## Chapter 5 Diagonalization

### 5.1 Eigenvalues and Eigenvectors

**Definition.** ***Diagonalizable*** *A linear operator $T$ on a finite-dimensional vector space $V$ is called diagonalizable if there is an ordered basis $\beta$ for $V$ such that $[T]_\beta$ is a diagonal matrix. A square matrix $A$ is called diagonalizable if $L_A$ is diagonalizable.*

*Remark.* Want to determine if an linear operator $T$ is diagonalizable and if so, ways to obtain the basis $\beta = \{v_1, v_2, \cdots, v_n\}$ for $V$ such that $[T]_\beta$ is a diagonal matrix. Note that if $D = [T]_\beta$ is a diagonal matrix, i.e. $D_{ij} = 0$ for $i \neq j$, then for each $v_j \in \beta$, we have

$$T(v_j) = \sum_{i}^{n} D_{ij} v_i = D_{jj} v_j = \lambda_j v_j$$

where $\lambda_j = D_{jj}$. Conversely, if $\beta = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis for $V$ such that $T(v_j) = \lambda_j v_j$ for some scalars $\lambda_1, \lambda_2, \cdots, \lambda_n$, then

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

**Definition.** ***Eigenvalue and Eigenvector (characteristic/proper value or vector)***
*Let $T$ be a linear operator on a vector space $V$. A nonzero vector $v \in V$ is called an eigenvector of $T$ if there exists a scalar $\lambda$ such that $T(v) = \lambda v$. The scalar $\lambda$ is called the eigenvalue corresponding to the eigenvector $v$.*
*Let $A$ be in $M_{n\times n}(F)$. A nonzero vector $v \in F^n$ is called an eigenvector of $A$ if $v$ is an eigenvector of $L_A$; that is, if $Av = \lambda v$ for some scalar $\lambda$. The scalar $\lambda$ is the eigenvalue of $A$ corresponding to the eigenvector $v$*

**Theorem.** ***5.1 Sufficient Condition for Diagonalizability***
*A linear operator $T$ on a finite-dimensional vector space $V$ is diagonalizable if and only if there exists an ordered basis $\beta$ for $V$ consisting of eigenvectors of $T$ (i.e. $v \in V$ is eigenvector if exists $\lambda$ such that $T(v) = \lambda v$). Furthermore, if $T$ is diagonalizable, $\beta = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis of eigenvectors of $T$, and $D = [T]_\beta$, then $D$ is diagonal matrix and $D_{jj}$ is the eigenvalue corresponding to $v_j$ for $1 \leq j \leq n$*

*Remark.* To diagonalize a matrix or linear operator is to find a basis of eigenvectors and the corresponding eigenvalues

**Theorem.** ***5.2 Computing Eigenvalues***
*Let $A \in M_{n\times n}(F)$. Then a scalar $\lambda$ is an eigenvalue of $A$ if and only if $det(A - \lambda I_n) = 0$*

*Proof.* A scalar is an eigenvalue if and only if exists a nonzero vector $v \in F^n$ such that $Av = \lambda v$, that is, $(A - \lambda I_n)(v) = 0$, which is true if and only if $A - \lambda$ is not invertible (invertible and one-to-one, or $N(A - \lambda) = \{0\}$ equivalent). This is equivalent to $det(A - \lambda I_n) = 0$ □

**Definition.** ***Characteristic Polynomial of a Matrix*** *Let $A \in M_{n\times n}(F)$. The polynoimal $f(t) = det(A - tI_n)$ is called the characteristic polynomial of A*

1. *The eigenvalues of a matrix are the zeros of its characteristic polynomial*

2. *To determine the eigenvalues of a matrix or linear operator, we normally compute its characteristic polynomial.*

**Definition.** ***Characteristic Polynomial of a Linear Operator*** *Let $T$ be a linear operator on an $n$-dimensional vector space $V$ with ordered basis $\beta$. We define the characteristic polynomial $f(t)$ of $T$ to be the characteristic polynomial of $A = [T]_\beta$. That is,*

$$f(t) = det(A - tI_n) \qquad P_T(t) = det([T]_\beta - tI_n) = P_{[T]_\beta}(t)$$

*We denote characteristic polynomial of an operator $T$ by $det(T - tI)$. Note the definition is independent of the choice of ordered basis $\beta$, the resulting characteristic polynomial is the same regardless the choice of basis.*

*Proof.* Let $\beta$ and $\beta'$ be basis of $V$, let $Q$ be change of basis matrix from $\beta'$ to $\beta$, then we have $[T]_{\beta'} = Q^{-1}[T]_\beta Q$, so then characeristic polynomial of linear operator invariant of choice of basis

$$det([T]_{\beta'} - tI_V) = det(Q^{-1}([T]_\beta - tI_V)Q) = det(Q^{-1})det([T]_\beta - tI_V)det(Q) = det([T]_\beta - tI_V)$$

□

**Theorem.** ***5.3 Properties of Characteristic Polynomial***
*Let $A \in M_{n\times n}(F)$*

1. *The characteristic polynomial of A is a polynomial of degree n with leading coefficients $(-1)^n$*

2. *A has at most n distinct eigenvalues.*

**Theorem.** ***5.4 Computing Eigenvectors***
*Let $T$ be a linear operator on a vector space $T$, and let $\lambda$ be an eigenvalue of $T$. A vector $v \in V$ is an eigenvector of $T$ corresponding to $\lambda$ if and only if $v \neq 0$ and $v \in N(T - \lambda I)$*

**Proposition.** ***Equivalent Eigenvector for Matrix and Linear Operators***
*Let $T : V \to V$ be a linear operator and $\beta$ be an ordered basis for $V$. Let $A = [T]_\beta$ and note $\phi_\beta(v) = [v]_\beta$, the cooredinate vector of $v$ relative to $\beta$. We could show that for all $v \in V$ an eigenvector of $T$ corresponding to an eigenvalue $\lambda$ if and only if $\phi_\beta(v)$ is an eigenvector of $A$ corresponding to $\lambda$. Now suppose $v$ is an eigenvector of $T$ corresponding to $\lambda$, then $T(v) = \lambda v$, then*

$$A\phi_\beta(v) = L_A\phi_\beta(v) = \phi_\beta T(v) = \phi_\beta(\lambda v) = \lambda\phi_\beta(v)$$

*Note $\phi_\beta(v) \neq 0$, since $\phi_\beta$ is an isomorphism, we have proved that $\phi_\beta(v)$ is an eigenvector of A. Conversely, if $\phi_\beta(v)$ is an eigenvector of A corresponding to $\lambda$. Equivalently, a* *vector $y \in F^n$ is an eigenvector of $A = [T]_\beta$ corresponding to $\lambda$ if and only if $\phi_\beta^{-1}(y)$ is an eigenvector of $T$ corresponding to $\lambda$. We have reduced the problem of finding the eigenvectors of a linear operator on a finite-dimensional vector space to the problem of finding the eigenvectors of a matrix.*

**Definition.** ***Geometric Description of how a linear operator $T$ acts on an eigenvector*** *in the context of a vector space $V$ over $\mathbb{R}$. Let $v$ be eigenvector of $T$ and $\lambda$ be corresponding eigenvalue. Let $W = span(\{v\})$, the one-dimensional subspace of $V$ spanned by $v$, a line passing through 0 and $v$. For any $w \in W$, $w = cv$ for some $c \in \mathbb{R}$*

$$T(w) = T(cv) = cT(v) = c\lambda v = \lambda w$$

*$T$ acts on the vector in $W$ by multiplying each such vector by $\lambda$*

## 5.2 Diagonalizability

**Theorem.** ***5.5 Set of Eigenvectors is Linearly Independent***
*Let $T$ be a linear operator on a vector space $V$, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be **distinct** eigenvalues of $T$. If $v_1, v_2, \cdots, v_k$ are eigenvectors of $T$ such that $\lambda_i$ corresponding to $v_i$ where $1 \leq i \leq k$ (choose one eigenvector corresponding to each eigenvalue.), then $\{v_1, v_2, \cdots, v_k\}$ is linearly independent.*

**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.* Suppose $T$ has $n$ distinct eigenvalues $\lambda_1, \cdots, \lambda_n$. For each $i$ choose eigenvector $v_i$ corresponding to $\lambda_i$. By previous theorem, $\{v_1, \cdots, v_n\}$ is linearly independent, and since $dim(V) = n$. the set is a basis for $V$. Thus, by theorem 5.1, $T$ is diagonalizable □

1. *Converse not true, if $T$ is diagonalizable, then it need not have $n$ distinct eigenvalues. For example, $I_V$ is diagonalizable even thouigh it has only 1 eigenvalue, $\lambda = 1$*

**Definition.** ***Splits Over*** *A polynomial $f(t) \in P(F)$ splits over $F$ if there are scalars $c, a_1, \cdots, a_n$ (not necessarily distinct) in $F$ such that*

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

*As an example $t^2 = (t + 1)(t - 1)$ splits over $\mathbb{R}$, but $(t^2 + 1)(t - 2)$ does not split over $\mathbb{R}$ but splits over $\mathbb{C}$ since it factors into $(t + i)(t - i)(t - 2)$.*

**Theorem.** ***5.6 Diagonalizability implies $f(t)$ Splits Completely***
*The characteristic polynomial of any diagonalizable linear operator splits. The converse is not true, i.e. that the characteristic polynomial of $T$ may split but need not be diagonalizable.*

*Proof.* Let $T$ be linear operator on $n$-dimensional vector space $V$, and let $\beta$ be an ordered basis for $V$ such that $[T]_\beta = D$ is a diagonal matrix. Suppose

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and let $f(t)$ be characteristic polynomial of $T$, then

$$f(t) = det(D - tI) = \begin{pmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{pmatrix} = (-1)^n (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n)$$

□

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

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

**Theorem.** ***5.7 Dimension of Eigenspace is Bounded by Multiplicity***
*Let $T$ be a linear operator on a finite-dimensional vector space $V$, and let $\lambda$ be an eigenvalue of $T$ having multiplicity $m$. Then $1 \leq dim(E_\lambda) \leq m$*

**Lemma.** *Let $T$ be a linear operator, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of $T$. For each $i = 1, 2, \cdots, k$, let $v_i \in E_{\lambda_i}$, the eigenspace corresponding to $\lambda_i$. If*

$$v_1 + v_2 + \cdots + v_k = 0$$

*then $v_i = 0$ for all $i$.*

**Theorem.** ***5.8 Union of l.i. Subsets of Eigenspaces are l.i.***
*Let $T$ be a linear operator on a vector space $V$, and elt $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of $T$. for each $i = 1, 2, \cdots, k$, let $S_i$ be a finite linearly independent subset of eigenspace $E_{\lambda_i}$. Then $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a linearly independent subset of $V$.*

**Theorem.** ***5.9 Construct Bases of Eigenvectors in Eigenspace to Form a Basis for the Entire Space***
*Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of $T$. Then*

1. $T$ *is diagonalizable if and only if the multiplicit of* $\lambda_i$ *is equal to* $dim(E_{\lambda_i})$ *for all* $i$

2. *If* $T$ *is diagonalizable and* $\beta_i$ *is an ordered basis for* $E_{\lambda_i}$ *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* $T$*.*

**Definition.** ***Test for Diagonalization*** *Let* $T$ *be a linear operator on* $n$*-dimensional vector space* $V$*. Then* $T$ *is diagonalizable if and only if both of conditions hold*

1. *characteristic polynomial of* $T$ *splits*

2. *For each eigenvalue* $\lambda$ *of* $T$*, the multiplicity of* $\lambda$ *equals* $n - rank(T - \lambda I) = dim(E_\lambda)$

*Same condition can be used to test a square matrix* $A$ *is diagonalizable because diagonalizability of* $A$ *is equivalent to diagonalizability of* $L_A$*. To test* $T$ *for diagonalizability, usually pick a basis* $\alpha$ *and let* $B = [T]_\alpha$*. If characteristic polynomial of* $B$ *splits, then use condition 2 to check if the mulitiplicity of each repeated eigenvalue of* $B$ *equals* $n - rank(B - \lambda I)$ *(dont need to check for eigenvalues with multiplicity 1 by theorem 5.7). If so, then* $B$*, and hence* $T$*, is diagonalizable. If* $T$ *is diagonalizable, we can find a basis* $\beta$ *for* $V$ *consisting of eigenvectors of* $T$ *by taking the union of basis for each eigenspace of* $B$*. Furthermore, if* $A$ *is* $n \times n$ *diagonalizable matrixd, we can find an invertible* $n \times n$ *matrix* $Q$*, and a diagonal matrix* $n \times n$ *matrix* $D$ *such that* $Q^{-1}AQ = D$ *with* $Q$ *having its columns the vectors in the basis of eigenvectors of* $A$*, and* $D$ *having its* $j$*th column entry the eigenvalue of* $A$ *corresponding to* $j$*th column of* $Q$*.*

### Application: Closed Formula for Exponential of Diagonalizable Matrix

### Application: System of Differential Equations

**Definition.** ***Matrix Exponential*** *For* $A \in M_{n\times n}(C)$*, define* $e^A = \lim_{m\to\infty} B_m$*, where*

$$B_m = I + A + \frac{A^2}{2!} + \cdots + \frac{A^m}{m!}$$

*So* $e^A$ *is the sum of infinite series*

$$I + A + \frac{A^2}{2!} + \cdots$$

*Note*

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \qquad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

## 5.4 Invariant Subspaces and the Cayley-Hamilton Theorem

**Definition.** ***T-invariant Subspace*** *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* $T(W) \subseteq W$*, that is, if* $T(v) \in W$ *for all* $v \in W$

**Definition.** ***T-cyclic Subspace of*** $V$ ***Generated by*** $x$ *Let* $T$ *be a linear operator on a vector space* $V$*, and let* $x$ *be a nonzero vector in* $V$*. The subspace*

$$W = span(\{x, T(x), T^2(x), \cdots\})$$

*is called the* $T$*-cyclic subspace of* $V$ *generated by* $x$*, denoted by* $\langle v \rangle_T$*.*

1. $W$ *is a* $T$ *invariant subspace*

2. $W$ *is the smallest subspace of* $V$ *containing* $x$*; any* $T$*-invariant subspace of* $V$ *containing* $x$ *must also contain* $W$

Proof on how $T_W$ is a linear operator on $W$

**Theorem.** ***5.21 Characteristic Polynomial of*** $T_W$ ***Divides That of*** $T$ *Let* $T$ *be a linear operator on a finite-dimensional vector space* $V$*, and let* $W$ *be a* $T$*-invariant subspace of* $V$*. Let* $T_W$ *be restriction of* $T$ *to* $W$*. Then the characteristic polynomial of* $T_W$ *divides the chracteristic polynomial of* $T$*.*

1. *If* $\lambda$ *is an eigenvalue of* $T_W$ *then it is also an eigenvalue of* $T$

5.4:12 Proof Supplementary to theorem 5.21
4.3:12 Proof Supplementary to theorem 5.21

**Theorem.** ***5.22 Basis and Characteristic Polynomial For T-Invariant Subspace are Readily Computable*** *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* $v \in V$*. Let* $k = dim(W)$*. Then*

1. $\{v, T(v), T^2(v), \cdots, T^{k-1}(v)\}$ *is a basis for* $W$

2. *If* $a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) = 0$*, then the characteristic polynomial of* $T_W$ *is*

$$f(t) = (-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$$

*Idea is we can easily compute the characteristic polynomial by computing* $T^k(v)$ *and express it as a linear combination of the basis* $\{v, T(v), \cdots, T^{k-1}(v)\}$*, and use the 2nd claim of the theorem. Of course, we can derive chracteristic polynomial by computing determinants. By theorem 5.21, we can use characteristic polynomial of* $T_W$ *to gain information about the characteristic polynomial of* $T$ *itself*

5.4:12 for theorem 5.22, proves the second claim by induction

**Definition.** ***Linear Operator over a Polynomial*** *Let $f(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial with coefficients from a field $F$. If $T$ is a linear operator on a vector space $V$ over $F$, or similarly for $A \in M_{n\times n}(F)$, we define*

$$f(T) = a_0I + a_1T + \cdots + a_nT^n \qquad f(A) = a_0I + a_1A + \cdots + a_nA^n$$

**Theorem.** ***5.23 Cayley-Hamilton Theorem*** *Let $T$ be a linear operator on a finite-dimensional vector space $V$, and let $f(t)$ be the characteristic polynomial of $T$. Then $f(T) = T_0$, the zero transformation. That is $T$ satisfies its characteristic equation.*

*Proof.* Prove $f(T)(v) = 0$ for all $v \in W$. Idea is to consider a T-invariant subspace $W$ for $v$ chosen. Write down $k$th element as a linear combination of preivous basis vectors and compute its characteristic polynomial using method in theorem 5.22 and notice $P_{T_W}(T) = T_0$, i.e. $P_{T_W}(T)(v) = 0$ for all $v \in W$. Also $P_{T_W}$ divides $P_T$, from here results follows $\square$

**Corollary.** ***Cayley-Hamilton Theorem for Matrices*** *Let $A$ be $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of $A$. THen $f(A) = 0$, the $n \times n$ zero matrix.*