✦ Member-only story

# Eigenvectors and Eigenvalues for Data Science

Linear Algebra for Data Science Series (7): Eigenvectors and Eigenvalues

Sunghyun Ahn    Follow    9 min read · Dec 6, 2024

240    💬 3
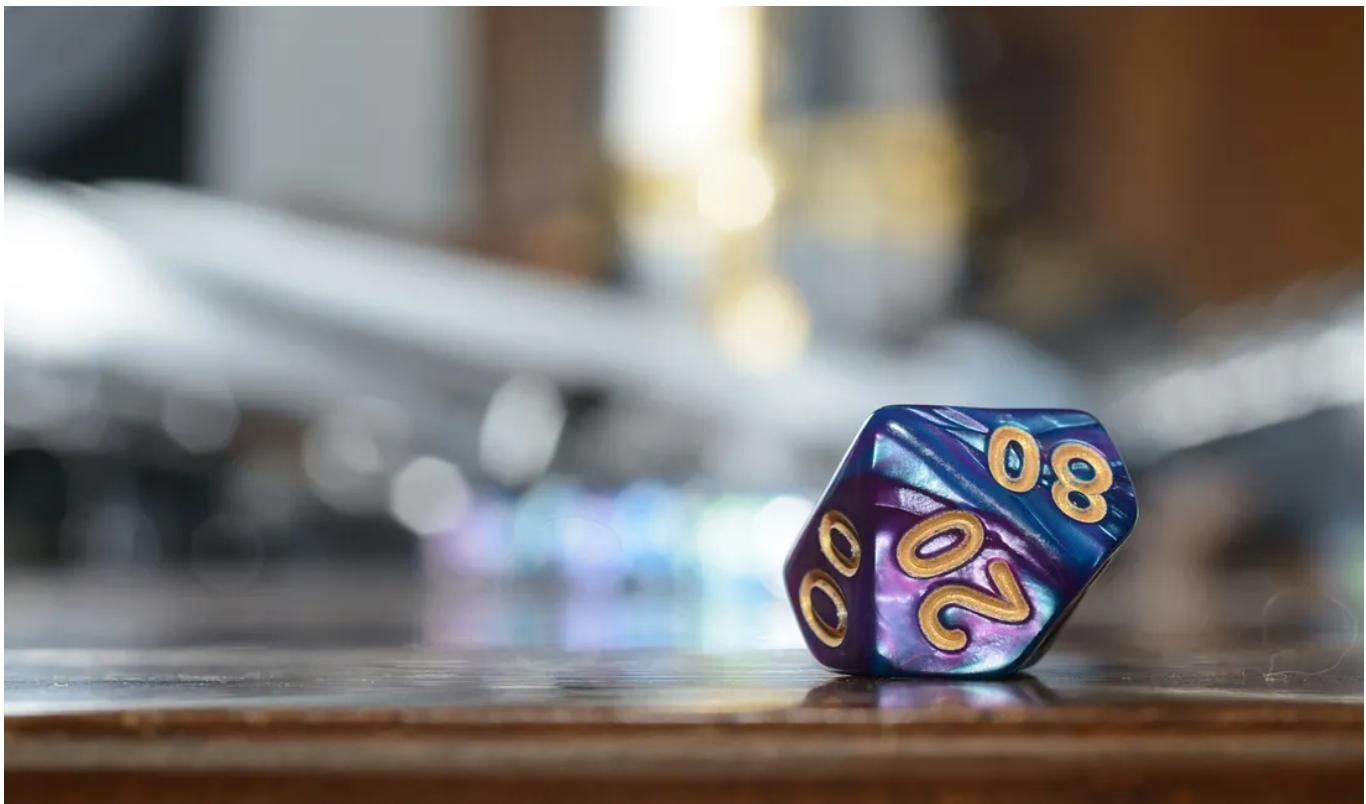


Photo by Carlos Felipe Ramírez Mesa on Unsplash

If you are not a paid member on Medium, I make my stories available for free: Friends link

If you've taken any college-level courses in engineering, advanced mathematics, and/or biostatistics, chances are you've come across the terms *eigenvectors* and *eigenvalues*.

For me, the first encounter was back in high school during AP Calculus BC. I'll admit, they were just fancy words I didn't quite understand until much later when I got to experience how much real-life implications they had talking to a connection I met at an AI convention for data science.

To save others from the same confusion I had, I will try to break down what eigenvectors and eigenvalues are in a clear and approachable way.

But before we begin, there's a fundamental rule about eigenvectors and eigenvalues you need to know:

**Rule for Eigenvectors and Eigenvalues**

Eigenvectors and eigenvalues are defined only for **_Square Matrices:_** $A \in \mathbb{R}^{n \times n}$

Rule for Eigenvectors and Eigenvalues

(We'll explore why this rule is true as we go along, so keep it in mind!)

## Table of Contents

## 1. Introduction to Eigenvectors and Eigenvalues

Let's say we have a square _matrix A_ and _some vector x._
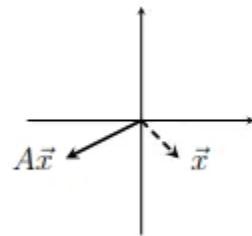
$$A = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

When we multiply *matrix A* by *x,* what happens? Well, matrix-vector multiplication always produces another vector. In most cases, this new vector has a different magnitude and direction compared to the original *vector x.* For example, suppose:

$$A\vec{x} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

We get Ax = [-2, -1]

On a 2D plane (the xy-plane), this would appear as a transformation of the original vector x.



## 1.1. Special Cases

Sometimes, though, there are special cases — special cases where certain *special vectors "x"* behave uniquely.

These special vectors only change their magnitude (by stretching, shrinking, or reflecting by a factor $\lambda$) when multiplied by *matrix A.* The critical feature of these vectors is that their **direction** remains unchanged after the multiplication.

**Surprisingly, these vectors are the stars of this discussion!**

- **Eigenvectors (x):** Special vectors that, when multiplied by A, only change in magnitude, not direction.

- **Eigenvalues ($\lambda$):** The factors by which the eigenvectors are stretched, shrunk, or reflected during multiplication.

An eigenvalue $|\lambda| > 1$ will stretch the vector, $|\lambda| < 1$ will shrink it, and $|\lambda| < 0$ will reflect it.

Eigenvectors are special vectors that, when multiplied by A, only change in magnitude, not direction.

## 2. Eigenvector Condition

Now that we've defined what an eigenvector is, let's unpack what it really means when we say, "an eigenvector only changes in magnitude, not direction."

Eigenvectors are special vectors such that, when multiplied by a matrix A, the resulting vector is simply a **scaled version of the original vector.** This scaling factor is the *eigenvalue λ.* If you can't get a visual representation of what that might look like, just wait! I am about to introduce you to the equation that must be satisfied for any *vector x* to be an eigenvector.

$$A\vec{x} = \lambda\vec{x}$$

Eigenvector Equation

Remember that the *eigenvalue λ* is just a scaling factor, meaning it's a scalar value that acts upon the *vector x.* Looking at this equation, you can see that a vector x multiplied by the matrix A is just the same as **vector x** multiplied by some scalr value λ!

### 2.1. Example

Let's revisit the earlier example with matrix A and vector x = [1,−1]. If we do a matrix-vector multiplication, we can get:

$$A\vec{x} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

We get Ax = [-2, -1]

We know that for x to be an eigenvector, the result of Ax should be equal to λx.

**So does this hold?**

To check, we'd need to find a scalar λ such that:

$$\lambda[1, -1] = [-2, -1]$$

Ax = λx

But is there any scalar λ that satisfies this equation? **Unfortunately, no.**
No matter what value of λwe try, we can't scale [1,−1] to produce [−2,−1].

## Thus we can say that our vector x is not an eigenvector!

### 2.2. Why Eigenvectors and Eigenvalues Are for Square Matrices

Some of you might already see why eigenvectors and eigenvalues require
square matrices — and the eigenvector equation explains it perfectly.

$$A\vec{x} = \lambda\vec{x}$$

The Eigenvector Equation

For this equation to work, matrix A must be square (n×n). **Why?**

**If matrix A is a square matrix:**

1. $A \in \mathbb{R}^{n \times n}$, meaning it has a dimension of $n \times n$
2. $\vec{x}$ is $\in \mathbb{R}^{n}$, meaning it has a dimension of $n \times 1$
3. Dot product of the two $(A \cdot \vec{x})$ has a dimension of $n \times 1$

Output vector results in n x 1

Thus, the output vector is the same size as the input vector. This is perfect
because we know that an *eigenvector "x"* that result from *Ax* is simply a
**scaled version** of x.

**If matrix A is a non-square matrix:**

1. $A \in \mathbb{R}^{m \times n}$, meaning it has a dimension of $m \times n$
2. $\vec{x}$ is $\in \mathbb{R}^{n}$, meaning it has a dimension of $n \times 1$
3. Dot product of the two ($A \cdot \vec{x}$) has a dimension of $m \times 1$

Output vector results in m x 1

When it is a non-square matrix, the output vector is a different size than the input vector ($Ax$). To handle non-square matrices, we use **Singular Value Decomposition (SVD)**, which generalizes these concepts. I'ved covered SVDs in detail in this next article:

**Singular Value Decomposition (SVD) in Data Science**

Linear Algebra for Data Science Interview Series (8): Singular Value Decomposition

medium.com

## 3. Solving Mathematically

Now that we've covered the conditions for eigenvectors and eigenvalues, how can we actually solve for them? While coding platforms like Python or R are typically used for efficiency, it's worth understanding the math behind the process.

We can rearrange the eigenvector equation into a solvable form:

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

New Eigenvector Equation

Here, I is the identity matrix of the same size as A. This equation tells us that *vector x* will be an eigenvector of A if it lies in the null space of $A - \lambda$.

### 3.1. Conditions for Eigenvectors and Eigenvalues

1. **Vector x must be non-zero:** A zero vector satisfies $(A - \lambda I)\, x = 0$ for any $\lambda$,

but this would make eigenvalues and eigenvectors meaningless (since any $\lambda$ would work). Hence, x must be non-zero.

2. Matrix $A - \lambda I$ must be singular (non-invertible): For the equation to have a non-trivial solution (x $\neq$ 0), A $- \lambda$ must not have full rank. This ensures that the determinant of A $- \lambda I$ is zero.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Full Rank: The system has exactly one unique solution.

## 3.2. Key Insight: Determinant = 0

Wait... we know if a the matrix is singular and cannot be inverted, the determinant of a matrix is zero. This gives us the characteristic equation:

$$det(A - \lambda I)\vec{x} = 0$$

We'll use this equation to find the **eigenvector** and **eigenvalue**

# 4. Calculation

Let's try to find the eigenvalues and eigenvalues of the below matrix A:

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

## 4.1. Eigenvalues

To find the eigenvalues, solve $det(A - \lambda I) = 0$.

In this equation, what is the identity matrix? Well the identity matrix is just a matrix that has 1's in it's diagonals and 0's everywhere else. We also know that the identity matrix will be the same shape as matrix A.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Identity matrix for A

Since we know this, we can now plug in and solve.

$$A - \lambda I = \begin{bmatrix} 3 - \lambda & 2 \\ 4 & 1 - \lambda \end{bmatrix}$$

We can see that **A — λI** is just **matrix A** subtracted by lambda across it's diagonal!

If we now take this and solve for the determinant, we'll get:

$$\det(A - \lambda I) = (3 - \lambda)(1 - \lambda) - (4)(2)$$
$$= \lambda^2 - 4\lambda - 5$$

**det(A — λI)**

Now we set $\lambda^2 - 4\lambda - 5$ equal to zero to solve for the eigenvalues:

$$\lambda_1 = 5, \quad \lambda_2 = -1$$

**Our Eigenvalues for Matrix A**

## 4.2. Eigenvectors

Next, find the eigenvectors corresponding to each eigenvalue by solving $(A - \lambda I)\, x = 0$. I'll show you the process for one eigenvalue, $\lambda_1 = 5$:

$$A - 5I = \begin{bmatrix} -2 & 2 \\ 4 & -4 \end{bmatrix}$$

$\lambda_1 = 5$

If we solve for $(A - 5I)\, x = 0$, it leads to:

$$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Eigenvector of A for eigenvalue = 5

If you made it this far into my article, I would appreciate it if you could leave a clap or a comment on my article!

## 5. Eigenvalue Decomposition

Now that we know how to calculate eigenvalues and eigenvectors, let's see what happens when we have *"n"* distinct eigenvalues, each corresponding to *"n"* independent eigenvectors. *This situation always occurs when there are no repeated eigenvalues.*

First, we should construct a matrix S, whose columns are the eigenvectors of A to denote the *"n"* independent eigenvectors.

$$S = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ | & | & & | \end{bmatrix}$$

Where $\mathbf{x}_n$ are the **eigenvectors of A**

Second, we should construct a diagonal matrix $\Lambda$ (it's capital lambda, not A), where the diagonal elements are all eigenvalues of matrix A.

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

If we compare this to the eigenvector equation:

$$A\vec{x} = \lambda\vec{x}$$

Eigenvector Equation

This equation can be re-written in Matrix form with our new matrix S and matrix $\Lambda$:

$$AS = S\Lambda$$

A: The matrix we're decomposing — S: A matrix of eigenvectors — Λ: A diagonal matrix of eigenvalues.

The above equation further reduces down to create our matrix A as...

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

Λ is the diagonal matrix of all eigenvalues

# This is referred to as the Eigenvalue Decomposition of A

## 5.1. Why does this matter?

Well okay, so we wrote this into a new form. At first glance, rewriting A into this form might not seem like a big deal. However, it unlocks a powerful way to compute higher powers of A, such as $A^k$:

Imagine that you have a matrix to the 3rd power, $A^3$. To calculate $A^3$, we would traditionally compute is as $A^3 = A \cdot A \cdot A$.

$$A^3 = A \cdot A \cdot A$$

However, if we begin needing to compute something more complex, like to the power of 100 or some huge number, then the computation will get super expensive.

Using the Eigenvalue Decomposition, the computation becomes much easier because Λ is the diagonal matrix. All we need to do is take the power of the diagonal elements (the eigenvalues). We only need to care about the diagonal matrix because...

> *The eigenvalues are what describe how much A stretches or compresses along each eigenvector direction.*

That's why we rewrite our matrix A!

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

## 5.2. What applications does this have?

Eigenvalue decomposition has applications across many fields because eigenvalues describe how matrix A stretches or compresses space along the directions defined by its eigenvectors. Here are a few key areas where this decomposition is often used:

**Physics and Engineering:**

- Solve differential equations efficiently.

- Analyze vibrations in mechanical systems.

- Study control systems using diagonalization techniques.

**Dynamical Systems:**

- Predict the long-term behavior of systems (e.g., population models, economic trends) by analyzing eigenvalues.

**Machine Learning:**

- Principal Component Analysis (PCA) uses eigenvalues and eigenvectors to reduce dimensionality and extract meaningful features from data.

That is why this Eigenvalue Decomposition is so useful.

## Conclusion

As I've shown, Eigenvectors and eigenvalues are not just random concepts from high school math, they create a fundamental basis for anything data science. From simplifying complex matrix operations to enabling dimensionality reduction in machine learning, their utility spans across various fields and applications.

With this foundational knowledge, you're now better equipped to learn about

more advanced topics like Singular Value Decomposition (SVD) and explore how these mathematical principles drive modern data science techniques.

**Singular Value Decomposition (SVD) in Data Science**

Linear Algebra for Data Science Interview Series (8): Singular Value Decomposition

medium.com

# Connect with me!

- [Linkedin](Linkedin), [Instagram](Instagram)

If you made it this far, I assume you are an aspiring data scientist, a teacher in the data science field, a professional looking to hone your craft, or just an avid learner in a different field! **I would love to have a chat with you on anything!**

Data Science        AI        Data        Machine Learning        Statistics

**Written by Sunghyun Ahn**                                    Follow

812 followers · 7 following

ex-Analyst @ Chemtopia | Graduate Research Assistant @ Seattle University |

# Responses (3)

Jagannath Rao

What are your thoughts?

Rmoraes
Feb 17

Very good introduction to the eigenworld. It is worth to remember that if all eigenvalue are different each other