

Eigenvalues and Orthogonality: math lesson, College – download the PDF

Some matrices have special directions that they only stretch, never turn. Finding those directions explains how a whole system evolves over time, and combining them with the idea of perpendicular vectors gives you some of the most useful tools in applied mathematics: projections, orthogonal bases and best-fit lines. In this chapter you will learn to compute eigenvalues, diagonalize a matrix, use the dot product to test for orthogonality, project vectors, build orthogonal bases with the Gram-Schmidt process, and fit data with least squares.

1. Eigenvalues and eigenvectors

When a square matrix A multiplies a vector $\mathbf{x}$, the result is usually a vector pointing in a new direction. For a few special vectors, the output points along the same line as the input.

D Definition: eigenvalue and eigenvector

Let A be an $n \times n$ matrix. A nonzero vector $\mathbf{v}$ is an **eigenvector** of A if

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

for some number λ . The number λ is the corresponding **eigenvalue**.

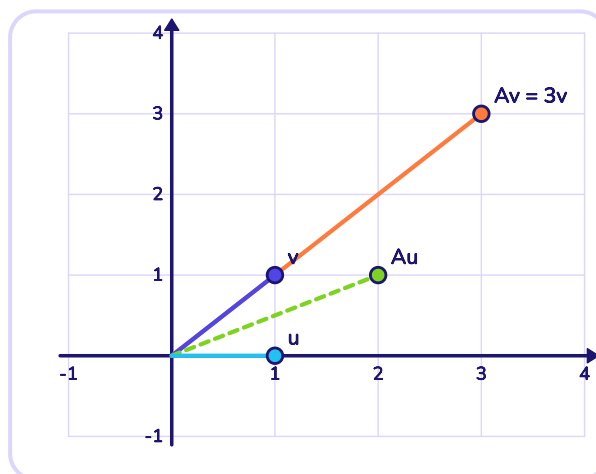
The vector $\mathbf{v}$ must be nonzero (the zero vector would satisfy the equation for every λ), but the eigenvalue λ is allowed to be 0 or negative. If $\lambda < 0$, the matrix flips the vector to the opposite direction while stretching or shrinking it.

E Example 1: checking an eigenvector

$$\text{Let } A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

$$A\mathbf{v} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\mathbf{v}, \text{ so } \mathbf{v} \text{ is an eigenvector with eigenvalue } 3.$$

$$A\mathbf{u} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \text{ which is not a multiple of } \mathbf{u}, \text{ so } \mathbf{u} \text{ is not an eigenvector.}$$



Zyro's tip

On my home planet we say an eigenvector is a “loyal direction”: the matrix may stretch it, shrink it or flip it, but it never bends it. Look at the figure: $\mathbf{v}$ and $A\mathbf{v}$ sit on the same line, while $\mathbf{u}$ and $A\mathbf{u}$ do not.

2. The characteristic polynomial

To find eigenvalues we rewrite $A\mathbf{v} = \lambda\mathbf{v}$ as $(A - \lambda I)\mathbf{v} = \mathbf{0}$. A nonzero solution $\mathbf{v}$ exists only when $A - \lambda I$ is not invertible, that is, when its determinant is zero.

P Property: characteristic equation

The number λ is an eigenvalue of A exactly when

$$\det(A - \lambda I) = 0.$$

The left side is a polynomial of degree n in λ , called the **characteristic polynomial**. For a 2×2 matrix it equals $\lambda^2 - (\text{trace})\lambda + \det A$.

Two checks save a lot of mistakes. The eigenvalues add up to the trace of A (the sum of the diagonal entries), and they multiply to give $\det A$. For a triangular matrix, the eigenvalues are simply the diagonal entries.

M Method: finding eigenvalues and eigenvectors

1. Write $\det(A - \lambda I) = 0$ and expand it.
2. Solve the polynomial equation to get the eigenvalues.
3. For each eigenvalue λ , solve $(A - \lambda I)\mathbf{v} = \mathbf{0}$ to find its eigenvectors.
4. Check by computing $A\mathbf{v}$ and comparing with $\lambda\mathbf{v}$.

E**Example 2: a full computation**

Let $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$. Its trace is **5** and its determinant is $4 + 2 = 6$, so

$$\det(A - \lambda I) = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

The eigenvalues are **2** and **3**.

For $\lambda = 2$: $A - 2I = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}$ gives $x = y$, so $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

For $\lambda = 3$: $A - 3I = \begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix}$ gives $x = 2y$, so $\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Check: $A\mathbf{v}_2 = \begin{pmatrix} 8 - 2 \\ 2 + 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} = 3\mathbf{v}_2$.

**Watch out**

A matrix can have no real eigenvalues. The rotation matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial $\lambda^2 + 1$, which has no real root: no real direction survives a quarter turn.

3. Diagonalization

If an $n \times n$ matrix has n linearly independent eigenvectors, we put them as the columns of a matrix P and the eigenvalues on the diagonal of a matrix D . Then $AP = PD$.

P**Property: diagonalization**

If P has eigenvectors of A as columns and D is the diagonal matrix of the matching eigenvalues, then

$$A = PDP^{-1} \quad \text{and} \quad A^k = PD^kP^{-1}.$$

A matrix with n distinct eigenvalues is always diagonalizable.

The power formula is the real payoff: raising a diagonal matrix to a power only means raising each diagonal entry to that power.

E**Example 3: diagonalizing the matrix of Example 2**

With $P = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, we get $\det P = -1$ and $P^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$. Multiplying out confirms

$$PDP^{-1} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} = A.$$

Therefore A^{10} has eigenvalues $2^{10} = 1024$ and $3^{10} = 59,049$, with the same eigenvectors as A .

⚠**Watch out**

Not every matrix is diagonalizable. For $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ the only eigenvalue is **3**, but every eigenvector is a multiple of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$: there are not enough independent eigenvectors.

4. Dot product and orthogonality

D Definition: dot product and length

For vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{w} = (w_1, \dots, w_n)$,

$$\mathbf{u} \cdot \mathbf{w} = u_1 w_1 + u_2 w_2 + \dots + u_n w_n, \quad \|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

Two vectors are **orthogonal** when $\mathbf{u} \cdot \mathbf{w} = 0$.

The dot product also gives the angle θ between two nonzero vectors: $\cos \theta = \frac{\mathbf{u} \cdot \mathbf{w}}{\|\mathbf{u}\| \|\mathbf{w}\|}$.

Orthogonal means $\theta = 90^\circ$. A set of vectors is **orthogonal** if every pair is orthogonal, and **orthonormal** if, in addition, every vector has length 1. To normalize a nonzero vector, divide it by its length.

Orthogonality is tied to eigenvalues by one important fact: if a matrix is symmetric (equal to its transpose), eigenvectors that belong to different eigenvalues are always orthogonal. That is why symmetric matrices can be diagonalized with a perfectly “square” set of axes.

E Example 4: dot products

For $\mathbf{u} = (2, -1, 2)$ and $\mathbf{w} = (1, 4, 3)$: $\mathbf{u} \cdot \mathbf{w} = 2 - 4 + 6 = 4$, so they are not orthogonal. Also $\|\mathbf{u}\| = \sqrt{4 + 1 + 4} = 3$, so $\mathbf{u}/3 = (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$ is a unit vector.

5. Orthogonal projections

To project a vector $\mathbf{b}$ onto the line through the origin spanned by a nonzero vector $\mathbf{a}$, we look for the point on that line closest to $\mathbf{b}$. It is the point where the leftover “error” is perpendicular to $\mathbf{a}$.

P

Property: projection onto a line

$$\text{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}.$$

The error $\mathbf{b} - \text{proj}_{\mathbf{a}} \mathbf{b}$ is orthogonal to $\mathbf{a}$, and its length is the distance from $\mathbf{b}$ to the line.

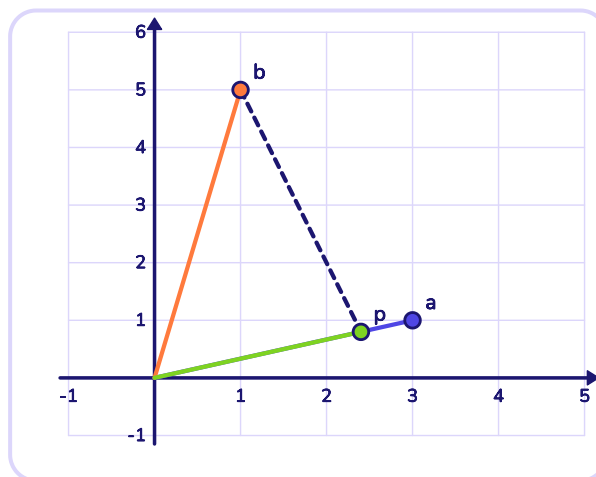
E

Example 5: projecting $\mathbf{b}$ onto $\mathbf{a}$

Let $\mathbf{a} = (3, 1)$ and $\mathbf{b} = (1, 5)$. Then $\mathbf{b} \cdot \mathbf{a} = 8$ and $\mathbf{a} \cdot \mathbf{a} = 10$, so

$$\text{proj}_{\mathbf{a}} \mathbf{b} = \frac{8}{10} (3, 1) = (2.4, 0.8).$$

The error is $(1, 5) - (2.4, 0.8) = (-1.4, 4.2)$, and $(-1.4)(3) + (4.2)(1) = 0$, as expected. The distance from $\mathbf{b}$ to the line is $\sqrt{1.96 + 17.64} = \sqrt{19.6} \approx 4.43$.



6. The Gram-Schmidt process

Orthogonal bases make computations easy, but a basis you are handed is rarely orthogonal. Gram-Schmidt repairs it one vector at a time: keep the first vector, then subtract from each new vector its projections onto the vectors already built.

M

Method: Gram-Schmidt

Start from independent vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$

1. $\mathbf{u}_1 = \mathbf{v}_1$.
2. $\mathbf{u}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1} \mathbf{v}_2$.
3. $\mathbf{u}_3 = \mathbf{v}_3 - \text{proj}_{\mathbf{u}_1} \mathbf{v}_3 - \text{proj}_{\mathbf{u}_2} \mathbf{v}_3$, and so on.
4. To get an orthonormal basis, divide each $\mathbf{u}_i$ by its length.

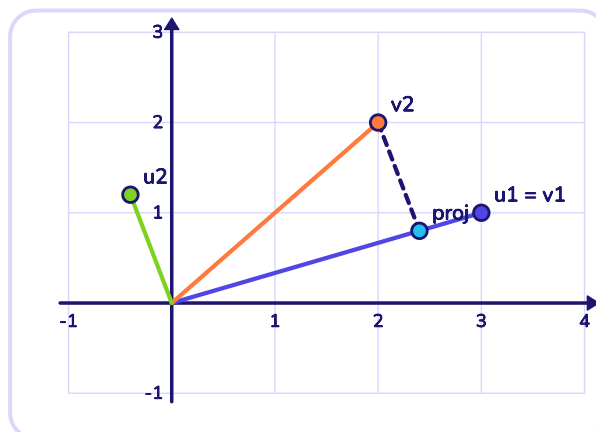
E

Example 6: two vectors in the plane

Take $\mathbf{v}_1 = (3, 1)$ and $\mathbf{v}_2 = (2, 2)$. Then $\mathbf{u}_1 = (3, 1)$ and $\mathbf{v}_2 \cdot \mathbf{u}_1 = 8$, $\mathbf{u}_1 \cdot \mathbf{u}_1 = 10$, so

$$\mathbf{u}_2 = (2, 2) - \frac{8}{10}(3, 1) = (2, 2) - (2.4, 0.8) = (-0.4, 1.2).$$

Check: $\mathbf{u}_1 \cdot \mathbf{u}_2 = -1.2 + 1.2 = 0$.



7. Least squares

Real data rarely fit a line exactly. A system $A\mathbf{x} = \mathbf{b}$ with more equations than unknowns usually has no solution, so we look for the vector $\hat{\mathbf{x}}$ that makes $A\hat{\mathbf{x}}$ as close as possible to $\mathbf{b}$. Closest means the error $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to every column of A , which is a projection of $\mathbf{b}$ onto the column space of A .

P Property: normal equations

The least squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b}.$$

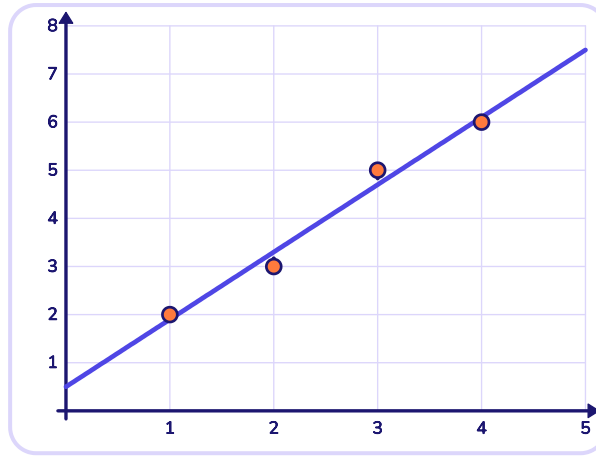
It minimizes the sum of the squared errors $\|\mathbf{b} - A\hat{\mathbf{x}}\|^2$.

E Example 7: a best-fit line

Fit $y = a + bx$ to the points $(1, 2), (2, 3), (3, 5), (4, 6)$. The columns of A are $(1, 1, 1, 1)$ and $(1, 2, 3, 4)$, and $\mathbf{b} = (2, 3, 5, 6)$. Then

$$A^T A = \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}, \quad A^T \mathbf{b} = \begin{pmatrix} 16 \\ 47 \end{pmatrix}.$$

Solving gives $a = 0.5$ and $b = 1.4$, so the line is $y = 0.5 + 1.4x$. The residuals are $0.1, -0.3, 0.3, -0.1$, and the sum of squared errors is 0.2 .



The same idea works for curves: a parabola $y = a + bx + cx^2$ only needs a third column $(x_1^2, \dots, x_n^2)$ in A . The method stays the same.

Key takeaways

- An eigenvector satisfies $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$; λ is its eigenvalue.
- Eigenvalues are the roots of $\det(A - \lambda I) = 0$; their sum is the trace and their product is the determinant.
- $A = PDP^{-1}$ when there are enough independent eigenvectors, and then $A^k = PD^kP^{-1}$.
- $\mathbf{u} \cdot \mathbf{w} = 0$ means orthogonal; eigenvectors of a symmetric matrix for different eigenvalues are orthogonal.
- $\text{proj}_{\mathbf{a}} \mathbf{b} = \frac{\mathbf{b} \cdot \mathbf{a}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a}$, and the error is perpendicular to $\mathbf{a}$.
- Gram-Schmidt subtracts projections to turn independent vectors into an orthogonal set.
- Least squares solves $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$ to find the best fit.



Test yourself: quick challenge for College

Change level

Level : College

6 random questions, instant feedback.

 **Start the challenge**

 **Speed drill for College: how many in 60 seconds?**

 **Keep exploring with Zyro**

 **Math practice**

Eigenvalues and Orthogonality: math practice, College

 **Math tests**

Eigenvalues and Orthogonality: math test, College

 Math quizzes

Eigenvalues and Orthogonality: math quiz, College

 Math practice

Logic, Proofs and Discrete Math: math practice, College

 Math practice

Limits and Continuity: math practice, College

 Math tests

Logic, Proofs and Discrete Math: math test, College

Zyro takes you to our partner sites

Blast off to other planets of knowledge and their apps.

Lift off