MA1140: Lecture 8
Eigenvalues and Eigenvectors

Dipankar Ghosh

Department of Mathematics
Indian Institute of Technology Hyderabad

January 28, 2020

Eigenvalues and eigenvectors (of linear operators)

Let $T : V \to V$ be a linear map, which we call linear operator.

Definition

1. A NON-ZERO vector $v \in V$ is called an **eigenvector** of T if $T(v) = \lambda v$ for SOME scalar λ.
2. A scalar λ is called an **eigenvalue** of T if there EXISTS a non-zero vector $v \in V$ such that $T(v) = \lambda v$.
3. If $T(v) = \lambda v$, then λ is called an eigenvalue of T corresponding to the eigenvector v.

- **Geometrically**, an eigenvector, corresponding to an eigenvalue, points in a direction that is stretched by the transformation, and the eigenvalue is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed.
- An eigenvalue can be positive, negative or zero.

Eigenvalues and eigenvectors (of square matrices)

Since there is a one to one correspondence between the set of all linear operators from V ($\cong \mathbb{R}^n$) to itself and the collection of all $n \times n$ matrices over $\mathbb{R}$, it is equivalent to define eigenvalues and eigenvectors of $n \times n$ matrices.

Definition

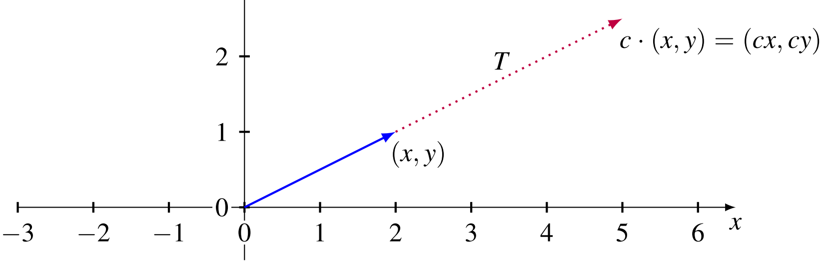
Let A be an $n \times n$ matrix over $\mathbb{R}$.

1. A NON-ZERO column vector $v \in \mathbb{R}^n$ is called an **eigenvector** of A if $Av = \lambda v$ for SOME $\lambda \in \mathbb{R}$.
2. A scalar λ is called an **eigenvalue** of A if there EXISTS a non-zero column vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$.
3. If $Av = \lambda v$, then λ is called an eigenvalue of A corresponding to the eigenvector v.

Let $c \in \mathbb{R}$. Define $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = c \begin{pmatrix} x \\ y \end{pmatrix} \text{ for every } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

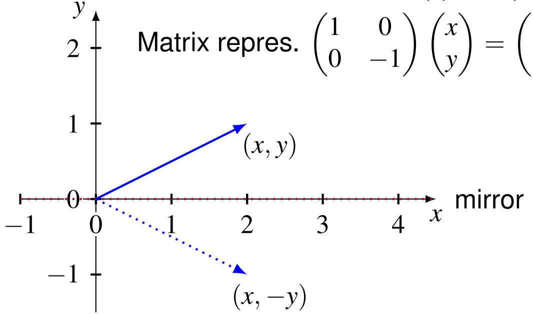
The matrix representation: $[T] = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$



Every $v \, (\neq 0) \in \mathbb{R}^2$ is an eigenvector of T with the eigenvalue c.

$T:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is defined by $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$.

Matrix repres. $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$



For $x \neq 0$, $\begin{pmatrix} x \\ 0 \end{pmatrix}$ is an eigenvector of T with eigenvalue 1.

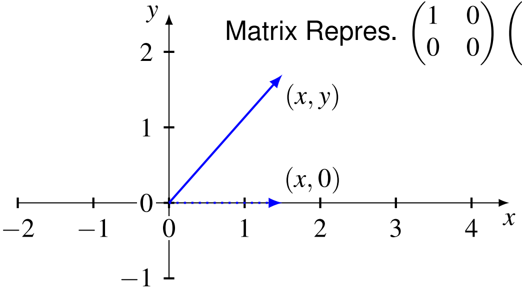
For $y \neq 0$, $\begin{pmatrix} 0 \\ y \end{pmatrix}$ is an eigenvector of T with eigenvalue -1.

These are ALL the eigenvectors of T. (Verify it!)

Example 3: eigenvalues and eigenvectors of projection

Define $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ by $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$

Matrix Repres. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$



projection on x-axis

For $x \neq 0$, $\begin{pmatrix} x \\ 0 \end{pmatrix}$ is an eigenvector of T with eigenvalue 1.

For $y \neq 0$, $\begin{pmatrix} 0 \\ y \end{pmatrix}$ is an eigenvector of T with eigenvalue 0.

These are ALL the eigenvectors of T. (Verify it!)

Example 4: A may not have eigenvalues and eigenvectors over a particular field

- Consider the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ over $\mathbb{R}$.
- Does A have eigenvalues and eigenvectors over $\mathbb{R}$?
- If yes, then there are $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \smallsetminus \{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\}$ and $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \\ \implies & \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ \implies & \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

  Since $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = 0$. Hence $\lambda^2 + 1 = 0$. But no such λ exists in $\mathbb{R}$.
- So A does not have eigenvalues and eigenvectors over $\mathbb{R}$.

The existence of eigenvalues and eigenvectors

- Consider the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ over $\mathbb{C}$, the set of complex numbers.
- Does A have eigenvalues and eigenvectors over $\mathbb{C}$? Ans. Yes.
- Note that $\lambda^2 + 1$ has solutions: $\pm i \in \mathbb{C}$.
- Then, for each $\lambda = \pm i$, in view of the previous slide, one should solve the system
$$\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
to get
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} i \\ -1 \end{pmatrix} = i \begin{pmatrix} i \\ -1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} i \\ 1 \end{pmatrix} = -i \begin{pmatrix} i \\ 1 \end{pmatrix}$$
- **Conclusion:** The matrix A has eigenvalues and eigenvectors over $\mathbb{C}$, but not over $\mathbb{R}$.

Characteristic polynomial of a matrix

Let A be an $n \times n$ matrix over $\mathbb{C}$. Denote the identity matrix by I_n.

Lemma

The following statements are equivalent:

- $\lambda \in \mathbb{R}$ is an eigenvalue of A.
- $\det(\lambda I_n - A) = 0$.

Proof. Note that λ is an eigenvalue of A $\Leftrightarrow$ there is $v \neq 0$ in $\mathbb{C}^n$ such that $Av = \lambda v$, i.e., $(A - \lambda I_n)v = 0 \Leftrightarrow (A - \lambda I_n)X = 0$ has a non-trivial solution $\Leftrightarrow \det(A - \lambda I_n) = 0 \Leftrightarrow \det(\lambda I_n - A) = 0$.

Definition

1. The **characteristic polynomial** of A, denoted by $p_A(x)$, is the polynomial defined by $p_A(x) := \det(xI_n - A)$.
2. Thus the eigenvalues of A are nothing but the roots of $p_A(x)$.
3. The **algebraic multiplicity** $\mathrm{AM}_A(\lambda)$ of the eigenvalue λ of A is its multiplicity as a root of the characteristic polynomial $p_A(x)$, that is, the largest integer k such that $(x - \lambda)^k$ is a factor of $p_A(x)$.

Eigenspace associated with an eigenvalue

Let A be an $n \times n$ matrix over $\mathbb{C}$. Denote the identity matrix by I_n.

Lemma

The following statements are equivalent:

1. $v \in \mathbb{C}^n$ is an eigenvector of A with the corr. eigenvalue λ.
2. $v \in \mathbb{C}^n$ is a non-trivial solution of the system $(A - \lambda I_n)X = 0$, i.e., $v \in \mathbb{C}^n \setminus \{0\}$ lies in $\mathrm{Null}(A - \lambda I_n)$.

Proof. Note that $Av = \lambda v$ if and only if $(A - \lambda I_n)v = 0$.

Definition

1. Given a particular eigenvalue λ of A. The set of all eigenvectors of A corresponding to λ, together with the zero vector, is called the **eigenspace** of A associated with λ.
   It is denoted by E_λ. Note that $E_\lambda = \mathrm{Null}(A - \lambda I_n)$.
2. The dimension of $E_\lambda = \mathrm{Null}(A - \lambda I_n)$ is referred to as the **geometric multiplicity of** λ, denoted by $\mathrm{GM}_A(\lambda)$.

Some inequalities on algebraic/geometric multiplicities

Theorem

Let A be an $n \times n$ matrix over $\mathbb{C}$. For every eigenvalue λ of A, we have

1. $1 \leqslant \mathrm{AM}_A(\lambda) \leqslant n$ and $1 \leqslant \mathrm{GM}_A(\lambda) \leqslant n$.
2. $\sum_{i=1}^{r} \mathrm{AM}_A(\lambda_i) = n$, the sum varies over all the eigenvalues of A.
3. $1 \leqslant \mathrm{GM}_A(\lambda) \leqslant \mathrm{AM}_A(\lambda) \leqslant n$.

Proof.

1. Note that $\deg(p_A(x)) = n$ and $p_A(x) = (x-\lambda)^{\mathrm{AM}_A(\lambda)} f(x)$ for some f. Since $\mathrm{GM}_A(\lambda) = \dim(\mathrm{Null}(A - \lambda I_n))$, we have $1 \leqslant \mathrm{GM}_A(\lambda) \leqslant n$.
2. It follows from $p_A(x) = \prod_{i=1}^{r}(x-\lambda_i)^{\mathrm{AM}_A(\lambda_i)}$ and $\deg(p_A(x)) = n$.
3. We will skip it.

$\square$

Example: Characteristic polynomial and eigenvalues

- Consider $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.
- The characteristic polynomial of A is given by
$$\begin{aligned} p_A(x) = \det(xI_2 - A) &= \det \begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix} \\ &= (x-1)(x-4) - 6 = x^2 - 5x - 2. \end{aligned}$$
- The roots of $p_A(x)$ are $\dfrac{5 \pm \sqrt{33}}{2}$.
- The eigenvalues of A are $\lambda_1 = \dfrac{5 - \sqrt{33}}{2}$ and $\lambda_2 = \dfrac{5 + \sqrt{33}}{2}$.
- The algebraic multiplicities of both λ_1 and λ_2 are 1.

How to compute eigenvalues and eigenvectors

- First compute the characteristic polynomial $p_A(x) = \det(xI_n - A)$ of A.
- Next compute the roots of $p_A(x)$ by factorizing it into linear factors. Which gives the eigenvalues.
- Then, for each eigenvalue λ, solve the homogeneous system:
$$(A - \lambda I_n)X = 0$$
to get eigenspace of A associated with λ.
- Recall that in order to solve a linear system, you may apply elementary row operations to make it into a system with row reduced echelon coefficient matrix.

Similarity of matrices

Definition

Two $n \times n$ matrices A and B are called **similar** if there exists an invertible $n \times n$ matrix P such that $B = P^{-1}AP$.

Some statements (without proof) about importance of similarity of matrices:

- Two matrices are similar if and only if they represent the same linear operator with respect to (possibly) different bases. (???)
- Two similar matrices A and B share many properties:
  - $\mathrm{rank}(A) = \mathrm{rank}(B)$ as operators from $\mathbb{R}^n$ to itself.
  - $\det(A) = \det(B)$; $\mathrm{tr}(A) = \mathrm{tr}(B)$ (sum of all diagonal entries).
  - A and B have same characteristic polynomial, $\det(xI_n - A)$.
  - Minimal polynomials of A and B are same. A monic polynomial $p(X) \in \mathbb{R}[X]$ is said to be a minimal polynomial of A if $p(A) = 0$ (zero matriz) and p has minimal possible degree.
  - Jordan canonical forms of A and B are same. (???)

Diagonalizable matrices

Motivation: For a matrix, eigenvalues and eigenvectors can be used to decompose the matrix, for example by diagonalizing it.

Definition

A matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix D, i.e., if there is an invertible matrix P such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{(a diagonal matrix).}$$

The set of eigenvectors helps us to test whether a matrix is diagonalizable or not.

The use of eigenvalues and eigenvectors on diagonalization

Theorem

Let A be an $n \times n$ matrix (over $\mathbb{C}$). The following are equivalent:

1. *A is diagonalizable.*
2. *The eigenvectors of A form a basis of $\mathbb{R}^n$, equivalently, A has n linearly independent eigenvectors $v_1, \ldots, v_n$ with associated eigenvalues $\lambda_1, \ldots, \lambda_n$ (which need not be distinct).*
3. *$\mathrm{GM}_A(\lambda) = \mathrm{AG}_A(\lambda)$ for every eigenvalue λ of A.*
4. *The minimal polynomial of A has distinct roots (equivalently, A satisfies a polynomial $p(x) \in \mathbb{C}[x]$ having distinct roots).*

Proof. (1) $\Rightarrow$ (2): There is an $n \times n$ invertible matrix P such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{for some } \lambda_1, \ldots, \lambda_n \in \mathbb{R}.$$

Hence multiply by P from the left side.

Proof of the theorem contd...

Proof. (1) $\Rightarrow$ (2): ... Thus $AP = P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$.

Write $P = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ for some $v_1, \ldots, v_n \in \mathbb{R}^n$.
Then $AP = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix}$ and

$$P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \left[P \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad P \begin{pmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad P \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{pmatrix} \right]$$

$= \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{bmatrix}$.
Therefore $Av_i = \lambda_i v_i$ for every $1 \leqslant i \leqslant n$.
Note that $v_1, \ldots, v_n$ are linearly independent, since P is invertible.

Proof. (2) $\Rightarrow$ (1): A has n linearly independent eigenvectors $v_1, \ldots, v_n \in \mathbb{R}^n$ with associated eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.
Set $P := \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$. Clearly P is an $n \times n$ matrix.
Since $v_1, \ldots, v_n$ are linearly independent, P is invertible.
Moreover
$AP = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{bmatrix}$
$= P \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$
Therefore

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

(1) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4): We will skip it. $\square$

Cayley-Hamilton Theorem

- Let A be an $n \times n$ matrix over $\mathbb{R}$.
- Write A^r for the matrix multiplication of r many copies of A.
- For $c \in \mathbb{R}$, cA is just component wise scalar multiplication.
- If $f(x) = a_r x^r + \cdots + a_2 x^2 + a_1 x + a_0 \in \mathbb{R}[x]$, then

  $f(A) = a_r A^r + \cdots + a_2 A^2 + a_1 A + a_0 I_n$ is an $n \times n$ matrix/$\mathbb{R}$.

Theorem (Cayley-Hamilton)

Consider the characteristic polynomial $p_A(x) := \det(xI_n - A)$. *Then* $p_A(A) = 0$ *(zero matrix of order* $n \times n$*).*

Warning: $p_A(A) \neq \det(AI_n - A)$. LHS is a matrix; RHS is a scalar.

Example

If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, then $p_A(x) = x^2 - 5x - 2$. The Cayley-Hamilton Theorem says that $A^2 - 5A - 2I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Similar matrices have same characteristic polynomial

Theorem

Let A and B be similar, i.e., $B = P^{-1}AP$ for some invertible matrix P. Then $\det(xI_n - A) = \det(xI_n - B)$.

Proof. $\det(xI_n - B) = \det(xP^{-1}I_nP - P^{-1}AP) = \det(P^{-1}(xI_n - A)P)$
$= (1/\det(P))\det(xI_n - A)\det(P) = \det(xI_n - A)$. □

Theorem

Let A and B be similar, i.e., $B = P^{-1}AP$. For a polynomial $f(x) \in \mathbb{R}[x]$, $f(A) = 0$ if and only if $f(B) = 0$ (zero matrix).

Proof. Note: $P^{-1}A^rP = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = B^r$,
and $P^{-1}(c_1D_1 + c_2D_2)P = c_1(P^{-1}D_1P) + c_2(P^{-1}D_2P)$.
Verify that $P^{-1}f(A)P = f(B)$ and $Pf(B)P^{-1} = f(A)$.
Hence the theorem follows. □

Proof of the Cayley-Hamilton Theorem for diagonalizable matrix

Let A be a diagonalizable matrix, i.e., there is P such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = B \text{ (say)}.$$

Note that $\det(xI_n - A) = \det(xI_n - B) = (x - \lambda_1)\cdots(x - \lambda_n)$.
By induction on n, one can verify that

$$(B - \lambda_1 I_n)(B - \lambda_2 I_n)\cdots(B - \lambda_n I_n) = 0.$$

Hence, multiplying P on left and P^{-1} on right, we get

$$(A - \lambda_1 I_n)(A - \lambda_2 I_n)\cdots(A - \lambda_n I_n) = 0$$

i.e., $p_A(A) = 0$ (zero matrix). $\square$

Remark

It can be observed from the proof that if A is diagonalizable, then A satisfies a polynomial having distinct roots.

Some real life applications of the use of eigenvalues and eigenvectors in science, engineering and computer science can be found here:
`https://www.intmath.com/matrices-determinants/8-applications-eigenvalues-eigenvectors.php`

Thank You!