (-1)^4 (-3) \cdot \det\begin{pmatrix} -3 & -5 \\ -4 & 4 \end{pmatrix} \\
=& 1[-5(-6) - 2(4)] - 3[-3(-6) - 2(-4)] - 3[-3(4) - (-5)(-4)] \\
=& 1(22) - 3(26) - 3(-32) \\
=& 40.
\end{aligned}$$

Properties of determinants

1. If B is a matrix obtained by interchanging any two rows or interchanging any two columns of an $n \times n$ matrix A, then $\det(B) = -\det(A)$.

2. If B is a matrix obtained by multiplying each entry of some row or column of an $n \times n$ matrix A by a scalar k, then $\det(B) = k \cdot \det(A)$.

3. If B is a matrix obtained from an $n \times n$ matrix A by adding a multiple of row i to row j or a multiple of column i to column j for $i \neq j$, then $\det(B) = \det(A)$.

   As an example of the use of these three properties in evaluating determinants, let us compute the determinant of the 4×4 matrix A considered previously. Our procedure is to introduce zeros into the second column of A by employing property 3, and then to expand along that column. (The elementary row operations used here consist of adding multiples of row 1 to rows 2 and 4.) This procedure yields

   $$\det(A) = \det \begin{pmatrix} 2 & 1 & 1 & 5 \\ 1 & 1 & -4 & -1 \\ 2 & 0 & -3 & 1 \\ 3 & 6 & 1 & 2 \end{pmatrix} = \det \begin{pmatrix} 2 & 1 & 1 & 5 \\ -1 & 0 & -5 & -6 \\ 2 & 0 & -3 & 1 \\ -9 & 0 & -5 & -28 \end{pmatrix}$$

   $$= 1(-1)^{1+2} \det \begin{pmatrix} -1 & -5 & -6 \\ 2 & -3 & 1 \\ -9 & -5 & -28 \end{pmatrix}.$$

   The resulting determinant of a 3×3 matrix can be evaluated in the same manner: Use type 3 elementary row operations to introduce two zeros into the first column, and then expand along that column. This results in the value -102. Therefore

   $$\det(A) = 1(-1)^{1+2}(-102) = 102$$

4. The determinant of an upper triangular matrix is the product of its diagonal entries. In particular, $\det(I) = 1$.

5. If two rows (or columns) of a matrix are identical, then the determinant of the matrix is zero.

   As an illustration of property 4, notice that

   $$\det \begin{pmatrix} -3 & 1 & 2 \\ 0 & 4 & 5 \\ 0 & 0 & -6 \end{pmatrix} = (-3)(4)(-6) = 72.$$

6. For any $n \times n$ matrices A and B, $\det(AB) = \det(A) \cdot \det(B)$.

7. An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$.

8. For any $n \times n$ matrix A, the determinants of A and A^t are equal.

   For example, property 7 guarantees that the matrix A on page 233 is invertible because $\det(A) = 102$.

   The final property, stated as Exercise 15 of Section 4.3, is used in Chapter 5. It is a simple consequence of properties 6 and 7.

9. If A and B are similar matrices, then $\det(A) = \det(B)$.

1.2 Roots of a polynomial

The process of finding the roots of a polynomial can be either very straight forward or very complicated. Over the time we've developed multiple tools that help us in this task, we'll list a few of them to help us work with the characteristic polynomial. The most useful method for finding roots is to factorize the polynomials into factors of the form $(x - c)$ where c is a real number, or find quadratic factors to then apply the quadratic formula. To find these factors we can use the synthetic division algorithm.

Theorem 1.2.1

Fundamental Theorem Of Algebra. Every polynomial with complex coefficients has at least one complex zero.

Theorem 1.2.2

Roots theorem. Every polynomial of degree $n \geq 1$ has exactly n zeros, provided that a zero of multiplicity k is counted k times.

Theorem 1.2.3

Rational Zeros Theorem. If the polynomial $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ has integer coefficients, then every rational zero of P is of the form
$$\frac{p}{q}$$
where p is a factor of the constant coefficient a_0 and q is a factor of the leading coefficient a_n.

> **Note**
>
> ***Finding the Rational Zeros of a Polynomial.***
>
> 1. List Possible Zeros. List all possible rational zeros using the Rational Zeros Theorem.
>
> 2. Divide. Use synthetic division to evaluate the polynomial at each of the candidates for rational zeros that you found in Step 1. When the remainder is 0 , note the quotient you have obtained.
>
> 3. Repeat. Repeat Steps 1 and 2 for the quotient. Stop when you reach a quotient that is quadratic or factors easily, and use the quadratic formula or factor to find the remaining zeros.

Chapter 2

Eigenvalues and Eigenvectors

Definition 2.0.1

Eigenvector and Eigenvalues. Definitions. Let T be a linear operator on a vector space V. A nonzero vector $v \in$ V is called an eigenvector of T if there exists a scalar λ such that $T(v) = \lambda v$. The scalar λ is called the eigenvalue corresponding to the eigenvector v.

Let A be in $\mathrm{M}_{n \times n}(F)$. A nonzero vector $v \in \mathrm{F}^n$ is called an eigenvector of A if v is an eigenvector of L_A; that is, if $Av = \lambda v$ for some scalar λ. The scalar λ is called the eigenvalue of A corresponding to the eigenvector v.

Theorem 2.0.2

Let T be a linear operator over a finite dimensional vector space V. And let λ be a scalar, then the following are equivalent

1. λ is an eigenvalue of T

2. The operator $(T - \lambda I)$ is singular (it doesn't have inverse)

3. $\det(T - \lambda I) = 0$

Definition 2.0.3

If A is a matrix over a field F, an eigenvalue of A in F is a scalar in F, such that $(A - \lambda I)$ is singular.

> **Definition 2.0.4**
>
> Let A be an squared matrix. The function $P_A : \mathbb{R} \to \mathbb{R}$ where $P_A(\lambda) = \det(A - \lambda I)$ is called the characteristic polynomial of A.

With these theorems and definitions, we can obtain eigenvalues of a matrix by finding the roots of the characteristic polynomial and obtain the eigenvector v corresponding to an eigenvalue λ by solving the system $(A - \lambda I)v = 0$

Example 2.0.5 Let

$$A = \begin{bmatrix} -1 & -3 & 9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{bmatrix}$$

To find the eigenvalues, eigenvectors, and the characteristic polynomial, we start with

$$f(\lambda) = \det (A - \lambda I)$$

Then,

$$\det (A - \lambda I) = \det \left(\begin{bmatrix} -1 & -3 & 9 \\ 0 & 5 & 18 \\ 0 & -2 & -7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

$$= \det \left(\begin{bmatrix} -1 - \lambda & -3 & 9 \\ 0 & 5 - \lambda & 18 \\ 0 & -2 & -7 - \lambda \end{bmatrix} \right)$$

$$= -\lambda^3 - 3\lambda^2 - 3\lambda - 1$$

Now, to find the eigenvalues, we need to find the roots of the polynomial. It can be observed that $-\lambda^3 - 3\lambda^2 - 3\lambda - 1$ is in the expanded form of a cube of a binomial, therefore, the polynomial can be factorized as follows:

$$-(\lambda + 1)^3$$

Thus, by setting $-(\lambda + 1)^3 = 0$, we can conclude that $\lambda_1 = -1, \lambda_2 = -1, \lambda_3 = -1$. Now let's find the eigenvectors, for this, we will solve the following homogeneous system: $Ax = 0$, where $x = (x_1, x_2, x_3)$ and $x_1, x_2, x_3 \in F$. So, the homogeneous system for $\lambda = -1$ would be as follows:

$$\begin{bmatrix} 0 & -3 & 9 \\ 0 & 6 & 18 \\ 0 & -2 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving the system, we find that the vector

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

is a solution to the system and therefore is an eigenvector for A.

<div style="border:1px solid">

Definition 2.0.6

Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Define $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$. The set E_λ is called the eigenspace of T corresponding to the eigenvalue λ. Analogously, we define the eigenspace of a square matrix A corresponding to the eigenvalue λ to be the eigenspace of L_A corresponding to λ.

</div>

Continuing with the previous example, its eigenspace E_λ would be

$$E_\lambda = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

2.1 Working with the characteristic polynomial

It's worth to remember some tools to find the determinants of a matrix and the roots of a polynomial. Let's first remember the definition of determinant.

We'll first state a theorem that will help us to find some parts of the characteristic polynomial.

<div style="border:1px solid">

Theorem 2.1.1

Denote the n eigenvalues of a $A \in M_n(c)$, by $\lambda_1, \lambda_2, \ldots, \lambda_n$ and its characteristic polynomial by

$$P_a(\lambda) = (-1)^n \lambda^n + C_{n-1}\lambda^{n-1} + \cdots + C_1\lambda + C_0$$

Then

$$C_0 = \det(A) = \lambda_1\lambda_2\cdots\lambda_n$$

and

$$(-1)^{n-1}C_{n-1} = tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

</div>

Chapter 3

Diagonalization of a Matrix

3.1 Multiplicity

> **Definition 3.1.1**
>
> ***Geometrical multiplicity of an eigenvalue.*** Let A be a squared matrix with the eigenvalue λ. The geometrical multiplicity of λ is the dimension of E_λ (The space generated by λ).

For example 1.2.5, the geometrical multiplicity of $\lambda = -1$ is 1 since

$$E_\lambda = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

> **Definition 3.1.2**
>
> Let A be a squared matrix with the eigenvalue λ. The algebraic multiplicity of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of $f(t)$.

Example 3.1.3 Let

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}$$

which has characteristic polynomial $f(t) = -(t-3)^2(t-4)$. Hence $\lambda = 3$ is an eigenvalue of A with multiplicity of 2, and $\lambda = 4$ is an eigenvalue of A with multiplicity 1.

> **Theorem 3.1.4**
>
> Let T be a linear operation on a finite-dimensional vector space V. And let λ be an eigenvalue of T. Having algebraic multiplicity of m. Then $1 \leq dim(E_\lambda) \leq m$.

3.2 Similar matrices

> **Definition 3.2.1**
>
> **Similar matrices**. Two square matrices A and B are said to be similar if there exists a non-singular matrix P such that
>
> $$B = P^{-1}AP$$

> **Theorem 3.2.2**
>
> Let $A \in \mathrm{M}_{n \times n}(F)$, and let γ be an ordered basis for F^n. Then $[\mathrm{L}_A]_\gamma = Q^{-1}AQ$, where Q is the $n \times n$ matrix whose j th column is the j th vector of γ.

> **Theorem 3.2.3**
>
> Similar matrices have the same eigenvalues counting multiplicity.

3.3 Diagonalization of a Matrix

> **Definition 3.3.1**
>
> Let $\mathcal{T}$ be the set of all linear transformations from V into W, and let $L(V, W) = \{T \mid T \in \mathcal{T}\}$, then $L(V, W)$ is a vector space where its operations are defined as:
>
> - **Vector addition**. Let $T_1, T_2 \in L(V, W)$ then for every $v_1 \in V$
>
> $$(T_1 + T_2)(v_1) = T_1(v_1) + T_2(v_1)$$
>
> - **Scalar multiplication**. Let $T \in L(V, W)$ and let $\alpha \in F$, then for all $v_1 \in V$
>
> $$(\alpha T)(v_1) = \alpha T(v_1)$$

Definition 3.3.2

Matrix left multiplication. Let A be an $m \times n$ matrix with entries from a field F. We denote by L_A the mapping $\mathrm{L}_A : \mathrm{F}^n \to \mathrm{F}^m$ defined by $\mathrm{L}_A(x) = Ax$ (the matrix product of A and x) for each column vector $x \in \mathrm{F}^n$. We call L_A a left-multiplication transformation.

Definition 3.3.3

Diagonal matrix. A matrix is called diagonal if all its entries A_{ij} are zero except when $i = j$.

Definition 3.3.4

Diagonalization of a Matrix. A linear operator T on a finite-dimensional vector space V is called diagonalizable if there is an ordered basis β for V such that $[\mathrm{T}]_\beta$ is a diagonal matrix.
A square matrix A is called diagonalizable if L_A is diagonalizable. In other words we say that an squared matrix A is diagonalizable if and only if there exists an invertible matrix P such that

$$D = P^{-1}AP$$

or

$$A = PDP^{-1}$$

Theorem 3.3.5

If A is similar to a diagonal matrix D, then the elements of D are the eigenvalues of A; And the values are independent of the form in which a linear transformation is represented.

Theorem 3.3.6

A nxn matrix is diagonalizable if and only if admits a linearly independent eigenvectors.

Theorem 3.3.7

The following are equivalent

1. $T : V \to V$ is diagonalizable

2. For every eigenvalue λ of T, the geometric multiplicity of λ coincides with the algebra multiplicity.

3.4 Change of Basis

Definition 3.4.1

Ordered basis. Let V be a finite-dimensional vector space. An ordered basis for such space is endowed with a specific order, that is, an ordered basis for V is a sequence of linearly independent vectors in V that generate V.

Example 3.4.2 Let $V = F^3$.

$$\beta = \{e_1, e_2, e_3\}$$
$$\alpha = \{e_2, e_1, e_3\}$$

Then $\beta \neq \alpha$. And β, α are ordered basis.

Definition 3.4.3

Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for V. For each $v \in V$, let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the unique scalars such that

$$v = \sum_{i=1}^{n} \alpha_i u_i$$

We define the coordinate vector of v relative to β, denoted by

$$[v]_\beta = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Example 3.4.4 Let $P_2(\mathbb{R})$ and let $\beta = \{1, x, x^2\}$ and $\alpha = \{x^2, x, 1\}$ if $v = f(x) = 4 + 6x - 7x^2$.

$$[v]_\beta = \begin{pmatrix} 4 \\ 6 \\ -7 \end{pmatrix}$$

$$[v]_\alpha = \begin{pmatrix} -7 \\ 6 \\ 4 \end{pmatrix}$$

> **Theorem 3.4.5**
>
> Let $\beta = \{u_1, u_2, \ldots, u_n\}$ an ordered basis for V, then a coordinate map
>
> $$[]_\beta : V \to \mathbb{R}^n$$
> $$V \to [v]_\beta$$
>
> Is a linear transformation from $V \to \mathbb{R}^n$

> **Definition 3.4.6**
>
> We call mxn matrix A defined by $A_{ij} = a_{ij}$. The matrix representation of T in the ordered basis β and β' and write $A = [T]_\beta$.
>
> If $V = W$ and $\beta = \beta'$, then $A = [T]_\beta$.

Example 3.4.7 Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for V, and let T be a coordinate mapping such that $T(v) = [v]_\beta$. Then T is a linear transformation.

Proof. First, let's show that $T(v + w) = T(v) + T(w)$ for all $v, w \in V$.
Let $v, w \in V$. We have:

$$T(v) = [v]_\beta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where $\alpha_1, \ldots, \alpha_n$ are scalars such that

$$v = \sum_{i=1}^{n} \alpha_i u_i$$

and similarly,

$$T(w) = [w]_\beta = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix}$$

where $\gamma_1, \ldots, \gamma_n$ are scalars such that

$$w = \sum_{i=1}^{n} \gamma_i u_i$$

Therefore,

$$T(v + w) = [v + w]_\beta$$

$$= \left[\sum_{i=1}^{n} (\alpha_i + \gamma_i) u_i \right]_\beta$$

$$= \begin{bmatrix} \alpha_1 + \gamma_1 \\ \alpha_2 + \gamma_2 \\ \vdots \\ \alpha_n + \gamma_n \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix}$$

$$= T(v) + T(w)$$

Next, let's show that $T(cv) = cT(v)$ for all $c \in \mathbb{R}$ and all $v \in V$.
Let $v \in V$ and $c \in \mathbb{R}$. Then we have:

$$T(v) = [v]_\beta = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where $\alpha_1, \ldots, \alpha_n$ are scalars such that

$$v = \sum_{i=1}^{n} \alpha_i u_i$$

So,

$$T(cv) = [cv]_\beta$$

$$= \left[\sum_{i=1}^{n}(c\alpha_i)u_i\right]_\beta$$

$$= \begin{bmatrix} c\alpha_1 \\ c\alpha_2 \\ \vdots \\ c\alpha_n \end{bmatrix}$$

$$= c\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$= cT(v)$$

Thus, T is a linear transformation.

Example 3.4.8 Sea $T : \mathcal{M}_{2x2}(\mathbb{R}) \to P_2(\mathbb{R})$ tal que

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$$

y sean β_1 y β_2 bases de $\mathcal{M}_{2x2}(\mathbb{R})$ y $P_2(\mathbb{R})$ tal que

$$\beta_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$$\beta_2 = \left\{ 1, x, x^2 \right\}$$

Encuentre $[T]_{\beta_1}^{\beta_2}$.

Para encontrar $[T]_{\beta_1}^{\beta_2}$ debemos aplicar la transformación a cada vector de β_1 y luego obtener la matriz de la transformación. En este orden de ideas, para $v_j \in \beta_1$ tenemos lo siguiente:

$$T(v_1) = T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = (1+0) + 2 \cdot 0x + 0x^2 = 1$$

$$T(v_2) = T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = (0+1) + 2 \cdot 0x + 1x^2 = 1 + x^2$$

$$T(v_3) = T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = (0+0) + 2 \cdot 0x + 0x^2 = 0$$

$$T(v_4) = T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = (0+0) + 2 \cdot 1x + 0x^2 = 2x$$

Luego

$$[T]_{\beta_1}^{\beta_2} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Example 3.4.9 Let $T : \mathcal{M}_{2 \times 2}(\mathbb{R}) \to P_2(\mathbb{R})$ be such that

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right)$$

and let β_1 and β_2 be bases of $\mathcal{M}_{2 \times 2}(\mathbb{R})$ and $P_2(\mathbb{R})$ respectively, given by

$$\beta_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

$$\beta_2 = \left\{ 1, x, x^2 \right\}$$

Find $[T]_{\beta_1}^{\beta_2}$.

To find $[T]_{\beta_1}^{\beta_2}$, we must apply the transformation to each vector in β_1 and then obtain the matrix of the transformation. In this vein, for $v_j \in \beta_1$ we have the following:

$$T(v_1) = T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = (1 + 0) + 2 \cdot 0x + 0x^2 = 1$$

$$T(v_2) = T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) = (0 + 1) + 2 \cdot 0x + 1x^2 = 1 + x^2$$

$$T(v_3) = T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right) = (0 + 0) + 2 \cdot 0x + 0x^2 = 0$$

$$T(v_4) = T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) = (0 + 0) + 2 \cdot 1x + 0x^2 = 2x$$

Then

$$[T]_{\beta_1}^{\beta_2} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Let $F : \mathbb{R}^3 \to \mathbb{R}^2$ be defined as

$$F(x, y, z) = (3x + 2y - 4z, x - 5y + 3z)$$

1. Find the matrix of F for the following bases of $\mathbb{R}^3$ and $\mathbb{R}^2$

$$\beta_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$$
$$\beta_2 = \{(1, 3), (2, 5)\}$$

2. Given $V \in \mathbb{R}^3$, verify
$$[F]_{\beta_1}^{\beta_2}[v]_{\beta_1} = [F(v)]_{\beta_2}$$

1. By definition of F,
$$F((1,1,1)) = (3(1) + 2(1) - 4(1), 1 - 5(1) + 3(1)) = (1, -1)$$
$$F((1,1,0)) = (3(1) + 2(1) - 4(0), 1 - 5(1) + 3(0)) = (5, -4)$$
and
$$F((1,0,0)) = (3(1) + 2(0) - 4(0), 1 - 5(0) + 3(0)) = (3, 1)$$

Then
$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}_{\beta_2} = \begin{bmatrix} -7 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 5 \\ -4 \end{bmatrix}_{\beta_2} = \begin{bmatrix} -33 \\ 19 \end{bmatrix}$$
$$\begin{bmatrix} 3 \\ -1 \end{bmatrix}_{\beta_2} = \begin{bmatrix} -13 \\ 8 \end{bmatrix}$$

Finally,
$$[F]_{\beta_1}^{\beta_2} = \begin{bmatrix} -7 & -33 & -13 \\ 1 & 19 & 8 \end{bmatrix}$$

2. Let $v = (a, b, c)$. We have
$$[F(v)]_{\beta_2} = \begin{bmatrix} a \\ b \end{bmatrix}_{\beta_2} = \begin{bmatrix} \begin{bmatrix} 1 & 2 & 3a + 2b - 4c \\ 3 & 5 & a - 5b + 3c \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -13a - 20b + 26c \\ 8a + 11b - 15c \end{bmatrix}$$

Similarly, we have
$$[v]_{\beta_1} = \begin{bmatrix} c \\ b - c \\ a - b \end{bmatrix}$$

Finally
$$[F]_{\beta_1}^{\beta_2}[v]_{\beta_1} = \begin{bmatrix} -7 & -33 & -13 \\ 1 & 19 & 8 \end{bmatrix} \begin{bmatrix} c \\ b - c \\ a - b \end{bmatrix}$$
$$= \begin{bmatrix} -13a - 20b + 26c \\ 8a + 11b - 15c \end{bmatrix}$$
$$= [F(v)]_{\beta_2}$$

> **Theorem 3.4.10**
>
> If $T : V \to V$ is a linear operator and $\beta = \{\alpha_1, \dots, \alpha_n\}$ and $\beta' = \{\alpha'_1, \dots, \alpha'_n\}$ two ordered bases for V, there exists a matrix $P \in M_{nxn}$ invertible such that
>
> $$[T]_{\beta'} = P^{-1}[T]P$$

> **Theorem 3.4.11**
>
> Let V and W be finite dimensional vector space V with the ordered bases β and γ, respectively and let $T : V \to W$ be a linear transformation, for every $v \in V$, we have
>
> $$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta$$

> **Theorem 3.4.12**
>
> If $T : V \to W$ is a linar transformation, such that $T \in L(V, W)$, and $\beta = \{\alpha_1, \dots \alpha_n\}$ a base for V, and $\beta' = \{\alpha'_1, \dots \alpha'_n\}$ be a base for W, then there exists a unique matrix $A \in M_{nxn}$ such that for each $x \in V$
>
> $$[T(x)]_\beta = A[x]_\beta$$

3.5 Sum of subspaces

> **Definition 3.5.1**
>
> ***Sum of sets***. Suppose that $U_1, \dots U_m$ are subsets of a vector space V. Then the sum of $U_1, \dots, U_m$ denoted by $U_1 + \dots + U_m$ is the set of all possible sums of the elements of $U_1, \dots, U_m$, more precisely
>
> $$U_1 + \dots U_m = \{u_1 + \dots u_m \mid u_1 \in U_1, \dots, u_m \in U_m\}$$

Example 3.5.2
$$U = \left\{(x, 0, 0) \in F^3 \mid x \in F\right\}$$
$$W = \left\{(0, y, 0) \in F^3 \mid x \in F\right\}$$

Then

$$U + W = \left\{(x, y, 0) \in F^3 \mid x, y \in F\right\}$$

Example 3.5.3
$$W_1 = \{m(1, 0) \mid m \in \mathbb{R}\}$$

$$W_2 = \{l(0,1) \mid l \in \mathbb{R}\}$$

Then

$$W_1 + W_2 = \{m(1,0) + l(1,0) \mid m, l \in \mathbb{R}\}$$

> **Theorem 3.5.4**
>
> ***The sum of subspaces is also a subspace.*** Let $V_1, \ldots, V_n$ subspaces of a vector space V, then
>
> $$V_1 + \cdots V_n$$

Chapter 4

Normed vector spaces

4.1 Inner products

> **Definition 4.1.1**
>
> Let V be a vector space over a field F. An inner product on V is a function that assigns to every ordered pair of vectors x and y in V a scalar in F, denoted $\langle x, y \rangle$, such that for all x, y, z in V and all c in F, the following hold:
> (a) $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$.
> (b) $\langle cx, y \rangle = c\langle x, y \rangle$.
> (c) $\overline{\langle x, y \rangle} = \langle y, x \rangle$, where the bar denotes complex conjugation.
> (d) If $x \neq 0$, then $\langle x, x \rangle$ is a positive real number.

This definition specifies the properties that an inner product must satisfy on a vector space, generalizing concepts like those of dot product in $\mathbb{R}^n$.

Example 4.1.2 For $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ in F^n, define

$$\langle x, y \rangle = \sum_{i=1}^{n} a_i \bar{b}_i$$

Thus, for $x = (1 + i, 4)$ and $y = (2 - 3i, 4 + 5i)$ in C^2,

$$\langle x, y \rangle = (1 + i)(2 + 3i) + 4(4 - 5i) = 15 - 15i$$

The inner product in this example is called the standard inner product on F^n.

This example shows how the standard inner product is calculated in F^n, demonstrating that it satisfies the properties of an inner product.

Example 4.1.3 If $\langle x, y \rangle$ is any inner product on a vector space V and $r > 0$, define another inner product by $\langle x, y \rangle' = r\langle x, y \rangle$. If $r \leq 0$, then condition (d) of the definition would not hold.

Shows how scaling an inner product by a positive scalar results in another valid inner product.

Example 4.1.4 Let $V = C([0,1])$, the vector space of real-valued continuous functions on $[0,1]$. For $f, g \in$ V, define $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$.

Inner product defined in terms of an integral, demonstrating application in function spaces.

Definition 4.1.5

Let $A \in \mathrm{M}_{m \times n}(F)$. We define the conjugate transpose or adjoint of A to be the $n \times m$ matrix A^* such that $(A^*)_{ij} = \bar{A}_{ji}$ for all i, j.

Defines the important concept of conjugate transpose, which is crucial in various linear algebra and functional analysis contexts.

Example 4.1.6 Let

$$A = \begin{pmatrix} i & 1 + 2i \\ 2 & 3 + 4i \end{pmatrix}$$

Then

$$A^* = \begin{pmatrix} -i & 2 \\ 1 - 2i & 3 - 4i \end{pmatrix}$$

Provides a concrete example of calculating the conjugate transpose of a matrix with complex entries.

Example 4.1.7 Let V $= \mathrm{M}_{n \times n}(F)$, and define $\langle A, B \rangle = \mathrm{tr}\,(B^*A)$ for $A, B \in$ V.

Defines an inner product on the vector space of matrices using the trace function, known as the Frobenius inner product.

Definition 4.1.8

Let V be an inner product space. For $x \in$ V, we define the norm or length of x by $\|x\| = \sqrt{\langle x, x \rangle}$.

Defines the norm of a vector in an inner product space, generalizing the concept of vector length.

> **Theorem 4.1.9**
>
> Theorem 6.2: Let V be an inner product space over F. Then for all $x, y \in$ V and $c \in F$, the following statements are true.
> (a) $\|cx\| = |c| \cdot \|x\|$.
> (b) $\|x\| = 0$ if and only if $x = 0$. In any case, $\|x\| \geq 0$.
> (c) (Cauchy-Schwarz Inequality) $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$.
> (d) (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$.

This theorem establishes key properties and inequalities for norms and inner products, including the important Cauchy-Schwarz and Triangle inequalities.

Example 4.1.10 For F^n, we may apply (c) and (d) of Theorem 6.2 to the standard inner product to obtain well-known inequalities:

$$\left| \sum_{i=1}^{n} a_i \bar{b}_i \right| \leq \left[\sum_{i=1}^{n} |a_i|^2 \right]^{1/2} \left[\sum_{i=1}^{n} |b_i|^2 \right]^{1/2}$$

and

$$\left[\sum_{i=1}^{n} |a_i + b_i|^2 \right]^{1/2} \leq \left[\sum_{i=1}^{n} |a_i|^2 \right]^{1/2} + \left[\sum_{i=1}^{n} |b_i|^2 \right]^{1/2}$$

Illustrates the application of Cauchy-Schwarz and Triangle inequalities in Euclidean space with standard inner products.

> **Definition 4.1.11**
>
> Let V be an inner product space. Vectors x and y in V are orthogonal (or perpendicular) if $\langle x, y \rangle = 0$. A subset S of V is orthogonal if any two distinct vectors in S are orthogonal. A vector x in V is a unit vector if $\|x\| = 1$. Finally, a subset S of V is orthonormal if S is orthogonal and consists entirely of unit vectors.

Defines concepts of orthogonality, unit vectors, and orthonormal sets in the context of inner product spaces.

Example 4.1.12 In F^3, $\{(1, 1, 0), (1, -1, 1), (-1, 1, 2)\}$ is an orthogonal set of nonzero vectors, but it is not orthonormal. If we normalize the vectors, we get the orthonormal set:

$$\left\{ \frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{3}}(1, -1, 1), \frac{1}{\sqrt{6}}(-1, 1, 2) \right\}$$

Shows how to normalize a set of orthogonal vectors to form an orthonormal set.

Example 4.1.13 Recall the inner product space H. For any integer n, let $f_n(t) = e^{int}$, where $0 \leq t \leq 2\pi$. Define $S = \{f_n : n$ is an integer $\}$. Clearly S is a subset of H. For $m \neq n$,

$$\langle f_m, f_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{imt} \overline{e^{int}} dt = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)t} dt = 0$$

Introduces an important orthonormal set in the context of continuous complex-valued functions on an interval, used in further examples and applications in analysis.

Definition 4.1.14

A linear space over the field $\mathbf{F}$ of scalars is a set V satisfying:

1. V is closed under vector addition.

2. Vector addition is commutative and associative.

3. There is a zero element in V such that $v + 0 = v$ for all v in V.

4. For each v in V, there is an additive inverse $-v$ such that $v + (-v) = 0$.

5. V is closed under scalar multiplication.

6. Scalar multiplication is associative and distributive.

7. $1v = v$ for all v in V.

Example 4.1.15 $\mathbf{R}^n$ with the usual operations is a vector space over $\mathbf{R}$.

Example 4.1.16 $\mathbf{C}^n$ with the usual operations is a vector space over $\mathbf{C}$.

Example 4.1.17 The set $C[a, b]$ of $\mathbf{F}$-valued continuous functions on $[a, b]$ is a linear space over $\mathbf{F}$.

Example 4.1.18 The set $C^k[a, b]$ of $\mathbf{F}$-valued k-times continuously differentiable functions on $[a, b]$ is a linear space over $\mathbf{F}$.

Example 4.1.19 The set $L^1(B)$ of functions $f : \mathbf{R}^n \to \mathbf{F}$ satisfying $\int_B |f(x)| dx < \infty$ is a linear space over $\mathbf{F}$.

Example 4.1.20 The set $L^2(B)$ of functions $f : \mathbf{R}^n \to \mathbf{F}$ satisfying $\int_B |f(x)|^2 dx < \infty$ is a linear space over $\mathbf{F}$.

Definition 4.1.21

Let V be a linear space. If $U \subseteq V$ is closed under vector addition and scalar multiplication, then U is a subspace of V. A subspace is itself a linear space.

Example 4.1.22 Let $V = \mathbf{R}^3$. If U is a subspace of V, then either

1. $U = \mathbf{R}^3$,

2. U is a plane through the origin,

3. U is a line through the origin,

4. $U = \{0\}$.

Example 4.1.23 Let $[a, b]$ be a finite interval. The set $C[a, b]$ is a subspace of $L^1[a, b]$.

Definition 4.1.24

A norm $\| \cdot \|$ on a linear space V is a mapping from V to $\mathbf{R}$ satisfying:

1. $\|v\| \geq 0$ for all $v \in V$;

2. $\|v\| = 0$ if and only if $v = 0$;

3. $\|\alpha v\| = |\alpha| \|v\|$ for all $\alpha \in \mathbf{C}$ and $v \in V$;

4. (Triangle inequality) $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in V$.

Example 4.1.25 $\mathbf{F}^n$ is a normed linear space with $\|z\| = \{|z_1|^2 + \cdots + |z_n|^2\}^{1/2}$.

Example 4.1.26 $C[a, b]$ is a normed linear space with the maximum (or L^∞) norm $\|f\| = \max_{[a,b]} |f(x)|$.

Example 4.1.27 $C^k[a, b]$ is a normed linear space with $\|f\| = \sum_{j=1}^{k} \max_{[a,b]} |f^{(j)}(x)|$.

Example 4.1.28 $L^1(B)$ is a normed linear space with $\|f\| = \int_B |f(x)| dx$.

Example 4.1.29 $L^2(B)$ is a normed linear space with $\|f\| = \{\int_B |f(x)|^2 dx\}^{1/2}$.

Definition 4.1.30

A sequence $\{v_k\}$ of vectors in a normed linear space V is convergent if there is a $v \in V$ such that $\|v_k - v\| \to 0$ as $k \to \infty$.

Definition 4.1.31

A sequence $\{v_k\}$ of vectors in a normed linear space V is Cauchy convergent if $\|v_m - v_n\| \to 0$ as $m, n \to \infty$.

Definition 4.1.32

A normed linear space is complete if all Cauchy convergent sequences are convergent. A complete normed linear space is called a Banach space.

Example 4.1.33 $C[a, b], C^k[a, b], L^1(B)$ and $L^2(B)$ are all Banach spaces with respect to the given norms.

Example 4.1.34 Let V be the set $C[0, 2]$ of real-valued functions with norm $\|f\| = \int_0^2 |f(x)|dx$. Although V is a normed linear space, it is not a Banach space. The sequence $\{f_k\}$ with $f_k(x) = x^k$ for $0 \le x < 1$ and 1 for $1 \le x \le 2$ is Cauchy convergent in V but does not converge to a function in V.

Definition 4.1.35

An inner product on a linear space V over $\mathbf{F}$ is a mapping $\langle \cdot, \cdot \rangle$ from $V \times V$ to $\mathbf{F}$ satisfying:

1. $\langle v, v \rangle \ge 0$ for all $v \in V$;

2. $\langle v, v \rangle = 0$ if and only if $v = 0$;

3. $\langle u, v \rangle = \langle v, u \rangle^*$ for all $u, v \in V$;

4. $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ for all $\alpha \in \mathbf{F}$ and $u, v \in V$;

5. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$.

Example 4.1.36 $\mathbf{F}^n$ is an inner product space: For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbf{F}^n$, $\langle x, y \rangle = x_1 y_1^* + \cdots + x_n y_n^*$. When $\mathbf{F} = \mathbf{R}$, this reduces to the usual dot product on $\mathbf{R}^n$.

Example 4.1.37 $L^2(B)$ is an inner product space with $\langle f, g \rangle = \int_B f(x)g(x)^* dx$.

Definition 4.1.38

An inner product space that is complete with respect to the norm $\|v\| = \sqrt{\langle v, v \rangle}$ is called a Hilbert space.

Example 4.1.39 $\mathbf{C}^n$ and $L^2(B)$ are Hilbert spaces with the given inner products.

4.2 Orthagonalization

Definition 4.2.1

Let V be an inner product space. A subset of V is an orthonormal basis for V if it is an ordered basis that is orthonormal.

Example 4.2.2 The standard ordered basis for $\mathbf{F}^n$ is an orthonormal basis for $\mathbf{F}^n$.

Example 4.2.3 The set

$$\left\{ \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right), \left(\frac{2}{\sqrt{5}}, \frac{-1}{\sqrt{5}} \right) \right\}$$

is an orthonormal basis for R^2.

Theorem 4.2.4

Theorem 6.3. Let V be an inner product space and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of nonzero vectors. If $y \in \mathrm{span}(S)$, then

$$y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{\|v_i\|^2} v_i$$

Theorem 4.2.5

Corollary 1. If, in addition to the hypotheses of Theorem 6.3, S is orthonormal and $y \in \mathrm{span}(S)$, then

$$y = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$$

Theorem 4.2.6

Corollary 2. Let V be an inner product space, and let S be an orthogonal subset of V consisting of nonzero vectors. Then S is linearly independent.

Example 4.2.7 By Corollary 2, the orthonormal set

$$\left\{ \frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{3}}(1, -1, 1), \frac{1}{\sqrt{6}}(-1, 1, 2) \right\}$$

obtained in Example 8 of Section 6.1 is an orthonormal basis for R^3. Let $x = (2, 1, 3)$. The coefficients given by Corollary 1 to Theorem 6.3 that express x as a linear combination of the basis vectors are:

$$a_1 = \frac{1}{\sqrt{2}}(2 + 1) = \frac{3}{\sqrt{2}}, \quad a_2 = \frac{1}{\sqrt{3}}(2 - 1 + 3) = \frac{4}{\sqrt{3}}$$

and

$$a_3 = \frac{1}{\sqrt{6}}(-2 + 1 + 6) = \frac{5}{\sqrt{6}}$$

Checking:

$$(2, 1, 3) = \frac{3}{2}(1, 1, 0) + \frac{4}{3}(1, -1, 1) + \frac{5}{6}(-1, 1, 2)$$

27

> **Theorem 4.2.8**
>
> **Theorem 6.4.** Let V be an inner product space and $S = \{w_1, w_2, \ldots, w_n\}$ be a linearly independent subset of V. Define $S' = \{v_1, v_2, \ldots, v_n\}$, where $v_1 = w_1$ and
>
> $$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j \quad \text{for } 2 \le k \le n$$
>
> Then S' is an orthogonal set of nonzero vectors such that $\text{span}(S') = \text{span}(S)$.

Example 4.2.9 In R^4, let $w_1 = (1, 0, 1, 0)$, $w_2 = (1, 1, 1, 1)$, and $w_3 = (0, 1, 2, 1)$. We use the Gram-Schmidt process to compute the orthogonal vectors v_1, v_2, and v_3. Take $v_1 = w_1 = (1, 0, 1, 0)$. Then

$$v_2 = (0, 1, 0, 1)$$

and

$$v_3 = (-1, 0, 1, 0)$$

Normalizing, we get:

$$u_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0), \quad u_2 = \frac{1}{\sqrt{2}}(0, 1, 0, 1), \quad u_3 = \frac{1}{\sqrt{2}}(-1, 0, 1, 0)$$

Example 4.2.10 Let $V = P(R)$ with the inner product $\langle f(x), g(x) \rangle = \int_{-1}^{1} f(t)g(t)dt$. Consider $P_2(R)$ with standard ordered basis β. We use Gram-Schmidt to obtain an orthonormal basis:

$$\left\{ 1, x, x^2 - \frac{1}{3} \right\} \to \left\{ \frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2 - 1) \right\}$$

> **Definition 4.2.11**
>
> Let β be an orthonormal subset of an inner product space V, and let $x \in$ V. We define the Fourier coefficients of x relative to β to be the scalars $\langle x, y \rangle$, where $y \in \beta$.

Example 4.2.12 Compute the Fourier coefficients of $f(t) = t$ relative to $S = \{ e^{int} : n \text{ is an integer} \}$.

$$\langle f, f_n \rangle = \frac{-1}{in} \text{ for } n \ne 0,$$

$$\langle f, 1 \rangle = \pi \text{ for } n = 0$$

Definition 4.2.13

Let S be a nonempty subset of an inner product space V. We define $S^\perp$ (read "S perp") to be the set of all vectors in V that are orthogonal to every vector in S; that is, $S^\perp = \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in S\}$. The set $S^\perp$ is called the orthogonal complement of S.

Example 4.2.14 Verify that $\{0\}^\perp = V$ and $V^\perp = \{0\}$ for any inner product space V.

Example 4.2.15 If $V = R^3$ and $S = \{e_3\}$, then $S^\perp$ equals the xy-plane.

Theorem 4.2.16

Theorem 6.6. Let W be a finite-dimensional subspace of an inner product space V, and let $y \in V$. Then there exist unique vectors $u \in W$ and $z \in W^\perp$ such that $y = u + z$. Furthermore, if $\{v_1, v_2, \ldots, v_k\}$ is an orthonormal basis for W, then

$$u = \sum_{i=1}^{k} \langle y, v_i \rangle \, v_i$$

Example 4.2.17 Let $V = P_3(R)$ with the inner product

$$\langle f(x), g(x) \rangle = \int_{-1}^{1} f(t)g(t)dt \quad \text{for all } f(x), g(x) \in V.$$

Chapter 5

Dual spaces

5.1 Notes

Definition 5.1.1

A linear functional on V is a linear map $V \to \mathbb{F}$. In other words, a linear functional on V is an element of $\mathcal{L}(V, \mathbb{F})$.

Definition 5.1.2

The dual space of V, denoted by V^*, is the space of all linear functionals on V; i.e. $V^* := \mathcal{L}(V, \mathbb{F})$.

Definition 5.1.3

Let $T : V \to W$ be linear. The dual map (or transpose) of T is the map $T^* : W^* \to V^*$ defined by $T^*g = gT$ for all $g \in W^*$.

Definition 5.1.4

Let U be a subspace of V. The annihilator of U in V^*, denoted by $U^\perp$, is the set of linear functionals on V which vanish on U; i.e. $U^\perp = \{f \in V^* \mid f(u) = 0 \text{ for all } u \in U\}$.

Theorem 5.1.5

Suppose that V is finite-dimensional and let $(v_1, \ldots, v_n)$ be a basis of V. For each $i = 1, \ldots, n$, define a linear functional $f_i : V \to \mathbb{F}$ by setting

$$f_i(v_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

and then extending f_i linearly to all of V. Then $(f_1, \ldots, f_n)$ is a basis of V^*, called the dual basis of $(v_1, \ldots, v_n)$. Hence, V^* is finite-dimensional and $\dim V^* = \dim V$.

Theorem 5.1.6

Suppose that V is finite-dimensional. The map $\mathrm{ev} : V \to V^{**}$ defined by

$$\mathrm{ev}(v)(f) := f(v)$$

is an isomorphism.

Theorem 5.1.7

Let V be an inner product space. Define a map $T : V \to V^*$ by

$$Tv = \langle v, \cdot \rangle$$

i.e. Tv is the linear functional on V whose value on $w \in V$ is $\langle v, w \rangle$. Then T is an isomorphism.

Theorem 5.1.8

Suppose that V is finite-dimensional and let $(v_1, \ldots, v_n)$ be a basis of V. Let $(f_1, \ldots, f_n)$ be the corresponding dual basis of V^*. For $T \in \mathcal{L}(V)$, the matrix of $T^* \in \mathcal{L}(V^*)$ with respect to $(f_1, \ldots, f_n)$ is the transpose of the matrix of T with respect to $(v_1, \ldots, v_n)$; i.e.

$$\mathcal{M}(T^*) = \mathcal{M}(T)^t$$

Theorem 5.1.9

There exists a natural isomorphism between $U^\perp$ and $(V/U)^*$. Hence we can identify linear functionals on V/U with elements of $U^\perp$.

Example 5.1.10 A force in physics is usually represented by a vector F. The main interest is what a force does when moving an object from point A to point B. If q is the position vector of the object, then F acts on it to produce a number

called "work"; that is, F can be thought of as a linear functional

$$F : q \mapsto F(q)$$

Thus, from this point of view, force should actually be thought of as an element of the dual space since it took a vector and produced a scalar.

5.2 Book

> **Definition 5.2.1**
>
> A linear transformation from a vector space V into its field of scalars F is called a *linear functional* on V.

Example 5.2.2 Let V be the vector space of continuous real-valued functions on the interval $[0, 2\pi]$. Fix a function $g \in V$. The function $h : V \to \mathbb{R}$ defined by

$$h(x) = \frac{1}{2\pi} \int_0^{2\pi} x(t)g(t)\, dt$$

is a linear functional on V. In the cases that $g(t)$ equals $\sin(nt)$ or $\cos(nt)$, $h(x)$ is often called the n-th Fourier coefficient of x.

Example 5.2.3 Let $V = M_{n \times n}(F)$, and define $f : V \to F$ by $f(A) = \text{tr}(A)$, the trace of A. Then f is a linear functional.

Example 5.2.4 Let V be a finite-dimensional vector space, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. For each $i = 1, 2, \ldots, n$, define $f_i(x) = a_i$, where

$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ is the coordinate vector of x relative to β. Then f_i is a linear

functional on V called the i-th coordinate function with respect to the basis β. Note that $f_i(x_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta.

> **Definition 5.2.5**
>
> For a vector space V over F, the *dual space* of V, denoted by V^*, is defined to be the vector space $\mathcal{L}(V, F)$, consisting of all linear functionals on V.

Theorem 5.2.6

Theorem 2.24: Suppose that V is a finite-dimensional vector space with the ordered basis $\beta = \{x_1, x_2, \ldots, x_n\}$. Let $f_i(1 \leq i \leq n)$ be the i-th coordinate function with respect to β, and let $\beta^* = \{f_1, f_2, \ldots, f_n\}$. Then β^* is an ordered basis for V^*, and for any $f \in V^*$,

$$f = \sum_{i=1}^{n} f(x_i) f_i$$

Definition 5.2.7

The ordered basis $\beta^* = \{f_1, f_2, \ldots, f_n\}$ of V^* that satisfies $f_i(x_j) = \delta_{ij} (1 \leq i, j \leq n)$ is called the *dual basis* of β.

Example 5.2.8 Let $\beta = \{(2, 1), (3, 1)\}$ be an ordered basis for $\mathbb{R}^2$. Suppose that the dual basis of β is given by $\beta^* = \{f_1, f_2\}$. To explicitly determine a formula for f_1, consider the equations

$$1 = f_1(2, 1) = f_1(2e_1 + e_2) = 2f_1(e_1) + f_1(e_2),$$
$$0 = f_1(3, 1) = f_1(3e_1 + e_2) = 3f_1(e_1) + f_1(e_2).$$

Solving these equations, we obtain $f_1(e_1) = -1$ and $f_1(e_2) = 3$; that is, $f_1(x, y) = -x + 3y$. Similarly, it can be shown that $f_2(x, y) = x - 2y$.

Theorem 5.2.9

Theorem 2.25: Let V and W be finite-dimensional vector spaces over F with ordered bases β and γ, respectively. For any linear transformation $T : V \to W$, the mapping $T^t : W^* \to V^*$ defined by $T^t(g) = gT$ for all $g \in W^*$ is a linear transformation with the property that $[T^t]_{\gamma^*}^{\beta^*} = \left([T]_\beta^\gamma\right)^t$.

Example 5.2.10 Define $T : P_1(\mathbb{R}) \to \mathbb{R}^2$ by $T(p(x)) = (p(0), p(2))$. Let β and γ be the standard ordered bases for $P_1(\mathbb{R})$ and $\mathbb{R}^2$, respectively. Clearly,

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}.$$

We compute $[T^t]_{\gamma^*}^{\beta^*}$ directly from the definition. Let $\beta^* = \{f_1, f_2\}$ and $\gamma^* = \{g_1, g_2\}$. Suppose that $[T^t]_{\gamma^*}^{\beta^*} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $T^t(g_1) = af_1 + cf_2$. So,

$$T^t(g_1)(1) = (af_1 + cf_2)(1) = af_1(1) + cf_2(1) = a(1) + c(0) = a.$$

But also,
$$(T^t(g_1))(1) = g_1(T(1)) = g_1(1, 1) = 1.$$

So $a = 1$. Similarly, $c = 0$, $b = 1$, and $d = 2$. Hence, a direct computation yields

$$\left[T^t\right]_{\gamma^*}^{\beta^*} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \left([T]_\beta^\gamma\right)^t$$

as predicted by Theorem 2.25.

Theorem 5.2.11

Theorem 2.26: Let V be a finite-dimensional vector space, and define $\psi : V \to V^{**}$ by $\psi(x) = \widehat{x}$. Then ψ is an isomorphism.

Theorem 5.2.12

Lemma: Let V be a finite-dimensional vector space, and let $x \in V$. If $\widehat{x}(f) = 0$ for all $f \in V^*$, then $x = 0$.

Theorem 5.2.13

Corollary: Let V be a finite-dimensional vector space with dual space V^*. Then every ordered basis for V^* is the dual basis for some basis for V.