Math 51
Lecture 23 – August 8, 2023

Topic(s): eigenvalues, eigenvectors

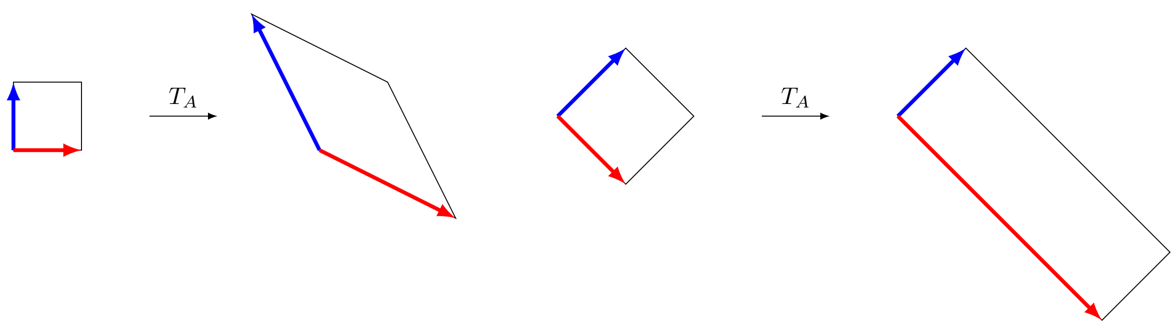
Let A be an $n \times n$ matrix. A vector $\mathbf{v} \in \mathbb{R}^n$ is called an **eigenvector** of A if it is *non-zero* and there is a scalar $\lambda \in \mathbb{R}$ for which $A\mathbf{v} = \lambda\mathbf{v}$. The prefix "eigen" is German for "characteristic" or "own" (in the sense of ownership).

If $\mathbf{v}$ is an eigenvector for an $n \times n$ matrix A, the scalar λ for which $A\mathbf{v} = \lambda\mathbf{v}$ is called the **eigenvalue** of A associated with $\mathbf{v}$.

Example 1. The 2×2 matrix $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ has eigenvectors $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. What is the eigenvalue associated with each eigenvector?

We see that $A\vec{u} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \vec{u}$ and $A\vec{v} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} = 3\vec{v}$.

So, $\lambda_{\vec{u}} = 1$ and $\lambda_{\vec{v}} = 3$.



Not all matrices have eigenvectors!

Example 2. For an angle θ that is neither $0°$ nor $180°$, the rotation matrix

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

has no eigenvectors.

$$A_\theta \vec{v} = \begin{cases} \vec{v} & \theta = 0° \\ -\vec{v} & \theta = 180° \end{cases}$$

For any other value of θ, $A_\theta \vec{v}$ cannot be a scalar multiple of $\vec{v}$.

Example 3. If the null space of an $n \times n$ matrix A contains a non-zero vector $\mathbf{w}$, is $\mathbf{w}$ an eigenvector of A?

If $\vec{0} \neq \vec{w} \in N(A)$, then $A\vec{w} = \vec{0} = 0\vec{w}$. So, $\vec{w}$ is an eigenvector of A with $\lambda_{\vec{w}} = 0$.

Hence, 0 is an eigenvalue of A *precisely when* A is not invertible.

Proposition 23.1.11. A scalar λ is an eigenvalue for an $n \times n$ matrix A precisely when $A - \lambda I_n$ is not invertible, or equivalently, the null space $N(A - \lambda I_n)$ contains a non-zero vector. The eigenvectors for A with eigenvalue λ are the non-zero vectors in $N(A - \lambda I_n)$ (this is called the **λ-eigenspace** for A).

Example 4. Identify the eigenvectors and associated eigenvalues for an $n \times n$ diagonal matrix

$$A = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}.$$

$A - \lambda I_n = \begin{bmatrix} d_1-\lambda & 0 & \cdots & 0 \\ 0 & d_2-\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n-\lambda \end{bmatrix}$ is not invertible (i.e., there is a nonzero vector in $N(A-\lambda I_n)$) only if $d_i - \lambda = 0$ for some i.

$\vec{e}_i$ is an eigenvector for A with eigenvalue d_i for all i ($A\vec{e}_i = d_i\vec{e}_i$).

The theory of eigenvectors gives a way to make a lot of $n \times n$ matrices A "look diagonal" when the effect of the corresponding linear transformation is expressed in an appropriate reference frame. More precisely, *if* there is a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors for A, say with respective eigenvalues $\lambda_1, \ldots, \lambda_n$ (some of which may be equal to each other), then the effect of A "looks diagonal" when n-vectors $\mathbf{x}$ are expressed in terms of this basis.

Example 5. Recall that, from Example 1, $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ has eigenvectors $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with corresponding eigenvalues 1 and 3. If $\mathbf{x} = a\mathbf{u} + b\mathbf{v}$ for some scalars a and b, write $A\mathbf{x}$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

We get

$$A\vec{x} = A(a\vec{u} + b\vec{v}) = aA\vec{u} + bA\vec{v} = a\vec{u} + 3b\vec{v}.$$

More generally, if $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ for a basis of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (when such a basis exists) with $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$, then

$$A\mathbf{x} = c_1 A\mathbf{v}_1 + \cdots + c_n A\mathbf{v}_n = \lambda_1 c_1 \mathbf{v}_1 + \cdots + \lambda_n c_n \mathbf{v}_n.$$

In other words, if $\mathbf{x}$ has coefficients $(c_1, \ldots, c_n)$ in terms of the $\mathbf{v}_j$'s, then $A\mathbf{x}$ has coefficients $(\lambda_1 c_1, \ldots, \lambda_n c_n)$ in terms of the $\mathbf{v}_j$'s. Thus, the effect of the matrix-vector product against A behaves like the diagonal matrix of λ_j's when general n-vectors are *written in terms of the basis of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for A!*

This seems very restrictive however – how do we check if such basis of eigenvectors exist? One of the truly great theorems of mathematics is that for symmetric matrices, such a basis always exists! In fact, such bases also exist for many non-symmetric matrices as well. We will discuss this amazing result and its awesome applications in the remaining lectures of this quarter.

Theorem 23.2.2. Let M be an $n \times n$ upper triangular (or lower triangular) matrix. The eigenvalues of M are exactly the diagonal entries. For each eigenvalue λ, the corresponding eigenvectors are the non-zero vectors in the null space $N(M - \lambda I_n)$: non-zero solutions $\mathbf{x}$ to the upper (or lower) triangular system

$$(M - \lambda I_n)\mathbf{x} = \mathbf{0}$$

of n linear equations in n variables (which we can solve via back-substitution).

Example 6. Let $M = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 8 & 2 \\ 0 & 0 & 2 \end{bmatrix}$. Find the eigenvectors and the corresponding eigenvalues of M.

Since M is upper triangular, the eigenvalues of M are 2 and 8.

$\lambda = 8$.

We see that $M - 8I_3 = \begin{bmatrix} -6 & 3 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & -6 \end{bmatrix}$, and $\begin{bmatrix} -6 & 3 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ leads to $2x_1 = x_2$ and $x_3 = 0$. Hence, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, and so, 8-eigenspace $= \operatorname{span}\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right)$.

$\lambda = 2$.

We get $M - 2I_3 = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 6 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, and $\begin{bmatrix} 0 & 3 & 1 \\ 0 & 6 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ leads to $x_2 = -3x_1$. Thus,

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -3x_1 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, and so, 2-eigenspace $= \operatorname{span}\left(\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$.

One of the ways to approach the problem of finding eigenvalues is by making use of the following observation. Since we are looking for λ's that make $A - \lambda I_n$ not invertible, an equivalent problem is: for which λ's is $\det(A - \lambda I_n) = 0$? It turns out that $\det(A - \lambda I_n)$ is a polynomial in λ (in fact, this polynomial is called the **characteristic polynomial** of A).

Theorem 23.3.1. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. Define its **trace** $\operatorname{tr}(A) = a + d$ to be the sum of its diagonal entries, and its **determinant** $\det(A) = ad - bc$.

(a) The eigenvalues of A in $\mathbb{R}$ are exactly the roots of $P_A(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$. This quadratic polynomial is called the **characteristic polynomial** of A.

(b) For each such root λ, the corresponding eigenvectors are the non-zero vectors in $N(A - \lambda I_2)$.

Corollary 23.3.2. A 2×2 matrix A has an eigenvalue precisely when $P_A(\lambda)$ has a real root, which is to say that its discriminant $\operatorname{tr}(A)^2 - 4\det(A)$ is greater than or equal to 0.

Example 7. Find the characteristic polynomial of A_θ.

Recall that $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Hence,

$$P_{A_\theta}(\lambda) = (\cos\theta - \lambda)^2 + \sin^2\theta = \lambda^2 - 2\cos\theta\,\lambda + 1.$$

The polynomial only has real roots if the discriminant

$$D = b^2 - 4ac = 4(\cos^2\theta - 1) \geq 0,$$

which happens only if $\cos\theta = \pm 1$ ($\theta = 0°, 180°$).

Example 8. Find the eigenvalues of $A = \begin{bmatrix} 3 & 16 \\ 2 & -1 \end{bmatrix}$.

Since $A - \lambda I_2 = \begin{bmatrix} 3-\lambda & 16 \\ 2 & -1-\lambda \end{bmatrix}$, we get

$$\begin{aligned} \det(A - \lambda I_2) &= (3-\lambda)(-1-\lambda) - 32 \\ &= \lambda^2 - 2\lambda - 35 \\ &= (\lambda+5)(\lambda-7). \end{aligned}$$

Thus, $\boxed{\lambda = 5, 7}$.

Example 9. Assume that A is a 2×2 matrix and that $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ are eigenvectors with respective eigenvalues 1 and 0. Calculate $A\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. (Hint: express $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$.)

Following the hint, we get $\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 3 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Hence,

$$A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{2}{3}A\begin{bmatrix} 3 \\ 1 \end{bmatrix} - \frac{1}{3}A\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 3 \\ 1 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \boxed{\begin{bmatrix} 2 \\ \frac{2}{3} \end{bmatrix}}.$$

Example 10. Suppose A has eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ with respective eigenvalues -2 and 1.

(a) What is $P_A(\lambda)$?

Since -2 and 1 are eigenvalues of A,

$$P_A(\lambda) = (\lambda+2)(\lambda-1) = \boxed{\lambda^2 + \lambda - 2}.$$

(b) Find A.

Motivated by Example 9, we see that

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{2}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Thus,

$$A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3}A\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1}{3}A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -2 \\ 2 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\frac{2}{3}A\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1}{3}A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = -\frac{2}{3}\begin{bmatrix} -2 \\ 2 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Hence, $A = \boxed{\begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix}}$.

(c) Show that $P_A(A) = 0$.

We compute

$$\begin{aligned} P_A(A) &= A^2 + A - 2I_2 \\ &= \begin{bmatrix} 2 & -2 \\ -1 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

In general, for any $n \times n$ matrix A, $P_A(A) = 0$. This is a theorem due to Cayley and Hamilton.

Example 11. Let

$$A = \begin{bmatrix} 2 & 0 & -1 & 1 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Check that $\mathbf{e}_1$ is an eigenvector of A. What is the corresponding eigenvalue?

Since $A\vec{e}_1 = 2\vec{e}_1$, we get $\lambda_{\vec{e}_1} = 2$.

(b) Determine the collection of eigenvectors for the eigenvalue 1 (write it as a span). What is its dimension?

We see that $A - I_4 = \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & -2 & 3 & 0 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Solving

$$\begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & -2 & 3 & 0 \\ 0 & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

via back-substitution gives $x_1 = x_4$, $x_2 = 3x_4$, and $x_3 = 2x_4$. Thus, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_4 \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}$,

and so, 1-eigenspace $= \text{span}\left(\begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}\right)$ is 1-dimensional.

(c) Determine the collection of eigenvectors for the eigenvalue -1 (write it as a span). What is its dimension?

$A + I_4 = \begin{bmatrix} 3 & 0 & -1 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ and solving $\begin{bmatrix} 3 & 0 & -1 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ gives $x_1 = 0$, $x_3 = 0$, and

$x_4 = 0$. Hence, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \vec{e}_2$, and so, (-1)-eigenspace $= \text{span}(\vec{e}_2)$ has dimension 1.

Exercise 23.7. Let A be a symmetric $n \times n$ matrix. Assume that $\mathbf{v}$ is an eigenvector with eigenvalue λ, and $\mathbf{w}$ is an eigenvector with eigenvalue μ, and that $\lambda \neq \mu$.

(a) Show that $(\lambda\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (\mu\mathbf{w})$.

(b) Explain why $\mathbf{v}$ and $\mathbf{w}$ are orthogonal. (This is a general feature of symmetric matrices!)