5.A Invariant Subspaces

Definition: *invariant subspace*

Suppose $T \in \mathcal{L}(V)$. A subspace U of V is called ***invariant*** under T if $u \in U$ implies $Tu \in U$.

Definition: *eigenvalue*

Suppose $T \in \mathcal{L}(V)$. A number $\lambda \in \mathbb{F}$ is called an ***eigenvalue*** of T if there exists $v \in V$ such that $v \neq 0$ and $Tv = \lambda v$.

Equivalent conditions to be an eigenvalue

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. Then TFAE:

(a) λ is an eigenvalue of T.

(b) $T - \lambda I$ is not injective.

(c) $T - \lambda I$ is not surjective.

(d) $T - \lambda I$ is not invertible.

Definition: *eigenvector*

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbb{F}$ is an eigenvalue of T. A vector $v \in V$ is called an ***eigenvector*** of T corresponding to λ if $v \neq 0$ and $Tv = \lambda v$.

Linearly independent eigenvectors

Let $T \in \mathcal{L}(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding eigenvectors. Then $v_1, \ldots, v_m$ is linearly independent.

Number of eigenvalues

Suppose V is finite-dimensional. Then each operator on V has at most $\dim V$ distinct eigenvalues.

Definition: $T|_U$ *and* T/U

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T.

- The ***restriction operation*** $T|_U \in \mathcal{L}(U)$ is defined by

$$T|_U(u) = Tu \ \forall\, u \in U.$$

- The ***quotient operator*** $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U \ \forall\, v \in V.$$

5.B Eigenvectors and Upper-Triangular Matrices

Definition: T^m
Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

- T^m is defined by
$$T^m = \underbrace{T \cdots T}_{m \text{ times}}.$$
- T^0 is defined to be the identity operator I on V.
- If T is invertible with inverse T^{-1}, then T^{-m} is defined by
$$T^{-m} = (T^{-1})^m.$$

Definition: $p(T)$
Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$ is a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for $z \in \mathbb{F}$. Then $p(T)$ is the operator defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

Definition: ***product of polynomials***
If $p, q \in \mathcal{P}(\mathbb{F})$, then $pq \in \mathcal{P}(\mathbb{F})$ is the polynomial defined by

$$(pq)(z) = p(z)q(z)$$

for $z \in \mathbb{F}$.

6.A Inner Products and Norms

Definition: ***norm***
The length of a vector x is called the **norm** of x denoted $\|x\|$ The norm is non-linear.

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Definition: ***dot product***
For $x, y \in \mathbb{R}^n$, the ***dot product*** of x and y, denoted $x \cdot y$ (or $x.y$) is defined by

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n,$$

where $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

Note that the dot product of two vectors in $\mathbb{R}^n$ is a scalar (number (i.e. 1, 7, 5.3, π)), not another vector.

$$x \cdot x = \|x\|^2 \;\; \forall \; x \in \mathbb{R}^n$$

dot product properties For a dot product on $\mathbb{R}^n$,

(a) $x \cdot x \geq 0 \; \forall \; x \in \mathbb{R}^n$

(b) $x \cdot x = 0 \iff x = 0$

(c) For a fixed $y \in \mathbb{R}^n$, the map from $\mathbb{R}^n \to \mathbb{R}$ that sends $x \in \mathbb{R}^n$ to $x \cdot y$ is linear

(d) $x \cdot y = y \cdot x \; \forall \; x, y \in \mathbb{R}^n$ (commutative)

Definition: $\operatorname{Re} z, \; \operatorname{Im} z$
Suppose $z = a + bi$ where $a, b \in \mathbb{R}$.

- The ***real part*** of z, denoted $\operatorname{Re} z$, is defined by $\operatorname{Re} z = a$.
- The ***imaginary part*** of z, denoted $\operatorname{Im} z$, is defined by $\operatorname{Im} z = b$.

Thus for every $z \in \mathbb{C}$, we have

$$z = \operatorname{Re} z + (\operatorname{Im} z)i.$$

Definition: ***complex conjugate***
Suppose $z \in \mathbb{C}$. The ***complex conjugate*** of $z \in \mathbb{C}$, denoted $\bar{z}$, is defined by

$$\bar{z} = \operatorname{Re} z - (\operatorname{Im} z)i.$$

Definition: ***absolute value***
Suppose $z \in \mathbb{C}$. The ***absolute value*** of $z \in \mathbb{C}$, denoted $|z|$, is defined by

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}.$$

Complex properties

- sum of z and $\bar{z}$
$$z + \bar{z} = 2 \operatorname{Re} z$$
- difference of z and $\bar{z}$
$$z - \bar{z} = 2(\operatorname{Im} z)i$$
- product of z and $\bar{z}$
$$z\bar{z} = |z|^2$$

- additivity and multiplicativity of complex conjugate

$$\overline{w+z} = \bar{w} + \bar{z} \qquad \text{and} \qquad \overline{wz} = \bar{w}\bar{z}$$

- conjugate of a conjugate

$$\bar{\bar{z}} = z$$

- real and imaginary parts are bounded by $|z|$

$$|\text{Re}\, z| \leq |z| \qquad \text{and} \qquad |\text{Im}\, z| \leq |z|$$

- absolute value of a complex conjugate

$$|\bar{z}| = \bar{z}$$

- multiplicativity of absolute value

$$|wz| = |w|\,|z|$$

- Triangle Inequality

$$|w+z| \leq |w| + |z|$$

Recall that for $\lambda = a + bi$ where $a, b \in \mathbb{R}$, then

- The absolute value of λ, denoted $|\lambda|$, is defined as

$$|\lambda| = \sqrt{a^2 + b^2}.$$

- The complex conjugate of λ, denoted $\bar{\lambda}$, is defined as

$$\bar{\lambda} = a - bi.$$

- The absolute value squared is defined as

$$|\lambda|^2 = \lambda\bar{\lambda}$$

For $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, we define the norm of z by

$$\|z\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

These absolute values are needed because we want $\|z\|$ to be a nonnegative scalar. Note that

$$\|z\|^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

We want to think of $\|z\|^2$ as an inner product of z with itself, like we did in $\mathbb{R}^n$. The equation above suggests that for $w = (w_1, \ldots, w_n) \in \mathbb{C}^n$, the inner product of w with z should be

$$\langle w, z \rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$$

$$\langle z, w\rangle = z_1\overline{w_1} + \cdots + z_n\overline{w_n}$$

If $\lambda \in \mathbb{C}$ then the notation $\lambda \geq 0$ means that λ is a non-negative real number.

We use the common notation $\langle u, v\rangle$ with angle brackets tod enote an inner product. Parenthesis can also be used, but then (u, v) becomes confusing.

Definition: ***inner product***
An ***inner product*** on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbb{F}$ and has the following properties:

positivity:
$\langle v, v\rangle \geq 0 \ \forall \ v \in V$

definiteness:
$\langle v, v\rangle = 0 \iff v = 0$

additivity in the first slot:
$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \ \ \forall \ u, v, w \in V$

homogeneity in the first slot:
$\langle \lambda u, v\rangle = \lambda \langle u, v\rangle \ \ \forall \ \lambda \in \mathbb{F}$ and $u, v \in V$

conjugate symmetry:
$\langle u, v\rangle = \overline{\langle v, u\rangle} \ \forall \ u, v \in V$

Note: $x = \overline{x} \ \forall \ x \in \mathbb{R}$

Examples:

A ***Euclidean inner product*** on $\mathbb{F}^n$ is defined by

$$\langle (w_1, \ldots, w_n),\ (z_1, \ldots, z_n)\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

If $c_1, \ldots, c_n$ are all positive numbers, then an inner product can be defined on $\mathbb{F}^n$ by

$$\langle (w_1, \ldots, w_n),\ (z_1, \ldots, z_n)\rangle = c_1w_1\overline{z_1} + \cdots + c_nw_n\overline{z_n}.$$

An inner product can be defined on the vector space of continuous real-valued functions on the interval $[-1,\ 1]$ by

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx.$$

An inner product can be defined on $\mathcal{P}(\mathbb{R})$ by

$$\langle p, q\rangle = \int_{0}^{\infty} p(x)q(x)e^{-x}\,dx$$

Definition: ***inner product space***
An ***inner product space*** is a vector space V along with an inner product on V.

Unless otherwise stated, you can assume the inner product on $\mathbb{F}^n$ is defined as

$$\langle w, z\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$$

Notation: For the rest of this chapter, V denotes an inner product space over $\mathbb{F}$.

Basic properties of an inner product

(a) For each fixed $u \in V$, the function that takes v to $\langle v, u\rangle$ is a linear map from V to $\mathbb{F}$.

(b) $\langle 0, u\rangle = 0 \;\; \forall \;\; u \in V$

(c) $\langle u, 0\rangle = 0 \;\; \forall \;\; u \in V$

(d) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle \;\; \forall \;\; u, v, w \in V$

(e) $\langle u, \lambda v\rangle = \bar{\lambda}\,\langle u, v\rangle \;\; \forall \;\; \lambda \in \mathbb{F}$ and $u, v \in V$

(f) $\langle u, v\rangle = 0 \iff \langle v, u\rangle = 0 \;\; \forall \;\; u, v \in V$

Definition: *norm*
For $v \in V$, the ***norm*** of v, denoted $\|v\|$, is defined by

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

In other words, $\|v\|^2 = \langle v, v\rangle$.

Basic properties of the norm

Suppose $v \in V$. Then

(a) $\|v\| = 0 \iff v = 0.$

(b) $\|\lambda v\| = |\lambda|\,\|v\| \;\; \forall \;\; \lambda \in \mathbb{F}.$

Examples:

(a) If $(z_1, \ldots, z_n) \in \mathbb{F}^n$ (with the usual Euclidean inner product), then

$$\|(z_1, \ldots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

(b) In the vector space of continuous real-valued functions on $[1, -1]$ (with the inner product defined as $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$), then

$$\|f\| = \sqrt{\int_{-1}^{1} \big(f(x)\big)^2\,dx}.$$

Definition: ***orthogonal***
Two vectors $u, v \in V$ are called ***orthogonal*** if $\langle u, v\rangle = 0$.

Instead of saying "u and v are orthogonal", we sometimes say "u is orthogonal to v".

For two vectors $u, v \in \mathbb{R}^2$, then

$$\langle u, v\rangle = \|u\|\,\|v\|\cos\theta,$$

where θ is the angle between the two vectors. If $\theta = 90^\circ = \frac{\pi}{2}$ radians then $\cos\theta = 0$. That is, if u and v are perpendicular the inner product is 0.

Orthogonality and 0

(a) 0 is orthogonal to every vector in V.

(b) 0 is the only vector in V that is orthogonal to itself.

Pythagorean Theorem

If u and v are orthogonal vectors in V. Then

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2$$

The converse is true if V is a real inner product space. That is, if $\|u+v\|^2 = \|u\|^2 + \|v\|^2$ then u and v are orthogonal.

An orthogonal decomposition

Suppose $u, v \in V$, with $v \neq 0$. Set $c = \dfrac{\langle u, v\rangle}{\|v\|^2}$ and $w = u - \dfrac{\langle u, v\rangle}{\|v\|^2}v$. Then

$$\langle w, v\rangle = 0 \qquad \text{and} \qquad u = cv + w.$$

Cauchy-Schwarz Inequality

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|\,.$$

Further, this is an *equality* if and only if u and v are scalar multiples of one another.

Triangle Inequality

Suppose $u, v \in V$. Then

$$\|u+v\| \leq \|u\| + \|v\|\,.$$

This is an equality if and only if u, v are a nonnegative multiples of the other.

Parallelogram Equality

Suppose $u, v \in V$. Then

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right).$$

6.B Orthonormal Bases

Definition: ***orthonormal***

- A lost of vectors is valled ***orthonormal*** if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.
- In other words, a list $e_1, \ldots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq j. \end{cases}$$

Example, the standard basis in $\mathbb{F}^n$.

The norm of an orthonormal linear combination

If $e_1, \ldots, e_m$ is an orthonormal list of vectors in V, then

$$\|a_1 e_e + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \ldots, a_m \in \mathbb{F}$.

Corollary: An orthonormal list is linearly independent

Every orthonormal list of vectors is linearly independent.

Definition: ***orthonormal basis***

An ***orthonormal basis*** of V is an orthonormal list of vectors in V that is also a basis of V.

For example, the standard basis is an orthonormal basis of $\mathbb{F}^n$.

An orthonormal list of the right length is an orthonormal basis

Every orthonormal list of vectors in V with length $\dim V$ is an orthonormal basis of V.

Writing a vector as linear combination of orthonormal basis

Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $v \in V$. Then

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 .$$

Gram-Schmidt Procedure

Suppose $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = \dfrac{v_1}{\|v_1\|}$.

For $j = 2, \ldots, m$, define e_j inductively by

$$e_j = \frac{v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}}{\|v_j - \langle v_j, e_1 \rangle e_1 - \cdots - \langle v_j, e_{j-1} \rangle e_{j-1}\|}.$$

Then $e_1, \ldots, e_m$ is an orthonormal list of vectors in V such that

$$\text{span}(v_1, \ldots, v_j) = \text{span}(e_1, \ldots, e_j)$$

for $j = 1, \ldots, m$.

Existance of orthonormal basis
Every finite-dimensional inner product space has an orthonormal basis.

Corollary: Orthonormal list extends to orthonormal basis
Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Upper-triangular matrix with respect to orthonormal basis
Suppose $T \in \mathcal{L}(V)$. If T has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Schur's Theorem
Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Because linear maps into a scalar field play a special role, we give them a special name.

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

Riesz Representation Theorem
Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $u \in V$ such that

$$\varphi(v) = \langle v, u \rangle \ \ \forall \ v \in V,$$

where

$$u = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n.$$

6.C Orthogonal Complements and Minimization Problems

Definition: ***orthogonal complement,*** $U^\perp$
If U is a subset of V, then the ***orthogonal complement*** of U, denoted $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 \text{ for every } u \in U\}.$$

For example, if U is a line in $\mathbb{R}^3$ then $U^\perp$ is the plain containing 0 that is perpendicular to U. If U was a plane in $\mathbb{R}^3$ then $U^\perp$ is a line passing through 0 that is perpendicular to U.

Basic properties of orthogonal complement

(a) If U is a subset of V, then $U^\perp$ is a subspace of V.

(b) $\{0\}^\perp = V$.

(c) $V^\perp = \{0\}$.

(d) If U is a subset of V, then $U \cap U^\perp \subseteq \{0\}$.

(e) If U and W are subsets of V and $U \subseteq W$, then $W^\perp \subseteq U^\perp$.

Direct sum of a subspace and its orthogonal complement

Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^\perp$$

Dimension of the orthogonal complement

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^\perp = \dim V - \dim U.$$

The orthogonal complement of the orthogonal complement

Suppose U is a finite-dimensional subspace of V. Then

$$U = \left(U^\perp\right)^\perp.$$

Definition: *orthogonal projection,* P_U

Suppose U is a finite-dimensional subspace of V. The ***orthogonal projection*** of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows:

For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then $P_U v = u$.

Properties of the orthogonal projection P_U

Suppose U is a finite-dimensional subspace of V and $v \in V$. Then

(a) $P_U \in \mathcal{L}(V)$

(b) $P_U u = u \ \forall \ u \in U$

(c) $P_U w = 0 \ \forall \ w \in U^\perp$

(d) $\operatorname{range} P_U = U$

(e) $\operatorname{null} P_U = U^\perp$

(f) $v - P_U v \in U^\perp$

(g) ${P_U}^2 = P_U$

(h) $\|P_U v\| \leq \|v\|$

(i) for every orthonormal basis $e_1, \ldots, e_m$ of U,

$$P_U v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m.$$

Minimizing the distance to a subspace

Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \leq \|v - u\|.$$

Furthermore, the inequality is an equality if and only if $u = P_U v$.

7.A Self-Adjoint and Normal Operators

Definition: ***adjoint***, T^*

Suppose $T \in \mathcal{L}(V, W)$. The ***adjoint*** of T is the function $T^* : W \to V$ such that

$$\langle Tv, w \rangle = \langle v, T^* w \rangle$$

for every $v \in V$ and every $w \in W$.

The adjoint is a linear map

If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.

Properties of the adjoint

(a) $(S + T)^* = S^* + T^*$ for all $S, T \in \mathcal{L}(V, W)$

(b) $(\lambda T)^* = \bar{\lambda} T^*$ for all $\lambda \in F$ and $T \in \mathcal{L}(V, W)$

(c) $(T^*)^* = T$ for all $T \in \mathcal{L}(V, W)$

(d) $I^* = I$, where I is the identity operator on V

(e) $(ST)^* = T^* S^*$ for all $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$ (here U is an inner product space over $\mathbb{F}$).

Null space and range of T^*

(a) $\operatorname{null} T^* = (\operatorname{range} T)^\perp$

(b) $\operatorname{range} T^* = (\operatorname{null} T)^\perp$

(c) $\operatorname{null} T = (\operatorname{range} T^*)^\perp$

(d) $\operatorname{range} T = (\operatorname{null} T^*)^\perp$

Definition: ***conjugate transpose***
The ***conjugate transpose*** of an m-by-n matrix is the n-by-m matrix obtained by transposing the matrix and taking the complex conjugate of each entry.

The matrix of T^*
Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Then

$$\mathcal{M}\big(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n)\big)$$

is the conjugate transpose of

$$\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m)\big)$$

Definition: ***self-adjoint***
An operator $T \in \mathcal{L}(V)$ is called ***self-adjoint*** if $T = T^*$. In other words, $T \in \mathcal{L}(V)$ is self-adjoint if and only if

$$\langle Tv, w\rangle = \langle v, Tw\rangle$$

for all $v, w \in V$.

Eigenvalues of self-adjoint operators are real
Every eigenvalue of a self-adjoint operator is real

Over $\mathbb{C}$, Tv is orthogonal to v for all v only for the 0 operator
Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Suppose

$$\langle Tv, v\rangle = 0$$

for all $v \in V$. Then $T = 0$.

Over $\mathbb{C}$, $\langle Tv, v\rangle$ is real for all v only for self-adjoint operators
Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then T is self-adjoint if and only if

$$\langle Tv, v\rangle \in \mathbb{R}$$

for every $v \in V$.

If $T = T^*$ and $\langle Tv, v\rangle = 0$ for all v, then $T = 0$
Suppose T is a self-adjoint operator on V such that

$$\langle Tv, v\rangle = 0$$

for all $v \in V$. Then $T = 0$.

Definition: ***normal***
An operator on an inner product space is called ***normal*** if it commutes with its adjoint. In other words, $T \in \mathcal{L}(V)$ is normal if

$$TT^* = T^*T.$$

T is normal if and only if $\|Tv\| = \|T^*v\|$ for all v

An operator $T \in \mathcal{L}(V)$ is normal if and only if

$$\|Tv\| = \|T^*v\|$$

for all $v \in V$.

For T normal, T and T^* have the same eigenvectors

Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. Then v is also an eigenvector of T^* with eigenvalue $\bar{\lambda}$.

Orthogonal eigenvectors for normal operators

Suppose $T \in \mathcal{L}(V)$ is normal. Then the eigenvectors of T corresponding to distinct eigenvalues are orthogonal.