MATH 311

Topics in Applied Mathematics I

Lecture 22: Eigenvalues and eigenvectors of a linear operator.

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 $\boxed{A\mathbf{v} = \lambda\mathbf{v}}$ 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.

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

Characteristic equation:

$$\begin{vmatrix} 1-\lambda & 1 & -1 \\ 1 & 1-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0.$$

Expand the determinant by the 3rd row:

$$(2-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0.$$

$$\big((1-\lambda)^2 - 1\big)(2-\lambda) = 0 \iff -\lambda(2-\lambda)^2 = 0$$

$$\implies \quad \lambda_1 = 0, \;\; \lambda_2 = 2.$$

$$A\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Convert the matrix to reduced row echelon form:

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A\mathbf{x} = \mathbf{0} \iff \begin{cases} x + y = 0, \\ z = 0. \end{cases}$$

The general solution is $(-t, t, 0) = t(-1, 1, 0)$, $t \in \mathbb{R}$. Thus $\mathbf{v}_1 = (-1, 1, 0)$ is an eigenvector associated with the eigenvalue 0. The corresponding eigenspace is the line spanned by $\mathbf{v}_1$.

$$(A - 2I)\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\iff \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff x - y + z = 0.$$

The general solution is $x = t - s$, $y = t$, $z = s$, where $t, s \in \mathbb{R}$. Equivalently,

$$\mathbf{x} = (t - s, t, s) = t(1, 1, 0) + s(-1, 0, 1).$$

Thus $\mathbf{v}_2 = (1, 1, 0)$ and $\mathbf{v}_3 = (-1, 0, 1)$ are eigenvectors associated with the eigenvalue 2. The corresponding eigenspace is the plane spanned by $\mathbf{v}_2$ and $\mathbf{v}_3$.

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

- The matrix A has two eigenvalues: 0 and 2.
- The eigenvalue 0 is *simple:* the corresponding eigenspace is a line.
- The eigenvalue 2 is of *multiplicity* 2: the corresponding eigenspace is a plane.
- Eigenvectors $\mathbf{v}_1 = (-1, 1, 0)$, $\mathbf{v}_2 = (1, 1, 0)$, and $\mathbf{v}_3 = (-1, 0, 1)$ of the matrix A form a basis for $\mathbb{R}^3$.
- Geometrically, the map $\mathbf{x} \mapsto A\mathbf{x}$ is the projection on the plane $\mathrm{Span}(\mathbf{v}_2, \mathbf{v}_3)$ along the lines parallel to $\mathbf{v}_1$ with the subsequent scaling by a factor of 2.

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 also 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.
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 $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ be a basis for V and $g : V \to \mathbb{R}^n$ 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

$$\det(A - \lambda I) = \det(UBU^{-1} - \lambda I)$$

$$= \det\big(UBU^{-1} - U(\lambda I)U^{-1}\big) = \det\big(U(B - \lambda I)U^{-1}\big)$$

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

Eigenspaces

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

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

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

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

If $V_\lambda \neq \{\mathbf{0}\}$ 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 $e^{\lambda x}$.

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

$Lf = \lambda f \iff f''(x) - \lambda f(x) = 0$ for all $x \in \mathbb{R}$.

It follows that each $\lambda \in \mathbb{R}$ is an eigenvalue of L and the corresponding eigenspace V_λ is two-dimensional. Note that $L = D^2$, hence $Df = \mu f \implies Lf = \mu^2 f$.

If $\lambda > 0$ then $V_\lambda = \mathrm{Span}(e^{\mu x}, e^{-\mu x})$, where $\mu = \sqrt{\lambda}$.

If $\lambda < 0$ then $V_\lambda = \mathrm{Span}\big(\sin(\mu x), \cos(\mu x)\big)$, where $\mu = \sqrt{-\lambda}$.

If $\lambda = 0$ then $V_\lambda = \mathrm{Span}(1, x)$.

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 $t_1, t_2, t_3 \in \mathbb{R}$. Then

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

It follows that

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

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 $\mathbb{R}^n$ 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 $\mathbb{R}^n$.