6/24

Corollary 4: Let $T \in \mathcal{L}(V)$ be selfadjoint (or even normal if $\mathbb{F} = \mathbb{C}$). Let $\lambda_1, \dots, \lambda_n$ be the distinct e-values of T. Then $E_{\lambda_i} \perp E_{\lambda_j}$ $\forall i \neq j$. and $V = \bigoplus_{k=1}^{n} E_{\lambda_k}$

proof:

That $E_{\lambda_i} \perp E_{\lambda_j}$ for $i \neq j$ is immediate from Cor. 3. By the spectral theorem, $\exists$ basis $(u_1, \dots, u_n)$ of e-vectors of T. Since it is a basis, any $v \in V$ may be written $v = \sum_{k=1}^{n} c_k u_k$, $c_k \in \mathbb{F}$. Since each u_k is in some E_{λ_j} so is $c_k u_k$ and so $v \in E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_n}$. Suppose we can write in two ways $v = \sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k$ where $x_k, y_k \in E_{\lambda_k}$ $\forall k = 1, \dots, n$. Let $w \in E_{\lambda_j}$ then $\langle v, w \rangle = \sum_{k=1}^{n} \langle x_k, w \rangle = \langle x_j, w \rangle$ and similarly $\langle v, w \rangle = \langle y_j, w \rangle$. So $\langle x_j, w \rangle = \langle y_j, w \rangle$ $\forall w \in E_{\lambda_j}$, so $x_j = y_j$ as desired. $\square$

HW Ch. 6 (4), 6, 7, 9, (10), (14), 15, 16, (27), 28, 31, 32
Ch. 7 1, 2, (4), 6, (8), (9), 10, 11, 13
Ch. 8 (1)(2) (14)(16) 21, 22, 23, 26, 27

( ) will be graded

CHAPTER 8: Operators on a vector space over $\mathbb{C}$

Leave i.p.s., normal, self adjoint operators. Back to setting: V is f.d., nonzero, v.s. over $\mathbb{F}$ in chapter below. Usually $\mathbb{F} = \mathbb{C}$.

Definition: If $T \in \mathcal{L}(V)$ and λ is an e-value of T, then we say v is a generalized e-vector (gevector, for short) corresponding to λ if $(T-\lambda I)^j(v) = 0$ for some $j = 1, 2, \ldots$

Let $G_\lambda = \{$ gevectors corresponding to $\lambda \}$ this is the generalized eigenspace

Remark: $E_\lambda \subseteq G_\lambda$ that is every e-vector is a gevector (since for evector, $(T-\lambda I)^1(v) = 0$

Example: $T: \mathbb{R}^3 \to \mathbb{R}^3$ be $T(x,y,z) = (y,0,z)$

Find the generalized e-vectors.

soln:

there exists a shortcut, learn later.

E-values: $T(x,y,z) = (y,0,z) = \lambda(x,y,z)$

$$\Longrightarrow \begin{cases} y = \lambda x \\ 0 = \lambda y \\ z = \lambda z \end{cases}$$

If $\lambda = 0 \Longrightarrow y = z = 0 \longrightarrow E_0 = \mathbb{F}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

If $\lambda \neq 0 \Longrightarrow y = 0,\ x = 0,$ and $\lambda = 1$ or $z \neq 0$

so $E_1 = \mathbb{F}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Lets look for gevectors corresponding to $\lambda = 0$: solve $(T - 0I)^j(v) = 0$ $j = 2, 3, \ldots$

When $j = 2 \Longrightarrow T^2(x,y,z) = T(y,0,z) = (0,0,z) = (0,0,0) \Longrightarrow z = 0$. So any $(x,y,0)$ is a gevector corresponding to $\lambda = 0$.

(Note $j = 3, 4, \ldots$ doesn't happen, $T^2 = T^3 = \ldots$ in this example)

Lets look for gevectors corresponding to $\lambda = 1$:

Solve $(T-I)^j(x,y,z) = (0,0,0)$

$j = 2 \Longrightarrow (T-I)(x,y,z) = (y-x, -y, 0)$ So $(T-I)^2(x,y,z) = (T-I)(y-x, -y, 0) = (-y-(y-x), y, 0) = (0,0,0) \Longrightarrow$

$y = 0 = x$ so the gevector is $(0,0,z) \in E_1$

$j = 3 \Longrightarrow (T-I)^3(x,y,z) = (\qquad\qquad) \longrightarrow$ nothing new and similarly for $j = 4, 5, \ldots$

Summarizing: $G_0 = \{(x, y, 0) : x, y \in \mathbb{F}\} = \text{Span}(\vec{i}, \vec{j})$.
$G_1 = \{(0, 0, z) : z \in \mathbb{F}\} = \text{Span}(\vec{k}) = E_1$

Note: $G_0 \oplus G_1 = \mathbb{F}^3$, $\exists$ "enough" (a basis of) generalized e-vectors.

<u>Fact</u>: If $T \in \mathcal{L}(V)$ then $\ker(T^k) \subseteq \ker(T^{k+1})$ so
$(0) = \ker(T^0) \subseteq \ker(T^1) \subseteq \ker(T^2) \subseteq \ldots$

proof:
$v \in \ker(T^k) \Longrightarrow T^k v = 0 \Longrightarrow T^{k+1} v = T(T^k v) = T(0) = 0$. $\square$

Similarly, $\text{Ran}(T^k) \supseteq \text{Ran}(T^{k+1})$ since $T^{k+1}(v) = T^k(Tv)$

Proposition 1: If $T \in \mathcal{L}(V)$ and $\ker(T^m) = \ker(T^{m+1})$, then $\ker(T^m) = \ker(T^{m+k})$ $\forall k = 1, 2, \ldots$

proof:

If $k = 1, 2, \ldots$ then if $v \in \ker T^{m+k+1}$ we have
$0 = T^{m+k+1} v = T^{m+1}(T^k v) \Longrightarrow T^k v \in \ker T^{m+1}$
So $T^k v \in \ker T^m \Longrightarrow T^{m+k} v = T^m(T^k v) = 0 \Longrightarrow$
$v \in \ker T^{m+k}$ so $\ker T^{m+k+1} = \ker T^{m+k}$ $\forall k = 1, 2, \ldots$

Proposition 2: If $T \in \mathcal{L}(V)$, $n = \dim V$, then $\ker(T^n) = \ker(T^{n+1}) = \ker(T^{n+2}) = \ldots$

proof:

Suppose $\ker(T^n) \neq \ker(T^{n+1})$. By Prop. 1, $\ker(T^k) \neq \ker(T^{k+1})$ for $k = 1, \ldots, n$. So $0 = \ker(T^0) \subsetneq \ker(T^1) \subsetneq \ker(T^2) \subsetneq \ldots \subsetneq \ker(T^n) \subsetneq \ker(T^{n+1})$
So $0 = \dim \ker(T^0) < \dim \ker(T^1) < \dim \ker(T^2) < \ldots < \dim \ker(T^{n+1})$. So the last place has dimension $\geq n+1$, which is impossible inside an n-dimensional space. Contradiction. So $\ker(T^n) = \ker(T^{n+1})$ and this

equals later terms by Prop. 1. $\square$

Corollary 1: If $T \in \mathcal{L}(V)$ and λ is an e-value of T, then $G_\lambda = \ker((T-\lambda I)^{\dim V})$

proof:

The generalized e-values are the vectors inside one of the following:

$\ker(T-\lambda I) \subseteq \ker(T-\lambda I)^2 \subseteq \ldots \subseteq \ker(T-\lambda I)^{\dim V} \underset{=}{\overset{\text{Prop 2}}{=}} \ker((T-\lambda I)^{(\dim V)+1}) = \ldots \; \square$

Proposition 3: $T \in \mathcal{L}(V) \Longrightarrow Ran(T^{\dim V}) = Ran(T^{(\dim V)+1}) = Ran(T^{(\dim V)+2}) = \ldots$

proof:

Let $m = \dim V + 1, \dim V + 2, \ldots$ By Ch. 3 Thm 1

$$\begin{aligned} \dim(Ran(T^m)) &= \dim V - \dim(\ker(T^m)) \\ &= \dim V - \dim(\ker(T^{\dim V})) \text{ by Prop. 2.} \\ &= \dim(Ran(T^{\dim V})) \quad \square \end{aligned}$$

<u>Theorem</u> 1: If $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{F}$, and suppose we fix a basis for V for which $\mathcal{M}(T)$ is upper triangular, then the number of times λ appears on the main diagonal of $\mathcal{M}(T)$ equals $\dim(G_\lambda)$.

Definition: If λ is an e-value of T, we call the number in the theorem, the <u>multiplicity</u> of λ. ($= \dim G_\lambda$)

Example: Revisit previous example

$T(x,y,z) = (y, 0, z)$. Let $\mathcal{B} = (\vec{k}, \vec{i}, \vec{j})$; notice

$T(\vec{k}) = \vec{k}$, $T(\vec{i}) = 0$, $T(\vec{j}) = \vec{i} = 0\vec{k} + 1\vec{i} + 0\vec{j} \Longrightarrow \mathcal{M}(T, \mathcal{B}, \mathcal{B}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

By Thm. 1, Eigenvalue, $\lambda=1$, has multiplicity 1 $= \dim(G_1)$ (see previous ex.)
Eigenvalue 0 has multiplicity $2 = \dim(G_2)$

proof of Thm. 1:

<u>Observation</u> 1: If the theorem is true when $\lambda=0$ then it's true for any λ. This is because, case $\lambda=0$ says: number of times 0 appears on main diagonal of $\mathcal{M}(T) = \dim(\ker(T^{\dim V}))$.
Replace T by $T-\lambda I$: the number of times 0 appears on diagonal of $\mathcal{M}(T-\lambda I) = \Theta(T-\lambda I) = \Theta(T) - \lambda\,\Theta(I) = \mathcal{M}(T) - \lambda I_n$, equals $\dim(\ker(T-\lambda I)^{\dim V}) = \dim(G_\lambda)$. That is $\dim(G_\lambda)$ equals the number of times 0 appears on diagonal of

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots \\ 0 & a_{22} & & \\ \vdots & & a_{33} & \\ \vdots & & & \ddots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}}_{\mathcal{M}(T)} - \lambda \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & 1 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11}-\lambda & \cdots & \cdots & * \\ 0 & a_{22}-\lambda & & \vdots \\ & & a_{33}-\lambda & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn}-\lambda \end{bmatrix}$$

But this equals the number of times λ appears on main diagonal of $\mathcal{M}(T)$. This finishes proof of observation.

Henceforth we can assume $\lambda = 0$.
We proved the theorem by induction on n, where $n = \dim V$. Check case $\dim V = 1$ yourself.
Assume the theorem is true for all spaces

of dimension k. Now let V be a space of dim $k+1$, $T \in \mathcal{L}(V)$, $\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & \cdots & & * \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{k+1} \end{bmatrix}$

upper triangular with respect to basis $(v_1, \ldots, v_{k+1})$. Let $U = \text{span}(v_1, \ldots, v_k)$, and let $R = T/U$ (Note: $Tv_j \in \text{span}(v_1, \ldots, v_{j-1})$ by Ch. 5 Prop. 1 so $T(U) \subseteq U$) and $\mathcal{M}(R) = \begin{bmatrix} \lambda_1 & \cdots & & * \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_k \end{bmatrix}$ $(*)$

By the inductive hypothesis 0 appears on diagonal of $(*)$, $\dim(\ker(R^k))$ times. Now $\dim(\ker(R^k)) = \dim(\ker(R^{k+1}))$ by Prop. 2 so 0 appears on diagonal of $(*)$ $\dim(\ker(R^{k+1}))$ times. I will assume for simplicity that $\lambda_{k+1} \neq 0$ (case $\lambda_{k+1} = 0$ is slightly different, see Axler). Claim: $\ker T^{k+1} \subseteq U$. If this claim is true we are done because: then if $T^{k+1}x = 0$ $\Longrightarrow x \in U \Longrightarrow x \in \ker(R^{k+1})$, so $\ker(T^{k+1}) = \ker(R^{k+1})$ so by above, 0 appears on diagonal of $\mathcal{M}(T)$ exactly $\dim(\ker(T)^{k+1})$ times.

Ending the proof of claim: $\mathcal{M}(T^n) = \mathcal{M}(T)^n$ (think in terms of θ) $= \begin{bmatrix} \lambda_1^n & \cdots & & * \\ 0 & \lambda_2^n & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n^n \end{bmatrix}$, $n = k+1$

This implies by looking at last column that $T^n v_n = \lambda_n^n v_n + u$, some $u \in U$. If $v \in \ker T^n$, $v = \hat{u} + c v_n$, $\hat{u} \in U$ so $0 = T^n v = T^n \hat{u} + c T^n v_n = T^n \hat{u} + cu + c\lambda_n^n v_n \Longrightarrow c = 0 \Longrightarrow v \in U$. $\square$

'25 G_λ = generalized e-space corresp. to e-value λ = $\ker((T-\lambda I)^{\dim V})$

Thm. 1 : $T \in \mathcal{L}(V)$, $\mathcal{M}(T)$ upper Δ with respect to some basis. Then $\dim(G_\lambda)$ = # of times λ appears on main diagonal of $\mathcal{M}(T)$

Prop. 4: V a complex v.s., $T \in \mathcal{L}(V) \Longrightarrow \dim V$ = sum of the multiplicities of the e-values of T.

proof:

By Ch.5 Thm 3, $\exists$ a basis with respect to which $\mathcal{M}(T)$ is upper Δ. By Thm. 1^1, or def. of multiplicity, the sum of the multiplicities of e-values equals the number of elements on the diagonal of $\mathcal{M}(T) = \dim V$. □

Definition: If V is a v.s. over $\mathbb{C}$, $T \in \mathcal{L}(V)$ whose distinct e-values are $\lambda_1, \dots, \lambda_m$, we define the characteristic polynomial of T to be $p(z) = (z-\lambda_1)^{d_1}(z-\lambda_2)^{d_2} \dots (z-\lambda_m)^{d_m}$ where d_k = multiplicity of λ_k

Remarks:

1) You may have seen different def. in 2331, later we'll prove they are the same.
2) The characteristic poly. has degree $\dim V$ by Prop. 4.
3) " " " " roots which are precisely the e-values of T.
4) If $\mathcal{M}(T) = \begin{bmatrix} c_1 & & * \\ & c_2 & \\ & \ddots & \\ 0 & & c_n \end{bmatrix}$ (upper Δ'r), then characteristic polynomial is $(z-c_1)(z-c_2)\dots(z-c_n)$.

Example: (from last class) $T(x,y,z) = (y,0,z)$ we saw if $B = (\vec{k}, \vec{i}, \vec{j})$ then $\mathcal{M}(T,B,B) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. So from 4)

characteristic poly. of T is $(z-1)z^2 = z^3 - z^2$

Theorem 2: (Cayley - Hamilton Theorem) If T is a v.s. over $\mathbb{C}$, $T \in \mathcal{L}(V)$ and p is the characteristic poly. of T then $p(T) = 0$.

proof:

Suppose $\mathcal{M}(T)$ is upper Δ'r with respect to basis $(v_1, \dots, v_n)$. Say $\mathcal{M}(T) = \begin{bmatrix} c_1 & & * \\ & c_2 & \\ & \ddots & \\ 0 & & c_n \end{bmatrix}$ If $p(T) v_k = 0$

$\forall k$ then for any $v \in V$, $v = \sum_{k=1}^{n} c_k v_k$ So $p(T)(v_k) =$

$\sum c_k p(T) v_k = \sum c_k 0 = 0$ and so $p(T) = 0$. We need to show $p(T) v_k = 0$ $\forall k$; that is $(T - c_1 I)(T - c_2 I) \dots (T - c_n I) v_k = 0$ $\forall k = 1, \dots, n$. using Remark 4) which follows if $(T - c_1 I) \dots (T - c_k I) v = 0$ $\forall k = 1, \dots, n$. We prove this by induction on k. For $k = 1$, we are asking if $(T - c_1 I) v_1 = Tv - c_1 v_1 = 0 \iff T v_1 = c_1 v_1$ which is just what first column of $\mathcal{M}(T)$ says. Now suppose statement is true for $k \leq j-1$. For $k = j$, note $\mathcal{M}(T - c_j I) = \mathcal{M}(T) - c_j I = \begin{bmatrix} c_1 - c_j & & * \\ & c_2 - c_j & \\ & \ddots & \\ 0 & & c_n - c_j \end{bmatrix}$ has 0 in

the j-j entry, which by definition of (j^{th} column of) $\mathcal{M}(T - c_j I)$ says $(T - c_j I)(v_j) \in \text{span}(v_1, \dots, v_{j-1})$ so $(T - c_j I) v_j = \sum_{i=1}^{j-1} d_i v_i$ and hence $(T - c_1 I) \dots (T - c_{j-1} I)(T - c_j I) v_j =$

$(T - c_1 I) \dots (T - c_{j-1} I)\left(\sum^{j-1} d_k v_k\right) = 0$ by case $k \leq j-1$. $\square$

Definition: If $T \in \mathcal{L}(V)$ and U is a subspace of V s.t. $T(U) = \{T(u) : u \in U\} \subseteq U$, then we say U is T-invariant, or say U is an <u>invariant subspace</u> for T.

Prop. 5 : If $T \in \mathcal{L}(V)$, p a polynomial, then $\ker p(T)$ is T-invariant.

proof:

Let $v \in \ker p(T) \Longrightarrow p(T)v = 0 \Longrightarrow p(T)(Tv) = Tp(T)v = T0 = 0$. So $Tv \in \ker p(T)$. We've shown $T(\ker p(T)) \subseteq \ker p(T)$. $\square$

Theorem 3 : If V is a v.s. over $\mathbb{C}$, $T \in \mathcal{L}(V)$ whose distinct e-values are $\lambda_1, \dots, \lambda_m$

(a) $V = G_{\lambda_1} \oplus G_{\lambda_2} \oplus \dots \oplus G_{\lambda_m}$

(b) each G_{λ_j} is T-invariant

proof (b)

Let $p(z) = (z - \lambda_j)^{\dim V}$, then $p(T) = (T - \lambda_j I)^{\dim V}$. By Prop. 5 the $\ker p(T) = \ker (T - \lambda_j I)^{\dim V} = G_{\lambda_j}$ is T-invariant.

(a)

Prop. 4 says $\dim(V) = \sum_{k=1}^{m} \dim(G_{\lambda_k})$. Let $U = G_{\lambda_1} + G_{\lambda_2} + \dots + G_{\lambda_m}$. Then U is T-invariant by (b), (since if $u = \sum x_k$, $x_k \in G_{\lambda_k} \Rightarrow T(u) = \sum_k T(x_k) \in U$ since $T(x_k) \in G_{\lambda_k}$ by (b)) Let $R \in \mathcal{L}(U)$ be defined by $R = T/_U$. It is easy to see R and T have same e-values, gevectors and gespaces. For example, if $v \in G_{\lambda_k} = \ker((T - \lambda_k I)^{\dim V})$ then $v \in U$, so $(R - \lambda_k I)^{\dim V}(v) = (T - \lambda_k I)^{\dim V}(v) = 0$, so $v \in \ker((R - \lambda_k I)^{\dim V}) = k^{th}$ gespace for R. The other direction is easier, so R and T have same k^{th} gespace. By Prop. 4, apply to R, $\dim(U) = \sum_{k=1}^{m} \dim(G_{\lambda_k})$. So $\dim U = \dim V$ so $U = V$ by HW 2,

so $V = G_{\lambda_1} + \dots + G_{\lambda_m}$. By Ch. 2 Prop. 8 $V = G_{\lambda_1} \oplus \dots \oplus G_{\lambda_m}$ $\square$

Corollary 2 : V is v.s. over $\mathbb{C}$ $T \in \mathcal{L}(V)$ then V has a basis consisting of gevectors of T.

proof:

Let $(u_1, \dots, u_{n_1})$ be a basis for G_{λ_1}, $(u_{n_1+1}, \dots, u_{n_2})$ be a basis for G_{λ_2}, etc. Then when you put them together $(u_1, \dots, u_{n_1}, u_{n_1+1}, \dots, u_{n_2}, \dots, u_{n_m})$ is a set of gevectors. As in proof of Ch. 2 Prop. 8, it is also a basis of V. $\square$

Corollary 3: If V is a v.s. over $\mathbb{C}$ and $T \in \mathcal{L}(V)$, $\exists$ a basis of V with respect to which $\mathcal{M}(T) =$

$$\begin{bmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & & & \vdots \\ \vdots & & A_3 & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_m \end{bmatrix}, \text{ where } A_j = \begin{bmatrix} \lambda_j & & * \\ & \lambda_j & \\ & & \ddots & \\ 0 & & & \lambda_j \end{bmatrix}$$

*"block diagonal"

(upper Δ'r with λ_j on each diagonal entry)

$\lambda_1, \dots, \lambda_m$ are the distinct e-values of T.

proof:

~~Similar to proof of Cor. 2. Let $R_j = T/G_{\lambda_j}$ (remember~~ by Thm 3 (b), G_{λ_j} is T-invariant. So $R_j \in \mathcal{L}(G_{\lambda_j})$ so by Ch. 5 Thm. 3, $\exists$ a basis B_j for G_{λ_j} s.t. $\mathcal{M}(R_j, B_j, B_j)$ is upper Δ'r $= \begin{bmatrix} \lambda_j & & * \\ & \ddots & \\ 0 & & \lambda_j \end{bmatrix}$ Put these bases together as in proof of Cor. 2 and then

CHECK $\longrightarrow$ it's easy to check by def. of $\mathcal{M}(T)$ that $\mathcal{M}(T)$ is of desired form. $\square$

So for example, if $v \in B_j$, note $T(v) \in G_{\lambda_j}$ by last theorem (b), so the column of $\mathcal{M}(T)$ corresponding to v has 0's in all positions corresponding to G_{λ_i} for $i \neq j$ and in positions corresponding to G_{λ_j} we get exactly what we got in $\mathcal{M}(R_j)$ above.

In fact, with a little more work (see Axler) you can improve Corollary 3 to get $\exists$ a basis for which $\mathcal{M}(T) =$

$$\begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{bmatrix} \quad \text{but} \quad A_j = \begin{bmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda_j \end{bmatrix}$$

This is called the Jordan normal form

CHAPTER 10: Matrices

• Change of basis matrices

Definition: Let B, C be two basis for a v.s. V. The 'change of basis matrix' is $\mathcal{M}(I_V, B, C)$

example 1:

$V = \mathbb{R}^2$, $B = (\vec{i}, \vec{j})$, $C = \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$

$$\mathcal{M}(I, C, B) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \mathcal{M}(I, B, C) = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix} \underset{\text{by Ch.3 Prop.5}}{=} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

example 2:

$V = \mathbb{F}^n$, $S = (\vec{e}_1, \ldots, \vec{e}_n) =$ standard basis, C another basis.

Just as in example 1, we get $\mathcal{M}(I, C, S) =$ matrix whose columns are the vectors in C.

Definition: Matrices $A, B \in M_n$ are called <u>similar</u>, if $\exists$ invertible $Q \in M_n$ s.t. $A = Q^{-1} B Q$ (or equivalently, $B = Q A Q^{-1}$)

Theorem 1: Let B, C be two bases for V, $T \in \mathcal{L}(V)$, then $\boxed{\mathcal{M}(T, B, B) = Q^{-1} \mathcal{M}(T, C, C)\, Q}$ where $Q = \mathcal{M}(I, B, C)$

proof:

By Ch.3 Prop.5 + Cor.7, $Q^{-1} \mathcal{M}(T, C, C) Q = \mathcal{M}(I, C, B) \mathcal{M}(T, C, C) \mathcal{M}(I, B, C)$
$= \mathcal{M}(I T I, B, B) = \mathcal{M}(T, B, B)$ $\square$

Prop. 1: If B, C are o.n.b. of an i.p.s. V then $Q = \mathcal{M}(I, B, C)$ is unitary (i.e. $Q^{-1} = Q^*$, or what is the same, $QQ^* = Q^*Q = I_n$)

proof:

By Ch. 6 Prop. 5, $Q^*Q = \mathcal{M}(I,B,C)^*\mathcal{M}(I,B,C) = \mathcal{M}(I,C,B)\mathcal{M}(I,B,C) \overset{\text{Ch. 3 Prop 5}}{=} \mathcal{M}(I) = I_n$ Similarly $QQ^* = I_n$. $\square$

6/26

- A, B in M_n are <u>similar</u> if $A = Q^{-1} B Q$ some invertible $Q \in M_n$
- U <u>unitary</u> if $U^* = U^{-1}$ $(U \in M_n)$

Definition: Say A and B are <u>unitary equivalent</u> if $\exists$ unitary $U \in M_n$ s.t. $A = U^* B U$ (or equivalently, $B = U A U^*$)

Rephrase some earlier results in terms of matrices:

Cor. 1: Every $A \in M_n(\mathbb{C})$ is unitarily equivalent to an upper Δ'r matrix.

proof:

Let $T\vec{x} = A\vec{x}$ for $\vec{x} \in \mathbb{C}^n$, then $\mathcal{M}(T,S,S) = A$ (see Ch. 3), where S = standard basis. By Ch 6 Cor. 6., $\exists$ an o.n.b. B with $\mathcal{M}(T,B,B)$ with $\mathcal{M}(T)$ being upper Δ'r. By Prop. 1 $\mathcal{M}(I,C,B)$ is unitary, call it U, by Thm. 1, $A = \mathcal{M}(T,S,S) = U^{-1}\mathcal{M}(T,B,B)U$. $\square$

Cor. 2: (Spectral Theorem for normal matrices) If $A \in M_n(\mathbb{C})$ then A is unitarily equivalent to a diagonal matrix iff $A^*A = AA^*$.

proof: $(\Rightarrow)$ If $A = U^* D U$, D diagonal, U unitary, then $A^*A = (U^*DU)^* U^*DU$
$= U^*D^*UU^*DU = U^*DD^*U = AA^*$ (using facts at end of CH6)

$(\Leftarrow)$ follow proof of Cor. 1 above, $\mathcal{M}(T,S,S) = A$, then $A^*A = AA^* \Rightarrow T^*T = TT^*$ by simpler version of the argument in $\mathbb{C}$-spectral theorem of Ch. 7 $((2) \Rightarrow (1))$. By that result, $\exists$ an o.n.b B s.t. $\mathcal{M}(T,B,B)$ is diagonal and then $A = U^{-1}\mathcal{M}(T,B,B)U$ as in proof of Cor. 1 □

Cor. 3: If $A \in M_n(\mathbb{F})$ then A is unitarily equivalent to a diagonal matrix with real numbers on diagonal iff $A = A^*$ (selfadjoint matrix).

proof: ($\Leftarrow$)
Similar to proof of Cor. 2, but use Thm. 2 Ch. 7 (spectral theorem for self-adjoint maps.)
($\Rightarrow$) If $A = U^* D U$, U unitary, D diagonal then
$A^* = (U^* D U)^* = U^* D^* U = U^* D U = A$. □

Cor. 4: Every $A \in M_n(\mathbb{C})$ is similar to a matrix of form
$$\begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A_n \end{bmatrix} \text{ with each } A_k = \begin{bmatrix} \lambda_k & & * \\ & \lambda_k & \\ 0 & & \ddots \\ & & & \lambda_k \end{bmatrix}$$

proof:
Similar to proof for Cor. 1 but use Ch. 8 Cor. 3 in place of Ch. 6 Cor. 6. □

Examples: See 2433 (prereq)

- V is a f.d. v.s. over $\mathbb{C}$, $V \neq (0)$. Recall the characteristic polynomial of $T \in \mathcal{L}(V)$, this is $p(z) = (z-\lambda_1)(z-\lambda_2)\ldots(z-\lambda_n)$. Here V is n-dimensional, $\lambda_1, \ldots, \lambda_n$ are the e-values, repeated according to multiplicity. Multiplying,
$p(z) = z^n - (\lambda_1 + \lambda_2 + \ldots + \lambda_n) z^{n-1} + \ldots + (-1)^n \lambda_1 \lambda_2 \ldots \lambda_n$

The negative of the coefficient of z^{n-1} is called $tr(T)$ in this case $= \lambda_1 + \lambda_2 + \dots + \lambda_n$.
The last (constant) term in $p(z)$, multiplied by $(-1)^n$ is called the $\det(T)$, in this case $\det(T) = \lambda_1 \lambda_2 \dots \lambda_n$

- Back to square matrices $A \in M_n$, define $\text{trace}(A) = \sum_{i=1}^{n} a_{ii}$, if

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & a_{22} & & \vdots \\ \vdots & & a_{33} & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}$$

Lemma 1: $\text{trace}(AB) = \text{trace}(BA)$ if $A, B \in M_n$

proof:

Let $A = [a_{ij}]$, $B = [b_{ij}]$, then $AB = [\sum_{k=1}^{n} a_{ik} b_{kj}]$, $BA = [\sum_{k=1}^{n} b_{ik} a_{kj}]$. So $\text{trace}(AB) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki}$

$= \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} a_{ik} = \text{trace}(BA)$. $\square$

Prop. 2: If $T \in \mathcal{L}(V)$ then $tr(T) = \text{trace}(\mathcal{M}(T, B, B))$ for any basis B for V.

proof:

By Ch. 5 Thm. 3, + Ch. 8 $\exists$ basis C s.t. $\mathcal{M}(T, C, C) = \begin{bmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}$

Notice $\text{trace}(\mathcal{M}(T, C, C)) = \lambda_1 + \dots + \lambda_n = tr(T)$. If B is any basis for T, then by Thm. 1, $\mathcal{M}(T, B, B) = A^{-1} \mathcal{M}(T, C, C) A$, some $A \in M_n$. So $\text{trace}(\mathcal{M}(T, B, B)) = \text{trace}(A^{-1}(\mathcal{M}(T, C, C))A)$
$\overset{\text{Lemma 1}}{=}$ $\text{trace}(\mathcal{M}(T, C, C)\cancel{AA^{-1}}) = \text{trace}(\mathcal{M}(T, C, C)) = tr(T)$. $\square$

- Determinants

Prop. 3: $T \in \mathcal{L}(V)$ is invertible iff $\det(T) \neq 0$.

proof: T is invertible iff 0 is not an e-value (ie. $\exists\, v$ s.t. $Tv \neq 0 \iff \ker(T) \neq (0)$). T is invertible iff:
- product of e-values is not 0.
- $\det(T) \neq 0$. $\square$

Theorem 2: If $T \in \mathcal{L}(V)$ then the characteristic poly. of T is $\det(zI - T)$

proof:
Let $\lambda_1, \ldots, \lambda_n$ be the e-values of T repeated according to multiplicity. Then the e-values of $zI - T$ are $z-\lambda_1, z-\lambda_2, \ldots, z-\lambda_n$, repeated according to multiplicity. [one way to see this: $\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

so $\mathcal{M}(zI-T) = z\mathcal{M}(I) - \mathcal{M}(T) = zI_n - \mathcal{M}(T) = \begin{bmatrix} z-\lambda_1 & & & * \\ & z-\lambda_2 & & \\ & & \ddots & \\ 0 & & & z-\lambda_n \end{bmatrix}$

and the numbers on the diagonal are the e-values repeated according to multiplicity] Hence by definition of $\det(zI-T) = (z-\lambda_1)(z-\lambda_2)\ldots(z-\lambda_n) =$ characteristic poly. of T. $\square$

END FOR TEST #2 MATERIAL **

Determinants of matrices

A *permutation* of $(1,2,\ldots,n)$ is a list $\sigma = (m_1, m_2, \ldots, m_n)$ that contains each of $1,\ldots,n$ exactly once. Eg. $\sigma = (3,2,4,1)$ is a permutation of $(1,2,3,4)$

The set of all permutations of $(1,2,\ldots,n)$ is written perm n. Ex. perm $3 = \{(1,2,3), (3,2,1), (2,1,3), (2,3,1), (3,1,2), (1,3,2)\}$

The sign of a permutation σ is defined to be $(-1)^{\text{number of pairs } (j,k) \text{ with } j<k \text{ s.t. } j \text{ appears after } k \text{ in } \sigma}$.

E.g. 1) sign $(2,1,3,4) = -1$
2) sign $(2,3,4,5,6,1) = (-1)^5 = -1$

(1,2) – flipped
(1,3)
(1,4)
(2,3)
(2,4)
(3,4)

Definition: If $A = [a_{ij}] \in M_n$, define determinant(A), or $|A|$, to be the number $\sum_{\sigma \in \text{perm } n} \text{sign}(\sigma)\, a_{\sigma(1),1},\ a_{\sigma(2),2}, \dots a_{\sigma(n),n}$

here $\sigma(k)$ is the k^{th} entry in the list σ.

Example: determinant$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$ $n=2$, perm $2 = \{(1,2),(2,1)\}$

$= \text{sign}(1,2)\, a_{1,1} a_{2,2} + \text{sign}(2,1)\, a_{2,1} a_{1,2}$
$= (+1)\, ad + (-1)\, cb$
$= ad - bc$

Example: $\left|\begin{bmatrix} a & b & c \\ d & e & f \\ x & y & z \end{bmatrix}\right| = \text{sign}(1,2,3)\, a_{1,1} a_{2,2} a_{3,3} + \text{sign}(1,3,2)\, a_{1,1} a_{3,2} a_{2,3} + \dots$

$= 1 \cdot aez - 1\,(ayf) + \dots$

pg. 230 in Axler

Prop. 4: determinant$\left(\begin{bmatrix} c_1 & & * \\ & \ddots & \\ 0 & & c_n \end{bmatrix}\right) = c_1 c_2 \dots c_n$, $c_k \in \mathbb{F}$

proof:
Divide perm n into 2 classes, $\{(1,2,\dots,n)\}$ and $\{$the rest$\}$
The 1st class makes a contribution to the determinant of sign$(1,2,\dots,n)\, a_{11} a_{22} \dots a_{nn} = 1 c_1 c_2 \dots c_n$. Let σ be a permutation in 2nd class.
Claim: $\sigma(k) > k$ for some $k = 1,2,\dots,n$

Because if this was false $\sigma(k) \le k \ \forall k$, so $\sigma(1)=1 \Rightarrow \sigma(2)=2 \Rightarrow$ etc, a contradiction. So claim is true, so for such k, $a_{\sigma(k),k} = 0$, so the contribution to determinant of this σ is 0. $\square$