MA 110
Linear Algebra and Differential Equations
Lecture 20

Prof. Sudhir R. Ghorpade
Department of Mathematics
IIT Bombay
http://www.math.iitb.ac.in/∼srg/

Spring 2025

Recall: We discussed the following.

- The notion of (abstract) vector space over $\mathbb{K}$.
- Examples: $\mathbb{K}^{n\times 1}$, $\mathbb{K}^{1\times n}$, $\mathbb{K}^{m\times n}$, $\mathbb{K}[x]$, $\mathcal{P}_n$, $C[a,b]$, $C^1[a,b]$, c, c_0
- Subspace of a vector space
- Linear combinations
- Span of a subset of a vector space
- Linear dependence and linear independence
- Crucial Result: Let S be a subset of s elements and R be a set of r elements of V. If $S \subset \operatorname{span} R$ and $s > r$, then S is linearly dependent.
- Notion of a finite dimensional vector space
- Basis and dimension

Definition

Any linearly independent subset of a finite dimensional vector space V which spans V is called a **basis** *for V.*

Here is the most important result about finite dimensional vector spaces. The proof is similar to that in the case of subspaces of $\mathbb{K}^n$.

Proposition

Let V be a finite dimensional vector space over $\mathbb{K}$. Then the following holds.

- *V has a basis.*
- *Every set that spans V has a subset which is a basis of V.*
- *Every linearly independent subset of V can be extended to a basis of V.*
- *Any two bases of V have the same cardinality, called the* **dimension** *of V and denoted by* $\dim V$.

Linear Transformations

Definition

Let V and W be vector spaces over $\mathbb{K}$. A **linear transformation** *or a* **linear map** *from V to W is a function $T : V \to W$ which 'preserves' the operations of addition and scalar multiplication, that is, for all $u, v \in V$ and $\alpha \in \mathbb{K}$,*

$$T(u+v) = T(u) + T(v) \quad \textit{and} \quad T(\alpha\, v) = \alpha\, T(v).$$

It is clear that if $T : V \to W$ is linear, then $T(0) = 0$. Also, T 'preserves' linear combinations of elements of V:

$$T(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 T(v_1) + \cdots + \alpha_k T(v_k)$$

for all $k \in \mathbb{N}$, $v_1, \ldots, v_k \in V$ and $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$.

Remark: Linear transformations $T : V \to V$ on a vector space V into itself are sometimes called **linear operators** on V.

Examples

1. Let $\mathbf{A}$ be an $m \times n$ matrix with entries in $\mathbb{K}$. Then the map $T : \mathbb{K}^{n\times 1} \to \mathbb{K}^{m\times 1}$ defined by $T(\mathbf{x}) := \mathbf{A}\,\mathbf{x}$ is linear. Similarly, the map $T' : \mathbb{K}^{1\times m} \to \mathbb{K}^{1\times n}$ defined by $T'(\mathbf{y}) := \mathbf{yA}$ is linear. More generally, the map

$$T : \mathbb{K}^{n\times p} \to \mathbb{K}^{m\times p} \quad \text{defined by} \quad T(\mathbf{X}) := \mathbf{A}\,\mathbf{X}$$

is linear, and the map

$$T' : \mathbb{K}^{p\times m} \to \mathbb{K}^{p\times n} \quad \text{defined by} \quad T'(\mathbf{Y}) := \mathbf{YA}$$

is linear.

2. $T : \mathbb{K}^{m\times n} \to \mathbb{K}^{n\times m}$ defined by $T(\mathbf{A}) := \mathbf{A}^{\mathsf{T}}$ is linear.

3. The map $T : \mathbb{K}^{n\times n} \to \mathbb{K}$ defined by $T(\mathbf{A}) := \operatorname{trace}\mathbf{A}$ is linear. But $\mathbf{A} \longmapsto \det \mathbf{A}$ does not define a linear map.

4. The map $T : \mathbb{K}[X] \to \mathbb{K}$ defined by $T(p(X)) = p(0)$ is linear.

5. Let $V := c_0$, the set of all sequences in $\mathbb{K}$ which converge to 0. Then the map $T : V \to V$ defined by

$$T(x_1, x_2, \ldots) := (0, x_1, x_2, \ldots)$$

is linear, and so is the map $T' : V \to V$ defined by

$$T'(x_1, x_2, \ldots) := (x_2, x_3, \ldots).$$

Note that $T' \circ T$ is the identity map on V, but $T' \circ T$ is not the identity map on V. The map T is called the **right shift operator** and T' is called the **left shift operator** on V.

6. Let $V := C^1([a, b])$, the set of all real-valued continuously differentiable functions, and let $W := C([a, b])$, the set of all real-valued continuous functions on $[a, b]$. Then the map $T' : V \to W$ defined by $T'(f) = f'$ is linear. Also, the map

$$T : W \to V \text{ defined by } \ T(f)(x) := \int_a^x f(t)dt \ \text{ for } x \in [a, b],$$

is linear. [Question. What are $T' \circ T$ and $T' \circ T$?]

Let V and W be vector spaces over $\mathbb{K}$, and let $T : V \to W$ be a linear map. Two important subspaces associated with T are

(i) $\mathcal{N}(T) := \{v \in V : T(v) = 0\}$, the **null space** of T, which is a subspace of V,

(ii) $\mathcal{I}(T) := \{T(v) : v \in V\}$, the **image space** of T, which is a subspace of W.

Suppose V is finite dimensional, and let $\dim V = n$. Since $\mathcal{N}(T)$ is a subspace of V, it is finite dimensional and $\dim \mathcal{N}(T) \leq n$

Let $v_1, \dots, v_n$ be a basis for V. If $v \in V$, then there are $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, so that $T(v) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n)$. This shows that $\mathcal{I}(T) = \text{span}\{T(v_1), \dots, T(v_n)\}$. Hence $\mathcal{I}(T)$ is also finite dimensional and $\dim \mathcal{I}(T) \leq n$.

Definition

The dimension of $\mathcal{N}(T)$ is called the **nullity** *of the linear map T, and the dimension of $\mathcal{I}(T)$ is called the* **rank** *of T.*

The Rank-Nullity Theorem for a matrix $\mathbf{A}$ that we proved earlier is a special case of the following result.

Proposition (Rank-Nullity Theorem for Linear Maps)

Let V and W be vector spaces over $\mathbb{K}$, and let $T : V \to W$ be a linear map. Suppose $\dim V = n \in \mathbb{N}$. Then

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

Proof (Sketch): Let $s := \text{nullity}(T)$ and let $\{u_1, \ldots, u_s\}$ be a basis of $\mathcal{N}(T)$. Extend the linearly independent set $\{u_1, \ldots, u_s\}$ to a basis $\{u_1, \ldots, u_s, u_{s+1}, \ldots, u_n\}$ of V. Check that the set $\{T(u_{s+1}), \ldots, T(u_n)\}$ is a basis of $\mathcal{I}(T)$. □

Corollary

Suppose V and W be finite dimensional vector spaces, and $T : V \to W$ is a linear map. Let $\dim V = n$ and $\dim W = m$.Then

$$T \text{ is one-one} \iff \text{rank}(T) = n.$$

In particular, if T is one-one, then $n \leq m$. If $m = n$, then

$$T \text{ is one-one} \iff T \text{ is onto}$$

Proof. The first assertion follows from the Rank-Nullity Theorem since T is one-one $\iff \mathcal{N}(T) = \{0\}$, that is, $\text{nullity}(T) = 0$. In particular, if T is one-one, then $n = \text{rank}(T) = \dim \mathcal{I}(T) \leq \dim W = m$. Finally, if $m = n$, then the last assertion follows from the first assertion, since $\text{rank}(T) = m \iff T$ is onto. □

As another application of the Rank-Nullity Theorem, we find an interesting relation between dimensions of finite dimensional subspaces of a vector space.

Proposition

Let W_1 and W_2 be finite dimensional subspaces of a vector space V. Then

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2).$$

Proof. Let $W_1 \times W_2 := \{(w_1, w_2) : w_1 \in W_1 \text{ and } w_2 \in W_2\}$. This is a vector space (w.r.t. componentwise addition and scalar multiplication) and $\dim(W_1 \times W_2) = \dim W_1 + \dim W_2$. Define $T : W_1 \times W_2 \to W_1 + W_2$ by $T(w_1, w_2) := w_1 - w_2$. Then T is linear, $\mathcal{N}(T) = \{(w, w) : w \in W_1 \cap W_2\}$ and $\mathcal{I}(T) = W_1 + W_2$. Hence by the Rank-Nullity Theorem,

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim(W_1 \times W_2). \quad \square$$

Matrix of a Linear Transformation

Suppose V, W are vector spaces over $\mathbb{K}$ and $E = (v_1, \ldots, v_n)$, $F = (w_1, \ldots, w_m)$ are their ordered bases, then

$$\text{linear map } T : V \to W \ \rightsquigarrow \ \text{matrix } \mathbf{M}_F^E(T) = A = [a_{jk}]$$

where $A = [a_{jk}]$ is the $m \times n$ matrix determined by

$$T(v_k) = \sum_{j=1}^{m} a_{jk} w_j \quad \text{for } k = 1, \ldots, n.$$

Example: If $I_V : V \to V$ is the identity map, then $\mathbf{M}_E^E(I_V) = \mathbf{I}$.

Basic Property: If U is another vector space, $D = (u_1, \ldots, u_p)$ an ordered basis of U, and $\mathbf{S} : U \to V$ a linear map, then

$$\mathbf{M}_F^D(T \circ S) = \mathbf{M}_F^E(T)\mathbf{M}_E^D(S).$$

Simple Exercise: If a linear map $T : V \to W$ is invertible, then $T^{-1} : W \to V$ is also a linear map.

The above example and the basic property implies that

$$T : V \to W \text{ is invertible } \iff \mathbf{M}_F^E(T) \text{ is invertible}$$

Moreover

$$\mathbf{M}_F^E(T)^{-1} = \mathbf{M}_E^F(T^{-1}).$$

Effect of Change of Basis: We can also use this and the basic property to relate $\mathbf{A} := \mathbf{M}_F^E(T)$ with $\mathbf{A}' := \mathbf{M}_{F'}^{E'}(T)$, where E', F' are some other ordered bases of V, W, as follows.

$$\mathbf{M}_{F'}^{E'}(T) = \mathbf{M}_{F'}^{F}(I_W)\mathbf{M}_F^E(T)\mathbf{M}_E^{E'}(I_V),$$

i.e., $\mathbf{A}' = \mathbf{QAP}$, where $\mathbf{Q} = \mathbf{M}_{F'}^{F}(I_W)$ and $\mathbf{P} = \mathbf{M}_E^{E'}(I_V)$ are invertible matrices of sizes $m \times m$ and $n \times n$, respectively.

Important Special Case: $W = V$ and $F = E$ and $F' = E'$. In this case, $\mathbf{Q} = \mathbf{M}_{E'}^{E}(I_V) = \mathbf{M}_E^{E'}(I_V)^{-1} = \mathbf{P}^{-1}$ and thus

$$\mathbf{A}' = \mathbf{M}_{E'}^{E'}(T) = \mathbf{M}_{E'}^{E}(I_V)\mathbf{M}_E^E(T)\mathbf{M}_E^{E'}(I_V) = \mathbf{P}^{-1}\mathbf{AP}.$$

In other words, $\mathbf{A}'$ is similar to $\mathbf{A}$.

Remark: The correspondence between an $m \times n$ matrix and a linear map from an n dimensional vector space V to an m dimensional vector space W allows us to obtain two versions of the same result such as the Rank-Nullity Theorem: a version using matrices, and another one using abstract vector spaces. Any one version can be derived from the other.

Example: For $n \in \mathbb{N}$, let $\mathcal{P}_n$ denote the vector space of all polynomials of degree less than or equal to n. Define $T : \mathcal{P}_n \to \mathcal{P}_{n-1}$ by $T(p(x)) = p'(x)$, the derivative of $p(x)$. Let $E := (1, x, \ldots, x^n)$ and $F := (1, x, \ldots, x^{n-1})$ be the ordered bases of $\mathcal{P}_n$ and $\mathcal{P}_{n-1}$ respectively. Then the $n \times (n+1)$ matrix of T with respect to these bases is

$$\mathbf{M}_F^E(T) := \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & n \end{bmatrix}.$$

Eigenvalue Problems for Linear Operators

Definition

Let V be a vector space over $\mathbb{K}$, and $T : V \to V$ a **linear operator**. *A scalar $\lambda \in \mathbb{K}$ is called an* **eigenvalue** *of T if there is a nonzero $v \in V$ such that $T(v) = \lambda v$, and then v is called an* **eigenvector** *or an* **eigenfunction** *of T corresponding to λ, and the subspace $\mathcal{N}(T - \lambda I)$ is called the* **eigenspace** *of T. The dimension of this eignspace if called the* **geometric multiplicity** *of λ as an eigenvalue of T*

Example: Let V denote the vector space of all real-valued infinitely differentiable functions on $\mathbb{R}$. Define $T(f) = f'$ for $f \in V$. Then T is a linear operator on V. Given $\lambda \in \mathbb{R}$, consider $f_\lambda(t) := e^{\lambda t}$ for $t \in \mathbb{R}$. Then $f_\lambda \in V$, $f_\lambda \neq 0$ and $T(f_\lambda) = \lambda f_\lambda$. Thus every $\lambda \in \mathbb{R}$ is an eigenvalue of T with f_λ as a corresponding eigenfunction. In fact, any eigenfunction of T corresponding to λ is a scalar multiple of f_λ.

We now consider a vector space V of finite dimension n and a linear operator $T : V \to V$. Fixing an an ordered basis $E = (v_1, \ldots, v_n)$ of V, we can associate to T an $n \times n$ matrix $\mathbf{A} := \mathbf{M}_E^E(T)$. Observe that if $\lambda \in \mathbb{K}$ and $v \in V$, then

$$T(v) = \lambda v \iff A\mathbf{x} = \lambda \mathbf{x},$$

where $\mathbf{x} = [x_1, \ldots, x_n]^T$ with $x_1, \ldots, x_n \in \mathbb{K}$ determined by writing $v = x_1 v_1 + \cdots + x_n v_n$. Thus, we see that

$$\lambda \text{ is an eigenvalue of } T \iff \lambda \text{ is an eigenvalue of } A.$$

With this in view, we define the **characteristic polynomial** of T to be the characteristic polynomials of $\mathbf{A}$. The **algebraic multiplicity** of an eigenvalue λ of T is defined to be the algebraic multiplicity of λ as an eigenvalue of $\mathbf{A}$. Further, the linear operator T is said to be **diagonalizable** if the matrix $\mathbf{A}$ is diagonalizable. The above definitions do not depend on the choice of the ordered basis E for V because if F is any other ordered basis of V, then $\mathbf{B} := \mathbf{M}_F^F(T)$ is similar to $\mathbf{A}$.

Inner Product Spaces

Let V be a vector space over $\mathbb{K}$. An **inner product** on V is a function $\langle \cdot\,, \cdot \rangle : V \times V \to \mathbb{K}$ satisfying the following properties. For $u, v, w \in V$ and $\alpha, \beta \in \mathbb{K}$,

1. $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0 \iff v = 0$, (positive definite)
2. $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$, (linear in 2nd variable)
3. $\langle v, u \rangle = \overline{\langle u, v \rangle}$. (conjugate symmetric)

From the above properties, conjugate linearity in the 1st variable follows: $\langle \alpha u + \beta v, w \rangle = \overline{\alpha} \langle u, w \rangle + \overline{\beta} \langle v, w \rangle$.

If $u, v \in V$ and $\langle u, v \rangle = 0$, then we say that u and v are **orthogonal**, and we write $u \perp v$.

For $v \in V$, we define the **norm** of v by $\|v\| := \langle v, v \rangle^{1/2}$.

If $v \in V$ and $\|v\| = 1$, then we say that v is a **unit vector** or a **unit function**. The set $\{v \in V : \|v\| \leq 1\}$ is called the **unit ball** of V.

Definition

A vector space V over $\mathbb{K}$ with a prescribed inner product on it is called an **inner product space**.

Examples

1. We have already studied the primary example, namely $V := \mathbb{K}^{n\times 1}$ with the **usual inner product** $\langle \mathbf{x}, \mathbf{y}\rangle := \mathbf{x}^*\mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$. There are other inner products on $\mathbb{K}^{n\times 1}$. For example, let $w_1, \ldots, w_n$ be positive real numbers, and define

$$\langle \mathbf{x}, \mathbf{y}\rangle_w := w_1\overline{x}_1 y_1 + \cdots + w_n \overline{x}_n y_n \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}.$$

Then this is an inner product on $V = \mathbb{K}^{n\times 1}$. On the other hand, the function on $\mathbb{R}^{4\times 1} \times \mathbb{R}^{4\times 1}$ defined by

$$\langle \mathbf{x}, \mathbf{y}\rangle_M := x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4 \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{4\times 1}$$

is not an inner product on $\mathbb{R}^{4\times 1}$. Note that for $\mathbf{x} \in \mathbb{R}^{4\times 1}$, $\langle \mathbf{x}, \mathbf{x}\rangle_M = x_1^2 + x_2^2 + x_3^2 - x_4^2$, and this can be negative.

2. Let $\mathbb{K} = \mathbb{R}$ and let $V := C([a, b])$, the vector space of all continuous real valued functions on $[a, b]$. Define

$$\langle f, g \rangle := \int_a^b f(t)g(t)dt \quad \text{for } f, g \in V.$$

It is easy to check that this is an inner product on V. We shall call this inner product the **usual inner product** on $C([a, b])$.

In this case, the norm of $f \in V$ is $\|f\| := \left(\int_a^b |f(t)|^2 dt \right)^{1/2}$.

This example gives a continuous analogue of the usual inner product on $\mathbb{K}^{n \times 1}$.

There are other inner products on V. For example, let $w : [a, b] \to \mathbb{R}$ be positive function, and define

$$\langle f, g \rangle_w := \int_a^b w(t)f(t)g(t)dt \quad \text{for } f, g \in V.$$

Then this is also an inner product on V.

Projection in the direction of a nonzero vector

Suppose V is any inner product space over $\mathbb{K}$ with a prescribed inner product given by $\langle \cdot\,, \cdot \rangle$.

Let w be a nonzero element of V. As earlier, define

$$P_w(v) := \frac{\langle w, v \rangle}{\langle w, w \rangle} w \quad \text{for } v \in V.$$

It is called the (perpendicular) **projection** of v in the direction of w. It is easy to see that $P_w : V \to V$ is a linear map and its image space is one dimensional. It is also clear from the definition that $P_w(w) = w$. This implies that

$$(P_w)^2 := P_w \circ P_w = P_w.$$

Note that $P_w(v)$ is a scalar multiple of w for every $v \in V$.

Two important properties of the projection of a vector in the direction of another (nonzero) vector are as follows.

Proposition

Let $w \in V$ be nonzero. Then for every $v \in V$,
(i) $\big(v - P_w(v)\big) \perp w$ and (ii) $\|P_w(v)\| \leq \|v\|$.

Proof. Let $v \in V$. For (i), we note that

$$\langle w,\ v - P_w(v)\rangle = \langle w,\ v\rangle - \langle w,\ P_w(v)\rangle = \langle w,\ v\rangle - \frac{\langle w,\ v\rangle}{\langle w,\ w\rangle}\langle w,\ w\rangle = 0.$$

For (ii), write $v = P_w(v) + u$, and note that $\langle u,\ P_w(v)\rangle = 0$ by (i). Hence

$$\|v\|^2 = \langle P_w(v) + u,\ P_w(v) + u\rangle = \|P_w(v)\|^2 + \|u\|^2,$$

Therefore, $\|v\|^2 \geq \|P_w(v)\|^2$, which yields (ii).

The following inequalities were proved earlier for vectors in $\mathbb{K}^{n\times 1}$. They hold in any inner product space.

Theorem

Let $\langle \cdot\,,\,\cdot \rangle$ be an inner product on a vector space V, and let $v, w \in V$. Then
(i) (Schwarz Inequality) $|\langle v,\, w\rangle| \le \|v\|\|w\|$.

(ii) (Triangle Inequality) $\|v + w\| \le \|v\| + \|w\|$.

Proof. (i) Let $w = 0$. Then $\langle v,\, 0\rangle = \langle v,\, 0+0\rangle = 2\langle v,\, 0\rangle$ implies $\langle v,\, w\rangle = 0$. Since $\|w\| = 0$, we obtain (i) if $w = 0$.

Now suppose $w \neq 0$. Then by (ii) of the previous proposition,

$$\left\| \frac{\langle w,\, v\rangle}{\langle w,\, w\rangle} w \right\| = \|P_w(v)\| \le \|v\|,$$

that is, $|\langle w,\, v\rangle|\|w\| \le \|v\|\langle w,\, w\rangle = \|v\|\|w\|^2$. Hence $|\langle v,\, w\rangle| \le \|v\|\|w\|$ in this case as well.

(ii) Since $\langle v, w \rangle + \langle w, v \rangle = 2\,\Re\,\langle v, w \rangle$, we see that

$$\begin{aligned} \|v+w\|^2 &= \langle v+w,\, v+w \rangle = \|v\|^2 + \|w\|^2 + 2\,\Re\,\langle v, w \rangle \\ &\le \|v\|^2 + \|w\|^2 + 2\,|\langle v, w \rangle| \\ &\le \|v\|^2 + \|w\|^2 + 2\,\|v\|\|w\| \quad \text{(by (i) above)} \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

Thus $\|v+w\| \le \|v\| + \|w\|$. □

As a consequence of the above theorem, we see that the norm function $\|\cdot\| : V \to \mathbb{K}$ on an inner product space V satisfies the following three basic properties:

(i) $\|v\| \ge 0$ for all $v \in V$ and $\|v\| = 0 \iff v = 0$,

(ii) $\|\alpha v\| = |\alpha|\|v\|$ for all $\alpha \in \mathbb{K}$ and $v \in V$,

(iii) $\|v+w\| \le \|v\| + \|w\|$ for all $v, w \in V$.

Orthogonal and Orthonormal Sets

Let V be an inner product space. Let E be a subset of V. Define

$$E^{\perp} := \{w \in V : w \perp v \text{ for all } v \in E\}.$$

It is easy to see that $E^{\perp}$ is a subspace of V. We call $E^{\perp}$ the **orthogonal complement** of E in V.

The set E is said to be **orthogonal** if any two (distinct) elements of E are orthogonal (to each other), that is, $v \perp w$ for all v, w in E with $v \neq w$. An orthogonal set whose elements are unit vectors is called an **orthonormal set.**

If E is orthogonal and does not contain 0, then it is easily seen that E is linearly independent. For example, consider $V := C([-\pi, \pi])$ with the usual inner product and let $E := \{\cos nt : n \in \mathbb{N}\} \cup \{\sin nt : n \in \mathbb{N}\}$. Check that E is orthogonal and $0 \notin E$. Hence, E is linearly independent.

Gram-Schmidt Orthogonalization Process

If we are given a sequence of linearly independent elements of V, then we can construct an orthogonal subset of V not containing 0, retaining the span of the elements so constructed at every step by the Gram-Schmidt Orthogonalization Process (G-S OP), just as discussed earlier.

Let (v_n) be a sequence of linearly independent elements in V. Define $w_1 := v_1$, and for $j \in \mathbb{N}$, define

$$\begin{aligned} w_{j+1} &:= v_{j+1} - P_{w_1}(v_{j+1}) - \cdots - P_{w_j}(v_{j+1}) \\ &= v_{j+1} - \frac{\langle w_1,\ v_{j+1}\rangle}{\langle w_1,\ w_1\rangle} w_1 - \cdots - \frac{\langle w_j,\ v_{j+1}\rangle}{\langle w_j,\ w_j\rangle} w_j. \end{aligned}$$

Then $\text{span}\{w_1, \ldots, w_{j+1}\} = \text{span}\{v_1, \ldots, v_{j+1}\}$, and the set $\{w_1, \ldots, w_{j+1}\}$ is orthogonal.

Now let $u_j := w_j / \|w_j\|$ for $j \in \mathbb{N}$, then $(u_1, u_2, \ldots)$ is an ordered orthonormal set such that for each $j \in \mathbb{N}$,

$$\text{span}\{v_1, \ldots, v_j\} = \text{span}\{w_1, \ldots, w_j\} = \text{span}\{u_1, \ldots, u_j\}.$$

Example

Let $V = C([-1, 1])$ with the usual inner product defined by

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

For $j = 0, 1, 2, \ldots$, consider the polynomial function $p_j(t) := t^j$, $t \in [-1, 1]$. Let us orthogonalize the set $\{p_0, p_1, p_2, p_3\}$. Define $q_0 := p_0$, and

$$q_1 := p_1 - \frac{\langle q_0, p_1 \rangle}{\langle q_0, q_0 \rangle} q_0 = p_1 - \left(\frac{1}{2} \int_{-1}^{1} t \, dt \right) p_0 = p_1.$$

Next, define

$$\begin{aligned} q_2 &:= p_2 - \frac{\langle q_0,\, p_2\rangle}{\langle q_0,\, q_0\rangle} q_0 - \frac{\langle q_1,\, p_2\rangle}{\langle q_1,\, q_1\rangle} q_1 \\ &= p_2 - \left(\frac{1}{2}\int_{-1}^{1} t^2 dt\right) q_0 - \left(\frac{3}{2}\int_{-1}^{1} t^3 dt\right) q_1 \\ &= p_2 - \frac{1}{3} p_0, \end{aligned}$$

and similarly,

$$\begin{aligned} q_3 &:= p_3 - \frac{\langle q_0,\, p_3\rangle}{\langle q_0,\, q_0\rangle} q_0 - \frac{\langle q_1,\, p_3\rangle}{\langle q_1,\, q_1\rangle} q_1 - \frac{\langle q_2,\, p_3\rangle}{\langle q_2,\, q_2\rangle} q_2 \\ &= p_3 - \frac{3}{5} p_1. \end{aligned}$$

Observe that $\|q_0\| = \sqrt{2}$, $\|q_1\| = \sqrt{2}/\sqrt{3}$, $\|q_2\| = 2\sqrt{2}/3\sqrt{5}$ and $\|q_3\| = 2\sqrt{2}/5\sqrt{7}$.

Hence if we let

$$u_j = \frac{q_j}{\|q_j\}} \quad \text{for } j = 0, 1, 2, 3,$$

then we obtain the following orthonormal subset of V having the same span as span$\{p_0, p_1, p_2, p_3\}$, namely, the space of all real-valued polynomial functions of degree at most 3:

$$u_0(t) := \frac{\sqrt{2}}{2}, \quad u_1(t) := \frac{\sqrt{6}}{2}\, t,$$
$$u_2(t) := \frac{\sqrt{10}}{4}\,(3t^2 - 1), \quad u_3(t) := \frac{\sqrt{14}}{4}\,(5t^3 - 3t).$$

In a similar manner, we can, in fact, obtain an infinite ordered orthornormal set $(u_0, u_1, \dots)$ of polynomials (in t) by applying G-S OP to $(p_0, p_1, \dots)$ The sequence of orthonormal polynomials thus obtained is known as the sequence of **Legendre polynomials**. It is of much use in many contexts.

Definition

Let V be a finite dimensional inner product space. An **orthonormal basis** *for V is a basis for V which is an orthonormal subset of V.*

We have proved the following results for subspaces of $\mathbb{K}^{n\times 1}$. Their proofs remain valid for any inner product space.

If $u_1, \ldots, u_k$ is an orthonormal set in V, then we can extend it to an orthonormal basis. As a consequence, every nonzero vector subspace V has an orthonormal basis.

The G-S OP enables us to improve the quality of a given basis for V by orthonormalizing it. For instance, if $\{u_1, \ldots, u_n\}$ is an orthonormal basis for V, and $v \in V$, then it is extremely easy to write v as a linear combination of $u_1, \ldots, u_n$; in fact

$$v = \langle u_1, v\rangle u_1 + \cdots + \langle u_n, v\rangle u_n.$$

Orthogonal Projections

Let W be a subspace of a finite dimensional inner product space V. The **Orthogonal Projection Theorem** says that for every $v \in V$, there are unique $w \in W$ and $\tilde{w} \in W^{\perp}$ such that $v = w + \tilde{w}$, that is, $V = W \oplus W^{\perp}$. The map $P_W : V \to V$ given by $P_W(v) = w$ is linear and satisfies $(P_W)^2 = P_W$. It is called the **orthogonal projection map** of V onto the subspace W.

In fact, if $u_1, \ldots, u_k$ is an orthonormal basis for W, then

$$P_W(v) = \langle u_1, v \rangle u_1 + \cdots + \langle u_k, v \rangle u_k \quad \text{for} \quad v \in V.$$

Given $v \in V$, its orthogonal projection $P_W(v)$ is the **unique best approximation to** v **from** W.

Further, $P_W(v)$ is the unique element of W such that $v - P_W(v)$ is orthogonal to W.

Now suppose V is an inner product space of dimension n. Fix an ordered orthonormal basis E of V. For a linear operator $T : V \to V$, let $\mathbf{A} = M_E^E(T)$ denote the matrix of T with respect to E. Then the linear map $T^* : V \to V$ whose matrix with respect to E is A^* is called the **adjoint** of T. The linear operator T^* is independent of the choice of an ordered orthonormal basis E, and it satisfies

$$\langle T(u), v \rangle = \langle u, T^*(v) \rangle \quad \text{for all } u, v \in V.$$

Define T to be **Hermitian** if $T = T^*$, and **skew-Hermitian** if $T = -T^*$. Thus, T is **Hermitian** if

$$\langle T(u), v \rangle = \langle u, T(v) \rangle \quad \text{for all } u, v \in V,$$

and is **skew-Hermitian** if

$$\langle T(u), v \rangle = -\langle u, T(v) \rangle \quad \text{for all } u, v \in V,$$

Define T to be **unitary** if $T \circ T^*$ and $T^* \circ T$ are both identity maps on V. And define T to be **normal** if $T \circ T^* = T^* \circ T$.

Note that T is unitary if and only if

$$\langle T(u),\ T(v)\rangle = \langle u,\ v\rangle \quad \text{for all } u, v \in V,$$

One can prove the spectral theorem for a normal operator on a finite dimensional inner product space V just as before. Moreover, one can also prove spectral theorems for self-adjoint operators on V just as before.

Remark The notion of adjoint of a linear operator can be generalized to the setting of certain infinite dimensional inner product spaces, and one also has analogous spectral theorems for normal operators in this more general set-up. These are quire useful in mathematics and physics, and they may be studied in some advanced courses.

THE END