

MATH 311

Topics in Applied Mathematics I

Lecture 16:
Diagonalization.

Eigenvalues and eigenvectors of a matrix

Definition. Let A be an $n \times n$ matrix. A number $\lambda \in \mathbb{R}$ is called an **eigenvalue** of the matrix A if $A\mathbf{v} = \lambda\mathbf{v}$ for a nonzero column vector $\mathbf{v} \in \mathbb{R}^n$.

The vector \mathbf{v} is called an **eigenvector** of A belonging to (or associated with) the eigenvalue λ .

If λ is an eigenvalue of A then the nullspace $N(A - \lambda I)$, which is nontrivial, is called the **eigenspace** of A corresponding to λ . The eigenspace consists of all eigenvectors belonging to the eigenvalue λ plus the zero vector.

Characteristic equation

Definition. Given a square matrix A , the equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** of A .

Eigenvalues λ of A are roots of the characteristic equation.

If A is an $n \times n$ matrix then $p(\lambda) = \det(A - \lambda I)$ is a polynomial of degree n . It is called the **characteristic polynomial** of A .

Theorem Any $n \times n$ matrix has at most n eigenvalues.

Eigenvalues and eigenvectors of an operator

Definition. Let V be a vector space and $L : V \rightarrow V$ be a linear operator. A number λ is called an **eigenvalue** of the operator L if $L(\mathbf{v}) = \lambda\mathbf{v}$ for a nonzero vector $\mathbf{v} \in V$. The vector \mathbf{v} is called an **eigenvector** of L associated with the eigenvalue λ . (If V is a functional vector space then eigenvectors are usually called **eigenfunctions**.)

If $V = \mathbb{R}^n$ then the linear operator L is given by $L(\mathbf{x}) = A\mathbf{x}$, where A is an $n \times n$ matrix (and \mathbf{x} is regarded a column vector). In this case, eigenvalues and eigenvectors of the operator L are precisely eigenvalues and eigenvectors of the matrix A .

Suppose $L : V \rightarrow V$ is a linear operator on a **finite-dimensional** vector space V .

Let $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be a basis for V and $g : V \rightarrow \mathbb{R}^n$ be the corresponding coordinate mapping. Let A be the matrix of L with respect to this basis. Then

$$L(\mathbf{v}) = \lambda \mathbf{v} \iff A g(\mathbf{v}) = \lambda g(\mathbf{v}).$$

Hence the eigenvalues of L coincide with those of the matrix A . Moreover, the associated eigenvectors of A are coordinates of the eigenvectors of L .

Definition. The characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ of the matrix A is called the **characteristic polynomial** of the operator L .

Then eigenvalues of L are roots of its characteristic polynomial.

Theorem. The characteristic polynomial of the operator L is well defined. That is, it does not depend on the choice of a basis.

Proof: Let B be the matrix of L with respect to a different basis $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Then $A = UBU^{-1}$, where U is the transition matrix from the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ to $\mathbf{u}_1, \dots, \mathbf{u}_n$. We have to show that $\det(A - \lambda I) = \det(B - \lambda I)$ for all $\lambda \in \mathbb{R}$. We obtain

$$\begin{aligned} \det(A - \lambda I) &= \det(UBU^{-1} - \lambda I) \\ &= \det(UBU^{-1} - U(\lambda I)U^{-1}) = \det(U(B - \lambda I)U^{-1}) \\ &= \det(U) \det(B - \lambda I) \det(U^{-1}) = \det(B - \lambda I). \end{aligned}$$

Basis of eigenvectors

Let V be a finite-dimensional vector space and $L : V \rightarrow V$ be a linear operator. Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be a basis for V and A be the matrix of the operator L with respect to this basis.

Theorem The matrix A is diagonal if and only if vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ are eigenvectors of L .

If this is the case, then the diagonal entries of the matrix A are the corresponding eigenvalues of L .

$$L(\mathbf{v}_i) = \lambda_i \mathbf{v}_i \iff A = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \dots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Theorem If $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are eigenvectors of a linear operator L associated with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$, then $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ are linearly independent.

Corollary 1 If $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct real numbers, then the functions $e^{\lambda_1 x}, e^{\lambda_2 x}, \dots, e^{\lambda_k x}$ are linearly independent.

Proof: Consider a linear operator $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ given by $Df = f'$. Then $e^{\lambda_1 x}, \dots, e^{\lambda_k x}$ are eigenfunctions of D associated with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. By the theorem, the eigenfunctions are linearly independent.

How to find a basis of eigenvectors

Corollary 2 Suppose $\lambda_1, \lambda_2, \dots, \lambda_k$ are all eigenvalues of a linear operator $L : V \rightarrow V$. For any $1 \leq i \leq k$, let S_i be a basis for the eigenspace associated to the eigenvalue λ_i . Then these bases are disjoint and the union $S = S_1 \cup S_2 \cup \dots \cup S_k$ is a linearly independent set.

Moreover, if the vector space V admits a basis consisting of eigenvectors of L , then S is such a basis.

Corollary 3 Let A be an $n \times n$ matrix such that the characteristic equation $\det(A - \lambda I) = 0$ has n distinct roots. Then **(i)** there is a basis for \mathbb{R}^n consisting of eigenvectors of A ; **(ii)** all eigenspaces of A are one-dimensional.

Diagonalization

Theorem 1 Let L be a linear operator on a finite-dimensional vector space V . Then the following conditions are equivalent:

- the matrix of L with respect to some basis is diagonal;
- there exists a basis for V formed by eigenvectors of L .

The operator L is **diagonalizable** if it satisfies these conditions.

Theorem 2 Let A be an $n \times n$ matrix. Then the following conditions are equivalent:

- A is the matrix of a diagonalizable operator;
- A is similar to a diagonal matrix, i.e., it is represented as $A = UBU^{-1}$, where the matrix B is diagonal;
- there exists a basis for \mathbb{R}^n formed by eigenvectors of A .

The matrix A is **diagonalizable** if it satisfies these conditions.

Example. $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

- The matrix A has two eigenvalues: 1 and 3.
- The eigenspace of A associated with the eigenvalue 1 is the line spanned by $\mathbf{v}_1 = (-1, 1)$.
- The eigenspace of A associated with the eigenvalue 3 is the line spanned by $\mathbf{v}_2 = (1, 1)$.
- Eigenvectors \mathbf{v}_1 and \mathbf{v}_2 form a basis for \mathbb{R}^2 .

Thus the matrix A is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad U = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Notice that U is the transition matrix from the basis $\mathbf{v}_1, \mathbf{v}_2$ to the standard basis.

Example. $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

- The matrix A has two eigenvalues: 0 and 2.
- The eigenspace for 0 is one-dimensional; it has a basis $S_1 = \{\mathbf{v}_1\}$, where $\mathbf{v}_1 = (-1, 1, 0)$.
- The eigenspace for 2 is two-dimensional; it has a basis $S_2 = \{\mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_2 = (1, 1, 0)$, $\mathbf{v}_3 = (-1, 0, 1)$.
- The union $S_1 \cup S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set, hence it is a basis for \mathbb{R}^3 .

Thus the matrix A is diagonalizable. Namely, $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad U = \begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

There are **two obstructions** to existence of a basis consisting of eigenvectors. They are illustrated by the following examples.

Example 1. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

$\det(A - \lambda I) = (\lambda - 1)^2$. Hence $\lambda = 1$ is the only eigenvalue. The associated eigenspace is the line $t(1, 0)$.

Example 2. $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$\det(A - \lambda I) = \lambda^2 + 1$.

\implies no real eigenvalues or eigenvectors

(However there are *complex* eigenvalues/eigenvectors.)

To *diagonalize* an $n \times n$ matrix A is to find a diagonal matrix B and an invertible matrix U such that $A = UBU^{-1}$.

Suppose there exists a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ for \mathbb{R}^n consisting of eigenvectors of A . That is, $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$, where $\lambda_k \in \mathbb{R}$.

Then $A = UBU^{-1}$, where $B = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ and U is a transition matrix whose columns are vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

Example. $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. $\det(A - \lambda I) = (4 - \lambda)(1 - \lambda)$.

Eigenvalues: $\lambda_1 = 4$, $\lambda_2 = 1$.

Associated eigenvectors: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Thus $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Suppose we have a problem that involves a square matrix A in the context of matrix multiplication.

Also, suppose that the case when A is a diagonal matrix is simple. Then the diagonalization may help in solving this problem (or may not). Namely, it may reduce the case of a diagonalizable matrix to that of a diagonal one.

An example of such a problem is, given a square matrix A , to find its power A^k :

$$A = \begin{pmatrix} s_1 & & & 0 \\ & s_2 & & \\ & & \ddots & \\ 0 & & & s_n \end{pmatrix} \implies A^k = \begin{pmatrix} s_1^k & & & 0 \\ & s_2^k & & \\ & & \ddots & \\ 0 & & & s_n^k \end{pmatrix}$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find A^5 .

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Then $A^5 = UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}UBU^{-1}$

$$= UB^5U^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1024 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1024 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1024 & 1023 \\ 0 & 1 \end{pmatrix}.$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find A^k ($k \geq 1$).

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} A^k &= UB^kU^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4^k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4^k & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4^k & 4^k - 1 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Problem. Let $A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}$. Find a matrix C such that $C^2 = A$.

We know that $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Suppose that $D^2 = B$ for some matrix D . Let $C = UDU^{-1}$. Then $C^2 = UDU^{-1}UDU^{-1} = UD^2U^{-1} = UBU^{-1} = A$.

We can take $D = \begin{pmatrix} \sqrt{4} & 0 \\ 0 & \sqrt{1} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

Then $C = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$.

Initial value problem for a system of linear ODEs:

$$\begin{cases} \frac{dx}{dt} = 4x + 3y, \\ \frac{dy}{dt} = y, \end{cases} \quad x(0) = 1, \quad y(0) = 1.$$

The system can be rewritten in vector form:

$$\frac{d\mathbf{v}}{dt} = A\mathbf{v}, \quad \text{where } A = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Matrix A is diagonalizable: $A = UBU^{-1}$, where

$$B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Let $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ be coordinates of the vector \mathbf{v} relative to the basis $\mathbf{v}_1 = (1, 0)$, $\mathbf{v}_2 = (-1, 1)$ of eigenvectors of A . Then $\mathbf{v} = U\mathbf{w} \implies \mathbf{w} = U^{-1}\mathbf{v}$.

It follows that

$$\frac{d\mathbf{w}}{dt} = \frac{d}{dt}(U^{-1}\mathbf{v}) = U^{-1}\frac{d\mathbf{v}}{dt} = U^{-1}A\mathbf{v} = U^{-1}AU\mathbf{w}.$$

$$\text{Hence } \frac{d\mathbf{w}}{dt} = B\mathbf{w} \iff \begin{cases} \frac{dw_1}{dt} = 4w_1, \\ \frac{dw_2}{dt} = w_2. \end{cases}$$

General solution: $w_1(t) = c_1 e^{4t}$, $w_2(t) = c_2 e^t$, where $c_1, c_2 \in \mathbb{R}$.

Initial condition:

$$\mathbf{w}(0) = U^{-1}\mathbf{v}(0) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Thus $w_1(t) = 2e^{4t}$, $w_2(t) = e^t$. Then

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = U\mathbf{w}(t) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2e^{4t} \\ e^t \end{pmatrix} = \begin{pmatrix} 2e^{4t} - e^t \\ e^t \end{pmatrix}.$$