Complex n-Space C^n, Complex Matrices, Spectral Theorem

COMPLEX NUMBERS

Definition

$\sqrt{-1} = i$ is the imaginary unit.

Definition

$a+bi, a, b \in \mathbf{R}$ is a complex number ($\mathbf{C}$). a is the real part, b is the imaginary part.

Properties of a Complex Number

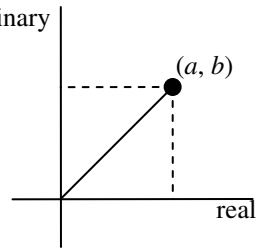
1) $(a+bi) = (a'+b'i)$ iff $a = a'$ and $b = b'$.
2) $(a+bi)+(a'+b'i) = (a+a')+(b+b')i$.
3) $(a+bi)-(a'+b'i) = (a-a')+(b-b')i$.
4) $(a+bi)(a'+b'i) = (aa'-bb')+(ab'+ba')i$.
5) $\dfrac{a+bi}{a'+b'i} = \dfrac{(aa'+bb')}{(a')^2+(b')^2} + \dfrac{(a'b-ab')}{(a')^2+(b')^2}i$.
6) $(a+bi)$ is a real number iff $b = 0$ ($\mathbf{R} \subset \mathbf{C}$).

Definition

Let us denote $z = a+bi$. The conjugate of z is $\bar{z} = a-bi$. The absolute value of z is $|z| - \sqrt{a^2+b^2}$.

Geometric Interpretation of a Complex Number

$z = a+bi \Leftrightarrow \mathbf{R}^2(a,b)$



- The absolute value $|z| = \sqrt{a^2+b^2}$ is just the distance from (a,b) to the origin.
- $|z_1 - z_2|$ is the distance from $z_1 = a_1 + b_1 i$ to $z_2 = a_2 + b_2 i$.

Polar Coordinates

- $\left.\begin{aligned} a &= r\cos\theta \\ b &= r\sin\theta \end{aligned}\right\} \Rightarrow z = a+bi = r\cos\theta + (r\sin\theta)i = r(\cos\theta + i\sin\theta) = r\cdot e^{i\theta}$.
  - Note: $r = |z| = \sqrt{a^2+b^2}$.
  - Note: $\theta = \arctan\dfrac{b}{a}$ the argument of z.

Definition

$e^{i\theta} = \cos\theta + (\sin\theta)i$.

Example

Write $z = -2 + 2i$ in polar form.

- Let $z = re^{i\theta}$.
  - $r = \sqrt{(-2)^2 + (2)^2} = \sqrt{8} = 2\sqrt{2}$.
  - $\theta = \arctan\left(-\frac{2}{2}\right) = \arctan(-1) = \frac{3\pi}{4}$.
- So $z = -2 + 2i = 2\sqrt{2} \cdot e^{i\frac{3\pi}{4}}$.

Theorem: Multiplication In Polar Coordinates

If $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_{21}}$ are complex numbers in polar form, then $z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \vartheta_2)}$.

Proof:

- $z_1 z_2 = \left(r_1 e^{i\theta_1}\right)\left(r_2 e^{i\theta_2}\right) = r_1 r_2 \cdot e^{i\theta_1} e^{i\theta_2}$.
- Want: $e^{i\theta_1} e^{i\theta_2} = e^{i(\theta_1 + \theta_2)}$.
- $$\begin{aligned} e^{i\theta_1} e^{i\theta_2} &= (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \\ &= (\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2) + (\sin\theta_1 \cos\theta_2 + \cos\theta_1 \sin\theta_2)i \\ &= \cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2) \\ &= e^{i(\theta_1 + \theta_2)} \end{aligned}$$.

Theorem: De Moivre's Theorem

If θ is any angle, then $\left(e^{i\theta}\right)^n = e^{i(n\theta)}$ holds for all integers n.

Proof (sketch):

- If $n \geq 0$, use induction.
- If $n < 0$, then $-n > 0$. So $\left(e^{i\theta}\right)^n = \left(\left(e^{i\theta}\right)^{-1}\right)^{-n}$. Since $\left(e^{i\theta}\right)^{-1} = \frac{1}{e^{i\theta}} = \frac{1}{\cos\theta + i\sin\theta} \cdot \frac{\cos\theta - i\sin\theta}{\cos\theta - i\sin\theta} = \cos\theta - i\sin\theta = \cos(-\theta) + i\sin(-\theta) = e^{i(-\theta)}$, so $\left(e^{i\theta}\right)^n = \left(e^{i(-\theta)}\right)^{-n}, n < 0$.

Example

Find $\left(-1 + \sqrt{3}i\right)^3$.

- Let $-1 + \sqrt{3}i = re^{i\theta} = 2e^{i\frac{2\pi}{3}}$.
  - $r = \sqrt{(-1)^2 + \left(\sqrt{3}\right)^2} = 2$.
  - $\theta = \arctan\left(-\sqrt{3}\right) = \frac{2\pi}{3}$.

- So $\left(-1+\sqrt{3}i\right)^3 = \left(2e^{i\frac{2\pi}{3}}\right)^3 = 8e^{i2\pi} = 8$.

Theorem: The n^{th} Root of Unity

If $n \geq 1$ is an integer, the n^{th} root of unity (the complex numbers z such that $z^n = 1$) are given by $z = e^{i2\pi\frac{k}{n}}, k = 0,1,\cdots,n-1$.

Proof:

- Let us denote $z = re^{i\theta}$.
- We want r and θ such that
$$r^n e^{in\theta} = 1 \Rightarrow r^n\left(\cos(n\theta)+i\sin(n\theta)\right) = 1 \Rightarrow \begin{cases} r^n\cos(n\theta) = 1 \Rightarrow r^n = 1 \Rightarrow r = 1 \\ r^n\sin(n\theta) = 0 \Leftrightarrow \theta = \dfrac{2\pi k}{n} \end{cases}.$$
- When $k = 0$, $\theta = 0$, and when $k = n$, $\theta = 2\pi = 0$. So $k = 0,\cdots,n-1$.

POLYNOMIALS

- Notice that the roots of ax^2+bx+c are $x_1 = \frac{-b}{2a}+\frac{\sqrt{4ac-b^2}}{2a}i$ and $x_2 = \frac{-b}{2a}-\frac{\sqrt{4ac-b^2}}{2a}i$. x_2 is the conjugate of x_1, that is $x_2 = \overline{x_1}$.

Complex Polynomials

Let us take $p(x) = x^2+ux+w, u, w \in \mathbf{C}$ a quadratic polynomial. Let us assume u_1, u_2 are the roots of $p(x)$. Then:

- $\overline{u_1}$ may not be equal to u_2.
- $u_1+u_2 = -u$.
- $u_1 \cdot u_2 = w$.

Theorem: Fundamental Theorem of Algebra

Every complex polynomial $p(x)$ of degree $n \geq 1$ has the form $p(x) = u\cdot(x-u_1)\cdots(x-u_n)$, where u, $u_1,\ldots, u_n$ are complex numbers ($u_i \neq 0$). The numbers $u_1,\ldots, u_n$ are the roots of $p(x)$. u is the coefficient of x^n.

Corollary

Every polynomial $p(x)$ of positive degree with real coefficients can be factored as a product of linear and irreducible quadratic factors (in real numbers).

Example

$p(x) = \prod_{i=1}^{m}(x-a_i)\cdot\prod_{j=1}^{m_1}\left(x^2+b_j+c_j\right)$. Using previous theorem, $p(x) = \prod_{i=1}^{m}(x-a_i)\cdot\prod_{j=1}^{m_1}\left(x-d_j\right)\cdot\prod_{j=1}^{m_1}\left(x-\overline{d_j}\right)$.

Proposition

If $p(x)$ is a real polynomial (with real coefficients), then if $u \in \mathbf{C}$ is a root of $p(x)$, then $\overline{u}$ will also be a root of $p(x)$.

Proof:

- $p(u)=0 \Rightarrow \overline{p(u)}=\overline{0}=0 \Rightarrow \overline{a_n u^n+\cdots+a_1 u+a_0}=0 \Rightarrow \overline{a_n u^n}+\cdots+\overline{a_1 u}+\overline{a_0}=0$
  $\Rightarrow \overline{a_n}\cdot\overline{u^n}+\cdots+\overline{a_1}\cdot\overline{u}+\overline{a_0}=0 \Rightarrow a_n\overline{u}^n+\cdots+a_1\overline{u}+a_0=0 \Rightarrow p(\overline{u})=0$.

COMPLEX SPACES

We denote the set of all n-tuples of complex numbers by $\mathbf{C}^n$.

Definition

Let $v_i, w_i, u \in \mathbf{C}$. The operations on $\mathbf{C}^n$ are:

- $[v_1 \quad \cdots \quad v_n]^T+[w_1 \quad \cdots \quad w_n]^T=[v_1+w_1 \quad \cdots \quad v_n+w_n]^T$.
- $u[v_1 \quad \cdots \quad v_n]^T=[u\,v_1 \quad \cdots \quad uv_n]^T$.

Definition

Let $v=[v_1 \quad \cdots \quad v_n]^T$ and $w=[w_1 \quad \cdots \quad w_n]^T$. The inner product of v and w is $\langle v,w\rangle = v_1\overline{w_1}+\cdots+v_n\overline{w_n}$.

Theorem: Properties of Inner Products

Let $z, z_1, w, w_1 \in \mathbf{C}^n$ and $\lambda \in \mathbf{C}$.

1) $\langle z+z_1, w\rangle = \langle z,w\rangle+\langle z_1,w\rangle$ and $\langle z, w+w_1\rangle = \langle z,w\rangle+\langle z,w_1\rangle$.
2) $\langle \lambda\cdot z, w\rangle = \lambda\langle z,w\rangle$ and $\langle z, \lambda\cdot w\rangle = \overline{\lambda}\langle z,w\rangle$.
3) $\langle z,w\rangle = \overline{\langle w,z\rangle}$.
4) $\langle z,z\rangle \geq 0$ and $\langle z,z\rangle = 0 \Leftrightarrow z=0$.

Definition

Let $z=[z_1 \quad \cdots \quad z_n]^T$. We define $\text{norm}(z)=\|z\|=\sqrt{\langle z,z\rangle}$.

COMPLEX MATRICES

Definitions

- A matrix $Z=[z_{ij}]$ is called a complex matrix if every entry z_{ij} is a complex number.
- The conjugate of a matrix $Z=[z_{ij}]$ is defined as $\overline{Z}=[\overline{z_{ij}}]$.
- The conjugate transpose of a matrix is defined as $(\overline{Z})^T=Z^*$.

Theorem: Properties of Conjugate Transpose

Let Z, W denote complex matrices, and $\lambda \in \mathbf{C}$. Then:

1) $\left(Z^*\right)^* = Z$.
2) $(Z+W)^* = Z^* + W^*$.
3) $(\lambda Z)^* = \overline{\lambda} \cdot Z^*$.
4) $(Z \cdot W)^* = W^* \cdot Z^*$.

Definition

A square complex matrix H is called Hermitian if $H = H^*$ (natural generalization of real symmetric matrices).

EIGENVALUES, EIGENVECTORS, AND ORTHOGONALITY

Definition

Let Z be an $n \times n$ matrix. A complex number λ is called an eigenvalue of z if $ZX = \lambda X$ holds for some column $X \neq 0$ in $\mathbf{C}^n$. X is called an eigenvector of Z corresponding to λ.

Definition

The characteristic polynomial $c_Z(x)$ of an $n \times n$ matrix Z is defined by $c_Z(x) = \det(xI - Z)$.

Definition

Two columns Z and W in $\mathbf{C}^n$ are said to be orthogonal if $\langle z, w \rangle = 0$.

Theorem: Properties of Hermitian Matrices

1) An $n \times n$ complex matrix H is Hermitian iff $\langle HZ, W \rangle = \langle Z, HW \rangle, \forall Z, W \in \mathbf{C}^n$.
2) The eigenvalues of a Hermitian matrix are real.
3) The eigenvectors corresponding to distinct eigenvalues are orthogonal.

Definition

A set of non-zero vectors $\{z_1, \ldots, z_n\}$ in $\mathbf{C}^n$ is called orthogonal if $\langle z_i, z_j \rangle = 0, i \neq j$, and is orthonormal if, in addition, $\|z_i\| = \langle z_i, z_i \rangle = 1$.

Theorem

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

1) $U^{-1} = U^*$.
2) The rows of U are an orthonormal set in $\mathbf{C}^n$.
3) The columns of U are an orthonormal set in $\mathbf{C}^n$.

Such a matrix is called unitary.

Definition

An $n \times n$ complex matrix Z is called unitary diagonalizable if $U^* ZU$ is diagonal for some unitary matrix U.

Definition

A complex matrix is called upper triangle if every entry below the main diagonal is zero.

Theorem: Schur's Theorem

If Z is any $n \times n$ complex matrix, there exists a unitary matrix U such that $U^* Z U = T$ is upper triangle. Moreover, the entries in the main diagonal are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of Z (including multiplicities).

- Notice: $\det Z = \lambda_1 \cdot \cdots \cdot \lambda_n$ and $\operatorname{tr} Z = \lambda_1 + \cdots + \lambda_n$.

Theorem: Spectral Theorem

If H is Hermitian, there is an unitary matrix U such that $U^* H U = D$ is diagonal.

Proof:

- By Schur's Theorem, $U^* H U = T$ is upper triangle.
- $T^* = \left(U^* H U\right)^* = U^* H^* U^{**} = U^* H U = T$.
- Since $T^* = T$, T is diagonal.

Definition

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

Theorem

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

Orthogonal Diagonalization, Quadratic Forms, Positive Definite Forms

ORTHOGONAL DIAGONALIZATION

Theorem

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

1) P is invertible and $P^{-1} = P^T$.
2) The rows of P are orthonormal.
3) The columns of P are orthonormal.

Such a matrix is called orthogonal.

Proof:

- **1⇔2**: Assume $P P^T = I$. Let $X_1, \ldots, X_n$ denote the rows of P. Then X_j^T is the j^{th} column of P^T. So the (i, j)-entry of $P P^T$ is $X_i \cdot X_j$. Thus $P P^T = I$ means that $X_i \cdot X_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$, which means the rows of P are orthonormal.
- **1⇔3**: Follows similarly.

Definition

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

Theorem: Principle Axis Theorem

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

1) A has an orthogonal set of n eigenvectors.
2) A is orthogonally diagonalizable.
3) A is symmetric.

Proof:

- **1⇔2**:
  - **2⇒1**: $A = P^{-1}DP = P^T DP \Leftrightarrow AP^T = P^T D$. This means the columns of P^T are eigenvectors, so A has an orthogonal set of n eigenvectors.
  - **1⇒2**: $P^T = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}$ where X_i are eigenvectors. So $AP^T = P^T D \Rightarrow AP^T P = P^T DP \Rightarrow PAP^T = D$.
- **2⇔3**:
  - **2⇒3**: If $P^T AP = D$ is diagonal and $P^T = P^{-1}$, then $A = PDP^T$. $A^T = \left(PDP^T\right)^T = P^{TT} D^T P^T = PDP^T = A$. Since $A^T = A$, A is symmetric.
  - **3⇒2**: Use induction on the size of A.
    - For $n = 1$, it is trivial.
    - Assume it is true for an $(n-1)\times(n-1)$ matrix. Show it is true for an $n \times n$ matrix.
      - Let λ_1 be a real eigenvalue of A and X_1 be the associated eigenvector such that $\|X_1\| = 1$. Now use the Gram-Schmidt Algorithm to find an orthonormal basis $\{X_1, \ldots, X_n\}$ for $\mathbf{R}^n$.
      - Let $P_1 = \begin{bmatrix} X_1 & \cdots & X_n \end{bmatrix}$, then $P_1^T AP_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix}$. Then there exists an $(n-1)\times(n-1)$ orthogonal matrix Q such that $Q^T A_1 Q = D_1$ is diagonal. Hence $P_2 = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$ is orthogonal.
      - $(P_1P_2)^T A(P_1P_2) = {P_2}^T \left({P_1}^T AP_1\right)P_2 = \begin{bmatrix} 1 & 0 \\ 0 & Q^T \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}$.
      - $= \begin{bmatrix} \lambda_1 & 0 \\ 0 & D_1 \end{bmatrix}$

Theorem

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

Proof:

- Let $AX = \lambda X$ and $AY = \mu Y$, where $\lambda \neq \mu$.
- $\lambda(X \cdot Y) = \lambda X \cdot Y = AX \cdot Y = X \cdot AY = X \cdot (\mu Y) = \mu(X \cdot Y) \Rightarrow (\lambda - \mu)(X \cdot Y) = 0 \Rightarrow (X \cdot Y) = 0$ since $\lambda \neq \mu$.

Example

Diagonalize $A=\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

- $c_A(\lambda)=\det(A-\lambda I)=\lambda(\lambda-9)^2$. So the eigenvalues are $\lambda_1=0$ and $\lambda_2=9$.
- For $\lambda_1=0$, the associated eigenvector is $\xi_1=\frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$.
- Find ξ_2, ξ_3 by Gram-Schmidt.
- $P=[\xi_1 \quad \xi_2 \quad \xi_3]$, and $P^TAP=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$.

Theorem: Triangulation Theorem

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

QUADRATIC FORMS

- $q:\mathbf{R}^n\to\mathbf{R}$.
- $q(x_1,\ldots,x_n)=a_{11}{x_1}^2+\cdots+a_{nn}{x_n}^2+a_{1n}x_1x_n+\cdots+a_{ij}x_ix_j+\cdots$.
- $q(x_1,x_2)=a_{11}{x_1}^2+a_{22}{x_2}^2+a_{12}x_1x_2+a_{21}x_2x_1=\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=X^TAX$.

  $q(x_1,x_2)=a_{11}{x_1}^2+a_{22}{x_2}^2+(a_{12}+a_{21})x_1x_2=a_{11}{x_1}^2+a_{22}{x_2}^2+\frac{(a_{12}+a_{21})}{2}x_1x_2+\frac{(a_{21}+a_{12})}{2}x_2x_1$
- $=\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & \frac{(a_{12}+a_{21})}{2} \\ \frac{(a_{21}+a_{12})}{2} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. So

  $q(X)=X^TAX$, A is a symmetric matrix.
- Let us diagonalize A. $A=PDP^T\Rightarrow q(X)=X^TPDP^TX$. Take $Y^T=X^TP$ (a change in variable). Then $q(Y)=Y^TDY=\lambda_1{y_1}^2+\cdots\lambda_n{y_n}^2$, λ_i are the eigenvalues of A.

Theorem

Let $q=X^TAX$ be a quadratic form in the variables $X=[x_1 \quad \cdots \quad x_n]^T$ and A is a symmetric $n\times n$ matrix.
Let P be an orthogonal matrix such that $P^TAP=D$ is diagonal, and define $Y=[y_1 \quad \cdots \quad y_n]^T$ by $Y=P^TX$.
If q is expressed in terms of these new variables Y, then the result is $q(Y)=\lambda_1{y_1}^2+\cdots\lambda_n{y_n}^2$, λ_i are the eigenvalues of A (repeated according to their multiplicities).

Theorem

Consider the quadratic form $q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ where a, b, c are not all zero. Then:

1) There is a counter-clockwise rotation of the coordinates axis about the origin such that, in the new system, q has no cross terms.
2) The graph of $ax_1^2 + bx_1x_2 + cx_2^2 = 1$ is an ellipse if $b^2 - 4ac < 0$, and a hyperbola if $b^2 - 4ac > 0$.

Proof:

- $q(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.
- $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. $c_A(x) = x^2 - (a+c) - \dfrac{b^2 - 4ac}{4}$, so the eigenvalues are $\lambda_1 = \dfrac{1}{2}\left(a + c - \sqrt{b^2 + (a-c)^2}\right)$ and $\lambda_2 = \dfrac{1}{2}\left(a + c + \sqrt{b^2 + (a-c)^2}\right)$. The eigenvectors are $F_1 = \dfrac{1}{\sqrt{b^2 + (a+c-d)^2}} \begin{pmatrix} a+c-d \\ b \end{pmatrix}$ and $F_2 = \dfrac{1}{\sqrt{b^2 + (a+c-d)^2}} \begin{pmatrix} -b \\ a+c-d \end{pmatrix}$ where $d = \sqrt{b^2 + (a-c)^2}$. So $ax_1^2 + bx_1x_2 + cx_2^2 = 1 \Leftrightarrow \lambda_1 y_1^2 + \lambda_2 y_2^2 = 1$.
- It follows that if $\lambda_1\lambda_2 > 0$, then it is an ellipse. Otherwise, if $\lambda_1\lambda_2 < 0$, then it is an hyperbola.

POSITIVE DEFINITE MATRICES

Definition

An $n \times n$ matrix is called positive definite if it is symmetric an all its eigenvalues are greater than 0.

Theorem

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

Proof:

- By Triangulation Theorem, $\det A = \lambda_1 \cdots\cdots \lambda_n > 0$, λ_i are the eigenvalues.

Theorem

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

Proof:

- Let $P^T AP = D = \text{diag}(\lambda_1, \ldots, \lambda_n)$, where $P^{-1} = P^T$ and λ_i are the eigenvalues of A.
- Assume A is positive definite. Given $X \in \mathbf{R}^n$, $X^T AX = X^T P^T APX = Y^T DY = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 > 0$.
- Assume $X^T AX > 0$ when $X \neq 0$. Let $X = PE_j \neq 0$, where E_j is the column j of I. Then $X^T AX = E_j^T P^T DPE_j = \lambda_j > 0$.

Vector Spaces and Subspaces

VECTOR SPACES

Definition

A vector space consists of a non-empty set V of object (called vectors) that can be added, that can be multiplied by a number (called a scalar), for which certain axioms hold. $(V, \mathbf{R}, +, \cdot)$.

Axioms for Addition

If v, w are two vectors in V, their sum is expressed as $v+w$ and it satisfies the following axioms:

- A_1: If u and v are in V, then $u+v$ is in V.
- A_2: $u+v=v+u$ for all $u, v \in V$.
- A_3: $u+(v+w)=(u+v)+w, \forall u, v, w \in V$.
- A_4: An element 0 in V exists such that $v+0=v=0+v, \forall v \in V$ (0 is called the zero vector).
- A_5: For each $v \in V$, an element $-v \in V$ exists such that $v+(-v)=0$ (-v is called the negative of v).

Axioms for Multiplication

- S_1: If v is in V, $a \cdot v \in V, \forall a \in \mathbf{R}$.
- S_2: $a \cdot (v+w)=a \cdot v + a \cdot w, \forall v, w \in V, a \in \mathbf{R}$.
- S_3: $(a+b) \cdot v = a \cdot v + b \cdot v, \forall v \in V, a, b \in \mathbf{R}$.
- S_4: $a \cdot (b \cdot v) = (ab) \cdot v, \forall v \in V, a, b \in \mathbf{R}$.
- S_5: $1v = v, \forall v \in V$.

Example

Let $V = \mathbf{R}^2$. Define + as $(x_1, y_1)+(x_2, y_2)=(x_1+x_2, y_1+y_2)$. Define $\cdot$ as $a \cdot (x, y)=(ay, ax)$. Is V a vector space?

- S_1: $a \cdot (x, y)=(ay, ax) \in \mathbf{R}^2$. Satisfied.
- S_2: $a \cdot (v+w)=a \cdot (x_1+x_2, y_1+y_2)=(a(y_1+y_2), a(x_1+x_2))$ and $a \cdot v + a \cdot w = (ay_1, ax_1)+(ay_2, ax_2)=(a(y_1+y_2), a(x_1+x_2))$. Satisfied.
- S_3: $(a+b) \cdot v = ((a+b)y, (a+b)x)$ and $a \cdot v + b \cdot v = (ay, ax)+(by, bx)=((a+b)y, (a+b)x)$. Satisfied.
- S_4: $a \cdot (b \cdot v)=a \cdot (by, bx)=(abx, aby)$ but $(ab) \cdot v = (aby, abx)$. Failed.
- S_5: $1 \cdot v = (y, x) \neq (x, y)$. Failed.

Theorem: Cancellation

Let $u, v, w \in V$. If $v+u=v+w$, then $u=w$.

Theorem

If $u, v \in V$, the equation $x+v=u$ has one and only one solution $x \in V$ given by $x=u-v$.

Proof:

- $x+v=(u-v)+v=[u+(-v)]+v \overset{A_3}{=} u+(-v)+v \overset{A_5}{=} u+0 \overset{A_4}{=} u$.

- Uniqueness: Suppose there is another solution $x_1 \neq x$. Then $x_1 + v = u = x + v \Rightarrow x_1 = x$. Contradiction!

Theorem

Let $v \in V$ and $a \in \mathbf{R}$. Then:

1) $0 \cdot v = 0$.
2) $a \cdot 0 = 0$.
3) If $a \cdot v = 0$, then either $a = 0$ or $v = 0$.
4) $(-1) \cdot v = -v$.
5) $(-a) \cdot v = -(a \cdot v) = a \cdot (-v)$.

Proof 2:

- $\left.\begin{aligned} a \cdot v &\overset{A_4}{=} a \cdot (0 + v) \overset{A_2}{=} a \cdot 0 + a \cdot v \\ a \cdot v &\overset{A_2}{=} a \cdot v + 0 \end{aligned}\right\} \Leftrightarrow a \cdot v + 0 = a \cdot 0 + a \cdot v$. By cancellation, $a \cdot 0 = 0$.

Proof 5:

- $(-a) \cdot v = ((-1)(a)) \cdot v \overset{S_4}{=} (-1)(a \cdot v) = -(a \cdot v)$.

Examples of Vector Spaces

- $\mathbf{R}^n$ with addition defined as $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$, and multiplication by a scalar defined as $a \cdot (x_1, \ldots, x_n) = (ax_1, \ldots, ax_n)$.
- P_n the set of polynomials of degree at most n, with addition defined as $\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i, a, b \in \mathbf{R}$, and multiplication by a scalar defined as $a \cdot \sum_{i=0}^{n} a_i x^i = \sum_{i=0}^{n} a a_i x^i$.
- $M_{m \times n}$ the set of all $m \times n$ matrices, with addition defined as $\lfloor a_{ij} \rfloor + \lfloor b_{ij} \rfloor = \lfloor a_{ij} + b_{ij} \rfloor$, and multiplication by a scalar defined as $a \cdot \lfloor a_{ij} \rfloor = \lfloor a a_{ij} \rfloor$.
- $F[a, b]$ the set of all functions defined in $[a, b]$, with addition defined as $(f + g)(x) = f(x) + g(x)$, and multiplication by a scalar defined as $(a \cdot f)(x) = a \cdot (f(x))$.

SUBSPACES

Definition

A subset U of a vector space V is called a subspace of V if U itself is a vector space that uses the vector addition and scalar multiplication of V.

Theorem: Subspace Test

Let U be a subset of a vector space V. Then U is a subspace if and only if it satisfies the following 3 conditions:

1) 0 lies in U, where 0 is the zero vector in V.
2) If u_1, u_2 lies in U, then $u_1 + u_2$ lies in U.

3) If u lies in U, then $a \cdot u$ lies in U for all $a \in \mathbf{R}$.

Example

Let $U = \{p(x) \text{ in } P \mid p(3) = 0\}$. Show that U is a subspace.

- $0(3) = 0$. So 0 is in U.
- $(u_1 + u_2)(3) = u_1(3) + u_2(3) = 0 + 0 = 0$. So $u_1 + u_2$ is in U.
- $(a \cdot u)(3) = a \cdot u(3) = a \cdot 0 = 0$. So $a \cdot u$ is in U.
- So U is a subspace.

Example

Show that the subset $D[a,b]$ of all differentiable functions on $[a,b]$ is a subspace of $F[a,b]$.

- $0(x) = 0$ is the zero function in $F[a,b]$. 0 is differentiable because constant functions are differentiable. So 0 is in $D[a,b]$.
- $(f+g)' = f' + g'$ exists, so $f+g$ is in $D[a,b]$.
- $(a \cdot f)' = a \cdot f'$ exists, so $a \cdot f$ is in $D[a,b]$.

LINEAR COMBINATIONS AND SPANNING SETS

Definition

Let $\{v_1, \ldots, v_n\}$ be a set of vectors in a vector space V. A vector v is called a linear combination of the vectors $v_1, \ldots, v_n$ if it can be expressed in the form $v = a_1 v_1 + \cdots + a_n v_n = \sum_{i=0}^{n} a_i v^i$, where a_i are scalars called the coefficients of v_i.

Definition

The set of all linear combinations of the vectors $v_1, \ldots, v_n$ is called their span, and is denoted $\text{span}\{v_1, \ldots, v_n\}$.

Theorem

Let $U = \text{span}\{v_1, \ldots, v_n\}$ in a vector space V. Then:

1) U is a subspace of v containing each of $v_1, \ldots, v_n$.
2) U is the smallest subspace containing these vectors in the sense that any subspace of V that contains each of $v_1, \ldots, v_n$ must contain U.

Proof 1:

- U is a subspace:
  - $0 = 0v_1 + \cdots + 0v_n \Rightarrow 0 \in \text{span}\{v_1, \ldots, v_n\}$.
  - Let $z = \sum_{i=0}^{n} a_i v_i$ and $w = \sum_{i=0}^{n} b_i v_i$ be in U. Then $z + w = \sum_{i=0}^{n} (a_i + b_i) v_i \Rightarrow (z+w) \in U$.
  - Let $z = \sum_{i=0}^{n} a_i v_i \in U$. Then $c \cdot z = c \cdot \sum_{i=0}^{n} (ca_i) v_i \Rightarrow c \cdot z \in U$.

- U contains each of $v_1,\ldots,v_n$: Take $a_1=0,\ldots,a_{i-1}=0,a_i=1,a_{i+1}=0,\ldots,a_n=0$, then $v_i\in U,\forall i$.

Proof 2:

- Let W be a subspace of V that contains each of $v_1,\ldots,v_n$. Because W is closed under scalar multiplication, each of $a_1v_1+\cdots+a_nv_n\in W,\forall a_i\in\mathbf{R}$. So $W\supset U$ (it contains all the elements in U).

The Dimension Theory

LINEAR INDEPENDENCE

Definition

A set of vectors $\{v_1,\ldots,v_n\}$ is called <u>linearly independent</u> if it satisfies the following condition:
$s_1v_1+\cdots+s_nv_n=0\Rightarrow s_1=\cdots=s_n=0$.
A set of vectors that is not linearly independent is said to be <u>linearly dependent</u>.

- The idea behind linear independence is that you cannot express one vector in the set $\{v_1,\ldots,v_n\}$ as a linear combination of the remain vectors ($v_i\neq a_1v_1+\cdots+a_{i-1}v_{i-1}+a_{i+1}v_{i+1}+\cdots a_nv_n$).

Example

Show that $\{1+x,3x+x^2,2+x-x^2\}$ is linearly independent in P_2.

$$s_1(1+x)+s_2(3x+x^2)+s_3(2+x-x^2)=0\Rightarrow(s_1+2s_3)+(s_1+3s_2+s_3)x+(s_2-s_3)x^2=0$$

- $$\Rightarrow\begin{cases}s_1+2s_3=0\\ s_1+3s_2+s_3=0\\ s_2-s_3=0\end{cases}\Rightarrow s_1=s_2=s_3=0$$.

Example

Show that $\{\sin x,\cos x\}$ is linearly independent in the vector space $F[0,\pi]$.

- Let $\alpha_1\sin x+\alpha_2\cos x=0$.
- Take $\begin{cases}x=0\Rightarrow\alpha_2=0\\ x=\dfrac{\pi}{2}\Rightarrow\alpha_1=0\end{cases}\Rightarrow\alpha_1=\alpha_2=0,\forall x\in[0,\pi]$.

Theorem

A set $\{v_1,\ldots,v_n\}$ of vectors in V is linearly dependent if and only if some v_i is a linear combination of the others.

Proof:

- Want $\{v_1,\ldots,v_n\}$ is linearly dependent $\Rightarrow$ some v_i is a linear combination of the others:
  - Some nontrivial combination vanishes $a_1v_1+\cdots+a_nv_n=0$ where some coefficient is not zero.

- Suppose $a_i \neq 0 \Rightarrow v_i = \frac{a_1}{a_i} v_1 + \cdots + \frac{a_{i-1}}{a_i} v_{i-1} + \frac{a_1}{a_{i+1}} v_{i+1} + \cdots + \frac{a_n}{a_i} v_n$ gives v_i a linear combination of the remaining vectors.

- Want some v_i is a linear combination of the others $\Rightarrow \{v_1, \ldots, v_n\}$ is linearly dependent:
  - $v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_n v_n$
    $\Rightarrow \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} - v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_n v_n = 0$.
  - There is a nontrivial linear combination, so it is linearly dependent.

Theorem: Uniqueness of Representation

Let $\{v_1, \ldots, v_n\}$ a linearly independent set of vectors. If a vector v has two representations, $v = \sum_{i=0}^{n} a_i v_i$ or $v = \sum_{i=0}^{n} b_i v_i$, then $a_i = b_i$ (there is only one linear combination).

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 $m \leq n$.

Proof:

- Let $V = \text{span}\{v_1, \ldots, v_n\}$. Suppose $\{v_1, \ldots, v_n\}$ is an independent set in V.
- Then $u_1 = \sum_{i=1}^{n} a_{1i} v_i$ and assume $a_{1i} \neq 0$ because $u_1 \neq 0$. So now $V = \text{span}\{u_1, v_2, \ldots, v_n\}$ (to see this, take any element $v \in V$,
  $v = \sum_{i=1}^{n} b_i v_i = \sum_{i=2}^{n} b_i v_i + b_1 \left(\frac{v_1}{a_{11}} - \sum_{i=2}^{n} \frac{a_{1i}}{a_{11}} v_i \right) = \frac{b_1}{a_{11}} u_1 + \sum_{i=2}^{n} \left(b_i - b_1 \frac{a_{1i}}{a_{11}} \right) v_i$, which means $\{u_1, v_2, \ldots, v_n\}$ spans V).
- Hence, write $u_2 = a_{21} v_1 + \sum_{i=2}^{n} a_{2i} v_i$ and assume $a_{2i} \neq 0, i = 2, \ldots, n$ because $\{u_1, u2\}$ are linearly independent.
- For convenience, say $a_{22} \neq 0$. So now, $V = \text{span}\{u_1, u_2, v_3, \ldots, v_n\}$.
- If $m > n$, this procedure will continue until all the vectors v_i are replaced by $u_1, \ldots, u_n$, and $V = \text{span}\{u_1, \ldots, u_n\}$. Then $u_{n+1}, \ldots, u_m$ can be represented as a linear combination of $\{u_1, \ldots, u_n\}$. This is a contradiction because u_i are linearly independent. So $m \leq n$.

Steinitz Exchange Lemma

If $V = \text{span}\{v_1, \ldots, v_n\}$ and if $\{u_1, \ldots, u_m\}$ is a independent set in V ($m \leq n$), then m of the vectors v_i can be replaced by $u_1, \ldots, u_m$ and the resulting set will still span V.

Definition

A set $\{e_1, \ldots, e_n\}$ of vectors in a vector space V is called a <u>basis</u> of V if it satisfies the following two conditions:

1) $\{e_1, \ldots, e_n\}$ is linearly independent.
2) $V = \text{span}\{e_1, \ldots, e_n\}$.

Theorem

Let $\{e_1,\ldots,e_n\}$ and $\{f_1,\ldots,f_m\}$ be two bases of *V*. Then $n=m$.

Definition

V is called finite dimensional if $V=0$ or *V* has a finite basis.

Examples

- M_{mn} has dimension $m\cdot n$.
- P_n has dimension $n+1$.
- $F[a,b]$ is not finite dimensional.

Example

Let $A=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$ and consider the subspace $U=\{X\in M_{22}\mid AX=XA\}$. Show that $\dim U=2$ and find a basis.

- Denote $X=\begin{bmatrix}a & b\\ c & d\end{bmatrix}$.
- $AX=XA\Rightarrow\begin{bmatrix}a+c & b+d\\ 0 & 0\end{bmatrix}=\begin{bmatrix}a & a\\ c & c\end{bmatrix}\Rightarrow\begin{cases}a+c=a\Rightarrow a=a\\ b+d=a\Rightarrow b=a-d\\ c=0\\ c=0\end{cases}\Rightarrow X=\begin{bmatrix}a & a-d\\ 0 & d\end{bmatrix}$. So

  $X=a\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}+d\begin{bmatrix}0 & -1\\ 0 & 1\end{bmatrix}$.
- Let $s_1\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}+s_2\begin{bmatrix}0 & -1\\ 0 & 1\end{bmatrix}=0\Rightarrow\begin{cases}s_1=0\\ s_1-s_2=0\Rightarrow s_1=s_2=0\\ s_2=0\end{cases}$. So $\left\{\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix},\begin{bmatrix}0 & -1\\ 0 & 1\end{bmatrix}\right\}$ is linearly independent.
- Therefore $\left\{\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix},\begin{bmatrix}0 & -1\\ 0 & 1\end{bmatrix}\right\}$ is a basis and $\dim U=2$.

Theorem

Let *V* be a vector space and assume that $\dim V=n>0$.

1) No set of more than *n* vectors can be linearly independent.
2) No set of fewer than *n* vectors can span *V*.

Proof 1: Fundamental Theorem.

Proof 2: Follows from Invariance Theorem. Assume $V=\text{span}\{f_1,\ldots,f_m\}$ is linearly independent and $m<n$. Then $\{f_1,\ldots,f_m\}$ is a basis. By Invariance Theorem, $m=n$. Contradiction.

PROCEDURE TO CREATE A BASIS

Given $S_k = \{v_1, \ldots, v_k\}$, we can complete a basis by taking an element v_{k+1} outside span S_k, then create a new set $S_{k+1} = \{v_1, \ldots, v_k, v_{k+1}\}$ and repeat the procedure until there is no element outside span S_{k+1}.

How can you find v_{k+1}?

1) Represent all possible linear combinations as $a_1 v_1 + \cdots + a_k v_k$.
2) Provide a general expression for an element in *V*, denoted *v* (ex: $b_0 + b_1 x + \cdots + b_n x^n$, $\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}$).
3) Create a system of equations $v = \sum_{i=1}^{n} a_i v_i$ where the unknowns will be a_i and the coefficients will be b_i.
4) Show this system has no solution for some b_i, then take this b_i as the extra vector.

EXISTENCE OF BASES

Theorem

Let $\{v_1, \ldots, v_n\}$ be a linearly independent set of vectors in a vector space *V*. The following conditions are equivalent for a vector *v* in *V*:

1) $\{v, v_1, \ldots, v_n\}$ is linearly independent.
2) *v* does not lie in $\text{span}\{v_1, \ldots, v_n\}$.

Theorem

Let $V \neq 0$ be a vector space spanned by *n* vectors.

1) Each set of linearly independent vectors is part of a basis.
2) Each spanning set for *V* contains a basis of *V*.
3) *V* has a basis and $\dim V = n$.

Proof 1: Follows from previous theorem because you can add elements to $\{v_1, \ldots, v_k\}$ where v_i is linearly independent until you span *V* (so until you get a basis).

Proof 2: Let $V = \text{span}\{v_1, \ldots, v_n\}, v_i \neq 0$. If $\{v_1, \ldots, v_n\}$ is linearly independent, then it is itself a basis. If not, then one of these vectors, say v_1, lie in the span of the other. Then $V = \text{span}\{v_2, \ldots, v_n\}$. Repeat until you get a linearly independent set.

Proof 3: It follows from 2 because if the spanning set of vectors is not linearly independent, then we can remove one in order to find a basis.

Theorem

Let *V* be a vector space and assume that $\dim V = n > 0$. Let *U* and *W* denote subspaces of *V*.

1) Any set of *n* linearly independent vectors in *V* is a basis.
2) Any spanning set of *n* non-zero vectors in *V* is a basis.
3) *U* is finite dimensional and $\dim U \leq n$.
4) Any basis of *U* is part of a basis of *V*.
5) If *U* is a subset of *W* and $\dim U = \dim W$, then $U = W$.

Example

Let a be a number, and let $W = \text{span}\{p(x) \mid p(x) \in P_n, p(a) = 0\}$. Show $S = \{(x-a), \ldots, (x-a)^n\}$ is a basis of W.

- Let $b_1(x-a) + \cdots + b_n(x-a)^n = 0 \Rightarrow b_n x^n + (b_{n-1} - ab_n)x^{n-1} + \cdots = 0 \Rightarrow b_n = \cdots = b_1 = 0$. So S is linearly independent.
- Because S is linearly independent, $\dim W \geq n$. But $\dim P_n = n+1$, so $\dim W = n$ or $\dim W = n+1$.
- Assume $\dim W = n+1$. Then $W = P_n$, but this a contradiction because $p(x) = b \in P_n(x)$ but $p(x) = b \notin W$. So $\dim W = n$.