Lecture 4. Eigenvectors and Eigenvalues
September 2022

EXAMPLE 4.1 Given a linear operator $f : \mathbb{R}^2 \to \mathbb{R}^2$,

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+3y \\ -x+5y \end{pmatrix},$$

describe the effect of it on the xy-plane.

Solution. Let's change the basis in $\mathbb{R}^2$ to a new basis

$$B = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}$$

and then find the matrix of the linear operator in this basis. Call this matrix C. Then $C = P^{-1}AP$, where A is the matrix of the linear operator in standard coordinates and P is the transition matrix from standard coordinates to B-coordinates:

$$A = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}.$$

Then

$$C = P^{-1}AP = \frac{1}{2}\begin{pmatrix} -1 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

The B-coordinates of the B-basis vectors are

$$[\mathbf{v}_1]_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad [\mathbf{v}_2]_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

so in B-coordinates, the linear operator can be described as a stretch in the direction of $\mathbf{v}_1$ by a factor of 4 and a stretch in the direction of $\mathbf{v}_2$ by a factor of 2:

$$[f(\mathbf{v}_1)]_B = 4[\mathbf{v}_1]_B, \qquad [f(\mathbf{v}_2)]_B = 2[\mathbf{v}_2]_B.$$

But the effect of f is the same no matter what basis is being used to describe it; it is only the matrices which change. So this statement must be true even in standard coordinates; that is, we must have

$$A\mathbf{v}_1 = 4\mathbf{v}_1, \qquad A\mathbf{v}_2 = 2\mathbf{v}_2.$$

In this example, we were told the basis to use in $\mathbb{R}^2$ in order to solve the question. In what follows, we will try to answer a question: "How did we know which basis would work?"

Eigenvalues and Eigenvectors

EXAMPLE 4.2 An **eigenvector** (or **characteristic vector**) of a linear operator is a nonzero vector that changes at most by a scalar factor when that linear operator is applied to it. The corresponding **eigenvalue**, often denoted by λ is the factor by which the eigenvector is scaled.

Geometrically, an eigenvector, corresponding to a real nonzero eigenvalue, points in a direction in which it is stretched by the operator and the eigenvalue is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed. Loosely speaking, the eigenvector is not rotated. Now we will give a formal definition.

DEFINITION 4.2' If f is a linear operator on a vector space V and $\mathbf{v}$ is a nonzero vector in V, then $\mathbf{v}$ is an **eigenvector** of f if $f(\mathbf{v})$ is a scalar multiple of $\mathbf{v}$. This can be written as

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

where λ is a scalar in $\mathbb{R}$, known as the **eigenvalue**. Any non-zero vector $\mathbf{x}$ for which this equation holds is called an **eigenvector for eigenvalue** λ or an **eigenvector of** f **corresponding to eigenvalue** λ or λ**-eigenvector** for short.

Now we need a method for computing eigenvalues and eigenvectors. Suppose the linear operator $f : V \to V$ has a matrix A with respect to some basis.

DEFINITION 4.3 The polynomial $p_A(\lambda) = |A - \lambda I|$ is called the **characteristic polynomial** of A, and the equation $p_A(\lambda) = 0$ is called the **characteristic equation** of A.

THEOREM 4.4 Let A be an $n \times n$ matrix.

1. The eigenvalues λ of A are the roots of the characteristic polynomial $p_A(\lambda)$ of A.
2. The λ-eigenvectors $\mathbf{x}$ are the nonzero solutions to the homogeneous system

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

Proof. To determine whether λ is an eigenvalue of A, we need to determine whether there are any non-zero solutions $\mathbf{x}$ to the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$. Note that the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$ is not of the standard form, since the right-hand side is not a fixed vector $\mathbf{b}$, but depends explicitly on $\mathbf{x}$. However, we can rewrite it in standard form. Note that $\lambda\mathbf{x} = \lambda I\mathbf{x}$, where I is, as usual, the identity matrix. So, the equation is equivalent to $A\mathbf{x} = \lambda I\mathbf{x}$, or $A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0}$, which is equivalent to $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

Now, a square linear system $B\mathbf{x} = \mathbf{0}$ has solutions other than $\mathbf{x} = \mathbf{0}$ precisely when $|B| = 0$. Therefore, taking $B = A - \lambda I$, λ is an eigenvalue if and only if the determinant of the matrix $A - \lambda I$ is zero. This determinant, $p_A(\lambda) = |A - \lambda I|$ is a polynomial of degree n in the variable λ. □

EXAMPLE 4.5 Find eigenvalues and eigenvectors for the matrix

$$A = \begin{pmatrix} 7 & -15 \\ 2 & -4 \end{pmatrix}.$$

Solution.

$$A - \lambda I = \begin{pmatrix} 7 & -15 \\ 2 & -4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 7-\lambda & -15 \\ 2 & -4-\lambda \end{pmatrix}$$

and the characteristic polynomial is

$$|A - \lambda I| = \lambda^2 - 3\lambda + 2.$$

So the eigenvalues are the solutions of $\lambda^2 - 3\lambda + 2 = 0$. To solve this for λ, we could use either the formula for the solutions to a quadratic equation, or simply observe that the characteristic polynomial factorises. We have $(\lambda - 1)(\lambda - 2) = 0$ with solutions $\lambda = 1$ and $\lambda = 2$. Hence the eigenvalues of A are 1 and 2, and these are the only eigenvalues of A.

To find an eigenvector for each eigenvalue λ, we have to find a nontrivial solution to $(A - \lambda I)\mathbf{x} = \mathbf{0}$, meaning a solution other than the zero vector. This is easy, since for a particular value of λ, all we need to do is solve a simple linear system.

To find the eigenvectors for eigenvalue 1, we solve the system $(A - I)\mathbf{x} = \mathbf{0}$. We do this by putting the coefficient matrix $A - I$ into reduced echelon form. This system has solutions

$$\mathbf{v} = t\begin{pmatrix} 5 \\ 2 \end{pmatrix} \quad \text{for any} \quad t \in \mathbb{R}.$$

There are infinitely many eigenvectors for 1: for each $t \neq 0$ $\mathbf{v}$ is an eigenvector of A corresponding to $\lambda = 1$. But we can't choose $t = 0$; for then $\mathbf{v}$ becomes the zero vector, and this is never an eigenvector, simply by definition.

To find the eigenvectors for 2, we solve $(A - 2I)\mathbf{x} = \mathbf{0}$ by reducing the coefficient matrix. Setting the non-leading variable equal to t, we obtain the solutions

$$\mathbf{v} = t\begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{for any} \quad t \in \mathbb{R}.$$

Any non-zero scalar multiple of the vector $(3, 1)^T$ is an eigenvector of A for eigenvalue 2.

EXAMPLE 4.6 Show that the eigenvalues of a triangular matrix are the main diagonal entries.

Solution. Assume that A is triangular. Then the matrix $A - \lambda I$ is also triangular and has diagonal entries $(a_{11} - \lambda)$, $(a_{22} - \lambda)$, $\ldots$, $(a_{nn} - \lambda)$ where $A = [a_{ij}]$. Hence

$$p_A(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda)$$

and the result follows because the eigenvalues are the roots of $p_A(\lambda)$.

THEOREM 4.7 If A and B are similar $n \times n$-matrices, then A and B have the same characteristic polynomial and eigenvalues.

Proof. For the characteristic polynomial,

$$\det(B - \lambda I) = \det\big(P^{-1}AP - \lambda(P^{-1}IP)\big) = \det\big(P^{-1}(A - \lambda I)P\big) = \det(A - \lambda I).$$

Finally, this shows that A and B have the same eigenvalues because the eigenvalues of a matrix are the roots of its characteristic polynomial. $\square$

From this theorem it follows that the characteristic polynomial for a linear operator $f : V \to V$ does not depend on basis and matrix representation, so we will denote it simply by $p(\lambda)$.

THEOREM 4.8 If A is a square matrix, then A and A^T have the same characteristic polynomial, and hence the same eigenvalues.

Proof. We use the fact that $A^T - \lambda I = (A - \lambda I)^T$. Then

$$p_{A^T}(\lambda) = |A^T - \lambda I| = |(A - \lambda I)^T| = |A - \lambda I| = p_A(\lambda).$$

Hence $p_{A^T}(\lambda)$ and $p_A(\lambda)$ have the same roots, and so A^T and A have the same eigenvalues. $\square$

Diagonalization

Recall that an $n \times n$-matrix D is called a **diagonal** matrix if all its entries off the main diagonal are zero.

DEFINITION 4.9 A square matrix A is **diagonalizable** if there exists an invertible matrix P such that $P^{-1}AP = D$ is a diagonal matrix, that is if A is similar to a diagonal matrix D. Here the invertible matrix P is called a **diagonalizing matrix** for A.

THEOREM 4.10 Let A be an $n \times n$ matrix.

1. A is diagonalizable if and only if $\mathbb{R}^n$ has a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ consisting of eigenvectors of A.
2. When this is the case, the matrix $P = (\mathbf{x}_1 \; \mathbf{x}_2 \; \ldots \; \mathbf{x}_n)$ is invertible and

$$P^{-1}AP = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

where, for each i, λ_i is the eigenvalue of A corresponding to $\mathbf{x}_i$.

Proof. 1) Let f be a linear operator induced by matrix A with respect to some basis. If A is diagonalizable, then there is some new basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, in which the matrix of f is diagonal: $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Since the columns of D are images of basis vectors $\mathbf{x}_i$ under f, then $f(\mathbf{x}_i) = \lambda_i \mathbf{x}_i$ for each i, $1 \leqslant i \leqslant n$. Thus, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a basis consisting of eigenvectors and λ_i are corresponding eigenvalues.

Conversely, if there is a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, consisting of eigenvectors, then the matrix describing f with respect to this basis is diagonal. Denote this matrix by $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Write P in terms of its columns as follows:

$$P = (\mathbf{x}_1 \; \mathbf{x}_2 \; \ldots \; \mathbf{x}_n)$$

and observe that $P^{-1}AP = D$. $\square$

DEFINITION 4.11 A linear operator $f : V \to V$ is **diagonalizable** if V has a basis of eigenvectors of f.

Diagonalization Algorithm To diagonalize an $n \times n$ matrix A:

1. Find the distinct eigenvalues λ of A.
2. Compute the basic eigenvectors corresponding to each of these eigenvalues λ as basic solutions of the homogeneous system $(A - \lambda I)\mathbf{x} = \mathbf{0}$.
3. The matrix A is diagonalizable if and only if there are n basic eigenvectors in all.
4. If A is diagonalizable, the $n \times n$ matrix P with these basic eigenvectors as its columns is a diagonalizing matrix for A, that is, P is invertible and $P^{-1}AP$ is diagonal.

Unfortunately, not all matrices are diagonalizable.

EXAMPLE 4.12 Show that

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

is not diagonalizable.

Solution. The characteristic polynomial is $p_A(\lambda) = (\lambda - 1)^2$, so A has only one eigenvalue $\lambda_1 = 1$ of multiplicity 2. But the system of equations $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has general solution $t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, so there is only one parameter, and so only one basic eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Hence A is not diagonalizable.