MATH 423

Linear Algebra II

Lecture 21:
Eigenvalues and eigenvectors (continued).
Diagonalization.

Eigenvalues and eigenvectors of a matrix

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

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 $\mathcal{N}(A - \lambda I)$, which is nontrivial, is called the **eigenspace** of A corresponding to λ (denoted $\mathcal{E}_\lambda$). The eigenspace $\mathcal{E}_\lambda$ consists of all eigenvectors belonging to the eigenvalue λ and 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 $L : V \to V$ be a linear operator. A number λ is called an **eigenvalue** of the operator L if $\boxed{L(\mathbf{v}) = \lambda\mathbf{v}}$ 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 space then eigenvectors are usually called **eigenfunctions**.

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

Suppose $L : V \to V$ is a linear operator on a **finite-dimensional** vector space V. Let α be an ordered basis for V. Then

$$L(\mathbf{v}) = \lambda\mathbf{v} \iff [L]_\alpha[\mathbf{v}]_\alpha = \lambda[\mathbf{v}]_\alpha.$$

Hence the eigenvalues of L coincide with those of the matrix $[L]_\alpha$. Moreover, the associated eigenvectors of $[L]_\alpha$ are coordinates of the eigenvectors of L. As a consequence, the number of eigenvalues of L cannot exceed $\dim V$.

Definition. The characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ of the matrix $A = [L]_\alpha$ 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 A and B be matrices of L with respect to different bases α and β. Then $B = UAU^{-1}$, where $U = [\mathrm{id}_V]_\alpha^\beta$ is the transition matrix that changes coordinates from the basis α to β. We have to show that $\det(B - \lambda I) = \det(A - \lambda I)$ for all $\lambda \in \mathbb{F}$. Indeed,

$$\det(B - \lambda I) = \det(UAU^{-1} - \lambda I)$$

$$= \det\big(UAU^{-1} - U(\lambda I)U^{-1}\big) = \det\big(U(A - \lambda I)U^{-1}\big)$$

$$= \det(U)\,\det(A - \lambda I)\,\det(U^{-1}) = \det(A - \lambda I).$$

Eigenspaces

Let $L : V \to V$ be a linear operator.

For any $\lambda \in \mathbb{F}$, let $\mathcal{E}_\lambda$ denotes the set of all solutions of the equation $L(\mathbf{x}) = \lambda\mathbf{x}$.

Then $\mathcal{E}_\lambda$ is a *subspace* of V since $\mathcal{E}_\lambda$ is the *nullspace* of a linear operator given by $\mathbf{x} \mapsto L(\mathbf{x}) - \lambda\mathbf{x}$.

$\mathcal{E}_\lambda$ minus the zero vector is the set of all eigenvectors of L associated with the eigenvalue λ. In particular, $\lambda \in \mathbb{F}$ is an eigenvalue of L if and only if $\mathcal{E}_\lambda \neq \{\mathbf{0}\}$.

If $\mathcal{E}_\lambda \neq \{\mathbf{0}\}$ then it is called the **eigenspace** of L corresponding to the eigenvalue λ.

Example. $V = C^\infty(\mathbb{R})$, $D : V \to V$, $Df = f'$.

A function $f \in C^\infty(\mathbb{R})$ is an eigenfunction of the operator D belonging to an eigenvalue λ if $f'(x) = \lambda f(x)$ for all $x \in \mathbb{R}$.

It follows that $f(x) = ce^{\lambda x}$, where c is a nonzero constant.

Thus each $\lambda \in \mathbb{R}$ is an eigenvalue of D.
The corresponding eigenspace is spanned by $e^{\lambda x}$.

Remark. If we consider D as an operator on the complex vector space $C^\infty(\mathbb{R}, \mathbb{C})$ then, similarly, each $\lambda \in \mathbb{C}$ is an eigenvalue of D and the corresponding eigenspace is spanned by $e^{\lambda x}$.

Let V be a vector space and $L : V \to V$ be a linear operator.

Proposition 1 If $\mathbf{v} \in V$ is an eigenvector of the operator L then the associated eigenvalue is unique.

Proof: Suppose that $L(\mathbf{v}) = \lambda_1 \mathbf{v}$ and $L(\mathbf{v}) = \lambda_2 \mathbf{v}$. Then
$$\lambda_1 \mathbf{v} = \lambda_2 \mathbf{v} \implies (\lambda_1 - \lambda_2)\mathbf{v} = \mathbf{0} \implies \lambda_1 - \lambda_2 = 0 \implies \lambda_1 = \lambda_2.$$

Proposition 2 Suppose $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of L associated with different eigenvalues λ_1 and λ_2. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Proof: For any scalar $t \neq 0$ the vector $t\mathbf{v}_1$ is also an eigenvector of L associated with the eigenvalue λ_1. Since $\lambda_2 \neq \lambda_1$, it follows that $\mathbf{v}_2 \neq t\mathbf{v}_1$. That is, $\mathbf{v}_2$ is not a scalar multiple of $\mathbf{v}_1$. Similarly, $\mathbf{v}_1$ is not a scalar multiple of $\mathbf{v}_2$.

Let $L : V \to V$ be a linear operator.

Proposition 3 If $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are eigenvectors of L associated with distinct eigenvalues λ_1, λ_2, and λ_3, then they are linearly independent.

Proof: Suppose that $t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + t_3\mathbf{v}_3 = \mathbf{0}$ for some $t_1, t_2, t_3 \in \mathbb{F}$. Then

$$\begin{gathered} L(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + t_3\mathbf{v}_3) = \mathbf{0}, \\ t_1 L(\mathbf{v}_1) + t_2 L(\mathbf{v}_2) + t_3 L(\mathbf{v}_3) = \mathbf{0}, \\ t_1\lambda_1\mathbf{v}_1 + t_2\lambda_2\mathbf{v}_2 + t_3\lambda_3\mathbf{v}_3 = \mathbf{0}. \end{gathered}$$

It follows that

$$\begin{gathered} t_1\lambda_1\mathbf{v}_1 + t_2\lambda_2\mathbf{v}_2 + t_3\lambda_3\mathbf{v}_3 - \lambda_3(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + t_3\mathbf{v}_3) = \mathbf{0} \\ \implies t_1(\lambda_1 - \lambda_3)\mathbf{v}_1 + t_2(\lambda_2 - \lambda_3)\mathbf{v}_2 = \mathbf{0}. \end{gathered}$$

By the above, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.
Hence $t_1(\lambda_1 - \lambda_3) = t_2(\lambda_2 - \lambda_3) = 0 \implies t_1 = t_2 = 0$
Then $t_3 = 0$ as well.

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 $D : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ 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.

Corollary 2 If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of a matrix A associated with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.

Corollary 3 Let A be an $n \times n$ matrix such that the characteristic equation $\det(A - \lambda I) = 0$ has n distinct real roots. Then $\mathbb{F}^n$ has a basis consisting of eigenvectors of A.

Proof: Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct real roots of the characteristic equation. Any λ_i is an eigenvalue of A, hence there is an associated eigenvector $\mathbf{v}_i$. By Corollary 2, vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Therefore they form a basis for $\mathbb{F}^n$.

Basis of eigenvectors

Let V be a finite-dimensional vector space and $L : V \to V$ 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.

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

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 $\mathbb{F}^n$ 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

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

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

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.)