Ex 6 Let T be the lin. op. from Ex 3, and let $W = \mathrm{Span}(\{e_1, e_2\})$, the T-cyclic subspace generated by e_1. We compute the char. poly $f(t)$ of T_W in two ways.

a) Using Thm 5.22.

By Ex 3 we have $\dim(W) = 2$, and $T^2(e_1) = -e_1$. So

$1e_1 + 0 \cdot T(e_1) + T^2(e_1) = 0.$

Hence by Thm 5.22(b), $f(t) = (-1)^2(1 + 0t + t^2) = t^2 + 1.$

b) By a direct calculation:

let $\beta = \{e_1, e_2\}$ - an ord. basis for W.

Since $T(e_1) = e_2$ and $T(e_2) = -e_1$, we have $[T_W]_\beta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, hence

$f(t) = \det\begin{pmatrix} -t & -1 \\ 1 & -t \end{pmatrix} = t^2 + 1.$

The Cayley - Hamilton Theorem

Def Let $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ be a polynomial with coefficients from a field F. Let T be a lin. op. on a v.s. V over F. We define

$f(T) = a_0 I_V + a_1 T + \ldots + a_n T^n,$

where $I_V \in \mathcal{L}(V)$ is the identity lin. op., and $T^n \in \mathcal{L}(V)$ is the n-fold composition of T, i.e. $T^n(x) = \underbrace{T(T(\ldots T(x))\ldots)}_{n \text{ times}}$ for all $x \in V$.

Rem Similarly, if $A \in M_{n \times n}(F)$, we define $f(A) = a_0 I_n + a_1 A + \ldots + a_n A^n$.

Thm E.3 Let $f(x) \in P(F)$ and $T \in \mathcal{L}(V)$ for V a v.s. over F. Then:

a) $f(T) \in \mathcal{L}(V)$,

b) If β is a finite ordered basis for V and $A = [T]_\beta$, then $[f(T)]_\beta = f(A)$.

Proof

a) Follow as the set of linear operators on $\mathcal{L}(V)$ is closed under composition, addition and scalar multiplication.

b) By the basic properties of matrix representation we have:

$[f(T)]_\beta = [a_0 I_V + a_1 T + \ldots + a_n T^n]_\beta = a_0 [I_V]_\beta + a_1 [T]_\beta + \ldots + a_n [T^n]_\beta = a_1 I_n + a_1 [T]_\beta + \ldots + a_n [T]_\beta^n =$
$= a_1 I_n + a_1 A + \ldots + a_n A^n = f(A).$

In view of Thm 5.22, cyclic spaces can be used to prove the following well-known result:

Thm 5.23 (Cayley-Hamilton) Let $T \in \mathcal{L}(V)$, V a fin. dim. v.s. over F. Let $f(t)$ be the char. poly. of T. Then $f(T) = T_0$, the zero-transformation (i.e. $T_0(x) = 0 \ \forall x \in V$). (That is, T satisfies its characteristic equation.)

Proof. We show that $f(T)(v) = 0$ for all $v \in V$.

This is obvious for $v = 0$ because $f(T)$ is linear by Thm E.3(a). So suppose $v \neq 0$.

Let W be the T-cyclic subspace of V generated by v, and let $k = \dim(W)$.

By Thm 5.22(a), there exist $a_0, a_1, \ldots, a_{k-1}$ s.t.

$a_0 v + a_1 T(v) + \ldots + a_{k-1} T^{k-1}(v) + T^k(v) = 0.$

Then Thm 5.22(b) implies that the char. poly. of T_W is

$g(t) = (-1)^k (a_0 + a_1 t + \ldots + a_{k-1} t^{k-1} + t^k).$

Combining these two equations we get

$g(T)(v) = (-1)^k (a_0 I + a_1 T + \ldots + a_{k-1} T^{k-1} + T^k)(v) = (-1)^k \big(\overbrace{a_0 v + a_1 T(v) + \ldots + a_{k-1} T^{k-1}(v) + T(v)}^{=0} \big) = 0.$

By Thm 5.21, $g(t)$ divides $f(t)$. Hence there exists a polynomial $q(t)$ s.t. $f(t) = q(t)\cdot g(t)$.

So $f(T)(v) = [q(T)\cdot g(T)](v) = q(t)\big(\underbrace{g(T)(v)}_{0 \text{ by above}}\big) = q(T)(0) = 0$ as $q(T): V \to V$ is linear.

Ex 7 Let $T \in \mathcal{L}(\mathbb{R}^2)$ defined by $T(a,b) = (a+2b, -2a+b)$, let $\beta = \{e_1, e_2\}$. Then:

$A = [T]_\beta = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$. The char. poly. of T is then $f(t) = \det(A - tI) = \det\begin{pmatrix} 1-t & 2 \\ -2 & 1-t \end{pmatrix} = t^2 - 2t + 5$.

We have $f(T) = T^2 - 2T + 5I_V$.

By Thm E.3: $[f(T)]_\beta = f(A) = A^2 - 2A + 5I_n = \begin{pmatrix} -3 & 4 \\ -4 & -3 \end{pmatrix} + \begin{pmatrix} -2 & -4 \\ 4 & -2 \end{pmatrix} + \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. So $f(T) = T_0$.

In general, using Thm 5.23 and Thm E.3 we have:

Corollary (Cayley-Hamilton for matrices) Let $A \in M_{n\times n}(F)$, let $f(t)$ be its char. poly. Then $f(A) = O$, the $n\times n$ zero matrix.

Proof.

Let $A \in M_{n\times n}(F)$ be given, consider $L_A \in \mathcal{L}(F^n)$, the lin. operator on F^n satisfying $[L_A]_\beta = A$, where β is the standard basis for F^n.

Recall that the char. poly. of L_A is by def. the char. poly. of A. Then

$$f(A) \underset{\text{Thm E.3}}{=} [f(L_A)]_\beta \underset{\text{Thm 5.23}}{=} [T_0]_\beta = O$$

T_0: the zero-transf. on F^n; O: the zero-matrix.

Direct sums of vector spaces

Let $T \in \mathcal{L}(V)$. There is a way of decomposing V into simpler subspaces that offers insight into the behavior of T.

In the case when T is diagonalizable, the simpler subspaces are its eigenspaces.

Def Let $W_1, W_2, \dots, W_k$ be subspaces of a v.s. V. The **sum** of these subspaces is the set of vectors $\{v_1 + v_2 + \dots + v_k : v_i \in W_i \text{ for } 1 \le i \le k\}$, which we denote by $W_1 + W_2 + \dots + W_k$ or $\sum_{i=1}^k W_i$. It is a subspace of V (Exercise!).

Ex Let $V = \mathbb{R}^3$, W_1 - the xy-plane, W_2 the yz-plane. Then $\mathbb{R}^3 = W_1 + W_2$, because for any vector $(a,b,c) \in \mathbb{R}^3$ we have

$$(a,b,c) = \underbrace{(a,0,0)}_{\in W_1} + \underbrace{(0,b,c)}_{\in W_2}.$$

But the presentation of (a,b,c) as a sum of vectors in W_1 and W_2 is not unique!

For example, $(a,b,c) = (a,b,0) + (0,0,c)$ is another presentation.

To fix this, we introduce a condition that assures unique presentation.

Def Let $W_1, \dots, W_k$ be subspaces of V.

We call V the **direct sum** of the subspaces $W_1, \dots, W_k$ and write $V = W_1 \oplus \dots \oplus W_k$ if $V = \sum_{i=1}^k W_i$ and $W_j \cap \sum_{i\ne j} W_i = \{0\}$ for each j, $1 \le j \le k$.

Ex Let $V = \mathbb{R}^4$, $W_1 = \{(a,b,0,0) : a,b \in \mathbb{R}\}$, $W_2 = \{(0,0,c,0) : c \in \mathbb{R}\}$, $W_3 = \{(0,0,0,d) : d \in \mathbb{R}\}$. For any $(a,b,c,d) \in V$ we have:

$(a,b,c,d) = (a,b,0,0) + (0,0,c,0) + (0,0,0,d) \in W_1 + W_2 + W_3$, so $V = \sum_{i=1}^3 W_i$.

We also have: $W_1 \cap (W_2 + W_3) = W_2 \cap (W_1 + W_3) = W_3 \cap (W_1 + W_2) = \{0\}$.

Thus $V = W_1 \oplus W_2 \oplus W_3$.

Thm 5.10 Let $W_1, \dots, W_k$ be subspaces of a fin. dim. v.s. V. The following are equivalent.

a) $V = W_1 \oplus \dots \oplus W_k$.

b) $V = \sum_{i=1}^{k} W_i$ and for any $v_1, \dots, v_k$ with $v_i \in W_i$, $1 \le i \le k$, if $v_1 + \dots + v_k = 0$, then $v_i = 0$ for all i.

c) Each vector $v \in V$ can be uniquely written as $v = v_1 + \dots + v_k$ for some $v_i \in W_i$.

d) If γ_i is an ordered basis for W_i, then $\gamma_1 \cup \dots \cup \gamma_k$ is an ordered basis for V.

e) For each $i = 1, \dots, k$ there exists an ordered basis γ_i for W_i s.t. $\gamma_1 \cup \dots \cup \gamma_k$ is an ordered basis for V.

Proof.

(a) $\Rightarrow$ (b) Clearly $V = \sum_{i=1}^{k} W_i$. Supp. that $v_i \in W_i$ and $v_1 + \dots + v_k = 0$. Then for any j we have $-v_j = \sum_{i \neq j} v_i \in \sum_{i \neq j} W_i$. But $-v_j \in W_j$, hence $-v_j \in W_j \cap \sum_{i \neq j} W_i \overset{\text{by (a)}}{=} \{0\}$.
So $v_j = 0$, hence (b) holds.

(b) $\Rightarrow$ (c) Let $v \in V$. By (b) $\exists\, v_1, \dots, v_k$ s.t. $v_i \in W_i$ and $v = v_1 + \dots + v_k$.
We must show that this representation is unique.
Supp. also that $v = w_1 + \dots + w_k$, with $w_i \in W_i$ for all i. Then
$(v_1 - w_1) + (v_2 - w_2) + \dots + (v_k - w_k) = 0$.
But $(v_i - w_i) \in W_i$ for all i, therefore $v_i - w_i = 0$ for all i by (b), thus $v_i = w_i$ for all i.

(c) $\Rightarrow$ (d) For each i, let γ_i be an ordered basis for W_i. Since $V = \sum_{i=1}^{k} W_i$ by (c), it follows that $\gamma_1 \cup \dots \cup \gamma_k$ generates V. To show that this set is lin. indep., consider vectors $v_{ij} \in \gamma_i$ ($j = 1, \dots, m_i$; $i = 1, \dots, k$) and scalars a_{ij} such that $\sum_{i,j} a_{ij} v_{ij} = 0$. For each i, set $w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$. Then for each i, $w_i \in \operatorname{Span}(\gamma_i) = W_i$ and
$w_1 + \dots + w_k = \sum_{i,j} a_{ij} v_{ij} = 0$.

Since $0 \in W_i$ for each i and $0 + 0 + \dots + 0 = w_1 + w_2 + \dots + w_k$, (c) implies that $w_i = 0$ $\forall i$.
Thus $0 = w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$ for each i.

But each γ_i is lin. indep., hence $a_{ij} = 0$ for all i and j. So $\gamma_1 \cup \dots \cup \gamma_k$ is lin. indep., so it is a basis for V.

(d) $\Rightarrow$ (e) Obvious.

(e) $\Rightarrow$ (a) For each i, let γ_i be an ordered basis for W_i s.t. $\gamma_1 \cup \dots \cup \gamma_k$ is an ord. basis for V. Then $V = \operatorname{Span}(\gamma_1 \cup \dots \cup \gamma_k) = \operatorname{Span}(\gamma_1) + \dots + \operatorname{Span}(\gamma_k) = \sum_{i=1}^{k} W_i$.
Fix j ($1 \le j \le k$) and supp. that, for some non-zero vector $v \in V$,
$v \in W_j \cap \sum_{i \neq j} W_i$. Then

$v \in W_j = \operatorname{Span}(\gamma_j)$ and $v \in \sum_{i \neq j} W_i = \operatorname{Span}\left(\bigcup_{i \neq j} \gamma_i\right)$.

Hence v is a non-trivial lin. comb. of both γ_j and $\left(\bigcup_{i \neq j} \gamma_i\right)$, so v can be expressed as a lin. comb. of $\gamma_1 \cup \dots \cup \gamma_k$ in more than one way, contradicting the property of a basis.
So $W_j \cap \sum_{i \neq j} W_i = \{0\}$, proving (a).

This allows us to characterize diagonalizability in terms of direct sums.

Thm 5.11 Given a v.s. V, $\dim(V)<\infty$, $T\in\mathcal{L}(V)$ is diagz. iff V is the direct sum of the eigenspaces of T.

Proof.
Let $\lambda_1,\ldots,\lambda_k$ be the distinct eigenvalues of T.
Suppose that T is diagz, and let γ_i be an ordered basis for E_{λ_i}.
Recall Thm 5.9 from 115A:
- If T is diagz, then $\gamma_1\cup\ldots\cup\gamma_k$ is an ordered basis for V consisting of eigenvectors of T.

Hence V is a direct sum of the E_{λ_i}'s by Thm 5.10 (e).

Conversely, supp. that V is a direct sum of the eigenspaces of T.
For each i, choose an ord. basis γ_i of E_{λ_i}.
By Thm 5.10(d), $\gamma_1\cup\ldots\cup\gamma_k$ is a basis for V.
Since this basis consists of eigenvectors ~~of~~ V, we conclude that T is diagz.

This gives a particular way of decomposing V as a direct sum of T-invariant subspaces when T is diagz. What can be said for general T?

Thm 5.24 Let $T\in\mathcal{L}(V)$, $\dim(V)<\infty$ and supp. $V=W_1\oplus\ldots\oplus W_k$, where W_i is a T-inv. subspace of V for each i $(1\le i\le k)$.
Let $f_i(t)$ be the char. poly. of T_{W_i}.
Then $f_1(t)\cdot\ldots\cdot f_k(t)$ is the char. poly. of T.

Proof. By induction on k.
Base step: supp. $k=2$.
Let β_i be an ord. basis for W_i, and $\beta=\beta_1\cup\beta_2$.
Then β is an ord. basis for V by Thm 5.10 (d).
Let $A=[T]_\beta$, $B_i=[T_{W_i}]_{\beta_i}$ for $i=1,2$.
Let $\beta_1=\{v_1,\ldots,v_m\}$, $\beta_2=\{v_{m+1},\ldots,v_n\}$, $\beta=\{v_1,\ldots,v_n\}$.

We have $$[T]_\beta \overset{\text{def}}{=} \begin{pmatrix} | & & | \\ [T(v_1)]_\beta & \cdots & [T(v_n)]_\beta \\ | & & | \end{pmatrix}.$$

Let $x\in V$ be any vector, say $x=a_1v_1+\ldots+a_nv_n$ for $a_i\in F$. Let $x_1:=\sum_{i=1}^m a_iv_i\in W_1$, $x_2=\sum_{i=m+1}^n a_iv_i\in W_2$.

If $x\in W_1$, then also $x_2=(x-x_1)\in W_1$, so $x_2\in W_1\cap W_2\overset{\text{as } V=W_1\oplus W_2}{=}\{0\}$, so $x_2=0$, then $a_i=0$ for all $m+1\le i\le n$.
If $x\in W_2$, then also $x_1=(x-x_2)\in W_2$, so $x_1\in W_1\cap W_2$, so $x_1=0$, then $a_i=0$ for all $1\le i\le m$.

As $v_i\in W_1$ for $1\le i\le m$ and W_1 is T-inv, we have $T(v_i)\in W_1$.
Similarly, $T(v_i)\in W_2$ for all $m+1\le i\le n$.
Hence for the coordinate vectors we have

For $1\le i\le m$: $[T(v_i)]_\beta=\begin{pmatrix} | \\ [T_{W_1}(v_i)]_{\beta_1} \\ | \\ 0 \\ \vdots \\ 0\end{pmatrix}\begin{matrix}\}m \\ \\ \}n-m\end{matrix}$; for $m+1\le i\le n$: $[T(v_i)]_\beta=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ | \\ [T_{W_2}(v_i)]_{\beta_2} \\ | \end{pmatrix}\begin{matrix}\}m \\ \\ \}n-m\end{matrix}$.

Hence $[T]_\beta=\begin{pmatrix} B_1 & O \\ O' & B_2\end{pmatrix}$, where O,O' are zero-matrices of appropriate size.

Then for the char. poly of T we have:

$f(t) = \det(A - tI) \underset{\substack{\uparrow \\ \text{Block} \\ \text{matrix}}}{=} \det(B_1 - tI) \cdot \det(B_2 - tI) = f_1(t) \cdot f_2(t)$.

Inductive step. Assume the theorem is valid for $k-1 \geq 2$, and we prove it for k.

Suppose $V = W_1 \oplus \ldots \oplus W_k$.

Let $W = W_1 + \ldots + W_{k-1}$.

Then W is a T-inv. subspace of V (if $x \in W_1 + \ldots + W_{k-1}$, say $x = x_1 + \ldots + x_{k-1}$ for $x_i \in W_i$, then $T(x) \overset{T \text{ is linear}}{=} T(x_1) + \ldots + T(x_{k-1})$, and $T(x_i) \in W_i$ as W_i is T-inv., so $T(x) \in W_1 + \ldots + W_{k-1}$.)

And $V = W \oplus W_k$ (clearly $V = W + W_k$; if $x \in W \cap W_k$, then $x \in \sum_{j \neq k} W_j \cap W_k = \{0\}$ as $V = W_1 \oplus \ldots \oplus W_k$).

So by the case for $k = 2$, we get $f(t) = g(t) \cdot f_k(t)$,

where $g(t)$ is the char. poly. of T_W.

Clearly $W = W_1 \oplus \ldots \oplus W_{k-1}$, therefore $g(t) = f_1(t) \cdot \ldots \cdot f_{k-1}(t)$ by the induction hypothesis.

So $f(t) = g(t) \cdot f_k(t) = f_1(t) \cdot \ldots \cdot f_k(t)$.

Cor Supp. that $T \in \mathcal{L}(V)$ is diagz, $\dim(V) < \infty$, and $\lambda_1, \ldots, \lambda_k$ are its distinct eigenvalues.

By Thm 5.11, V is a direct sum of the eigenspaces of T.

Since each eigenspace is T-inv., we can apply Thm 5.24.

For each λ_i, consider $T_{E_{\lambda_i}}$. Let γ_i be an ord. basis for E_{λ_i}. Let $\dim(E_{\lambda_i}) = m_i$, write $\gamma_i = \{v_1, \ldots, v_{m_i}\}$.

Then for each $1 \leq j \leq m_i$ we have $T_{E_{\lambda_i}}(v_j) = \lambda_i \cdot v_j$, so

$[T_{E_{\lambda_i}}(v_j)]_{\gamma_i} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{pmatrix} \leftarrow$ the j^{th} position, so $[T_{E_{\lambda_i}}]_{\gamma_i} = \underbrace{\begin{pmatrix} \lambda_i & & & 0 \\ & \lambda_i & & \\ & & \ddots & \\ 0 & & & \lambda_i \end{pmatrix}}_{m_i} \Big\} m_i$, so the char. poly of $T_{E_{\lambda_i}}$ is

$\det([T_{E_{\lambda_i}}]_{\gamma_i} - tI) = (\lambda_i - t)^{m_i}$.

So by Thm 5.11 we get $f(t) = (\lambda_1 - t)^{m_1} (\lambda_2 - t)^{m_2} \cdot \ldots \cdot (\lambda_k - t)^{m_k}$.

(So the multiplicity of each eigenval. is equal to the dimension of the corresponding eigenspace, as expected).

Next we define an operation on matrices which corresponds to direct sum of subspaces.

Def Let $B_1 \in M_{m \times m}(F)$, $B_2 \in M_{n \times n}(F)$. We define the direct sum of B_1 and B_2, denoted $B_1 \oplus B_2$, as the $(m+n) \times (m+n)$-matrix A s.t.

$$A_{ij} = \begin{cases} (B_1)_{ij} & \text{for } 1 \leq i, j \leq m, \\ (B_2)_{(i-m),(j-m)} & \text{for } m+1 \leq i, j \leq n+m, \\ 0 & \text{otherwise.} \end{cases}$$

If $B_1, \ldots, B_k$ are square matrices over F, possibly of different sizes, we define the direct sum of $B_1, \ldots, B_k$ recursively by

$B_1 \oplus \ldots \oplus B_k = (B_1 \oplus \ldots \oplus B_{k-1}) \oplus B_k$.

If $A = B_1 \oplus \ldots \oplus B_k$, then it is of the form $A = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k \end{pmatrix}$.

Ex Let $B_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, $B_2 = (3)$, $B_3 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$. Then $B_1 \oplus B_2 \oplus B_3 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$.

Thm 5.25 Let $T \in \mathcal{L}(V)$ be a lin. op., $\dim(V) < \infty$. Let $W_1, \dots, W_k$ be T-inv. subspaces of V s.t. $V = W_1 \oplus \dots \oplus W_k$. For each i, let β_i be an ordered basis for W_i, and let $\beta = \beta_1 \cup \dots \cup \beta_k$. Let $A = [T]_\beta$ and $B_i = [T_{W_i}]_{\beta_i}$, for $i = 1, \dots, k$. Then
$A = B_1 \oplus \dots \oplus B_k$.

Proof Homework 3.