Ch3 (contd) 6/9/08

~~note: $(v_1, \dots, v_n)$ is a basis for V.~~

~~Chapter 3: Linear maps~~

Here, V, W are vector spaces.

$T: V \to W$ is linear if:

linear map/transformation/operator

1) $T(v_1 + v_2) = T(v_1) + T(v_2) \ \forall v_1, v_2 \in V$

2) $T(cv) = cT(v) \ \forall c \in \mathbb{F}, v \in V$.

note: (1) $\Rightarrow T(0) = 0$

$\mathcal{L}(V,W) = \{T: V \to W \text{ linear}\}$

$\mathcal{L}(V) = \mathcal{L}(V,V)$

Ex1

Zero map: $0(x) = 0 \ \forall x$

Ex2

Identity map: $I_V: V \to V : x \to x$

Ex3

Inclusion map: $i_U: U \to V$ if U subspace of V, $i_U x = x \ \forall x \in U$.

Ex4

$T\vec{x} = A\vec{x} \ \forall \vec{x} \in \mathbb{F}^n$ where $A \in M_{m \times n}$

Ex 5

$T(p) = p' \ \forall p \in P(\mathbb{R})$

Ex6

$T(f) = \int_0^1 f\,dx$, $f: [0,1] \to \mathbb{R}$ continuous

Ex7

$T(p) = x^2 p$ for $p \in P(\mathbb{R})$

~~Ex 8~~ Backwards shift: $T: \underbrace{\mathbb{F}^\infty}_{\text{scalars}} \to \mathbb{F}^\infty$,

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$$

All eight are linear. Let's just show this for Ex 8 today.

Check (1) $T((x_1, x_2, \dots) + (y_1, y_2, \dots)) = T((x_1+y_1, x_2+y_2, \dots))$

$= (x_2+y_2, x_3+y_3, \dots) =$

$(x_2, x_3, \dots) + (y_2, y_3, \dots) =$

$T((x_1, x_2, \dots) + T(y_1, y_2, \dots))$.

(2) $T(c(x_1, x_2, \dots)) = T((cx_1, cx_2, \dots)) =$

$(cx_2, cx_3, \dots) = c(x_2, x_3, \dots) =$

$c\,T(x_1, x_2, \dots)$.

Def:

$T: V \to W$ is called an <u>isomorphism</u> if it is one-to-one (1-1) and onto (surjective). If there exists an isomorphism $T: V \to W$, we say V and W are <u>isomorphic</u>, and we think of them as being the same, and write

$$V \cong W \quad \text{or} \quad V \cong W \text{ linearly}$$

f is 1-1: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$

$f: V \to W$: $W = \{ f(v) : v \in V \}$

Item 2 in examples above is an isomorphism; the others usually are not.

Ex 9

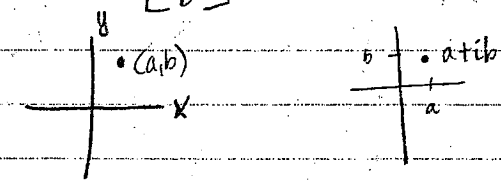
$$T: \mathbb{R}^3 \to P_2(\mathbb{R}) : \begin{bmatrix} a \\ b \\ c \end{bmatrix} \to a + bx + cx^2$$

(* Check T is linear)

and T is an isomorphism. So $\mathbb{R}^3 \cong P_2(\mathbb{R})$.

Ex 10

$$T: \mathbb{R}^2 \to \mathbb{C} : \begin{bmatrix} a \\ b \end{bmatrix} \to a + ib, \quad a, b \in \mathbb{R}.$$



Onto since any z in the complex number can be written $z = a+ib$, $a, b \in \mathbb{R}$, so $z = T((a,b))$.

1-1: $T((a,b)) = T((c,d)) \Leftrightarrow a+ib = c+id$.
$\Rightarrow$ real part $a = c$, imag part $b = d \Rightarrow (a,c) = (b,d)$

$T((a,b) + (c,d)) = T(a+c, b+d) =$

$a+c + i(b+d) = (a+ib) + (c+id) =$

$T(a,b) + T(c,d)$

$T(k(a,b)) = T(ka, kb) = ka + ikb =$

$k(a+ib) = k(T(a,b)), k \in \mathbb{R}.$

So $\mathbb{R}^2 \cong \mathbb{C}$, as vector spaces over $\mathbb{R}$.

Similarly $\mathbb{R}^{2n} \cong \mathbb{C}^n$ as vector spaces over $\mathbb{R}$

via $T: \mathbb{R}^{2n} \to \mathbb{C}^n$: $\begin{bmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ \vdots \\ a_n \\ b_n \end{bmatrix} \to \begin{bmatrix} a_1 + ib_1 \\ a_2 + ib_2 \\ \vdots \\ a_n + ib_n \end{bmatrix}$

Def:

If $T \in \mathcal{L}(V, W)$, set $\ker(T) = \{x \in V; T(x) = 0\}$
This is the kernel or null space of T
(book: null T).

Exs above

1. kernel is the entire domain V
2. $\ker I = (0)$.
3. Same as ex 2.
4. $\ker(T)$ you meet in 2331 = nullspace of matrix A.
5. $\ker(T) = \{$constant functions$\}$.
6. we did not talk about (it is harder).
7. $\ker(T) = (0)$
8. has kernel $\{(c, 0, 0, \ldots), c \in \mathbb{F}\}$.

9+10: $\ker(T) = (0)$.

Prop 1

If $T \in \mathcal{L}(V, W)$ then $\ker(T)$ is a subspace of V.

Proof: We said at start of Ch 3 that $T(0) = 0$ so $0 \in \ker T$.

If $v, w \in \ker T$, $c \in \mathbb{F}$,
$T(cv + w) = T(cv) + T(w) = cT(v) + T(w) =$

$c0 + 0 = 0$. So $cv + w \in \ker(T)$ so $\ker(T)$ is a subspace by Chap 1, prop 7.

Prop 2

If $T \in \mathcal{L}(V, W)$, then T is one-to-one iff $\ker(T) = 0$.

Proof: ($\Longrightarrow$)

If $v \in \ker T$, then $T(v) = 0 = T(0)$, so if T is one-to-one, $v = 0$.
So $\ker(T) = (0)$.

($\Longleftarrow$)

Suppose $\ker(T) = 0$. If $T(v_1) = T(v_2)$
$\Rightarrow 0 = T(v_1) - T(v_2) = T(v_1) + T(-v_2)$
$= T(v_1 - v_2)$ So $v_1 - v_2 \in \ker(T) = 0$
$\Rightarrow v_1 = v_2$.

Def: $\mathrm{Ran}(T)$ = range of $T = \{T(v) ; v \in V\} = T(V)$, if $T: V \to W$.

T is onto (surjective) means $\mathrm{Ran}(T) = W$, if $T: V \to W$

Prop 3

$\mathrm{Ran}(T)$ is a subspace of W if $T: V \to W$ is linear.

Proof: $T(0) = 0 \in \mathrm{Ran}(T)$. If $v, w \in \mathrm{Ran}(T)$, $c \in \mathbb{F}$, by Chap 1, Prop 7, we need to show $cv + w \in \mathrm{Ran}(T)$. Now $v = T(x)$, $w = T(y)$, $x, y \in V$, so $cv + w = cT(x) + T(y) = T(cx) + T(y) = T(cx + y) \in \mathrm{Ran}(T)$.

In Ex's 1–10 above, the ranges are

1. 0
2. V
3. U
4. $T\vec{x} = A\vec{x}$ has range = span of the columns of A (its dimension = rank(A)).
5. $P(\mathbb{R})$ (T is surjective).
6. $\mathbb{R}$ (surjective).
7.

Structure of the space $\mathcal{L}(V, W)$

It is a vector space:

$S, T \in \mathcal{L}(V, W)$, define $(S+T)(v) = S(v) + T(v)$.

$(cT)(v) = cT(v)$, $v \in V, c \in \mathbb{F}$.

Check 10 conditions to be a vector space

1. $(S+T)(v_1+v_2) = S(v_1+v_2) + T(v_1+v_2)$
$S(v_1) + S(v_2) + T(v_1) + T(v_2)$ (since linear)
$= (S+T)(v_1) + (S+T)(v_2)$

$(S+T)(cv) = S(cv) + T(cv) = cS(v) + cT(v)$
$= c((S+T)(v))$

2. cT is linear

* finish proof of this, and of items 3-10 of def of a vector space (HW).

The zero is the zero map (Ex 1, above). The additive inverse is $(-1)(cT)$.

Note that if $T \in \mathcal{L}(V, W)$ and $R\ \mathcal{L}(W, Z)$, where V, W, Z are vector spaces, then

$$R \circ T : V \to Z$$

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[rr, bend left, "R \circ T"] & W \arrow[r, "R"] & Z \end{tikzcd}$$

Ex 7

(Ex 4 and Ex 7 above)

If $T(p) = p'$, $R(p) = x^2 p$, then

$(R \circ T)(p) = x^2 p'$. eg $(R \circ T)(1 + x^2) = 2x^3$.

Claim:

$R \circ T$ is linear: $R \circ T(v_1 + v_2) = R(T(v_1 + v_2)) =$
$R(T(v_1) + T(v_2)) = R(T(v_1)) + R(T(v_2)) =$
$(R \circ T)(v_1) + (R \circ T)(v_2)$

$(R \circ T)(cv) = R(T(cv)) = R(cT(v)) = cR(T(v)) =$
$c(R \circ T)(v)$.

We have rules such as

$(S + T) \circ R = S \circ R + T \circ R$
$S \circ (R + P) = S \circ R + S \circ P$ } distributive laws
$S \circ (R \circ T) = (S \circ R) \circ T$
$R \circ (cT) = cR \circ T$.

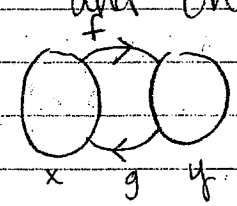
6/10/08

Prop 4

$T: V \to W$ is an isomorphism iff $\exists\, S \in \mathcal{L}(W, V) \ni S \circ T = I_V,\ T \circ S = I_W$.

Proof: $S \circ T = I_V,\ T \circ S = I_W$ (note $S = T^{-1}$).

($\Leftarrow$) $f \circ g = I,\ g \circ f = I$ for any functions f, g imply f 1-1 and onto and $g = f^{-1}$.



$fg = I_Y$
$gf = I_X$

f onto: any $y \in Y = f(g(x))$

f 1-1: $f(x_1) = f(x_2) = g(f(x_1)) = g(f(x_2))$
$\Rightarrow x_1 = x_2$

$fg = I \Rightarrow fgg^{-1} = I \circ g^{-1} \Rightarrow f = g^{-1}$

($\Rightarrow$) Since T is one-to-one and onto, it has an inverse, $S = T^{-1}$. We need to show $S = T^{-1}$ be linear. So let $w_1, w_2 \in W$, $c \in \mathbb{F}$, $w = T(v_1)$ say, $w_2 = T(v_2)$.

So $w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2)$, so applying S we get
$S(w_1 + w_2) = ST(v_1 + v_2) = v_1 + v_2 = S(w_1) + S(w_2)$
Also, $cw_1 = cT(v_1) = T(cv_1)$, so applying S, we get $S(cw_1) = ST(cv_1) = cv_1 = cS(w_1)$.

Let's now connect what we've done so far in Ch. 3, with material in Ch. 2

Theorem 1

V f.d., $T: V \to W$ linear $\Rightarrow$ Ran(T) is f.d. and $\dim(V) = \dim(\ker T) + \dim(\operatorname{Ran} T)$.

Proof:

Let $(u_1, \dots, u_m)$ be a basis for $\ker(T)$, so $m = \dim(\ker T)$. By Chap 2, Thm 3, $\exists$ vectors $(v_1, \dots, v_n)$ $\ni$ $(u_1, \dots, u_m, v_1, \dots, v_n)$ is a basis for V. So $\dim(V) = m+n$.
So we are done if we can prove $\dim(\operatorname{Ran} T) = n$. But this follows immediately from:

Claim 1: $(T(v_1), T(v_2), \dots, T(v_n))$ is a basis for Ran T. This in turn follows from

(a) $(T(v_1), \dots, T(v_n))$ is l.i.:

$$\sum_{k=1}^{n} c_k T(v_k) = 0, \; c_k \in \mathbb{F} \Rightarrow T\left(\sum_{k=1}^{n} c_k v_k\right)$$

$$\underset{\text{(since T linear)}}{=} \sum_k c_k T(v_k) = 0.$$

So $\sum_{k=1}^{n} c_k v_k \in \ker T$.

$$\Rightarrow \sum_{k=1}^{n} c_k v_k = \sum_{i=1}^{m} d_i u_i, \text{ some } d_i \in \mathbb{F},$$

Hence $c_k = 0$ $\forall$ k since $(u_1, \dots, u_m, v_1, \dots, v_n)$ is l.i.

(b) $(T(v_1), \dots, T(v_n))$ spans Ran T.

Let $z \in \operatorname{Ran} T$, so $z = T(v)$ say.
We can write $v = \sum^{m} c_k u_k + \sum^{n} d_k v_k$
apply T:

$$z = T(v) = T(\sum c_k \overset{0}{\cancel{u_k}} + \sum d_k v_k) =$$

$$\sum_k d_k(T(v_k)).$$

Corollary 1 If V, W, f.d. with $\dim(V) > \dim(W)$, then $\nexists$ one-to-one linear $T: V \to W$.

Corollary 2 V, W f.d. with $\dim(V) < \dim(W)$, then $\nexists$ surjective linear $T: V \to W$.

Proof (1):
By Thm 1, $\dim \ker T =$
$\dim V - \dim(\operatorname{Ran} T) \geq \dim V - \dim W > 0$
Hence $\ker T \neq (0)$, so T is not one-to-one by Prop 2.

SIDEBAR: $\dim((0)) = 0$

Proof (2): By Thm 1, $\dim(\operatorname{Ran} T) =$
$\dim(V) - \dim(\ker T) \leq \dim(V) < \dim(W)$
so $\operatorname{Ran}(T) \neq W$.

Corollary 3

If V f.d and $T \in \mathcal{L}(V)$ then T is one-to-one iff T is surjective

Proof: T 1-1 $\overset{\text{Prop 2}}{\Longrightarrow}$ $\ker T = (0) \overset{\text{Thm 1}}{\Longrightarrow}$
$\dim V = 0 + \dim(\operatorname{Ran}(T)) \overset{\text{Ch2 HW11}}{\Longrightarrow} \operatorname{Ran}(T) = V$.

Conversely, T surjective $\overset{\text{Thm 1}}{\Longrightarrow}$
$\dim(\ker T) = \dim(V) - \dim(\operatorname{Ran} T) =$
$\dim(V) - \dim(V) = 0 \Rightarrow \ker(T) = (0)$

So T is 1-1 by Prop. 2.

Corollary 4

Two f.d. vector spaces are isomorphic iff they have the same dimensions.

Proof: ($\Rightarrow$) let $T: V \to W$ be an isomorphism.
By Thm 1, $\dim(V) = 0 + \dim(\operatorname{Ran} T)$
$= \dim(W)$

($\Leftarrow$) If $\dim(V) = \dim(W)$, $\exists$ basis $(v_1, \dots, v_n)$ of V, $(w_1, \dots, w_n)$ of W, same n.

General principle: If V, W are any v.s.'s and $(v_1, \dots, v_n)$ is a basis for V, and $(w_1, \dots, w_n)$ is any <u>list</u> in W, and if we define $T: V \to W$ by

$T\left(\sum_{k=1}^{n} c_k v_k\right) = \sum_{k=1}^{n} c_k w_k$, $c_1, \dots, c_n \in \mathbb{F}$, then T is well defined and linear (and of course $T(v_k) = w_k \ \forall k = 1, \dots, n$.

To prove this principle, note by Ch. 2 Prop 3, any $v \in V$ can be written as $\sum_k c_k v_k$ in one and only one way, so T is well-defined.

T is linear: If $c, c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$, then $T\left(\sum_k c_k v_k + \sum d_k v_k\right) = T\left(\sum_k (c_k + d_k) v_k\right) =$

$$\sum_k (c_k + d_k) w_k = \sum_k c_k w_k + \sum_k d_k w_k =$$

$$T\left(\sum_k c_k v_k\right) + T\left(\sum_k d_k v_k\right)$$

Also, $T\left(c \sum_k c_k v_k\right) = T\left(\sum_k c c_k v_k\right) =$

$$\sum_k c c_k w_k = c \sum_k c_k w_k = c\, T\left(\sum_k c_k v_k\right)$$

This ends proof of general principle Back to proof of Corollary 4

By general principle above, we can define $T: V \to W$ by $T\left(\sum_{k=1}^n c_k v_k\right) = \sum_{k=1}^n c_k w_k$ and T is linear. T is onto, since any $w \in W$ is of form $\sum_{k=1}^n c_k w_k = T\left(\sum_k c_k v_k\right)$.

T is 1-1 because if $\sum_{k=1}^n c_k v_k \in \ker(T)$, then $0 = T\left(\sum_k c_k v_k\right) = \sum_{k=1}^n c_k w_k$. Since $(w_1, \dots, w_n)$ is l.i., we get $c_k = 0 \ \forall k$, so

$$\sum_k c_k v_k = 0.$$

These results have many very important applications to solving linear equations

Recall in 2331, the 'main thing' was solving

$$(*) \quad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Here, a_{ij} are constants, so we have m equations in n unknowns. If $b_1 = b_2 = \dots = b_m = 0$, $(*)$ is called homogeneous.

$(*)$ rewritten $A\vec{x} = \vec{b}$ where $A = [a_{ij}]$, $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, which has associated homogeneous equation $A\vec{x} = \vec{0}$.

Many of the main facts from 2331 follow from what we've done above, by writing $T: \mathbb{R}^n \to \mathbb{R}^m$ for $T(\vec{x}) = A(\vec{x})$. So $(*)$ or $(**)$ becomes $T(\vec{x}) = \vec{b}$, and the solutions of a homogeneous equation $A\vec{x} = \vec{0}$ is precisely $\ker(T)$.

The set of $\vec{b}$ for which $(**)$ has a solution is $\mathrm{Ran}(T)$.

So, Thm 1 above gives the formula $n = \mathrm{nullity}(A) + \mathrm{rank}(A)$ from 2331, where nullity of A is the dimension of the set of solutions to $A\vec{x} = \vec{0}$, and $\mathrm{rank}(A)$ is the dimension of span of columns of A. The fact that $\mathrm{rank}(A) = \dim(\mathrm{Ran}\,T)$ follows from fact $A\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \vec{b}$ iff $c_1\vec{A}_1 + \dots + c_n\vec{A}_n = b$, where a_k is the k^{th} column of A.

Indeed, $A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1\vec{A_1} + c_2\vec{A_2} + \dots c_n\vec{A_n}$.

Corollary 1 in 2331 language says:

If $m<n$ in (*) then $A\vec{x} = \vec{0}$ has nontrivial solutions (since existence of nontrivial solutions to $A\vec{x} = \vec{0} \Leftrightarrow \ker T \neq (0) \Leftrightarrow T$ not 1-1.

Corollary 2 in 2331 language says:

If $m > n$ in (*), then there are some values for $b_1, b_2, \dots, b_n$ $\ni$ (*) has no solution. (Since (*) having no solution $\Leftrightarrow \vec{b} \notin Ran(T)$ so this happens when T is not surjective).

Corollary 3 and 4 in 2331 language says:

If $n = m$ in (*), then A is invertible iff nullity$(A) = 0$ iff (*) can be solved $\forall b_1, \dots, b_n$ iff the solution to (*) is unique $\forall b_1, \dots, b_n$ iff rank$(A) = n$. (This is because nullity$(A)=0$ iff $\ker(T) = 0$ iff T is 1-1 iff (by Corollary 4) T is onto iff $Ran(T)$ is $\mathbb{R}^n$ iff (*) can be solved $\forall b_1, b_2, \dots, b_n$ iff rank$(A) = n$.
By 2331, this also happens iff A is invertible.
If A is invertible, the solution to (**) is unique: $\vec{x} = A^{-1}\vec{b}$.

DAVID BLECHER

Notes

6/11 <u>The matrix of a linear map</u>

If $T \in \mathcal{L}(V,W)$ and $B = (v_1, \ldots, v_n)$ is a basis for V, and $C = (w_1, \ldots, w_m)$ is a basis for W, then for any $j = 1, \ldots, n$, we can write uniquely (in one and only one way)

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

The matrix $[a_{ij}]$ whose i-j-entry is a_{ij}, is called the <u>matrix of</u> T with respect to the bases B, C, and it is written $\mathcal{M}(T, B, C)$ or just $\mathcal{M}(T)$ if B, C are understood.

Example 1: $T: \mathbb{R}^2 \to \mathbb{R}^3$, $T(x,y) = (x+3y, 2x+7y, 7x+9y)$, and let B and C be the 'canonical / standard' bases for $\mathbb{R}^2$ and $\mathbb{R}^3$. $B = (\vec{i}, \vec{j})$, $C = (\vec{i}, \vec{j}, \vec{k})$. Note
$T(\vec{i}) = T(1,0) = (1,2,7) = 1\vec{i} + 2\vec{j} + 7\vec{k}$ and
$T(\vec{j}) = T(0,1) = (3,7,9) = 3\vec{i} + 7\vec{j} + 9\vec{k}$. Hence

$$\mathcal{M}(T, B, C) = \begin{bmatrix} 1 & 3 \\ 2 & 7 \\ 7 & 9 \end{bmatrix}$$

Sometimes write this as $M(T)$.

Example 2: V any v.s. with basis $B = (v_1, \ldots, v_n)$ let $T = I_V$.
$I_V(v_j) = v_j = 0v_1 + 0v_2 + \ldots + 0v_{j-1} + 1v_j + 0v_{j+1} + \ldots + 0v_n$.
So

$$\mathcal{M}(I_V, B, B) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & 0 & & 0 \\ 0 & 0 & & 1 \end{bmatrix} = I_n$$

, $n \times n$ identity matrix.

Example 3: Let $T: \mathbb{F}^n \to \mathbb{F}^m$ be $T(\vec{x}) = A\vec{x}$, where $A \in M_{m,n}$
Let B, C be standard bases for $\mathbb{F}^n, \mathbb{F}^m$. Then

$$T(\vec{e}_j) = A\vec{e}_j = A \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} j = \vec{A}_j$$

, the j^{th} column of A

$$* \; \vec{A}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = a_{1j}\vec{e}_1 + a_{2j}\vec{e}_2 + \ldots + a_{mj}\vec{e}_m$$

$$\text{So } \mathcal{M}(T,B,C) = \begin{array}{c} \text{j^{th} column} \\ \downarrow \\ \begin{bmatrix} a_{11} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \end{array} = A$$

<u>Theorem 2</u>: If V, W are f.d. v.s. over $\mathbb{F}$ with bases $(v_1, \ldots, v_n)$ for V, and $(w_1, \ldots, w_m)$ for W, then the map $\theta : \mathcal{L}(V,W) \longrightarrow M_{m,n}$ defined by $\theta(T) = \mathcal{M}(T)$, the matrix of T with respect to these bases, is an isomorphism. So $\mathcal{L}(V,W) \cong M_{m,n}$.

proof:

Suppose $S, T \in \mathcal{L}(V,W)$, $c \in \mathbb{F}$, $\mathcal{M}(T) = [a_{ij}]$, $\mathcal{M}(S) = [b_{ij}]$. $(S+T)(v_j) = S(v_j) + T(v_j) = \sum_{i=1}^m b_{ij} w_j + \sum_{i=1}^m a_{ij} w_j = \sum_{i=1}^m (b_{ij} + a_{ij}) w_j$. So $\mathcal{M}(S+T) = [b_{ij} + a_{ij}] =$

$[b_{ij}] + [a_{ij}] = \mathcal{M}(S) + \mathcal{M}(T)$.
$(cT)(v_j) = cT(v_j) = c \sum_{i=1}^m a_{ij} w_j = \sum_{i=1}^m (ca_{ij})(w_i)$. So

$\mathcal{M}(cT) = [ca_{ij}] = c[a_{ij}] = c\mathcal{M}(T)$. So θ is linear.
<u>θ is 1-1</u>: If $\theta(T) = 0$ then $[a_{ij}] = 0 \implies a_{ij} = 0$ $\forall i,j \implies T(v_j) = \sum_i a_{ij} w_i = 0$ $\forall j$, so $T(\sum_{j=1}^n c_j v_j) =$

$\sum_j c_j T(v_j) = \sum_j c_j 0 = 0$. Hence $T(v) = 0$ $\forall v \in V$, so $T = 0$.

So $\ker T = (0)$ and T is 1-1.
<u>T is onto</u>: Pick $[d_{ij}] \in M_{m,n}$, By 'General Principle' in proof of Cor. 4 $\exists T \in \mathcal{L}(V,W)$ s.t. $T(v_j) = (\sum_{i=1}^m d_{ij} w_i)$ $\forall j = 1, \ldots, n$. By definition of $\mathcal{M}(T)$, we

have $\theta(T) = \mathcal{M}(T) = [d_{ij}]$, so θ is onto. □

<u>Corollary 5</u>: If $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ then T is of the form $T(\vec{x}) = A\vec{x}$ for $\forall \vec{x} \in \mathbb{R}^n$ for some $A \in M_{m,n}$

proof:

Let $\theta: \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}$ be as in last theorem, and let $A = \theta(T)$. To show $T(\vec{x}) = A\vec{x}$ $\forall \vec{x} \in \mathbb{R}^n$, it is enough to prove this for $\vec{x}$ in $(\vec{e}_1, \dots, \vec{e}_n)$, the standard basis (since $T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = T(\sum x_k \vec{e}_k) = \sum_k x_k T(\vec{e}_k)$, and similarly $A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_k x_k A\vec{e}_k$), but by definition of $A = \theta(T) = \mathcal{M}(T)$, we have $T(\vec{e}_j) = \sum_{i=1}^{m} a_{ij} \vec{e}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = \vec{A_j} = A\vec{e}_j$. $\square$

<u>Proposition 5</u>: *Note $\mathcal{M}(R \circ T) = \mathcal{M}(R)\mathcal{M}(T)$* . If $T \in \mathcal{L}(V, W)$ $R \in \mathcal{L}(W, Z)$, where V, W, Z are v.s.'s with bases B, C, D respectively then $\mathcal{M}(R \circ T) = \mathcal{M}(R)\mathcal{M}(T)$, where these $\mathcal{M}$'s are the matrices with respect to the given bases.

proof:

Let $\mathcal{M}(T) = [a_{ij}]$ and $\mathcal{M}(R) = [c_{ij}]$. Then $R(w_k) = \sum_{i=1}^{r} c_{ik} z_i$ $\forall k = 1, \dots, m$. where $(w_1, \dots, w_m)$ is basis C for W, $(z_1, \dots, z_r)$ is basis D for Z, and $(v_1, \dots, v_n) = B$. Then $RT(v_j) = R\left(\sum_{k=1}^{m} a_{kj} w_k\right) = \sum_{k=1}^{m} a_{kj} R(w_k) = \sum_{k=1}^{m} a_{kj} \sum_{i=1}^{r} c_{ik} z_i = \sum_{i=1}^{r} \sum_{k=1}^{m} c_{ik} a_{kj} = \sum_{i=1}^{r} d_{ij} z_i$, where $d_{ij} = \sum_{k=1}^{m} c_{ik} a_{kj}$. So $\mathcal{M}(RT) = [d_{ij}] = [c_{ij}][a_{ij}] = \mathcal{M}(R)\mathcal{M}(T)$. $\square$

<u>Corollary 6</u>: If A is an $n \times n$ matrix with nullity 0 (or equivalently with rank n), then A is invertible. (i.e. $\exists$ matrix B s.t. $AB = I_n$, $BA = I_n$)

proof:

Let $T(\vec{x}) = A\vec{x}$ $\forall \vec{x} \in \mathbb{R}^n$, then as we said yesterday, nullity $0 \Rightarrow T$ is 1-1 $\overset{\text{Cor. 3}}{\Longrightarrow}$ T is onto so T is an isomorphism. By prop. 4, $\exists$ $S \in \mathcal{L}(\mathbb{R}^n)$ s.t. $S \circ T = I$, $T \circ S = I$. Let $B = \mathcal{M}(S)$, by Ex. 3 above, $\mathcal{M}(T) = A$ (with respect to standard bases).

$BA = \mathcal{M}(S)\mathcal{M}(T) \overset{\text{Prop. 5}}{=} \mathcal{M}(S \circ T) = \mathcal{M}(I) \overset{\text{Ex. 2}}{=} I_n$. Similarly

$AB = \mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(T \circ S) = \mathcal{M}(I) = I_n$. □

<u>Corollary 1</u>: If $T \in \mathcal{L}(V, W)$ is an isomorphism and B is a basis for V, C is a basis for W, then $\mathcal{M}(T, B, C)^{-1} = \mathcal{M}(T^{-1}, C, B)$

proof: $\mathcal{M}(T, B, C)\mathcal{M}(T^{-1}, C, B) \overset{\text{Prop. 5}}{=} \mathcal{M}(T \circ T^{-1}) = \mathcal{M}(I) = I_n$.

Similarly $\mathcal{M}(T^{-1}, C, B)\mathcal{M}(T, B, C) = \mathcal{M}(T^{-1} \circ T) = \mathcal{M}(I) = I_n$ so the one matrix is the inverse matrix of the other. □

- <u>New notation</u>: If $B = (v_1, \ldots, v_n)$ is a basis for V and if $v \in V$ then we can write uniquely $v = \sum_{k=1}^{n} c_k v_k$ with $c_k \in \mathbb{F}$. Write $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ as $[v]_B$ or $[v]$ if B is understood. Sometimes called the <u>coordinate</u> <u>vector</u> of v with respect to B.

<u>Remark</u>: In the setup above $V \cong \mathbb{R}^n$ via the map $T: \mathbb{R}^n \to V$ define by $T\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) = \sum_{k=1}^{n} c_k v_k \in V$.

As an exercise, check T is linear. Clearly T is onto, since $(v_1, \ldots, v_n)$ is spanning V, and T is 1-1 since if $T\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) = 0$ then $\sum c_k v_k = 0 \Longrightarrow c_1 = c_2 = \ldots = 0$ since $(v_1, \ldots, v_n)$ is L.I.. Thus T is an isomorphism.

Now $T^{-1}: V \longrightarrow \mathbb{R}^n$ is nothing but the map $v \mapsto [v]_B$ in notation above. (Why?)

<u>Proposition 6</u>: If $T \in \mathcal{L}(V,W)$, and B is a basis for V, C is a basis for W then

$$[T(v)]_C = \mathcal{M}(T)\,[v]_B \quad \forall v \in V$$

where $\mathcal{M}(T)$ is the matrix T with respect to these bases.

proof:

Suppose $\mathcal{M}(T) = [a_{ij}]$, $[v]_B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, $B = (v_1, \dots, v_n)$, $C = (w_1, \dots, w_n)$ so $v = \sum_{j=1}^{n} b_j v_j$ and $T(v) = T(\sum b_j v_j) = \sum_j^n b_j T(v_j) = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij} w_j$ (by def.^{n} of $\mathcal{M}(T)$). This equals $\sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} b_j) w_i = \sum_{i=1}^{m} d_i w_i$, where $d_i = \sum_{j=1}^{n} a_{ij} b_j$. So $[T(v)]_C = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix} = [a_{ij}] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} =$

$\mathcal{M}(T)[u]_B$. $\square$

<u>Proposition 7</u>: V, W f.d. $\Longrightarrow \mathcal{L}(V,W)$ f.d. and $\dim(\mathcal{L}(V,W)) = (\dim V)(\dim W)$.

proof:

Suppose $\dim V = n$, $\dim W = m$, then by Theorem 2, $\mathcal{L}(V,W) \cong M_{m,n}$. But we said $\dim(M_{m,n}) = m \cdot n = (\dim V)(\dim W)$. And $\dim(\mathcal{L}(V,W)) = \dim(M_{m,n}) = (\dim V)(\dim W)$

$\square$

END OF CHAPTER 3

DAVID BLECHER

Some corrections to HW 3.
Delete last thing in Q5 [H-K p.83] "Can you ~~describe~~..."
Delete Q2 [H-K p.86]
Add Q6 in Axler
To be graded: H-K: first 3 questions [1,9,2] and last: [1a,b]
Axler: 5, 6, 12, 22, 26

CHAPTER 4 Polynomials

$p(z) = a_0 + a_1 z + \dots + a_m z^m$ degree of $p = m$ if $a_m \neq 0$.
zero OR root of p: number λ s.t. $p(\lambda) = 0$

Proposition 4.1: λ is a root $\iff$ you can factor
$p(z) = (z-\lambda)\, q(z)$, $q \in \mathcal{P}(\mathbb{F})$.

Corollary 4.3: A polynomial of degree m has at most m distinct roots.
* distinct means objects $O_1, O_2, \dots, O_m$ are distinct means $O_i \neq O_j$

4.5 Division Algorithm: $p, q \in \mathcal{P}(\mathbb{F})$, $p \neq 0 \implies \exists\ s, r \in \mathcal{P}(\mathbb{F})$ s.t. $q = sp + r$ and $\deg r < \deg p$.

4.7 Fundamental Theorem of Algebra: every non-constant polynomial has a root, possibly complex.
Eg: $1 + x^2$ only roots are complex: $\pm i$

4.8 If $p \in \mathcal{P}(\mathbb{F})$ is not constant, then we have a unique factorization (up to order of factors).
$p(z) = c(z-\lambda_1)(z-\lambda_2)\dots(z-\lambda_m)$, where $c \neq 0$, $c, \lambda_1, \dots, \lambda_m \in \mathbb{C}$
The λ_i are exactly the roots.

• $\mathbb{C}$ numbers : $z = a+ib$ $a, b \in \mathbb{R}$, uniquely $a = Re(z)$, $b = Im(z)$
• complex conjugate : $\bar{z} = a - ib$
• absolute value : $|z| = \sqrt{a^2+b^2}$
• $z + \bar{z} = 2 \, Re \, z$
• $z - \bar{z} = 2i \, Im \, z$
• $\bar{z} z = |z|^2 \longrightarrow z\left(\frac{\bar{z}}{|z|^2}\right) = 1$
• $1/z = \frac{\bar{z}}{|z|^2} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i$
• $\overline{w+z} = \bar{w} + \bar{z}$, $\overline{wz} = \bar{w}\bar{z}$
• $\bar{\bar{z}} = z$, $|wz| = |w| \, |z|$

<u>4.10</u> p has real coefficients, $p(\lambda) = 0 \Longrightarrow p(\bar{\lambda}) = 0$
proof:

$$\text{If } p = a_0 + a_1 z + \ldots + a_n z^n, \quad 0 = \bar{0} = \overline{p(\lambda)} =$$
$$\overline{a_0 + a_1\lambda + \ldots + a_n\lambda^n} = \bar{a}_0 + \overline{a_1\lambda} + \ldots + \overline{a_n\lambda^n}$$
$$= a_0 + a_1\bar{\lambda} + \ldots + a_n\overline{\lambda^n}$$
$$= a_0 + a_1\bar{\lambda} + \ldots + a_n(\bar{\lambda})^n$$
$$= p(\bar{\lambda}) \quad \square$$

• Polynomials of operators: If $T \in \mathcal{L}(V)$ define $T^0 = I_V = I$

$$T^m = \underbrace{T \circ T \circ \ldots \circ T}_{m=1,2,3,\ldots} \qquad T^{-m} = \underbrace{T^{-1} \circ T^{-1} \circ \ldots \circ T^{-1}}_{m=1,2,3,\ldots}$$

if T is <u>invertible</u>, by which $\underset{(\text{Prop. 4 Ch.3})}{}$ we mean that T is an isomorphism, or equivalent, $\exists \, S \in \mathcal{L}(V)$ s.t. $ST = TS = I$

Ex. $T^n T^m = T^{n+m}$; $(T^n)^m = T^{nm}$

(idea of proof: $(T^3)^2 = T^3 T^3 = TTTTTT = T^6 = T^{3 \cdot 2}$.

If $p \in \mathcal{P}(\mathbb{F})$ with $p = a_0 + a_1 z + \ldots + a_n z^n$, define
$p(T) = a_0 I + a_1 T + \ldots + a_n T^n$

Ex. $p = x^2 - 2x + 3$, then $p(T) = T^2 - 2T + 3I$

The function $\pi : \mathcal{P}_m(\mathbb{F}) \to \mathcal{L}(V)$ defined by $\pi(p) = p(T)$, can be defined in an equivalent way using 'general principle' met in the proof of Ch. 3 corollary 4.

$$\pi\left(\sum_{k=0}^{m} c_k z^k\right) = \underbrace{\sum_{k=0}^{m} c_k T^k}_{p(T)} \text{ is well defined and linear.}$$

Consequence: $(p+q)(T) = p(T) + q(T)$, $(cp)(T) = c\,p(T)$ $\forall\, p, q \in \mathcal{P}(\mathbb{F})$ $c \in \mathbb{F}$

We also have $(pq)(T) = p(T)\,q(T)$. * composition

If energetic, prove last one. Here is the idea: Suppose p, q have degree ≤ 1, so $p = a_0 + a_1 z$ $q = b_0 + b_1 z$, then $pq = (a_0 + a_1 z)(b_0 + b_1 z) = a_0 b_0 + (a_1 b_0 + a_0 b_1) z + a_1 b_1 z^2$

So $(pq)(T) = a_0 b_0 I + (a_1 b_0 + a_0 b_1) T + a_1 b_1 T^2$

$$= (a_0 I + a_1 T)(b_0 I + b_1 T)$$

$$= p(T)\, q(T)$$

Consequence: $p(T)\,q(T) = q(T)\,p(T)$

proof:

$$\begin{array}{ccc} p(T)\,q(T) & & q(T)\,p(T) \\ \| & & \| \\ (pq)(T) & = & (qp)(T) \end{array}$$

* Memorize these facts, since we'll be using them silently below.

Lemma: Suppose $T \in \mathcal{L}(V)$, $v \in V$, $c \in \mathbb{F}$ and $T(v) = cv$. Then $p(T)(v) = p(c)v$ for any $p \in \mathcal{P}(\mathbb{F})$

* $T^2(v) = T(T(v)) = T(cv) = cT(v) = c^2 v$

$T^3(v) = T(T^2(v)) = T(c^2 v) = c^2 T(v) = c^3 v$ and more generally $T^k(v) = c^k v$

proof:

If $p = a_0 + a_1 z + \ldots + a_m z^m$ then $p(T)(v) = (a_0 I + a_1 T + \ldots + a_m T^m)(v)$

$= a_0 v + a_1 c v + \ldots + a_m c^m v = p(c) v$.

CHAPTER 5 : Eigenvalues and Eigenvectors

* For the rest of this $\boxed{\text{course}}$, V is f.d. vector space, $V \neq (0)$

<u>Definition</u>: If $T \in \mathcal{L}(V)$ then a number λ is called an <u>eigenvalue</u> of T if $\exists$ <u>nonzero</u> $v \in V$ s.t. $T(v) = \lambda v$.

* eigenvalue (e-value)

We say that v is an eigenvector (e-vector) corresponding to e-value λ. We also say 0 is an e-vector since $T(0) = 0 = \lambda 0$.

* <u>Clearly</u>: v is an e-vector corresponding to $\lambda \iff$ $Tv = \lambda v \iff (T - \lambda I)(v) = 0 \iff v \in \ker(T - \lambda I)$

So $\ker(T - \lambda I) = \{$all e-vectors corresponding to $\lambda\}$ <u>and this is called eigenspace written E_λ.</u>

* <u>Also clearly</u>: λ is an e-value $\iff \ker(T - \lambda I) \neq (0) \overset{\text{CH.3 Prop. 2}}{\iff}$ $T - \lambda I$ not <u>1-1</u> $\underset{\text{CH.3 Cor. 3}}{\iff}$ $T - \lambda I$ <u>not</u> onto

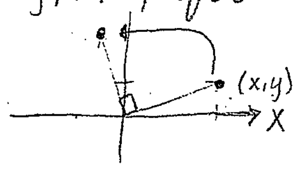
Examples:

(1) I_V has only one e-value, $\lambda = 1$ (since $v = \lambda v \implies (1-\lambda)v = 0 \implies 1 - \lambda = 0 \implies \lambda = 1$)

The e-space $E_1 = V$.

(2) $T(x,y) = (-y, x)$, for $(x,y) \in \mathbb{F}^2$

If $\mathbb{F} = \mathbb{R}$ 90° counter-clockwise rotation



a non-zero e-vector v does not exist, thus there are no e-values, but there are complex e-values.

$T(v) = \lambda v$ says $(-y, x) = \lambda(x, y)$, $(x,y) \neq (0,0)$

$\iff \begin{cases} -y = \lambda x \\ x = \lambda y \end{cases} \implies -y = \lambda(\lambda y) \implies -y = \lambda^2 y \implies (\lambda^2 + 1) y = 0$

Now $y \neq 0$ because $y = 0 \implies x = 0 \implies (x,y) = (0,0)$

this is a contradiction. So $\lambda^2+1=0 \Rightarrow$
$\lambda = \pm i$.
$E_i = \{(x, -ix) : x \in \mathbb{F}\}$, $E_{-i} = \{(x, ix) : x \in \mathbb{F}\}$
*No non-zero e-vectors lie in $\mathbb{R}^2$. If $\mathbb{F}=\mathbb{R}$ there are <u>no</u> e-values.

<u>Theorem 1</u>: If $\lambda_1, \ldots, \lambda_m$ are distinct e-values of $T \in \mathcal{L}(V)$, and if v_k is a non-zero e-vector corresponding to λ_k, $\forall k$, then $(v_1, v_2, \ldots, v_m)$ is L.I.

proof:

Suppose $\sum_{k=1}^{m} c_k v_k = 0$, $c_k \in \mathbb{F}$. Define

$p = (z-\lambda_2)(z-\lambda_3)\ldots(z-\lambda_m) \in \mathcal{P}(\mathbb{F})$, $0 = p(T)\left(\sum_{k=1}^{m} c_k v_k\right) =$

$\sum_{k=1}^{m} c_k\, p(T)(v_k) \overset{\text{Ch.4 Lemma}}{=\!=} \sum_{k=1}^{m} c_k\, p(\lambda_k) v_k = c_1 p(\lambda_1) v_1 \overset{\text{HW!}}{\Longrightarrow}$

$c_1 p(\lambda_1) = 0 \Rightarrow c_1 = 0$. Argue similarly with $p = (z-\lambda_1)(z-\lambda_3)\ldots(z-\lambda_m)$ to see $c_2 = 0$ and similarly $c_3 = c_4 = \ldots = c_m = 0$ $\square$

DAVID BLECHER

6/13 <u>Theorem 1</u>: If $\lambda_1, \dots, \lambda_m$ are distinct e-values of $T \in \mathcal{L}(V)$ and if v_j is an e-vector corresponding to λ_j, $\forall j$, then $(v_1, \dots, v_m)$ is L.I.

<u>Corollary 1</u>: $T \in \mathcal{L}(V) \Longrightarrow T$ has at most $\dim V$ distinct e-values

proof:

Let $(v_1, \dots, v_m)$ be as in Thm. 1, then by Thm. 1, these are L.I., so $m \le \dim V$ by Ch. 2 Thm. 1 □

<u>Theorem 2</u>: Every $T \in \mathcal{L}(V)$ has an e-value if V is a f.d. v.s. over $\mathbb{C}$, $V \neq (0)$.

proof:

Suppose $\dim V = n$ and $0 \neq v \in V$. Then $(v, Tv, T^2v, \dots, T^n v)$ has $n+1$ elements, so is linear dependent by Ch. 2 Thm 1, So $\sum_{k=0}^{n} c_k T^k v = 0$ for scalars c_k not all zero. Let $p(z) = \sum_{k=0}^{n} c_k z^k \in \mathcal{P}(\mathbb{C})$. By result in Ch. 4 we can factor $p = c(z-\lambda_1)(z-\lambda_2)\dots(z-\lambda_m)$, $\lambda_k \in \mathbb{C}$, $c \neq 0$. Hence $0 = p(T)v = c(T-\lambda_1 I)(T-\lambda_2 I)\dots(T-\lambda_m I)v$. So $(T-\lambda_1 I)(T-\lambda_2 I)\dots(T-\lambda_m I)$ is not one-to-one. So by HW 3 Q6, $\exists j$ s.t. $(T-\lambda_j I)$ is not one-to-one, ie $\ker(T-\lambda_j I) \neq (0)$ so λ_j is an e-value. □

* Thm 1, Corollary 1, Thm 2 will be on Test *

Recall: the 'main diagonal' of a square matrix $[a_{ij}]$ are the numbers $a_{11}, a_{22}, \dots, a_{nn}$ $\begin{bmatrix} 3 & 1 & 5 & 7 \\ 0 & 2 & 8 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -2 \end{bmatrix}$. A square matrix is <u>upper triangular</u> (Δ'r) if all numbers below main diagonal are zero. A square matrix is <u>lower triangular</u> if all numbers above main diagonal are zero. A square matrix is <u>diagonal</u> if all numbers above and below the main diagonal are zero. eg. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Proposition 1: If $T \in \mathcal{L}(V)$ and $(v_1, \dots, v_n)$ is a basis for V then $\mathcal{M}(T)$, the matrix of T with respect to this basis, if and only if $T(v_j) \in \operatorname{span}(v_1, \dots, v_j)$, $\forall j = 1, \dots, n$.

proof:

The second condition (the part after the if and only if) holds if and only if $T(v_j) = c_1 v_1 + c_2 v_2 + \dots + c_j v_j + 0 v_{j+1} + \dots + 0 v_n$. So by definition of $\mathcal{M}(T)$, this is saying that the j'th column of $\mathcal{M}(T)$ has 0's in jth row and below, $\forall j$, which happens iff $\mathcal{M}(T)$ is upper triangular. □

Theorem 3: V, a complex v.s. (f.d. $\neq (0)$), $T \in \mathcal{L}(V)$, then $\exists$ basis $\mathcal{B}$ of V s.t. $\mathcal{M}(T) = \mathcal{M}(T, \mathcal{B}; \mathcal{B})$ is upper triangular.

proof:

By induction on $\dim(V)$. If $\dim(V) = 1$, then $\mathcal{M}(T)$ is a 1×1 matrix, so diagonal. Suppose result is true $\forall$ v.s.'s of dimension $< k$. Suppose V has dim k, $T \in \mathcal{L}(V)$. By Thm. 2, T has an e-value λ. Let $U = \operatorname{Ran}(T - \lambda I)$. Now $\ker(T - \lambda I) \neq (0)$, so $(T - \lambda I)$ is not one-to-one so not surjective by Cor. 3 Ch. 3. So $U \neq V$, $\dim(U) < \dim(V)$. Note: $T(T - \lambda I)v = (T^2 - \lambda T)v = (T - \lambda I)Tv \in U$, which shows that $T(U) \subseteq U$. Let $R = T_U$ (this is the function from U to U defined by $R(u) = T(u)$ $\forall u \in U$.) By the inductive hypothesis, $\exists$ a basis $(u_1, \dots, u_m)$ of U s.t. $\mathcal{M}(R)$ is upper triangular. By Prop. 1

we deduce $T(u_j) = R(u_j) \in \operatorname{span}(u_1, \dots, u_j)$ $\forall j = 1, \dots, n$. Enlarge $(u_1, \dots, u_m)$ to a basis $(u_1, \dots, u_m, v_1, \dots, v_n)$ of V. $\forall k = 1, \dots, n$, $T(v_k) = (T - \lambda I)(v_k) + \lambda v_k$. But $(T-\lambda I)(v_k) \in U = \operatorname{span}(u_1, \dots, u_m)$. So $T(v_k) \in \operatorname{span}(u_1, \dots, u_m, v_1, \dots, v_k)$. By Prop. 1, $\mathcal{M}(T)$ is upper Δ'r. □

<u>Lemma 1</u>: If A is an upper triangular matrix, then A is invertible if and only if the numbers on its main diagonal are all non-zero.

proof:

Suppose $A = [\vec{A}_1 \mid \vec{A}_2 \mid \dots \mid \vec{A}_n] = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & \vdots \\ \vdots & 0 & \lambda_3 & \cdots & \vdots \\ \vdots & & 0 & \ddots & * \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{bmatrix}$

$*$ is some number. Some $\lambda_j = 0 \iff \vec{A}_j \in \operatorname{span}(\vec{A}_1, \dots \vec{A}_{j-1}) \underset{\substack{\text{Ch.1} \\ \text{Lemma 1}}}{\iff}$ columns of A are not L.I. $\iff \operatorname{rank}(A) < n$

$\iff A$ not invertible. □

<u>Lemma 2</u>: If $T \in \mathcal{L}(V)$, then T is invertible if and only if $\mathcal{M}(T)$ is an invertible matrix.

proof:

$(\Longrightarrow)$ Ch. 3 Cor. 7

$(\Longleftarrow)$ If $\mathcal{M}(T)B = B\mathcal{M}(T) = I_n$ write $B = \mathcal{M}(S)$ by Ch. 2 Thm 3, let $\theta: \mathcal{L}(V) \to M_n$ be the isomorphism in that theorem, then $\theta(ST) = \theta(S)\theta(T) = \mathcal{M}(S)\mathcal{M}(T) = B\mathcal{M}(T) = I_n = \theta(I_V) \xRightarrow{\theta \text{ 1-1}} ST = I_V$

$\theta(TS) = \theta(T)\theta(S) = \mathcal{M}(T)\mathcal{M}(S) = I_n = \theta(I_V) \Longrightarrow TS = I_V$.

So T is invertible, $T^{-1} = S$. □

<u>Proposition 2</u>: Suppose $T \in \mathcal{L}(V)$ and $\mathcal{B}$ is a basis for which $\mathcal{M}(T)$ is upper triangular then

(1) the e-values of T are the numbers on the main diagonal of $\mathcal{M}(T)$

(2) T is invertible if and only if none of the numbers on the main diagonal are zero. (the numbers referred to here are the e-values of T)

proof:

Note (2) follows from Lemma 1 and Lemma 2. T is invertible $\overset{\text{Lemma 2}}{\Longleftrightarrow}$ $\mathcal{M}(T)$ is invertible $\overset{\text{Lemma 1}}{\Longleftrightarrow}$ diagonal entries are non-zero.

(1) λ is an e-value of $T \Longleftrightarrow \ker(T-\lambda I) \neq 0 \Longleftrightarrow$ $T-\lambda I$ is not invertible $\overset{\text{Lemma 2}}{\Longleftrightarrow}$ $\mathcal{M}(T-\lambda I)$ is not invertible. But $\mathcal{M}(T-\lambda I) \overset{\text{Ch.3 Thm 3}}{=} \theta(T-\lambda I) = \theta(T) - \lambda\,\theta(T)$ $= \mathcal{M}(T) - \lambda \mathcal{M}(I) = \mathcal{M}(T) - \lambda I_n$. So λ is an e-value of T. $\Longleftrightarrow \mathcal{M}(T) - \lambda I_n$ is not an invertible matrix. Since $\mathcal{M}(T) - \lambda I_n$ is also upper Δ'r matrix. So by Lemma 1, latter happens. $\Longleftrightarrow$ at least one number on the main diagonal of $\mathcal{M}(T) - \lambda I_n$ is 0. $\Longleftrightarrow$ $\lambda =$ one of the entries on the main diagonal of $\mathcal{M}(T)$. $\square$

$$\mathcal{M}(T) - \lambda I_n = \begin{bmatrix} a_1 & * & \cdots & \cdots & * \\ 0 & a_2 & * & \cdots & \vdots \\ \vdots & 0 & a_3 & * & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & 0 & a_n \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & & & \vdots \\ \vdots & & \lambda & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix} =$$

$$\begin{bmatrix} a_1-\lambda & * & \cdots & \cdots & * \\ 0 & a_2-\lambda & & & \vdots \\ \vdots & & a_3-\lambda & & \vdots \\ \vdots & & & \ddots & * \\ 0 & \cdots & \cdots & 0 & a_n-\lambda \end{bmatrix}$$

Proposition 3: If $T \in \mathcal{L}(V)$ and if B is a basis for V then $\mathcal{M}(T) = \mathcal{M}(T, B, B)$ is a diagonal matrix if and

only if $\mathcal{B}$ consists entirely of e-vectors of T.
"Summary: you can make $\mathcal{M}(T)$ diagonal iff you can find a basis consisting entirely of e-vectors of T."

proof:

Very minor adjustment to the proof of Prop. 1 (Exercise) $\square$

$T(v_j) = 0v_1 + 0v_2 + \ldots + 0v_{j-1} + cv_j + 0v_{j-1} + \ldots + 0v_n$ is the main change in proof of Prop 1.

orollary 3: If $T \in \mathcal{L}(V)$ has $\dim(V)$ distinct e-values then $\exists$ basis with respect to which $\mathcal{M}(T)$ is a diagonal matrix.

proof:

Let $\lambda_1, \ldots, \lambda_n$ be distinct e-values, $n = \dim V$, let v_k be e-vector corresponding to λ_k. By Thm. 1, $(v_1, \ldots, v_n)$ is L.I. By Ch. 2 Prop. 7, this is a basis. So $\mathcal{M}(T)$ is diagonal by Prop. 3. $\square$

ome back to rest of Ch. 5 if we have time. Ch. 5 + 4 HW due 6/18

HAPTER 6 Inner product spaces

nner product: $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$

on $\mathbb{R}^3$ is just $\left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\rangle$ = dot product = $xa + yb + zc$

$$\left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\rangle = 3 - 2 = 1$$

on $\mathbb{C}^3$

$$\left\langle \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\rangle = x\bar{a} + y\bar{b} + z\bar{c}$$

DAVID BLECHER

6/16

An example illustrating some important things from end of Ch. 5. HW Ch. 5 due Wed. 6/18 (handout + extra question on website) Ch. 4 due as well.

Example: Let $T: M_2(\mathbb{R}) \to M_2(\mathbb{R})$ be given by $T(A) = \begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix} A$

Lets first find the e-values/e-vectors of T. Seek $\lambda \in \mathbb{R}$ s.t. $\begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix} A = \lambda A$ for some nonzero $A \in M_2$

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ this last equation becomes $\begin{bmatrix} 1 & -1 \\ -4 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$= \begin{bmatrix} a-c & b-d \\ -4(a-c) & -4(b-d) \end{bmatrix} = \begin{bmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{bmatrix} \iff \begin{cases} a-c = \lambda a & ① \\ b-d = \lambda b & ② \\ -4(a-c) = \lambda c & ③ \\ -4(b-d) = \lambda d & ④ \end{cases}$$

so $4①+③$: $0 = 4\lambda a + \lambda c = \lambda(4a+c)$

so either $\lambda = 0$ or $c = -4a$

$4②+④$: either $\lambda = 0$ or $d = -4b$

If $\lambda = 0$, then $a = c$ and $b = d$ so $A = \begin{bmatrix} a & b \\ a & b \end{bmatrix} =$

$\begin{bmatrix} a & 0 \\ a & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & b \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. so e-value $\lambda = 0$ has $\left(\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right)$ as basis for e-space E_0.

If $\lambda \neq 0$, then $c = -4a$ and $d = -4b$ and

$\left.\begin{array}{l} ① \text{ becomes } 5a = \lambda a \\ ② \text{ becomes } 5b = \lambda b \end{array}\right\}$ $\lambda = 5$ or $a = \cancel{0} = b$ impossible

$a = 0 = b$ impossible since else $A = 0$.

If $\lambda = 5$: then $A = \begin{bmatrix} a & b \\ -4a & -4b \end{bmatrix} = \begin{bmatrix} a & 0 \\ -4a & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & -4b \end{bmatrix} =$

$a\begin{bmatrix}1&0\\-4&0\end{bmatrix}+b\begin{bmatrix}0&1\\0&-4\end{bmatrix}$ so e-value $\lambda=5$ has

$\left(\begin{bmatrix}1&0\\-4&0\end{bmatrix},\begin{bmatrix}0&1\\0&-4\end{bmatrix}\right)$ as a basis for e-space E_5.

The list $B=\left(\overset{u_1}{\begin{bmatrix}1&0\\1&0\end{bmatrix}},\overset{u_2}{\begin{bmatrix}0&1\\0&1\end{bmatrix}},\overset{u_3}{\begin{bmatrix}1&0\\-4&0\end{bmatrix}},\overset{u_4}{\begin{bmatrix}0&1\\0&-4\end{bmatrix}}\right)$ are all

e-vectors of T, and they are L.I. (no one is a linear combination of the others) Since $\dim(M_2)=4$, they form a basis of M_2. By Prop. 3, $M(T)=M(T,B,B)$ is a diagonal matrix. Lets compute it.

$T\left(\begin{bmatrix}1&0\\1&0\end{bmatrix}\right)=0=0u_1+0u_2+0u_3+0u_4$ so $M(T)=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&5&0\\0&0&0&5\end{bmatrix}$

$T\left(\begin{bmatrix}0&1\\0&1\end{bmatrix}\right)=0=0u_1+0u_2+0u_3+0u_4$

$T\left(\begin{bmatrix}1&0\\-4&0\end{bmatrix}\right)=5\begin{bmatrix}1&0\\-4&0\end{bmatrix}=0u_1+0u_2+5u_3+0u_4$

$T\left(\begin{bmatrix}0&1\\0&-4\end{bmatrix}\right)=5\begin{bmatrix}0&1\\0&-4\end{bmatrix}=0u_1+0u_2+0u_3+5u_4$

So $T(A)=\begin{bmatrix}1&-1\\-4&4\end{bmatrix}A$ transformation on M_2 "is just"

$\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&5&0\\0&0&0&5\end{bmatrix}$. What "is just" means is that any computation about T can be done instead by the 4×4 matrix which is much simpler via the equation we proved in Ch. 3 Prop. 6 $[T(A)]_B=M(T)[A]_B$ or $[T(A)]_B=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&5&0\\0&0&0&5\end{bmatrix}[A]_B$

Eg. Compute $T\left(\begin{bmatrix}2&0\\-3&0\end{bmatrix}\right)$, using above

$A=\begin{bmatrix}2&0\\-3&0\end{bmatrix}=1u_1+1u_3=[A]_B=\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$ so $[T(A)]_B=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&5&0\\0&0&0&5\end{bmatrix}\begin{bmatrix}1\\0\\1\\0\end{bmatrix}=$

$$\begin{bmatrix}0\\0\\5\\0\end{bmatrix} \quad \text{So} \quad T(A) = 0u_1 + 0u_2 + 5u_3 + 0u_4 = 5\begin{bmatrix}1 & 0\\-4 & 0\end{bmatrix} = \begin{bmatrix}5 & 0\\-20 & 0\end{bmatrix}$$

CHAPTER 6 Inner product spaces

Definition: An <u>inner product</u> on a v.s. V is a function $\langle \cdot, \cdot \rangle$ of 2 variables so that if $v, w \in V$ then $\langle v, w \rangle \in \mathbb{F}$ and

"positive definite" $\begin{cases} \text{(i) } \langle v, v \rangle \geq 0 \quad \forall v \in V \\ \text{(ii) } \langle v, v \rangle = 0 \iff v = 0 \end{cases}$

"linear in 1st variable" $\begin{cases} \text{(iii) } \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle \quad \forall v_1, v_2, w \in V \\ \text{(iv) } \langle cv, w \rangle = c\langle v, w \rangle \quad \forall c \in \mathbb{F} \end{cases}$

(v) $\langle v, w \rangle = \overline{\langle w, v \rangle}$ (this reads $\langle v, w \rangle = \langle w, v \rangle$ if $\mathbb{F} = \mathbb{R}$)

An <u>inner product space</u> is a v.s. with an inner product.

Ex. 1. $\mathbb{R}^3$ and define $\langle \vec{v}, \vec{w} \rangle = \vec{v} \cdot \vec{w}$ (dot product)
$= v_1 w_1 + v_2 w_2 + v_3 w_3$

met (i) - (v) in Calc. 3 eg. (iii): $(\vec{v_1} + \vec{v_2}) \cdot \vec{w} = \vec{v}_1 \cdot \vec{w} + \vec{v}_2 \cdot \vec{w}$

More generally $\mathbb{R}^n$ is an inner product space with $\langle \vec{v}, \vec{w} \rangle = \vec{v} \cdot \vec{w} = \sum_{k=1}^{n} v_k w_k$.

$\mathbb{C}^n$ is an inner product space with $\langle \vec{v}, \vec{w} \rangle = \sum_{k=1}^{n} v_k \overline{w_k}$

note: $\langle \vec{v}, \vec{v} \rangle = \sum_{k=1}^{n} v_k \overline{v_k} = \sum_{k=1}^{n} |v_k|^2 \geq 0$

and $= 0 \iff |v_k| = 0 \ \forall k \iff v_k = 0 \ \forall k \iff \vec{v} = \vec{0}$

check <u>(iii) and (iv)</u> as exercise.

(v) $\langle \vec{v}, \vec{w} \rangle = \sum_{k=1}^{n} v_k \overline{w_k} = \overline{\sum_{k=1}^{n} \overline{v_k w_k}} = \overline{\sum_{k=1}^{n} \overline{v_k} w_k} = \overline{\langle \vec{w}, \vec{v} \rangle}$

Example 2: $V = \mathbb{F}^n$, $c_1, \dots, c_n > 0$, define $\langle \vec{v}, \vec{w} \rangle = \sum_{k=1}^{n} c_k v_k \overline{w}_k$

note: $\langle \vec{v}, \vec{v} \rangle = \sum_{k=1}^{n} c_k |v_k|^2 \geq 0$ and as above it is easy to see (ii) - (v).

Example 3: On $\mathcal{P}(\mathbb{F})$, define $\langle p, q \rangle = \sum_{k=0}^{n} a_k \overline{b}_k$ if $p = \sum_{k=0}^{n} a_k z^k$, $q = \sum_{k=0}^{n} b_k z^k$

$$\langle 1 + x^2, -3 + 2x - x^2 + x^3 \rangle = 1 \cdot (-3) + 0 \cdot (2) + 1 \cdot (-1) + 0 \cdot (1) = -4$$

* if energetic, check (i) - (v) *

Example 4: On $\mathcal{P}(\mathbb{F})$ define $\langle p, q \rangle = \int_0^1 p \overline{q} \, dx$. This also defines an inner product on $\{ f : [0,1] \to \mathbb{F} \text{ continuous} \}$

can take $\mathbb{F} = \mathbb{R}$ if you like

(i) $\langle p, p \rangle = \int_0^1 |p|^2 dx \geq 0$ (by Calc. 1)

(ii) $\langle p, p \rangle = 0 \iff \int_0^1 |p|^2 dx = 0 \overset{3333}{\iff} p = 0$

(iii) $\int_0^1 (f+g) \overline{q} \, dx = \int_0^1 f \overline{q} \, dx + \int_0^1 g \overline{q} \, dx = \langle f, q \rangle + \langle g, q \rangle$

(iv) $\langle cp, q \rangle = \int_0^1 cp\overline{q} \, dx = c \int_0^1 p \overline{q} \, dx = c \langle p, q \rangle$

(v) $\langle p, q \rangle = \int_0^1 pq \, dx = \int_0^1 qp \, dx = \langle q, p \rangle$ (in $\mathbb{F} = \mathbb{R}$)