MATHEMATICS FOR SCIENCE STUDENTS

An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

$a^b = e^{b\log(a)}$

$(a+b)^n = \sum\limits_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$T\left(\alpha\vec{u} + \beta\vec{v}\right) = \alpha T\left(\vec{u}\right) + \beta T\left(\vec{v}\right)$

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

$\Gamma(z) = \int\limits_{0}^{\infty} t^{z-1} e^{-t}\,\mathrm{d}t$

$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$

$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim\limits_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$

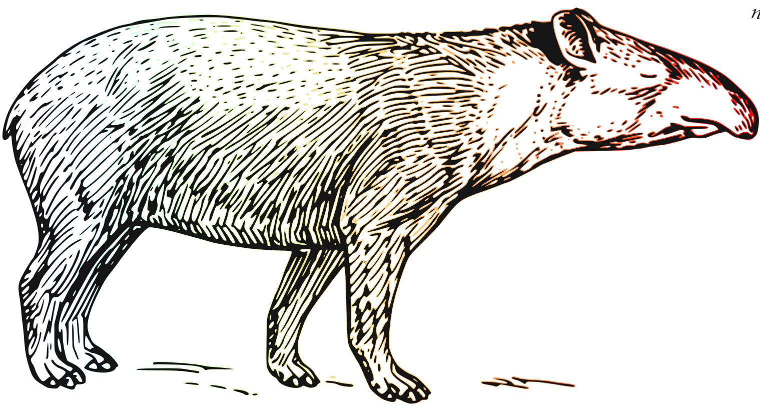
$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$

$\vec{v} = \sum\limits_{i=1}^{n} \alpha_i \hat{e}_i$

$e^{\pi i} + 1 = 0$

$\int\limits_{a}^{b} f(x)\,\mathrm{d}x = F(b) - F(a)$

$\cos(x) = \sum\limits_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}$



PUBLISHED IN THE WILD

! **To be written**: Rights, lefts, etc. will be written here in the future !

HERE BE TABLE

CHAPTER 0

LINEAR ALGEBRA
(INTUITIVE APPROACH)

Linear algebra is one of the most important and often used fields, both in theoretical and applied mathematics. It brings together the analysis of systems of linear equations and the analysis of linear functions (in this context usually called linear transformations), and is employed extensively in almost any modern mathematical field, e.g. approximation theory, vector analysis, signal analysis, error correction, 3-dimensional computer graphics and many, many more.

In this book, we divide our discussion of linear algebra into to chapters: the first (this chapter) deals with a wider, birds-eye view of the topic: it aims to give an intuitive understanding of the major ideas of the topic. For this reason, in this chapter we limit ourselves almost exclusively to discussing linear algebra using 2- and 3-dimensional analysis (and higher dimensions when relevant) using real numbers only. This allows us to first create an intuitive picture of what is linear algebra all about, and how to use correctly the tools it provides us with.

The next chapter takes the opposite approach: it builds all concepts from the ground-up, defining precisely (almost) all basic concepts and proving them rigorously, and only then using them to build the next steps. This approach has two major advantages: it guarantees that what we build has firm foundations and does not fall apart at any future point, and it also allows us to generalize the ideas constructed during the process to such extent that they can be used as foundation to build ever newer tools we can apply in a wide range of cases.

0.1 EIGENVECTORS AND EIGENVALUES

0.1.1 Definition

Some linear transformations have special directions which only scale by the application of the transformation and are not mapped to different directions. Take for example the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, which scales space by 2 in the y-direction. All vectors pointing in the y-direction get scaled by T (namely by a factor of 2) and still point in the y-direction after the application of T. All vectors pointing in the x-direction do not change at all (i.e. they are "scaled" by a factor of 1), and of course still point in the x-direction after the application of T. Any other vector - i.e. those that have both components different than zero - change their direction after the application of T (see Figure 0.1).

We call such vectors the **eigenvectors** of the transformation. The amount by which the are scaled is then their respective **eigenvalues**.

Note 0.1 Pronounciation

The word *eigen* is a German word meaning "own" (as in "own rules"), or "self" (as in self-made). We will see how this meaning fits the concept later in the section. The *ei* part it pronounced the same as the English word "eye", and the *g* is pronounced like the *g* in the English word dog (i.e. unlike the *g* in *generation*).

!

Example 0.1 Eigenvectors and eigenvalues

Text here

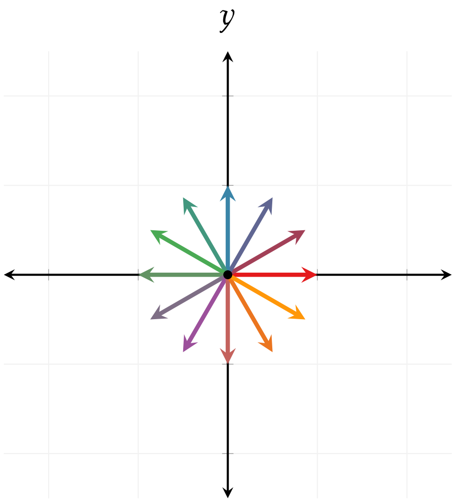
In matrix form, a vector $\vec{v}$ is an eigenvector of a transformation represented by the matrix A, if

$$A\vec{v} = \lambda\vec{v}, \tag{0.1.1}$$

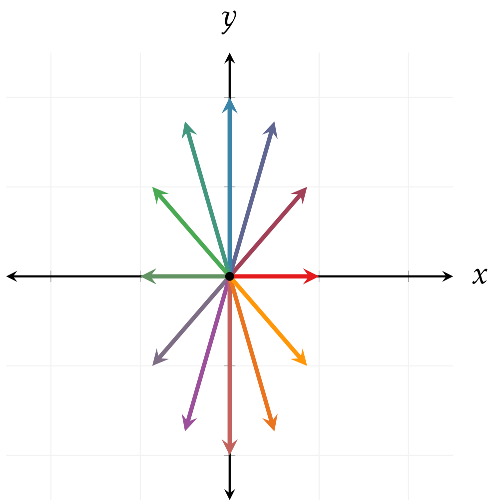
where $\lambda \in \mathbb{R}$, $\lambda \neq 0$. This kind of equation is typically called an **eigenvector equation**. When there are several eigenvectors for a transformation, each with its distinct eigenvalue, we simply add indeces to all relevant parts:

$$A\vec{v}_i = \lambda_i\vec{v}_i, \tag{0.1.2}$$

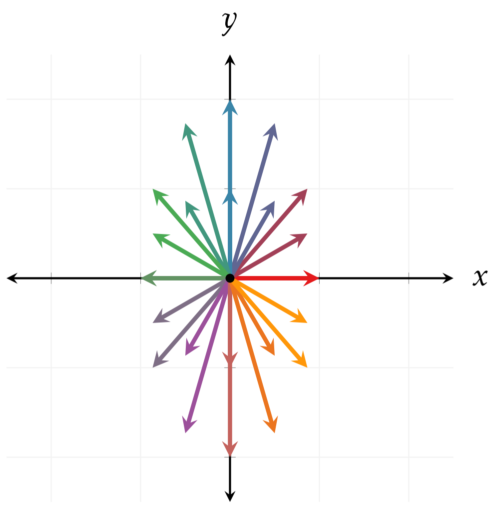
where again $\lambda_i \in \mathbb{R}$ and $\lambda_i \neq 0$.



(a) Some vectors.



(b) Same vectors after the application of T.



(c) The vectors before and after the application of T layered on top of eachother.

Figure 0.1 Some vectors before and after application of the y-scaling transformation T. Note how only the vectors pointing in the direction of the x- and y-axes stay in the same direction, while all the other vectors change their directions.

Note 0.2 The zero-vector

Although technically the zero vector is indeed "scaled" by any linear transformation - by infinitely many scalars - it is **not** considered an eigenvector, exactly because of the fact it has no unique eigenvalue.

!

Before continuing to explore some more examples of eigenvectors, there are two properties[1] of eigenvectors that are important to mention. Given a linear transformation T,

- A scale of any eigenvector $\vec{v}$ of T is also an eigenvector of T, with the same eigenvalue.

  Proof 0.1 Eigenvector scale

  Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation represented by the square matrix A, with eigenvector $\vec{v}$ and its respective eigenvalue λ. Then

  $$A\vec{v} = \lambda\vec{v}.$$

  Replacing $\vec{v}$ with a scale of itself, i.e. $\vec{u} = \alpha\vec{v}$, then applying A to $\vec{u}$ gives us

  $$A\vec{u} \overset{(1)}{=} A\left(\alpha\vec{v}\right) \overset{(2)}{=} \alpha A\vec{v} \overset{(3)}{=} \alpha\lambda\vec{v} \overset{(4)}{=} \lambda\alpha\vec{v} \overset{(5)}{=} \lambda\vec{u}.$$

  where

  (1) Substitution of $\vec{u}$ by its definition $\vec{u} = \alpha\vec{v}$.

  (2) Due to the linearity of A we can bring α out of the product.

  (3) Resulting due to $\vec{v}$ being an eigenvector of A.

  (4) The product of real numbers is commutative.

  (5) Substituting back $\alpha\vec{v} = \vec{u}$.

  Therefore, $\vec{u}$ is also an eigenvector of A (and thus T) with the same eigenvalue λ as $\vec{v}$.

  QED

  Since a linear transformation never has just a single eigenvector but infinitely many (i.e. its entire span), we will refer from now on the **families** of eigenvectors, all pointing in the same direction, represented by a single vector (usually a unit vector, but not necesseraly).

- The linear combination of two eigenvectors of T is **not neccessarily an eigenvector of T**! For example, consider the above transformation which scales all vectors by 2 in the y-direction: as we saw, any vector in the x-direction is an eigenvector of the transformation, and so does any vector in the y-direction. Specifically, the vectors $\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are two separate eigenvectors of the transformation

[1] actually one property and one non-property

(with eigenvalues 1 and 2, respectively), however the vector

$$\vec{c} = \vec{a} + \vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is **NOT** an eigenvector of the transformation, since

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \overset{T}{\mapsto} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

0.1.2 Some examples

Example 0.2 Eigenvectors and eigenvalues of common linear transformations

Let's find the eigenvectors and their respective eigenvalues for the some basic linear transformations in 2-dimensions.

- **Skew (shear) in the x-direction**: any vector pointing in the x-direction is unchanged (i.e. scaled by a factor of 1), while any other vector changes its direction. Thus, the transformation has only a single family of eigenvectors, which we will represent by the vector $\hat{x}$, and their respective eigenvalue $\lambda = 1$.

- **Rotation around the origin**: no vector (except the zero-vector) is scaled by rotation, and therefore rotations have no eigenvectors. However, in 3-dimensions any rotation transformation does have an eigenvector, with eigenvalue $\lambda = 1$: it is of course the family of vectors pointing in the direction of the axis of rotation. We will discuss this further later in the section.

- **Reflection across the x-axis**: in this case, any vector pointing in the x-axis is not being changed - an eigenvector with eigenvalue $\lambda = 1$, and any vector pointing in the y-direction is reflected verically, i.e.

$$\begin{bmatrix} 0 \\ \beta \end{bmatrix} \overset{T}{\mapsto} \begin{bmatrix} 0 \\ -\beta \end{bmatrix} = -\begin{bmatrix} 0 \\ \beta \end{bmatrix},$$

  and is therefore an eigenvector with eigenvalue $\lambda = -1$. The transformation has no other eigenvectors.

- **Reflection across a line going through the origin**: much like the previous case, any vector lying on the reflection line will not change ($\lambda = 1$), and any vector pointing in an orthogonal direction to the line will flip ($\lambda = -1$). No other eigenvectors exist for this transformation.

0.1.3 Calulating eigenvectors

Calculating the eigenvectors of a given transformation, and their respective eigenvalues, is a rather easy procedure to perform once one has the transformation in matrix form. However, in order to understand *why* the procedure works it is useful to derive it first, which is what we'll do now.

We take the eigenvector equation (Equation 0.1.1) and rearrange it slightly:

$$A\vec{v} - \lambda\vec{v} = \vec{0}. \tag{0.1.3}$$

We can then group together all parts which include $\vec{v}$, but we must be careful: A is a matrix while λ is a scalar. This means that the term $A - \lambda$ has no meaning, since we haven't defined how to add or subtract matrices and real numbers. We therefore change Equation 0.1.3 a bit without changing its validity, by replacing $\lambda\vec{v}$ with $\lambda I\vec{v}$: i.e. instead of scaling $\vec{v}$ by a scalar, we scale it using the matrix λI, which yields the same result:

$$A\vec{v} - \lambda I\vec{v} = \vec{0}. \tag{0.1.4}$$

Note 0.3 That one weird trick

If you are not convinced the above trick works, consider the following:

$$\lambda = 3, \vec{v} = \begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix}.$$

Then the direct scaling of $\vec{v}$ by λ is

$$\lambda\vec{v} = \begin{bmatrix} 3 \\ 15 \\ -21 \end{bmatrix},$$

and scaling it using λI yields

$$\lambda I\vec{v} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ -7 \end{bmatrix} = \begin{bmatrix} 3 \\ 15 \\ -7 \end{bmatrix},$$

i.e. we get exactly the same result.

!

Grouping together the parts with $\vec{v}$ gives:

$$(A - \lambda I)\vec{v} = \vec{0}. \tag{0.1.5}$$

Note that $A - \lambda I$ is a matrix, with the following form:

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}. \end{aligned} \tag{0.1.6}$$

Equation 0.1.5 tells us that $A - \lambda I$ sends the vector $\vec{v}$ to $\vec{0}$, and therefore $\vec{v}$ is in the kernel of $(A - \lambda I)$. Sinve by the definition of an eigenvector $\vec{v} \neq \vec{0}$, this means that the kernel of $A - \lambda I$ has more then just the zero vector, and thus

$$|A - \lambda I| = 0. \tag{0.1.7}$$

(this is derived from **??** and **??**)

Therefore, if we solve Equation 0.1.7 for λ, we will get all the values of λ for which the eigenvector equation holds, and in turn we get all the eigenvalues λ_i of the linear transformation represented by A. We can then substitute each λ_i into the eigenvector equation and find its respective eigenvector family.

Example 0.3 Eigenvectors and eigenvalues of a 2×2 matrix

The matrix representing the y-scaling transformation discussed in the beginning of this section is

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

Therefore, in order to find the eigenvectors of A we solve the equation

$$0 = |A - \lambda I| = \begin{vmatrix} 1 - \lambda & 0 \\ 0 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda).$$

In this case there are two solutions: $\lambda_1 = 1$ and $\lambda_2 = 2$. To find their corresponding eigenvectors, we subtitute them into the eigenvectors equation. First λ_1:

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},$$

which corresponds to the following system of equations:

$$\begin{cases} 1 \cdot x + 0 \cdot y = x, \\ 0 \cdot x + 2 \cdot y = y, \end{cases}$$

for which the solution is $x \in \mathbb{R}$ and $y = 0$ - i.e. any vector pointing in the x-direction. Now for $\lambda_2 = 2$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix},$$

i.e. the following system of equations:

$$\begin{cases} 1 \cdot x + 0 \cdot y = 2x, \\ 0 \cdot x + 2 \cdot y = 2y. \end{cases}$$

The solution is of course $x = 0$ and $y \in \mathbb{R}$, meaning any vector pointing in the y-direction.

Example 0.4 Eigenvectors and eigenvalues of a 3×3 matrix

Let us now calculate the eigenvectors and their respective eigenvalues for the following 3×3 matrix:

$$A = \begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}.$$

We start by calculating the determinant $|A - \lambda I|$:

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 5-\lambda & -6 & -3 \\ 0 & -1-\lambda & 0 \\ 0 & 2 & 2-\lambda \end{vmatrix} \\ &= (5-\lambda) \begin{vmatrix} -1-\lambda & 0 \\ 2 & 2-\lambda \end{vmatrix} - (-6) \cancel{\begin{vmatrix} 0 & 0 \\ 0 & 2-\lambda \end{vmatrix}} + 3(-3) \cancel{\begin{vmatrix} 0 & -1-\lambda \\ 0 & 2 \end{vmatrix}} \\ &= (5-\lambda)(-1-\lambda)(2-\lambda) \end{aligned}$$

The solutions of $|A - \lambda I| = 0$ are therefore

$$\lambda_1 = 5,\ \lambda_2 = -1,\ \lambda_3 = 2.$$

- $\boldsymbol{\lambda_1 = 5}$: solving $A\vec{v} = 5\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 5 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5x \\ 5y \\ 5z \end{bmatrix},$$

i.e.

$$\begin{cases} 5x - 6y - 3z &= 5x, \\ 0x - 1y + 0z &= 5y, \\ 0x + 2y + 2z &= 5z, \end{cases}$$

for which the solution is

$$x \in \mathbb{R},\ y = 0,\ z = 0.$$

The first family of eigenvectors are the vectors pointing in the x-direction, and their respective eigenvalue is $\lambda = 5$. We can verify this:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x - 6 \cdot 0 - 3 \cdot 0 \\ 0x - 1 \cdot 0 + 0 \cdot 0 \\ 0x + 2 \cdot 0 + 2 \cdot 0 \end{bmatrix} = \begin{bmatrix} 5x \\ 0 \\ 0 \end{bmatrix} = 5 \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}.$$

- $\boldsymbol{\lambda_2 = -1}$: solving $A\vec{v} = -\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = - \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -x \\ -y \\ -z \end{bmatrix},$$

i.e.

$$\begin{cases} 5x - 6y - 3z & = -x, \\ 0x - 1y + 0z & = -y, \\ 0x + 2y + 2z & = -z, \end{cases}$$

for which the solution is

$$x = \frac{2}{3}y,\ y \in \mathbb{R},\ z = -x.$$

Verifying the solution using the representative vector $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}$:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \cdot 2 - 6 \cdot 3 + 3 \cdot 2 \\ 0 \cdot 2 - 1 \cdot 3 + 0 \cdot (-2) \\ 0 \cdot 2 + 2 \cdot 3 + 2 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 10 - 18 + 6 \\ -3 \\ 6 - 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \\ 2 \end{bmatrix} = - \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}.$$

- $\boldsymbol{\lambda_3 = 2}$: solving $A\vec{v} = 2\vec{v}$ yields

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix},$$

i.e.

$$\begin{cases} 5x - 6y - 3z & = 2x, \\ 0x - 1y + 0z & = 2y, \\ 0x + 2y + 2z & = 2z, \end{cases}$$

for which the solution is

$$x = z,\ y = 0.$$

Verifying the solution using the representative vector $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$:

$$\begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5\cdot 1 - 6\cdot 0 - 3\cdot 1 \\ 0\cdot 1 - 1\cdot 0 + 0\cdot 0 \\ 0\cdot 1 + 2\cdot 0 + 2\cdot 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

To summarize: the linear transformation represented by the matrix

$$A = \begin{bmatrix} 5 & -6 & -3 \\ 0 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

has three families of eigenvectors:

Eigenvalue	**Eigenvector**
$\lambda_1 = 5$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
$\lambda_2 = -1$	$\begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix}$
$\lambda_3 = 2$	$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Example 0.5 The eigenvectors of reflections in $\mathbb{R}^2$, polar form

In $\mathbb{R}^2$, The reflection transformation across the x-axis has two families of eigenvectors: vectors pointing in x-direction (with eigenvalue $\lambda = 1$), and vectors pointing in the y-direction (with eigenvalue $\lambda = -1$). The reflection transformation across the y-axis has the same families, but their eigenvalues are flipped: the vectors pointing in the x-direction have eigenvalue $\lambda = -1$, and those pointing in the y-direction have eigenvalu $\lambda = 1$.

Generalizing to reflections in any direction, we expect to see a similar result: one family of eigenvectors being vectors pointing in the direction of the reflection line with eigenvalue $\lambda = 1$, and the other family of vectors pointing in the orthogonal direction, with eigenvalue $\lambda = -1$. Let us now calculate this directly from the matrix representations of the transformation. We start with the angle-based representation, where the reflection line is represented by its angle θ relative to the x-axis (**??**):

$$\text{Ref}_{\theta} = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}.$$

Using **??**, we expect the vectors with eigenvalue $\lambda = 1$ to have the following form:

$$\begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix},$$

since this is a structure of a vector pointing in the direction with angle θ relative to the x-axis. The ratio of the components of these vectors is therefore expected to be

$$\frac{x}{y} = \frac{\cos(\theta)}{\sin(\theta)}.$$

The vectors with eigenvalue $\lambda = -1$ should be orthogonal to the previous family, i.e. have the structure

$$\begin{bmatrix} -r\sin(\theta) \\ r\cos(\theta) \end{bmatrix},$$

or component ratio of

$$\frac{x}{y} = -\frac{\sin(\theta)}{\cos(\theta)}.$$

Now that we know what to expect, let's start with the actual calculations:

$$\begin{aligned} 0 &= \left|\text{Ref}_{\theta} - \lambda I\right| \\ &= (\cos(2\theta) - \lambda)(-\cos(2\theta) - \lambda) - \sin^2(\theta) \\ &= -\cos^2(2\theta) - \cancel{\lambda\cos(2\theta)} + \cancel{\lambda\cos(2\theta)} + \lambda^2 - \sin^2(\theta) \\ &= \lambda^2 - 1. \end{aligned}$$

(the last equality stems from **??**)
We see that the solutions for the above equation in λ are $\lambda_1 = 1$ and $\lambda_2 = -1$, as expected. Substituting each of these eigenvalues into the eigenvector equation yields:

- $\lambda = 1$: the eigenvector equation

  $$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

  gives rise to the system

  $$\begin{cases} \cos(2\theta)x + \sin(2\theta)y &= x, \\ \sin(2\theta)x - \cos(2\theta)y &= y. \end{cases}$$

  The solution of the system is

  $$y = \frac{x\sqrt{1 - \cos^2(2\theta)}}{\cos(2\theta) + 1}$$

$$
\begin{aligned}
&= \frac{x\sin(2\theta)}{\cos(2\theta)+1} \\
&= \frac{\cancel{2}x\sin(\theta)\cancel{\cos(\theta)}}{\cancel{2}\cos^{\cancel{2}}(\theta)} \\
&= \frac{x\sin(\theta)}{\cos(\theta)}.
\end{aligned}
$$

Rearranging this yields

$$
\frac{x}{y} = \frac{\cos\theta}{\sin\theta},
$$

as expected.

- $\lambda = -1$: the eigenvector equation

$$
\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}
$$

gives rise to the system

$$
\begin{cases} \cos(2\theta)x + \sin(2\theta)y & = -x, \\ \sin(2\theta)x - \cos(2\theta)y & = -y. \end{cases}
$$

The solution of the system is

$$
\begin{aligned}
y &= \frac{x\sqrt{1-\cos^2(2\theta)}}{\cos(2\theta)-1} \\
&= \frac{x\sin(2\theta)}{\cos(2\theta)-1} \\
&= \frac{2x\sin(\theta)\cos(\theta)}{\cancel{1}-2\sin^2(\theta)-\cancel{1}} \\
&= -\frac{\cancel{2}x\cancel{\sin(\theta)}\cos(\theta)}{\cancel{2}\sin^{\cancel{2}}(\theta)} \\
&= -\frac{x\cos(\theta)}{\sin(\theta)}.
\end{aligned}
$$

And rearranging this yields

$$
\frac{x}{y} = -\frac{\sin(\theta)}{\cos(\theta)},
$$

also as expected.

Example 0.6 The eigenvectors of reflections in $\mathbb{R}^2$, slope form

In this example we will again calculate the eigenvectors and eigenvalues of the general reflection transformation, this time by using the slope m of the reflection line, and the matrix representation

$$\mathrm{Ref}_m = \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.$$

Again we expect the two eigenvalues $\lambda = \pm 1$. The eigenvector family corresponding to $\lambda = 1$ should have a component ratio $\frac{y}{x} = m$, i.e. $y = mx$ (the line equation) - and the eigenvector family corresponding to the eigenvalue $\lambda = -1$ should be orthogonal to the other family, i.e. has component ratio $\frac{y}{x} = -\frac{1}{m}$ (i.e. $y = -\frac{1}{m}x$).
<u>The calculation</u>: the determinant $\left|\mathrm{Ref}_m - \lambda\right|$ is then

$$\begin{aligned}
\left|\mathrm{Ref}_m - \lambda\right| &= \begin{vmatrix} \frac{1-m^2}{1+m^2} - \lambda & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} - \lambda \end{vmatrix} \\
&= \left(\frac{1-m^2}{1+m^2} - \lambda\right)\left(\frac{m^2-1}{1+m^2} - \lambda\right) - \frac{4m^2}{\left(1+m^2\right)^2} \\
&= \frac{\left(1-m^2\right)\left(m^2-1\right)}{\left(1+m^2\right)^2} - \lambda\frac{1-m^2}{1+m^2} - \lambda\frac{m^2-1}{1+m^2} + \lambda^2 - \frac{4m^2}{\left(1+m^2\right)^2}.
\end{aligned}$$

Setting $a = 1-m^2$ and $b = 1+m^2$ we get that $m^2 - 1 = -a$, and we continue:

$$\begin{aligned}
\ldots &= \frac{a(-a)}{b^2} - \cancel{\lambda\frac{a}{b}} - \cancel{\lambda\frac{-a}{b}} + \lambda^2 - \frac{4m^2}{b^2} \\
&= \frac{-a^2 - 4m^2}{b^2} + \lambda^2 \\
&= \frac{-(1-m^2) - 4m^2}{\left(1+m^2\right)^2} + \lambda^2 \\
&= \frac{-\left(1-2m^2+m^4\right) - 4m^2}{\left(1+m^2\right)^2} + \lambda^2 \\
&= \frac{-1+2m^2-m^4-4m^2}{\left(1+m^2\right)^2} + \lambda^2 \\
&= \frac{-1-2m^2-m^4}{\left(1+m^2\right)^2} + \lambda^2 \\
&= -\frac{1+2m^2+m^4}{\left(1+m^2\right)^2} + \lambda^2
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{\left(1+m^2\right)^2}{\left(1+m^2\right)^2} + \lambda^2 \\
&= -1 + \lambda^2,
\end{aligned}
$$

i.e. $\lambda^2 = 1$ and thus $\lambda = \pm 1$ as expected. We now subtitute these eigenvalues in the eigenvector equation:

- $\lambda = 1$: the matrix equation is

$$
\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},
$$

for which the system of equations is

$$
\begin{cases} \frac{1-m^2}{1+m^2}x + \frac{2m}{1+m^2}y & = x, \\ \frac{2m}{1+m^2}x + \frac{m^2-1}{1+m^2}y & = y, \end{cases}
$$

and its solution is

$$
y = mx,
$$

i.e. $\frac{y}{x} = m$ as expected.

- $\lambda = -1$: the matrix equation is

$$
\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix},
$$

for which the system of equations is

$$
\begin{cases} \frac{1-m^2}{1+m^2}x + \frac{2m}{1+m^2}y & = -x, \\ \frac{2m}{1+m^2}x + \frac{m^2-1}{1+m^2}y & = -y, \end{cases}
$$

and its solution is

$$
y = -\frac{x}{m},
$$

i.e. $\frac{y}{x} = -\frac{1}{m}$ as expected.

0.1.4 Characteristic polynomial

Did you notice that in all of the above examples the expression $|A - \lambda I|$ is a polynomial in λ? This is not a coincidence: any expression of such form is a polynomial in λ, its degree depending on the form of A. We call this polynomial the **characteristic polynomial** of A. As we saw in the examples, the roots of the characteristic polynomial are the eigenvectors of A. A more precise definition of the characteristic polynomial is given below.

Definition 0.1 Characteristic polynomial of a matrix A

Let A be a square matrix representing some transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Then

$$P(\lambda) = |A - \lambda I|, \tag{0.1.8}$$

is called the characteristic polynomial of A (and T). The roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of P are the eigenvalues of A (and T).

π

Let us examine some values and coefficients of the characteristic polynomial $P(\lambda)$. Substituting $\lambda = 0$ into P produces

$$P(0) = |A - 0I| = |A|, \tag{0.1.9}$$

i.e. the value $P(0)$ is always the determinant of A. Recall that for a polynomial, this value is the free coefficient of P. We can see this fact clearly when we consider a generic 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Its characteristic polynomial is then

$$\begin{aligned} P(\lambda) &= |A - \lambda I| \\ &= \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} \\ &= (a-\lambda)(d-\lambda) - bc \\ &= ad - a\lambda - d\lambda + \lambda^2 - bc \\ &= \underbrace{\boxed{(ad-bc)}}_{|A|} + \underbrace{\boxed{(a+d)}}_{\mathrm{Tr}(A)}\lambda + \lambda^2. \end{aligned}$$

We see that not only is the free coefficient of P equal to the determinant of A, but also that the coefficient of λ is the trace of A. This is not just the case in 2×2 matrices, but all square matrices: the coefficient of λ^{n-1} is the trace of the matrix (up to a sign, i.e. $\pm\mathrm{Tr}(A)$). For example, the characteristic polynomial of a generic 3×3 matrix[2]

$$B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is

$$\begin{aligned} P(\lambda) = |B - \lambda I| &= \begin{vmatrix} a-\lambda & b & c \\ d & e-\lambda & f \\ g & h & i-\lambda \end{vmatrix} \\ &= (a-\lambda)\Big((e-\lambda)(i-\lambda) - fh\Big) - b\Big(d(i-\lambda) - fg\Big) + c\Big(dh - (e-\lambda)g\Big) \end{aligned}$$

[2] note that e and i are just some real numbers, and **not** the constants e and i, respectively.

$$
\begin{aligned}
&= (a-\lambda)\left(ei - e\lambda - \lambda i + \lambda^2 - fh\right) - b\left(di - d\lambda - fg\right) + c\left(dh - eg + \lambda g\right) \\
&= aei - ae\lambda - ai\lambda + a\lambda^2 - afh - ei\lambda - i\lambda^2 - \lambda^3 - fh\lambda - bdi + bd\lambda + bfg + cdh - ceg + cg\lambda \\
&= aei - ae\lambda - afh - ai\lambda + a\lambda^2 - bdi + bd\lambda + bfg + cdh - ceg + cg\lambda - ei\lambda + e\lambda^2 + fh\lambda \\
&+ i\lambda^2 - \lambda^3 \\
&= \boxed{aei - afh - bdi + bfg - cdh - ceg} + (-ae - ai - ei - fh + bd + cg - ei + fh)\lambda \\
&+ \left(\boxed{a + e + i}\right)\lambda^2 - \lambda^3. \qquad (0.1.10)
\end{aligned}
$$

We see that highlighted parts are exactly $\boxed{|B|}$ and $\boxed{\mathrm{Tr}(B)}$, respectively.

0.1.5 Special cases

In the case of upper- or lower-triangular matrices, the eigenvalues are simply the main diagonal elements, e.g. the eigenvalues of the matrix

$$
A = \begin{bmatrix}
a_{11} & 0 & 0 & \cdots & 0 \\
a_{21} & a_{22} & 0 & \cdots & 0 \\
a_{21} & a_{22} & a_{23} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}
\end{bmatrix} \qquad (0.1.11)
$$

are

$$
\lambda = \{a_{11}, a_{22}, a_{33}, \ldots, a_{nn}\},
$$

i.e. written simply,

$$
\lambda_i = a_{ii}. \qquad (0.1.12)
$$

! **To be written**: rest of the section !