Lesson - Eigenvalues and eigenvectors

Year: University Year 1

Eigenvectors reveal the natural axes of a linear transformation. They appear in vibrations, population models, Google's PageRank, principal component analysis and quantum mechanics.

Starter activity

Take the matrix that doubles the horizontal coordinates and leaves the vertical ones alone: $(x, y) \mapsto (2x, y)$. Draw the unit square and its image. Which directions are only stretched? Now try the matrix which rotates by 90° and see whether any direction is left pointing the same way. What difference do you notice between the two examples?

Eigenvalues and the characteristic equation

Definition: eigenvalue and eigenvector

Let A be a square matrix. A non-zero vector $\mathbf{v}$ is an **eigenvector** of A with **eigenvalue** λ if $A\mathbf{v} = \lambda\mathbf{v}$. The map $\mathbf{x} \mapsto A\mathbf{x}$ only stretches (or flips) $\mathbf{v}$, without changing its line.

Rule: characteristic equation

$A\mathbf{v} = \lambda\mathbf{v}$ is equivalent to $(A - \lambda I)\mathbf{v} = \mathbf{0}$, which has a non-zero solution if and only if $\det(A - \lambda I) = 0$. For a 2×2 matrix this reads $\lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0$.

Property: trace and determinant

The eigenvalues sum to the trace and multiply to the determinant. For a triangular matrix the eigenvalues are the diagonal entries.

Method: finding eigenvalues and eigenvectors

1. Write $\det(A - \lambda I) = 0$ and solve for λ.
2. For each λ, solve $(A - \lambda I)\mathbf{v} = \mathbf{0}$.
3. Give one non-zero vector for each independent solution.
4. Check with the trace and determinant.

Example 1

Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

$\operatorname{tr} A = 7$ and $\det A = 10$, so $\lambda^2 - 7\lambda + 10 = 0$, giving $\lambda = 2$ or 5. For $\lambda = 2$: $\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}\mathbf{v} = \mathbf{0}$ gives $2x + y = 0$, eigenvector $(1, -2)$. For $\lambda = 5$: $\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}\mathbf{v} = \mathbf{0}$ gives $x = y$, eigenvector $(1, 1)$.

Definition: multiplicities

The eigenspace for λ is $\ker(A - \lambda I)$. Its dimension is the *geometric multiplicity*; the multiplicity of λ as a root of the characteristic polynomial is the *algebraic multiplicity*.

Example 2: complex eigenvalues

Find the eigenvalues of the rotation matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$\lambda^2 + 1 = 0$, so $\lambda = \pm i$. There is no real eigenvector: a rotation by 90° fixes no real direction.

Diagonalisation

Theorem: diagonalisation

If A has n linearly independent eigenvectors, put them as the columns of P and the eigenvalues on the diagonal of D. Then $A = PDP^{-1}$ and A is **diagonalisable**, and $A^k = PD^kP^{-1}$. Eigenvectors for *distinct* eigenvalues are always independent, so n distinct eigenvalues guarantee diagonalisability.

Method: computing a power of a matrix

1. Find the eigenvalues and independent eigenvectors.
2. Build P from the eigenvectors and D from the eigenvalues.
3. Compute P^{-1}.
4. Evaluate $A^k = PD^kP^{-1}$.

Example 3

Find A^5 for $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

$\lambda^2 - 4\lambda + 3 = 0$, so $\lambda = 1, 3$, with eigenvectors $(1, -1)$ and $(1, 1)$. So $P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$, and $A^5 = \frac{1}{2}\begin{pmatrix} 3^5 + 1 & 3^5 - 1 \\ 3^5 - 1 & 3^5 + 1 \end{pmatrix} = \begin{pmatrix} 122 & 121 \\ 121 & 122 \end{pmatrix}$.

Watch out

A repeated eigenvalue needs as many independent eigenvectors as its algebraic multiplicity. $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ has eigenvalue 2 twice but only one independent eigenvector $(1, 0)$, so it is *not* diagonalisable. But repeated eigenvalues do not always spoil it: I is already diagonal.

Further facts

Property: eigenvalues of related matrices

- If λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k, and $1/\lambda$ of A^{-1} when A is invertible.
- A matrix is invertible if and only if 0 is not an eigenvalue.
- A real **symmetric** matrix has real eigenvalues, eigenvectors for distinct eigenvalues are orthogonal, and it can be diagonalised by an orthogonal matrix.

Theorem: Cayley-Hamilton

Every matrix satisfies its own characteristic equation. For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $A^2 - 4A + 3I = 0$, so $A^{-1} = \frac{1}{3}(4I - A)$.

Applications

Example 4: Markov chain

Find the long-run state for the transition matrix $\begin{pmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{pmatrix}$.

The eigenvalues are 1 and 0.7. The eigenvector for $\lambda = 1$, scaled to sum 1, is $(\frac{2}{3}, \frac{1}{3})$: the steady state.

Note: Fibonacci

The recurrence $F_{n+1} = F_n + F_{n-1}$ is governed by $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, whose eigenvalues $\frac{1 \pm \sqrt{5}}{2}$ explain the golden-ratio growth.

Interpreting eigenvalues

Key fact: what eigenvalues tell you

An eigenvector of A is a direction that the matrix leaves unchanged apart from stretching, shrinking or reversing it, and the eigenvalue is the stretch factor. If $|\lambda| > 1$ the direction grows under repeated application of A; if $|\lambda|$ it shrinks; if $\lambda = 1$ it is a steady state; and if λ is negative the direction flips at each step. When a matrix can be diagonalised, all of its behaviour is described by independent stretches along its eigenvectors, which is why eigenvalues predict the long-term outcome of Markov chains, population models, vibrating structures and systems of differential equations.

Method: eigenvalues and eigenvectors

1. Form $A - \lambda I$ and set $\det(A - \lambda I) = 0$; solve the characteristic polynomial for λ.
2. For each eigenvalue, solve $(A - \lambda I)\mathbf{v} = \mathbf{0}$ by row reduction. The solutions form a line (or plane) of eigenvectors.
3. Check your work: the trace of A equals the sum of the eigenvalues, and the determinant equals their product.
4. If there are n independent eigenvectors, put them as columns of P and the eigenvalues on the diagonal of D, so that $A = PDP^{-1}$.

Worked example 5: A triangular matrix with a repeated eigenvalue

Find the eigenvalues and eigenvectors of $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ and of $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$.

Both matrices have the characteristic equation $(3 - \lambda)^2 = 0$, so the eigenvalue 3 is repeated. For the first matrix, which is $3I$, every non-zero vector is an eigenvector, so there are two independent eigenvectors and the matrix is already diagonal. For the second, $(A - 3I)\mathbf{v} = \mathbf{0}$ gives $y = 0$, so only multiples of $(1, 0)$ are eigenvectors, and the matrix cannot be diagonalised. The algebraic multiplicity is the same, but the geometric multiplicities differ, and that is what decides diagonalisability.

Worked example 6: A matrix power by diagonalisation

Find a formula for A^n when $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

We found eigenvalues 2 and 5 with eigenvectors $(1, -2)$ and $(1, 1)$. Writing $P = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}$ and $D = \operatorname{diag}(2, 5)$, we have $A^n = PD^nP^{-1}$. The inverse is $P^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$. Multiplying out gives $A^n = \frac{1}{3}\begin{pmatrix} 2^n + 2 \cdot 5^n & 5^n - 2^n \\ 2 \cdot 5^n - 2 \cdot 2^n & 2 \cdot 2^n + 5^n \end{pmatrix}$. Setting $n = 1$ recovers A, a useful check. For large n the 5^n terms dominate.

Worked example 7: Solving a system of differential equations

Solve $x' = x + 2y$, $y' = 3x + 2y$ with $x(0) = 1$, $y(0) = 0$.

The matrix is $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ with trace 3 and determinant -4, so $\lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0$. For $\lambda = 4$ the eigenvector satisfies $-3x + 2y = 0$, so $(2, 3)$; for $\lambda = -1$ it satisfies $2x + 2y = 0$, so $(1, -1)$. The general solution is $\mathbf{x} = c_1 e^{4t}(2, 3) + c_2 e^{-t}(1, -1)$. The initial condition gives $2c_1 + c_2 = 1$ and $3c_1 - c_2 = 0$, so $c_1 = \frac{1}{5}$ and $c_2 = \frac{3}{5}$. The solution grows like e^{4t}, because the dominant eigenvalue is positive.

Worked example 8: Eigenvalues of a projection

Find the eigenvalues of the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ that projects the plane onto the x-axis.

The matrix is diagonal, so the eigenvalues are 1 and 0. The eigenvector for 1 is $(1, 0)$: points on the x-axis are left unchanged. The eigenvector for 0 is $(0, 1)$: points on the y-axis are squashed to the origin. A zero eigenvalue always means that the matrix is singular, with determinant zero and a non-trivial kernel.

Worked example 9: Symmetric matrices have real eigenvalues

Show that the eigenvalues of $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ are always real.

The characteristic equation is $\lambda^2 - (a + d)\lambda + (ad - b^2) = 0$. Its discriminant is $(a + d)^2 - 4(ad - b^2) = (a - d)^2 + 4b^2$, which is a sum of squares and therefore at least zero. So both roots are real. Moreover, the eigenvectors for different eigenvalues are perpendicular, which is why symmetric matrices can be diagonalised with a rotation.

Key fact: Markov chains and steady states

A Markov chain describes a system that moves between states with fixed probabilities. Its transition matrix has non-negative entries and each column adds up to one, which guarantees that 1 is always an eigenvalue. The eigenvector for the eigenvalue 1, scaled so that its entries sum to one, is the steady state: the long-term proportion of time spent in each state. All the other eigenvalues have modulus less than or equal to one, and for a regular chain they are strictly smaller than one, so their contributions fade away. The size of the second largest eigenvalue tells you how quickly the system forgets its starting point: a value close to one means slow convergence, and a value close to zero means fast convergence.

Key fact: stability of linear systems

For the differential equation $\dot{\mathbf{x}} = A\mathbf{x}$ each eigenvalue $\lambda = a + bi$ contributes a term that behaves like e^{at} multiplied by an oscillation of frequency b. The system is stable, meaning that every solution tends to zero, exactly when all eigenvalues have negative

real part. If at least one eigenvalue has positive real part, most solutions blow up. For a two-dimensional system the trace and determinant are enough to decide: stability needs a negative trace and a positive determinant. A negative determinant gives a saddle, where solutions approach along one direction and escape along another.

Note: complex eigenvalues of real matrices

A real matrix has a characteristic polynomial with real coefficients, so complex eigenvalues come in conjugate pairs $a \pm bi$. They describe a rotation by the angle $\arg(a+bi)$ combined with a scaling by $\sqrt{a^2+b^2}$. In a linear dynamical system the real part a controls growth or decay, and the imaginary part b controls the speed of oscillation. This is why a damped spring has eigenvalues of the form $-\gamma \pm i\omega$.

Watch out: frequent errors

- The zero vector is never an eigenvector, even though $A\mathbf{0} = \lambda\mathbf{0}$ for every λ.
- Eigenvalues of $A+B$ are not the sums of eigenvalues of A and B (unless the matrices share eigenvectors); the same applies to products.
- A repeated eigenvalue does not guarantee two independent eigenvectors.
- When diagonalising, the order of columns in P must match the order of eigenvalues in D.

Did you know?

The word "eigen" is German for "own" or "characteristic": an eigenvector is a characteristic direction of a matrix. The idea first appeared in the study of vibrating strings and rotating bodies in the eighteenth century. Today Google ranks web pages with an eigenvector of a matrix with billions of rows, and engineers check the stability of bridges by studying eigenvalues.

Otto's tip

Check with trace and determinant: the eigenvalues must add to the trace and multiply to the determinant. And never take $\mathbf{v} = \mathbf{0}$ as an eigenvector.