

Eigenvalues and Eigenvectors

[Read this note online](#)

Eigenvalues and eigenvectors describe the special directions a matrix transformation preserves and the factor it scales them by. They're among the most consequential concepts in linear algebra. Once you understand them, you have a structural vocabulary for principal component analysis, dynamical systems, quantum mechanics, vibration analysis, recommender systems, and Google's original PageRank algorithm.

The defining equation is short, $A\mathbf{v} = \lambda\mathbf{v}$, but the implications are enormous. Eigenvectors are the "axes" of a matrix transformation; eigenvalues are how much the transformation stretches along each axis. Diagonalizing a matrix means rewriting it in its eigenvector basis, where its action becomes pure scaling.

This study note covers the definition, the characteristic polynomial, how to compute eigenvalues and eigenvectors by hand, the geometric interpretation, diagonalization, the spectral theorem, and applications across science, engineering, and machine learning.

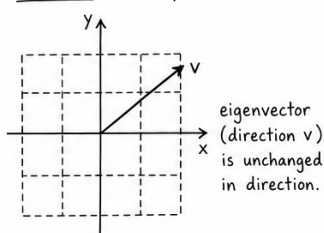
An eigenvector keeps its direction under a linear transformation, possibly reversing it. The eigenvalue gives its scale factor.

EIGENVALUES AND EIGENVECTORS

IDEA: An eigenvector is a direction that is unchanged by a linear transformation (only stretched or flipped).

WHERE: A is an $n \times n$ matrix
 $v \neq 0$ is an eigenvector
 λ is the corresponding eigenvalue (scalar)

BEFORE: identity transformation



$$A v = \lambda v$$

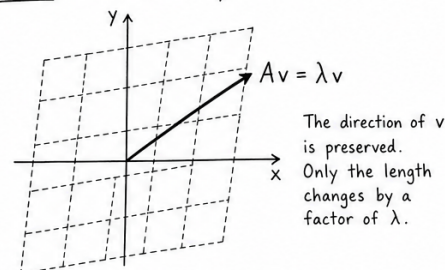
matrix (pointing to A), eigenvalue (scalar) (pointing to λ), eigenvector (pointing to v), eigenvector (same direction) (pointing to v)

CHARACTERISTIC EQUATION:

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

Solve for λ to find eigenvalues.

AFTER: transformation by A



EXAMPLE (2×2):

Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

Step 1: Characteristic equation

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0 \\ (2-\lambda)(2-\lambda) - 1 &= 0 \\ \lambda^2 - 4\lambda + 3 &= 0 \\ (\lambda - 1)(\lambda - 3) &= 0 \end{aligned}$$

Step 2: Eigenvalues

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 3 \end{aligned}$$

Step 3: Eigenvectors

For $\lambda_1 = 1$:
 $(A - I)v = 0 \Rightarrow v_1 \propto \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

For $\lambda_2 = 3$:
 $(A - 3I)v = 0 \Rightarrow v_2 \propto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

TAKEAWAY: Eigenvectors are special directions that a linear transformation does not rotate — only stretches (or flips). Eigenvalues tell us by what factor.

Eigenvalues and eigenvectors equation transformation diagram modern textbook illustration

The Definition

For a square matrix A , a nonzero vector $\mathbf{v}$ is an *eigenvector* if:

for some scalar λ , called the corresponding *eigenvalue*.

In words: an eigenvector is a vector whose direction the matrix preserves. The matrix may stretch or compress it, but it doesn't rotate it. The factor by which it's stretched is the eigenvalue.

Most **vectors** are not eigenvectors of a given matrix. They get rotated and stretched in arbitrary ways. Eigenvectors are the special exceptions, the directions that the transformation scales without rotating.

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

The Characteristic Polynomial

To find eigenvalues, rearrange $A\mathbf{v} = \lambda\mathbf{v}$ as $(A - \lambda I)\mathbf{v} = \mathbf{0}$. For a nonzero $\mathbf{v}$ to satisfy this, the matrix $A - \lambda I$ must be singular:

This is the *characteristic equation*, and the polynomial $p(\lambda) = \det(A - \lambda I)$ is the *characteristic polynomial*. Its degree equals the matrix size; an $n \times n$ matrix has a characteristic polynomial of degree n and therefore exactly n eigenvalues, counted with multiplicity (over the **complex numbers**).

Solving $p(\lambda) = 0$ finds the eigenvalues. Once eigenvalues are known, plug each back into $(A - \lambda I)\mathbf{v} = \mathbf{0}$ and solve for $\mathbf{v}$ to get the corresponding eigenvectors.

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

Worked Example: 2×2 Matrix

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

Step 1: characteristic polynomial.

Solving: $\lambda_1 = 5, \lambda_2 = 2$.

Step 2: eigenvector for $\lambda_1 = 5$.

$$(A - 5I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \mathbf{v} = \mathbf{0}. \text{ Both rows give } v_1 = v_2, \text{ so } \mathbf{v}_1 = (1, 1).$$

Step 3: eigenvector for $\lambda_2 = 2$.

$$(A - 2I)\mathbf{v} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0}. \text{ Both rows give } 2v_1 + v_2 = 0, \text{ so } \mathbf{v}_2 = (1, -2).$$

$$\det(A - \lambda I) = (4 - \lambda)(3 - \lambda) - 2 = \lambda^2 - 7\lambda + 10 = 0$$

Geometric Meaning

Imagine the matrix A as a transformation that takes 2D vectors to 2D vectors. Most input vectors come out pointing in different directions. But the eigenvectors come out pointing in the same direction (or directly opposite, if the eigenvalue is negative), only stretched.

For the example above, every vector along $(1, 1)$ gets scaled by 5 (and stays on that line). Every vector along $(1, -2)$ gets scaled by 2 (and stays on that line). Vectors not on these two lines get warped into combinations of the two stretches.

This decomposition is the heart of what eigenvalues and eigenvectors do: they identify the "stable directions" of a linear transformation and the scaling factor along each.

Diagonalization

If a square matrix A has n linearly independent eigenvectors, it can be written as:

where the columns of P are the eigenvectors and Λ is the diagonal matrix of eigenvalues. A is then called *diagonalizable*.

Diagonalization makes many matrix computations easier. Powers: $A^k = P\Lambda^k P^{-1}$, and Λ^k is just the eigenvalues raised to the power k . Matrix exponentials: $e^A = Pe^{\Lambda}P^{-1}$. Solving differential equations $\dot{\mathbf{x}} = A\mathbf{x}$ reduces to scalar equations along eigenvector directions.

Not every matrix is diagonalizable. Defective matrices (missing eigenvectors) require Jordan canonical form instead. In practice, many matrices are diagonalizable enough that this concern rarely matters in applied work.

$$A = P\Lambda P^{-1}$$

The Spectral Theorem

For real symmetric matrices, the *spectral theorem* guarantees:

So $A = Q\Lambda Q^T$ where Q is an orthogonal matrix (columns are orthonormal eigenvectors). The transpose replaces the inverse, simplifying many computations.

The spectral theorem is one of the cleanest results in linear algebra and is the backbone of principal component analysis, spectral graph theory, and quantum mechanics. Symmetric matrices appear constantly in applied math (covariance matrices, Hessians, Laplacians), so the spectral theorem applies to nearly every applied problem.

All eigenvalues are real. Eigenvectors corresponding to distinct eigenvalues are orthogonal. The matrix is diagonalizable, and there exists an orthonormal basis of eigenvectors.

Properties of Eigenvalues

These properties let you compute many derived quantities (determinants, traces, condition numbers) from eigenvalues without re-doing the full matrix work.

Trace: the sum of eigenvalues equals the trace of the matrix (sum of diagonal entries). **Determinant:** the product of eigenvalues equals the determinant. **Inverse:** if A is invertible, the eigenvalues of A^{-1} are $1/\lambda_i$, with the same eigenvectors. **Powers:** the eigenvalues of A^k are λ_i^k , with the same eigenvectors. **Transpose:** A^T has the same eigenvalues as A (but generally different eigenvectors). **Polynomial:** if $p(\lambda)$ is any polynomial, the eigenvalues of $p(A)$ are $p(\lambda_i)$.

Complex Eigenvalues

Real matrices can have complex eigenvalues. They come in conjugate pairs: if $a + bi$ is an eigenvalue, so is $a - bi$. Geometrically, complex eigenvalues correspond to rotational components of the transformation.

A pure rotation matrix $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ has eigenvalues $e^{\pm i\theta}$. The angle θ of rotation appears directly in the complex eigenvalues. Magnitude 1 means no scaling, only rotation.

Complex eigenvalues show up in oscillatory dynamical systems, AC circuit analysis, and the analysis of Markov chains with periodic structure. Working with them requires comfort with complex numbers but produces no new fundamental difficulties.

PCA: Eigenvalues for Data Science

Principal Component Analysis (PCA) decomposes the covariance matrix of a dataset into eigenvectors and eigenvalues. The eigenvectors give the principal directions of variance; the eigenvalues give the variance along each direction.

Project the data onto the first few eigenvectors (those with the largest eigenvalues) to get a low-dimensional representation that preserves most of the variance. This is the standard dimensionality reduction technique in **data science** and **machine learning**.

PCA is used for visualization (project to 2D for plotting), feature reduction (use top k components instead of all features), noise reduction (small eigenvalues are usually noise), and exploratory analysis. The fact that all of this works elegantly traces back to the spectral theorem applied to the symmetric covariance matrix.

PageRank: Eigenvalues for Search

Google's original PageRank algorithm computes the principal eigenvector of the web's link matrix. Web pages are nodes; links are weighted edges. The PageRank score of each page is the corresponding entry of the eigenvector associated with the largest eigenvalue.

The interpretation: PageRank is the long-run probability that a random surfer is at a given page. The link matrix is a Markov chain transition matrix; its dominant eigenvector is the stationary distribution. Eigenvalue analysis gave Google a principled, scalable way to rank web pages and helped launch the modern search era.

Vibration Analysis and Dynamical Systems

Engineers analyze structural vibrations by computing eigenvalues and eigenvectors of mass and stiffness matrices. Eigenvalues give the natural frequencies; eigenvectors give the corresponding mode shapes.

Bridges, buildings, aircraft, and engines all depend on knowing where these resonant frequencies sit relative to expected loads. The Tacoma Narrows bridge collapse in 1940 is the textbook cautionary tale of eigenfrequency mismatch with environmental forcing.

The same math underlies stability analysis of control systems and ordinary differential equation systems. Negative real eigenvalues mean stable equilibria; positive real eigenvalues mean instability; complex eigenvalues with negative real parts mean damped oscillations.

Quantum Mechanics

In quantum mechanics, observables are represented by Hermitian operators (the complex-matrix analogue of symmetric matrices). The possible measurement outcomes are the eigenvalues of the operator; the corresponding eigenvectors are the eigenstates.

Energy levels of an atom are eigenvalues of the Hamiltonian operator. Spin states are eigenvectors of spin operators. The entire formalism of quantum mechanics is built on eigenvalue equations like $H\psi = E\psi$, the Schrödinger equation in time-independent form.

Numerical Computation

Computing eigenvalues by hand is fine for 2×2 and 3×3 matrices. Beyond that, characteristic polynomials become unwieldy and high-degree polynomial roots are numerically unstable.

Modern algorithms compute eigenvalues iteratively. The QR algorithm (Francis, 1961) is the standard for general dense matrices. The Lanczos and Arnoldi methods handle large sparse matrices, computing only a few eigenvalues at a time. Power iteration is the simplest method and works well when only the largest eigenvalue is needed (PageRank uses a variant of this).

NumPy, MATLAB, and LAPACK all implement these algorithms efficiently. For most practical work, you call a library function and trust the implementation; understanding the algorithms matters when you hit numerical issues or work with structured matrices that admit faster methods.

Common Mistakes with Eigenvalues

Forgetting that eigenvectors are nonzero by definition. The zero vector trivially satisfies $A\mathbf{0} = \lambda\mathbf{0}$ for any λ , so we exclude it. **Treating eigenvectors as unique.** Any nonzero scalar multiple of an eigenvector is also an eigenvector. The eigenvector direction is unique; the magnitude is conventional (often normalized to unit length). **Assuming all matrices are diagonalizable.** Defective matrices have fewer eigenvectors than their dimension and require Jordan form instead. **Computing eigenvalues by characteristic polynomial for large matrices.** The polynomial roots are numerically unstable; use iterative algorithms (QR, Lanczos) instead. **Forgetting that complex eigenvalues come in conjugate pairs for real matrices.** If $a + bi$ is an eigenvalue, so is $a - bi$. **Confusing eigenvalues of A with eigenvalues of $A + cI$.** The latter are $\lambda + c$, shifted by the same constant.

Why Eigenvalues Are So Useful

Eigenvalues and eigenvectors decompose a transformation into scaling factors along independent directions. Almost every "structural" question about a matrix reduces to its eigenvalue structure: stability, condition number, principal directions of variation, natural frequencies, asymptotic behavior, dimensionality reduction.

The unifying insight is this: a generic linear transformation looks complicated in arbitrary coordinates, but in the eigenvector basis it's just diagonal, pure scaling. Diagonal matrices are trivially understood. Eigenvalue analysis is the systematic search for the coordinate system in which any transformation becomes simple.

A Brief History

Eigenvalues appeared first in mid-19th-century work on conic sections and quadratic forms. Augustin-Louis Cauchy proved the spectral theorem for symmetric matrices in 1829. David Hilbert generalized eigenvalue concepts to infinite-dimensional spaces in the early 20th century, paving the way for quantum mechanics. The terms *eigenvalue* and *eigenvector* come from the German *Eigenwert* and *Eigenvektor* (literally "own-value" and "own-vector"), introduced by David Hilbert.

The 20th century made eigenvalue computation practical. The QR algorithm (1961) enabled efficient computation for general dense matrices. Modern numerical linear algebra continues to extend the toolkit to massive sparse matrices and structured problems.

Singular Value Decomposition (SVD)

Singular Value Decomposition generalizes eigenvalue decomposition to non-square matrices. Any $m \times n$ matrix A factors as $A = U\Sigma V^T$, where U and V are orthogonal matrices and Σ is a diagonal matrix of nonnegative singular values.

SVD always exists. Its singular values reveal the matrix's effective rank, condition number, and intrinsic structure. Truncating SVD to keep only the top k singular values gives the best rank- k approximation under several natural error norms, the basis for low-rank methods, image compression, and PCA.

SVD is one of the most useful results in applied linear algebra. It underpins recommender systems (Netflix challenge), latent semantic analysis in NLP, and the majority of dimensionality reduction techniques in machine learning.

Spectral Graph Theory and Network Analysis

The eigenvalues of a graph's adjacency matrix or Laplacian reveal structural properties of the network. The second-smallest eigenvalue of the Laplacian (Fiedler eigenvalue) measures graph connectivity. The corresponding eigenvector partitions the graph into well-connected clusters, the foundation of spectral clustering.

Spectral methods for graph analysis power community detection in social networks, recommender system clustering, image segmentation, and many other modern data science tasks. They translate combinatorial graph problems into linear algebra problems that scale to enormous datasets via sparse eigenvalue solvers.

Power Iteration: the Simplest Eigenvalue Algorithm

Power iteration finds the dominant eigenvector of a matrix. Start with a random vector $\mathbf{x}_0$. Repeatedly compute $\mathbf{x}_{k+1} = A\mathbf{x}_k / \|A\mathbf{x}_k\|$. The sequence converges to the eigenvector with the largest absolute eigenvalue.

The convergence rate depends on the gap between the largest and second-largest eigenvalues. Bigger gap, faster convergence. Power iteration is simple, scales to enormous sparse matrices, and is the algorithm behind PageRank's web-scale eigenvalue computation. More sophisticated methods (inverse iteration, Lanczos, Arnoldi) extend the same basic idea to compute multiple eigenvalues efficiently.

The related [OpenStax mathematics textbooks](#) provide additional explanations and exercises for the mathematical definitions used here.

Finding Preserved Directions

For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, the vectors $(1,1)$ and $(1,-1)$ are eigenvectors with eigenvalues 3 and 1. Multiplication confirms $A(1, 1) = (3, 3)$ and $A(1, -1) = (1, -1)$.

Any nonzero multiple of either vector is also an eigenvector for the same eigenvalue. The zero vector is excluded because it would satisfy the equation for every scalar without identifying a direction. A matrix can fail to have enough independent eigenvectors for diagonalization, even when its characteristic polynomial has roots.

The points to check in this example are:

- An eigenvector must be nonzero.
- Repeated eigenvalues do not automatically provide enough independent eigenvectors.
- For a real symmetric matrix, an orthonormal eigenvector basis can be chosen.

Check Your Understanding

Try the questions before reading the answers.

Question 1. What is the determinant of the example matrix?

Answer. 3, agreeing with the product of its eigenvalues.

Question 2. What is its trace?

Answer. 4, agreeing with their sum.

FAQs

What does it mean for a vector to be an eigenvector?

An eigenvector of a matrix A is a nonzero vector whose direction the matrix preserves: $A v = \lambda v$ for some scalar λ . The matrix may stretch or compress the eigenvector but doesn't rotate it. Most vectors are not eigenvectors; eigenvectors are the special directions a transformation preserves.

How do I find the eigenvalues of a matrix?

Solve the characteristic equation $\det(A - \lambda I) = 0$. The solutions are the eigenvalues. For 2×2 and 3×3 matrices this is straightforward by hand; for larger matrices use numerical methods (QR algorithm, power iteration). Software libraries like NumPy and MATLAB compute eigenvalues in one function call.

How do I find eigenvectors once I have the eigenvalues?

For each eigenvalue λ , solve $(A - \lambda I) \mathbf{v} = 0$ to find the corresponding eigenvectors. The solutions form a subspace called the eigenspace for that eigenvalue. Pick any nonzero vector in that subspace; it's an eigenvector.

What is diagonalization?

Writing a matrix as $A = P \Lambda P^{-1}$ where P contains eigenvectors as columns and Λ is the diagonal matrix of eigenvalues. Diagonalization makes powers, exponentials, and many other matrix operations trivially fast. Not every matrix is diagonalizable; those that aren't are called defective.

What does the spectral theorem say?

For real symmetric matrices, all eigenvalues are real, eigenvectors for distinct eigenvalues are orthogonal, and there's always an orthonormal basis of eigenvectors. Equivalently, $A = Q \Lambda Q^T$ where Q is orthogonal. This makes symmetric matrices the cleanest case in linear algebra.

Can a matrix have complex eigenvalues?

Yes, real matrices can have complex eigenvalues, which come in conjugate pairs. They correspond to rotational components of the transformation. Pure rotation matrices have eigenvalues $e^{\pm i\theta}$. Complex eigenvalues appear in oscillatory systems, AC circuits, and dynamical systems with periodic behavior.

How are eigenvalues used in PCA?

PCA finds the eigenvectors of the data covariance matrix. The eigenvectors point in the directions of greatest variance; the corresponding eigenvalues quantify the variance along each direction. Projecting data onto the top eigenvectors gives the lowest-rank approximation that preserves the most variance.

What's the geometric meaning of an eigenvalue?

It's the scaling factor along the eigenvector direction. Eigenvalue 2 means the matrix stretches that direction by a factor of 2. Eigenvalue -1 means it reflects across the origin (same line, opposite direction). Eigenvalue 0 means it collapses that direction to zero (the matrix is singular).

How do eigenvalues relate to the determinant and trace?

The product of all eigenvalues equals the determinant. The sum equals the trace (sum of diagonal entries). These two relationships let you compute determinant and trace from eigenvalues without doing the full matrix work, and they sometimes give shortcut answers for analysis problems.

Why is the matrix called defective if it has fewer eigenvectors than its dimension?

Because it can't be diagonalized in the standard way. The matrix lacks enough eigenvectors to form a basis. Defective matrices need Jordan canonical form instead, a slightly more complex decomposition that captures the algebraic structure even when full diagonalization fails.

How is PageRank related to eigenvalues?

PageRank computes the principal eigenvector (associated with eigenvalue 1) of the link transition matrix of the web. The eigenvector entries are the steady-state probabilities a random surfer visits each page, which becomes the page's rank. This is the power iteration algorithm applied at web scale.

What's the difference between eigenvalues and singular values?

Eigenvalues come from square matrices via $Av = \lambda v$. Singular values come from the SVD $A = U\Sigma V^T$ and apply to any rectangular matrix; they're always nonnegative. For a symmetric positive semidefinite matrix, the eigenvalues equal the singular values. In general, singular values are eigenvalues of the related matrix $A^T A$, square-rooted.

Are eigenvectors unique?

The direction of an eigenvector is unique (up to sign), but the magnitude isn't. Any nonzero scalar multiple of an eigenvector is also an eigenvector with the same eigenvalue. Conventionally, eigenvectors are normalized to unit length. For repeated eigenvalues, the eigenspace can be multidimensional, and any orthonormal basis of that space works.

What is the characteristic polynomial?

The polynomial $p(\lambda) = \det(A - \lambda I)$. Its roots are the eigenvalues. The degree equals the matrix dimension. Computing eigenvalues by finding polynomial roots is the standard approach for small matrices but is numerically unstable for larger ones; iterative algorithms work better in practice.