MATH 136
Linear Algbera 1

Sachin Kumar*
University of Waterloo

Fall 2022[†]

*skmuthuk@uwaterloo.ca
[†]Last updated: February 26, 2023

Contents

1 Eigenvalues and Diagonalization

1.1 Eigenpair

Definition 1.1.1 (Eigenvector, Eigenvalue and Eigenpair)
Let $A \in M_{n \times n}(\mathbb{F})$, a non-zero vector $\vec{x}$ is an eigenvector of A over $\mathbb{F}$, if there exists a scalar $\lambda \in \mathbb{F}$ such that
$$A\vec{x} = \lambda \vec{x}$$
The scalar λ is then called an eigenvalue of A over $\mathbb{F}$, and the pair $(\lambda, \vec{x})$ is an eigenpair of A over $\mathbb{F}$.

Note: Given any matix $A \in M_{n \times n}(\mathbb{F})$, we will seek vectors $\vec{x}$ such that $A\vec{x}$ is a scalar multiple of $\vec{x}$. For any such matrix A, $\vec{x} = \vec{0}$ is such a vector because $A\vec{0} = \vec{0} = c\vec{0}$ for any constant c, i.e., $A\vec{0}$ is always a scalar multiple of $\vec{0}$. This fact is so trivial that we focus our attention on vectors $\vec{x} \neq \vec{0}$ such that $A\vec{x}$ is a scalar multiple of $\vec{x}$.

1.2 Characteristic Polynomial and Eigenvalue

Definition 1.2.1 (Eigenvalue Equation or Eigenvalue Problem)
Let $A \in M_{n \times n}(\mathbb{F})$, we refer to the equation
$$A\vec{x} = \lambda \vec{x}$$
or
$$(A - \lambda I)\vec{x} = \vec{0}$$
as the eigenvalue equation for the matrix A over $\mathbb{F}$. It is also sometimes refered to as the eigenvalue problem.

Note: This is an unusual equation to solve since we want to solve it for both the vector $\vec{x} \in \mathbb{F}^n$ and the scalaer $\lambda \in \mathbb{F}$. We will approach the problem by first identifying eligible values of λ. We can then determine corresponding sets of vectors $\vec{x}$ that solve the equation for each λ we identify.
As mentioned in the previous note, a trivial solution to this equation is $\vec{x} = \vec{0}$. We try to obtain a non-trivial $(\vec{x} \neq \vec{0})$ solution to the eigenvalue equation. This is possible if and only if the RREF of matrix $A - \lambda I$ has fewer than n pivots, which occurs if and only if it is not invertible, i.e, $\det(A - \lambda I) = 0$.

Computing $\det(A - \lambda I) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix}$ by expanding along the first row, we sum up n terms, each of which is the product of entries in $A - \lambda I$. Since each entry is either a constant or a linear term in a variable λ, and all the products present in the expanded determinant calculation contains n terms, we can infer that $\det(A - \lambda I)$ is a polynomila in λ of degree n.

Definition 1.2.2 (Characteristic Polynomial and Characteristic Equation)
Let $A \in M_{n \times n}(\mathbb{F})$ and $\lambda \in \mathbb{F}$, the characteristic polynomial of A, denoted by $C_A(\lambda)$, is

$$C_A(\lambda) = \det(A - \lambda I)$$

The characteristic equation of A is
$$C_A(\lambda) = 0$$

Note: The eigenvalues of A over $\mathbb{F}$ are the roots of the characteristic polynomial of A in $\mathbb{F}$. If $C_A(\lambda)$ has no real roots then A will have neither eigenvalues nor eigenpairs over $\mathbb{F}$.

1.2.1 Properties of the Characteristic Polynomial

Proposition 1.2.1
Let $A \in M_{n \times n}(\mathbb{F})$. Then A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof

$$
\begin{aligned}
A \text{ is invertible iff } & \det(A) \neq 0 \\
\text{iff } & \det(A - 0I_n) \neq 0 \\
\text{iff } & 0 \text{ is not a root of the characteristic polynomial} \\
\text{iff } & 0 \text{ is not an eigenvalue of the matrix } A
\end{aligned}
$$

Definition 1.2.3 (Trace)
Let $A \in M_{n \times n}(\mathbb{F})$. we define the trace of A by,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Proposition 1.2.2 (Features of the characteristic polynomial)
Let $A \in M_{n \times n}(\mathbb{F})$ have the characteristic polynomial $C_A(\lambda) = \det(A - \lambda I)$. Then $C_A(\lambda)$ is the degree n polynomial in λ of the form

$$C_A(\lambda) = c_n \lambda^n + c_{n-1} \lambda^{n-1} + \cdots + c_1 \lambda + c_0$$

where,

1. $c_n = (-1)^n$

2. $c_{n-1} = (-1)^{n-1} \mathrm{tr}(A)$

3. $c_0 = \det(A)$

Proof
By performing cofactor expansion along the first row of A and along the first row of every subsequent $(1,1)$-submatrix, we obtain an expression of the form,

$$C_A(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda) + f(\lambda)$$

where f is the polynomial that is the sum of the other terms of the determinant. In the above, we have singled out the contribution of the product of the diagonal entries, and have grouped all other terms. This latter grouping $f(\lambda)$ is a polynomial of degree at most $n-2$. To see this, note that unless the cofactor calculation is done with respect to a diagonal entry, it will use an entry (i,j) that corresponds to deleting the entry $(a_{ii} - \lambda)$ from the i-th row and the entry $(a_{jj} - \lambda)$ from the j-th column. This leaves us with at most $n-2$ entries that contain a λ.

Expanding out the first term in the expression $C_A(\lambda)$ to obtain,

$$C_A\lambda = a_{11} \ldots a_{nn} + \cdots + (-1)^{n-1}(a_{11} + \cdots + a_{nn})\lambda^{n-1} + (-1)^n \lambda^n + f(\lambda)$$

we see that $C_A(\lambda)$ is a polynomial of degree n in λ because we have proved that this degree is not exceeded by the polynomial f, which is of degree at most $n-2$.

1. The co-efficient of λ^n is $c_n = (-1)^n$

2. The co-efficient of λ^{n-1} is

$$c_{n-1} = (-1)^{n-1}(a_{11} + \cdots + a_{nn}) = (-1)^{n-1}\mathrm{tr}(A)$$

3. From the definition of $C_A(\lambda)$, we have the constant term is,

$$c_0 = C_A(0) = \det(A - 0I) = \det(A)$$

Proposition 1.2.3 (Characteristic Polynomial and Eigenvalues over $\mathbb{C}$)
Let $A \in M_{n \times n}(\mathbb{F})$ have the characteristic polynomial,

$$C_A(\lambda) = c_n \lambda^n + c_{n-1}\lambda^{(n-1)} + \cdots + c_1 \lambda + c_0$$

and n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$. Then

$$c_{n-1} = (-1)^{(n-1)} \sum_{i=1}^{n} \lambda_i$$

and

$$c_0 = \prod_{i=1}^{n} \lambda_i$$

Note that if A has repeated eigenvalues over $\mathbb{C}$, then we include each eigenvalue in the list $\lambda_1, \lambda_2, \ldots, \lambda_n$ as many times as its corresponding linear factor appears in the characteristic polynomial $C_A(\lambda)$.

Proof
The eigenvalues of A over $\mathbb{C}$ are the n complex roots of the characteristic polynomial, so its characteristic polynomial has the form,

$$C_A(\lambda) = k(\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n)$$

for some $k \in \mathbb{C}$

1. Consider c_{n-1}, the coefficient of λ^{n-1} in $C_A(\lambda)$. By expanding out the expression for $C_A(\lambda)$, we find that there are n terms involving λ^{n-1}, each of which is the product of $(-1)^n$ with one of the constants $-\lambda_1, -\lambda_2, \ldots, -\lambda_n$ and with λ from each of the other $n-1$ linear factors. By taking the

sum of these terms, we find that

$$c_{n-1}\lambda^n = (-1)^n[-\lambda_1(\lambda)^{n-1} - \lambda_2(\lambda)^{n-1} - \cdots - \lambda_n(\lambda)^{n-1}]$$

$$= \left((-1)^{(n-1)}\sum_{i=1}^{n}\lambda_i\right)\lambda^n$$

i.e.,

$$c_{n-1} = (-1)^{(n-1)}\sum_{i=1}^{n}\lambda_i$$

2. Expanding the terms of $C_A\lambda$, its constant term must be $(-1)^n$ times the product of the constant terms in each of the n linear factors, which are $-\lambda_1, -\lambda_2, \ldots, -\lambda_n$. Therefore,

$$c_0 = (-1)^n\prod_{i=1}^{n}(-\lambda_i) = (-1)^{2n}\prod_{i=1}^{n}\lambda_i = \prod_{i=1}^{n}\lambda_i$$

Excercise 1: Prove the following corollary 1 on Eigenvalues and trace/determinant

Let $A \in M_{n \times n}(\mathbb{F})$ have n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$. Show that

$$\sum_{i=1}^{n}\lambda_i = \mathrm{tr}(A)$$

and

$$\prod_{i=1}^{n}\lambda_i = \det(A)$$

1.3   Eigenvectors

Once we have found an eigenvalue λ of a matrix A over $\mathbb{F}$, we can examine the eigenvalue equation in the form $(A - \lambda I)\vec{x} = 0$ in order to obtain a corresponding eigenvector x.

Note: Consider $(\lambda, \vec{v})$ is an eigenpair of A. We can get other eigenpairs by pairing non-zero scalar multiples of the eigenvector $\vec{v}$ with the eigenvalue λ.

1.4   Eigenspaces

Proposition 1.4.1 (Linear combinations of Eigenvectors)
Let $c, d \in \mathbb{F}$ and suppose that $(\lambda_1, \vec{x})$ and $(\lambda_1, \vec{y})$ are eigenpairs of the matrix A over $\mathbb{F}$ with the same eigenvalue λ_1. if $c\vec{x} + d\vec{y} \neq \vec{0}$, then $(\lambda_1, c\vec{x} + d\vec{y})$ is also an eigenpair for A with the eigenvalue λ_1.

Proof
We see that
$$A(c\vec{x} + d\vec{y}) = c(A\vec{x}) + d(A\vec{y}) = c(\lambda_1\vec{x}) + d(\lambda_1\vec{y}) = \lambda_1(c\vec{x} + d\vec{y})$$

Thus, since $c\vec{x} + d\vec{y} \neq 0$, it follows that $c\vec{x} + d\vec{y}$ is an eigenvector of A with eigenvalue λ_1i.e., $(\lambda_1, c\vec{x} + d\vec{y})$ is an eigenpair for A.

Since every non-zero linear combination of eigenvectors with respect to a fixed eigenvalue λ are also eigenvectors with respect to λ, this suggests collecting all possible eigenvectors of A with respect to a value λ into a set.

Definition 1.4.1 (Eigenspace)