# Eigenvectors and Eigenvalues

—Joaquín Gómez—

—Written: November 22, 2024—

—1st edition: February 19, 2025—

Suppose we have an equation $u(\mathbf{t})$ whose solution vector has to be in the same direction as $A$. The main idea of eigenvectors is to find a vector $x$ in the same direction as $A$.

We set the general equation as

$$Ax = \lambda x$$

This is the same as saying that the $Ax$ is the same as $x$ multiplied by some number $\lambda$.

If we rewrite the equation, and multiply by the identity $I$

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

This equation holds when $x$ is the zero vector. However, if we discard that solution, we know that $A - \lambda I$ must be singular (because there exists some combination of columns that give zero), meaning its determinant is zero. This means our matrix has a row of zeros, or its columns vectors are dependent from each other, etc.

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

If we want to take the determinant for a $2 \times 2$ matrix $A$, we get the following equation

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}$$

Now we can set the equation to zero and expand, also note that $\lambda_1 + \lambda_2 + ... + \lambda_n = T$, where $T$ is the trace of $A$ (in this case $T = a_{11} + a_{22}$)

$$a_{11}a_{22} - a_{11}\lambda - a_{22}\lambda + \lambda^2 - a_{12}a_{21} = 0$$

$$\lambda^2 - \lambda(a_{11} + a_{22}) + (a_{11}a_{22} - a_{12}a_{21}) = 0$$

Now, note that the second term is the product of $\lambda$ and the *trace* of the matrix, which is the sum of the main diagonal. The third term is the determinant of the matrix.

Now, what happens if we shift the matrix by any scalar $c$?

$$\begin{gathered}(A - cI)x = (\lambda - c)x \\ (A - cI - (\lambda - c)I)x = 0 \\ (A - cI - \lambda I + cI)x = 0 \\ (A - \lambda I)x = 0\end{gathered}$$

This means that shifting $A$ by some quantity will also shift the eigenvalue, but the eigenvector will remain the same (this can somehow hint that the eigenvector forms a basis for $A$, such that shifting $A$ will not modify its orientation, also this can be implied because $A$ is a linear transformation). For example, let's take a permutation matrix $P$

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

We want to find its eigenvalues and eigenvectors, in fact, there are two values since the equation we derived is quadratic. So, solving for the eigenvalue we get

$$\begin{gathered}\lambda^2 - \lambda(0 + 0) + (-1) = 0 \\ \lambda^2 - 1 = 0 \\ (\lambda - 1)(\lambda + 1) = 0\end{gathered}$$

this means the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$.

Now, for the eigenvectors we have to solve $Px = \lambda x$, for both values of $\lambda$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x = x \text{ and } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x = -x$$

this matrix performs permutation. If a vector $v = \langle a, b \rangle$ is multiplied we get that $Pv = \langle b, a \rangle$. The solution for the first eigenvector, a vector that has not to change direction  when multiplied by $P$ is just $x_1 = \langle 1, 1 \rangle$. The second one however, has to come negative. The eigenvector that satisfies is $x_2 = \langle -1, 1 \rangle$, if you reverse the order of the components you get $-x_2$.

# Upper triangular matrices

If $A$ is a square matrix, and is not singular, then is reducible to an upper triangular matrix $U$. We know that the determinant of an upper triangular matrix is

$$\det U = u_{11} u_{22} \cdot \cdot \cdot u_{nn}$$

which is the product of the components in the main diagonal. This is especially useful if we try to set the equation for eigenvectors for $U$

$$\det(U - \lambda I) = \begin{vmatrix} u_{11} - \lambda & - & - & ... & - \\ 0 & u_{22} - \lambda & - & ... & - \\ 0 & 0 & u_{33} - \lambda & ... & - \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & u_{nn} - \lambda \end{vmatrix}$$

$$= (u_{11} - \lambda)(u_{11} - \lambda) \cdot \cdot \cdot (u_{nn} - \lambda) = 0$$

Such that $u_{11} = \lambda_1, u_{22} = \lambda_2, ..., u_{nn} = \lambda_n$. Knowing that $U$ is the upper triangular form of $A$, then

$$\det(A) = \lambda_1 \lambda_1 \cdot \cdot \cdot \lambda_n$$

*The determinant of $A$ is the product of its eigenvalues.*

# Diagonalization

Suppose you have an invertible matrix $A$ and a matrix $S$ whose columns are *independent* eigenvectors of $A$.

In the chapter above we've seen that the determinant of $A$ is the product of the eigenvalues, such that $A$ is reducible to an upper triangular matrix with eigenvalues in the main diagonal.

Take the product of $A$ and $S$

$$SA = \begin{bmatrix} . & . & & . \\ . & . & & . \\ x_1 & x_2 & ... & x_n \\ . & . & & . \\ . & . & & . \end{bmatrix} \cdot A = \begin{bmatrix} . & . & & . \\ . & . & & . \\ \lambda_1 x_1 & \lambda_2 x_2 & ... & \lambda_n x_n \\ . & . & & . \\ . & . & & . \end{bmatrix} = \Lambda S$$

Where $A$ is

$$A \to \Lambda = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$$

With the equation above, we can get $A$ or $\Lambda$ (capital lambda), in the case only when $S$ is invertible

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

Diagonalization is another form of factorization, instead of LU-Factorization or QR from Gram-Schmidt.

But how can you prove that $S$ is a matrix of eigenvectors and $\Lambda$ a diagonal matrix of eigenvalues?

Well, take the following approach. We have basically applied the definition of similar matrix

$$AP = PD$$

where $D$ is a diagonal matrix. This can be written as a linear system like

$$A\mathbf{p_1} = d_1\mathbf{p_1}$$
$$A\mathbf{p_2} = d_2\mathbf{p_2}$$
$$...$$
$$A\mathbf{p_n} = d_n\mathbf{p_n}$$

Note that this is the same as $Ax = \lambda x$, replacing $D = diag(\lambda_1, ..., \lambda_n)$ where $\lambda_i$ is the $i^{th}$ eigenvalue of $A$ and $P = [\mathbf{p_1}, ..., \mathbf{p_n}] \longrightarrow [\mathbf{x_1}, ..., \mathbf{x_n}]$ where each

$\mathbf{x_i}$ is a column vector, namely, the $i^{th}$ eigenvector of $A$. Then we have that $P \rightarrow S$ and $D \rightarrow \Lambda$ and

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

# Powers of $A$

If $A$ is a square matrix, and $Ax = \lambda x$ then

$$A^2 x = \lambda A x = \lambda^2 x$$

The eigenvectors seem to remain the same as we take powers of $A$. If we do the same for the diagonalization formula we get

$$A^2 = S\Lambda S^{-1} S\Lambda S^{-1}$$

$S^{-1}S = I$, so the result is $A^2 = S\Lambda^2 S^{-1}$.

Both results indicate that the eigenvectors remain the same doesn't matter the power of $A$ taken. This diagonalization makes the power of a matrix easier to compute, because taking the powers of a diagonal matrix is just taking the power of the pivots, so

$$\Lambda^2 = \begin{bmatrix} (\lambda_1)^2 & 0 & ... & 0 \\ 0 & (\lambda_2)^2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & (\lambda_n)^2 \end{bmatrix}$$

In general

$$A^k x = \lambda^k x \longrightarrow A^k = S\Lambda^k S^{-1}$$

## Example theorem

Suppose you have an invertible matrix $A$ with independent eigenvectors, arranged in $S$ as columns (eigenvector matrix). When do the powers of $A$ go to zero?

The question is resumed as the following equation

$$\lim_{n \to \infty} A^n = 0$$

has to hold for some values of $A$.

First, we diagonalize $A$, so

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

Now, knowing that the eigenvalue matrix $\Lambda$ is diagonal, and contains $\lambda_i$ in each position, for some $i = 1, 2, ..., n$, and taking the $n^{th}$ power of $A$, we get that

$$\lim_{n \to \infty} A^n = \lim_{n \to \infty} S^{-1} \Lambda^n S$$

Such that $A \to 0$ if and only if

$$|\lambda_i| < 1$$

# Diagonalizable matrices

A matrix is diagonalizable if the eigenvectors are all independent. We know that, if all the eigenvalues $\lambda_i$ of a matrix $A$ are different, then the eigenvectors $x_i$ must be different.

To demonstrate that if all eigenvalues are different, the eigenvectors are independent, we will define a linear transformation $T(x)$

$$T(x) = \lambda_i x_i$$

Now, the independence of eigenvectors $x_i$ for $i = 1, 2, ..., n$ can be proven if

$$\sum_{i=1}^{n} c_i x_i = 0$$

holds if and only if $c_1 = 0, c_2 = 0, ..., c_n = 0$. Now, applying $T$ in the sum

$$T\left(\sum_{i=1}^{n} c_i x_i\right) = \sum_{i=1}^{n} c_i T\left(x_i\right) = \sum_{i=1}^{n} c_i \lambda_i x_i = 0$$

by the linearity of $T$. We now have a set of eigenvalues, not necessarily complete. If we take the $k^{th}$ eigenvalue and multiply that by $\sum_{i=1}^{n} c_i x_i = 0$, then subtract the latest equation to form

$$\lambda_k \sum_{i=1}^{n} c_i x_i - \sum_{i=1}^{n} c_i \lambda_i x_i = \sum_{i=1}^{n} (\lambda_k - \lambda_i) c_i x_i = 0$$

This means that if two eigenvalues are equal, then $c_i$ is not necessarily zero, thus the eigenvalues are dependent. Which proves that a sufficient condition for eigenvectors to be independent is a full set of different eigenvalues, none of them zero.

But what if there are repeated eigenvalues? We'll first prove that there is some matrix that will hold the equation despite having repeated eigenvalues. This matrix is the identity, because it's a diagonal matrix with just ones in the main diagonal. This means that the eigenvalues of $I$ are $\lambda_i = 1$.

The eigenvectors are all independent, if you take a $2 \times 2$ identity

$$I_2 x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x = 0$$

such that $x_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}$. For an $n$ dimension identity, it's all the same. The eigenvectors are all independent, meaning that $S$, the eigenvector matrix, is invertible.

Now, there will be cases in which the null space of $A - \lambda I$ will have less dimensions than $A$, meaning not all the eigenvectors are independent. An example for this is the following matrix

$$R = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

Note that this matrix has eigenvalues $\lambda_i = 2$, it's easy to see because the matrix is in upper triangular form. Now, taking the determinant of the eigenvalue shifted matrix

$$\det(R - \lambda I) = (2 - \lambda)(2 - \lambda) = 0$$

makes it clearer. Now, if you try to solve the main equation $Ax = \lambda x$ you will get

$$(A-2I)x = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x = 0$$

but there's only one unique solution for this system, which is $x_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$. The rest of the solutions are arbitrary, because the matrix is singular. This ultimately means that the matrix only has one eigenvector, the eigenvector matrix is not invertible.

# Difference equations

If $A$ is any $n \times n$ invertible matrix, and there is a given vector $u_0$ of $n$ components, then the following equation can be written

$$u_{k+1} = Au_k$$

Starting with the given vector, we have

$$u_1 = Au_0, u_2 = Au_1 = A^2u_0, ..., u_k = A^ku_0$$

for these type of equations, the diagonalization process is useful. Taking the powers of $A$ can be challenging and unnecessarily long. If we diagonalize the matrix, supposing we have an eigenvector matrix $S$, non-singular, then

$$u_k = A^ku_0 = S\Lambda^kS^{-1}u_o$$

which is a much simpler form of the original matrix.

Now, if all eigenvectors are independent, there exists a combination of eigenvector such that, if $x_i$ is the $i^{th}$ eigenvector, then

$$u_0 = c_1x_1 + c_2x_2 + ... + c_nx_n$$

Now, because $A$ is reducible to the upper triangular form, we have that it is composed of the $\lambda$'s on the main diagonal. Then

$$Au_0 = c_1\lambda_1x_1 + c_2\lambda_2x_2 + ... + c_n\lambda_nx_n$$

if $C = \langle c_1, c_2, ..., c_n \rangle$ and $\Lambda$ is the eigenvalue matrix, then

$$u_k = A^ku_0 = C\Lambda^kS$$

# Solving differential equations

Eigenvalues and eigenvectors can be used to solve systems of differential equations, we will explore first order differential equations, to start we will get the solution for a sole equation given by

$$\frac{dy}{dx} + P(x)y = Q(x)$$

In the case where $Q(x) = 0$ for all $x$, we can write this as

$$\frac{dy}{dx} + P(x)y = 0$$

and solve for $y$

$$\begin{aligned} &\int \frac{dy}{y} = \int -P(x)\,dx \\ &ln(y) = -\int P(x)\,dx \\ &y = e^{-\int P(x)\,dx} \end{aligned}$$

such that

$$y = P(x)e^{-A(x)} \qquad \text{if } A(x) = \int P(x)dx$$

Now suppose that $P$ is continuous in an open interval $l$. Let's choose any point in $l$ and let $b$ be any real number. Then exists one function $y = f(x)$ that satisfies the initial value problem

$$y' + P(x)y = 0, \quad \text{with} \quad f(a) = b$$

and it is given by

$$f(x) = be^{-A(x)}, \quad \text{where} \quad A(x) = \int_a^x P(t)dt$$

A good derivation of the formulas can be found in "Calculus Vol.1, Tom M. Apostol". Given that $P$ and $Q$ are continuous functions on an open interval $l$.

We choose any point in $l$ and let $b$ be any real number. Then, there exists a function and only one $y = f(x)$ that satisfies the equation

$$y' + P(x)y = Q(x), \quad \text{with} \quad f(a) = b$$

and it is given by

$$f(x) = be^{-A(x)} + e^{-A(x)} \int_a^x Q(t)e^{A(t)}\, dt,$$

where $A(x) = \int_a^x P(t)dt$.

Now, for a system of differential equations where $u$ is a transformation $u$ : $\mathbb{R} \to \mathbb{R}^2$

$$\begin{cases} \frac{du_1}{dt} = -u_1 + 2u_2 \\ \frac{du_2}{dt} = 2u_1 - u_2 \end{cases} \qquad \text{where } u(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ and } u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

can be represented in matrix form, where $A$ is the coefficient matrix, putting the equation in the form

$$\frac{du}{dt} = Au$$

shields the solution to a pure exponential as $u(t) = e^{At}u(0)$, but what does this mean? Essentially we should note that $At = \lambda t$ and, by substituting it to the solution, we get a system of solutions

$$u_i(t) = u_i(0)e^{\lambda_i t}x_i$$

where $\lambda_i$ and $x_i$ will be the $i^{th}$ eigenvalue and eigenvector, for $i = 1, 2, ..., n$. This would give $n$ solutions that should satisfy the linear equation $\frac{du}{dt} = Au$, then if each $u_i$ satisfies, the sum $u_1 + u_2 + ... + u_n$ also does it, so we get that the total solution $u(t) = \sum_{i=1}^{n} u_i(t)$.

For the example above we have that the coefficient matrix is

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

Now to solve the initial conditions, note that the solution for $u(t)$ is given by

$$u(t) = c_1 x_1 e^{\lambda_1 t} + c_2 x_2 e^{\lambda_2 t}$$

This because $(A - \lambda I)x = 0$ solves for $x_1, x_2$, the constant value is from the solution to the differential equation and corresponds to the initial value (when $t = 0$), and from the equation we also have that $u_1 = \lambda_1, u_2 = \lambda_2$.

So this problem reduces to finding the eigenvalues

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

then $\lambda_1 = 0, \lambda_2 = -3$. Now to find the eigenvectors we look for $(A - \lambda I)x = 0$

$$(A - \lambda I)x = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} x = \begin{bmatrix} -1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 1 \end{bmatrix} x_2 =$$

so $x_1 = \langle 2, 1 \rangle$ and $x_2 = x_1 = \langle -1, 1 \rangle$.

With this result we have solved the system of differential equations. The function $u(t)$ is given by

$$u(t) = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{-3t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{0t} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{-3t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

To find $c_1$ and $c_2$ we know the initial condition $u(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, so we solve

$$u(0) = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

If $c_1 = 1$ and $c_2 = 2$ we get $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$, so the values for both constants must divide by $3$, getting

$$c_1 = \frac{1}{3}, \quad c_2 = \frac{2}{3}$$

# Eigenvector matrix

Now, given that $A$ is square and can be written in the form

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

we can use Taylor series to expand the general solution to the following differential equation:

$$\frac{du}{dt} = Au \qquad \text{where } u(t) = e^{At}u(0)$$

The Taylor series for the exponential is given by

$$e^x = \sum_n^\infty \frac{x^n}{n!},$$

and we can do $x = At$, the definition of Taylor series for matrices remains the same, noting that $A$ is a square matrix

$$e^{At} = \sum_n^\infty \frac{(At)^n}{n!} = I + At + \frac{1}{2}A^2t^2 + ... + \frac{1}{n!}A^nt^n$$

Now, if we rewrite $A = S\Lambda S^{-1}$

$$SS^{-1} + S\Lambda S^{-1}t + \frac{1}{2}(S\Lambda S^{-1})^2t^2 + ... + \frac{1}{n!}(S\Lambda S^{-1})^nt^n$$

by the properties mentioned above, we have that $A^k = S\Lambda^kS^{-1}$, so we can simplify the terms

$$SS^{-1} + S\Lambda S^{-1}t + \frac{1}{2}S\Lambda^2S^{-1}t^2 + ... + \frac{1}{n!}S\Lambda^nS^{-1}t^n$$

Which is the same as

$$S\left(\sum_n^\infty \frac{(\Lambda t)^n}{n!}\right)S^{-1} = Se^{\Lambda t}S^{-1}$$

Finally, the general solution can be rewritten as

$$u(t) = u(0) \cdot Se^{\Lambda t}S^{-1},$$

which is a more simple form. We can rewrite $e^{\Lambda t}$ as

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & ... & \\ & & & e^{\lambda_n t} \end{bmatrix}$$

This general form of $u(t)$ is very similar to the previous example. This last one is just

$$u(t) = u(0) \cdot S(e^{\lambda_1 t} + e^{\lambda_2 t} + ... + e^{\lambda_n t})S^{-1}$$