MATHEMATICS FOR SCIENCE STUDENTS

An open-source book

Written, illustrated and typeset (mostly) by

PELEG BAR SAPIR

with contributions from others

$a^b = e^{b\log(a)}$ $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ $T\left(\alpha\vec{u}+\beta\vec{v}\right) = \alpha T\left(\vec{u}\right) + \beta T\left(\vec{v}\right)$

$A = Q\Lambda Q^{-1}$ $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t$

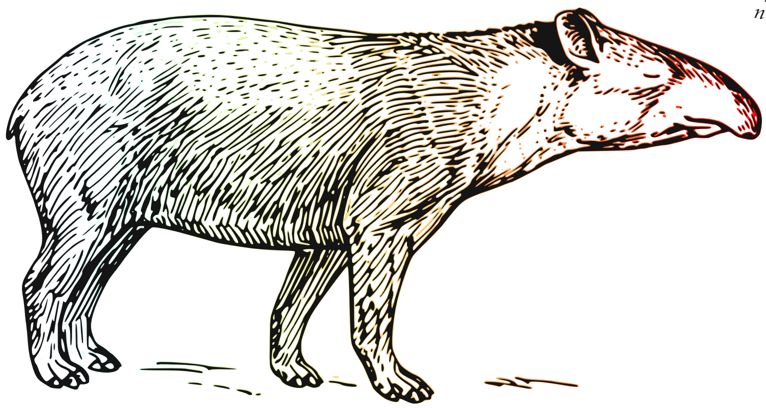
$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ $\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$

$\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{ij}$ $\vec{v} = \sum_{i=1}^{n} \alpha_i \hat{e}_i$

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

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

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



PUBLISHED IN THE WILD

The definitions of intersections and unions can be easily extended to any whole number of sets.

Example 0.13 Intersection and union of 3 sets

The intersection of 3 sets $A = \{1, 2, 3, 4, 5\}$, $B = \{-2, -1, 0, 1, 2\}$ and $C = \{2, 3, 4, 5, 6\}$ is the set of all elements that are in A and in B and in C, i.e. the set $A \cap B \cap C = \{2\}$. The union of these sets is the set of all elements that are in either of the sets, i.e. $A \cup B \cup C = \{-2, -1, 0, 1, 2, 3, 4, 5, 6\}$.

The most general definition of an intersection of n sets (where n is a whole number), which we will call $A_1, A_2, A_3, \cdots, A_n$ is

$$A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n = \Big\{x \,|\, (x \in A_1) \wedge (x \in A_2) \wedge (x \in A_3) \wedge \cdots \wedge (x \in A_n)\Big\}. \tag{0.1.7}$$

The left hand side of Equation 0.1.7 can be written as

$$A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i. \tag{0.1.8}$$

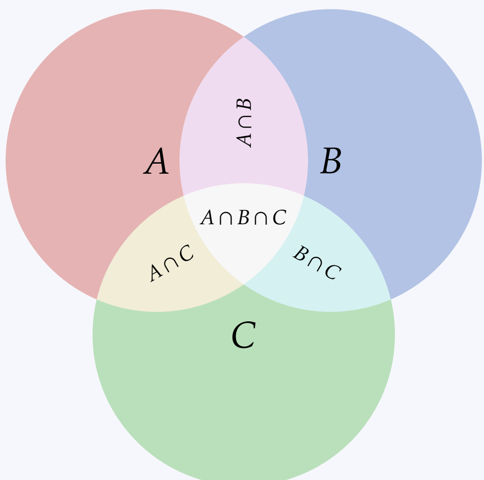
(clarifying the notation? i.e. indexing, etc.)

Similarly, the union of n different sets is defined as

$$\begin{aligned} \bigcup_{i=1}^{n} A_i &= A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n \\ &= \Big\{x \,|\, (x \in A_1) \vee (x \in A_2) \vee (x \in A_3) \vee \cdots \vee (x \in A_n)\Big\}. \end{aligned} \tag{0.1.9}$$

Example 0.14 Venn diagrams: intersection and union of 3 sets

The following Venn diagram shows all possible intersections between three sets:



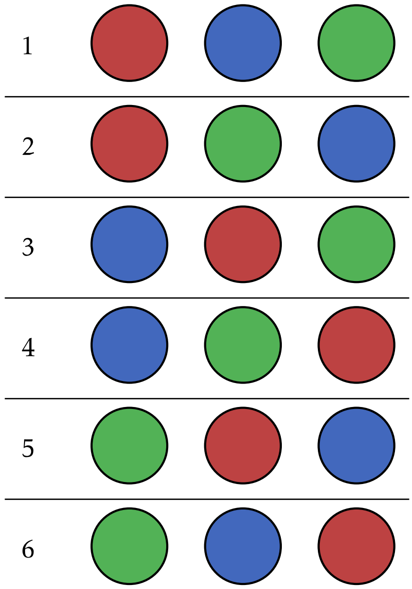


Figure 0.4 All the ways of arranging a set of 3 differently colored spheres.

Going about the options systematically, we can describe them as follows: we have three options for placing the first sphere - red, blue or green. Once we have chosen the first sphere, we're left with only two option for the second sphere: if we chose red, then we're left with choosing between blue and green. The choice of last sphere is then dictated by the previous choices: if for the second sphere we chose green, then we are left with only the blue sphere for the third position (as in option 2 above). We call each of these ways to organize the spheres a **combination**.

Quantitatively the number of ways we have of organizing the spheres is

$$k = 3 \times 2 \times 1 = 6. \tag{0.6.1}$$

We can expand this logic to however many $n \in \mathbb{N}$ different spheres we wish: for n different spheres we have n options for placing the first sphere, then $n-1$ options for placing the second sphere, then $n-2$ options for placing the third sphere... all the way to the last sphere. The number of total combinations is therefore

$$k = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1. \tag{0.6.2}$$

The function used to represent k in the above general form is called the **factorial**, and is denoted using an exclamation mark:

$$k = n! \tag{0.6.3}$$

A somewhat more rigouros (and quite intereseting) way of defining the factorial is as follows:

$$n! := \begin{cases} 1 & \text{if } n = 1, \\ n \times (n-1)! & \text{if } n > 1. \end{cases} \tag{0.6.4}$$

Again, we see that the coefficients in the expansion are exactly the binomial coefficients for $n = 3$ (namely $1, 3, 3, 1$), and that the powers of a and b decrease and increase, respectively, just as for $n = 2$: a goes from 3 to 0, and b from 0 to 3. This time the power of a is always $3 - k$ (where k is the power of b).

Therefore, in the most general case (i.e. $(a+b)^n$) we expect each term to be of the following structure:

$$a^k b^{n-k},$$

and the coefficients being the binomial coefficients $\binom{n}{k}$. Putting this into a formula:

$$\begin{aligned}(a+b)^n &= \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-2}a^2b^{n-2} + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n \\ &= \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k.\end{aligned} \tag{0.6.8}$$

Example 0.46 Exapnding $(a+b)^4$ using binomial coefficients

Let's use the above patterns to expand $(a+b)^4$. Looking at Pascal's triangle, the coefficients in row $n = 4$ are $1, 4, 6, 4, 1$. Starting with the powers a^4 and $b^0 = 1$ and then decreasing and increasing the powers of a and b, respectively, we write all terms:

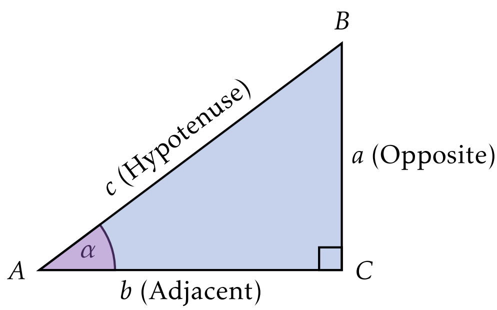
$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$$

You can check this expansion manually and see that it is indeed correct.

0.7 TRIGONOMETRIC FUNCTIONS

0.7.1 Basic Definitions

Consider a **right triangle** $\triangle ABC$ with sides a, b, and Hypotenuse c, where the angle $\angle ACB$ is $90°$, and the angle $\angle BAC$ is denoted as α:



We use the ratios between the three sides of the triangle to define three functions of α:

smaller than some constant $\underline{M} \in \mathbb{R}$. For example, consider the simple sequence $a_n = n$, where $n = \{1, 2, 3, 4, \dots\}$: there is no element in the sequence that is smaller than 1. Therefore, a_n is bounded from below by 1. Of course, one may argue that b_n is also bounded from below by 0, or -6, or in fact any negative number. This is true, however we are usually interested in the *maximal* number $\underline{M}$ that bounds the sequence from below, which in this case is $\underline{M} = 1$. We call that number the **infimum** of the sequence, and denote it as $\inf a_n$.

Similarly, a sequence a_n can be bounded from above by some number $\overline{M} \in \mathbb{R}$, i.e. there exist no n for which $a_n > \overline{M}$. We call the *minimal* such number the **supremum** of the sequence a_n, denoted $\sup a_n$. For example, the sequence $b_n = \frac{1}{n}$ is bounded from above by any real number $x \geq 1$, and therefore $\sup b_n = 1$. In fact, b_n is also bounded from below by $\underline{M} = 0$, and therefore we say that it is **bounded**. Another example of a sequence that is bounded is $e_n = \sin(n)$, which is bounded from below by $\underline{M} = -1$ and from above by $\overline{M} = 1$.

Example 1.6 Bounded and unbounded sequences

The following table shows some examples of sequences that are bounded from below, from above, or neither:

a_n	**First 5 elements**	$\inf a_n$	$\sup a_n$
$n^2 - n$	$0, 2, 6, 12, 20, \dots$	0	-
$\frac{n}{n+1}$	$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \dots$	$\frac{1}{2}$	1
e^{-n}	$e^{-1}, e^{-2}, e^{-3}, e^{-4}, e^{-5}, \dots$	0	e^{-1}
$\log(n)$	$0, \log(2), \log(3), \log(4), \log(5), \dots$	0	-
$(-1)^n$	$-1, 1, -1, 1, -1, \dots$	-1	1
$(-1)^n n$	$-1, 2, -3, 4, -5, \dots$	-	-
$(-2)^n$	$-2, 4, -8, 16, -32, \dots$	-	-

1.1.3 Subsequences

Given any sequence a_n, we can remove from it any number of its elements (including 0 elements) and get a new sequence b_n which is a **subsequence** of a_n. For example, let $a_n = n^2 - 5n$. We can remove each 2nd element from a_n (i.e. those with indices $2, 4, 6, 8, \dots$) and get the following sequence b_n:

$$a_n = -4, -6, -6, -4, 0, 6, 14, 24, 36, 50, 66, 84, \dots$$

$$b_n = -4, -6, 0, 14, 36, 66, \dots$$

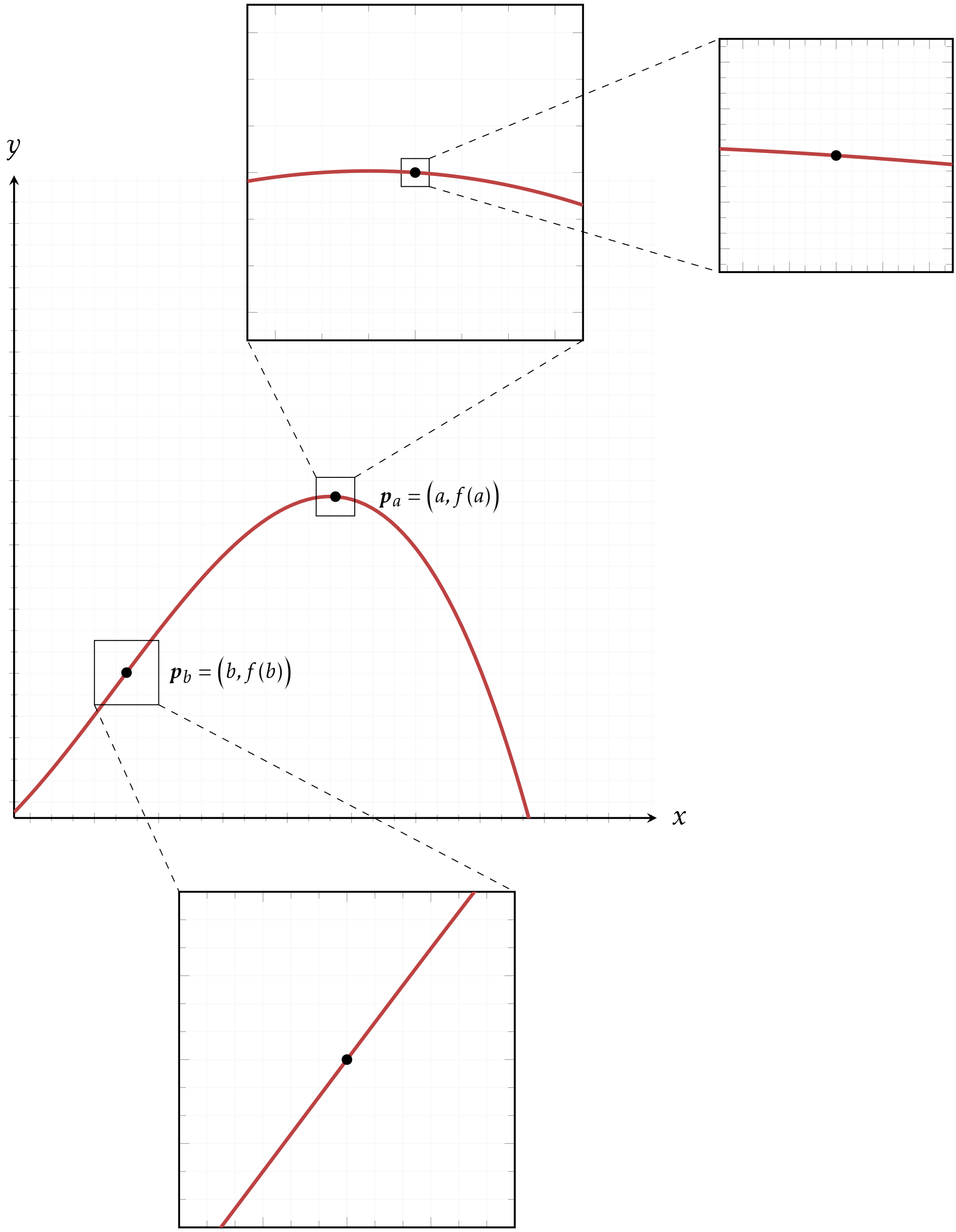


Figure 1.13 Zooming in on a real function f at two points: $\boldsymbol{p}_a = \big(a, f(a)\big)$ (upper right) and $\boldsymbol{p}_b = \big(b, f(b)\big)$ (bottom right). Note how around each of the points, the function looks somewhat linear: this is more pronounced around $\boldsymbol{p}_b$ where the function looks linear in the entire zoomed-in area, while near $\boldsymbol{p}_a$ it looks linear only near the point itself even though the zoom factor is higher.

Let's define linear combinations a little more formaly:

Definition 2.2 Linear combinations

A linear combination of n vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of the same dimension, using n scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, is an expression of the form

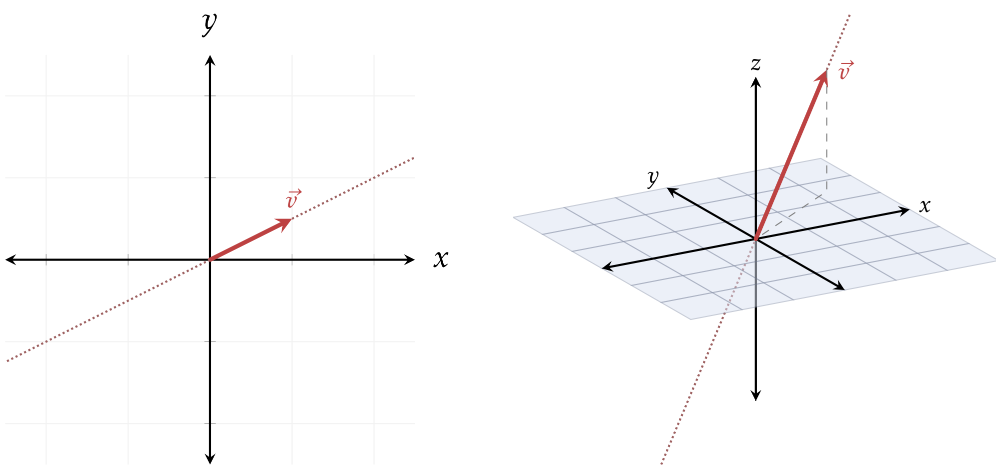
$$\vec{w} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n = \sum_{i=1}^{n} \alpha_i \vec{v}_i. \tag{2.1.15}$$

π

Linear combinations of real vectors have geometric meaningsc: we start with the set of all linear combinations of a single vector $\vec{v} \in \mathbb{R}^n$, i.e.

$$V = \left\{ \alpha \vec{v} \mid \alpha \in \mathbb{R} \right\}. \tag{2.1.16}$$

The set V represents a line in the direction of $\vec{v}$ going through the origin (see Figure 2.6). The set V is itself a vector space of dimension 1, and as such a **subspace** of $\mathbb{R}^n$. We say that it is the **span** of the vector $\vec{v}$ (i.e. the vector $\vec{v}$ **spans** the subspace V).



1 $\mathbb{R}^2$ 2 $\mathbb{R}^3$

Figure 2.6 The span of a single vector $\vec{v}$, shown as a dashed line: in $\mathbb{R}^2$ (left) and $\mathbb{R}^3$ (right).

Similarily, the set of all linear combinations of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ that are not scales of each other (i.e. there is no such $\alpha \in \mathbb{R}$ for which $\vec{v} = \alpha \vec{u}$),

$$V = \left\{ \alpha \vec{u} + \beta \vec{v} \mid \alpha, \beta \in \mathbb{R} \right\}, \tag{2.1.17}$$

is a plane that goes through the origin (see Figure 2.7). Such vectors are also said to be **non-collinear**.

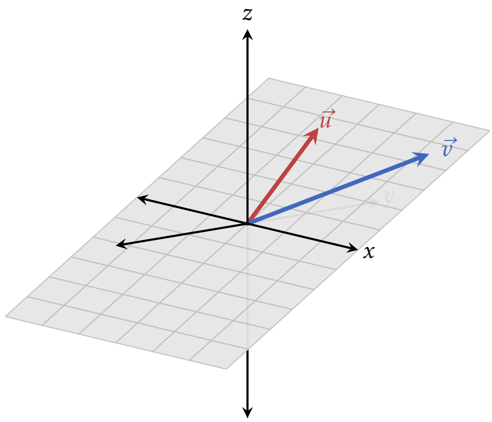


Figure 2.10 The angle between two linearly independent vectors lies on the plane spanned by the vectors.

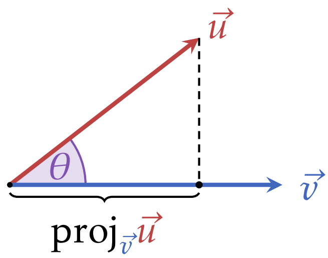


Figure 2.11 The projection of a vector $\vec{u}$ onto another vector $\vec{v}$ in the plane spanned by the two vectors.

2.1.6 The scalar product

When given two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ it is often useful to know the angle between them: if the two vectors are linearly dependent then the angle is either $\theta = 0$ if they point in the same direction, or $\theta = \pi$ if the point in opposite directions (remember: we measure angles in radians). Otherwise, the angle θ can take any value in $(0, \pi)$. Angles are always measured on a plane, and in the case of two linearly independent vectors that plane is of course the one spanned by the two vectors (Figure 2.10).

If considering only the plane the vectors span, we can rotate it such that one of the vectors, say $\vec{u}$, lies horizotally (see Figure 2.11). We then drop a perpendicular line from the head of the $\vec{u}$ to the horizontal vector $\vec{v}$. We call the length from the origin to the intersection point of $\vec{v}$ and the perpendicular line the **projection** of $\vec{u}$ onto $\vec{v}$, and denote it as $\mathrm{proj}_{\vec{v}}\vec{u}$.

Since the origin, the head of $\vec{u}$ and the intersection point of the perpendicular line with $\vec{v}$ form a right triangle, using basic trigonometry we find that the cosine of the angle θ is

$$\cos(\theta) = \frac{\mathrm{proj}_{\vec{v}}\vec{u}}{\|\vec{u}\|}. \tag{2.1.30}$$

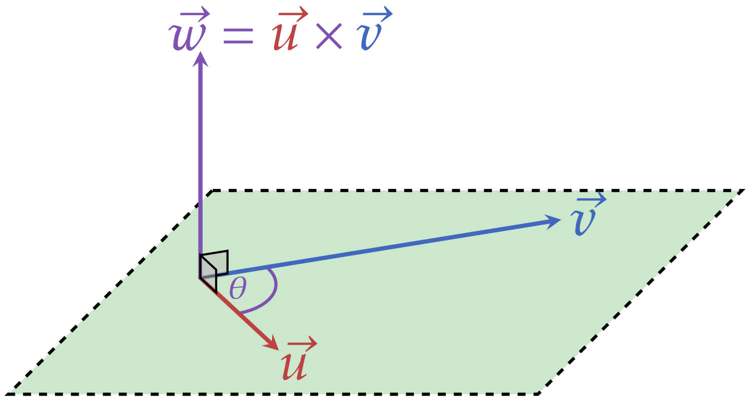


Figure 2.14 The cross product of the vectors $\vec{u}$ and $\vec{v}$ relative to the plane spanned by the two vectors.

while choosing the latter gives

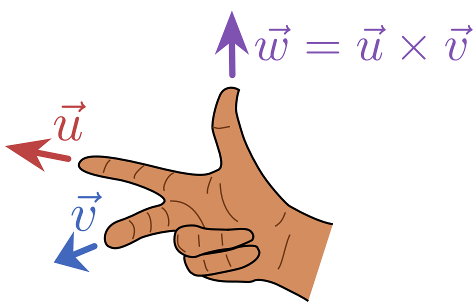
$$\vec{u} \times \vec{v} = \begin{bmatrix} b \\ -a \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = bc - ad. \tag{2.1.45}$$

These two forms are the opposite of each other - i.e. if one yields the value 4, the other yields the value -4. We will see which one is used in a moment.

On to $\mathbb{R}^3$: geometrically, the cross product of two vectors $\vec{u}, \vec{v} \in \mathbb{R}^3$ is defined as a **vector** $\vec{w} \in \mathbb{R}^3$ which is **orthogonal to both** $\vec{u}$ and $\vec{v}$, and with norm of the same magnitude as the product would have in $\mathbb{R}^2$, i.e.

$$\left\| \vec{w} \right\| = \left\| \vec{u} \right\| \left\| \vec{v} \right\| \sin(\theta). \tag{2.1.46}$$

The direction of $\vec{u} \times \vec{v}$ is determined by the **right-hand rule**: using a person's right hand, when $\vec{u}$ points in the direction of their index finger and $\vec{v}$ points in the direction of their middle finger, then vector $\vec{w} = \vec{u} \times \vec{v}$ points in the direction of their thumb:



The cross product is **anti-commutative**, i.e. changing the order of the vectors results in inverting the product:

$$\vec{u} \times \vec{v} = -\left(\vec{v} \times \vec{u} \right).$$

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

Challenge 2.4 Proof of linearity

Prove that the transformation T in Equation 2.3.4 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:

$$\mathbb{R}^n \ni \quad T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}, \quad \in \mathbb{R}^m \tag{2.3.9}$$

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

Challenge 2.5 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:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{2.3.10}$$

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

$$A\vec{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \tag{2.3.11}$$

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

Note 2.7 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.

$$I_n = \begin{array}{c} \overbrace{\hphantom{1 \quad 0 \quad 0 \quad \cdots \quad 0}}^{n \text{ columns}} \\ \left.\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}\right\} n \text{ rows} \end{array} \tag{2.3.35}$$

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

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

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:

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \tag{2.3.36}$$

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 2.28 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