Contents

5.2 Diagonalizability

Theorem 5.5. Let T be a linear operator on a vector space V, and let λ_1, λ_2, ..., λ_k be distinct eigenvalues of T. If $v_1,\ v_2,\ \ldots,\ v_k$ are eigenvectors of T such that λ_i corresponds to v_i $(1 \leq i \leq k)$, then $\{v_1,\ v_2,\ \ldots,\ v_k\}$ is linearly independent.

Proof.

Corollary. Let T be a linear operator on an $n-$dimensional vector space V. If T has n distinct eigenvalues, then T is diagonalizable.

Proof.

Definition. A polynomial $f(t)$ in $\mathsf{P}(F)$ **splits over** F if there are scalars c, $a_1,\ a_2,\ \ldots,\ a_n$ (not necessarily distinct) in F such that

$$f(t) = c(t - a_1)(t - a_2)\cdots(t - a_n).$$

Definition. Let λ be an eigenvalue of a linear operator or matrix with characteristic polynomial $f(t)$. The **(algebraic) multiplicity** of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of $f(t)$.

Definition. Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Define $\mathsf{E}_\lambda = \{x \in \mathsf{V} \ :\ \mathsf{T}(x) = \lambda x\} = \mathsf{N}(\mathsf{T} - \lambda \mathsf{I}_\mathsf{V})$ The Set E_λ is called the **eigenspace** of T corresponding to the eigenvalue λ. Analogously, we define the **eigenspace** of a square matrix A to be the eigenspace of L_A.

Theorem 5.6. The characteristic polynomial of any diagonalizable linear operator splits.

Proof.

Theorem 5.7. Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T having multiplicity m. Then $1 \leq \dim(\mathsf{E}_\lambda) \leq m$.

Proof.

Lemma. Let T be a linear operator, and let λ_1, λ_2, ..., λ_k be distinct eigenvalues of T. For each $i = 1,\ 2,\ \ldots,\ k$, let $v_i \in \mathsf{E}_{\lambda_i}$, the eigenspace corresponding to λ_i. If

$$v_1 + v_2 + \cdots + v_k = \mathit{0},$$

then $v_i = \mathit{0}$ for all i.

Proof.

Theorem 5.8. Let T be a linear operator on a vector space V, and let λ_1, λ_2, ..., λ_k be distinct eigenvalues of T. For each $i = 1,\ 2,\ \ldots,\ k$, let S_i be a finite linearly independent subset of the eigenspace E_{λ_i}. Then $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent subset of V.

Proof.

Theorem 5.9. Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. Then

(a) T is diagonalizable if and only if the multiplicity of λ_i is equal to $\dim(\mathsf{E}_{\lambda_i})$ for all i.

(b) if T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T.

(c) State and prove results analogous to (a) and (b) for matrices.

Proof.

Corollary. Let T be a linear operator on an $n-$dimensionala vector space V. Then T is diagonalizable if and only if both of the following conditions hold.

1. The characteristic polynomial of T splits.
2. For each eigenvalue λ of T, the multiplicity of λ equals $n - \mathrm{rank}(\mathsf{T} - \lambda \mathsf{I})$.

Direct sums

Definition. Let W_1, W_2, …, W_k be subspaces of a vector space V. We define the **sum** of these subspaces to be the set

$$\{v_1 + v_2 + \cdots + v_k \,:\, v_i \in \mathsf{W}_i \text{ for } 1 \leq i \leq k\},$$

which we denote by $\mathsf{W}_1 + \mathsf{W}_2 + \cdots + \mathsf{W}_k$ or $\displaystyle\sum_{i=1}^{k} \mathsf{W}_i$.

Definition. Let W_1, W_2, …, W_k be subspaces of a vector space V. We call V the **direct sum** of the subspaces W_1, W_2, …, W_k and write $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$, If

$$\mathsf{V} = \sum_{i=1}^{k} \mathsf{W}_i$$

and

$$\mathsf{W}_j \cap \sum_{i \neq j} \mathsf{W}_i = \{0\} \quad \text{for each } j \ (1 \leq j \leq k).$$

Theorem 5.10. Let W_1, W_2, …, W_k be subspaces of a finite-dimensional vector space V. The following conditions are equivalent.

(a) $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$

(b) $\mathsf{V} = \displaystyle\sum_{i=1}^{k} \mathsf{W}_i$ and, for any vectors v_1, v_2, …, v_k such that $v_i \in \mathsf{W}_i$ $(1 \leq i \leq k)$, if $v_1 + v_2 + \cdots + v_k = \mathit{0}$, then $v_i = \mathit{0}$ for all i.

(c) Each vector $v \in \mathsf{V}$ can be uniquely written as $v = v_1 + v_2 + \cdots + v_k$, where $v_i \in \mathsf{W}_i$.

(d) If γ_i is an ordered basis for W_i $(1 \leq i \leq k)$, then $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis for V.

(e) For each $i = 1, 2, \ldots, k$, there exists an ordered basis γ_i for W_i such that $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis for V.

Proof.

Proof. Continued...

Theorem 5.11. A linear operator T on a finite-dimensional vector space V is diagonalizable if and only if V is the direct sum of the eigenspaces of T.

Proof.

5.6 Invariant subspaces and the Cayley–Hamilton Theorem

Definition. Let T be a linear operator on a vector space V. A subspace W of V is called a **T-invariant subspace** of V if $\mathsf{T}(\mathsf{W}) \subseteq \mathsf{W}$, that is, if $\mathsf{T}(v) \in \mathsf{W}$ for all $v \in \mathsf{W}$.

Theorem 5.21. Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Then the characteristic polynomial of T_{W} divides the characteristic polynomial of T.

Proof.

Theorem 5.22. Let T be a linear operator on a finite-dimensional vector space V, and let W denote the T-cyclic subspace of V generated by a nonzero vector $v \in \mathsf{V}$. Let $k = \dim(\mathsf{W})$. Then

(a) $\left\{v,\ \mathsf{T}(v),\ \mathsf{T}^2(v), \ldots,\ \mathsf{T}^{k-1}(v)\right\}$ is a basis for W.

(b) If $a_0 v + a_1 \mathsf{T}(v) + \cdots + a_{k-1}\mathsf{T}^{k-1}(v) + \mathsf{T}^k(v) = \mathit{0}$, then the characteristic polynomial of T_{W} is $f(t) = (-1)^k(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k)$.

Proof.

The Cayley-Hamilton Theorem

Theorem 5.23 (Cayley-Hamilton). Let T be a linear operator on a finite-dimensional vector space V, and let $f(t)$ be the characteristic polynomial of T. Then $f(\mathsf{T}) = \mathsf{T}_0$, the zero transformation. That is, T "satisfies" its characteristic equation.

Proof.

Corollary (Cayley-Hamilton Theorem for Matrices). Let A be an $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of A. Then $f(A) = O$, the $n \times n$ zero matrix.

Theorem 5.24. Let T be a linear operator on a finite-dimensional vector space V, and suppose that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$, where W_i is a T-invariant subspace of V for each i $(1 \leq i \leq k)$. Suppose that $f_i(t)$ is the characteristic polynomial of $\mathsf{T}_{\mathsf{W}_i}$ $(1 \leq i \leq k)$. Then $f_1(t) \cdot f_2(t) \cdot \cdots \cdot f_k(t)$ is the characteristic polynomial of T.

Proof.

Definition. Let $B_1 \in \mathsf{M}_{m\times m}(F)$, and let $B_2 \in \mathsf{M}_{n\times n}(F)$. 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 such that

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

If $B_1,\ B_2,\ \ldots,\ B_k$ are square matrices with entries from F, then we define the **direct sum** of $B_1,\ B_2,\ \ldots,\ B_k$ recursively by

$$B_1 \oplus B_2 \oplus \cdots \oplus B_k = (B_1 \oplus B_2 \oplus \cdots \oplus B_{k-1}) \oplus B_k.$$

If $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$, then we often write

$$A = \begin{pmatrix} B_1 & O & \cdots & O \\ O & B_2 & \cdots & O \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & B_k \end{pmatrix}.$$

Theorem 5.25. Let T be a linear operator on a finite-dimensional vector space V, and let $\mathsf{W}_1,\ \mathsf{W}_2,\ \ldots,\ \mathsf{W}_k$ be T-invariant subspaces of V such that $\mathsf{V} = \mathsf{W}_1 \oplus \mathsf{W}_2 \oplus \cdots \oplus \mathsf{W}_k$. For each i, let β_i be an ordered basis for W_i, and let $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Let $A = [\mathsf{T}]_\beta$ and $B_i = [\mathsf{T}_{\mathsf{W}_i}]_{\beta_i}$ for $i = 1,\ 2,\ \ldots,\ k$. Then $A = B_1 \oplus B_2 \oplus \cdots \oplus B_k$.

Proof.