MATH 146 Linear Algebra 1, Lecture Notes by Stephen New

Chapter 8. Eigenvalues, Eigenvectors and Diagonalization

8.1 Definition: For a square matrix $D \in M_n(R)$ with entries in a ring R, we say that D is a **diagonal** matrix when $D_{k,l} = 0$ whenever $k \neq l$. For $\lambda_1, \lambda_2, \cdots, \lambda_n \in R$, we write $D = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ for the diagonal matrix D with $D_{k,k} = \lambda_k$ for all indices k.

8.2 Definition: Let $L \in \text{End}(U)$ where U is a finite dimensional vector space over a field F. We say that that L is **diagonalizable** when there exists an ordered basis $\mathcal{A}$ for U such that $[L]_\mathcal{A}$ is diagonal.

8.3 Note: Let $L \in \text{End}(U)$ where U is a finite dimensional vector space over a field F. When $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ is an ordered basis for U and $\lambda_1, \lambda_2, \cdots, \lambda_n \in F$, we have

$$[L]_\mathcal{A} = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \iff \big[L(u_k)\big]_\mathcal{A} = \lambda_k e_k \text{ for all } k \iff L(u_k) = \lambda_k u_k \text{ for all } k$$

Thus L is diagonalizable if and only if there exists an ordered basis $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ for U and there exist $\lambda_1, \lambda_2, \cdots, \lambda_n \in F$ such that $L(u_k) = \lambda_k u_k$ for all k.

8.4 Definition: Let $L \in \text{End}(U)$ where U is a vector space over a field F. For $\lambda \in F$, we say that λ is an **eigenvalue** (or a **characteristic value**) of L when there exists a nonzero vector $0 \neq u \in F^n$ such that $L(u) = \lambda u$. Such a vector $0 \neq u \in U$ is called an **eigenvector** (or **characteristic vector**) of L for λ. The **spectrum** of L is the set

$$\text{Spec}(L) = \big\{\lambda \in F \big| \lambda \text{ is an eigenvalue of } L\big\}.$$

For $\lambda \in F$, the **eigenspace** of L for λ is the subspace

$$E_\lambda = \big\{u \in U \big| L(u) = \lambda u\big\} = \big\{u \in U \big| (L - \lambda I)u = 0\big\} = \text{Ker}(L - \lambda I) \subseteq U.$$

Note that E_λ consists of the eigenvectors for λ together with the zero vector.

8.5 Note: Let $L \in \text{End}(U)$ where U is a finite dimensional vector space over a field F. For $\lambda \in F$

$$\begin{aligned} \lambda \text{ is an eigenvalue of } L &\iff \text{there exists } 0 \neq u \in \text{Ker}(L - \lambda I) \\ &\iff (L - \lambda I) \text{ is not invertible} \\ &\iff \det(L - \lambda I) = 0 \\ &\iff \lambda \text{ is a root of } f(x) = \det(L - xI). \end{aligned}$$

Note that when $\mathcal{A}$ is any ordered basis for U, we have

$$f(x) = \det(L - xI) = \det\big(\big[L - xI\big]_\mathcal{A}\big) = \det\big([L]_\mathcal{A} - xI\big) \in P_n(F).$$

8.6 Definition: Let $L \in \text{End}(U)$ where U is an n-dimensional vector space over a field F. The **characteristic polynomial** of L is the polynomial

$$f_L(x) = \det(L - xI) \in P_n(F).$$

Note that $\text{Spec}(L)$ is the set of roots of $f_L(x)$.

8.7 Note: Let $L \in \text{End}(U)$ where U is an n-dimensional vector space over a field F. Recall that L is diagonalizable if and only if there exists an ordered basis $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ for U such that each u_k is an eigenvector for some eigenvalue λ_k. The eigenvalues of L are the roots of $f_L(x)$, so there are at most n possible distinct eigenvalues. For each eigenvalue, the largest number of linearly independent eigenvectors for λ is equal to the dimension of E_λ. We can try to diagonalize L be finding all the eigenvalues λ for L, then finding a basis for each eigenspace E_λ, then selecting an ordered basis $\mathcal{A}$ from the union of the bases of the eigenspaces. In particular, note that if $\sum\limits_{\lambda \in \text{Spec}(L)} \dim(E_\lambda) < n$ then L cannot be diagonalizable.

8.8 Definition: Let F be a field and let $A \in M_n(F)$. By identifying A with the linear map $L = L_A \in \text{End}(F^n)$ given by $L(x) = Ax$, all of the above definitions and remarks may be applied to the matrix A. The matrix A is **diagonalizable** when there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$ (or equivalently $P^{-1}AP = D$). An **eigenvalue** for A is an element $\lambda \in F$ for which there exists $0 \neq x \in F^n$ such that $Ax = \lambda x$, and then such a vector x is called an **eigenvector** of A for λ. The set of eigenvalues of A, denoted by $\text{Spec}(A)$, is called the **spectrum** of A. For $\lambda \in F$, the **eigenspace** for λ is the vector space $E_\lambda = \text{Null}(A - \lambda I)$. The **characteristic polynomial** of A is the polynomial $f_A(x) = \det(A - xI) \in P_n(F)$,

8.9 Example: Let $A = \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix} \in M_2(F)$ where $F = \mathbf{R}$ or $\mathbf{C}$. The characteristic polynomial of A is

$$f_A(x) = \det(A - xI) = \begin{vmatrix} 3-x & -1 \\ 4 & -1-x \end{vmatrix} = (x-3)(x+1) + 4 = x^2 - 2x + 1 = (x-1)^2$$

so the only eigenvalue of A is $\lambda = 1$. When $\lambda = 1$ we have

$$(A - \lambda I) = (A - I) = \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix}$$

so the eigenspace $E_1 = \text{Null}(A - I)$ has basis $\{u\}$ where $u = (1,2)^T$. Since

$$\sum_{\lambda \in \text{Spec}(A)} \dim(E_\lambda) = \dim(E_1) = 1 < 2,$$

we see that A is not diagonalizable.

8.10 Example: Let $A = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \in M_2(F)$ where $F = \mathbf{R}$ or $\mathbf{C}$. The characteristic polynomial of A is

$$f_A(x) = \begin{vmatrix} 1-x & -2 \\ 2 & 1-x \end{vmatrix} = (x-1)^2 + 4 = x^2 - 2x + 5.$$

For $x \in \mathbf{C}$, we have $f_A(x) = 0 \iff x = \frac{2 \pm \sqrt{4-20}}{2} = 1 \pm 2i$. When $F = \mathbf{R}$, A has no eigenvalues (in $\mathbf{R}$) and so A is not diagonalizable. When $F = \mathbf{C}$, the eigenvalues of A are $\lambda_1 = 1 + 2i$ and $\lambda_2 = 1 - 2i$. As an exercise, show that when $\lambda = \lambda_1$ the eigenspace E_{λ_1} has basis $\{u_1\}$ where $u_1 = (i,1)^T$, and when $\lambda = \lambda_2$ the eigenspace E_{λ_2} has basis $u_2 = (-i,1)^T$, then verify that the matrix $P = (u_1, u_2) \in M_2(\mathbf{C})$ is invertible and that $P^{-1}A\,P = \text{diag}(\lambda_1, \lambda_2)$ thus showing that A is diagonalizable.

8.11 Theorem: *Let $L \in \mathrm{End}(U)$ where U is a vector space over a field F. Let $\lambda_1, \lambda_2, \cdots, \lambda_\ell \in F$ be distinct eigenvalues of L. For each index k, let $0 \neq u_k \in U$ be an eigenvector of L for λ_k. Then $\{u_1, u_2, \cdots, u_\ell\}$ is linearly independent.*

Proof: Since $u_1 \neq 0$ the set $\{u_1\}$ is linearly independent. Suppose, inductively, that the set $\{u_1, u_2, \cdots, u_{\ell-1}\}$ is linearly independent. Suppose that $\sum_{i=1}^{\ell} t_i u_i = 0$ with $t_i \in F$. Note that

$$0 = (L - \lambda_\ell I)\Big(\sum_{i=1}^{\ell} t_i u_i\Big) = \sum_{i=1}^{\ell} t_i(L(u_i) - \lambda_\ell u_i) = \sum_{i=1}^{\ell} t_i(\lambda_i - \lambda_\ell)u_i = \sum_{i=1}^{\ell-1} t_i(\lambda_i - \lambda_\ell)u_i$$

and so $t_i = 0$ for $1 \leq i < \ell$ since $\{u_1, u_2, \cdots, u_{\ell-1}\}$ is linearly independent. Since $t_i = 0$ for $1 \leq i \leq \ell - 1$ and $\sum_{i=1}^{\ell} t_i u_i = 0$, we also have $t_\ell = 0$. Thus $\{u_1, u_2, \cdots, u_\ell\}$ is linearly independent.

8.12 Corollary: *Let $L \in \mathrm{End}(U)$ where U is a vector space over a field F. Let $\lambda_1, \lambda_2, \cdots, \lambda_\ell$ be distinct eigenvalues of L. For each index k, let $\mathcal{A}_k$ be a linearly independent set of eigenvectors for λ_k. Then $\bigcup_{k=1}^{\ell} \mathcal{A}_k$ is linearly independent.*

Proof: Suppose that $\sum_{k=1}^{\ell}\sum_{i=1}^{m_k} t_{k,i}u_{k,i} = 0$ where each $t_{k,i} \in F$ and for each k, the vectors $u_{k,i}$ are distinct vectors in $\mathcal{A}_k$. Then we have $\sum_{i=1}^{\ell} u_k = 0$ where $u_k = \sum_{i=1}^{m_k} t_{k,i}u_{k,i} \in E_{\lambda_k}$. From the above theorem, it follows that $u_k = 0$ for all k, because if we had $u_k \neq 0$ for some values of k, say the values $k_1, k_2, \cdots, k_r$, then $\{u_{k_1}, u_{k_2}, \cdots, u_{k_r}\}$ would be linearly independent but $\sum_{i=1}^{r} u_{k_r} = 0$, which is impossible. Since for each index k we have $0 = u_k = \sum_{i=1}^{m_k} t_{k,i}u_{k,i}$ it follows that each $t_{k,i} = 0$ because $\mathcal{A}_k$ is linearly independent.

8.13 Corollary: *Let $L \in \mathrm{End}(U)$ where U is a finite dimensional vector space over a field F. Then*

$$L \text{ is diagonalizable if and only if } \sum_{\lambda \in \mathrm{Spec}(L)} \dim(E_\lambda) = \dim U.$$

In this case, if $\mathrm{Spec}(L) = \{\lambda_1, \lambda_2, \cdots, \lambda_\ell\}$ and, for each k, $\mathcal{A}_k = \{u_{k,1}, u_{k,2}, \cdots, u_{k,m_k}\}$ is an ordered basis for E_{λ_k}, and then

$$\mathcal{A} = \bigcup_{k=1}^{\ell} \mathcal{A}_k = \big\{u_{1,1}, u_{1,2}, \cdots, u_{1,m_1}, u_{2,1}, u_{2,2}, \cdots, u_{2,m_2}, \cdots, u_{\ell,1}, u_{\ell,2}, \cdots, u_{\ell,m_\ell}\big\}$$

is an ordered basis for U such that $[L]_{\mathcal{A}}$ is diagonal.

8.14 Definition: Let F be a field. For $f \in F[x]$ and $a \in F$, the **multiplicity** of a as a root of f, denoted by $\text{mult}(a, f(x))$, is the smallest $m \in \mathbf{N}$ such that $(x-a)^m$ is a factor of $f(x)$. Note that a is a root of f if and only if $\text{mult}(a, f) > 0$. For a non-constant polynomial $f \in F[x]$, we say that f **splits** (over F) when f factors into a product of linear factors in $F[x]$, that is when f is of the form $f(x) = c \prod_{i=1}^{n} (x - a_i)$ for some $a_i \in F$.

8.15 Theorem: *Let $L \in \text{End}(U)$ where U is a finite dimensional vector space over a field F. Let $\lambda \in \text{Spec}(L)$ and let $m_\lambda = \text{mult}\big(\lambda, f_L(x)\big)$. Then*

$$1 \leq \dim(E_\lambda) \leq m_\lambda.$$

Proof: Since λ is an eigenvalue of L we have $E_\lambda \neq \{0\}$ so $\dim(E_\lambda) \geq 1$. Let $m = \dim(E_\lambda)$ and let $\mathcal{A} = (u_1, u_2, \cdots, u_m)$ be an ordered basis for E_λ. Extend $\mathcal{A}$ to an ordered basis $\mathcal{B} = (u_1, \cdots, u_m, \cdots, u_n)$ for U. Since $L(u_i) = \lambda u_i$ for $1 \leq i \leq m$, the matrix $[L]_\mathcal{B}$ is of the form

$$[L]_\mathcal{B} = \begin{pmatrix} \lambda I & A \\ 0 & B \end{pmatrix} \in M_n(F)$$

where $I \in M_m(F)$. The characteristic polynomial of L is

$$f_L(x) = \begin{vmatrix} (\lambda - x)I & A \\ 0 & B - xI \end{vmatrix} = (\lambda - x)^m f_B(x).$$

Thus $(x-\lambda)^m$ is a factor of $f_L(x)$ and so $m_\lambda = \text{mult}\big(\lambda, f_L(x)\big) \geq m$.

8.16 Corollary: *Let $L \in \text{End}(U)$ where U is a finite dimensional vector space over a field F. Then L is diagonalizable if and only if $f_L(x)$ splits and $\dim(E_\lambda) = \text{mult}(\lambda, f_L(x)\big)$ for every $\lambda \in \text{Spec}(L)$.*

Proof: Suppose that L is diagonalizable. Choose an ordered basis $\mathcal{A}$ so that $[L]_\mathcal{A}$ is diagonal, say $[L]_\mathcal{A} = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. Note that $f_L(x) = f_D(x) = \prod_{k=1}^{n} (\lambda_k - x)$, and so $f_L(x)$ splits. For each $\lambda \in \text{Spec}(L)$, let $m_\lambda = \text{mult}(\lambda, f_L(x))$. Then, by the above theorem together with Corollary 8.13, we have

$$n = \dim(U) = \sum_{\lambda \in \text{Spec}(L)} \dim(E_\lambda) \leq \sum_{\lambda \in \text{Spec}(L)} m_\lambda = \deg(f_L) = n$$

which implies that $\dim(E_\lambda) = m_\lambda$ for all λ. Conversely, if f_L splits and $\dim(E_\lambda) = m_\lambda$ for all λ then

$$\sum_{\lambda \in \text{Spec}(L)} \dim(E_\lambda) = \sum_{\lambda \in \text{Spec}(L)} m_\lambda = \deg(f_L) = n = \dim(U)$$

and so L is diagonalizable.

8.17 Corollary: *Let $A \in M_n(F)$ where F is a field. Then A is diagonalizable if and only if $f_A(x)$ splits and $\dim(E_\lambda) = \text{mult}(\lambda, f_A(x))$ for all $\lambda \in \text{Spec}(A)$.*

8.18 Note: To summarize the above results, given a matrix $A \in M_n(F)$, where F is a field, we can determine whether A is diagonalizable as follows. We find the characteristicc polynomial $f_A(x) = \det(A - xI)$. We factor $f_A(x)$ to find the eigenvalues of A and the multiplicity of each eigenvalue. If $f_A(x)$ does not split then A is not diagonalizable. If $f_A(x)$ does split, then for each eigenvalue λ with multiplicity $m_\lambda \geq 2$, we calculate $\dim(E_\lambda)$. If we find one eigenvalue λ for which $\dim(E_\lambda) < m_\lambda$ then A is not diagonalizable. Otherwise A is diagonalizable. In particular we remark that if $f_A(x)$ splits and has n distinct roots (so the eigenvalues all have multiplicity 1) then A is diagonalizable.

In the case that A is diagonalizable and $f_A(x) = (-1)^n \prod_{k=1}^{\ell} (x - \lambda_k)^{m_k}$, if we find an ordered basis $\mathcal{A}_k = \{u_{k,1}, u_{k_2}, \cdots, u_{k,m_k}\}$ for each eigenspace, then we have $P^{-1}AP = D$ with

$$\begin{aligned} P &= \big(u_{1,1}, u_{1,2}, \cdots, u_{1,m_1}, u_{2,1}, u_{2,2}, \cdots, u_{2,m_2}, \cdots, u_{\ell,1}, u_{\ell,2}, \cdots, u_{\ell,m_\ell}\big) \\ D &= \operatorname{diag}\big(\lambda_1, \lambda_1, \cdots, \lambda_1, \lambda_2, \lambda_2, \cdots, \lambda_2, \cdots, \lambda_\ell, \lambda_\ell, \cdots, \lambda_\ell\big) \end{aligned}$$

where each λ_k is repeated m_k times.

8.19 Example: Let $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix} \in M_3(\mathbf{Q})$. Determine whether A is diagonalizable and, if so, find an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Solution: The characteristic polynomial of A is

$$\begin{aligned} f_A(x) = \big|A - xI\big| &= \begin{vmatrix} 3-x & 1 & 1 \\ 2 & 4-x & 2 \\ -1 & -1 & 1-x \end{vmatrix} \\ &= -(x-3)(x-4)(x-1) - 2 - 2 - 2(x-3) + 2(x-1) - (x-4) \\ &= -(x^3 - 8x^2 + 19x - 12) - x + 4 = -(x^3 - 8x^2 + 20x - 16) \\ &= -(x-2)(x^2 - 6x + 8) = -(x-2)^2(x-4) \end{aligned}$$

so the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 4$ of multiplicities $m_1 = 2$ and $m_2 = 1$. When $\lambda = \lambda_1 = 2$ we have

$$A - \lambda I = A - 2I = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so the eigenspace E_2 has basis $\{u_1, u_2\}$ with $u_1 = (-1, 0, 1)^T$ and $u_2 = (-1, 1, 0)^T$. When $\lambda = \lambda_2 = 4$ we have

$$A - \lambda I = A - 4I = \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 2 \\ -1 & -1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & -1 \\ 0 & 2 & 4 \\ 0 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

so the eigenspace E_4 has basis $\{u_3\}$ where $u_3 = (-1, -2, 1)^T$. Thus we have $P^{-1}AP = D$ where

$$P = \big(u_1, u_2, u_3\big) = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad D = \operatorname{diag}\big(\lambda_1, \lambda_1, \lambda_2\big) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

8.20 Definition: For a square matrix $T \in M_n(R)$ with entries in a ring R, we say that T is **upper triangular** when $T_{k,l} = 0$ whenever $k > l$.

8.21 Definition: Let $L \in \mathrm{End}(U)$ where U is a finite dimensional vector space over a field F. We say that L is (upper) **triangularizable** when there exists an ordered basis $\mathcal{A}$ for U such that $[L]_{\mathcal{A}}$ is upper triangular.

8.22 Definition: For a square matrix $A \in M_n(F)$, where F is a field, we say that A is (upper) **triangularizable** when there exists an invertible matrix $P \in GL_n(F)$ such that $P^{-1}AP$ is upper triangular.

8.23 Theorem: *(Schur's Theorem) Let F be a field.*

(1) Let $L \in \mathrm{End}(U)$ where U is a finite dimensional vector space over F. Then L is triangularizable if and only if $f_L(x)$ splits, and
(2) Let $A \in M_n(F)$. Then A is triangularizable if and only if $f_A(x)$ splits.

Proof: We shall prove Part (2), and we leave it as an exercise to show that Part (1) holds if and only if Part (2) holds. Suppose first that A is triangularizable. Choose an invertible matrix P and an upper triangular matrix T with $P^{-1}AP = T$. Then

$$f_A(x) = f_T(x) = \prod_{k=1}^{n} (T_{k,k} - x)$$

and so $f_A(x)$ splits.

Conversely, suppose that $f_A(x)$ splits. Choose a root λ_1 of $f_A(x)$ and note that λ_1 is an eigenvalue of A. Choose an eigenvector u_1 for λ_1, so we have $Au_1 = \lambda_1 u_1$. Since $u_1 \neq 0$ the set $\{u_1\}$ is linearly independent. Extend the set $\{u_1\}$ to a basis $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ for F^n. Let $Q = (u_1, u_2, \cdots, u_n) \in M_n(F)$, and note that Q is invertible because $\mathcal{A}$ is a basis for F^n. Since $Q^{-1}Q = I$, the first column of $Q^{-1}Q$ is equal to e_1, so we have

$$\begin{aligned} Q^{-1}AQ &= Q^{-1}A(u_1, u_2, \cdots, u_n) = Q^{-1}(Au_1, Au_2, \cdots, Au_n) \\ &= Q^{-1}\big(\lambda_1 u_1, A(u_2, \cdots, u_n)\big) = \big(\lambda_1 Q^{-1}u_1, Q^{-1}A(u_2, \cdots, u_n)\big) \\ &= \big(\lambda_1 e_1, Q^{-1}A(u_2, \cdots, u_n)\big) = \begin{pmatrix} \lambda_1 & x^T \\ 0 & B \end{pmatrix} \end{aligned}$$

with $x \in F^{n-1}$ and $B \in M_{n-1}(F)$.. Note that $f_A(x) = (x - \lambda_1) f_B(x)$ and so $f_B(x)$ splits. We suppose, inductively, that B is triangularizable. Choose $R \in GL_{n-1}(F)$ so that $R^{-1}BR = S$ with S upper-triangular. Let $P = Q \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \in M_n(F)$. Then P is invertible with $P^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & R^{-1} \end{pmatrix} Q^{-1}$ and

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & 0 \\ 0 & R^{-1} \end{pmatrix} Q^{-1}AQ \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & x^T \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & x^T R \\ 0 & BR \end{pmatrix} = \begin{pmatrix} \lambda_1 & x^T R \\ 0 & R^{-1}BR \end{pmatrix} = \begin{pmatrix} \lambda_1 & x^T R \\ 0 & S \end{pmatrix} \end{aligned}$$

which is upper triangular.