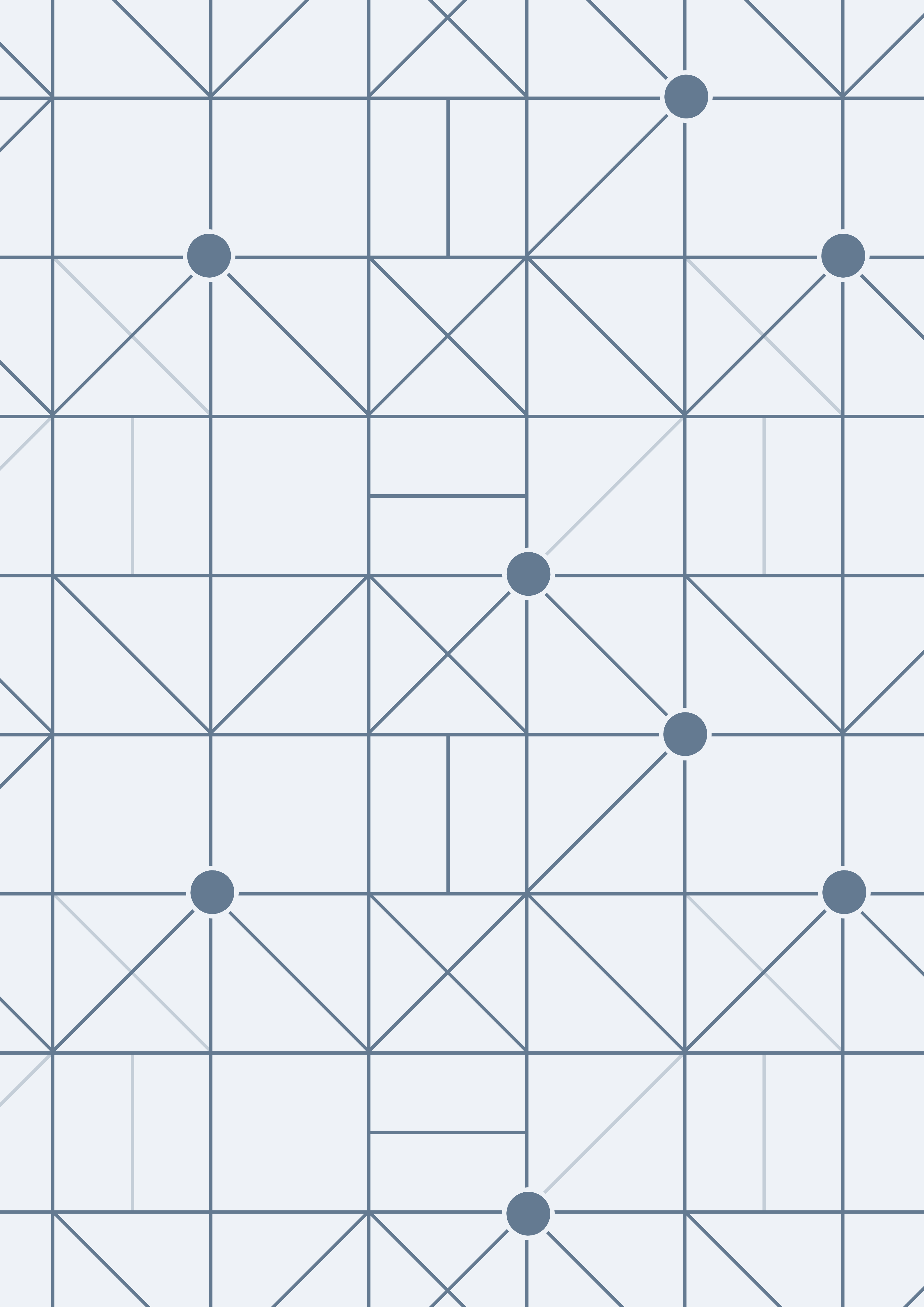


1 Abstract Linear Algebra

VECTOR SPACES

To do linear combinations, we need to be able to scale and add

Examples of Linear Combinations

1) Polynomials

$\mathbb{R}_n[X]$: set of polynomials in x with real coefficients of degree $\leq n$

$\mathbb{R}_n[X] = \{\alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \mid \alpha_j \in \mathbb{R},\ 0 \leq j \leq n\}$

- <u>Linear Combination</u>: $\forall a, b \in \mathbb{R}$

$$a(\alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n) + b(\beta_0 + \beta_1 x + \cdots + \beta_n x^n)$$
$$= \sum_{j=0}^{n} (a\alpha_j + b\beta_j) x^j.$$

- <u>Zero Polynomial</u>: Let $\alpha_j = 0 \quad \forall j$

Note: Also works for $\mathbb{C}_n[X]$

2) Functions:

$\mathcal{F}([a,b], \mathbb{R})$: set of real valued functions

$f : [a,b] \to \mathbb{R} \qquad ([a,b] \subseteq \mathbb{R}$ is an interval$)$

- <u>Linear Combination</u>: For $f, g \in \mathcal{F}([a,b], \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$, define function

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) \qquad \forall x \in [a,b]$$

- <u>Zero function</u>: $\underline{0}(x) = 0 \quad \forall x \in [a,b]$

Note: $\mathbb{R}_n[X] \subseteq \mathcal{F}([a,b], \mathbb{R})$

For any n and linear combination rule in $\mathbb{R}_n[X]$ agrees with rule in $\mathcal{F}([a,b], \mathbb{R})$

3) Matrices: $M_{p \times n}(\mathbb{F})$: set of $p \times n$ matrices with entries in $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$

- <u>Linear Combination</u>: matrix addition and scalar multiplication of matrices

Let $A = (a_{jk})$, $B = (b_{jk})$, $\alpha, \beta \in \mathbb{F}$ (j,k)-entry given by

$$(\alpha A + \beta B)_{jk} = \alpha a_{jk} + \beta b_{jk}$$

- <u>Zero Matrix</u>: $0_{p \times n} \in Mat_{p \times n}(\mathbb{F})$

Definition of a Vector Space

Definition Vector Space

Let $\mathbb{F}$ be a field (usually $\mathbb{R}$ or $\mathbb{C}$). A **vector space** over $\mathbb{F}$ is a set V together with binary operations

vector addition

$$V \times V \longrightarrow V$$
$$(u,v) \longmapsto (u+v)$$

scalar multiplication

$$\mathbb{F} \times V \longrightarrow V$$
$$(\alpha, v) \longmapsto \alpha v$$

(A1) commutativity over addition

$$u+v = v+u \quad \forall u,v \in V$$

(A2) associativity over addition

$$u+(v+w) = (u+v)+w \quad \forall u,v,w \in V$$

(A3) $\underline{0}$ vector

$$\exists\, \underline{0} \in V \text{ such that } \underline{0}+v = v \quad \forall v \in V$$

(A4) Inverse

Given any $v \in V$, $\exists\, -v \in V$ with $(-v)+v = \underline{0}$

(M1) Distributivity

$$\alpha(u+v) = \alpha u + \beta v \quad \forall \alpha \in \mathbb{F},\ u,v \in V$$

(M2) Scalar Multiplication

$$\alpha(\beta v) = (\alpha\beta)v \quad \forall \alpha,\beta \in \mathbb{F} \text{ and } v \in V$$

(M3) Distributivity

$$(\alpha+\beta)v = \alpha v + \beta v \quad \forall \alpha,\beta \in \mathbb{F},\ v \in V$$

(M4) Multiplicative Identity

$$1v = v \quad \forall v \in V \qquad (\text{where } 1 \in \mathbb{F} \text{ is the usual } 1)$$

- $\underline{0}$ is the **zero vector**
- **Real Vector Space**: Vector Space over $\mathbb{R}$

  Complex Vector Space: Vector Space over $\mathbb{C}$

- A **vector** is an element of a vector space
- Given a vector space V over a field $\mathbb{F}$, any $\alpha \in \mathbb{F}$ is a **scalar**

Note

i) Being binary operation implies V is closed under linear combination

$\forall\, u, v \in \mathbb{F}$ and any $\alpha \in \mathbb{F}$, $u+v \in V$, $\alpha v \in V$

ii) Axioms A1 - A4 together with binary operation addition is an abelian group

Examples of Vector Space

1) $\mathbb{F}^n$ is a vector space over $\mathbb{F}$ with obvious definitions of vector addition and multiplication

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} \in \mathbb{F}^n$$

$$s \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} s\alpha_1 \\ \vdots \\ s\alpha_n \end{pmatrix} \in \mathbb{F}^n \qquad \forall\, s \in \mathbb{F}$$

2) All examples on page 2 🔗

3) The **trivial vector space** $\{\underline{0}\}$ is a vector field over any field $\mathbb{F}$

$$\underline{0} = \underline{0} + \underline{0}, \qquad \alpha \underline{0} = \underline{0} \quad \forall \alpha \in \mathbb{F}$$

4) **Field of order 2** (order is cardinality)

$$\mathbb{F} = \mathbb{F}_2 = \{\underline{0}, 1\}$$

order

Field operation: modular arithmetic

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Note: Any finite field must have order of a prime power

$\mathbb{F} = \mathbb{F}_n$ where $n = p^a$ $a \in \mathbb{N}$, p prime

order

4) $\mathbb{F}_2^3$: vector space of 3-dimensional column vectors with entries in $\mathbb{F}_2$

$$|\mathbb{F}_2^3| = 2\cdot 2\cdot 2 = 8$$

$$\mathbb{F}_2^3 = \left\{ \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix} \right\}$$

5) **Diagonal Matrix Space:**

Let V_1 be the set of $n\times n$ diagonal matrix

For 2 elements of V_1,

$$u = \begin{pmatrix}\alpha_1 & & 0\\ & \alpha_2 \ddots & \\ 0 & & \alpha_n\end{pmatrix} \qquad v = \begin{pmatrix}\beta_1 & & 0\\ & \beta_2 \ddots & \\ 0 & & \beta_n\end{pmatrix}$$

<u>Addition</u>: $u+v = \begin{pmatrix}\alpha_1 & & 0\\ & \alpha_2 \ddots & \\ 0 & & \alpha_n\end{pmatrix} + \begin{pmatrix}\beta_1 & & 0\\ & \beta_2 \ddots & \\ 0 & & \beta_n\end{pmatrix} = \begin{pmatrix}\alpha_1+\beta_1 & & \\ & \alpha_2+\beta_2 \ddots & \\ & & \alpha_n+\beta_n\end{pmatrix}$

<u>Scalar Multiplication</u>: $\forall \gamma \in \mathbb{R}$, $\gamma\begin{pmatrix}\alpha_1 & & 0\\ & \alpha_2 \ddots & \\ 0 & & \alpha_n\end{pmatrix} = \begin{pmatrix}\gamma\alpha_1 & & 0\\ & \ddots & \\ 0 & & \gamma\alpha_n\end{pmatrix}$

So V_1 is a vector space over $\mathbb{R}$

<u>Note</u>: This is basically same as $\mathbb{R}^n$, just change of notation

Can show

$$\varphi : V_1 \longrightarrow \mathbb{R}$$

$$\begin{pmatrix}\alpha_1 & & \\ & \ddots & \\ & & \alpha_n\end{pmatrix} \longmapsto \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$$ is an isomorphism

6) $V_2 \subseteq V_1$ be the set of matrices with positive diagonal entries

Let

$$u = \begin{pmatrix}\alpha_1 & & 0\\ & \alpha_2 \ddots & \\ 0 & & \alpha_n\end{pmatrix} \qquad v = \begin{pmatrix}\beta_1 & & 0\\ & \beta_2 \ddots & \\ 0 & & \beta_n\end{pmatrix} \qquad \alpha_i, \beta_i > 0$$

Define "vector addition" and "scalar multiplication" by

$$u+v=\begin{pmatrix}\alpha_1\beta_1 & & & 0\\ & \alpha_2\beta_2 & & \\ & & \ddots & \\ 0 & & & \alpha_n\beta_n\end{pmatrix}$$

$$\gamma u=\begin{pmatrix}\alpha_1^{\gamma} & & & 0\\ & \alpha_2^{\gamma} & & \\ & & \ddots & \\ 0 & & & \alpha_n^{\gamma}\end{pmatrix} \qquad \gamma\in\mathbb{R}$$

Recall: $\alpha_j^{\gamma}=\exp(\gamma\log(\alpha_j))$ so only makes sense for $\alpha_j>0$

Then V_2 is a vector space

proof:

(A1) $u+v=\begin{pmatrix}\alpha_1\beta_1 & & \\ & \ddots & \\ & & \alpha_n\beta_n\end{pmatrix}=\begin{pmatrix}\beta_1\alpha_1 & & \\ & \ddots & \\ & & \beta_n\alpha_n\end{pmatrix}=v+u$

(M3) For $\gamma,\lambda\in\mathbb{R}$, $(\gamma+\lambda)u=\gamma u+\lambda v$

since $\alpha_j^{\gamma+\lambda}=\alpha_j^{\lambda+\gamma}$

(A4) Identity: $\underline{0}=I_n$

Matrices need to be diagonal

Note: Operations from V_1 would NOT work on V_2 since for $v\in V_2$, need negative entries to form $-v$, $-v\notin V_2$, i.e. A4 fails

- V_2 is not a subspace of V_1 since $\underline{0}\in V_1$ not in V_2, $\underline{0}\notin V_2$

Linear Combination

Definition Linear Combination

Given vectors $\underline{v}_1,\dots,\underline{v}_q\in V$ and scalars $\alpha_1,\dots,\alpha_q\in\mathbb{F}$, the sum

$$\alpha_1\underline{v}_1+\dots+\alpha_q\underline{v}_q=\sum_{j=1}^{q}\alpha_j\underline{v}_j$$

is called the linear combination

Linear Subspace

Definition Subspaces

A subset $S \subseteq V$ is called a subspace (or linear subspace) of V if

(S1) $S \neq \emptyset$

(S2) $\underline{0} \in S$

(S3) $\forall\, \underline{v}_1, \dots, \underline{v}_q \in S,\ a_1\underline{v}_1 + \cdots + a_q\underline{v}_q \in S$ (closed under linear combination)

Linear Dependance/Independence

Definition Linear dependence

A collection of vectors $\mathcal{C} = \{\underline{v}_1, \dots, \underline{v}_q\} \subseteq V$ is linearly dependant if $\exists\, (\alpha_1, \dots, \alpha_q) \in \mathbb{F}^q \setminus \{(0, \dots, 0)\}$ s.t

$$\alpha_1\underline{v}_1 + \cdots + \alpha_q\underline{v}_q = \underline{0}$$

Otherwise, we say $\underline{v}_1, \dots, v_q$ are linearly independant

Definition Linear independence

$\underline{v}_1, \dots, \underline{v}_q$ are linearly independent if

$$\alpha_1\underline{v}_1 + \cdots + \alpha_q\underline{v}_q = \underline{0} \implies \alpha_1 = 0, \dots, \alpha_q = 0$$

Spans

Definition Span

Let $\mathcal{C} \subset V$ be a non-empty collection of vectors.

The span of $\mathcal{C}$ denoted

$$Sp(\mathcal{C})$$

is the set of all linear combination of $\mathcal{C}$

$$Sp(\mathcal{C}) = \{\underline{u} \in \mathbb{F}^n \mid \underline{u} = \alpha_1\underline{v}_1 + \cdots + \alpha_n\underline{v}_n \text{ for some } \alpha_i \in \mathbb{F},\ \underline{v}_i \in S\}$$

By convention, $\boxed{Sp(\emptyset) = \{\underline{0}\}}$

Basis

Definition Basis

Let $S \leq V$ be a non-trivial $S \neq \{0\}$ subspace of $\mathbb{F}^n$,

A collection $B = \{\underline{v}_1, \cdots, \underline{v}_q\} \subseteq S$ forms a basis if

i) $\underline{v}_1, \cdots, \underline{v}_q$ is linearly independent

ii) $sp(\underline{v}_1, \cdots, \underline{v}_q) = S$

By definition,

$$\boxed{\text{basis of } \{0\} \text{ is } \emptyset}$$

Lemma

For any $\mathcal{C} \subset \mathbb{F}^n$, $\mathcal{C} \neq 0$,

$sp(\mathcal{C})$ is a subspace of $\mathbb{F}^n$

In fact, $sp(\mathcal{C})$ is the smallest subspace of $\mathbb{F}^n$ containing $\mathcal{C}$, i.e.

if $S \leq \mathbb{F}^n$ is any subspace with $\mathcal{C} \subseteq S$, then $sp(\mathcal{C}) \subseteq S$

<u>Proof</u>: See part 1

Examples of subspaces

1) Every vector space V contains a $\underline{0} \in V$

Trivial Subspace: $\{\underline{0}\} \leq V$ is a subspace

2) $\boxed{\mathbb{R}[X] : \text{set of polynomials in } X \text{ with real coefficients of any degree}}$

$\forall n \in \mathbb{N}$, $\mathbb{R}_n[X] \subsetneq \mathbb{R}[X]$ is a subspace for every n

<u>Basis</u>: $\mathbb{R}_n[X] . \{1, x, x^2, \cdots, x^n\}$

deg $\leq n$

Basis for $\mathbb{R}[X]$: $\{1, x, x^2, \cdots\}$ infinite basis

3) Matrices:

$$\boxed{M_{p \times n}(\mathbb{F}) : \text{vector space of } p \times n \text{ matrices}}$$

Standard Basis:

$\{E_{jk} \mid 1 \leq j \leq p, 1 \leq k \leq n\}$ where E_{jk} is the matrix with 1 in jk-entry, 0 elsewhere

4) $M_{2\times 2}(\mathbb{F})$:

$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $\quad E_{21} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$E_{12} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ $\quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

5) Let $\boxed{\mathcal{C}([a,b],\mathbb{R}) \text{ be the set of continuous functions: } f:[a,b] \longrightarrow \mathbb{R}}$

Then $\mathcal{C}([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})$

Identity: The $\underline{0}$ function: $\underline{0}:[a,b] \to \mathbb{R}$

$$\underline{0}(x) = 0 \quad \forall x \in [a,b]$$

is constant, hence continuous

Linear Combination: If $f, g \in \mathcal{C}([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})$, then

$\alpha f + \beta g \in \mathcal{F}([a,b],\mathbb{R})$ is continuous

$\Longrightarrow \alpha f + \beta g \in \mathcal{C}([a,b])$

Hence $\mathcal{C}([a,b],\mathbb{R})$ is a subspace

6) Let $I([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})$ is the set of all integrable functions

$$f:[a,b] \to \mathbb{R}$$

such that integral

$$\int_a^b f(x)\,dx \quad \text{exists}$$

$\underline{0}$ is integrable

Linear combination of integrable functions is integrable

Hence $I([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})$ is a subspace

Continuous functions are integrable.

$C([a,b],\mathbb{R})$ is a subspace of $I([a,b],\mathbb{R})$

Note: $\mathbb{R}_n[X] \subseteq \mathbb{R}[X] \subseteq \mathcal{C}([a,b],\mathbb{R}) \subseteq \mathcal{I}([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})$

These are all $\subsetneq$ as $a<b$

7) In $\mathcal{F}([0,2\pi],\mathbb{R})$ are $x, \sin(x), e^x$ linearly independent

Consider $\alpha x + \beta \sin x + \gamma e^x = 0$

Need to hold for all $x \in [0,2\pi]$

At $x=0$, $\gamma e^0 = 0 \implies \gamma = 0$

$$\implies \alpha x + \beta \sin(x) = 0$$

$$\implies \alpha + \beta\cos(x) = 0 \qquad \text{differentiating}$$

$x = \frac{\pi}{2} : \alpha = 0$

Hence

$$\beta\sin(x) = 0 \quad \forall x \in [0,2\pi] \implies \beta = 0$$

$$\mathbb{R}[X] \subseteq \underbrace{C([a,b],\mathbb{R}) \subseteq \mathcal{I}([a,b],\mathbb{R}) \subseteq \mathcal{F}([a,b],\mathbb{R})}_{\text{can't have finite basis}}$$

Aside: Basis for $C([a,b],\mathbb{R})$?

Can't use Taylor series; not linear combinations of $\{1, x, x^2, \dots\}$ since linear combinations are finite sum

Upshot: every vector space has a basis, but can be impossible to describe it for ∞ dim spaces

Dimensions

Definition Dimensions

For any subspace $S \subseteq V$, we define dimension of S by

$$\dim(S) = \#(\text{basis of } S) \qquad \text{cardinality}$$

Theorem

Let V be a vector space of over $\mathbb{F}$ with a finite basis. Then every basis of V has the same number of elements

Note: Steinitz Exchange Lemma holds for any vector space with a finite basis

Properties of dimensions and basis

Lemma

Let V be an n-dimensional vector space $S \subseteq V$ be a subspace. Then S has a finite basis.

Let $q = \dim(S)$

(0) Every linear independent set of vectors $\{\underline{u}_1, \cdots, \underline{u}_t\} \subset S$ can be extended to a basis of S

(i) Any linearly independent subset $\mathcal{C}$ has no more than q elements

(ii) Any linearly independent subset $\mathcal{C} \subseteq \mathbb{F}^n$ can be extended to a basis of $\mathbb{F}^n$

(iii) Any finite spanning set for S contains a basis of S

Hence no subset containing fewer than q elements span S

(iv) Any linearly independent subset of S containing q elements spans S so it is a basis of S

Similarly if a set of size q spans S then it is linearly independent and its a basis.

(v) If $q=0$, then $S = \{\underline{0}\}$. If $q=n$, then $S = \mathbb{F}^n$

Examples:

i) $\mathbb{R}_n[x]$: $n+1$ dimension, since it has basis $\{1, x, \cdots, x^n\}$

ii) $M_{p \times n}(\mathbb{F})$: has dim $p \cdot n$, standard basis

iii) Let $V = M_{2\times 2}(\mathbb{R})$

$S \subseteq V$: Symmetric matrix is a subspace

proof:

i) $\forall A, B \in S, \ \alpha, \beta \in \mathbb{R}$

$$(\alpha A + \beta B)^T = \alpha A^T + \beta B^T$$

$$= \alpha A + \beta B \qquad \text{Symmetric} \Longrightarrow A^T = A$$

$$\Rightarrow \alpha A + \beta B \in S$$

ii) $0_{n \times n}$ is symmetric $\Rightarrow 0 \in S$

$M_{2\times 2}(\mathbb{R})$ has dim 4 $\Longrightarrow$ $\dim(S) \leq 4$

We can write elements of S uniquely as

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$\Longrightarrow$ $\mathcal{B} = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is a basis for S

$\Longrightarrow$ $\dim(S) = 3$

Another way to see this, is

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$$

has 3 free variables $\Longrightarrow$ 3 basis and $\dim S = 3$

We can extend basis of S to V by adding one more linearly independent matrix. To find just find one not in span of $\mathcal{B}$

$$M \notin Sp(\mathcal{B}) \Longrightarrow M \notin S$$
$$\Longrightarrow M \text{ is not symmetric}$$

Let M = any matrix outside S, example

$$E_{12} - E_{21} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

DIRECT SUMS

Direct Sums

Definition Sum of Subspaces

Let V be a vector space

Let $S_1, \dots, S_q \subset V$ be subspaces. Then sum

$$S_1 + S_2 + \dots + S_q = Sp(S_1 \cup \dots \cup S_q) = \{\alpha_1 \underline{v}_1 + \dots + \alpha_q \underline{v}_q \mid \alpha_j \in \mathbb{F}, \underline{v}_j \in S_j\}$$

When

$$S_j \cap \left(\sum_{k \neq j} S_k \right) = \{\underline{0}\} \qquad \forall\, 1 \leq j \leq q$$

we call this the direct sum denoted

$$S_1 \oplus S_2 \oplus \dots \oplus S_q = \bigoplus_{j=1}^{q} S_j$$

Theorem

For any subspaces $S_1, \dots, S_q \in V$

i) $S_1 \cap \dots \cap S_q$ is a subspace

(ii) $S_1 + \dots + S_q$ is a subspace

Proof: Part 1

Lemma

Let S_1, S_2 be subspaces of vector space V. Then

$$\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)$$

In particular for direct sum

$$\dim(S_1 \oplus S_2) = \dim(S_1) + \dim(S_2)$$

Lemma

Let $S_1 \oplus S_2 \oplus \dots \oplus S_q$ be a direct sum of subspaces and

$$\underline{v}_j \in S_j \setminus \{\underline{0}\} \text{ (non-zero)} \quad \text{for } j = 1, \dots, q$$

Then $\underline{v}_1, \dots, \underline{v}_q$ are linearly independent

LINEAR MAPS

Definition Linear Maps

Let V, W be vector spaces over the same field $\mathbb{F}$.

A map $L: V \longrightarrow W$ is called linear map if

$$L(\alpha \underline{u} + \beta \underline{v}) = \alpha L(\underline{u}) + \beta L(\underline{v}) \qquad \forall \alpha, \beta \in \mathbb{F}, \ \forall \underline{u}, \underline{v} \in V$$

In abstract algebra, linear maps are referred to as vector space homomorphism, since they like other homomorphisms, they are structure-preserving maps.

Therefore we denote the set of all linear maps from V to W by

$$\mathrm{Hom}(V, W)$$

Lemma

Let U, V, W be vector spaces over the same field $\mathbb{F}$.

If $L, M \in \mathrm{Hom}(V, W)$ and $\alpha, \beta \in \mathbb{F}$. Then $\alpha L + \beta M$ defined by

$$(\alpha L + \beta M)(v) = \alpha L(v) + \beta M(v) \qquad v \in V$$

is also a linear map.

Also if $L \in \mathrm{Hom}(V, W)$ and $K \in \mathrm{Hom}(U, V)$, then the composite

$$L \circ K \in \mathrm{Hom}(U, W)$$

Matrix representing linear maps

Let V and W be finite dimensional vector space.

Pick an ordered basis for V and W

- $(v_1, v_2, \cdots, v_n)$ be a basis for V
- $(w_1, w_2, \cdots, w_n)$ be a ordered basis W.

The linear map $L: V \longrightarrow W$ can be represented by an $m \times n$ matrix whose jth column is given by $L(v_j)$

$$\boxed{L(v_j) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_m w_m \longmapsto \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}}$$

co-efficients

Examples of matrices representing Linear maps.

1) $V = \mathbb{R}_2[X]$

$$L : V \longrightarrow V$$
$$p(x) \longmapsto p'(x)$$

This is a linear map since differentiation is a linear operation

$$(f+g)' = f' + g'$$
$$(\alpha f)' = \alpha f'$$

Explicitly $\quad L(\alpha_0 + \alpha_1 x + \alpha_2 x^2) = \alpha_1 + 2\alpha_2 x$

An ordered basis for V is

$$(v_1, v_2, v_3) = (1, x, x^2)$$

Calculating effect of L on basis

$$L(v_1) = L(1) = 0.1 + 0 \cdot x + 0 \cdot x^2$$
$$L(v_2) = L(x) = 1.1 + 0 \cdot x + 0 \cdot x^2$$
$$L(v_3) = L(x^2) = 0.1 + 2 . x + 0 . x^2$$

So matrix representing L is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

j th column consists of co-efficients of $L(v_j)$ w.r.t this basis

2) Let $V = M_{2\times2}(\mathbb{R})$

$$L : V \longrightarrow V \text{ given by}$$
$$L(A) = A^T$$

We saw above, transpose respects linear combination $\Longrightarrow$ L is a linear map.

Pick an ordered basis for V

$$v_1 = E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad v_2 = E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Calculating effect of L on basis

$$L(v_1) = v_1 = 1\cdot v_1 + 0\cdot v_2 + 0\cdot v_3 + 0\cdot v_4$$

$$L(v_2) = v_3 = 0\cdot v_1 + 0\cdot v_2 + 1\cdot v_3 + 0\cdot v_4$$

$$L(v_3) = v_2 = 0\cdot v_1 + 1\cdot v_2 + 0\cdot v_3 + 0\cdot v_4$$

$$L(v_4) = v_4 = 0\cdot v_1 + 0\cdot v_2 + 0\cdot v_3 + 1\cdot v_4$$

So matrix representing L is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3) $V = \mathbb{R}[x]$

$$L : V \longrightarrow V; \quad p(x) \longmapsto p'(x)$$

L is still linear.

$\mathbb{R}[x]$ does not have a finite basis $\Longrightarrow$ no matrix representation.

4) $V = I([a,b], \mathbb{R})$ and define

$$L : V \longrightarrow V$$

$$L(f) = \int_a^x f(t)\, dt$$

Fundamental Theorem of Calculus tells us that

$L(f)$ itself is integrable $\Longrightarrow$ $L(f) \in V$

$\Longrightarrow$ well-defined map

Integration is a linear map $\Longrightarrow$ L is a linear map.

But $I([a,b], \mathbb{R})$ is not finite dimensional $\Longrightarrow$ cannot represent L by a matrix.

Images and Kernels

Definition Image and Kernel

Let L be a Linear map from V to W; $L: V \longrightarrow W$

Image of L: $Im(L) = \{ w \in W \mid w = L(v) \text{ for some } v \in V \}$

Kernel of L: $Ker(L) = \{ v \in V \mid L(v) = 0 \}$ also called null-space

Lemma

Suppose L is a linear map $L: V \longrightarrow W$

- $Im(L)$ is a subspace of W
- $Ker(L)$ is a subspace of V

If V has a finite basis $\{v_1, \dots, v_n\}$ then $Im(L)$ is spanned by

$$L(v_1), \dots, L(v_n)$$

Proof: proof of last part

Assume $L: V \longrightarrow W$ is a linear map, V has basis

Let $w \in Im(L) \implies w = L(v)$ for some $v \in V$

We can write $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ for some $\alpha_1, \dots, \alpha_n \in \mathbb{F}$

$$\begin{aligned} w = L(v) &= L(\alpha_1 v_1 + \dots + \alpha_n v_n) \\ &= \alpha_1 L(v_1) + \dots + \alpha_n L(v_n) \end{aligned}$$

$$\implies w \in Span(L(v_1), \dots, L(v_n))$$

$$\implies Im(L) \subseteq Sp(L(v_1), \dots, L(v_n))$$

The other inclusion in the other direction is true since by definition, each

$$L(v_j) \in Im(L)$$

So we get

$$Im(L) = Sp(L(v_1), \dots, L(v_n))$$

$\blacksquare$

Injective, Surjective, isomorphisms

Definition Let $L: V \longrightarrow W$ be a linear map

- L is one-to-one (injective) if $L(u_1) = L(u_2) \implies u_1 = u_2$
- L is onto (surjective) if $\forall\, w \in W \;\exists\, v \in V$ s.t $L(v) = w$ $(Im(L) = W)$
- L is bijective if f is both one to one and onto

Lemma

A linear map $L: V \longrightarrow W$ between vector spaces over the same field $\mathbb{F}$

i) one to one $\iff Ker(L) = \{\underline{0}\}$

ii) Hence L is bijective when $Ker(L) = \{\underline{0}\}$ and $W = Im(L)$

iii) When L is bijective, it has an inverse

$$L^{-1}: W \longrightarrow V$$

which is also a linear map

<u>Proof</u>: of (iii)

Let $\alpha, \beta \in \mathbb{F}$, $w_1, w_2 \in W$. We want to show that

$$L^{-1}(\alpha w_1 + \beta w_2) = \alpha L^{-1}(w_1) + \beta L^{-1}(w_2). \qquad (*)$$

Apply L to LHS of $(*)$

$$L(L^{-1}(\alpha w_1 + \beta w_2)) = \alpha w_1 + \beta w_2$$

since L and L^{-1} are inverses and $L \circ L^{-1} = I_W$

Applying L to RHS of $(*)$

$$L(\alpha L^{-1}(w_1) + \beta L^{-1}(w_2)) = \alpha L(L^{-1}(w_1)) + \beta L(L^{-1}(w_2)) \quad \text{since } L \text{ is linear}$$

$$= \alpha w_1 + \beta w_2 \quad L \circ L^{-1} = I_W$$

Therefore, we have shown

$$L(L^{-1}(\alpha w_1 + \beta w_2)) = L(\alpha L^{-1}(w_1) + \beta L^{-1}(w_2))$$

L is bijective $\implies L$ is injective

$\implies L^{-1}(\alpha w_1 + \beta w_2) = \alpha L^{-1}(w_1) + \beta L^{-1}(w_2)$ ∎

Definition Isomorphism

An invertible linear map

$$L: V \longrightarrow W$$

is called a vector space isomorphism

We say V is isomorphic to W denoted $V \cong W$ if such a map exists

Remark: Composition of 2 linear (bijective) maps gives a linear (bijective) map

Composition of a linear map is linear and composition of bijections is a bijection.

Hence

1) $U \cong V$ and $V \cong W \implies U \cong W$
2) $V \cong V$ via identity map: $I_V : V \to V$; $I_V(v) = v$
3) $V \cong W \implies W \cong V$ since isomorphisms are invertible

$\Big\}$ Equivalence relation

Rank-Nullity

Definition Rank/Nullity

Let L be a linear map.

Rank of L, $rk(L)$ is the dimension of $Im(L)$

$$rk(L) = \dim(Im(L))$$

Nullity of L, $null(L)$ is the dimension of $ker(L)$

$$null(L) = \dim(ker(L))$$

Theorem Rank-Nullity Theorem

Let V, W be finite dimensional vector spaces over same field $\mathbb{F}$

For a linear map $L: V \to W$

$$\dim(V) = rk(L) + null(L)$$

Example:

Let $V = W = \mathbb{R}_n[x]$ and define

$$L : V \longrightarrow V;$$
$$p(x) \longmapsto p'(x)$$

for any arbitrary $p(x) \in \mathbb{R}_n[x]$

$$p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n \quad \text{for some } \alpha_j \in \mathbb{R}$$

Then

$$L(p(x)) = p'(x) = \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 + \cdots + n\alpha_n x^{n-1} \in \mathbb{R}_{n-1}[x]$$

$$\Longrightarrow \mathrm{Im}(L) = \mathbb{R}_{n-1}[x]$$

But note that every polynomial in $\mathbb{R}_{n-1}[x]$ can be obtained like this

$$\Longrightarrow \mathrm{Im}(L) = \mathbb{R}_{n-1}[x]$$

Therefore $\boxed{\mathrm{rank}(L) = \dim(\mathrm{Im}(L)) = n}$

Finding kernel

$$L(p(x)) = p'(x) = 0 \iff p(x) \text{ is a constant polynomial}$$
$$\iff p(x) = \alpha_0 \quad \forall x, \text{ for some } \alpha_0 \in \mathbb{F}$$
$$\iff \ker(L) = \mathrm{sp}(1) = \{\alpha_0 \mid \alpha_0 \in \mathbb{R}\}$$

Therefore $\boxed{\mathrm{null}(L) = \dim(\ker L) = 1}$

By rank-nullity theorem,

$$\dim(V) = n + 1$$

Constructing matrix, using ordered basis for $V = \mathbb{R}_n[x]$

$$(v_1, v_2, \dots, v_{n+1}) = (1, x, \dots, x^n)$$

Observe that

$$L(v_j) = L(x^{j-1}) = (x^{j-1})' = (j-1)x^{j-2} = (j-1)v_{j-1}$$

We get an $(n+1) \times (n+1)$ matrix wrt to ordered basis

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$\text{rank}(A) = n = n$ linearly independent column vectors.

$\text{null}(A) = (n+1) - \text{rank} A = 1$

Corollary

If V and W are vector spaces over the same field and $\dim(V) = \dim(W)$, then

$$V \cong W \quad \text{isomorphic}$$

In particular, every n-dimensional vector space over $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$

Proof:

Sufficient to find an isomorphism $\psi : V \longrightarrow \mathbb{F}^n$ whenever $\dim(V) = n$

Then $V \cong \mathbb{F}^n$ and $W \cong \mathbb{F}^n \Longrightarrow V \cong W$

Take any ordered basis $B = (v_1, \dots, v_n)$ of V and define

$$\boxed{\begin{gathered} \psi_B : V \to \mathbb{F}^n \\ \psi_B\left(\sum_{j=1}^{n} \alpha_j v_j\right) = \sum_{j=1}^{n} \alpha_j \underline{e}_j = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \end{gathered}}$$

Checking ψ_B is linear, check addition and scalar multiplication

<u>Addition</u>:

$$\psi_B\left(\sum_{j=1}^{n} \alpha_j v_j + \sum_{j=1}^{n} \beta v_j\right) = \psi_B\left(\sum_{j=1}^{n} (\alpha_j + \beta_j) v_j\right)$$

$$= \sum_{j=1}^{n} (\alpha_j + \beta_j)\, \underline{e}_j = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_j + \beta_j \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

$$= \psi_B\left(\sum_{j=1}^{n} \alpha_j v_j\right) + \psi_B\left(\sum_{j=1}^{n} \beta_j v_j\right)$$

Scalar Multiplication:

$$\psi_B\left(\gamma\left(\sum_{j=1}^{n}\alpha_j v_j\right)\right) = \psi_B\left(\sum_{j=1}^{n}(\gamma\alpha_j)v_j\right) = \sum_{j=1}^{n}(\gamma_j\alpha_j)\underline{e}_j$$

$$= \begin{pmatrix}\gamma\alpha_1\\ \vdots \\ \gamma\alpha_n\end{pmatrix} = \gamma\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}$$

$$= \gamma\left(\psi_B\left(\sum_{j=1}^{n}\alpha_j v_j\right)\right)$$

Calculating Kernel $\operatorname{Ker}(\psi_B)$:

$$\sum_{j=1}^{n}\alpha_j v_j \in \operatorname{Ker}(\psi_B) \iff \psi_B\left(\sum_{j=1}^{n}\alpha_j v_j\right) = \underline{0}$$

$$\iff \begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = \begin{pmatrix}0\\ \vdots \\ 0\end{pmatrix}$$

$$\iff \alpha_j = 0 \quad \forall j$$

Therefore

$$\operatorname{Ker}(\psi_B) = \{\underline{0}\} \implies \operatorname{null}(L) = 0$$

By rank-nullity theorem,

$$\operatorname{rank}(\psi_B) = n \implies \operatorname{Im}(\psi_B) = \mathbb{F}^n$$

since only n dimensional subspace of $\mathbb{F}^n$ is $\mathbb{F}^n$ itself.

Hence by Lemma on pg 18, ψ_B is an isomorphism ■

Given an ordered basis $B = (v_1, \dots, v_n)$, we define co-ordinate map

$$\psi_B : V \longrightarrow \mathbb{F}^n$$

$$\psi_B(\alpha_1 v_1 + \dots + \alpha_n v_n) = \sum_{i=1}^{n}\alpha_n \underline{e}_n = \begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}$$

This is an isomorphism, unique for which

$$\psi(v_j) = \underline{e}_j$$

Conversely vector space isomorphisms matches bases to bases. So given vector space isomorphism

$$\psi : V \longrightarrow \mathbb{F}^n$$

its inverse $\psi^{-1} : V \to \mathbb{F}^n$ is also an isomorphism and maps standard basis $(\underline{e}_1, \cdots, \underline{e}_n)$ of $\mathbb{F}^n$ to an ordered basis $(v_1, \dots, v_n)$ of V, $v_j = \psi^{-1}(\underline{e}_j)$

Lemma

For a finite dimensional vector space V, there is a bijective correspondence between coordinate maps (isomorphism)

$$\psi : V \longrightarrow \mathbb{F}^n$$

and ordered basis $B = (v_1, \dots, v_n)$ of V

CHANGE OF BASIS MATRIX

Let V be a vector space over $\mathbb{F}$ of dimension n; $\dim(V) = n$

Let $\mathcal{A}$ and $\mathcal{B}$ be 2 ordered basis for V

$$\mathcal{A} = (w_1, \dots, w_n)$$

$$\mathcal{B} = (v_1, \dots, v_n)$$

Have co-ordinate maps

$$\psi_{\mathcal{B}} : V \longrightarrow \mathbb{F}^n$$

$$\psi_{\mathcal{A}} : V \longrightarrow \mathbb{F}^n$$

$\psi_{\mathcal{A}}$ is an isomorphism, hence invertible

$$\psi_{\mathcal{A}}^{-1} \circ \psi_{\mathcal{A}} = I_V$$

Hence

$$\psi_{\mathcal{B}} = \underbrace{\psi_{\mathcal{B}} \circ \psi_{\mathcal{A}}^{-1}}_{\text{Linear map } \mathbb{F}^n \to \mathbb{F}^n} \circ \psi_{\mathcal{A}}$$

So can represent $\psi_{\mathcal{B}} \circ \psi_{\mathcal{A}}^{-1}$ by $n \times n$ matrix called change of basis matrix

↳ $C_{\mathcal{A}}^{B}$: from A to B

Notation: Change of basis matrix : Transition matrix from $\mathcal{A}$ to $\mathcal{B}$

$$C_{\mathcal{A}}^{\mathcal{B}}$$

Writing as a matrix

$$\psi_{\mathcal{B}} = C_{\mathcal{A}}^{\mathcal{B}} \psi_{\mathcal{A}}$$

i.e. $\psi_{\mathcal{B}}(v) = C_{\mathcal{A}}^{\mathcal{B}} \psi_{\mathcal{A}}(v)$

To find matrix, apply $\psi_B \circ \psi_A^{-1}$ to standard basis $(\underline{e}_1, \cdots, \underline{e}_n)$ of $\mathbb{F}^n$

$$(\psi_B \circ \psi_A^{-1})(\underline{e}_j) = \psi_B(\psi_A^{-1}(\underline{e}_j))$$
$$= \psi_B(w_j) \quad \text{since } \psi_A(w_j) = \underline{e}_j$$

$\Longrightarrow$ j^{th} column in matrix is given by co-ordinates of w_j from A written in terms of basis of B

Given 2 ordered basis for vector space V

$$A = (w_1, \cdots, w_n)$$
$$B = (v_1, \cdots, v_n)$$

The change matrix is given by writing each $w_j \in A$ in terms of basis B

j^{th} column $\quad w_j = C_{1j} v_1 + \cdots + C_{nj} v_n \qquad$ for some $C_{1j}, \cdots, C_{1n}$

Then we see that the j-th column of the matrix C_A^B is given by

$$\psi_B(w_j) = C_{1j} \underline{e}_j + \cdots + C_{nj} \underline{e}_n = \begin{pmatrix} C_{1j} \\ \vdots \\ C_{nj} \end{pmatrix}$$

$$C_A^B = \begin{pmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{pmatrix}$$

Multiplication by C_A^B converts co-ordinates w.r.t A into co-ordinates w.r.t B

$$\psi_B(v) = C_A^B \psi_A(v) \qquad \forall\, v \in V$$

Lemma

Let ψ_A and ψ_B be 2 co-ordinate maps on a finite dimensional vector space V, corresponding to ordered basis

$$A = (w_1, \cdots, w_n)$$
$$B = (v_1, \cdots, v_n)$$

Then

$$\psi_B = C_A^B \psi_A$$

where C_A^B is the change of basis matrix defined above, whose columns are co-ordinate basis of A in terms of B

Change of basis matrices posess some natural properties, which are easily proven from the defining equation

$$\psi_{\mathcal{B}}(w_j) = C^{\mathcal{B}}_{\mathcal{A}} \underline{e}_j$$

Lemma

For 3 bases $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{G}$, we have

$$C^{\mathcal{G}}_{\mathcal{A}} = C^{\mathcal{G}}_{\mathcal{B}} C^{\mathcal{B}}_{\mathcal{A}}$$

Since

$$C^{\mathcal{A}}_{\mathcal{A}} = I_n \quad \text{identity matrix}$$

it follows $C^{\mathcal{A}}_{\mathcal{B}} = (C^{\mathcal{B}}_{\mathcal{A}})^{-1}$

Examples

1) $V = \mathbb{R}_2[X]$

$\mathcal{B} = (1, x, x^2)$

$\mathcal{A} = (1+x, x, 1+x^2)$

Check that $\mathcal{A}$ is a basis (check linear independence)

$$\alpha(1+x) + \beta x + \gamma(1+x^2) = 0 \quad \text{for some } \alpha, \beta, \gamma \in \mathbb{F}$$

$$\iff (\alpha+\gamma) + (\alpha+\beta)x + \gamma x^2 = 0$$

$$\iff \alpha+\beta = 0, \quad \gamma = 0, \quad \alpha+\gamma = 0$$

$$\iff \alpha = \gamma = \beta = 0$$

$\dim(V) = 3$, 3 linearly independent vectors $\Longrightarrow$ forms a basis for V

Co-ordinate map for $\mathcal{B}$

$$\psi_{\mathcal{B}} : V \longrightarrow \mathbb{R}^3$$

$$\psi_{\mathcal{B}}(\alpha_0 + \alpha_1 x + \alpha_2 x^2) = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}$$

To work out $C^{\mathcal{A}}_{\mathcal{B}}$, write elements of $\mathcal{B}$ in terms of $\mathcal{A}$

$$1 = 1(1+x) + (-1)\cdot x + 0\,x^2$$

$$x = 0\cdot(1+x) + 1\cdot x + 0(1+x^2)$$

$$x^2 = (-1)(1+x) + 1\cdot x + 1(1+x^2)$$

$$C^A_B = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

To find C^B_A, either compute inverse of C^A_B or follow same method

Express the vectors in A in terms of B

$$1+x = 1\cdot 1 + 1\cdot x + 0\cdot x^2$$

$$x = 0\cdot 1 + 1\cdot x + 0\cdot x^2$$

$$1+x^2 = 1\cdot 1 + 0\cdot x + 1\cdot x^2$$

Thus

$$C^B_A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Check that $C^A_B C^B_A = I_3$

$$C^A_B C^B_A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Since $\psi_A = C^A_B \psi_B$, it follows that the co-ordinate map $\psi_B : V \longrightarrow \mathbb{R}^3$ is given by:

Since $\psi_A(v) = C^A_B \psi_B(v)$, we get

$$\psi_A : V \longrightarrow \mathbb{R}^3;$$

$$\psi_A(\alpha_0 + \alpha_1 x + \alpha_2 x^2) = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 - \alpha_2 \\ -\alpha_0 + \alpha_1 + \alpha_2 \\ \alpha_2 \end{pmatrix}$$

Multiplication by C^A_B takes co-ordinate vectors written in terms of B to co-ordinate vectors in terms of A.

For a concrete example, let $p(x) = 1 + 2x + 3x^2$

$p(x)$ in terms of B:

$$\psi_B(p(x)) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Multiplying by C^A_B gives us

$$C^A_B \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$$

which is co-ordinate vector of $p(x)$ in terms of A since

$$(-2)\cdot(1+x)+4\cdot x+3\cdot(1+x^2) = -2-2x+4x+3+3x^2 = 1+2x+3x^2 = p(x)$$

We have verified

$$C_B^A \psi_B(p(x)) = C_B^A\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}-2\\4\\3\end{pmatrix} = \psi_A(p(x))$$

2) Let linear map on $V=\mathbb{R}_2[x]$ be

$$L: V \longrightarrow V$$

$$p(x) \longmapsto p'(x)$$

L represented by matrix w.r.t $\mathcal{B}$ be

$$M_B(L) = \begin{pmatrix}0&1&0\\0&0&2\\0&0&0\end{pmatrix}$$

Finding matrix w.r.t $\mathcal{A}$

$$L(1+x) = (1+x)' = 1 = 1(1+x)+(-1)(x)+0(1+x^2)$$

$$L(x) = (x)' = 1 = 1(1+x)+(-1)(x)+0(1+x^2)$$

$$L(x^2) = (x^2)' = 2x = 0(1+x)+2(x)+0(1+x^2)$$

Hence

$$M_A(L) = \begin{pmatrix}1&1&0\\-1&-1&2\\0&0&0\end{pmatrix}$$

Matrix representing Linear map revisited

$L: V \longrightarrow V$ is a linear map and $\mathcal{B}=(v_1,\cdots,v_n)$ an ordered basis for V,

Matrix representing L w.r.t basis $\mathcal{B}$ is denoted $M_B(L)$

L is uniquely determined by its action on the basis vectors of $\mathcal{B}$, so the j^{th} column of $M_B(L)$ can be computed by applying L to basis vector v_j and writing co-efficients w.r.t $\mathcal{B}$ as a column vector.

$\Longrightarrow$ matrix obtained by applying L to v_j writing in terms of $\mathcal{B}$ and writing co-ordinates as jth column. (i.e. applying ψ_B)

$$M_B(L): \mathbb{F}^n \longrightarrow \mathbb{F}^n$$

Equivalently basis $\mathcal{B}$ gives V the co-ordinate map $\psi_B: V \longrightarrow \mathbb{F}^n$

So we take $\underline{e}_j$, apply ψ_B^{-1} to get v_j, apply L, then apply ψ_B to get co-ordinate vector

$$M_B(L): \mathbb{F}^n \xrightarrow{\psi_B^{-1}} V \xrightarrow{L} V \xrightarrow{\psi_B} \mathbb{F}^n$$

Hence define $\boxed{M_{\mathcal{B}}(L) = \psi_{\mathcal{B}} \circ L \circ \psi_{\mathcal{B}}^{-1}}$

Also use notation $M_{\mathcal{B}}(L)$ to denote matrix representing this map w.r.t standard basis of $\mathbb{F}^n$

These descriptions are equivalent.

The matrix $M_{\mathcal{B}}(L)$ for a given basis $\mathcal{B} = (v_1, \dots, v_n)$ is obtained by applying L to v_j, writing result in terms of $\mathcal{B}$ and then writing coordinates obtained as j^{th} column vector.

We can describe this in terms of the co-ordinate map $\psi_{\mathcal{B}}$

So we take $\underline{e}_j$, apply $\psi_{\mathcal{B}}^{-1}$ to get v_j, apply L, then apply $\psi_{\mathcal{B}}$ to get co-ordinate in terms of $\mathcal{B}$.

More concretely, since $\mathcal{B}$ is a basis,

$$\boxed{L(v_j) = A_{1j} v_1 + \cdots + A_{nj} v_n \quad \text{for some } A_{ij}, \dots, A_{nj} \in \mathbb{F}}$$

Then $M_{\mathcal{B}}(L)$ is the $n \times n$ matrix (A_{ij})

<u>proof:</u> Recall that $M_{\mathcal{B}}(L)\underline{e}_j$ gives the j^{th} column of $M_{\mathcal{B}}(L)$.

Now

$$\begin{aligned} M_{\mathcal{B}}(L)\underline{e}_j &= (\psi_{\mathcal{B}} \circ L \circ \psi_{\mathcal{B}}^{-1})(\underline{e}_j) \\ &= \psi_{\mathcal{B}}(L(\psi_{\mathcal{B}}^{-1}(\underline{e}_j))) \\ &= \psi_{\mathcal{B}}(L(v_j)) \\ &= \psi_{\mathcal{B}}(A_{1j} v_1 + \cdots + A_{nj} v_n) \\ &= A_{1j} \psi_{\mathcal{B}}(v_1) + \cdots + A_{nj} \psi_{\mathcal{B}}(v_n) \\ &= A_{1j} \underline{e}_j + \cdots + A_{nj} \underline{e}_j \end{aligned}$$

Therefore

$$M_{\mathcal{B}}(L) = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

Properties of matrix representing a Linear map

Proposition

Let V be a vector space over field $\mathbb{F}$, $L_1, L_2 : V \longrightarrow V$ be 2 linear transformations and $\mathcal{B}$ be a basis for V. Then:

(i) $\forall$ scalars $\alpha, \beta \in \mathbb{F}$,

$$M_{\mathcal{B}}(\alpha L_1 + \beta L_2) = \alpha M_{\mathcal{B}}(L_1) + \beta M_{\mathcal{B}}(L_2)$$

(ii) $M_{\mathcal{B}}(L_1 \circ L_2) = M_{\mathcal{B}}(L_1) M_{\mathcal{B}}(L_2)$

In particular,

$$L : V \longrightarrow V \text{ is invertible} \implies M_{\mathcal{B}}(L^{-1}) = M_{\mathcal{B}}(L)^{-1}$$

Proof:

(ii) We use $M_{\mathcal{B}}(L) = \psi_{\mathcal{B}} \circ L \circ \psi_{\mathcal{B}}^{-1}$

Hence matrix multiplication corresponds to composition of Linear maps (and we can interpret them as $n \times n$ matrix as a linear map $\mathbb{F}^n \to \mathbb{F}^n$)

We have

$$\begin{aligned} M_{\mathcal{B}}(L_1) M_{\mathcal{B}}(L_2) &= (\psi_{\mathcal{B}} \circ L_1 \circ \psi_{\mathcal{B}}^{-1}) \circ (\psi_{\mathcal{B}} \circ L_2 \circ \psi_{\mathcal{B}}^{-1}) \\ &= \psi_{\mathcal{B}} \circ L_1 \circ (\psi_{\mathcal{B}} \circ \psi_{\mathcal{B}}^{-1}) \circ L_2 \circ \psi_{\mathcal{B}}^{-1} \quad \text{composition of functions is associative} \\ &= \psi_{\mathcal{B}} \circ L_1 \circ L_2 \circ \psi_{\mathcal{B}}^{-1} \\ &= M_{\mathcal{B}}(L_1 \circ L_2) \end{aligned}$$

When L is invertible, we have

$$L^{-1} \circ L = L \circ L^{-1} = I_V$$

Apply to above to get

$$\left.\begin{aligned} M_{\mathcal{B}}(L^{-1}) M_{\mathcal{B}}(L) &= M_{\mathcal{B}}(L \circ L^{-1}) = M_{\mathcal{B}}(I_V) = I_n \\ M_{\mathcal{B}}(L) M_{\mathcal{B}}(L^{-1}) &= M_{\mathcal{B}}(L \circ L^{-1}) = M_{\mathcal{B}}(I_V) = I_n \end{aligned}\right\} \implies M_{\mathcal{B}}(L^{-1}) = M_{\mathcal{B}}(L)^{-1}$$

(i) $M_{\mathcal{B}}(L) : \mathbb{F}^n \longrightarrow \mathbb{F}^n$ is a linear map.

We have $M_{\mathcal{B}}(L) = \psi_{\mathcal{B}} \circ L \circ \psi_{\mathcal{B}}^{-1}$

Let $\underline{v}$ be any arbitrary column vector in $\mathbb{F}^n$, and v be the corresponding vector V w.r.t co-ordinate map $\psi_{\mathcal{B}}$, i.e. we have $v = \psi_{\mathcal{B}}^{-1}(\underline{v}) \implies \underline{v} = \psi_{\mathcal{B}}(v)$

Then

$$\begin{aligned}
M_{\mathcal{B}}(\alpha L_1+\beta L_2) &= (\psi_{\mathcal{B}}\circ(\alpha L_1+\beta L_2)\circ\psi_{\mathcal{B}}^{-1})(\underline{v}) \\
&= (\psi_{\mathcal{B}}\circ(\alpha L_1+\beta L_2)\circ\psi_{\mathcal{B}}^{-1})(\psi_{\mathcal{B}}(v)) \\
&= (\psi_{\mathcal{B}}\circ(\alpha L_1+\beta L_2)\circ\psi_{\mathcal{B}}^{-1}\circ\psi_{\mathcal{B}})(v) \\
&= (\psi_{\mathcal{B}}\circ(\alpha L_1+\beta L_2))(v) \\
&= \psi_{\mathcal{B}}(\alpha L_1(v)+\beta L_2(v)) \\
&= \alpha\psi_{\mathcal{B}}(L_1(v)) + \beta\psi_{\mathcal{B}}(L_2(v)) \\
&= \alpha\psi_{\mathcal{B}}(L_1(\psi_{\mathcal{B}}^{-1}(\underline{v}))) + \beta\psi_{\mathcal{B}}(L_2(\psi_{\mathcal{B}}^{-1}(\underline{v}))) \\
&= \alpha(\psi_{\mathcal{B}}\circ L_1\circ\psi_{\mathcal{B}}^{-1})(\underline{v}) + \beta(\psi_{\mathcal{B}}\circ L_2\circ\psi_{\mathcal{B}}^{-1})(\underline{v}) \\
&= \alpha M_{\mathcal{B}}(L_1)(\underline{v}) + \beta M_{\mathcal{B}}(L_2)(\underline{v}) \\
&= \alpha(M_{\mathcal{B}}(L_1)+\beta M_{\mathcal{B}}(L_2))(\underline{v})
\end{aligned}$$

True $\forall\, \underline{v}\in\mathbb{F}^n \implies$ we have an equality of linear maps $\mathbb{F}^n\longrightarrow\mathbb{F}^n$

$$M_{\mathcal{B}}(\alpha L_1+\beta L_2) = \alpha M_{\mathcal{B}}(L_1)+\beta M_{\mathcal{B}}(L_2)$$

■

Theorem

Let $L: V\longrightarrow V$ be a linear transformation, $\mathcal{A}$, $\mathcal{B}$ be 2 bases for V. Then

$$M_{\mathcal{B}}(L) = C_{\mathcal{A}}^{\mathcal{B}}\, M_{\mathcal{A}}(L)\,(C_{\mathcal{A}}^{\mathcal{B}})^{-1} = (C_{\mathcal{B}}^{\mathcal{A}})^{-1} M_{\mathcal{A}}(L)\, C_{\mathcal{B}}^{\mathcal{A}}$$

In particular, $M_{\mathcal{A}}(L)$ and $M_{\mathcal{B}}(L)$ are similar matrices

__proof__:

We have

$$M_{\mathcal{B}}(L) = \psi_{\mathcal{B}}\circ L\circ\psi_{\mathcal{B}}^{-1}$$
$$M_{\mathcal{A}}(L) = \psi_{\mathcal{A}}\circ L\circ\psi_{\mathcal{A}}^{-1}$$

Also $\psi_{\mathcal{B}}(v) = C_{\mathcal{A}}^{\mathcal{B}}\psi_{\mathcal{A}}(v) \quad \forall v\in V \implies \psi_{\mathcal{B}} = C_{\mathcal{A}}^{\mathcal{B}}\circ\psi_{\mathcal{A}}$

Hence

$$\begin{aligned}
M_{\mathcal{B}}(L) &= \psi_{\mathcal{B}}\circ L\circ\psi_{\mathcal{B}}^{-1} \\
&= (C_{\mathcal{A}}^{\mathcal{B}}\circ\psi_{\mathcal{A}})\circ L\circ(C_{\mathcal{A}}^{\mathcal{B}}\circ\psi_{\mathcal{A}})^{-1} \\
&= C_{\mathcal{A}}^{\mathcal{B}}\circ\underbrace{\psi_{\mathcal{A}}\circ L\circ\psi_{\mathcal{A}}^{-1}}_{M_{\mathcal{A}}(L)}\circ(C_{\mathcal{A}}^{\mathcal{B}})^{-1}
\end{aligned}$$

$$= C_A^B \circ M_A(L) \circ (C_A^B)^{-1}$$ viewed as linear maps $\mathbb{F}^n \to \mathbb{F}^n$

$$= C_A^B M_A(L)(C_A^B)^{-1}$$ viewed as matrices

Previous Example continued

$V = \mathbb{R}_2[x]$

$\mathcal{B} = (1, x, x^2)$

$A = (1+x, x, 1+x^2)$

$$C_A^B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad C_B^A = (C_A^B)^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$L: V \longrightarrow V$ given by $p(x) \longmapsto p'(x)$

$$M_B(L) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$M_A(L) = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Verifying theorem

$$C_A^B M_A(L)(C_A^B)^{-1}$$

$$= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = M_B(L)$$

Remark: If W and V are finite dimensional vector spaces with

$A = (w_1, \dots, w_n)$ a basis for W

$\mathcal{B} = (v_1, \dots, v_n)$ a basis for V

Then any linear map $L: W \to V$ can be represented by a matrix

(given by linear map)

$$M_A^B(L): \mathbb{F}^n \longrightarrow \mathbb{F}^n; \quad M_A^B(L) = \psi_B \circ L \circ \psi_A^{-1}$$

If A' is another basis for W

If B' is another basis for V

then we have change of basis formula

$$M_{A'}^{B'}(L) = C_B^{B'} M_A^B(L) C_{A'}^A$$

EIGENVECTORS AND EIGENVALUES

Notation:
$$L: V \rightarrow V \qquad (L: V \circlearrowleft)$$

Definition

A linear map $L: V \circlearrowleft$

An eigenvector of L is a non-zero vector $\underline{v} \in V$ such that

$$L\underline{v} = \lambda \underline{v} \quad \text{where } \lambda \in \mathbb{F} \text{ scalar}$$

In this case λ is an eigenvalue of L

The same definition applicable to matrices

$$A\underline{v} = \lambda \underline{v}$$

The set of all eigenvalues of L is called the spectrum of L : Spec L

$$\operatorname{Spec} L = \{\lambda \in \mathbb{F} \mid L - \lambda I_n \text{ is not invertible}\}$$

Indeed

$$L\underline{v} = \lambda \underline{v} \iff (L - \lambda \mathrm{Id}_N)\underline{v} = 0$$

Remark: Similar matrices have same eigenvalues

Example:

Recall $V = \mathbb{R}_2[X]$

$$L: V \longrightarrow V, \quad p(x) \longmapsto p'(x)$$

$B = (1, x, x^2)$

$$M_B(L) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Upper Triangular $\Longrightarrow$ eigenvalues are diagonal elements

$\Longrightarrow$ $\lambda = 0,\ a_0 = 3$

The eigenvectors are $\begin{pmatrix}1\\0\\0\end{pmatrix} \Longrightarrow g_\lambda = 1$

L and $M_B(L)$ have same eigenvalues/eigenvectors

Theorem

Let $L: V \circlearrowleft$ be a linear map, V be a finite dimensional vector space over a field $\mathbb{F}$.

Then for each matrix representation $A = M_B(L)$ of L

v is an eigenvector $\Longleftrightarrow$ co-ordinate vector $\underline{v} = \psi_B(v)$ is an eigenvector with eigenvalue λ

Moreover, the characteristic polynomial $\det(\lambda I_n - A)$ depends only on L, not B.
Hence we can define this to be the characteristic polynomial

$$c_L(\lambda) \text{ of } L$$

proof:

For $B = (v_1, \dots, v_n)$ a basis of V

Recall that $\psi_B : V \longrightarrow \mathbb{F}^n$ is the co-ordinate map which is the vector space isomorphism that satisfies

$$\psi_B(v_j) = \underline{e}_j$$

For $v \in V$, let $\underline{v} = \psi_B(v)$

Recall that viewed as linear maps $\mathbb{F}^n \longrightarrow \mathbb{F}^n$, we have

$$M_B(L) = \psi_B \circ L \circ \psi_B^{-1}$$

Let $\lambda \in \mathbb{F}$

Note that applying ψ_B, we have

$$\begin{aligned}\psi_B(L(v)) &= (\psi_B \circ L)(v)\\ &= (\psi_B \circ L \circ \psi_B^{-1} \circ \psi_B)(v)\\ &= (M_B(L) \circ \psi_B)(v)\end{aligned}$$

$$= M_{\mathcal{B}}(L)(\psi_{\mathcal{B}}(v))$$
$$= M_{\mathcal{B}}(L)(\underline{v})$$

and

$$\psi_{\mathcal{B}}(\lambda v) = \lambda \psi_{\mathcal{B}}(v) = \lambda \underline{v}$$

So $L(v) = \lambda v \iff \psi_{\mathcal{B}}(L(v)) = \psi_{\mathcal{B}}(\lambda v)$ bijective

$\iff M_{\mathcal{B}}(L)(\underline{v}) = \lambda \underline{v}$

$\iff A(\underline{v}) = \lambda \underline{v}$

So v is an eigenvector of L with eigenvalue $\lambda \iff \underline{v} = \psi_{\mathcal{B}}(v)$ is an eigenvector of $A = M_{\mathcal{B}}(L)$ with eigenvalue λ

To show $c_A(\lambda)$ does not depend on $\mathcal{B}$, can argue that its roots are eigenvalues of L and only depend on L (and $c_A(\lambda)$ is monic)

Alternative for $\mathcal{A}$ another ordered basis of V, we saw that

$M_{\mathcal{B}}(L)$ and $M_{\mathcal{A}}(L)$ are similar matrices

($\exists$ an invertible matrix $P = C_{\mathcal{A}}^{\mathcal{B}}$ s.t $M_{\mathcal{B}}(L) = P^{-1} M_{\mathcal{A}}(L) P$)

and we saw that similar matrices have same characteristic polynomial (Lemma 2.17) ■

Diagonalizable Linear maps

Definition

We say a linear map $L: V \longrightarrow V$ is diagonalizable when V admits a basis $\mathcal{B}$ for which the matrix $M_{\mathcal{B}}(L)$ representing it is diagonal.

Recall that $n \times n$ matrix A is diagonalizable if

$\exists$ an invertible matrix P for which $P^{-1}AP$ is diagonal.

This happens when eigenvectors of $\mathbb{F}^n$ form a basis of $\mathbb{F}^n$

Using isomorphism $\psi_{\mathcal{B}}: V \longrightarrow \mathbb{F}^n$,

A linear map $L: V \longrightarrow V$ is diagonalizable $\iff$ V admits a basis $\mathcal{B} = (v_1, \dots, v_n)$ where v_j is an eigenvector

Example of infinite dimensions

If V not finite dimensional, situation more complicated

1) Let $V = \mathbb{R}[X]$

For $p(x) \in V$, define

$$L(p(x)) = \int_0^x p(t)\,dt$$

Note that $L(p)$ is a polynomial in x, and integration is linear

No eigenvectors:

$$\text{if } L(p(x)) = \lambda p(x)$$
$$\int_0^x p(t)\,dt = \lambda p(x)$$

Differenting both sides and using fundamental theorem of calculus

$$p(x) = (\lambda p(x))' = \lambda p'(x)$$

if $\lambda = 0$ then $p(x) = 0 \ \forall x \in R \implies$ p is zero polynomial / vector

$\implies$ But 0 vector is NEVER an eigenvector so there is no eigenvector for $\lambda = 0$

$$\text{if } \lambda \neq 0 \implies p(x) = \lambda p'(x)$$
$$\implies \frac{1}{\lambda} = \frac{p'(x)}{p'(x)}$$
$$\implies p(x) = \alpha e^{x/\lambda} \text{ for some } \alpha \in \mathbb{R}$$

not a polynomial

Hence $\operatorname{spec} L = \emptyset$

(2) $V = C^\infty([a,b], \mathbb{R})$: vector space of infinitely differentiable functions $f:[a,b] \longrightarrow \mathbb{R}$

$L: V \longrightarrow V$;

$L(f) = f'$

is a linear map on V

We have

$$L(e^{\lambda x}) = \lambda e^{\lambda x} \quad \forall \lambda \in \mathbb{R}$$

$\Longrightarrow e^{\lambda x}$ is an eigenvector of L for every $\lambda \in \mathbb{R}$

$\text{spec}(L) = \mathbb{R}$

(3) Legendre equation

$$(1-x^2)y'' - 2xy' - \lambda y = 0, \quad \lambda \in \mathbb{R}, \ y \text{ a function of } x$$

Define $L(y) = (1-x^2)y'' - 2xy' \Longrightarrow$ by properties of differentation L is linear

Could view L as a linear map on the space $C^\infty([a,b], \mathbb{R})$

Then Legendre equation becomes

$$L(y) = \lambda y \quad \text{(an eigenvector problem)}$$

If y is a polynomial of degree n, then so is $(1-x^2)y''$ and $-2xy$, we restrict L to $\mathbb{R}_n[x]$ ← finite dim

$L: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x]$ is a linear map

$L(y) = \lambda y$ an eigenvector problem

Represent L by an $(n+1) \times (n+1)$ matrix

For example if $n=2$, use $\mathcal{B} = (1, x, x^2)$ for $V = \mathbb{R}_2[X]$

$$L(1) = (1-x^2)1' - 2x1' = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2$$

$$L(x) = (1-x^2)x'' - 2xx' = -2x = 0 \cdot 1 + (-2) \cdot x + 0 \cdot x^2$$

$$L(x^2) = (1-x^2)(x^2)'' - 2x(x^2)' = (1-x^2)2 - 2x\,2x = 2 - 6x^2 = 2 \cdot 1 + 0 \cdot x + (-6)x^2$$

So w.r.t $\mathcal{B}$, L is represented by the matrix

$$A = M_{\mathcal{B}}(L) = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -2 & 0 \\ 0 & 0 & -6 \end{pmatrix}$$

Eigenvectors:

$\lambda_3 = -6$:

$$\begin{pmatrix} 0-(-6) & 0 & 0 \\ 0 & -2-(-6) & 0 \\ 0 & 0 & -6-(-6) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{cases} 6x_1 + 2x_3 = 0 \\ 4x_2 = 0 \end{cases}$$

$$\Longrightarrow \begin{cases} x_3 = -3x_1 \\ x_2 = 0 \end{cases}$$

So eigenvector of A, $\lambda_3 = -6$ is $v_3 = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}$

Similarly for $\lambda_1 = 0$, $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$\lambda_2 = -2$, $v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

To get corresponding eigenvectors of L, we apply ψ_B

$$\psi_B^{-1}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = 1.1 + 0\cdot x + 0\cdot x^2 = 1$$

$$\psi_B^{-1}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = 0.1 + 1.x + 0\cdot x^2 = x$$

$$\psi_B^{-1}\left(\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}\right) = 1.1 + 0\cdot x - 3x^2 = 1 - 3x^2$$

Get eigenvectors

$$\boxed{p_1(x) = 1, \quad p_2(x) = x, \quad p_3(x) = 1 - x^2}$$

Legendre polynomial

customary scaling $p_3(x) = \frac{1}{2}(3x^2 - 1)$

Matrix A is diagonalizable since has 3 distinct eigenvalues

$\Longrightarrow$ L is diagonalizable

Can also see since $p_1(x)$, $p_2(x)$, $p_3(x)$ are linearly independent in $\mathbb{R}_2[x]$ so they form a basis of eigenvectors

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \qquad P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -6 \end{pmatrix}$$

For higher n, keep same eigenfunctions and get new ones

e.g for $n=3$, also have eigenvalue $\lambda = -12 \implies$ eigenvector $v_4 = \begin{pmatrix} 0 \\ -3 \\ 0 \\ 5 \end{pmatrix}$

$$\implies p_4(x) = -3x + 5x^2 \qquad \left(\text{or } \tfrac{1}{2}(5x^3 - 3x) \text{ rescaled}\right)$$

Legendre polynomials are orthogonal (inner product is 0)

DUAL SPACES

Linear Functional

Definition Linear Functional

For a vector space V over a field $\mathbb{F}$, a linear functional is a linear map

$$L : V \longrightarrow \mathbb{F}$$

i.e. an element of $\mathrm{Hom}(V, \mathbb{F})$

Dual Spaces

Definition Dual Spaces

The space $\mathrm{Hom}(V, \mathbb{F})$ of linear functionals form a vector space over $\mathbb{F}$ called the dual space of V denoted V^*

Example: $V = \mathbb{R}^3$ (column vectors)

$V = \mathbb{R}^3$ (column vectors), then V^* can be viewed as row vectors since these "act" on column vectors by matrix multiplication

Matrix multiplication is linear and outputs in $\mathbb{R}$

$$(y_1, y_2, y_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = y_1x_1 + y_2x_2 + y_3x_3 \in \mathbb{R}$$

Standard basis for $V = \mathbb{R}^3$: $\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ $\quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ $\quad \underline{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

Dual basis for V^* is $f_1 = (1\ 0\ 0)$ $\quad f_2 (0 1 0)$ $\quad f_3 (0 0 1)$

e.g

$$f_1(\underline{e}_1) = (1 0 0)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 \qquad f_1(\underline{e}_2) = \underline{0}$$

Isomorphism between V and V^* for finite dimensional V

Proposition

Let V be a finite dimensional vector space, basis $\mathcal{B} = (v_1, \dots, v_n)$

Then V^* has a basis given by linear functionals

$$v_j^* : V \longrightarrow \mathbb{F}$$

$$v_j^*(v_k) = \delta_{jk} \quad \text{for } j = 1, \dots, n$$

Hence $\dim(V) = \dim(V^*)$ and $V \cong V^*$ (isomorphic)

An isomorphism is given by the linear map

$$L : V \longrightarrow V^*$$

$$L(v_j) = v_j^* \quad \text{for } j = 1, \dots, n$$

proof:

To show v_j^* form a basis, we need to show they span V^* and they are linearly independent

So need to show that every linear functional $f : V \longrightarrow \mathbb{F}$ is a linear combination of $v_1^*, \dots, v_n^*$

Every linear map f is competely determined by action on $v_1, \dots, v_n$

Define

$$\gamma_j = f(v_j) \quad \text{where } v_j \in V \text{ for } j = 1, \dots, n$$

Claim: $f = \sum_{j=1}^{n} \gamma_j v_j^*$

This is because

$$\left(\sum_{j=1}^{n} \gamma_j v_j^*\right)(v_k) = \sum_{j=1}^{n} \gamma_j (v_j^*(v_k)) = \gamma_k = f(v_k)$$

Same action on all basis vectors hence have same map $\Longrightarrow$ spans V^*

Have seen that each $F : V \longrightarrow \mathbb{F}$ in V^* is a linear combination of

$$v_1^*, \dots, v_n^*$$

$\Longrightarrow$ $v_1^*, \dots, v_n^*$ span V linearly independent: Assume $\exists\, \alpha_1, \dots, \alpha_n \in \mathbb{F}$ s.t

$$\sum_{j=1}^{n} \alpha_j v_j^* = \underline{0}$$ 0 map $\underline{0}: V \to \mathbb{F}$
$\underline{0}(v) = 0 \quad \forall v \in V$

Then $$\left(\sum_{j=1}^{n} \alpha_j v_j^*\right)(v_k) = \underline{0}(v_k) = 0 \implies \sum_{j=1}^{n} \alpha_j (v_j^*(v_k)) = \underline{0}$$

$$\implies \sum_{j=1}^{n} \alpha_j \delta_{jk} = \underline{0}$$

$$\implies \alpha_k = 0 \qquad \forall k$$

Hence linearly independent and span $V \implies$ forms basis

$$\dim(V^*) = n = \dim(V)$$ number of elements in basis

V, V^* vector spaces over same field with same dimension $\implies$ isomorphic $V \cong V^*$ corollary pg 21

If $\mathcal{B} = (v_1, \dots, v_n)$ and $\mathcal{B}^* = (v_1^*, \dots, v_n^*)$, we have 2 isomorphisms (co-ordinate maps)

$$\psi_{\mathcal{B}} : V \to \mathbb{F}^n ; \; v_j \mapsto \underline{e}_j$$

$$\psi_{\mathcal{B}^*} : V^* \to \mathbb{F}^n ; \; v_j^* \mapsto \underline{e}_j$$

So isomorphism $L: V \to V^*$ given by

$$L: V \to V^*$$

$$L: V \xrightarrow{\psi_{\mathcal{B}}} \mathbb{F}^n \xrightarrow{(\psi_{\mathcal{B}}^*)^{-1}} V^*$$

$$v_j \longmapsto \underline{e}_j \longmapsto v_j^*$$

■

Dual of a Linear map

Definition Dual Map

If $L: V \to W$ is a linear map, then the dual map of L is the linear map

$$L^*: W^* \to V^*$$

$$L^*(f) = f \circ L \quad \text{for } f \in W^*$$

In the example $V = \mathbb{R}^n$ viewed as column vectors with V^* being interpreted as row vectors and $W = \mathbb{R}_m$

The linear map $L: V \to W$ can be represented by an $m \times n$ matrix A

The linear map $L: W^* \to V^*$ is represented by the $n \times m$ matrix transpose A

2. Inner Products

Dot Product in $\mathbb{R}^n$

In $\mathbb{R}^n$, we can use dot (or scalar) product. If $\underline{u} = \sum_{j=1}^{n} \alpha_j \underline{e}_j$ and $\underline{v} = \sum_{j=1}^{n} \beta_j \underline{e}_j$

$$\underline{u} \cdot \underline{v} = \underline{u}^T \underline{v} = \sum_{j=1}^{n} \alpha_j \beta_j$$

The length of $\underline{u}$ is given by $\|\underline{u}\| = \sqrt{\underline{u} \cdot \underline{u}}$ and angle between $\underline{u}$ and $\underline{v}$ given by

$$\cos\theta = \frac{\underline{u} \cdot \underline{v}}{\|\underline{u}\|\|\underline{v}\|}$$

Properties of $\mathbb{R}^n$

1) $\underline{u} \cdot \underline{v} = \underline{v} \cdot \underline{u}$
2) $(\alpha \underline{u} + \beta \underline{v}) \cdot \underline{w} = \alpha(\underline{u} \cdot \underline{w}) + \beta(\underline{v} \cdot \underline{w})$
3) $\|\underline{u}\| \geq 0$
4) $\|\underline{u}\| \iff \underline{u} = 0$

Hermitian/Complex inner product on $\mathbb{C}^n$

In $\mathbb{C}^n$, usual dot product $\sqrt{u \cdot v}$ is not useful, since

$$\sqrt{u \cdot v} \in \mathbb{C} \notin \mathbb{R}$$

For example

$$u = \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}$$

$$\sqrt{u \cdot u} = \sqrt{i \cdot i + 0 + 0} = i$$

$$\sqrt{v \cdot v} = \sqrt{i \cdot i + 1 + 0} = 0 \quad \text{but } \underline{v} \neq 0$$ not like length

So we use complex conjugate $|z| = \sqrt{z\bar{z}}$ (non-negative)

Define Hermitian (or complex) inner product

$$\boxed{\langle u, v \rangle = \bar{u}^T v = \sum_{j=1}^{n} \bar{\alpha}_j \beta_j}$$

REAL INNER PRODUCT SPACES

Inner Product

Definition Inner Product

Let V be a vector space. An inner product on V is a function

$$\langle \, , \rangle : V \times V \longrightarrow \mathbb{R}$$

such that $\forall u,v,w \in V$ and $\forall \alpha \in \mathbb{R}$

i) $\langle u,v\rangle = \langle v,u\rangle$ — symmetry

ii) $\langle u+v, w\rangle = \langle u,w\rangle + \langle v,w\rangle$ — linear in first variable

iii) $\langle \alpha u, v\rangle = \alpha\langle u,v\rangle$

iv) $\langle u,u\rangle \geq 0$ — positive definite

v) $\langle u,u\rangle = 0 \iff u=0$

Vector space over $\mathbb{R}$ an inner product is also called a real inner product space (also called Euclidean space)

Examples of Inner Product Spaces

1) $\forall n \in \mathbb{N}$, $\mathbb{R}^n$ with usual dot product is an inner product

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = (x_1 \; x_2 \cdots x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Remark: Note that (ii) and (iii) imply that in any inner product space, for $u,v,w \in V$, $\alpha,\beta \in \mathbb{R}$

$$\langle \alpha u + \beta v, w\rangle = \alpha\langle u,w\rangle + \beta\langle v,w\rangle$$

Combined with symmetry (i), we get

$$\langle u, \alpha v + \beta w\rangle = \langle \alpha v + \beta w, u\rangle = \alpha\langle v,u\rangle + \beta\langle w,u\rangle$$

$$= \alpha\langle u,v\rangle + \beta\langle u,w\rangle$$

$\Longrightarrow \langle , \rangle$ linear in both variables, i.e. bilinear

2) $V = C([0,1], \mathbb{R})$: Space of continuous real valued functions on $[0,1]$

Define inner product

$$\langle f, g \rangle = \int_0^1 f(t)\, g(t)\, dt$$

We know continuous functions are integrable and product of continuous functions is continuous.

$\Longrightarrow$ outputs a real number $\forall f, g \in V$

$\Longrightarrow \langle , \rangle$ a function from $V \times V \longrightarrow \mathbb{R}$

Let $f, g, h \in V$ and $\alpha \in \mathbb{R}$

Symmetry:

$$\langle f, g \rangle = \int_0^1 f(t) g(t)\, dt = \int_0^1 g(t) f(t)\, dt = \langle g, f \rangle$$

Linearity: In first variable

$$\langle f+g, h \rangle = \int_0^1 (f+g)(t)\, h(t)\, dt$$

$$= \int_0^1 (f(t) + g(t))\, h(t)\, dt$$

$$= \int_0^1 (f(t)h(t) + g(t)h(t))\, dt$$

$$= \int_0^1 f(t)h(t) + \int_0^1 g(t)h(t)\, dt = \langle f, g \rangle + \langle g, h \rangle$$

$$\langle \alpha f, g \rangle = \int_0^1 (\alpha f)(t) = \int_0^1 \alpha f(t) g(t) = \alpha \int_0^1 f(t) g(t) = \alpha \langle f, g \rangle$$

Positive definite:

$$\langle f, f \rangle = \int_0^1 (f(t))^2 \geq 0 \quad \text{since } (f(t))^2 \geq 0 \quad \forall t \in [0,1]$$

If $\langle f, f \rangle = \int_0^1 (f(t))^2 = 0 \Longleftrightarrow f(t) = 0 \ \forall t \in [0,1]$ since $(f(t))^2 \geq 0$

Remark: proof did not rely on $[0,1]$, define inner product on $C([a,b],\mathbb{R})$ by

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt \qquad a<b$$

3) $V=\mathbb{R}^2$, product given by

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 3x_1y_1 + 2x_2y_2$$

<u>Symmetry</u>: $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 3x_1y_1 + 2x_2y_2$

$$= 3y_1x_1 + 2y_2x_2 = \left\langle \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle$$

<u>Linear in first variable</u>: $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle = 3(x_1+y_1)z_1 + 2(x_2+y_2)z_2$

$$= 3x_1z_1 + 2x_2z_2 + 3y_1z_1 + 2y_2z_2$$

$$= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle$$

$$\left\langle \alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 3(\alpha x_1)y_1 + 2(\alpha x_2)y_2$$

$$= \alpha 3x_1y_1 + \alpha 2x_2y_2$$

$$= \alpha(3x_1y_1 + 2x_2y_2)$$

<u>Positive definite</u>: $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle = 3x_1^2 + 2x_2^2 \geq 0$ as $x^2 \geq 0 \ \forall x \in \mathbb{R}$

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle = 0 \iff \underbrace{3x_1^2}_{\geq 0} + \underbrace{2x_2^2}_{\geq 0} = 0 \iff x_1 = 0 \,;\ x_2 = 0$$

Norm of a vector

Definition Norm

In an inner product space V, the norm (or length) of a vector v is

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

where non-negative square root is taken

Remark: The norm is a map $\|\cdot\| : V \longrightarrow \mathbb{R}$ but is NOT a linear map

For $\alpha \in \mathbb{R}$, $v \in V$, $\|\alpha v\| = \sqrt{\langle \alpha v, \alpha v \rangle} = \sqrt{\alpha^2 \langle v, v \rangle}$

$= |\alpha| \, \|v\|$, not $\alpha \|v\|$

Unit Vectors

Definition Unit Vector

If a vector $v \in V$ has norm 1

$$\|v\| = 1$$

then v is called a unit vector

If $v \neq 0$ is any non-zero vector, vector

$$\frac{v}{\|v\|}$$

has norm 1 $\Longrightarrow$ $\frac{v}{\|v\|}$ is a unit vector

Examples of Norms

1) In $C([0,1], \mathbb{R})$ with inner product

$$\langle f, g \rangle = \int_0^1 f(t) g(t) \, dt$$

vector $f(x) = x$ has norm $\|f\| = \sqrt{\langle f, f \rangle}$

$$\langle f, f \rangle = \int_0^1 f(t) f(t) \, dt = \int_1^2 t^2 \, dt = \left[\frac{t^3}{3} \right]_{t=0}^{t=1} = \frac{1}{3}$$

$$\Longrightarrow \|f\| = \frac{1}{\sqrt{3}}$$

Hence $\frac{f(x)}{\|f(x)\|} = \sqrt{3}x$ is a unit vector

2) $V = \mathbb{R}^2$ with inner product

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 3x_1y_1 + 2x_2y_2$$

Vector $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ has norm

$$\|v\| = \sqrt{3 \cdot 0 \cdot 0 + 2 \cdot 1 \cdot 1} = \sqrt{2}$$

$\frac{v}{\|v\|} = \frac{v}{\sqrt{2}} = \begin{pmatrix} 0 \\ 1/\sqrt{2} \end{pmatrix}$ a unit vector

BILINEAR FORMS

Definition Bilinear Forms

Let V, W, U be 3 vector spaces over same field $\mathbb{F}$. A map

$$f : V \times W \longrightarrow U$$

is said to be a bilinear map if it is linear in each of its arguments.

In the special case where $V = V$, $U = \mathbb{F}$, a bilinear map

$$f : V \times V \longrightarrow \mathbb{F}$$

is called a bilinear form.

In detail, a bilinear map $f : V \times V \longrightarrow \mathbb{F}$ that satisfies $\forall u, v, w \in V$, $\alpha \in \mathbb{F}$

i) $\langle u+w, v \rangle = \langle u, w \rangle + \langle v, w \rangle$

ii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$

iii) $\langle u, v+w \rangle = \langle u, v \rangle + \langle u, w \rangle$

iv) $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$

Examples of Bilinear Forms

1) $V = \mathbb{R}^2$, a bilinear form that is not an inner product is given by

$$f : V \times V \longrightarrow \mathbb{R}^2$$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 3x_1y_1 + x_1y_2 + 2x_2y_2$$

map bilinear, but not symmetric $\Longrightarrow$ not an inner product

2) $V = \mathbb{R}^2$

$$g : V \times V \longrightarrow \mathbb{R}$$

$$g\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 3x_1y_1 + 2x_2y_2$$

This is bilinear, symmetric, not positive definite

Matrix representing a bilinear form on $\mathbb{R}^n$

Proposition

A map $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ is a bilinear form

$$\Updownarrow$$

$\exists\, A \in M_{n\times n}(\mathbb{R})$ such that

$$f(u, v) = u^T A v \qquad \forall\, u, v \in \mathbb{R}^n$$

The entries A_{jk} of the matrix are given by

$$A_{jk} = f(\underline{e}_j, \underline{e}_k)$$

The matrix A is known as the matrix representing the bilinear form of f

Examples of matrix representing bilinear form

1) $V = \mathbb{R}^2$

Using $A_{jk} = f(\underline{e}_j, \underline{e}_k)$, the dot product represented by I_2 since

$$A_{11} = f(\underline{e}_1, \underline{e}_2) = \underline{e}_1 \cdot \underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1^2 + 0^2 = 1$$

$$A_{21} = f(\underline{e}_2, \underline{e}_1) = \underline{e}_2 \cdot \underline{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

2) $V = \mathbb{R}^2$

Using bilinear form $f : V \times V \longrightarrow \mathbb{R}$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 3x_1y_1 + x_1y_2 + 2x_2y_2$$

$$A_{11} = f(\underline{e}_1, \underline{e}_2) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 3\cdot1\cdot1 + 1\cdot0 + 0\cdot0 = 3$$

$$A_{12} = f(\underline{e}_1, \underline{e}_2) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 3\cdot1\cdot0 + 1\cdot1 + 1\cdot0 = 1$$

$$A_{21} = f(\underline{e}_2, \underline{e}_1) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 3\cdot1\cdot0 + 0\cdot0 + 2\cdot1\cdot0 = 0$$

$$A_{22} = f(\underline{e}_2, \underline{e}_2) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 3\cdot0\cdot0 + 1\cdot0 + 2\,1\cdot1 = 2$$

Matrix representing A is

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$$ co-efficients

2) Similarly for the bilinear form

$g : V \times V \longrightarrow \mathbb{R}$

$$g\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = -3x_1y_1 + 2x_2y_2$$

$$B = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}$$ co-efficients

Symmetric positive definite matrices

Definition Symmetric matrices

For any $n \times n$ matrix $A \in Mat(n, \mathbb{F})$, A is symmetric if

$$A^T = A \quad \text{or} \quad A_{ij} = A_{ji}$$

Definition Positive Definite

A real symmetric $n \times n$ matrix is said to be positive definite if

$$v^T A v \geq 0 \quad \forall \text{ column vectors } v \in \mathbb{R}^n$$

$$v^T A v = 0 \iff v = 0$$

Leading Principle Major

Definition Leading Principal Minor

For any $n \times n$ matrix, a leading principal minor of A is the determinant of the submatrix formed by taking the top left $k \times k$ submatrix of A for any $1 \leq k \leq n$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Lemma

Let A be a real symmetric $n \times n$ matrix. Then the following are equivalent

i) A is positive definite

ii) All eigenvalues are positive

iii) All the leading principal minors are positive (Sylvester's Critereon)

Proposition

A bilinear form $\langle , \rangle$ on $\mathbb{R}^n \times \mathbb{R}^n$ is a real inner product

$$\Updownarrow$$

matrix representing $\langle , \rangle$ is a real symmetric positive definite matrix

Examples

1) $V = \mathbb{R}^2$

The matrix representing dot product

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

I_2 is real, symmetric, positive eigenvalues $\Longrightarrow$ positive definite

2) $V = \mathbb{R}^2$

Using bilinear form $f : V \times V \longrightarrow \mathbb{R}$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = 3x_1y_1 + x_1y_2 + 2x_2y_2$$

Represented by matrix $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$

Since

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3y_1 + y_2 \\ 2y_2 \end{pmatrix}$$

$$= x_1(3y_1 + y_2) + x_2 2y_2$$

$$= 3x_1y_1 + x_1y_2 + 2x_2y_2$$

Note: Can instantly see matrix from co-efficients

$$A_{ij} = \text{coefficient of } x_iy_j$$

A not symmetric $\Longrightarrow$ form not symmetric

$\Longrightarrow$ not an inner product

Also saw $B = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}$ non-positive eigenvalues

$\Longrightarrow$ not positive definite

3) $V = \mathbb{R}^3$

$h : V \times V \longrightarrow \mathbb{R}^3$ given by

$$h\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) = 2x_1y_1 + x_2y_2 - 2x_2y_3 - 2x_3y_2 + Kx_3y_3 \quad K \in \mathbb{R}$$

This is a bilinear form on $\mathbb{R}$ since we can represent by matrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & K \end{pmatrix} \quad \text{real, symmetric}$$

<u>Positive definite</u>: using (iii) Sylvesters Critereon

Calculating determinants of

$$\begin{array}{c} \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & K \end{pmatrix} \end{array}$$

(1) $|2| = 2 > 0$

(2) $\begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} = 2 > 0$

(3) $\begin{vmatrix} 2 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & k \end{vmatrix} = 2\begin{vmatrix} 1 & -2 \\ -2 & k \end{vmatrix} = 2(k-4)$

$$2(k-4) \geq 0 \iff k-4 \geq 0 \iff k \geq 4$$

Matrix positive definite $\iff K \geq 4$

Therefore h is an inner product iff $K \geq 4$

Matrix Form of a bilinear map on real vector space V

Generalise result to any finite dimensional vector space over $\mathbb{R}$

Theorem

Let V be a finite dimensional vector space over $\mathbb{R}$, let $\mathcal{B} = \{v_1, \dots, v_n\}$ be any basis for V.

For $u, v \in V$, let $\underline{u} = \psi_{\mathcal{B}}(u)$ and $\underline{v} = \psi_{\mathcal{B}}(v)$ (so $\underline{u}$ and $\underline{v}$ are co-ordinate column vectors of u and v with respect to $\mathcal{B}$

A map $\langle , \rangle : V \times V \to \mathbb{R}$ is a bilinear form $\iff \exists$ a matrix $A \in M_{n \times n}(\mathbb{R})$ such that

$$\langle u, v \rangle = \underline{u}^T A \underline{v} \qquad \forall u, v \in V$$

The entries of A_{jk} of the matrix is given by

$$A_{jk} = \langle v_j, v_k \rangle$$

A is called matrix representing the bilinear form $\langle , \rangle$ w.r.t basis $\mathcal{B}$

The bilinear form $\langle , \rangle$ is a real inner product on $V \iff A$ is real, symmetric positive definite matrix

Example

We saw $\langle f, g \rangle = \int_0^1 f(t) g(t) \, dt$ is an inner product on infinite dimension vector space $C([0,1], \mathbb{R})$

Let $V = \mathbb{R}_1[x]$. Then $V \subseteq C([0,1], \mathbb{R})$

$\implies$ So this is an inner product on V as well.

Matrix w.r.t $\mathcal{B} = (\underset{v_1}{1}, \underset{v_2}{x})$ standard basis:

$$A_{11} = \langle v_1, v_2 \rangle = \langle 1, 1 \rangle = \int_0^1 1^2 \, dt = [x]_0^1 = 1$$

$$A_{12} = \langle v_1, v_2 \rangle = \langle 1, x \rangle = \int_0^1 1 \cdot t \, dt = \frac{1}{2}$$

$A_{12} = A_{21}$ is inner product is symmetric

$$A_{22} = \langle v_2, v_2 \rangle = \langle x, x \rangle = \int_0^1 t \cdot t \, dt = \frac{1}{3}$$

We get matrix

$$A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}$$

Check this is positive definite using (iii) using Sylvester's Critereon

$$\begin{array}{cc} 1 & 2 \\ \multicolumn{2}{c}{\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}} \end{array}$$

(1) $|1| = 1$

(2) $\begin{vmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{vmatrix} = \frac{1}{12} > 0$

Check matrix represents $\langle , \rangle$:

A polynomial in $\mathbb{R}_1[x]$ has form

$$\alpha_0 + \alpha_1 x \quad , \alpha_0, \alpha_1 \in \mathbb{R}$$

$$\varphi_B(\alpha_0 + \alpha_1 x) = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}$$

$$\langle f, f \rangle = \langle \alpha_0 + \alpha_1 x, \alpha_0 + \alpha_1 x \rangle = \int_0^1 (\alpha_0 + \alpha_1 t)^2 \, dt$$

$$= \int_0^1 (\alpha_0^2 + 2\alpha_0\alpha_1 t + \alpha_1^2 t^2) \, dt$$

$$= \left[\alpha_0^2 t + \alpha_0 \alpha_1 t + \frac{\alpha_1^2 t^3}{3} \right]_0^1$$

$$= \alpha_0^2 + \alpha_0\alpha_1 + \frac{1}{3}\alpha_1^2$$

$$(\alpha_0 \ \alpha_1) \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = (\alpha_0 \ \alpha_1) \begin{pmatrix} \alpha_0 + 1/2\,\alpha_1 \\ 1/2\,\alpha_0 + 1/3\,\alpha_1 \end{pmatrix}$$

$$= \alpha_0^2 + \alpha_0\alpha_1 + \frac{1}{3}\alpha_1$$

Matrices Representing the same bilinear form

How are matrices representing the same bilinear form with respect to different bases related?

Proposition

Let V be a finite dimensional vector space over $\mathbb{R}$ and let $f: V \times V \to \mathbb{R}$ be a bilinear form.

Let $\mathcal{A}$ and $\mathcal{B}$ be 2 bases for V. Let B be the matrix representing f w.r.t $\mathcal{B}$ and A be the matrix representing f w.r.t $\mathcal{A}$.

Then $\exists$ an invertible matrix P such that

$$B = P^T A P$$

In fact, we have $P = C_{\mathcal{B}}^{\mathcal{A}}$, the change of basis from $\mathcal{B}$ to $\mathcal{A}$

Proof:

If $\underline{u}_{\mathcal{B}} = \psi_{\mathcal{B}}(u)$ $\quad \underline{v}_{\mathcal{B}} = \psi_{\mathcal{B}}(v)$

$\underline{u}_{\mathcal{A}} = \psi_{\mathcal{A}}(u)$ $\quad \underline{v}_{\mathcal{A}} = \psi_{\mathcal{A}}(v)$

Then

$$\underline{u}_{\mathcal{A}} = \psi_{\mathcal{A}}(u) = (\underbrace{\psi_{\mathcal{A}} \circ \psi_{\mathcal{B}}^{-1}}_{C_{\mathcal{B}}^{\mathcal{A}}} \circ \psi_{\mathcal{B}})(u)$$

$$= C_{\mathcal{B}}^{\mathcal{A}} \psi_{\mathcal{B}}(u)$$

$$= C_{\mathcal{B}}^{\mathcal{A}} \underline{u}_{\mathcal{B}}$$

Hence

$$\langle u, v \rangle = \underline{u}_{\mathcal{A}}^T A \underline{v}_{\mathcal{A}} = (C_{\mathcal{B}}^{\mathcal{A}} \underline{u}_{\mathcal{B}})^T A (C_{\mathcal{B}}^{\mathcal{A}} \underline{v}_{\mathcal{B}})$$

$$= \underline{u}_{\mathcal{B}}^T \underbrace{(C_{\mathcal{B}}^{\mathcal{A}})^T A\, C_{\mathcal{B}}^{\mathcal{A}}}_{B} \underline{v}_{\mathcal{B}}$$

$\blacksquare$

Definition Congruent Matrices

Matrices A and B that satisfy the condition

$$B = P^T A P$$

for some invertible matrix P are called congruent matrices

Important!!!

Show that if A, B congruent then

i) A symmetric $\iff$ B symmetric

ii) A positive definite $\iff$ B positive definite

Congruence is a equivalence relation on matrices

COMPLEX INNER PRODUCT SPACES

Definition Complex inner product spaces

Let V be a vector space over $\mathbb{C}$. A Hermitian inner product on V is a function

$$\langle , \rangle : V \times V \longrightarrow \mathbb{C}$$

such that $\forall u, v, w \in V$ and $\alpha \in \mathbb{C}$,

i) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ Hermitian (conjugate) symmetry

ii) $\langle u, v+w \rangle = \langle u, v \rangle + \langle u, w \rangle$

iii) $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle \quad \forall \alpha \in \mathbb{C}$ linearity in second argument

iv) $\langle u, u \rangle \geq 0$ (in particular $\langle u, u \rangle \in \mathbb{R}_{\geq 0}$)   positive definiteness

v) $\langle u, u \rangle = 0 \iff u = 0$

Recall:

1) Complex conjugate

For any $z \in \mathbb{C}$, $z = x + iy$

$\bar{z} = x - iy$ - complex conjugate

2) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$

3) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

Recall: For $\langle , \rangle : V \times V \to \mathbb{C}$, a Hermitian inner product

$u, v, w \in V$, $\alpha \in \mathbb{C}$

1)
$$\begin{aligned}
\langle u+v, w \rangle &= \overline{\langle w, u+v \rangle} && \text{by (i)} \\
&= \overline{\langle w, u \rangle + \langle w + u \rangle} && \text{by (ii)} \\
&= \overline{\langle w, u \rangle} + \overline{\langle w, v \rangle} && \text{since } \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \\
&= \langle u, w \rangle + \langle v, w \rangle
\end{aligned}$$

2)
$$\begin{aligned}
\langle \alpha u, v \rangle &= \overline{\langle v, \alpha u \rangle} && \text{by (i)} \\
&= \overline{\alpha \langle v, u \rangle} && \text{by (ii)} \\
&= \bar{\alpha}\, \overline{\langle v, u \rangle} && \text{since } \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2
\end{aligned}$$

$= \bar{\alpha} \langle u, v \rangle$ by (i)

Norm in a complex inner product

Definition Norm

Let V be a complex vector space

The norm (or length) is a function

$$\|\cdot\| : V \longrightarrow \mathbb{R}_{>0}$$

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

Vectors of norm 1 are called unit vectors

Note: Norm is NOT linear

$$\|\alpha v\| = |\alpha| \|v\|$$

Hermitian inner product using matrices

Definition Conjugate Transpose

For any $p \times n$ matrix $A = (A_{jk})$, we define its conjugate transpose to be

$$A^{\dagger} = (\bar{A})^{T}$$

i.e.

$$A^{\dagger}_{kg} = (\bar{A}_{kj})$$

Definition Hermitian

We say a square matrix A is Hermitian when

$$A^{\dagger} = A$$

and positive definite when $u^{\dagger} A u > 0$ for all u

Theorem

Let V be a complex finite dimensional vector space, let $\mathcal{B} = (v_1, \cdots, v_n)$ be a basis for V.

An operation $\langle , \rangle$ on $V \times V$ is an Hermitian inner product

$$\Updownarrow$$

$\exists$ a Hermitian positive definite matrix $A \in M_{n\times n}(\mathbb{C})$ for which

$$\langle u, v \rangle = \underline{u}^{\dagger} A \underline{v}$$

$\forall u, v \in V$, where $\underline{u} = \psi_{\mathcal{B}}(u)$ and $\underline{v} = \psi_{\mathcal{B}}(v)$ are the corresponding co-ordinate (column) vectors in $\mathbb{C}^n$

The A_{jk} of entries A are given by

$$A_{jk} = \langle v_j, v_k \rangle$$

Note: A is real $\Longleftrightarrow$ $A^{\dagger} = A^{T}$

Symmetric matrices are Hermitian, not vice verca

Examples of inner products

1) $V = \mathbb{C}$, $\langle z, w \rangle \quad \bar{z} w$

Note that $\langle z, z \rangle = \bar{z} z = |z|^2 \in \mathbb{R}_{\geq 0}$

2) **Standard Inner Product on $\mathbb{C}^n$:**

$V = \mathbb{C}^n$:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n = \sum_{i=1}^{n} \bar{x}_i y_i$$

Checking axioms:

(i)
$$\left\langle \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle = \bar{y}_1 x_1 + \cdots + \bar{y}_n x_n$$
$$= x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n$$
$$= \overline{\bar{x}_1 y_1} + \cdots + \overline{\bar{x}_n y_n}$$
$$= \overline{\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle}$$

(ii) (iii) same as $\mathbb{R}^n$

(iv) $$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right\rangle = \bar{x}_1 x_1 + \cdots + \bar{x}_n x_n$$
$$= |x_1|^2 + \cdots + |x_n|^2 \in \mathbb{R}_{\geq 0}$$

and $|x_1|^2 + \cdots + |x_n|^2 = 0 \iff x_1 = 0, \cdots, x_n = 0$

Matrix representing $\langle , \rangle = I_n$

$$\langle u, v \rangle = \underline{u}^{\dagger} I_n \underline{v} = \underline{u}^{\dagger} \underline{v}$$

3) The space of continuous functions $\mathcal{C}([0,1], \mathbb{C})$

$$f : [0,1] \longrightarrow \mathbb{C} \qquad ([0,1] \subseteq \mathbb{R})$$

with inner product given by $\langle f, g \rangle = \int_0^1 \overline{f(t)}\, g(t)\, dt$

4) $V = \mathbb{C}^2$ with inner product

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 3\bar{x}_1 y_1 + 2\bar{x}_2 y_2$$

Matrix representing $\langle , \rangle$: by looking coefficients

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

Theorem

All eigenvalues of a Hermitian matrix are real.

Proof:

$$A\vec{v} = \lambda \vec{v} \implies (A\vec{v})^{\dagger} = (\lambda \vec{v})^{\dagger}$$
$$\implies \vec{v}^{\dagger} A^{\dagger} = \bar{\lambda} \vec{v}^{\dagger}$$
$$\implies \vec{v}^{\dagger} A^{\dagger} \vec{v} = \bar{\lambda} \vec{v}^{\dagger} \vec{v} \qquad \text{multiply both sides by } \vec{v}$$
$$\implies \vec{v}^{\dagger} A \vec{v} = \bar{\lambda} \vec{v}^{\dagger} \vec{v} \qquad A^{\dagger} = A$$
$$\implies \vec{v}^{\dagger} \lambda \vec{v} = \bar{\lambda} \vec{v}^{\dagger} \vec{v}$$
$$\implies \lambda \vec{v}^{\dagger} \vec{v} = \bar{\lambda} \vec{v}^{\dagger} v$$

$\vec{v}^{\dagger} \vec{v} \neq \underline{0}$ since $\vec{v} \neq \underline{0} \implies \lambda = \bar{\lambda} \implies \lambda \in \mathbb{R}$ ∎

5) Consider

$$A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$$

$$A^{\dagger} = \bar{A}^{T} = \begin{pmatrix} 2 & \bar{i} \\ \bar{-i} & 2 \end{pmatrix}^{T} = \begin{pmatrix} 2 & -i \\ i & 2 \end{pmatrix}^{T} = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} = A$$

$\Rightarrow$ A Hermition

In fact lemma on pg 49 still applies in Complex case

Using (iii) Sylvester's Critereon

$$\begin{array}{cc} 1 & 2 \\ \multicolumn{2}{c}{\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}} \end{array}$$

(1) $|2| = 2 > 0$

(2) $\begin{vmatrix} 2 & i \\ -i & 2 \end{vmatrix} = 2^2 + i^2 = 1 > 0$

So matrix positive definite

$$\Longrightarrow \langle u, v \rangle = \underline{u}^T A \underline{v} \quad \text{defines inner product}$$

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = (\bar{x}_1, \bar{x}_2) \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$= 2\bar{x}_1 y_1 + i\bar{x}_1 y_2 - i\bar{x}_2 y_1 + 2\bar{x}_2 y_2$$

CAUCHY-SCHWARTZ INEQUALITIES AND METRIC SPACES

Let V be any inner product space (real or complex)

Theorem Cauchy-Schwartz inequality

If u and v are any vectors in an inner product space V, then

$$|\langle u,v\rangle| \leq \|u\|\,\|v\|$$

Proof:

Let V be any real inner product space.

For any $w\in V$, $\langle w,w\rangle \geq 0$

Let $w=\alpha u+\beta v$ where $\alpha=-\langle u,v\rangle$, $\beta=\langle u,u\rangle$

Then $\beta\geq 0$ and

$$\begin{aligned}\langle \alpha u+\beta v, \alpha u+\beta v\rangle &= \langle \alpha u,\alpha u\rangle + \langle \alpha u,\beta v\rangle + \langle \beta v,\alpha u\rangle + \langle \beta v,\beta v\rangle \\ &= \alpha^2\langle u,u\rangle + \alpha\beta\langle u,v\rangle + \beta\alpha\langle v,u\rangle + \beta^2\langle v,v\rangle \\ &= \alpha^2\beta - \alpha^2\beta - \alpha^2\beta + \beta^2\langle v,v\rangle \\ &= \beta\left(\langle u,u\rangle\langle v,v\rangle - \langle u,v\rangle^2\right)\end{aligned}$$

If $\beta=0 \iff u=0$, then Cauchy-Schwartz inequality trivial.

Otherwise $\beta>0 \implies \beta(\langle u,u\rangle\langle v,v\rangle - \langle u,v\rangle^2)\geq 0$

$$\implies \langle u,u\rangle\langle v,v\rangle \geq \langle u,v\rangle^2$$

$$\implies \langle u,v\rangle \leq \sqrt{\langle u,u\rangle}\sqrt{\langle v,v\rangle} \qquad \text{(non-negative square roots)}$$

$$\implies \langle u,v\rangle \leq \|u\|\,\|v\|$$

Let V be a complex inner product space. For any $w\in V$ and

$$\langle w,w\rangle \geq 0$$

Let $w=\alpha u+\beta v$ where $\alpha=-\langle u,v\rangle$ and $\beta=\langle u,u\rangle$, $\beta\in\mathbb{R}$, $\beta\geq 0$. Then

$$\begin{aligned}\langle \alpha u+\beta v, \alpha u+\beta v\rangle &= \alpha\bar{\alpha}\langle u,u\rangle + \bar{\alpha}\beta\langle u,v\rangle + \bar{\beta}\alpha\langle v,u\rangle + \bar{\beta}\beta\langle v,v\rangle \\ &= \alpha\bar{\alpha}\langle u,u\rangle + \bar{\alpha}\beta\langle u,v\rangle + \beta\alpha\langle v,u\rangle + \beta^2\langle v,v\rangle \\ &= \alpha\bar{\alpha}\langle u,u\rangle + \bar{\alpha}\beta\langle u,v\rangle + \beta\alpha\langle v,u\rangle + \beta^2\langle v,v\rangle \\ &= \bar{\alpha}\alpha\beta + \bar{\alpha}\beta(-\alpha) + \beta\alpha(-\bar{\alpha}) + \beta^2\langle v,v\rangle\end{aligned}$$

$$= -\beta\alpha\bar{\alpha} + \beta^2\langle v,v\rangle$$
$$= \beta(-|\alpha|^2 + \beta\langle v,v\rangle)$$
$$= \beta(\langle u,u\rangle\langle v,v\rangle - |\langle u,v\rangle|^2)$$

If $\beta = \langle u,u\rangle = 0 \implies u = 0$, Cauchy-Schwartz inequality is trivially true, both sides 0.

Otherwise $\beta > 0$ and $\langle \alpha u + \beta v, \alpha u + \beta v\rangle \geq 0$, so we have

$$\langle u,u\rangle\langle v,v\rangle - |\langle u,v\rangle|^2 \geq 0. \implies |\langle u,v\rangle|^2 \leq \langle u,u\rangle\langle v,v\rangle$$
$$\implies |\langle u,v\rangle| \leq \|u\|\|v\| \quad \text{(non-negative square roots)}$$

■

Triangle inequality

Theorem Triangle Inequality

If V is an inner product space, $u, v \in V$, then

$$\|u+v\| \leq \|u\| + \|v\|$$

Proof:

i) Over $\mathbb{R}$: By definition

$$\|u+v\|^2 = \langle u+v, u+v\rangle$$
$$= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle$$
$$= \|u\|^2 + 2\langle u,v\rangle + \|v\|^2$$
$$\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \quad \text{Cauchy-Schwartz inequality}$$
$$= (\|u\| + \|v\|)^2$$

Taking non-negative square root

$$\|u+v\| \leq \|u\| + \|v\|$$

ii) Over $\mathbb{C}$

$$\|u+v\|^2 = \langle u+v, u+v\rangle$$
$$= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle$$
$$= \|u\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + \|v\|^2$$
$$= \|u\|^2 + 2\mathrm{Re}(\langle u,v\rangle) + \|v\|^2 \quad \text{since } z + \bar{z} = \mathrm{Re}(z) \quad \forall z \in \mathbb{C}$$
$$\leq \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \quad \text{since } \mathrm{Re}(z) \leq |z| \quad \forall z \in \mathbb{C}$$

$$\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \quad \text{by Cauchy Schwartz inequality}$$

$$= (\|u\| + \|v\|)^2$$

$$\Longrightarrow \|u+v\|^2 \leq (\|u\| + \|v\|)^2$$

$$\Longrightarrow \|u+v\| \leq \|u\| + \|v\| \quad \text{non-negative square root}$$

∎

Cosine Angle

On real inner product space

$$-\|u\|\,\|v\| < \langle u, v\rangle \leq \|\underline{u}\|\,\|\underline{v}\|$$

$$\Longrightarrow \quad -1 \leq \frac{\langle u, v\rangle}{\|u\|\,\|v\|} \leq 1$$

Can use this to define the cosine of the angle between u and v

$$\boxed{\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}}$$

Metric on an inner product space

Definition

In any inner product space V, the corresponding **metric** or **distance function** is a function

$$d : V \times V \longrightarrow \mathbb{R}$$

$$d(u,v) = \|u-v\|$$

The metric on an inner product space V satisfies

$\forall u,v,w \in V$:

i) (positivity) $d(u,v) \geq 0$

ii) (symmetry) $d(u,v) = d(v,u)$

iii) (triangle inequality): $d(u,w) \leq d(u,v) + d(v,w)$

iv) $d(u,v) = 0 \iff u = v$

Remark: We defined norm with help of inner product. Then used norm to define a metric.

$$\{\text{metric spaces}\} \subseteq \{\text{normed spaces}\} \subseteq \{\text{inner product spaces}\}$$

vector space for which a norm is defined

Examples of normed spaces

In $\mathbb{R}^n$, with dot product

if $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $u = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\|v\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

$$d(u,v) = \sqrt{(y_1-x_1)^2 + \cdots + (y_n-x_n)^2}$$ standard Euclidean metric

ORTHOGONALITY

Let V be an inner product space

Definition Orthogonal

Two vectors u, v of an inner product space are said to be orthogonal iff

$$\langle u, v\rangle = 0$$

Orthogonal vectors denoted by $u \perp v$

Orthonormal Vectors

Definition

A set S of non-zero vectors in an inner product space is said to be orthogonal if $u \perp v$ for all distinct pair of vectors in S

$$\langle v_i, v_j\rangle = 0 \quad i \neq j \quad \forall\, v_i, v_j \in S$$

If all $u \in S$ is a unit vector, then S is said to be orthonormal.

Note if $\langle u, v\rangle = 0$, then $\langle v, u\rangle = 0$ since $\langle u, v\rangle = \langle v, u\rangle$ (over $\mathbb{R}$)

$$\langle v, u\rangle = \overline{\langle u, v\rangle} = \bar{0} = 0 \quad (\text{over } \mathbb{C})$$

For any $v \in V$, we have $\langle v, \underline{0}\rangle = \underline{0}$ since

$$\begin{aligned} \langle v, \underline{0}\rangle &= \langle v, w - w\rangle \quad \text{for } \forall w \in W \\ &= \langle v, w\rangle + \langle v, -w\rangle \\ &= \langle v, w\rangle - \langle v, w\rangle \\ &= 0 \end{aligned}$$

Similarly $\langle 0, \underline{v}\rangle = 0 \ \forall v \in V$

$\Longrightarrow$ $\underline{0}$ vector orthogonal to every vector

Theorem

Any orthogonal set in an inner product space V is linearly independent.

Hence V has dimension n and S has n elements, S is a basis on V

<u>Proof</u>:

Suppose $S \subseteq V$, subset of non-zero vectors $v \in S$ $v \neq \underline{0}$ in V such that

$$\langle u, v \rangle = 0 \qquad \forall u, v \in S, \; u \neq v$$

Let $v_1, \cdots, v_k$ be k distinct vectors in S

$$S = \{v_1, \cdots, v_k\}$$

Suppose $\exists \, \alpha_1, \cdots, \alpha_n \in \mathbb{F}$ s.t

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \underline{0}$$

Take inner product with v_i

$$0 = \langle v_i, \underline{0} \rangle = \langle v_i, \alpha_1 v_1 + \cdots + \alpha_k v_k \rangle$$

$$= \langle v_i, \alpha_1 v_1 \rangle + \cdots + \langle v_i, \alpha_k v_k \rangle$$

$$= \alpha_1 \langle v_i, v_1 \rangle + \cdots + \alpha_k \langle v_i, \alpha_k \rangle$$

$$= \alpha_i \langle v_i, v_i \rangle$$

Since $v_i \neq 0$, $\langle v_i, v_i \rangle \neq 0 \implies \alpha_i = 0 \quad \forall i$

$\implies v_i$ linearly independent.

<u>Note</u>: if $(v_1, \cdots, v_n)$ orthogonal basis for V,

Then any $v \in V$ can be rewritten as

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n \quad \text{for some } \alpha_1, \cdots, \alpha_n \in \mathbb{F}$$

Easy to find co-efficients α_i

$$\langle v_i, v \rangle = \langle v_i, v \rangle = \langle v_i, \alpha_1 v_1 + \cdots + \alpha_n v_n \rangle$$

$$= \alpha_1 \langle v_i, v_1 \rangle + \cdots + \alpha_n \langle v_i, v_n \rangle$$

$$= \alpha_i \langle v_i, v_i \rangle$$

$$\implies \boxed{\alpha_i = \frac{\langle v_i, v \rangle}{\|v_i\|^2}}$$

Examples

1) In $\mathbb{R}^n$ (with standard inner dot product)

The standard basis is orthonormal

2) In $\mathbb{R}^3$, with standard inner product

$$\begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}5\\4\\-1\end{pmatrix}, \begin{pmatrix}-1\\2\\3\end{pmatrix}$$

form an orthogonal basis but not orthormal since not normal

Can get orthormal basis by dividing each vector by norm

$$\left\| \begin{pmatrix}1\\-1\\1\end{pmatrix} \right\| = \sqrt{(1)^2+(-1)^2+1^2} = \sqrt{3} \implies \begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}$$

$$\left\| \begin{pmatrix}5\\4\\-1\end{pmatrix} \right\| = \sqrt{5^2+4^2+(-1)^2} = \sqrt{42} \implies \begin{pmatrix}5/\sqrt{42}\\4/\sqrt{42}\\-1/\sqrt{42}\end{pmatrix}$$

$$\left\| \begin{pmatrix}-1\\2\\3\end{pmatrix} \right\| = \sqrt{(-1)^2+(2)^2+(3)^2} = \sqrt{14} \implies \begin{pmatrix}-1/\sqrt{14}\\2/\sqrt{14}\\3/\sqrt{14}\end{pmatrix}$$

3) In $\mathbb{C}^2$ (with standard Hermitian inner product)

$$\begin{pmatrix}\frac{i}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}, \begin{pmatrix}1/\sqrt{2}\\ i/\sqrt{2}\end{pmatrix}$$

form orthonormal basis

$$\left\langle \begin{pmatrix}i/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}, \begin{pmatrix}1/\sqrt{2}\\i/\sqrt{2}\end{pmatrix} \right\rangle = \overline{\left(\frac{-i}{\sqrt{2}}\right)}\frac{1}{\sqrt{2}} + \left(\frac{i}{\sqrt{2}}\right)\frac{i}{\sqrt{2}} = 0$$

Orthogonal and Unitary Matrices

Definition Orthogonal/Unitary

We say a real matrix $Q \in M_{n\times n}(\mathbb{R})$ is orthogonal when

$$Q^T Q = I_n \quad (Q^{-1} = Q^T)$$

We say a complex matrices $P \in M_{n\times n}(\mathbb{C})$ is unitary when

$$P^\dagger P = I_n \quad (P^\dagger = P^{-1})$$

Lemma

i) A basis $v_1, \dots, v_n$ of $\mathbb{R}^n$ is orthonormal for the standard real inner product

$$\Updownarrow$$

these are the columns of an orthogonal matrix Q.

ii) A basis $v_1, \dots, v_n$ of $\mathbb{C}^n$ is orthonormal for the standard Hermitian inner product

$$\Updownarrow$$

these are the columns of a unitary matrix Q.

Proof:

1) On $\mathbb{R}^n$, we have $\underbrace{\langle v_j, v_k \rangle}_{\text{dot product}} = v_j^T I_n v_k = v_j^T v_k$

Basis orthonormal $\iff v_j^T v_k = \delta_{jk}$ (like matrix multiplication)

So $\iff$ the v_i are columns of a matrix Q with $Q^T Q = I_n$

2) Similar for $\mathbb{C}^n$

$\blacksquare$

Remark: Orthogonal matrices preserve the standard real inner product on $\mathbb{R}^n$

Let Q be an $n\times n$ orthogonal $n\times n$ matrix, $u, v \in \mathbb{R}^n$

$$\begin{aligned}\langle Qu, Qv \rangle &= (Qu)^T I_n (Qv) \\ &= (Qv)^T Qu \\ &= u^T \underbrace{Q^T Q}_{I_n} u \\ &= u^T v = \langle u, v \rangle\end{aligned}$$

Similarly, unitary matrices preserve standard Hermitian inner product on $\mathbb{C}^n$

if $P \in M_{n\times n}(\mathbb{C})$ is unitary, then

$$\langle Pu, Pv\rangle = \langle u, v\rangle \quad \forall\, u, v \in \mathbb{C}^n$$

Example

Writing last example in matrix form

$$Q = \begin{pmatrix} 1/\sqrt{3} & 5/\sqrt{42} & -1/\sqrt{14} \\ -1/\sqrt{3} & 4/\sqrt{42} & 2/\sqrt{14} \\ 1/\sqrt{3} & -1/\sqrt{42} & 3/\sqrt{14} \end{pmatrix} \quad \text{orthogonal matrix}$$

$$P = \begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{pmatrix} \quad \text{unitary matrix}$$

Projection of a vector

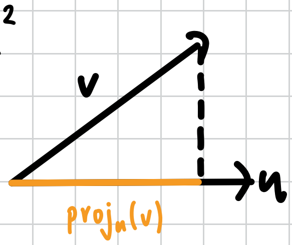
Definition

Let V be an inner product space, let $u \in V$ be a non-zero vector. The vector

$$\text{proj}_u(v) = \frac{\langle u, v\rangle}{\langle u, u\rangle} u$$

is called the projection of v in direction of u or projection of v onto $Sp(u)$

In $\mathbb{R}^2$



Note $v - \text{proj}_u(v) \perp u$

Lemma

Suppose that $S = \{w_1, \dots, w_k\}$ is an orthogonal set in an inner product space V and that v is any vector in V. Then the vector

$$w = v - \sum_{i=1}^{k} \text{proj}_{w_i}(v) = v - \sum_{i=1}^{k} \frac{\langle w_i, v\rangle}{\langle w_i, w_i\rangle} w_i$$

is orthogonal to each vector in S and consequently to each vector in the span of S in V

Proof

For each $j = 1, \dots, k$, we have

$$\langle w_j, w \rangle = \left\langle w_j, v - \sum_{i=1}^{k} \operatorname{proj}_{w_i}(v) \right\rangle$$

$$= \left\langle w_j, v - \sum_{i=1}^{k} \frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle} w_i \right\rangle$$

linearity $$= \langle w_j, v \rangle - \sum_{i=1}^{k} \left\langle w_j, \underbrace{\frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle}}_{\in \mathbb{F}} w_i \right\rangle$$

linearity $$= \langle w_j, v \rangle - \sum_{i=1}^{k} \frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle} \langle w_j, w_i \rangle$$

$$= \langle w_j, v \rangle - \frac{\langle w_j, v \rangle}{\cancel{\langle w_j, w_j \rangle}} \cancel{\langle w_j, w_j \rangle}$$ S orthogonal set

$$= 0$$

$\Longrightarrow w \perp w_j$ for each $w_j \in S$

If $u \in Sp(S)$ then $u \in \sum_{i=1}^{n} \alpha_i w_i$ for some $\alpha_i \in \mathbb{F}$

So $$\langle w, u \rangle = \left\langle w, \sum_{i=1}^{k} \alpha_i w_i \right\rangle$$

$$= \sum_{i=1}^{k} \alpha_i \langle w, w_i \rangle$$

$$= 0$$ $\Longrightarrow w \perp u$ for each $u \in Sp(S)$

GRAM-SCHMIDT PROCESS

Theorem

Any finite dimensional inner product space V has an orthonormal basis

Proof: (Algorithm, important)

Start with any ordered basis $(u_1, \cdots, u_n)$ for V. Define

For each $j = 1, \cdots, k$, we have

$$v_1 = u_1$$

$$v_2 = u_2 - \mathrm{proj}_{v_1}(u_2) = u_2 - \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1$$

$$v_3 = u_3 - \mathrm{proj}_{v_1}(u_3) - \mathrm{proj}_{v_2}(u_3) = u_3 - \sum_{j=1}^{2} \mathrm{proj}_{v_i}(u_3) = u_3 - \sum_{i=1}^{2} \frac{\langle v_i, u_3 \rangle}{\langle v_i, v_i \rangle} v_i$$

$$\vdots$$

$$v_n = u_n - \sum_{i=1}^{n-1} \mathrm{proj}_{v_i}(u_n) = u_n - \sum_{j=1}^{n-1} \frac{\langle v_i, u_n \rangle}{\langle v_i, v_i \rangle} v_i$$

Then $(v_1, \cdots, v_n)$ is an orthogonal basis for V

Indeed for $1 \le k \le n$, we have

$$v_k = u_k - \sum_{i=1}^{k-1} \frac{\langle v_i, u_k \rangle}{\langle v_i, v_i \rangle} v_i$$

Rewrite as

$$u_k = v_k + \sum_{i=1}^{k-1} \underbrace{\frac{\langle v_i, u_k \rangle}{\langle v_i, v_i \rangle}}_{\in \mathbb{F} \text{ scalar}} v_i$$

So any u_k can be written as a linear combination of the v_j with $j \le k$

Since $u_1, \cdots, u_n$ span $V \Longrightarrow v_1, \cdots, v_n$ span V

$\Longrightarrow \{v_1, \cdots, v_n\}$ is a spanning set of size n and $\dim(V) = n$

$\Longrightarrow$ a basis for V

By lemma 7.33 v_k is orthogonal to $v_1, \cdots, v_{k-1}$

Final step: normalize vectors

$\hat{v}_k = \dfrac{v_k}{\|v_k\|}$ is a unit vector $\Longrightarrow (\hat{v}_1, \cdots, \hat{v}_n)$ is an orthonormal basis

Note: Gram-Schmidt process depends on ordering of $u_1, \dots, u_n$

Change order $\rightsquigarrow$ different orthogonal basis

Example applying Gram-Schmidt process

1) Use Gram-Schmidt process to turn the basis

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

into an orthonormal basis for $\mathbb{R}^3$, standard inner product (dot product)

i) $v_1 = u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

ii) $v_2 = u_2 - \text{proj}_{v_1}(u_2) = u_2 - \dfrac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1$

$$= \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} - 2/2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Longrightarrow v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

iii) $v_3 = u_3 - \text{proj}_{v_1}(u_3) - \text{proj}_{v_2}(u_3) = u_3 - \dfrac{\langle v_1, u_3 \rangle}{\langle v_1, v_1 \rangle} v_1 - \dfrac{\langle v_2 - u_3 \rangle}{\langle v_2, v_2 \rangle} v_2$

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1/6 \\ -1/6 \\ 1/3 \end{pmatrix}$$

Normalising

$$\hat{v}_1 = \frac{v_1}{\|v_1\|} = \frac{\begin{pmatrix}1\\1\\0\end{pmatrix}}{\sqrt{1^2+1^2+0}} = \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}$$

$$\hat{v}_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}$$

$$\hat{v}_3 = \frac{v_3}{\|v_3\|} = \sqrt{6}\begin{pmatrix}1/6\\-1/6\\1/3\end{pmatrix}$$

2) Use Gram-Schmidt to construct orthonormal basis for $\mathbb{C}^2$

Starting with

$$u_1 = \begin{pmatrix}1\\i\end{pmatrix}, \quad u_2 = \begin{pmatrix}2\\-i\end{pmatrix}$$

w.r.t standard Hermitian inner product

$$v_1 = u_1 = \begin{pmatrix}1\\i\end{pmatrix}$$

$$v_2 = u_2 - \mathrm{proj}_{v_1}(u_2) = u_2 - \frac{\langle v_1, u_2\rangle}{\langle v_1, v_1\rangle} = \begin{pmatrix}2\\-i\end{pmatrix} - \frac{\left\langle \begin{pmatrix}1\\i\end{pmatrix}, \begin{pmatrix}2\\-i\end{pmatrix}\right\rangle}{\left\langle \begin{pmatrix}1\\i\end{pmatrix}, \begin{pmatrix}1\\i\end{pmatrix}\right\rangle}\begin{pmatrix}1\\i\end{pmatrix}$$

$$= \begin{pmatrix}2\\i\end{pmatrix} - \frac{1\cdot 2 + \bar{i}\cdot(-i)}{\bar{1}\cdot 1 + \bar{i}\cdot i}\begin{pmatrix}1\\i\end{pmatrix}$$

$$= \begin{pmatrix}2\\-i\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\i\end{pmatrix} = \begin{pmatrix}3/2\\-3/2\, i\end{pmatrix}$$

$\{v_1, v_2\}$ is an orthogonal basis for $\mathbb{C}^2$

Normalizing

$$\|v_1\| = \sqrt{\left\langle \begin{pmatrix}1\\i\end{pmatrix}, \begin{pmatrix}1\\i\end{pmatrix}\right\rangle} = \sqrt{2}$$

$$\|v_2\| = \sqrt{\left\langle \begin{pmatrix}3/2\\-3/2\, i\end{pmatrix}, \begin{pmatrix}3/2\\-3/2\, i\end{pmatrix}\right\rangle} = \sqrt{\overline{\tfrac{3}{2}}\cdot\tfrac{3}{2} + \left(\overline{\tfrac{-3}{2}i}\right)\cdot\left(\tfrac{-3}{2}i\right)} = \frac{3}{2}\sqrt{2}$$

$$\hat{v}_1 = \frac{v_1}{\|v_1\|} = \begin{pmatrix}1/\sqrt{2}\\i/\sqrt{2}\end{pmatrix} \qquad \hat{v}_2 = \frac{v_2}{\|v_2\|} = \begin{pmatrix}1/\sqrt{2}\\-i/\sqrt{2}\end{pmatrix}$$

$\{\hat{v}_1, \hat{v}_2\}$ is an orthonormal basis

3) Legendre Polynomials

$V = \mathbb{R}_2[x]$

$u_1 = 1, \quad u_2 = x, \quad u_3 = x^2, \quad u_4 = x^3$ (standard basis)

Use Gram-Schmidt process to get an orthogonal basis w.r.t inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(t) g(t) \, dt$$

$p_1 = u_1$

$$p_2 = u_2 - \mathrm{proj}_{p_1}(u_2) = u_2 - \frac{\langle p_1, u_2 \rangle}{\langle p_1, p_1 \rangle} p_1 = x - \frac{\overset{0}{\cancel{\int_{-1}^{1} 1 t \, dt}}}{\int_{-1}^{1} 1^2 \, dt} \cdot 1 = x$$

$$p_3 = u_3 - \mathrm{proj}_{p_1}(u_3) - \mathrm{proj}_{p_2}(u_3) = x^2 - \frac{\int_{-1}^{1} 1 \cdot t^2 \, dt}{\int_{-1}^{1} 1 \cdot 1 \, dt} \cdot 1 - \frac{\overset{0}{\cancel{\int_{-1}^{1} t \cdot t^2 \, dt}}}{\int_{-1}^{1} t \cdot t \, dt} x = x^3 - 1/3$$

$$p_4 = x^3 - \frac{3}{5} x$$

We get scaled Legendre polynomials

(standard scaling redefines $p_n(1) = 1$)

Calculations with respect to an orthonormal basis

Theorem

Let $\{v_1, \dots, v_n\}$ be an orthonormal basis for an inner product space V over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$)

For $u, v \in V$, let $u = \sum_{i=1}^{n} \alpha_i v_i$, $v = \sum_{i=1}^{n} \beta_i v_i$ for some $\alpha_i, \beta_i \in \mathbb{F}$

Then

i)
$$\langle u, v \rangle = \begin{cases} \sum_{i=1}^{n} \alpha_i \beta_i & \mathbb{F} = \mathbb{R} \\ \sum_{i=1}^{n} \overline{\alpha_i} \beta_i & \mathbb{F} = \mathbb{C} \end{cases}$$

ii) (Parseval's identity)
$$\|u\|^2 = \begin{cases} \sum_{i=1}^{n} \alpha_i^2 & \mathbb{F} = \mathbb{R} \\ \sum_{i=1}^{n} |\alpha_i|^2 & \mathbb{F} = \mathbb{C} \end{cases}$$

Proof:

If $u = \sum_{i=1}^{n} \alpha_i v_i$ $\quad v = \sum_{i=1}^{n} \beta_i v_i$

Then
$$\alpha_i = \langle v_i, u \rangle$$

If $\mathbb{F} = \mathbb{R}$:

$$\langle u, v \rangle = \left\langle u, \sum_{i=1}^{n} \beta_i v_i \right\rangle = \sum_{i=1}^{n} \beta_i \langle u, v_i \rangle \quad \text{by linearity in second argument}$$

$$= \sum_{i=1}^{n} \beta_i \langle v_i, u \rangle \quad \text{symmetric}$$

$$= \sum_{i=1}^{n} \beta_i \alpha_i$$

$$= \sum_{i=1}^{n} \alpha_i \beta_i$$

So $\|u\|^2 = \langle u, u \rangle = \sum_{i=1}^{n} \alpha_i \alpha_i = \sum_{i=1}^{n} \alpha_i^2$

Similar for $\mathbb{C}$

$\blacksquare$

Example

$V = C([-\pi, \pi], \mathbb{R})$ and inner product

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, g(x)\, dx$$

Given an orthonormal set $\left\{ \frac{1}{\sqrt{2}}, \cos(2x) \right\}$ in V, find

$$\int_{-\pi}^{\pi} \sin^4 x \, dx \quad \text{without computing antiderivative}$$

$\left\{ \frac{1}{\sqrt{2}}, \cos(2x) \right\}$ orthonormal basis for subspace W of V given by

$$W = Sp\left(\frac{1}{\sqrt{2}}, \cos(2x) \right)$$

$\langle , \rangle$ also an inner product on W.

We have $\sin^4 x = (\sin^2 x)^2$

$$\sin^2 x = \frac{1 - \cos(2x)}{2} = \underbrace{\frac{1}{\sqrt{2}}}_{\alpha_1} \cdot \frac{1}{\sqrt{2}} + \underbrace{\left(\frac{-1}{2} \right)}_{\alpha_2} \cos(2x)$$

So
$$\begin{aligned}
\int_{-\pi}^{\pi} \sin^4 x \, dx &= \pi \langle \sin^2 x, \sin^2 x \rangle \\
&= \pi \|\sin^2 x\|^2 \\
&= \pi \left(\left(\frac{1}{\sqrt{2}} \right)^2 + \left(\frac{-1}{2} \right)^2 \right) \\
&= \frac{3\pi}{4}
\end{aligned}$$

PROJECTIONS

Projection maps

Definition Complements/Projections

Let V be vector space, $U \subseteq V$ a subspace

i) A complement of U is a subspace W of V such that

$$V = U \oplus W$$

and $\forall v \in V$ can be uniquely written in form

$$v = u + w \quad \text{with } u \in U \text{ and } w \in W$$

ii) For V, U, W, the unique linear map

$$L : V \longrightarrow U \, ;$$

$$L(v) = u \qquad \forall v \in V$$

is called the projection onto U along W

iii) Furthermore $L : V \to V$ be a linear transformation. Then given a subspace U of V, L is called a projection onto U if it is the projection onto U along some complement of U.

iv) L is called a projection (map) if L is a projection from V to U for some subspace U

Lemma

Let $L : V \to V$ be a linear map. Then

$$L \text{ is a projection} \iff L^2 = L$$

Proof:

($\Rightarrow$): If L is a projection onto U along W, where $V = U \oplus W$

Then any $\forall v \in V$, $v = u + w$, $u \in U$, $w \in W$

So we have $L(v) = u$ and $L(L(v)) = L(u) = u$

$$\Longrightarrow L(L(v)) = L(v)$$

$$\Longrightarrow L^2 = L$$

($\Leftarrow$): Assume $L : V \to V$ is a linear map with $L^2 = L$

Let $W = \ker(L)$ and $U = \ker(\underbrace{I_V}_{\text{identity on } V} - L)$

Then W and U are both subspaces of V

Need to show $V = U \oplus W$, i.e. $\forall v \in V$, we can write v as

$$v = u + w \quad \text{for some } u, w \in V$$

and $U \cap W = \{\underline{0}\}$

Note if $v \in U \cap W \implies v \in U = \ker(I_V - L)$ and $u \in \ker(L)$

(1) $v \in U = \ker(I_V - L) \implies (I_V - L)(v) = 0$

$$\implies I(v) - L(v) = 0$$

$$\implies v - L(v) = 0$$

$$\implies L(v) = v$$

(2) $v \in W = \ker(L) \implies L(v) = 0$

$$\implies v = 0$$

Hence $U \cap W = \{0\}$

Also each $v \in V$ can be written as

$$v = \underbrace{L(v)}_{u \in U} + \underbrace{v - L(v)}_{w \in W}$$

Note that $u = L(v) \in U = \ker(I_V - L)$ since

$$\begin{aligned}(I_V - L)(u) &= (I_V - L)(L(v)) \\ &= L(v) - L(L(v)) \\ &= L(v) - L(v) \qquad L^2 = L \\ &= \underline{0} \implies L(u) = u\end{aligned}$$

and $w = v - L(v) \in W = \ker(L)$ since

$$\begin{aligned}L(w) &= L(v - L(v)) \\ &= L(v) - L(L(v)) \\ &= L(v) - L(v) \\ &= \underline{0}\end{aligned}$$

This proves that $V = U \oplus W$

Since $L(v) = L(u+w) = L(u) + \underset{=0}{L(w)} = u$

$\Longrightarrow$ L is the projection from V to U along W ■

Example of Projections

$\mathbb{R}^2 = U \oplus W$

$U = x\text{-axis}$

$W = y\text{-axis}$

Then map $f: \mathbb{R}^2 \to U$;

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

is a projection of $\mathbb{R}^2$ onto x axis (along y-axis)

Note: $\begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} x-y \\ 0 \end{pmatrix}}_{x\text{-axis}} + \underbrace{\begin{pmatrix} y \\ y \end{pmatrix}}_{\text{line } x=y}$

So $\mathbb{R}^2 = U \oplus W$ where $U = x$ axis

$W = \text{line } x = y$

The map $g: \mathbb{R}^2 \longrightarrow U$

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x-y \\ 0 \end{pmatrix}$$

projection of $\mathbb{R}^2$ onto x-axis (along $x=y$)

Lemma

Every projection map is diagonalizable and has eigenvalues 0 and 1 only

Orthogonal Complements and orthogonal projections

Definition Orthogonal Complement

If V is an inner product space and U is a subspace of V, then the orthogonal complement of V in U is

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

$U^{\perp}$ is a subspace of V

Definition

Let V be an inner product space. Let L be a projection of V to U along W

If $W = U^{\perp}$, then we say L is the orthogonal projection from V to U.

Example.

1) $V = \mathbb{R}^2$

$U = x$-axis and $W = y$ axis, then

$$W = U^{\perp}$$

w.r.t standard inner product, since

$$\begin{pmatrix} x \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ y \end{pmatrix} = 0 \quad \forall x, y \in \mathbb{R}$$

The map $f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$ is the orthogonal projection of $\mathbb{R}^2$ onto U

2) $V = \mathbb{R}^2$

inner product $= \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3 x_2 y_2$

This is a bilinear form, we can represent by a matrix

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

So $\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = (x_1 \ x_2) A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

inner product since A symmetric and positive definite

Sylvester's critereon

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

- $|1| > 0$
- $\begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} = 1 \cdot 3 - (-1)(-1) = 1 > 0$

Note $\underset{\in \mathbb{R}}{\begin{pmatrix} x \\ y \end{pmatrix}} = \underset{\substack{\in U \\ x\text{-axis}}}{\begin{pmatrix} x-y \\ 0 \end{pmatrix}} + \underset{\substack{\in W \\ \text{line } x=y}}{\begin{pmatrix} y \\ y \end{pmatrix}}$

So map $g: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x-y \\ 0 \end{pmatrix}$ is a projection onto U along W

But $\left\langle \begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} y \\ y \end{pmatrix} \right\rangle = xy - xy - 0 \cdot y + 3 0 \cdot y = 0 \quad \forall x, y \in \mathbb{R}$

$\Longrightarrow W = U^\perp$

Hence g is the orthogonal projection onto U w.r.t this inner product

Example

Let V be an inner product space, $u \in V$ a non-zero vector. Then map

$\text{proj}_u : V \longrightarrow V$ (used in Gram-Schmidt process)

given by $\text{proj}_u(v) = \dfrac{\langle u, v \rangle}{\langle u, u \rangle} u$

so a projection from V to $Sp(v)$

Clearly $Im(\text{proj}_u) \subseteq Sp(v)$

Also for $x \in Sp(v)$, $x = \lambda u$ for some $\lambda \in \mathbb{F}$

$$\begin{aligned} \text{proj}_u(x) &= \text{proj}_u(\lambda u) \\ &= \frac{\langle u, \lambda u \rangle}{\langle u, u \rangle} u \\ &= \lambda \frac{\langle u, u \rangle}{\langle u, u \rangle} u \qquad \text{linearity} \\ &= \lambda u = x \end{aligned}$$

$\Longrightarrow \text{proj}_u$ is a projection map.

3. Matrix Decomposition

QR Decomposition

Theorem QR Decomposition

If A is a real $m\times n$ matrix with linearly independent columns (i.e. $\text{rank}(A)=n$), then A can be factored as

$$A = QR$$

where Q is the $m\times n$ matrix with orthonormal column vectors, and R is an $n\times n$ invertible upper triangular matrix

The decomposition can be found by applying Gram-Schmidt process to column vectors

$u_1, \cdots, u_n$ of A.

Then Q consists of the columns $\hat{v}_1, \cdots, \hat{v}_n$

Entries of R are given by

$$R_{jk} = \langle \hat{v}_j, u_k \rangle = \left\langle \frac{v_j}{\|v_k\|}, u_k \right\rangle = \frac{\langle v_j, u_k \rangle}{\sqrt{\langle v_j, v_j \rangle}} = \frac{\langle v_j, u_k \rangle}{\langle v_j, v_j \rangle} \|v_j\|$$

By construction, R has 0's below the diagonal

Example of QR Decomposition

Using example on page 71

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad u_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \quad u_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = (u_1 \;\; u_2 \;\; u_3)$$

Applying Gram-Schmidt

$$\hat{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \quad \hat{v}_2 = \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \quad \hat{v}_3 = \begin{pmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}$$

$$Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{pmatrix} = (\hat{v}_1, \hat{v}_2, \hat{v}_3)$$

$R_{jk} = \langle \hat{v}_j, u_k \rangle$, therefore

- $R_{11} = \langle \hat{v}_1, u_1 \rangle = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2}$

- $R_{12} = \langle \hat{v}_1, u_2 \rangle = \langle \hat{v}_1, u_2 \rangle = \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_2 \rangle} \|v_1\| = 1 \cdot \sqrt{2} = \sqrt{2}$

- $R_{13} = \langle \hat{v}_1, u_3 \rangle = \frac{1}{\sqrt{2}}$

- $R_{21} = \langle \hat{v}_2, u_3 \rangle = 0$
- $R_{22} = \langle \hat{v}_2, u_2 \rangle = \sqrt{2}$
- $R_{23} = \langle \hat{v}_2, u_1 \rangle = \frac{2}{\sqrt{3}}$
- $R_{31} = \langle \hat{v}_3, u_1 \rangle = 0$
- $R_{32} = \langle \hat{v}_3, u_2 \rangle = 0$
- $R_{33} = \langle \hat{v}_3, u_3 \rangle = \frac{1}{\sqrt{6}}$

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3} & 2/\sqrt{3} \\ 0 & 0 & 1/\sqrt{6} \end{pmatrix}$$

4. Spectral Theorems

SELF ADJOINT LINEAR MAPS

Definition Adjoint

Let V be an inner product space, real or complex.

Let $L: V \longrightarrow V$ be a linear map. The adjoint of L^* is the linear map

$$L^*: V \longrightarrow V$$

$$\langle L^*(u), v \rangle = \langle u, L(v) \rangle \qquad \forall\, u, v \in V$$

If $L = L^*$, then we say that L is a self-adjoint linear map

Adjoint vs Dual Map

Recall that if $L: V \longrightarrow W$ is a linear map, then it has a corresponding dual map

$$L^*: W^* \longrightarrow V^*$$

between the corresponding dual spaces, given by

$$L^*(f) = f \circ L \quad , \quad f \in W^*$$

In the example, $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, if L is represented (w.r.t standard basis) by matrix A, then L^* is represented (with respect to the dual of the standard basis) by the matrix A^T.

Using the same notation is not a coincidence;

Every finite dimensional vector space V is isomorphic to its dual space

$$V \cong V^*$$

So in the case of a linear map $L: V \rightarrow V$, its dual map is $L^*: V^* \rightarrow V^*$ but using an isomorphism b/w V and V^*, we can choose to view the dual map L^* as a map from V to V

$$L^*: V \longrightarrow V$$

So the dual map can be identified with the adjoint map, hence we use the same notation for it both represented by A^T.

Many isomorphisms between V and V^*. If V is an inner product space, can also use an inner product to define an isomorphism b/w V and V^*

Examples of self adjoint linear maps

$V = \mathbb{R}^n$, standard inner product

$L: V \to V$ given by

$$L: \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{where } A \in M_{n \times n}(\mathbb{R})$$

matrix represent L w.r.t standard basis

Then its adjoint $L^*: V \to V$ is given by

$$L^*: \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

<u>Check</u>: $\underline{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $\quad \underline{v} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\langle L^*(u), v \rangle = \left\langle L^*\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right), \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle$$

$$= \left\langle A^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle$$

standard inner product $\langle u, v \rangle = u^T v$

$$= \left(\left(A^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right)$$

$$= \left(A^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)^T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= (x_1 \cdots x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle$$

$$= \langle u, Av \rangle$$

$$= \langle u, L(v) \rangle$$

$\Longrightarrow L^*$ adjoint of L

L self-adjoint $\iff L = L^*$

$\iff A = A^T$

$\iff A$ symmetric

For V finite dimensional, many isomorphisms $V \longrightarrow V^*$ (one for each choice of basis)

If V is also an inner product space, can define

$$T : V \longrightarrow V^* \text{ by}$$

$$T(v) = \langle v, - \rangle$$

i.e. T takes $v \in V$ and outputs the linear functional $f = T(v) \in V^*$ where $f : V \longrightarrow \mathbb{F}$ given by

$$f(w) = \langle v, w \rangle \quad \forall w \in V$$

T is an isomorphism

Example

$V = \mathbb{C}^2$ with standard Hermitian product and let

$$L : V \to V \text{ be given by}$$

$$L : \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a + ib \\ -ai \end{pmatrix}$$

Showing L is self adjoint

L is self adjoint $\iff L = L^*$

$\iff \langle L(u), v \rangle = \langle u, L(v) \rangle \quad \forall u, v \in V$

Let $u = \begin{pmatrix} a \\ b \end{pmatrix} \quad v = \begin{pmatrix} c \\ d \end{pmatrix}$

$$\langle L(u), v \rangle = \left\langle L\left(\begin{pmatrix} a \\ b \end{pmatrix}\right), \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle$$

$$= \left\langle \begin{pmatrix} a + ib \\ -ai \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle$$

$$= \overline{a + ib} \cdot c + \overline{(-ai)} \cdot d$$

$$= \bar{a}c + i\bar{b}c + aid$$

$$\langle u, L(v) \rangle = \left\langle \begin{pmatrix} a \\ b \end{pmatrix}, L\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \right\rangle = \left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c + id \\ -ci \end{pmatrix} \right\rangle$$

$$= \bar{a} \cdot (c + id) + \bar{b} \cdot (-ci)$$

$= \bar{a}c - \bar{b}ci + adi$

Expressions equal $\Longrightarrow$ self adjoint

Apply L to standard basis, for $\mathbb{C}^2$ to find matrix representing L (w.r.t basis)

$$L\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1+0i\\-i\end{pmatrix} = \begin{pmatrix}1\\-i\end{pmatrix}$$

$$L\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0+i1\\0\end{pmatrix} = \begin{pmatrix}i\\0\end{pmatrix}$$

$$A = \begin{pmatrix}1 & i\\-i & 0\end{pmatrix} \qquad A \text{ Hermitian}$$

Eigenvalues

$$\det\begin{pmatrix}\lambda-1 & -i\\ i & \lambda\end{pmatrix} = \lambda^2 - \lambda^2 - 1 = 0$$

$$\Longrightarrow \lambda = \frac{1 \pm \sqrt{5}}{2} \quad \text{real}$$

Self-adjoint operators vs Hermitian matrices

Proposition

Let V be an inner product space, $L: V \to V$ be a linear map and $\mathcal{B}$ be an orthonormal ordered basis for V

Then L is self-adjoint $\Longleftrightarrow$ $M_{\mathcal{B}}(L)$, the matrix representing L w.r.t $\mathcal{B}$ is Hermitian

Proof:

Let $\mathcal{B} = (v_1, \dots, v_n)$ be an orthonormal ordered basis for V

Let $A = M_{\mathcal{B}}(L)$ be the matrix representing L w.r.t basis $\mathcal{B}$

$B = M_{\mathcal{B}}(L^*)$ be the matrix representing L^* w.r.t basis $\mathcal{B}$

($\Rightarrow$): Recall

$$L(v_j) = \sum_{k=1}^{n} A_{kj} v_k$$

so by linearity in 2nd argument, we have

$$\langle v_i, L(v_j) \rangle = \langle v_i, \sum_{k=1}^{n} A_{kj} v_k \rangle$$

$$= \sum_{k=1}^{n} A_{kj} \langle v_i, v_k \rangle$$ only non-zero term is $\langle v_i, v_i \rangle = 1$, $k=i$

$$= A_{ij}$$

$$L^*(v_i) = \sum_{k=1}^{n} B_{ki} v_k$$

So

$$\langle L^*(v_j), v_j \rangle = \left\langle \sum_{k=1}^{n} B_{ki} v_k , v_j \right\rangle$$

$$= \sum_{k=1}^{n} \bar{B}_{ki} \langle v_k, v_i \rangle$$ only non-zero term

$$= \bar{B}_{ji}$$

So $A_{ij} = \bar{B}_{ji} \ \forall \, ij \implies A = A^{+}$

$\implies A$ is Hermitian

$(\Longleftarrow)$: Assume A Hermitian and $\mathcal{B} = (v_1, \dots, v_n)$ be an orthonormal basis of V. Let

$$L : V \to V$$

be the linear map represented by A w.r.t A.

By going backwards direction of "$\implies$" proof, we get

$$\langle L(v_i), v_j \rangle = \langle v_i, L(v_j) \rangle \quad \forall \, v_i, v_j \in \mathcal{B}$$

Let $u, v \in V$ with $u = \sum_{i=1}^{n} \alpha_i v_i$

$$v = \sum_{j=1}^{n} \beta_i v_j \qquad \alpha_j, \beta_j \in \mathbb{C}$$

Then $\langle L(u), v \rangle = \left\langle L\left(\sum_{i=1}^{n} \alpha_i v_i \right), \sum_{j=1}^{n} \beta_j v_j \right\rangle$

$$= \left\langle \sum_{i=1}^{n} \alpha_i L(v_i), \sum_{j=1}^{n} \beta_j v_j \right\rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{\alpha}_i \beta_j \langle L(v_i), v_j \rangle$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{\alpha}_i \beta_i \langle v_i, L(v_j) \rangle$$

$$= \left\langle \sum_{i=1}^{n} \alpha_i v_i , L\left(\sum_{j=1}^{n} \beta_j v_j\right) \right\rangle$$

$$= \langle u, L(v) \rangle$$

Proposition

The eigenvalues of self-adjoint maps are real

Proof:

Let $L: V \to V$ be self adjoint

An eigenvalue λ of L satisfies

$$L(v) = \lambda v \quad \text{for some non-zero } v \in V$$

We have $\langle v, L(v) \rangle = \langle v, \lambda v \rangle = \lambda \langle v, v \rangle$

Since L self adjoint, we have $\langle v, L(v) \rangle = \langle L(v), v \rangle$

$$= \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle$$

$v \neq 0 \implies \langle v, v \rangle \neq 0$ (positive definite)

$\implies \lambda = \bar{\lambda}$

$\implies \lambda \in \mathbb{R}$

Diagonalizability of self-adjoint maps, and orthogonal eigenvectors

Theorem

A linear map $L: V \to V$ on a finite dimensional vector space is said to be diagonalizable if

(1) V has a basis of eigenvectors

$$\Updownarrow$$

(2) minimal polynomial $d_L(x)$ of L factoring into distinct linear factors (no repeat roots)

Proposition

Every self-adjoint $L: V \longrightarrow V$ on a finite dimensional complex inner product space V is diagonalizable

<u>**Proof:**</u> (by contradiction):

Since we are working over $\mathbb{C}$, every non-constant polynomial is a product of linear factors

So $d_L(x)$, minimal polynomial of L is a product of linear factors

$$L \text{ diagonalizable} \iff d_L(x) \text{ has no repeated factors}$$

If $d_L(x)$ has repeated factors, then

$$d_L(x) = (x-\lambda)^2 p(x) \text{ for some polynomial } p(x)$$

Hence

$$d_L(L) = (L - \underbrace{I}_{\text{identity map}}\lambda)^2 p(L) = \underbrace{0}_{0 \text{ map}}$$

but $\exists\, v \in V$ such that $((L-\lambda I)\, p(L))(v) \neq \underline{0}$

Hence

$$* \quad \langle ((L-\lambda I)\, p(L))(v), ((L-\lambda I)\, p(L))(v) \rangle \neq 0$$

Note that since L is self adjoint, for any $u_1, u_2 \in V$, we have

$$\begin{aligned}
\langle (L-\lambda I)(u_1), u_2 \rangle &= \langle L(u_1) - \lambda u_1, u_2 \rangle \\
&= \langle L(u_1), u_2 \rangle - \bar{\lambda} \langle u_1, u_2 \rangle \\
&= \langle L(u_1), u_2 \rangle - \lambda \langle u_1, u_2 \rangle \\
&= \langle u_1, L(u_2) - \lambda u_2 \rangle \\
&= \langle u_1, (L-\lambda I)(u_2) \rangle
\end{aligned}$$

since L is self adjoint and λ eigenvalue of L, hence real

So $*$ becomes

$$0 \neq \langle (p(L))(v), (L-\lambda I)^2 (p(L))(v) \rangle = \langle (p(L))(v), \underline{0} \rangle = 0$$

Contradiction ※

Hence $d_L(x)$ factors into distinct linear factors $\Longrightarrow$ diagonalizable

■

Proposition

Every self-adjoint linear map $L: V \longrightarrow V$ on a finite dimensional real inner product space is diagonalizable

proof:

Let V be a real finite dimensional inner product space

$L: V \rightarrow V$ be self adjoint

Let A be the matrix of L w.r.t some orthonormal basis of V

Then by prop pg 86 A is Hermitian (Also A is real $\Longrightarrow$ symmetric)

Let $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ be the linear map whose matrix w.r.t the standard basis for $\mathbb{C}^n$ is A.

Then by prop pg 86 T is self adjoint

prop pg 88, eigenvalues are real

prop pg 89 T is diagonalizable $\Longrightarrow$ minimal polynomial $d_T(x)$ is a product of distinct linear factors

(of form $x - \lambda \in \mathbb{R}$)

d_T is also the minimal polynomial of A $\Longrightarrow$ minimal polynomial of L

$\Longrightarrow$ L is diagonalizable

Proposition

If $L: V \longrightarrow V$ is a self adjoint linear map, then any 2 eigenvectors of L associated to distinct eigenvalues are orthogonal

Proof:

Let v_1, v_2 be eigenvectors of L with eigenvalues of L with eigenvalues λ_1, λ_2, $\lambda_1 \neq \lambda_2$

$$\langle v_1, L(v_2) \rangle = \langle v_1, \lambda v_2 \rangle$$
$$= \lambda \langle v_1, v_2 \rangle$$

L self adjoint

$$\langle v_1, L(v_2) \rangle = \langle L(v_1), v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle$$
$$= \bar{\lambda} \langle v_1, v_2 \rangle$$

$= \lambda \langle v_1, v_2 \rangle$ eigenvalues real, self adjoint map

$\Rightarrow \lambda_1 \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$

$\Rightarrow \underset{\neq 0}{(\lambda_1 - \lambda_2)} \langle v_1, v_2 \rangle = 0$

Spectral Theorems

Theorem Spectral Theorem for self-adjoint linear transformations

Let V be finite dimensional vector space (complex or real) inner product space, and let

$$L : V \longrightarrow V$$

be a self-adjoint map on V. Then $\exists$ an orthonormal basis for V such that the matrix representing L w.r.t that basis is diagonal with all entries real

Proof:

Since L is diagonalizable, we have

$$V = \bigoplus_{\lambda} \ker(L - \lambda I)$$ eigenspace

and by prop pg 90, each kernel is orthogonal to all others.

Using Gram-Schmidt process, we may choose an orthonormal basis for each kernel.

Hence, we have an orthonormal basis of eigenvectors for V which diagonalizes L

So matrix w.r.t this basis is diagonalizable, real

$\blacksquare$

Corollary Spectral Theorem for Hermitian matrices

Any complex Hermitian nxn matrix A is diagonalizable, all its eigenvalues are real.

The basis of eigenvectors diagonalizing A can be chosen to be orthonormal for the standard Hermitian inner product on $\mathbb{C}^n$.

Hence $\exists$ a unitary matrix U such that

$$U^{-1} A U \text{ is diagonal}$$

Corollary

Any real symmetric $n \times n$ matrix A is diagonalizable, and all its eigenvalues are real

The basis of eigenvectors diagonalizing A can be chosen to be orthonormal for the standard inner product on $\mathbb{R}^n$

Hence $\exists$ an orthogonal matrix Q s.t

$$Q^{-1}AQ \text{ is diagonal}$$

Examples

Find a unitary matrix A that diagonalizes

$$A = \begin{pmatrix} 1 & 0 & i \\ 0 & 2 & 0 \\ -i & 0 & 1 \end{pmatrix}$$

A is Hermitian $\iff A^{\dagger} = A$

$\implies$ we can find such a matrix

Finding eigenvalues of A,

$$\det(\lambda I - A) = (\lambda - 2)^2 \lambda = 0 \implies \lambda = 0, \ \lambda = 2$$

Eigenspace of $\cdot\ \lambda = 2$: $\operatorname{Ker}(A - 2I) = \operatorname{Sp}\left\{ \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix}, \begin{pmatrix} i \\ 1 \\ 1 \end{pmatrix} \right\}$

$\lambda = 0$: $\operatorname{Ker}(A) = \operatorname{Sp}\left\{ \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix} \right\}$

Diagonalizable

$$P = \begin{pmatrix} 1 & i & 1 \\ 0 & 1 & 0 \\ i & 1 & i \end{pmatrix}$$

with

$$P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

But P is not unitary (since columns not orthonormal)

Let u_1, u_2, u_3 be columns of D.

By prop 9.8, $u_1 \perp u_2$, $u_2 \perp u_3$ (since eigenvalues distinct)

Apply G-S process to u_1, u_2

$v_1 = u_1$

$$v_2 = u_2 - \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Then $\{v_1, v_2, u_3\}$ is an orthogonal basis

Normalise. Hence

$$U = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -i/\sqrt{2} & 0 & i/\sqrt{2} \end{pmatrix}$$

ISOMETRIES AND NORMAL MATRICES

Isometries

Definition Isometry

A linear map $L: V \to V$ on an inner product space V is an isometry if it preserves the inner product i.e.

$$\langle L(u), L(v)\rangle = \langle u, v\rangle$$

Proposition

The eigenvalues of an isometry have modulus 1

Proof:

real case follows

Let V be a complex inner product space.

Let $L: V \to V$ be an isometry

Let v be an eigenvector of L, eigenvalue λ

$$\underline{v} \neq \underline{0} \quad \text{and} \quad L(v) = \lambda v$$

We have

$$\langle v, v\rangle \underset{\text{isometry}}{=} \langle L(v), L(v)\rangle = \langle \lambda v, \lambda v\rangle = \bar{\lambda}\lambda\langle v, v\rangle = |\lambda|^2\langle v, v\rangle$$

Since $v \neq \underline{0}$, $\langle v, v\rangle \neq 0 \implies |\lambda|^2 = 1$

$$\implies |\lambda| = 1$$

■

Proposition

Let V be a finite dimensional complex inner product space Let

$$L: V \to V$$

be an isometry. Then there is a basis for V diagonalizing L

Spectral Theorem for isometries

Theorem

Let L be an isometry on a finite dimensional complex inner product space V.

There exists an orthonormal basis for V relative to which the matrix of L is diagonal, with all eigenvalues having modulus 1

Spectral Theorem for unitary matrices

A square complex square matrix is unitary

$$\Updownarrow$$

$U^{-1} = U^{\dagger}$ $\iff$ columns are orthonormal w.r.t standard Hermitian inner product.

Proposition

Let V be a complex inner product space,

$$L : V \longrightarrow V$$

be a linear map and $\mathcal{B} = (v_1, \dots, v_n)$ be an orthonormal ordered basis for V

Then L is an isometry $\iff$ $M_{\mathcal{B}}(L)$ is unitary

Lemma

An eigenvalue λ of a unitary matrix satisfies $|\lambda| = 1$

Corollary

The eigenvalues of an orthogonal matrix (over $\mathbb{R}$), if they exist, are 1 or -1

Theorem Spectral Theorem for unitary matrices

Any unitary matrix U is diagonalizable, and all its eigenvalues have absolute value 1.

The basis of eigenvectors diagonalizing U can be chosen to be orthonormal for the standard Hermitian inner product on $\mathbb{C}^n$.

Normal Matrices and commuting linear maps

Definition Normal

A complex square matrix A is said to be **normal** if it commutes with its conjugate transpose

$$AA^{\dagger} = A^{\dagger}A$$

Hermitian, real symmetric, unitary and (real) orthogonal matrices are all normal

Definition Invariant Subspace

If $L: V \longrightarrow V$ be a linear map, U is a subspace of V.

U is said to be an **invariant subspace** for L if

$$L(u) \in U \qquad \forall u \in U$$

This is equivalent to saying we can restrict the map L to U

$$L|_U : U \longrightarrow U \quad \text{defined by}$$

$$L|_U(u) = L(u) \quad \forall u \in U$$

defines a linear map

Lemma

Let $A, B: V \longrightarrow V$ be commuting linear operators, i.e.

$$AB = BA$$

Then any eigenspace for A is an invariant subspace for B

__Proof__:

Let $A, B: V \longrightarrow V$ be linear maps with $AB = BA$

Let λ be an eigenvalue of A, v be an eigenvector of λ

$$v \in V_{\lambda} = \ker(A - \lambda I) \quad \text{corresponding eigenspace}$$

$$\begin{aligned} AB(v) &= A(B(v)) \\ &= (BA)(v) \\ &= B(A(v)) \end{aligned}$$

$$= B(\lambda v)$$

$$= \lambda B(v)$$

$\Longrightarrow$ $B(v)$ is an eigenvector for A with eigenvalue λ

$$B(v) \in V_\lambda$$

V_λ is an invariant subspaces for B ■

Simultaneous Diagonalisability and diagonalizability by unitary matrices

Theorem

Let V be a finite dimensional vector space

Let $\{A_i\}_{i \in I}$ be a family of commuting linear operators $V \to V$, and assume A_i is diagonalizable for every i

Then the A_i are simultaneously diagonalizable, meaning there exists a basis of V w.r.t which all A_i are represented by diagonal matrices

Theorem

A square matrix A can be diagonalized by a unitary matrix

$$\Updownarrow$$

A is normal