MA110
Lecture 21

Saurav Bhaumik
Department of Mathematics
IIT Bombay

Spring 2025

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: A linear transformation from a vector space V to itself is often called a linear operator on V.

Examples

$\boxed{1}$. 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.

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

$\boxed{3}$. 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.

$\boxed{4}$. 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 $v_1, \ldots, v_n$ be a basis for V. If $v \in V$, then there are $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ 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 $\mathcal{I}(T) = \text{span}\{T(v_1), \ldots, T(v_n)\}$. 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

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

Proof (Sketch): Let $s := \mathrm{nullity}(T)$ 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 $\mathcal{I}(T)$. $\square$

Corollary

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

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

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

$$\textit{if } m = n, \textit{ then } \quad T \textit{ is one-one} \iff T \textit{ is onto.}$$

Proof. The first assertion follows from the Rank-Nullity Theorem since

$$T \text{ is one-one} \iff \mathcal{N}(T) = \{0\} \iff \operatorname{nullity}(T) = 0.$$

If T is one-one, then $n = \operatorname{rank}(T) = \dim \mathcal{I}(T) \leq \dim W = m$. Further, if $m = n$, then $\operatorname{rank}(T) = n \iff T$ is onto. $\square$

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. The dimension of the vector space $W_1 \times W_2$ is equal to $\dim W_1 + \dim W_2$. 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, and

$$\mathcal{N}(T) = \{(w, w) : w \in W_1 \cap W_2\} \text{ and } \mathcal{I}(T) = W_1 + W_2.$$

Hence by the Rank-Nullity Theorem,

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

To a linear map from a finite dimensional vector space to another finite dimensional space, one can associate a matrix exactly as before.

Let V be a vector space of dimension n, and let $E := (v_1, \ldots, v_n)$ be an ordered basis for V. Also, let W be a vector space of dimension m, and let $F := (w_1, \ldots, w_m)$ be an ordered basis for W. Let $T : V \to W$ be a linear map. Then for each $k = 1, \ldots, n$, we can uniquely write

$$T(v_k) = a_{1k} w_1 + \cdots + a_{mk} w_m = \sum_{j=1}^{m} a_{jk} w_j \quad \text{for some } a_{jk} \in \mathbb{K}.$$

The $m \times n$ matrix $\mathbf{A} := [a_{jk}]$ is called the **matrix of the linear transformation** $T : V \to W$ with respect to the ordered basis $E := (v_1, \ldots, v_n)$ of V and the ordered basis $F := (w_1, \ldots, w_m)$ of W. It is denoted by $\boxed{\mathbf{M}_F^E(T)}$.

Examples 1. Define $T : \mathcal{P}_n \to \mathcal{P}_{n-1}$ by $T(p) = p'$, the derivative of p. Consider the ordered bases $E := (1, t, \ldots, t^n)$ and $F := (1, t, \ldots, t^{n-1})$ of $\mathcal{P}_n$ and $\mathcal{P}_{n-1}$ respectively. Then the $n \times (n+1)$ matrix of the linear map T with respect to these bases is

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

2. Let $T : \mathbb{K}^{2\times 2} \to \mathbb{K}^{2\times 2}$ be the linear transformation defined by $T(A) = A^{\mathsf{T}}$. Then the matrix of T with respect to the basis $E := \{\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}\}$ is $\mathbf{M}_E^E(T) := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Eigenvalues and Eigenvectors of Linear Operators

Definition

Let V be a vector space over $\mathbb{K}$, and let $T : V \to V$ be 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.*

Example: Let V denote the vector space $C^\infty(\mathbb{R})$ 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. Let E be an ordered basis for V, and let $\mathbf{A} := \mathbf{M}_E^E(T)$, the matrix of the linear operator T with respect to E. We remark that if F is another ordered basis for V, and $\mathbf{B} := \mathbf{M}_F^F(T)$, the matrix of the linear operator T with respect to F, then $\mathbf{B}$ is similar to $\mathbf{A}$; in fact we have seen that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, where $\mathbf{P} = \mathbf{M}_F^E(I)$, and where $I : V \to V$ is the identity map.

Definition

The **geometric multiplicity** *of an eigenvalue of T is the dimension of the corresponding eigenspace. It equals the geometric multiplicity of λ as an eigenvalue of the associated matrix* $\mathbf{A}$. *The* **characteristic polynomial** *of T is defined to be the characteristic polynomials of* $\mathbf{A}$. *Further, T is called* **diagonalizable** *if the matrix* $\mathbf{A}$ *is diagonalizable.*

Definition

The **algebraic multiplicity** *of an eigenvalue of the linear operator T is defined to be the algebraic multiplicity of the associated matrix* $\mathbf{A}$.

The relationships between the geometric multiplicity and the algebraic multiplicity of an eigenvalue of a square matrix hold for a linear operator as well.

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 the matrix $\mathbf{B} := \mathbf{M}_F^F(T)$ is similar to the matrix $\mathbf{A} := \mathbf{M}_E^E(T)$ as we have seen earlier.

Results about the linear independence of eigenvectors corresponding to distinct eigenvalues hold in the general case.

Inner Product Spaces

Definition

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: $\langle \alpha u + \beta v,\, w\rangle = \overline{\alpha}\langle u,\, w\rangle + \overline{\beta}\langle v,\, w\rangle$.

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

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.

Examples

$\boxed{1}$. 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 $w_1, \dots, w_n$ be positive real numbers, and define

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

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 $\mathbf{x} \in \mathbb{R}^{4\times 1}$, $\langle \mathbf{x},\ \mathbf{x}\rangle_M = x_1^2 + x_2^2 + x_3^2 - x_4^2$. (This is used in defining the **Minkowski space**)

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

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

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

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

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 (orthogonal) **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. Also, $P_w(w) = w$, so that $(P_w)^2 := P_w \circ P_w = P_w$.

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) $\|P_w(v)\| \le \|v\|$.

The proof of (i) is an easy verification, and (ii) follows from the formula $\|v\|^2 = \|P_w(v)\|^2 + \|v - P_w(v)\|^2$, which is a consequence of (i).

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) $|\langle v,\, w\rangle| \leq \|v\|\|w\|$.

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

Proof. (i) First, suppose $w = 0$. Then for any $v \in V$,
$\langle v,\, w\rangle = \langle v,\, 0\rangle = \langle v,\, 0+0\rangle = 2\langle v,\, 0\rangle$, and so $\langle v,\, w\rangle = 0$.
Also, $\|w\| = 0$. Hence we are done.
Now suppose $w \neq 0$. Then by (ii) of the previous proposition,

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

that is,

$$|\langle w,\, v\rangle| \|w\| \leq \|v\| \langle w,\, w\rangle = \|v\| \|w\|^2.$$

Hence $|\langle v,\, w\rangle| \leq \|v\|\|w\|$.

(ii) Since $\langle v,\, w\rangle + \langle w,\, v\rangle = 2\,\mathfrak{R}\,\langle v,\, w\rangle$, we see that

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

Thus $\|v+w\| \leq \|v\| + \|w\|$. $\square$

We observe that the norm function $\|\cdot\| : V \to \mathbb{K}$ satisfies the following three crucial properties:

(i) $\|v\| \geq 0$ 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) $\|v+w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

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.

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 E is linearly independent. For example, let $V := C[-\pi, \pi]$ and $E := \{\cos nt : n \in \mathbb{N}\} \cup \{\sin nt : n \in \mathbb{N}\}$. Since E is orthogonal and $0 \notin E$, the set E is linearly independent.

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 stepby 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 be the set of all real-valued polynomial functions on $[-1, 1]$ along with the inner product defined by

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

For $j = 0, 1, 2, \ldots$, let $p_j(t) := t^j$, $t \in [-1, 1]$. 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,

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

.

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

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

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

The sequence of orthonormal polynomials thus obtained by orthonormalizing the monomials by the G-S OP is known as the sequence of **Legendre polynomials**. These are of much use in many contexts.

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.

Definition

Suppose V is an inner product space of dimension n. For a linear operator $T : V \to V$, define its **adjoint** $T^* : V \to V$ *by*

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

Define T to be **Hermitian** *or* **self-adjoint** *if $T = T^*$, and* **skew-Hermitian** *or* **skew self-adjoint** *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 T is **skew-Hermitian** if

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

Note that for $\mathbf{A} \in \mathbb{K}^{n\times n}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{K}^{n\times 1}$,

$$\langle \mathbf{A}\,\mathbf{x},\ \mathbf{y}\rangle = (\mathbf{A}\,\mathbf{x})^*\mathbf{y} = \mathbf{x}^*(\mathbf{A}^*\mathbf{y}) = \langle \mathbf{x},\ \mathbf{A}^*\mathbf{y}\rangle.$$

Hence a matrix $\mathbf{A}$ is self-adjoint, that is, $\mathbf{A}^* = \mathbf{A}$ if and only if $\langle \mathbf{A}\,\mathbf{x},\ \mathbf{y}\rangle = \langle \mathbf{x},\ \mathbf{A}\mathbf{y}\rangle$. The following result, therefore, is natural.

Proposition

Let V be a finite dimensional inner product space, and let $T : V \to V$ be a linear operator. Then T is Hermitian if and only if the matrix of T with respect to any ordered orthonormal basis of V is self-adjoint.

An operator T which commutes with its adjoint T^* will be called **normal operator** on V. One can prove the spectral theorem for a normal operator on a finite dimensional inner product space V just as before.

THE END