-1.1 MATRICES

In the previous section we described linear transformations in a rather abstract way: what they are, how they behave qualitatively and how they look like in 2- and 3-dimensions. In this section we introduce a numerical method of representing linear transformations: matrices.

-1.1.1 Linear transformation of basis vectors

Recall that any vector $\vec{v} \in \mathbb{R}^n$ can be written as a linear combination of basis vectors $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n$:

$$\vec{v} = \sum_{i=1}^{n} \alpha_i \vec{b}_i = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \cdots + \alpha_n \vec{b}_n. \tag{-1.1.1}$$

Applying a linear transformation T on $\vec{v}$ yields, using the properties of linear transformations,

$$\begin{aligned} T(\vec{v}) &= T\left(\alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \cdots + \alpha_n \vec{b}_n\right) \\ \text{additivity} \longrightarrow &= T\left(\alpha_1 \vec{b}_1\right) + T\left(\alpha_2 \vec{b}_2\right) + \cdots + T\left(\alpha_n \vec{b}_n\right) \\ \text{scalability} \longrightarrow &= \alpha_1 T\left(\vec{b}_1\right) + \alpha_2 T\left(\vec{b}_2\right) + \cdots + \alpha_n T\left(\vec{b}_n\right). \end{aligned} \tag{-1.1.2}$$

This result is pretty neat: it means that by knowing how a linear transformation T changes the basis vectors, we know exactly how any vector is transformed by T. This true for any basis, and thus specifically to the standard basis, where the coefficients $\alpha_1, \alpha_2, \dots, \alpha_n$ are actually the components of the vector, i.e. $v_1, v_2, \dots, v_n$. Thus in the standard basis:

$$T(\vec{v}) = v_1 T(\hat{e}_1) + v_2 T(\hat{e}_2) + \cdots + v_n T(\hat{e}_n). \tag{-1.1.3}$$

Example -1.1 Vector transformation via a basis

Applying the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$, defined as

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y-2z \\ 2x+z \\ -x-y-z \end{bmatrix}$$

on the vector $\vec{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ yields the following vector:

$$T(\vec{v}) = T\left(\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 2+(-1)-2\cdot 3 \\ 2\cdot 2+3 \\ -2-(-1)-3 \end{bmatrix} = \begin{bmatrix} 2-1-6 \\ 4+3 \\ -2+1-3 \end{bmatrix} = \begin{bmatrix} -5 \\ 7 \\ -4 \end{bmatrix}.$$

Now, let us apply T first to the three standard basis vectors $\hat{x}, \hat{y}, \hat{z}$:

$$T(\hat{x}) = T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1+0-\cancel{2\cdot 0}\\2\cdot 1+0\\-1-0-0\end{bmatrix} = \begin{bmatrix}1\\2\\-1\end{bmatrix},$$
$$T(\hat{y}) = T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}0+1-\cancel{2\cdot 0}\\\cancel{2\cdot 0}+0\\-0-1-0\end{bmatrix} = \begin{bmatrix}1\\0\\-1\end{bmatrix},$$
$$T(\hat{z}) = T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0+0-2\cdot 1\\\cancel{2\cdot 0}+1\\-0-0-1\end{bmatrix} = \begin{bmatrix}-2\\1\\-1\end{bmatrix}.$$

Taking these results and applying Equation -1.1.2 yields

$$\begin{aligned}
T(\vec{v}) &= 2T(\hat{x}) - T(\hat{y}) + 3T(\hat{z})\\
&= 2\begin{bmatrix}1\\2\\-1\end{bmatrix} - \begin{bmatrix}1\\0\\-1\end{bmatrix} + 3\begin{bmatrix}-2\\1\\-1\end{bmatrix}\\
&= \begin{bmatrix}2\\4\\-2\end{bmatrix} - \begin{bmatrix}1\\0\\-1\end{bmatrix} + \begin{bmatrix}-6\\3\\-3\end{bmatrix}\\
&= \begin{bmatrix}2-1-6\\4-0+3\\-2-(-1)+(-3)\end{bmatrix}\\
&= \begin{bmatrix}-5\\7\\-4\end{bmatrix},
\end{aligned}$$

which is indeed what we got when we applied T directly to $\vec{v}$.

-1.1.2 From transformations to matrices

The most general linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ has the following form:

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}ax+by\\cx+dy\end{bmatrix}, \tag{-1.1.4}$$

where $a, b, c, d \in \mathbb{R}$. If we apply this transformation to $\hat{x}$ and $\hat{y}$ we get, respectively,

$$T(\hat{x}) = \begin{bmatrix}a\\c\end{bmatrix}, \quad T(\hat{y}) = \begin{bmatrix}b\\d\end{bmatrix}. \tag{-1.1.5}$$

We can now collect these two vectors to form a new structure, which we call a **matrix** (in this specific casr a 2×2 matrix):

$$A = \begin{bmatrix}a & b\\c & d\end{bmatrix}. \tag{-1.1.6}$$

We then define the product of M with a vector $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ to yield $T(\vec{v})$, i.e.

$$A\vec{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}. \tag{-1.1.7}$$

This defintion can be re-written as following:

$$A\vec{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A_1 \cdot \vec{v} \\ A_2 \cdot \vec{v} \end{bmatrix}, \tag{-1.1.8}$$

i.e. the i-th component of the resulting vector is the scalar product of the i-th **row** of the matrix with the vector $\vec{v}$.

Example -1.2 Matrix-vector product

Some matrix-vector products:

$$\begin{bmatrix} 1 & -2 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-3) + (-2) \cdot 2 \\ \cancel{0 \cdot (-3)} + 5 \cdot 2 \end{bmatrix} = \begin{bmatrix} -7 \\ 10 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot (-4) \\ 1 \cdot 5 + 2 \cdot (-4) \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 + \cancel{0 \cdot (-2)} \\ \cancel{0 \cdot 2} + 2 \cdot (-2) \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \end{bmatrix}.$$

Challenge -1.1 Proof of linearity

Prove that the transformation T in Equation -1.1.4 is indeed linear.

The most general form of a linear transformation is $T : \mathbb{R}^n \to \mathbb{R}^m$, i.e. a transformation which takes n-dimensional vectors as input and returns m-dimensional vectors as output:

$$\mathbb{R}^n \ni \quad T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}, \quad \in \mathbb{R}^m \tag{-1.1.9}$$

where $a_{ij} \in \mathbb{R}$, $i = 1, 2, 3, \ldots, m$ and $j = 1, 2, 3, \ldots, n$.

Challenge -1.2 Proof of linearity

Prove that the above transformation T is indeed linear.

Respectively, we define an $m \times n$ matrix (m rows by n columns) by collecting all the coefficients a_{ij} into a single structure:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{-1.1.10}$$

The product $M\vec{v}$ (where $\vec{v} \in \mathbb{R}^n$) is then defined as

$$A\vec{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \tag{-1.1.11}$$

Again, note that the i-th component of the resulting vector is the scalar product $A_i \cdot \vec{v}$.

Note -1.1 When is a matrix-vector product defined

In order for a matrix-vector product to be defined, the vector must be of the same dimension as the number of **columns** in the matrix - i.e. given an $a \times b$ matrix, a vector must be b-dimensional for the product to be defined.

Example -1.3 Some matrix-vector products

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

The structure of an $m \times n$ matrix A has a nice property: given that the transformation in represented in some basis $B = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right\}$, the i-th column of the matrix always shows how $\vec{b}_i$ is transformed by the product $A\vec{b}_n$. This is easy to see in the case of the standard basis, which we anyway use througout this chapter:

$$
\begin{array}{cc}
 & \begin{array}{ccccc} T(\hat{e}_1) & T(\hat{e}_2) & & T(\hat{e}_n) \\ \downarrow & \downarrow & & \downarrow \end{array} \\
A = & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.
\end{array}
$$

Example -1.4 Matrices

The product of the following matrix A with each of the vectors $\hat{e}_1, \hat{e}_2, \hat{e}_3$ (i.e. $\hat{x}, \hat{y}$ and $\hat{z}$, respectively) returns the respective column of the matrix:

$$A\hat{e}_1 = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + \cancel{2 \cdot 0} + \cancel{0 \cdot 0} \\ -1 \cdot 1 + \cancel{3 \cdot 0} + \cancel{4 \cdot 0} \\ \cancel{0 \cdot 1} + \cancel{1 \cdot 0} + \cancel{3 \cdot 0} \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix},$$

$$A\hat{e}_2 = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cancel{1 \cdot 0} + 2 \cdot 1 + \cancel{0 \cdot 0} \\ \cancel{-1 \cdot 0} + 3 \cdot 1 + \cancel{4 \cdot 0} \\ \cancel{0 \cdot 0} + 1 \cdot 1 + \cancel{3 \cdot 0} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},$$

$$A\hat{e}_3 = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 3 & 4 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \cancel{1 \cdot 0} + \cancel{2 \cdot 0} + \cancel{0 \cdot 1} \\ \cancel{-1 \cdot 0} + \cancel{3 \cdot 0} + 4 \cdot 1 \\ \cancel{0 \cdot 0} + \cancel{1 \cdot 0} + 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}.$$

-1.1.3 Matrix representation of the basic linear transformations (2D)

We can now represent all of the basic linear transformations in $\mathbb{R}^2$ mentioned in the previous section (**??**) as 2×2 matrices. We do this by observing how the basis vectors $\hat{x}$ and $\hat{y}$ change after the application of each transformation.

- **Identity**: both basis vectors remain the same: $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Therefore the matrix I representing the identity transformation is

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{-1.1.12}$$

- **Scaling by s in the x-direction**: the basis vector $\hat{x}$ is streched by s: $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} s \\ 0 \end{bmatrix}$. The basis vector $\hat{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ stays the same. Therefore the matrix S_x representing the transformation is

$$S_x = \begin{bmatrix} s & 0 \\ 0 & 1 \end{bmatrix}. \tag{-1.1.13}$$

- **Scaling by s in the y-direction**: much like with S_x, now the basis vector $\hat{y}$ is the one getting streched, by β: $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ s \end{bmatrix}$. The basis vector $\hat{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ stays the same. Therefore the matrix S_y representing the transformation is

$$S_y = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}. \tag{-1.1.14}$$

- **Rotating by θ counter-clockwise about the origin**: Figure -1.1 shows how do $\hat{x}$ and $\hat{y}$ transformed by the rotation. In the case of $\hat{x}$, the resulting vector is $R_\theta(\hat{x}) = \left[\cos(\theta), \sin(\theta)\right]$, since thiese are the respective sides of a right triangle of hypotenous 1 and angle θ. The components of $R_\theta(\hat{y})$ can be calculated by rotating $\hat{x}$ by $\theta + \frac{\pi}{2}$ ($\theta + 90°$): $\cos\left(\theta + \frac{\pi}{2}\right) = -\sin(\theta)$, and $\sin\left(\theta + \frac{\pi}{2}\right) = \cos(\theta)$. Therefore we get

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}. \tag{-1.1.15}$$

Altogether the rotation matrix R_θ is

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{-1.1.16}$$

- **Skew by k in the x-direction**: what differentiates this transformation from scaling in the x-direction is that a skew changes only $\hat{y}$ by adding to it some horizontal displacement $\vec{K} = k\hat{x}$ (see Figure -1.2). Therefore $\hat{x}$ remains the same while $\hat{y}$ is transformed as $\hat{y} \rightarrow \hat{y} + \vec{k} = \hat{y} + k\hat{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} k \\ 0 \end{bmatrix} = \begin{bmatrix} k \\ 1 \end{bmatrix}$, and altogether the matrix is

$$K_x = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}. \tag{-1.1.17}$$

- **Skew by k in the y-direction**: same idea, except the roles of the axes are reveresed:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ k \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus the matrix is

$$K_y = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}. \tag{-1.1.18}$$

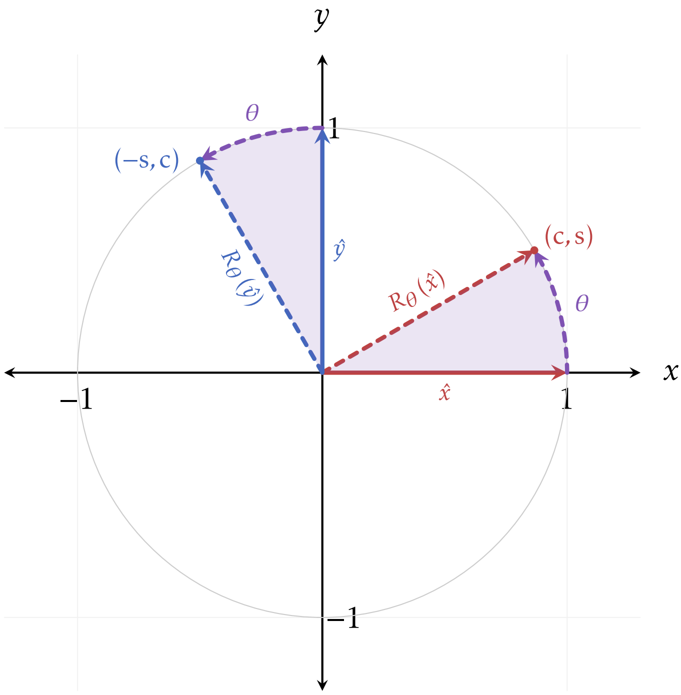


Figure -1.1 Rotation of $\hat{x}$ and $\hat{y}$ by an angle θ counter-clockwise about the origin. The notations c, s stand for $\cos(\theta)$ and $\sin(\theta)$, respectively.

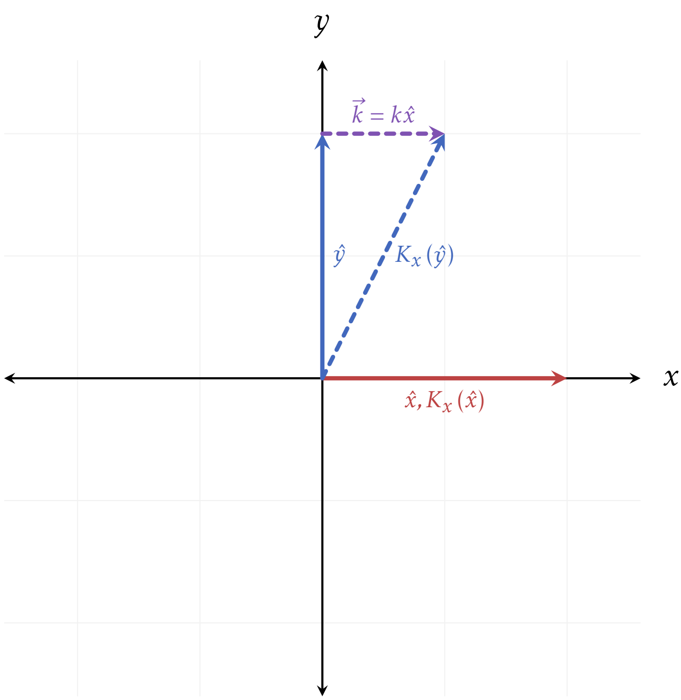


Figure -1.2 Skew in the x-direction.

- **Reflections across a line going through the origin**: in the case of reflections across the x-axis, $\hat{x}$ stays the same, while $\hat{y}$ is flipped (see Figure -1.3(a)), i.e. $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \to -\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. Therefore the matrix is

$$\mathrm{Ref}_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{-1.1.19}$$

Similarily, a reflection across the y-axis flipps $\hat{x}$ while keeping $\hat{y}$ the same (see Figure -1.3(b)), i.e.

$$\mathrm{Ref}_y = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{-1.1.20}$$

Another special case of these kinds of reflections is done across the line rotated by $\frac{\pi}{4} = 45°$ relative to the x-axis, i.e the line $y = x$. In this case $\hat{x}$ and $\hat{y}$ are swapped, giving

$$\mathrm{Ref}_{\frac{\pi}{4}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{-1.1.21}$$

The most general reflection is made across a line of angle θ relative to the x-axis (see Figure -1.3(c)):

$$\mathrm{Ref}_\theta = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}. \tag{-1.1.22}$$

A way to calculate this matrix will be shown later in the chapter.

We can translate the matrix to be based on the slope m of the line instead of its angle θ relative to the x-axis by using the relation $m = \tan(\theta)$ and the two trigonomentric identities for double angles (**??**):

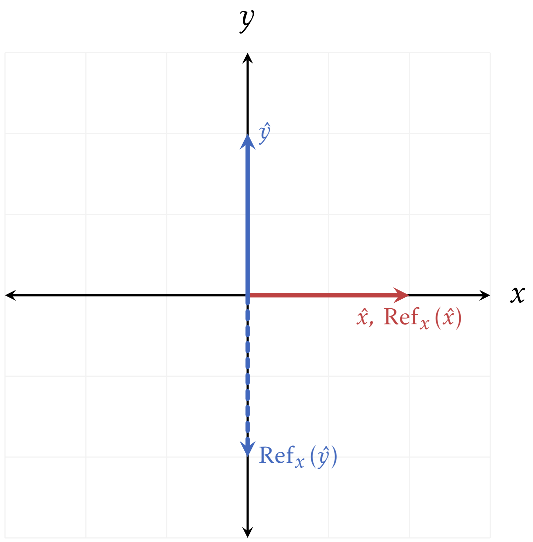
$$\begin{aligned}
\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} &= \begin{bmatrix} \frac{1-\tan^2(\theta)}{1+\tan^2(\theta)} & \frac{2\tan(\theta)}{1+\tan^2(\theta)} \\ \frac{2\tan(\theta)}{1+\tan^2(\theta)} & \frac{\tan^2(\theta)-1}{1+\tan^2(\theta)} \end{bmatrix} \\
&= \frac{1}{1+\tan^2(\theta)} \begin{bmatrix} 1-\tan^2(\theta) & 2\tan(\theta) \\ 2\tan(\theta) & \tan^2(\theta)-1 \end{bmatrix} \\
&= \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}.
\end{aligned}$$

- **Reflection across the origin**: in this case both $\hat{x}$ and $\hat{y}$ are flipped, i.e.

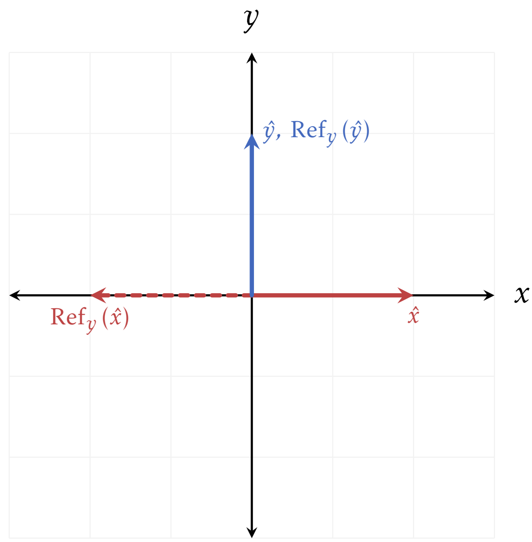
$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

and the matrix is essentially a rotation by π (180°) around the origin:

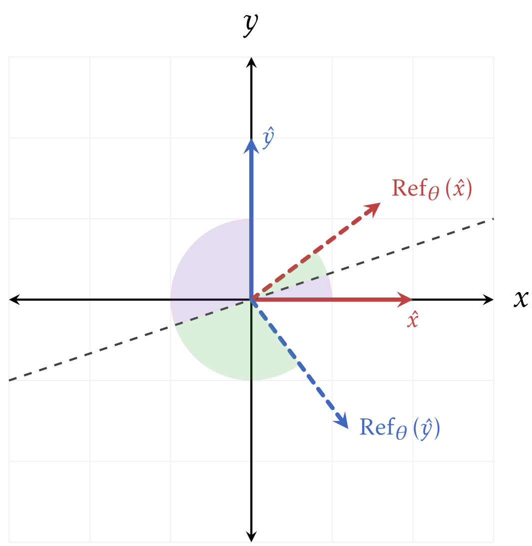
$$R = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{-1.1.23}$$



(a) Reflection across the x-axis.



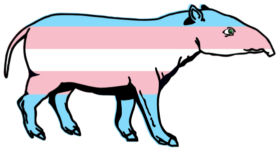
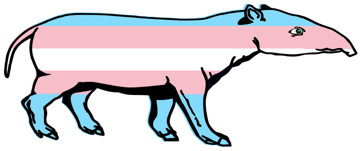
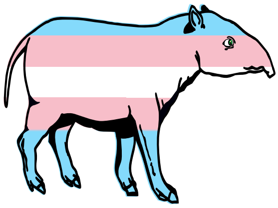
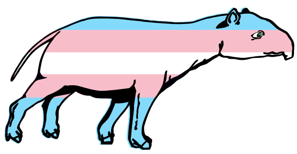
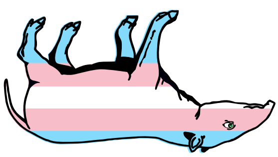
(b) Reflection across the y-axis.

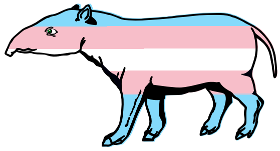
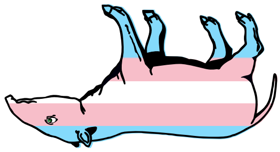


(c) Reflection across a line going through the origin. Notice how in both cases the purple and green angles are the same: this shows that both $\hat{x}$ and $\hat{y}$ are reflected across the line.

Figure -1.3 Reflections across different lines going through the origin.

Table ?? summarizes all the matrices of the basic linear transformations.

Transformation	Trans Tapir	$T(\hat{x})$	$T(\hat{y})$	Matrix
Identity		$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
Scale in x		$\begin{bmatrix}s\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}s & 0\\0 & 1\end{bmatrix}$
Scale in y		$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\s\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & s\end{bmatrix}$
Rotation		$\begin{bmatrix}\mathrm{c}\\-\mathrm{s}\end{bmatrix}$	$\begin{bmatrix}\mathrm{s}\\\mathrm{c}\end{bmatrix}$	$\begin{bmatrix}\mathrm{c} & -\mathrm{s}\\\mathrm{s} & \mathrm{c}\end{bmatrix}$
Skew in x		$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}k\\1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$
Skew in y		$\begin{bmatrix}1\\k\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$
Reflection by x		$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\-1\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$

Reflection by y		$\begin{bmatrix}-1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$ $\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}$
Reflection by line	$y = mx$ 	$\begin{bmatrix}c_2\\s_2\end{bmatrix}$ $\begin{bmatrix}s_2\\-c_2\end{bmatrix}$ $\begin{bmatrix}c_2 & s_2\\s_2 & -c_2\end{bmatrix}$
Reflection about origin		$\begin{bmatrix}-1\\0\end{bmatrix}$ $\begin{bmatrix}0\\-1\end{bmatrix}$ $\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$

-1.1.4 Matrix representation of the basic linear transformations (3D)

In 3-dimensions, the respective matrices are very similar. For example, the matrix for scaling by α in the x-direction, β in the y-direction and γ in the z-direction is

$$S = \begin{bmatrix}\alpha & 0 & 0\\0 & \beta & 0\\0 & 0 & \gamma\end{bmatrix}. \tag{-1.1.24}$$

As mentioned in the previous section, in 3-dimensions there are infinitely many rotations: the axis of rotation can be any line going through the origin (i.e. any vector except $\vec{0}$ can represent an axis of rotation). Let us start with constructing rotations around the three axes x, y and z first. When rotating around the x axis it stays stationary, while the rotation itself is done in the yz-plane. This means that we can take the 2×2 rotation matrix (Equation -1.1.16) and expand it such that it affects only the yz-plane:

$\hat{x}$ doesn't change

$$R_{\theta}^{x} = \begin{bmatrix}1 & 0 & 0\\0 & \cos(\theta) & -\sin(\theta)\\0 & \sin(\theta) & \cos(\theta)\end{bmatrix}. \tag{-1.1.25}$$

2D rotation matrix

A graphical representation of the rotation can be seen in Figure -1.4.

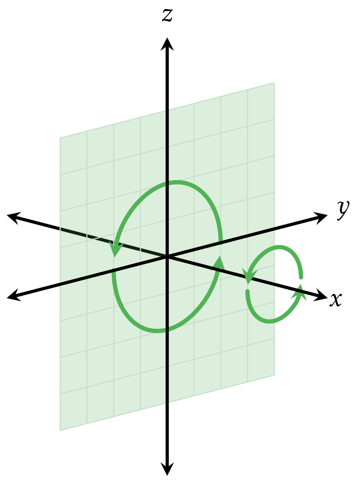


Figure -1.4 In $\mathbb{R}^3$, rotation around the x-axis is a rotation in the yz-plane (i.e. $x = 0$).

The rotation matrices around the y- and z-axes follow the same structure:

$$R^y_\varphi = \begin{bmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ 0 & 1 & 0 \\ -\sin(\varphi) & 0 & \cos(\varphi) \end{bmatrix}, \tag{-1.1.26}$$

$$R^z_\psi = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{-1.1.27}$$

Note -1.2 Direction of the y-axis

The signs of $\sin(\varphi)$ in R^y_φ are flipped compared to R^x_θ and R^z_ψ, for the same reason a similar thing happens in the y-component of the cross product: it is due to the use of a right-handed system.

!

The most general rotation in $\mathbb{R}^3$, i.e. around an axis represented by the unit vector $\hat{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}$ counter-clockwise by an angle θ, is given in matrix form as

$$R_\theta = \begin{bmatrix} \cos(\theta) + u_x^2[1-\cos(\theta)] & u_x u_y[1-\cos(\theta)] - u_z\sin(\theta) & u_x u_z[1-\cos(\theta)] + u_y\sin(\theta) \\ u_y u_x[1-\cos(\theta)] + u_z\sin(\theta) & \cos(\theta) + u_y^2[1-\cos(\theta)] & u_y u_z[1-\cos(\theta)] - u_x\sin(\theta) \\ u_z u_x[1-\cos(\theta)] - u_y\sin(\theta) & u_z u_y[1-\cos(\theta)] + u_x\sin(\theta) & \cos(\theta) + u_z^2[1-\cos(\theta)] \end{bmatrix}. \tag{-1.1.28}$$

For the moment the derivation of this matrix is not presented.

TBW: REFLECTIONS IN 3D.

-1.1.5 Matrix operations

An important operation that can be performed on a matrix is the **transpose**: this operation "rotates" all rows of the matrix to columns, and wise-versa:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \xrightarrow{\text{transpose}} \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}. \tag{-1.1.29}$$

Mathematically, the transpose takes any element a_{ij} of the matrix and exchanges its indeces, yielding a_{ji}. If the original matrix has dimensions $m \times n$, then the transposed matrix has dimensions $n \times m$. The notation for the transpose of a matrix A is $A^{\top}$.

Example -1.5 Transposing matrices

The following presents three matrices each with its transpose. The elements in each matrix on the left hand side are highlighted column-wise, and these colors remain with the elements after the transpose. That way, the effect of the transpose is clear: columns in the original matrix become rows in the transposed matrix and vice-versa. In addition, the dimensions of each matrix are written below it.

$$\underset{3\times 3}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}^{\top}} = \underset{3\times 3}{\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}}$$

$$\underset{2\times 3}{\begin{bmatrix} 0 & 1 & -1 \\ 2 & -3 & 5 \end{bmatrix}^{\top}} = \underset{3\times 2}{\begin{bmatrix} 0 & 2 \\ 1 & -3 \\ -1 & 5 \end{bmatrix}}$$

$$\underset{6\times 1}{\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \\ 7 \\ -4 \end{bmatrix}^{\top}} = \underset{1\times 6}{\begin{bmatrix} 1 & 2 & -1 & 0 & 7 & -4 \end{bmatrix}}$$

Since for the main diagonal elements of a matrix the row and column have equal indeces, the transpose operation does not affect their position in the matrix, i.e. $a_{ii} \xrightarrow{\text{transpose}} a_{ii}$. This means that $\operatorname{tr} A = \operatorname{tr} A^{\top}$. Also, diagonal matrices are not affected by a transpose. The transpose of a transposed matrix is the original matrix, i.e. $\left(A^{\top}\right)^{\top} = A$.

Much like vectors, a matrix can be scaled by a real number, and two matrices can be added together if their dimensions are identical. The rules for scaling a matrix by a scalar and the addition of two matrices are the same as with vectors, namely everything is dome element wise:

- **Scaling**: given a matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

  and a scalar $\gamma \in \mathbb{R}$, their product is

$$\gamma A = \begin{bmatrix} \gamma \cdot a_{11} & \gamma \cdot a_{12} & \cdots & \gamma \cdot a_{1n} \\ \gamma \cdot a_{21} & \gamma \cdot a_{22} & \cdots & \gamma \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma \cdot a_{m1} & \gamma \cdot a_{m2} & \cdots & \gamma \cdot a_{mn} \end{bmatrix}. \tag{-1.1.30}$$

- **Addition**: given two matrices,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix},$$

  their sum is

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}. \tag{-1.1.31}$$

Note -1.3 Matrix addition

Since matrix addition is done **element wise** it is comutative, i.e. for any two $m \times n$ matrices A and B,

$$A + B = B + A.$$

!

-1.1.6 Types of matrices

Any matrix A which represents a transformation of the type $\mathbb{R}^n \to \mathbb{R}^n$ (i.e. from a space onto itself) has the same number of rows and columns (i.e. its dimension is $n \times n$):

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}. \tag{-1.1.32}$$

Due to their shape, such matrices are called **square matrices**. The elements $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ of a square matrix jointly form its **main diagonal** (also: **principal diagonal**):

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}. \tag{-1.1.33}$$

The sum of the main diagonal elements is called the **trace** of the matrix:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}. \tag{-1.1.34}$$

Triangular matrices are matrices in which the elements above or below the main diagonal are all zeros, e.g.

$$\underset{\text{upper triangular}}{U = \begin{bmatrix} 1 & 6 & 6 & -3 \\ 0 & 2 & 7 & 1 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & -4 \end{bmatrix}}, \qquad \underset{\text{lower triangular}}{L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 5 & 1 & -5 & 0 \\ -4 & 1 & 2 & -3 \end{bmatrix}}.$$

A somewhat formal way of defining the elements ”above” the main diagonal is all elements a_{ij} for which $j < i$. Similarily, when $j > i$ the element a_{ij} is ”below” the main diagonal. Note that the transpose of an upper triangular matrix is a lower triangular matrix and vice-versa.

Challenge -1.3 Upper/lower triangular matrices

Show that if A is an upper triangular matrix then $A^{\top}$ is a lower triangular matrix, and if B is a lower triangular matrix then $B^{\top}$ is an upper triangular matrix.

?

A **diagonal matrix** A is a matrix in which all the non-main diagonal elements, i.e. a_{ij} where $i \neq j$, equal zero. These matrices can be thought of as scaling matrices: each entry a_{ii} tells us how the sapce is scaled in the i-th dimension.

Example -1.6 Diagonal matrices

Text.

As we saw in the cases of $\mathbb{R}^2$ and $\mathbb{R}^3$, diagonal matrices are **scaling matrices**: each entry a_{ii} tells us by how much space is scaled in the i-th direction.

A very important family of **square** matrices are the **identity matrices**. These matrices have a unique structure: their main diagonal elements are all 1, while the rest of the elements (the **off-diagonal elements**) are all 0:

$$I_n = \overbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}}^{n \text{ columns}} \left.\vphantom{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}\right\} n \text{ rows} \tag{-1.1.35}$$

Sometimes for clarity large areas of zero-value elements in a matrix are depicted together. In that form, the identity matrix is written as

$$I_n = \begin{bmatrix} 1 & & & \Large 0 \\ & 1 & & \\ & & 1 & \\ \Large 0 & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

In such a depiction, the off-diagonal elements are each written using a single zero. This kind of notation will come in handy in later sections. Yet another way of defining the identity matrix is by using the **Kronecker delta**, which takes two integers i, j and returns 1 if they are equal, otherwise it returns 0:

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{-1.1.36}$$

Using the Kronecker delta, each element a_{ij} of the identity matrix I_n simply equals δ_{ij}.

An identity matrix of dimension n represents the identity transformation in $\mathbb{R}^n$: each standard basis vector $\hat{e}_i$ is left unchanged by the transformation.

Example -1.7 Identity matrices

The following are the identity matrices of $\mathbb{R}^2, \mathbb{R}^3, \ldots, \mathbb{R}^6$, where in each matrix the main diagonal is highlighted:

$$\underset{I_2}{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \quad \underset{I_3}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \quad \underset{I_4}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}} \quad \underset{I_5}{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}} \quad \underset{I_6}{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}}$$

In the next section we will see the importance of the identity matrices.

Another important family of matrices are the **orthogonal matrices** (also **orthonormal matrices**): we say that a matrix Q is an orthogonal matrix if all of its columns, when viewed as column vectors, form an orthonormal set. For example, the identity matrices are all orthogonal matrices. Another orthogonal matrix is the matrix

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{-1.1.37}$$

since both $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$ are unit vectors, and they are orthogonal to each other (as seen in REF).

A **symmetric matrix** is a square matrix for which

$$A^\top = A. \tag{-1.1.38}$$

”Graphically”, the symmetry of such matrices can be seen in respect to their main diagonal: if we imagine placing a mirror on the main diagonal, each element a_{ij} would be ”reflected” across the mirror, and thus be equal to a_{ji} (see example below).

Example -1.8 Symmetric matrix

The following matrix S is a symmetric 4×4 matrix, in which the elements a_{ij}, a_{ji} are higlighted with the same color:

$$S = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 0 & 1 & 3 \\ 5 & 1 & 4 & 2 \\ 7 & 3 & 2 & 6 \end{bmatrix}$$

Note -1.4 Transpose of a symmetric matrix

A symmetric matrix is its own transpose, i.e. if A is a symmetric matrix then $A^\top = A$.

A rather non-interesting family of matrices are the **zero matrices**: these are matrices which have only zero-elements, i.e.

$$\mathbf{0}_n = \overbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}^{n \text{ columns}} \left.\vphantom{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}\right\} m \text{ rows} \tag{-1.1.39}$$

The zero matrices are called that way since for a given matrix A,

$$\begin{aligned} A + \mathbf{0} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11}+0 & a_{12}+0 & \cdots & a_{1n}+0 \\ a_{21}+0 & a_{22}+0 & \cdots & a_{2n}+0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+0 & a_{m2}+0 & \cdots & a_{mn}+0 \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A. \tag{-1.1.40}$$

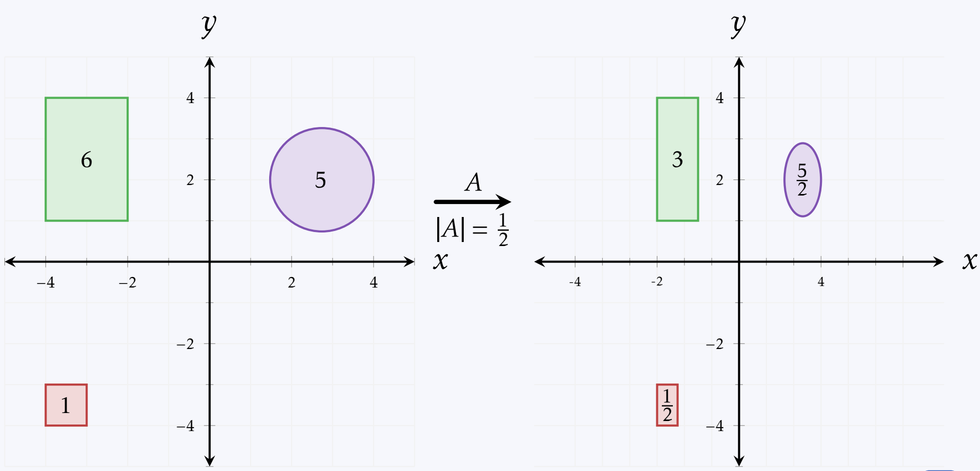
I.e. much like the number zero and the zero vector, the zero matrix is neutral in respect to addition.

-1.1.7 The determinant

As mentioned in **??**, linear transformation scale all volumes by the same amount[1]. This scaling factor is encapsulated in the matrix representing the transformation by a number called the **determinant** of the matrix. The determinant of a matrix A is written as $|A|$ and sometimes $\det(A)$.

Example -1.9 The determinant as a scaling factor

In the following transformation, represented by the matrix $A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$, areas are scaled by a factor of $\frac{1}{2}$ and therefore $|A| = \frac{1}{2}$ (the number inside each shapes is its area):



Since there is not much sense in discussing volume changes between different spaces (e.g. $\mathbb{R}^5 \to \mathbb{R}^7$), only square matrices, which as you recall represent linear transformations from a space onto itself, have determinants. Determinants can take any real number as values, including zero and negative numbers.

What does a zero determinant mean? Since in $\mathbb{R}^2$ determinants tells us the scaling factor of areas by the transformation, if the matrix representing the linear transformation has a

[1]remember that 2-dimensional volumes are areas.

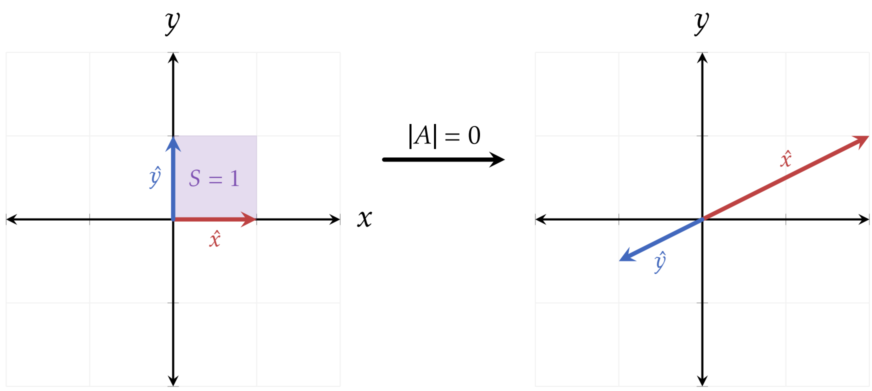


Figure -1.5 A transformation which "squashes" all areas into a line is represented by a matrix A with $|A| = 0$. Note how the unit volume defined by $\hat{x}$ and $\hat{y}$ is transformed into a shape of zero area: a line going through the origin. Also note that $\hat{x}$ and $\hat{y}$ are linearly dependent, since they lie on the same line. Cf. Figure -1.6.

zero determinant, it means that somehow all areas are "squashed" by the transformation to zero. There are two possible relevant shapes of zero area: a line going through the origin, or the origin itself which is a point. See Figure -1.5 for a visualization.

Similarily, in $\mathbb{R}^3$ the determinant tells us how volumes are scaled by a linear transformation, and thus a 3×3 matrix with zero determinant means that all the transformation represented by the matrix "squashes" all volumes to one of three relevant shapes with zero volume: a plane going through the origin, a line going through the origin, or the origin point itself. See **??** for a visualization.

DISCUSSION OF NEGATIVE DETERMINANTS...

To calculate the determinant of a matrix, we start with the simplest case: 2×2 matrices. Since all areas are equaly scaled by a linear transformation, we look at the unit sqaure defined by $\hat{x}$ and $\hat{y}$ (see Figure -1.6). After the application of the transformation represented by the generic matrix $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ (where $a, b, c, d \in \mathbb{R}$), these basis vectors are transformed into the vectors

$$A\hat{x} = \begin{bmatrix} a \\ b \end{bmatrix} \text{ and } A\hat{y} = \begin{bmatrix} c \\ d \end{bmatrix}, \tag{-1.1.41}$$

respectively.

The unit square defined by $\hat{x}$ and $\hat{y}$ is therefore transformed into the parallelogram defined by $A\hat{x}$ and $A\hat{y}$. **??** tells us that the area of the parallelogram is $S = ad - bc$. Therefore, the determinant - which equals the change in area after application of A, is

$$|A| = ad - bc \tag{-1.1.42}$$

as well.

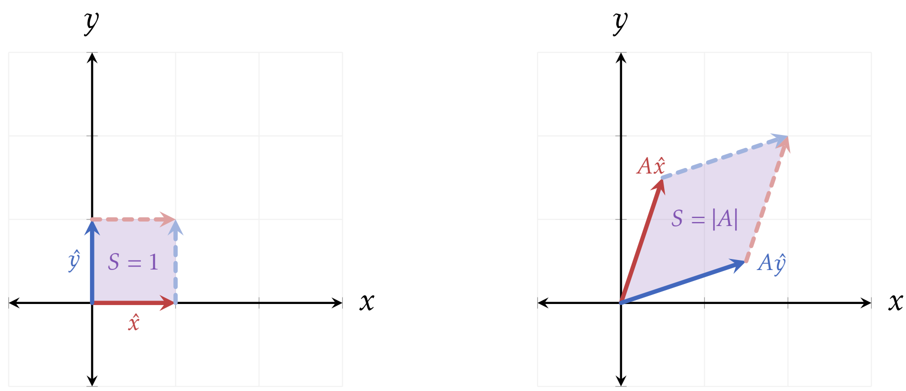


Figure -1.6 Unit area defined by the vectors $\hat{x}$ and $\hat{y}$ before application of a linear transformation represented by the matrix A (left) and the parallelogram defined by the vectors $A\hat{x}$ and $A\hat{y}$ after application of the transformation (right).

Example -1.10 Determinants of 2×2 matrices

Some 2×2 matrices and their determinants:

$$\begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \longrightarrow 1\cdot 3 - (-2)\cdot 0 = 3.$$
$$\begin{bmatrix} 1 & 5 \\ 1 & 3 \end{bmatrix} \longrightarrow 1\cdot 3 - 1\cdot 5 = -2.$$
$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \longrightarrow 1\cdot 2 - 1\cdot 2 = 0.$$
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \longrightarrow 1\cdot 4 - 2\cdot 2 = 0.$$
$$\begin{bmatrix} 0 & 7 \\ 0 & -3 \end{bmatrix} \longrightarrow \cancel{0\cdot(-3)} - \cancel{0\cdot 7} = 0.$$

Calculating the determinant of a 3×3 matrix is based on the calculation of the determinant of a 2×2 matrix. First, we should define an idea called a **minor** of a matrix. The ij-minor of a 3×3 matrix A is the determinant of the 2×2 matrix resulting by the removal of the i-th row and j-th column of A, e.g. let

$$A = \begin{bmatrix} 2 & -5 & 4 \\ -3 & 0 & 2 \\ 3 & 3 & 2 \end{bmatrix},$$

then Table -1.2 shows all the minors of A.

Using the its minors, the determinant of a 3×3 matrix can be calculated using the fol-

Table -1.2 All the minors of the matrix A.

i	j	3×3-matrix	2×2 determinant	value
1	1	$\begin{bmatrix} \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & 0 & 2 \\ \blacksquare & 3 & 2 \end{bmatrix}$	$\begin{vmatrix} 0 & 2 \\ 3 & 2 \end{vmatrix}$	-6
1	2	$\begin{bmatrix} \blacksquare & \blacksquare & \blacksquare \\ -3 & \blacksquare & 2 \\ 3 & \blacksquare & 2 \end{bmatrix}$	$\begin{vmatrix} -3 & 2 \\ 3 & 2 \end{vmatrix}$	-12
1	3	$\begin{bmatrix} \blacksquare & \blacksquare & \blacksquare \\ -3 & 0 & \blacksquare \\ 3 & 3 & \blacksquare \end{bmatrix}$	$\begin{vmatrix} -3 & 0 \\ 3 & 3 \end{vmatrix}$	-9
2	1	$\begin{bmatrix} \blacksquare & -5 & 4 \\ \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & 3 & 2 \end{bmatrix}$	$\begin{vmatrix} -5 & 4 \\ 3 & 2 \end{vmatrix}$	-22
2	2	$\begin{bmatrix} 2 & \blacksquare & 4 \\ \blacksquare & \blacksquare & \blacksquare \\ 3 & \blacksquare & 2 \end{bmatrix}$	$\begin{vmatrix} 2 & 4 \\ 3 & 2 \end{vmatrix}$	-8
2	3	$\begin{bmatrix} 2 & -5 & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare \\ 3 & 3 & \blacksquare \end{bmatrix}$	$\begin{vmatrix} 2 & -5 \\ 3 & 3 \end{vmatrix}$	21
3	1	$\begin{bmatrix} \blacksquare & -5 & 4 \\ \blacksquare & 0 & 2 \\ \blacksquare & \blacksquare & \blacksquare \end{bmatrix}$	$\begin{vmatrix} -5 & 4 \\ 0 & 2 \end{vmatrix}$	-10
3	2	$\begin{bmatrix} 2 & \blacksquare & 4 \\ -3 & \blacksquare & 2 \\ \blacksquare & \blacksquare & \blacksquare \end{bmatrix}$	$\begin{vmatrix} 2 & 4 \\ -3 & 2 \end{vmatrix}$	16
3	3	$\begin{bmatrix} 2 & -5 & \blacksquare \\ -3 & 0 & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare \end{bmatrix}$	$\begin{vmatrix} 2 & -5 \\ -3 & 0 \end{vmatrix}$	-15

lowing formula:

$$|A| = a_{11}m_{11} - a_{12}m_{12} + a_{13}m_{13}, \tag{-1.1.43}$$

where a_{ij} and m_{ij} are the elements and minors of the matrix, respectively. For example, using the above matrix A, we get that

$$\begin{aligned}|A| &= a_{11}m_{11} - a_{12}m_{12} + a_{13}m_{13} \\ &= 2\begin{vmatrix}0 & 2 \\ 3 & 2\end{vmatrix} - (-5)\begin{vmatrix}-3 & 2 \\ 3 & 2\end{vmatrix} + 4\begin{vmatrix}-3 & 0 \\ 3 & 3\end{vmatrix} \\ &= 2\cdot(-6) - (-5)\cdot(-12) + 4\cdot(-9) \\ &= -12 - 60 - 36 = -108.\end{aligned}$$

Example -1.11 Determinants of 3×3 matrices

...

The determinant of a 4×4 matrix follows the same pattern, i.e.

$$|A| = a_{11}m_{11} - a_{12}m_{12} + a_{13}m_{13} - a_{14}m_{14}, \tag{-1.1.44}$$

where again a_{ij} and m_{ij} are, respectively, the elements and minors of a matrix. Much like with the case of a minor of a 3×3 matrix being a determinant of a 2×2 matrix, the minor of a 4×4 matrix is a determinant of a 3×3 matrix, itself calculated using determinants of 2×2 matrices. This pattern continues to higher dimensions, i.e. the calculation of the determinant of a 5×5 matrix uses determinants of 5 4×4 matrices, the calculation of the determinant of a 6×6 matrix uses 6 determinants of 5×5 matrices, and so forth. Therefore, the total number of 2×2 determinants needed for the calculation of the determinant of an $n \times n$ matrix is

$$d = n \times (n-1) \times (n-2) \times \cdots \times 5 \times 4 \times 3 = \frac{n!}{2}. \tag{-1.1.45}$$

Some properties of determinants:

- In any case where the columns of a matrix form a linearly dependent set, the determinant is zero. This is due to the loss of dimensionality (i.e. at least one basis vector is mapped to a vector which can be written as a linear combination of the other vectors). One obvious case is where there is one or more columns of zeros in the matrix.
- The determinant of the transpose of a matrix is the same as the determinant of the original matrix, i.e. $|A| = |A^\top|$. This is due to the fact that all ideas discussed here can be applied directly to row vectors (as mentioned in the previous sections), and the transpose operation essentially switches between these forms: the columns of the original matrix become the rows in its transposed format. Therefore, the previous property applies to the rows of a matrix as well: e.g. a row of zeros means that the determinant is zero.

-1.1.8 Matrix-vector products

As discussed in the first part of this section, matrices represent linear transformations - in fact, we define a matrix in such a way that its product with a vector gives the result of applying the transformation the matrix represents on the vector (see Equation -1.1.11). Let us now review this idea and elaborate a bit on the process of calculating matrix-vector products.

Given an $m \times n$ matrix A and an n-dimensional vector $\vec{v}$, the product $A\vec{v}$ is an m-dimensional vector, in which each element v_i is the scalar product between the i-th **row** of A (interpreted as a vector) and the vector $\vec{v}$ itself. To illustrate this, we use the following 4×3 matrix A and 3-dimensional vector $\vec{v}$:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & -1 \\ 2 & 4 & 0 \\ 6 & 1 & -3 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix}$$

(note that the number of **columns** in A is the same as the number of elements of $\vec{v}$, namely 3).

The resulting vector $\vec{u}$ is then given by the following formula:

$$u_i = \underset{\substack{\uparrow \\ i\text{-th row of } A}}{A^i} \cdot \vec{v}.$$

In the following illustration, each row of A is scalar multiplied with the vector $\vec{v}$, yielding the respective element of $\vec{u}$. The respective rows of A and elements of $\vec{v}$ are color-coded for clarity.

$$\underset{A}{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 2 & 4 & 0 \\ 6 & 1 & 3 \end{bmatrix}} \quad \underset{\vec{v}}{\begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}} \quad \underset{\vec{u}}{\begin{bmatrix} 1 \cdot 2 + \cancel{0 \cdot 5} + \cancel{2 \cdot 0} \\ \cancel{0 \cdot 2} + 3 \cdot 5 + \cancel{1 \cdot 0} \\ 2 \cdot 2 + 4 \cdot 5 + \cancel{0 \cdot 0} \\ 6 \cdot 2 + 1 \cdot 5 + \cancel{3 \cdot 0} \end{bmatrix}} \to \underset{\vec{u}}{\begin{bmatrix} 2 \\ 15 \\ 24 \\ 17 \end{bmatrix}}$$

Example -1.12 Matrix-vector products

The following are some examples of matrix-vector products. Note how in each product the number of columns in the matrix is the same as the number of elements of the vector.

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + \cancel{0 \cdot (-7)} + 2 \cdot 4 \\ 3 \cdot 2 - 1 \cdot (-7) + 5 \cdot 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 33 \end{bmatrix}.$$

$$\begin{bmatrix} 0 & 4 & -5 \\ 4 & 6 & -2 \\ -2 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} \cancel{0 \cdot 0} + 4 \cdot (-3) - 5 \cdot (-2) \\ \cancel{4 \cdot 0} + 6 \cdot (-3) - 2 \cdot (-2) \\ -\cancel{2 \cdot 0} + 2 \cdot (-3) + \cancel{0 \cdot (-2)} \end{bmatrix} = \begin{bmatrix} -2 \\ -14 \\ -6 \end{bmatrix}.$$

$$\begin{bmatrix} -2 & -1 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -\cancel{2 \cdot 0} - 1 \cdot (-1) \\ -\cancel{2 \cdot 0} + 1 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$\begin{bmatrix} 6 & -1 \\ 5 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \cdot 2 - 1 \cdot 2 \\ 5 \cdot 2 + \cancel{0 \cdot 2} \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}.$$

$$\begin{bmatrix} 2 & -1 & -2 & -2 \\ 2 & 0 & 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} \cancel{2 \cdot 0} - 1 \cdot (-2) - 2 \cdot 2 - 2 \cdot 4 \\ \cancel{2 \cdot 0} + \cancel{0 \cdot (-2)} + \cancel{0 \cdot 2} + 5 \cdot 4 \end{bmatrix} = \begin{bmatrix} -10 \\ 20 \end{bmatrix}.$$

$$\begin{bmatrix} 3 & 4 & -1 & 1 \\ -2 & 5 & 6 & -2 \\ 4 & 5 & -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 3 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \cdot 6 + 4 \cdot 3 - \cancel{1 \cdot 0} + 1 \cdot 4 \\ -2 \cdot 6 + 5 \cdot 3 + \cancel{6 \cdot 0} - 2 \cdot 4 \\ 4 \cdot 6 + 5 \cdot 3 - \cancel{1 \cdot 0} + 3 \cdot 4 \end{bmatrix} = \begin{bmatrix} 34 \\ -5 \\ 51 \end{bmatrix}.$$

-1.1.9 Matrix-matrix products

Since the product of a matrix and a vector is itself a vector, one can take the resulting vector and multiply it by another matrix, i.e. given the matrices A, B and a vector $\vec{v}$, the expression

$$B \cdot (A \cdot \vec{v})$$

is a vector as well.

Of course, the dimensions of all participating objects must align for the products to be properly defined: if A is an $m \times n$ matrix, then $\vec{v}$ must be an n-dimensional vector. The result of the product $A \cdot \vec{v}$ is then an m-dimensional vector which we can call $\vec{u}$. Thus, for the product $B \cdot \vec{u}$ to be properly defined, B must have the same number of columns as $\vec{u}$ has elements, namely m columns. The number of rows is free, and can be any natural number k. Therefore, B is a $k \times n$ matrix, and the product $B \cdot \vec{u}$ is a k-dimensional vector.

Example -1.13 Multiple matrix-vector products

Let

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 5 & 4 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 2 \\ 6 & -7 \\ -1 & 0 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix},$$

then

$$\vec{u} = A \cdot \vec{v} = \begin{bmatrix} 2 \cdot 1 - 1 \cdot 2 + \cancel{0 \cdot 5} \\ 5 \cdot 1 + 4 \cdot 2 + 3 \cdot 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 28 \end{bmatrix}.$$

The product $B \cdot \vec{u}$ is defined, since $\vec{u}$ has the same number of elements as B has columns (namely 2). Its result is the 3-dimensional vector

$$\vec{w} = B \cdot \vec{u} = \begin{bmatrix} \cancel{5 \cdot 0} + 2 \cdot 28 \\ \cancel{6 \cdot 0} - 7 \cdot 28 \\ \cancel{-1 \cdot 0} + \cancel{0 \cdot 28} \end{bmatrix} = \begin{bmatrix} 56 \\ -196 \\ 0 \end{bmatrix}.$$

Multiple matrix-vector product therefore represent application of multiple linear transformations on an initial vector, in the order the matrix-vector products are performed. For example, consider the vector $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Rotating $\vec{v}$ by $\frac{\pi}{2}$ (= 90°) counter-clockwise around the origin and then scaling the result by 2 should yield the vector $\vec{w} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \xrightarrow{\text{rotation by } \frac{\pi}{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \xrightarrow{\text{scaling by 2}} \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Using the respective matrix representation of each transformation, we get the following:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \cdot \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

(note that the matrix-vector products are performed from right to left)

More generally, there could be several products made successionally, i.e.

$$A_n \left(A_{n-1} \left(A_{n-2} \cdots \left(A_2 \left(A_1 \vec{v} \right) \right) \right) \right).$$

Matrix-vector products are most commonly written without the paranthesis nor the dot symbol, i.e. as

$$A_n A_{n-1} A_{n-2} \cdots A_2 A_1 \vec{v},$$

and the order of multiplication is from left to right, i.e. A_1 is the first matrix to be multiplied by $\vec{v}$, then A_2 is multiplied by the result of the product $A_1 \vec{v}$, then A_3 is multiplied by the result of $A_2 A_1 \vec{v}$ and so on.

At this point one should wonder whether instead of doing this long chain of products on each individual vector, perhaps the matrices themselves could be multiplied first, yielding a matrix representing the total transformation applied to a vector, as a composition of the separate transformations in the correct order. The answer is of course yes![2]

Let us define the product of two matrices: given the an $m \times n$ matrix A and an $n \times k$ matrix B, the product $C = AB$ is itself a matrix, having the dimension $m \times k$, in which

[2] otherwise this subsection would not be called ”Matrix-matrix products”, after all.

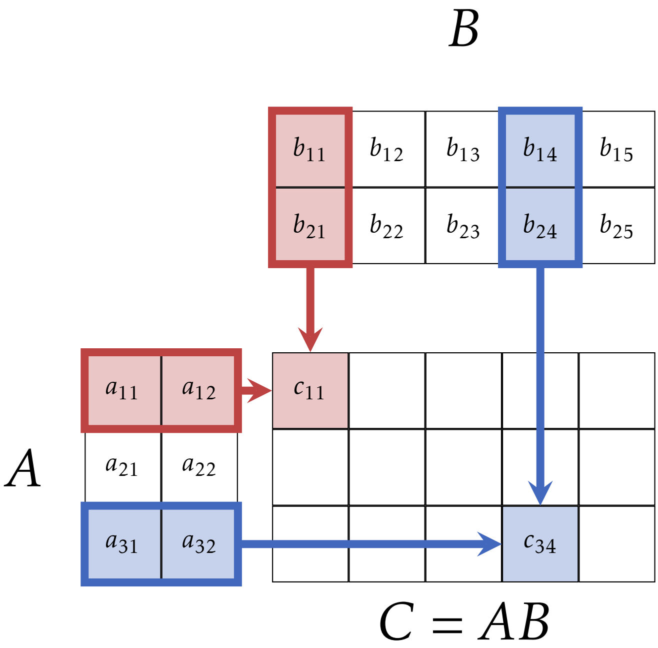


Figure -1.7 The element c_{ij} of the matrix $C = AB$ is the scalar product of the i-th row of A and the j-th column of B.

every element c_{ij} is the scalar product of the row A^i and the column B_j, i.e.

$$c_{ij} = A^i \cdot B_j = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}. \tag{-1.1.46}$$

Figure -1.7 illustrates this idea graphically.

Using the previous example, instead of calculating the $\frac{pi}{2}$ rotation of the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and then scaling it by 2, we calculate the product of the two matrices representing these transformations, and then apply them to the vector:

$$C = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \cancel{0 \cdot 2} - \cancel{1 \cdot 0} & \cancel{0 \cdot 0} - 1 \cdot 2 \\ 1 \cdot 2 + \cancel{0 \cdot 0} & \cancel{1 \cdot 0} + \cancel{0 \cdot 2} \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}.$$

We then apply C to the vector and get the expected result:

$$C\vec{v} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \cancel{0 \cdot 1} - 2 \cdot 1 \\ 2 \cdot 1 + \cancel{0 \cdot 1} \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Example -1.14 Matrix-matrix products

TBW

Recall that the order of composition of linear transformations matters: $T_1 \circ T_2 \neq T_2 \circ T_1$. Therefore, matrix-matrix products, which represent such compositions, are non-commutative, i.e.

$$AB \neq BA. \tag{-1.1.47}$$

Example -1.15 Non-commutativity of matrix-matrix products

Of course, there are special cases where $AB = BA$, but these are the exception and not the norm. An example is the the rotation and scaling of the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ we saw above: if we flip the order of application of the two linear transformations we get the same result. This is true for any vector:

$$\begin{tikzcd}
 & \begin{bmatrix} -y \\ x \end{bmatrix} \arrow[dr, "\text{scale}"] & \\
\begin{bmatrix} x \\ y \end{bmatrix} \arrow[ur, "\text{rotate}"] \arrow[dr, "\text{scale}"'] & & \begin{bmatrix} -2y \\ 2x \end{bmatrix} \\
 & \begin{bmatrix} 2x \\ 2y \end{bmatrix} \arrow[ur, "\text{rotate}"'] &
\end{tikzcd}$$

We can see that fact by multiplying the two matrices directly, in both directions:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \cancel{0 \cdot 2} - \cancel{1 \cdot 0} & \cancel{0 \cdot 0} - 1 \cdot 2 \\ 1 \cdot 2 + \cancel{0 \cdot 0} & \cancel{1 \cdot 0} + \cancel{0 \cdot 2} \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix},$$
$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \cancel{2 \cdot 0} + \cancel{0 \cdot 1} & 2 \cdot (-1) + \cancel{0 \cdot 0} \\ 1 \cdot 2 + \cancel{0 \cdot 0} & \cancel{1 \cdot 0} + \cancel{0 \cdot 2} \end{bmatrix} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}.$$

Later in this chapter we will analyze the conditions for such commutativity to occur.

The determinant of a matrix-matrix product AB equals the product of the determinants of the separate matrices, i.e.

$$|AB| = |A| \times |B|. \tag{-1.1.48}$$

The reason is that the change in volume after application of two consecutive transformations is the product of the change in volume for each separate transformation. This also mean that $|AB| = |BA|$. The trace of a matrix-matrix product behaves the same as well:

$$\operatorname{tr}(AB) = \operatorname{tr}(BA). \tag{-1.1.49}$$

Proof -1.1 Trace of a matrix-matrix product

Prove the above behaviour of the trance operator.

QED

On the other hand, the traspose operator doesn't behave so "nicely":

$$(AB)^{\top} = B^{\top}A^{\top}. \tag{-1.1.50}$$

-1.1.10 Inverse matrices

-1.1.11 Kernel and null space