Math 24: Winter 2021
Lecture 22

Dana P. Williams

Dartmouth College

Friday, February 26, 2021

Let's Get Started

1. We should be recording.
2. Remember, it is more comfortable for me if you turn on your video so that I feel like I am talking to real people.
3. But first, are there any questions from last time?

Review

Definition

Let T be a linear operator on a vector space V. A subspace W of V is called a T-invariant subspace, or just invariant, if $v \in W$ implies $T(v) \in W$. Alternatively, $T(W) \subset W$.

Definition

Suppose $T \in \mathcal{L}(V)$ and $v \in V$. Then

$$W = \operatorname{Span}\{\, v, T(v), T^2(v), \dots \}$$

is called the T-cyclic subspace of V generated by v.

Remark

You showed W is the smallest T-invariant subspace containing v.

Review

Definition

If $T \in \mathcal{L}(V)$ and if W is a T-invariant subspace of V, then the operator $T_W : W \to W$ in $\mathcal{L}(W)$ is called the restriction of T to W.

Theorem

Suppose that $T \in \mathcal{L}(V)$ for a finite-dimensional vector space V and that W is a T-invariant subspace of V. Then the characteristic polynomial $p_W(\lambda)$ of T_W divides the characteristic polynomial $p(\lambda)$ of T.

Review

Theorem

Suppose that $T \in \mathcal{L}(V)$ for a finite-dimensional vector space V. Suppose that $v \in V$ is nonzero and let $W = \operatorname{Span}\{\, v, T(v), \dots \}$ be the T-cyclic subspace generated by v. Let $k = \dim(W)$.

1. $\beta = \{\, v, T(v), \dots, T^{k-1}(v) \,\}$ *is an ordered basis for W.*
2. *If*
$$a_0 v + a_1 T(v) + \cdots a_{k-1} T^{k-1}(v) + T^k(v) = 0_V,$$
then the characteristic polynomial of T_W is
$$p_W(\lambda) = (-1)^k \big(a_0 + a_1\lambda + \cdots a_{k-1}\lambda^{k-1} + \lambda^k.\big)$$

Review

Theorem (Cayley-Hamilton Theorem)

Suppose that V is a finite-dimensional vector space and that $T \in \mathcal{L}(V)$. If $p(\lambda)$ is the characteristic polynomial of T, then $p(T) = T_0$.

Remark

I the language of the text, we way that T "satisfies" its characteristic polynomial.

Corollary

If $A \in M_{n\times n}(\mathbf{F})$, and $p(\lambda)$ is the characteristic polynomial of A, then $p(A) = O$ where O is the zero matrix.

Motivation

Remark

I want to pause for a brief interlude to go back to the days of bliss where vectors were just directed line segments in $\mathbf{R}^2$. This material isn't really part of the course, but I wanted to use is as motivation for what is to come in Chapter 6.

Remark

If $u = (u_1, u_2) \in \mathbf{R}^2$, the we think of u as the directed line segment from $(0, 0)$ to (u_1, u_2) in the plane. Its length is $\|u\| = \sqrt{u_1^2 + u_2^2}$. Then vector addition and subtraction obey the "parallelogram law". (A picture is good here.)

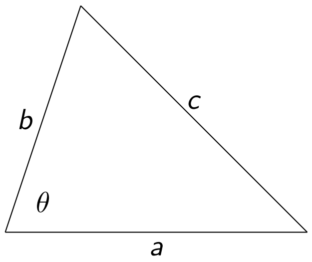
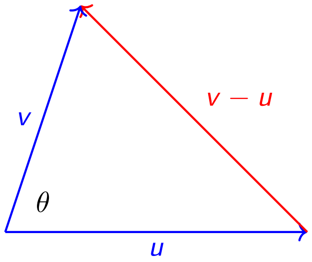


Figure: The Law of Cosines: $c^2 = a^2 + b^2 - 2ab\cos(\theta)$

We will also want to “recall” the law of cosines from high school geometry. It is a generalization of the Pythagorean Theorem which can be proved with the Pythagorean Theorem and some trigonometry.

Dot Products



Now we let u and v be vectors and θ the angle between them. Then the law of cosines gives us

$$\|v - u\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos(\theta).$$

Angles

Thus if $u = (u_1, u_2)$ and $v = (v_1, v_2)$, then

$$(v_1 - u_1)^2 + (v_2 - u_1)^2 = u_1^2 + u_2^2 + v_1^2 + v_2^2 - 2\|u\|\|v\|\cos(\theta).$$

A little algebra shows that $-2(u_1v_1 + u_2v_2) = -2\|u\|\|v\|\cos(\theta)$. In Math 8 or our physics classes, we call $u \cdot v = u_1v_1 + u_2v_2$ the dot product of u and v and brag about the formula

$$\cos(\theta) = \frac{u \cdot v}{\|u\|\|v\|}.$$

In paricular, u and v are perpendicular (shortly we will say orthogonal) if $u \cdot v = 0$.

Remark

The cool thing is that the innocuous dot product allows us to see the geometry algebraically. We are going to want to do something similar in general vector spaces over $\mathbf{R}$ or $\mathbf{C}$.

Remark

At this point, we are going to start making more use of the field $\mathbf{C}$ of complex numbers. The wise student will be re-reading, ok reading, Appendix D to refresh their memory. Today, let's just recall that $\mathbf{C} = \{\, x + iy : x, y \in \mathbf{R} \,\}$ with the operations $(x + iy) + (x' + iy') = (x + x') + i(y + y')$ and $(x + iy)(x' + iy') = (xx' - yy') + i(xy' + x'y)$. On cool tool is the complex conjugate. If $z = x + iy$, the the complex conjugate of z is $\overline{z} = x - iy$. We have $z\overline{z} = |z|^2 = x^2 + y^2$. This is handy when doing complex arithmetic:

$$\frac{1+2i}{3+4i} = \frac{1+2i}{3+4i} \cdot \frac{3-4i}{3-4i} = \frac{3+8+i(6-4)}{9+16} = \frac{1}{25}(11+2i).$$

Break Time

Time for a break and some questions

Definition

Suppose that V is a vector space over $\mathbf{F}$ where $\mathbf{F}$ is either $\mathbf{R}$ or $\mathbf{C}$. Then an inner product on V is a function $\langle \cdot \,, \cdot \rangle : V \times V \to \mathbf{F}$ such that for all $x, y, z \in V$ and all $c \in \mathbf{F}$ we have

1. $\langle cx + y \,, z \rangle = c\langle x \,, z \rangle + \langle y \,, z \rangle$,
2. $\langle y \,, x \rangle = \overline{\langle x \,, y \rangle}$ where the bar is complex conjugation, and
3. $\langle x \,, x \rangle > 0$ if $x \neq 0$.

Remark

Item (1), says that for fixed $z \in V$, $x \mapsto \langle x \,, z \rangle$ is linear from V to $\mathbf{F}$. If $\mathbf{F} = \mathbf{R}$, then the complex conjugate in item (2) goes away and we just have $\langle x \,, y \rangle = \langle y \,, x \rangle$.

Remark

Since we have only defined inner products over $\mathbf{R}$ and $\mathbf{C}$, when working with inner product, $\mathbf{F}$ will always be either $\mathbf{R}$ or $\mathbf{C}$. Then the understanding will be that, as in item (2) above, the complex conjugate plays no role when $\mathbf{F} = \mathbf{R}$.

The Example

Example (Standard Inner Product)

If $V = \mathbf{F}^n$, then the standard inner product on $\mathbf{F}^n$ is given by

$$\langle x \, , \, y \rangle = \sum_{k=1}^{n} x_k \overline{y_k}.$$

I will leave the verification that the standard inner product is an inner product to you. Axioms (1) and (2) are fairly routine. For item (3), note that

$$\langle x \, , \, x \rangle = \sum_{k=1}^{n} x_k \overline{x_k} = \sum_{k=1}^{n} |x_k|^2.$$

Of course, in the real case, $|x_k|^2$ is just x_k^2.

The Dot Product Again

Remark

If $V = \mathbf{R}^2$, then the standard inner product on $\mathbf{R}^2$ is just what we called the dot product back in the day:
$\langle x\,,\,y\rangle = x \cdot y = x_1y_1 + x_2y_2$.

An Infinite-Dimensional Example

Example

Let $V = C([0,1])$ be the real vector-space of real-valued continuous functions on $[0,1]$. Then if $f, g \in V$, we can define

$$\langle f \ , g \rangle = \int_0^1 f(x)g(x)\,dx.$$

Checking axioms (1) and (2) is pretty straightforward. For axiom (3), note that

$$\langle f \ , g \rangle = \int_0^1 |f(x)|^2\,dx > 0$$

if f is not the zero function.

Complex Matrices

Remark

When working with real matrices, the transpose is a key tool. For complex matrices, we often need to pay attention to the complex structure which leads to the following definition.

Definition

If $A \in M_{m\times n}(\mathbf{F})$, then the conjugate transpose of A is the matrix $A^* \in M_{n\times n}(\mathbf{F})$ where $(A^*)_{ij} = \overline{A_{ji}}$.

Example

Let $A = \left(\begin{smallmatrix} 1 & i & 1-i \\ 2+3i & 0 & 2i \end{smallmatrix}\right)$ then $A^* = \begin{pmatrix} 1 & 2-3i \\ -i & 0 \\ 1+i & -2i \end{pmatrix}$. Of course, if $\mathbf{F} = \mathbf{R}$ or all the entries of A are real, then $A^* = A^t$.

The Standard Inner Product

Remark

If $V = \mathbf{F}^n$, then the standard inner product satisfies $\langle x\,,\,y\rangle = y^* x$ where as usual we view $x \in \mathbf{F}^n$ as a $1 \times n$-matrix if we feel like it. Then we can go further as the next example shows.

Example

Let $V = M_{n\times n}(\mathbf{F})$. Then we can define $\langle A\,,\,B\rangle = \operatorname{tr}(B^*A)$ where $\operatorname{tr}(A) = \sum_{k=1}^n A_{kk}$. To see that this is an inner product on $M_{n\times n}(\mathbf{F})$, start with

$$\begin{aligned}\langle A + rB\,,\,C\rangle &= \operatorname{tr}(C^*(A + rB)) = \operatorname{tr}(C^*A + rC^*B)\\ &= \operatorname{tr}(C^*A) + k\operatorname{tr}(C^*B) = \langle A\,,\,C\rangle + k\langle B\,,\,C\rangle.\end{aligned}$$

Thus axiom (1) holds. Axiom (2) is not so hard to check, and

$$\begin{aligned}\langle A\,,\,A\rangle = \operatorname{tr}(A^*A) = \sum_{j=1}^n (A^*A)_{jj} &= \sum_{j=1}^n\sum_{k=1}^n (A^*)_{jk}A_{kj}\\ &= \sum_{j=1}^n\sum_{k=1}^n |A_{kj}|^2\end{aligned}$$

which is > 0 if $A \neq O$.

Inner Product Spaces

Definition

A vector space V endowed with a given inner product is called and inner product space.

Example

For example, $M_{n\times n}(\mathbf{F})$ endowed with the inner product from the previous slide is called the Frobenius inner product space and we will assume that $M_{n\times n}(\mathbf{F})$ is this inner product space unless otherwise noted. Similarly, unless stated otherwise, we will always endow $\mathbf{F}^n$ with the standard inner product.

Theorem

Suppose that V is an inner product space. Then for all $x, y, z \in V$ and $c \in \mathbf{F}$ we have

1. $\langle x\,,\, cy + z\rangle = \overline{c}\langle x\,,\, y\rangle + \langle x\,,\, z\rangle$,
2. $\langle x\,,\, 0_V\rangle = 0 = \langle 0_V\,,\, x\rangle$,
3. $\langle x\,,\, x\rangle = 0$ *if and only if* $x = 0_V$, *and*
4. *if* $\langle x\,,\, y\rangle = \langle x\,,\, z\rangle$ *for all* $x \in V$, *then* $y = z$.

Remark

Item (1) says that for fixed x, $y \mapsto \langle x\,,\, y\rangle$ is conjugate linear. Of course, if $\mathbf{F} = \mathbf{R}$, then $y \mapsto \langle x\,,\, y\rangle$ is actually linear. I will leave the proof of this theorem for homework.

Length

Definition

If V is an inner product space and $x \in V$, then we call $\|x\| = \sqrt{\langle x\,,\,x \rangle}$ the norm (or length) of x.

Theorem

Let V be an inner product space. Then for all $x, y \in V$ and $c \in \mathbf{F}$, we have

1. $\|cx\| = |c|\|x\|$,
2. $\|x\| \geq 0$ *and* $\|x\| = 0$ *if and only if* $x = 0_V$,
3. *(Cauchy-Schwarz Inequality)* $|\langle x\,,\,y \rangle| \leq \|x\|\|y\|$, *and*
4. *(Triangle Inequality)* $\|x + y\| \leq \|x\| + \|y\|$.

Proof.

Items (1) and (2) are straightforward and I'll leave those to you.

Proof

Proof Continued.

(3) If $y = 0_V$, the inequality is immediate. So we assume $y \neq 0_V$. Then for all $c \in \mathbf{F}$,

$$\begin{aligned} 0 &\le \langle x - cy \,,\, x - cy \rangle = \langle x \,,\, x - cy \rangle + c\langle y \,,\, x - cy \rangle \\ &= \langle x \,,\, x \rangle - \overline{c}\langle x \,,\, y \rangle + c\langle y \,,\, x \rangle - c\overline{c}\langle y \,,\, y \rangle. \qquad (\ddagger) \end{aligned}$$

Now let $c = \frac{\langle x,y \rangle}{\langle y,y \rangle}$. Then $\overline{c}\langle x \,,\, y \rangle$, $c\langle y \,,\, x \rangle$, and $c\overline{c}\langle y \,,\, y \rangle$ are all equal to

$$\frac{\langle x \,,\, y \rangle \langle y \,,\, x \rangle}{\langle y \,,\, y \rangle} = \frac{|\langle x \,,\, y \rangle|^2}{\|y\|^2}.$$

Plugging into $(\ddagger)$ gives us

$$0 \le \|x\|^2 - \frac{|\langle x \,,\, y \rangle|^2}{\|y\|^2}.$$

And item (3) follows after multiplying through by $\|y\|^2$ and taking square roots.

Proof

Proof Continued.

(4) For the triangle inequality, we use item (3):

$$\|x+y\|^2 = \langle x+y \, , x+y\rangle = \langle x \, , x\rangle + \langle x \, , y\rangle + \langle y \, , x\rangle + \langle y \, , y\rangle$$

$$\begin{aligned}
&= \|x\|^2 + 2\operatorname{Re}\bigl(\langle x \, , y\rangle\bigr) + \|y\|^2 \\
&\le \|x\|^2 + 2|\langle x \, , y\rangle| + \|y\|^2 \\
&\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= \bigl(\|x\| + \|y\|\bigr)^2
\end{aligned}$$

Now take square roots. $\square$

Orthogonality

Definition

Let V be an inner product space. They $x, y \in V$ are orthogonal (or perpendicular) if $\langle x \,, y\rangle = 0$. A subset $S \subset V$ is called orthogonal if any two distinct vectors in S are orthogonal. A vector $x \in V$ is called unit vector if $\|x\| = 1$. A subset $S \subset V$ is called orthonormal if it is an orthogonal set consisting of unit vectors.

Enough

1. That is enough for today.