Basic algebra for data analysis (refresher)

Marco Corneli - M1DS

September 15, 2020

Contents

1 Vector spaces and linear maps

1.1 Basics

Definition. *A (sub)* ***vector space*** *is a set V whose elements called* vectors *can be added and multiplied by a real (or complex) scalar in such a way that*

i) $v_1, v_2 \in V \Rightarrow v_1 + v_2 \in V$,

ii) $\lambda \in \mathbb{R}, v \in V \Rightarrow \lambda v \in V$.

Example. The set of points $(x, y) \in \mathbb{R}^2$ on the line of equation $y = x$ forms a vector space. The set of points $(x, y) \in \mathbb{R}^2$ on the parable of equation $y = x^2$ is *not* a vector space.

Definition. *A subset of vectors $\{v_1, v_2, \ldots, v_N\} \subset V$ is a* ***generating system*** *(GS) for V if $\forall v \in V$ it holds that*

$$v = \alpha_1 v_1 + \cdots + \alpha_N v_N, \qquad \exists \quad (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N$$

Exercise. Prove that $\{(1, 0), (0, 1), (1, 2)\}$ is a GS for $\mathbb{R}^2$.

Definition. *The vectors $v_1, \ldots, v_N$ are* ***linearly independent*** *if the only solution of the system $\alpha_1 v_1 + \cdots + \alpha_N v_N = 0$ is*

$$\alpha_1 = \cdots = \alpha_N = 0.$$

Exercise. Show that the vectors $(1, 0), (0, 1), (1, 2)$ are *not* linearly independent.

Definition. *A generating system of linearly independent vectors of V forms a* ***basis****. The number of vectors in the the basis is the dimension of the vector space V.*

Examples.

1. $\mathbb{R}^3$ is a vector space. The canonical basis is given by $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

2. $R_2[x]$ denotes the set of real polynomials of order 2

$$R_2[x] = \{a_0 + a_1 x + a_2 x^2 | (a_0, a_1, a_2) \in \mathbb{R}^3\}.$$

It is a vector space and the canonical basis of $R_2[x]$ is $\{1, x, x^2\}$.

Proposition 1. *Given a vector space V and a basis $\{v_1, \ldots, v_N\}$, any vector $v \in V$ can be expressed in a* unique *way as a linear combination of $v_1, \ldots, v_N$.*

Proof. Since a basis is also a GS, there exist $\alpha_1, \ldots, \alpha_N$ real numbers such that

$$v = \sum_{i=1}^{N} \alpha_i v_i.$$

Assume now that also exist $\beta_1, \ldots, \beta_N$ real numbers such that $v = \sum_{i=1}^{N} \beta_i v_i$, then

$$0 = \sum_{i=1}^{N} \alpha_i v_i - \sum_{i=1}^{N} \beta_i v_i = \sum_{i=1}^{N} (\alpha_i - \beta_i) v_i.$$

Since $v_1, \ldots, v_N$ are linearly independent, $\alpha_i - \beta_i = 0$ for all i and the uniqueness is proven. $\square$

Consider now two vector spaces V, W.

Definition. *A map $f : V \to W$ is **linear** if*

i) $\forall v_1, v_2 \in V$, it holds that $f(v_1 + v_2) = f(v_1) + f(v_2)$,

ii) $\forall \lambda \in \mathbb{R}$, it holds that $f(\lambda v) = \lambda f(v)$.

The *set* of the linear transformations from V to W is denoted by $\mathcal{L}(V, W)$ and a linear transformation from V into V itself is called **endomorphism**.

Examples.

i) $id(v) = v$ is a linear transformation.

ii) Consider $f : \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$f(x, y, z) = (2x + 2y, z, x - z).$$

It is immediate to verify that it is linear. For example:

$$\begin{aligned} f(\lambda x, \lambda y, \lambda z) &= (2\lambda x + 2\lambda y, \lambda z, \lambda x - \lambda z) \\ &= \lambda(2x + 2y, z, x - z) \\ &= \lambda f(x, y, z). \end{aligned}$$

There is a link between linear maps and matrices. In more details, given a linear map $f : V \to W$ **and** a pair of bases, of V and W, respectively, then f is uniquely identified by a matrix (say A_f).

For instance, in the previous example, we can express $f(x, y, z)$ as

$$f(x, y, z) = A_f \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

where

$$A_f = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix},$$

where the **columns** of A_f, denoted A_f^1, A_f^2 and A_f^3 are obtained as images (via $f(\cdot)$) of the vectors of the canonical basis. Thus

$$\begin{aligned} A_f^1 &= [f(1,0,0)]^T = (2,0,1)^T \\ A_f^2 &= [f(0,1,0)]^T = (2,0,0)^T \\ A_f^3 &= [f(0,0,1)]^T = (0,1,-1)^T \end{aligned}$$

iii) Given $f, g \in \mathcal{L}(V, W)$ we define the operations of sum and product by scalar as

$$\begin{aligned} (f+g)(v) &:= f(v) + g(v), \qquad v \in V \\ (\lambda f)(v) &:= \lambda f(v) \qquad \lambda \in \mathbb{R} \end{aligned} \tag{1}$$

With these definitions, the maps $f + g$ and λf are still linear (exercise).

1.2 Kernel and image

Definition. *Given $f \in \mathcal{L}(V, W)$, the* ***kernel*** *of f is the the subset*

$$Ker(f) := \{v \in V | f(v) = 0_W\} \subseteq V,$$

whereas the ***image*** *of f is the subset*

$$Im(f) := \{w \in W | w = f(v), \exists v \in V\}.$$

Proposition 2. *With the above notations, it holds that:*

i) $Ker(f)$ is a vector space in V,

ii) $Im(f)$ is a vector space in W,

iii) f is <u>injective</u> if and only if $Ker(f) = O_V$,

iv) f is <u>surjective</u> if and only if $Im(f) = W$.

Proof. $(*)$ In the order:

i) Given $v_1, v_2 \in Ker(f)$ it holds that

$$f(v_1 + v_2) = f(v_1) + f(v_2) = 0_W + 0_W = 0_W.$$

Thus $v_1 + v_2 \in Ker(f)$. Similarly that, if $v \in Ker(f)$, then $\lambda v \in Ker(f)$.

ii) Exercise.

iii) Assume f is injective: if $f(v_1) = f(v_2) \Rightarrow v_1 = v_2$. If $v \in Ker(f)$, then $f(v) = O_w = f(0_V)$, where the last equality comes from linearity. By injectivity, $v = 0_V$. Conversely, assume $Ker(f) = 0_V$. Then

$$f(v_1) = f(v_2) \Rightarrow 0_W = f(v_1) - f(v_2) = f(v_1 - v_2)$$

by linearity and thus $v_1 - v_2 \in Ker(f)$. But, since $Ker(f)$ only contains O_V, $v_1 = v_2$.

iv) Trivial.

□

Proposition 3. *Given a basis $\{v_1, \dots, v_N\}$ of V and a linear map $f : V \to W$, then $\{f(v_1), \dots, f(v_N)\}$ is a generating system for $Im(f)$.*

Proof. $(*)$ Given $w \in Im(f)$ it holds

$$w = f(v) = f\left(\sum_{i=1}^{N} \alpha_i v_i\right) = \sum_{i=1}^{N} \alpha_i f(v_i),$$

for some $\alpha_1, \dots, \alpha_N$. □

Definition. *Given a linear transformation $f \in \mathcal{L}(V, W)$, we call **rank** of f (a.k.a. $rk(f)$) the dimension of $Im(f)$.*

Theorem 1. *Given a vector space V of dimension N and $f \in \mathcal{L}(V, W)$, then*

$$N = \dim (ker(f)) + rk(f). \tag{2}$$

Proof. Given a basis $\{u_1, \dots, u_r\}$ of $Ker(f)$, it can be augmented to a basis of V, say $\{u_1, \dots, u_r, v_1, \dots, v_{N-r}\}$ for some $v_1, \dots, v_{N-r}$ (admitted). Since $f(u_i) = 0$ for all $i \le r$, by Proposition 3 we know that $f(v_1), \dots, f(v_{N-r})$ is a generating system for $Im(f)$. Now

$$\begin{aligned} \sum_{i=1}^{N-r} \alpha_i f(v_i) = 0 \quad &\iff \quad f\left(\sum_{i=1}^{N-r} \alpha_i v_i\right) = 0 \\ &\iff \quad \sum_{i=1}^{N-r} \alpha_i v_i \in Ker(f) \\ &\iff \quad \sum_{i=1}^{N-r} \alpha_i v_i = \sum_{j=1}^{r} \beta_j u_j, \end{aligned}$$

for some real parameters $\beta_1, \dots, \beta_r$. The last equality reads

$$\alpha_1 v_1 + \cdots + \alpha_{N-r} v_{N-r} - \beta_1 u_1 - \beta_r u_r = 0$$

and, since $\{u_1, \dots, u_r, v_1, \dots, v_{N-r}\}$ is a basis, then

$$\alpha_1 = \cdots = \alpha_{N-r} = \beta_1 = \cdots = \beta_r = 0$$

and $v_1, \dots, v_{N-r}$ is a basis for $Im(f)$. $\square$

Exercise. Given the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$

$$f(x, y, z) = f(x + y, y + z, x + 2y + z)$$

find a basis for both $Ker(f)$ and $Im(f)$.

Solution. First, we identify the matrix A_f associated with f (according to the canonical basis). Since $f(1,0,0) = (1,0,1)$, $f(0,1,0) = (1,1,2)$ and $f(0,0,1) = (0,1,1)$, then

$$A_f = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

By definition of kernel we need to look for (x, y, z) whose image (via f) is null. It boils down to solve the following system

$$\begin{cases} x + y = 0 \\ y + z = 0 \\ x + 2y + z = 0 \end{cases}$$

whose solution is $(x, y, z) = (-y, y, -y) = y(-1, 1, -1)$, for all $y \in \mathbb{R}$. Thus, $\{(-1, 1, -1)\}$ is a generating system for $Ker(f)$ and since $y(-1, 1, -1) = 0$ iff $y = 0$ it is also linearly independent. Then, $\{(-1, 1, -1)\}$ is a basis of $ker(f)$.

Since the basis only contains *one* vector, $\dim(Ker(f)) = 1$ and by Theorem 1 we know that $rk(f) = 3 - 1 = 2$. Therefore, a basis of $Im(f)$ contains *two* vectors. How to find them? We know that $Im(f)$ contains the points of $\mathbb{R}^3$ that can be written as

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} y + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} z.$$

Thus the columns of A_f (let us call them A_f^1, A_f^2, A_f^3) form a generating system for $Im(f)$. Since its dimension is two, we know that these three (column) vectors must be linearly dependent. We only need to select two of them being linearly *independent* in order to have a basis of $Im(f)$. Let us consider A_f^1 and A_f^2. The equation $\alpha_1 A_f^1 + \alpha_2 A_f^2 = 0$ reduces to

$$\begin{cases} \alpha_1 + \alpha_2 = 0 \\ \alpha_2 = 0 \\ \alpha_1 + 2\alpha_2 = 0 \end{cases}$$

whose solution clearly is $\alpha_1 = \alpha_2 = 0$. Thence $\{(1, 0, 1), (1, 1, 2)\}$ is a basis of $Im(f)$.

Exercise. Given the linear map $f : \mathbb{R}^3 \to \mathbb{R}^2$

$$f(x, y, z) = (x, y + z)$$

find a basis for both $Ker(f)$ and $Im(f)$.

1.3 Isomorphisms and matrix inversion

The set of the linear maps from V to W, $\mathcal{L}(V, W)$ equipped with the operations of sum and product by scalar in Eq. (1) is itself a vector space. Indeed, given $f, g \in \mathcal{L}(V, W)$ and $v_1, v_2 \in V$ it holds that

$$\begin{aligned}(f + g)(v_1 + v_2) &= f(v_1 + v_2) + g(v_1 + v_2)\\ &= f(v_1) + g(v_1) + f(v_2) + g(v_2)\\ &= (f + g)(v_1) + (f + g)(v_2).\end{aligned}$$

In words, if f and g are linear maps (from V to W), and thence additive, then $f + g$ is still additive. Moreover, for any $\delta \in \mathbb{R}$

$$\begin{aligned}(f + g)(\delta v) &= f(\delta v) + g(\delta v)\\ &= \delta(f(v) + g(v)).\\ &= (\delta(f + g))(v)\end{aligned}$$

In conclusion, if f, g are linear maps, $f + g$ is *still* a linear map and thus belongs to $\mathcal{L}(V, W)$.

Similarly, it can be shown that, if $f \in \mathcal{L}(V, W)$, then, for any $\lambda \in \mathbb{R}$, λf is still linear.

Definition. *When $W = \mathbb{R}$, $\mathcal{L}(V, \mathbb{R})$ is called **dual** space.*

Definition. *Consider two linear maps $g : V \to W$ and $f : W \to U$, where V, W and U are three vector spaces. The **composition** $f \bullet g : V \to U$ is defined as*

$$(f \bullet g)(v) := f(g(v)).$$

Proposition 4. *$f \bullet g$ is a linear map.*

Proof. $(*)$ Exercise. □

We introduce the **inverse** of a linear map according to the following definition:

Definition. *$f \in \mathcal{L}(V, W)$ is invertible if it exists $g : W \to V$ such that*

$$g \bullet f = id_V \qquad\qquad f \bullet g = id_W$$

. The map g (often denoted f^{-1}) is the inverse of f.

First, notice that if the inverse exists, it is unique. Indeed, let's say g, g' are two inverse of f. Then

$$g' = id_V \bullet g' = (g \bullet f) \bullet g' = g \bullet (f \bullet g') = g \bullet id_W = g.$$

Second, recall that a function (not necessary linear!) is invertible if and only if it is injective and surjective. Thus:

1. $f \in \mathcal{L}(V, W)$ is injective if and only if $Ker(f) = \{0_V\}$ (*cfr.* Proposition 2),

2. f is surjective if and only if $rk(f) = \dim(W)$ (*cfr.* Proposition 2), thence

3. by Theorem 1, if f is invertible, then

$$\dim(V) = 0 + rk(f) = \dim(W).$$

   Conversely, if we know that $\dim(V) = \dim(W)$, then f is invertible if and only if $Ker(f) = \{0_V\}$.

Proposition 5. *The inverse of a linear map is linear.*

Proof. With the same notation used so far, consider $w_1, w_2 \in W$:

$$f(g(w_1 + w_2)) = w_1 + w_2 = f(g(w_1)) + f(g(w_2)) = f(g(w_1) + g(w_2)),$$

where used $f \bullet g = id_W$ and the linearity of f. Since f is injective

$$g(w_1 + w_2) = g(w_1) + g(w_2).$$

Similarly, we see that $g(\lambda w) = \lambda g(w)$, $\lambda \in \mathbb{R}, w \in W$. $\square$

We saw that any linear map $f : V \to W$ is linked with a matrix A_f. If f is invertible it makes sense to look for the inverse f^{-1} by inspecting the associated matrix $A_{f^{-1}}$ which is the inverse of A_f. In formulas

$$A_{f^{-1}} = (A_f)^{-1}.$$

Exercise. Consider the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$

$$f(x, y, z) = (x + y + 2z, x - z, 3y).$$

Determine if f is invertible and, in case it is, find the inverse map f^{-1}.

Solution. The matrix A_f is

$$A_f = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & 3 & 0 \end{pmatrix}.$$

Since here $\dim(V) = \dim(W) = 3$ we know that f is invertible if and only if $Ker(f) = 0$. The linear system

$$\begin{cases} x + y + 2z = 0 \\ x - z = 0 \\ 3y = 0 \end{cases}$$

clearly admits the unique solution $(x, y, z) = (0, 0, 0)$, thence $Ker(f) = 0_{\mathbb{R}^3}$ and f is invertible.

Recall. *Another common way to asses whether a square matrix A is invertible or not it to look at its **determinant**[1]. Indeed A is invertible if and only if $|A| \neq 0$.*

In this case

$$|A_f| = -3 \begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3(-1-2) = 9 \neq 0.$$

Since $(A_f)^{-1}$ is

$$(A_f)^{-1} = \begin{pmatrix} 1/3 & 2/3 & -1/9 \\ 0 & 0 & 1/3 \\ 1/3 & -1/3 & -1/9 \end{pmatrix}.$$

the inverse linear map f^{-1} is[2]

$$f^{-1}(x, y, z) = (A_f)^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \left(\frac{x}{3} + \frac{2y}{3} - \frac{z}{9}, \frac{z}{3}, \frac{x}{3} - \frac{y}{3} - \frac{z}{9} \right).$$

[1] `https://en.wikipedia.org/wiki/Determinant`

[2] Any decent statistical software can compute the determinant of a matrix A and the inverse matrix A^{-1} (if it exists). In R, for instance, the instructions det(A) and solve(A) do the job.

2 Scalar products

2.1 Definition and properties

Given a vector space V we consider a map $g : V \times V \to \mathbb{R}$. The map g is **bilinear** if it is linear in both variables

$$\begin{array}{rll} i) & g(v_1 + v_2, w) = g(v_1, w) + g(v_2, w) & g(\lambda v, w) = \lambda g(v, w) \\ ii) & g(v, w_1 + w_2) = g(v, w_1) + g(v, w_2) & g(v, \lambda w) = \lambda g(v, w) \end{array}$$

with $v, w, v_1, v_2, w_1, w_2 \in V$ and $\lambda \in \mathbb{R}$.

Definition. *A **scalar (or dot) product** on V is any bilinear and symmetric map (i.e. $g(v, w) = g(w, v)$, for all $v, w \in V$). By convention a scalar product is denoted by $\langle \cdot, \cdot \rangle$.*

Examples.

i) The standard dot product in $\mathbb{R}^N$:

$$\langle v, w \rangle := w^T v = \sum_{i=1}^{N} v_i w_i, \quad v, w \in \mathbb{R}^N.$$

It is clearly symmetric and hence bilinear since

$$\langle v_1 + v_2, w \rangle = \sum_{i=1}^{N} (v_{1i} + v_{2i}) w_i = \sum_{i=1}^{N} v_{1i} w_i + \sum_{i=1}^{N} v_{2i} w_i = \langle v_1, w \rangle + \langle v_2, w \rangle$$

and

$$\langle \lambda v, w \rangle = \sum_{i=1}^{N} \lambda v_i w_i = \lambda \sum_{i=1}^{N} v_i w_i = \lambda \langle v, w \rangle.$$

ii) The map $\langle \cdot, \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$

$$\langle v, w \rangle := v_1 w_1 + v_1 w_2 + v_2 w_1 + v_2 w_2 + v_3 w_3$$

is clearly symmetric, indeed

$$\langle w, v \rangle = w_1 v_1 + w_1 v_2 + w_2 v_1 + w_2 v_2 + w_3 v_3.$$

It is also bilinear (exercise) and so it is a scalar product on $\mathbb{R}^3$.

Here we only focus on **not degenerated** scalar products, i.e.

$$\forall v \in V, \langle v, w \rangle = 0 \Rightarrow w = O_V.$$

Definition. *A dot product* $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ *is*

i) positive definite: $\langle v, v \rangle > 0 \quad \forall v \in V$,

ii) negative definite: $\langle v, v \rangle < 0 \quad \forall v \in V$,

iii) positive semi-definite $\langle v, v \rangle \geq 0 \quad \forall v \in V$ *with equality for some* $v \neq 0_V$,

iv) negative semi-definite $\langle v, v \rangle \leq 0 \quad \forall v \in V$ *with equality for some* $v \neq 0_V$,

Definition. *A* ***positive definite*** *dot product on* V *induces a* ***norm*** $\| \cdot \|$: $V \to \mathbb{R}^+$ *defined as*

$$\| v \| = \sqrt{\langle v, v \rangle}. \tag{3}$$

For instance, the standard dot product on $\mathbb{R}^N$ induces the well known Euclidean norm

$$\| v \| = \left(\sum_{i=1}^{N} v_i^2 \right)^{1/2},$$

whereas the second dot product considered in the example above is *not* positive definite. Indeed:

$$\langle v, v \rangle = (v_1 + v_2)^2 + v_3^2 \geq 0$$

with the equality verified for all $v = (x, -x, 0)$ with $x \in \mathbb{R}$.

Proposition 6. *Consider a positive definite scalar product* $\langle \cdot, \cdot \rangle$ *on* V *and its norm* $\| \cdot \|$. *Then for all* $\lambda \in \mathbb{R}$ *and* $v, w \in V$

i) $\| v \| \geq 0$ *and* $\| v \| = 0$ *if and only if* $v = 0_V$,

ii) $\| \lambda V \| = |\lambda| \, \| v \|$ *and* $\| v + w \|^2 = \| v \|^2 + \| w \|^2 + 2 \langle v, w \rangle$,

iii) $| \langle v, w \rangle | \leq \| v \| \| w \|$, (***Cauchy-Swartz*** *inequality*)

iv) $| \, \| v \| - \| w \| \, | \leq \| v - w \| \leq \| v \| + \| w \|$, (***Triangle*** *inequality*)

v) $\langle v, w \rangle = \frac{1}{4} [\| v + w \|^2 - \| v - w \|^2]$.

Proof. $(*)$

i) This trivially comes form the definition of positive definiteness of $\langle \cdot, \cdot \rangle$.

ii) It is a consequence of the definition of norm, in Eq. (3).

iii) The inequality turns into an equality in the trivial case where either v or w is null. Thus, assume $v, w \neq 0_v$. For all $a, b \in \mathbb{R}$, from point i) and ii) we know that

$$0 \leq\| av + bw \|^2 = a^2 \| v \|^2 + b^2 \| w \|^2 + 2ab \langle v, w \rangle .$$

In particular, for $a = \| w \|^2$ and $b = -\langle v, w \rangle$ the above inequality reduces to

$$0 \leq \| w \|^2 \left(\| w \|^2 \| v \|^2 - (\langle v, w \rangle)^2 \right)$$

and the sentence is proven.

Notice also that for these choices of a and b, the equality holds if and only if

$$\| w \|^2 v - \langle v, w \rangle w = 0_V \qquad \Longleftrightarrow \qquad v = \frac{\langle v, w \rangle}{\| w \|^2} w$$

In this case the vectors v, w are linearly dependent (or aligned).

iv) Thanks to Cauchy-Swartz we know that

$$\begin{aligned}
(\| v \| - \| w \|)^2 &= \| v \|^2 + \| w \|^2 - 2 \| v \| \| w \| \\
&\leq \| v \|^2 + \| w \|^2 + 2 \langle v, w \rangle = \| v + w \|^2 \\
&\leq \| v \|^2 + \| w \|^2 + 2 \| v \| \| w \| = (\| v \| + \| w \|)^2 .
\end{aligned}$$

Taking the square root of the blue terms proves iv).

v) It suffices to remark that

$$\begin{aligned}
\| v + w \|^2 &= \| v \|^2 + \| w \|^2 + 2 \langle v, w \rangle \\
\| v - w \|^2 &= \| v \|^2 + \| w \|^2 - 2 \langle v, w \rangle
\end{aligned}$$

and subtract.

$\square$

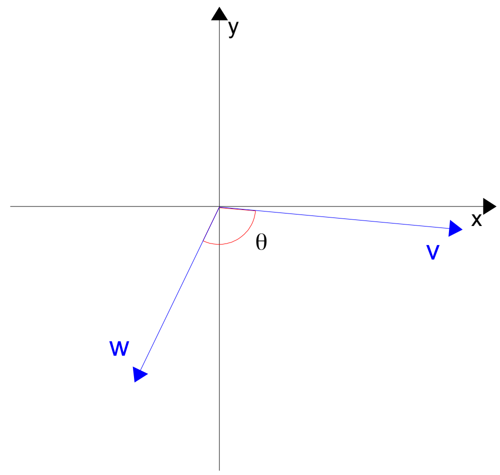

Figure 1: The angle θ betwen two vectors in $\mathbb{R}^2$

2.2 Orthogonality

Definition. *Given a vector space V and two vectors $v, w \in V$, the* ***angle*** *between them is the real number $\theta \in [0, \pi]$ defined by*

$$\cos\theta := \frac{\langle v, w\rangle}{\| v \| \| w \|} \tag{4}$$

Remarks.

1. The cosine function is bijective on $[0, \pi]$, so the angle θ is well defined. This definition always correspond to the *acute* angle between two vectors (not the *obtuse* one).

2. Cauchy-Swartz inequality guarantees that the above definition always provides us with a cosine in $[-1, 1]$.

3. We saw that the CS inequality turns into an equality where two vectors are linearly dependent (i.e. aligned): in that case the above definition provides us with $\cos\theta = 1$ and hence $\theta = 0$. We also know from

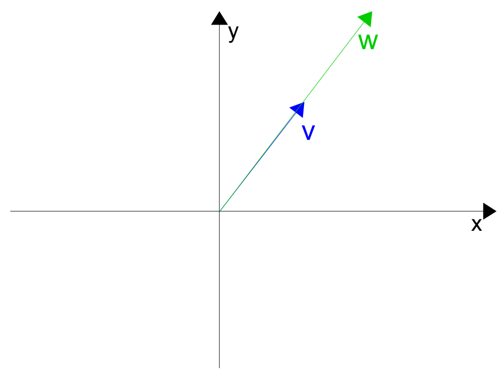

Figure 2: Two linearly dependent or aligned vectors.

trigonometry that when $\theta = \pi$, $\cos\theta = 0$, which is aligned with the above definition if and only if $\langle v, w\rangle = 0$. This motivates the following

Definition. *Two vectors v, w in a vector space V, equipped with a dot product $\langle\cdot,\cdot\rangle$ are said **orthogonal** $(v \perp w)$ if $\langle v, w\rangle = 0$. Given a subspace $S \subset V$ we say $v \perp S$ if $\langle v, s\rangle = 0$ for all $s \in S$.*

Now consider a vector space of dimension 2 (for instance $\mathbb{R}^2$ as in Figure 1), for simplicity. If two vectors v, w are linearly dependent, then the equation

$$\alpha_1 v + \alpha_2 w = 0$$

is satisfied for some α_1 or α_2 different from zero. Assume wlog that $\alpha_1 \neq 0$. Then

$$v = -\frac{\alpha_2}{\alpha_1} w,$$

meaning that v is a multiple of w and the two vectors are aligned (as in Figure 2).

Since $\mathbb{R}^2$ has dimension 2 it suffices to chose *any* two vectors not aligned (namely $\theta \neq 0$, as in Figure 1) to have a basis. If in particular we chose $v \perp w$ $(\theta = \pi)$ we have an **orthogonal basis**. This intuition is formalized in the following

Proposition 7. *Given a vector space V of dimension N and $v_1, \ldots, v_N$ non-null vectors, pairwise orthogonal $(v_i \perp v_j \quad \forall j \neq i)$, they form a basis for V.*

Proof. $(*)$ Assume

$$\alpha_1 v_1 + \cdots + \alpha_N v_N = 0_V.$$

Then for all $i \leq N$

$$0 = \langle 0_V, v_i \rangle = \left\langle \sum_{j=1}^{N} \alpha_j v_j, v_i \right\rangle = \sum_{j=1}^{N} \alpha_j \langle v_j, v_i \rangle = \alpha_i \parallel v_i \parallel^2$$

and the above equality is clearly satisfied if and only if $\alpha_i \neq 0$ $(v_i \neq 0)$. Thence the N vectors are linearly independent and form a basis. □

An orthogonal basis $\{v_1, \ldots, v_N\}$ is formed by pairwise orthogonal vectors. If, moreover, $\parallel v_i \parallel = 1$, for all i the basis is **orthonormal**.

Example. In $\mathbb{R}^2$, $\{(3,0),(0,3)\}$ forms an orthogonal basis, however

$$\parallel (3,0) \parallel = \sqrt{0^2 + 3^2} = 3 \neq 1.$$

By multiplying each vector by the inverse of its norm (here $1/3$) we get $\{(1,0),(0,1)\}$ which is orthonormal.

Theorem 2. *If $\{v_1, \ldots, v_N\}$ is an orthogonal basis of V, then, for all $v \in V$*

$$v = \frac{\langle v, v_1 \rangle}{\parallel v_1 \parallel^2} v_1 + \cdots + \frac{\langle v, v_N \rangle}{\parallel v_N \parallel^2} v_N. \tag{5}$$

Proof. $(*)$ We know that

$$v = \alpha_1 v_1 + \cdots + \alpha_N v_N, \qquad \exists \quad \alpha_1, \ldots, \alpha_N \in \mathbb{R}$$

For all $i \in \{1, \ldots, N\}$ it holds that

$$\langle v_i, v \rangle = \left\langle v_i, \sum_{j=1}^{N} \alpha_j v_j \right\rangle = \sum_{j=1}^{N} \alpha_j \langle v_i, v_j \rangle = \alpha_i \parallel v_i \parallel^2$$

and thence

$$\alpha_i = \frac{\langle v, v_i \rangle}{\parallel v_i \parallel^2}.$$

□

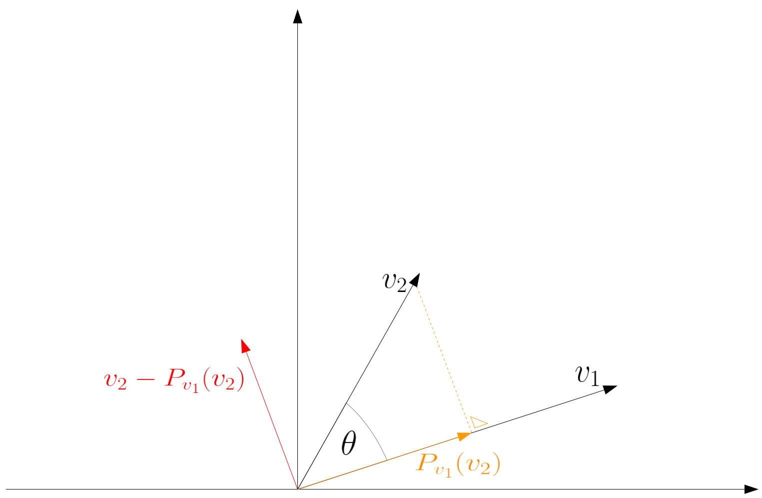

Figure 3: An illustration of the Gram-Schmidt procedure to obtain an orthogonal basis in the Euclidean plane.

Remark. Notice that, if $\{v_1, \ldots, v_N\}$ is orthonormal, then Eq. (5) reduces to

$$v = \sum_{i=1}^{N} \langle v, v_i \rangle \, v_i.$$

It can be proven than in a vector space V of dimension N, an orthogonal (and hence orthonormal) basis can always be obtained from any basis $\{v_1, \ldots, v_N\}$. This result is now informally motivated in $\mathbb{R}^2$. Consider two vectors v_1, v_2 in the Euclidean plane, as in Figure 3. Since v_1 and v_2 are not aligned the form a basis in $\mathbb{R}^2$. An orthogonal basis can be obtained as follows:

i) Consider the **orthogonal projection** of v_2 on v_2, denoted $P_{v_1}(v_2)$ (the orange vector in Figure 3). It is aligned to v_1, thus

$$P_{v_1}(v_2) = \alpha v_1, \qquad \exists \alpha \in \mathbb{R}^+.$$

Moreover, by <u>trigonometry</u> we know that

$$\| \, P_{v_1}(v_2) \, \| \; (= \alpha \, \| \, v_1 \, \|) = \| \, v_2 \, \| \cos\theta,$$

thence

$$\alpha = \frac{\| v_2 \|}{\| v_1 \|} \cos\theta = \frac{\cancel{\| v_2 \|} \langle v_1, v_2 \rangle}{\| v_1 \|^2 \cancel{\| v_2 \|}}$$

and finally

$$P_{v_1}(v_2) = \frac{\langle v_1, v_2 \rangle}{\| v_1 \|^2} v_1.$$

ii) Then we define

$$w_2 := v_2 - P_{v_1}(v_2) = v_2 - \frac{\langle v_1, v_2 \rangle}{\| v_1 \|^2} v_1,$$

which is the red vector in Figure 3 and $\{v_1, w_2\}$ forms an orthogonal basis in $\mathbb{R}^2$.

The procedure outlined so far (a.k.a. Gram-Schmidt) can be generalized to vector spaces of higher dimension.

Proposition 8. *Given a vector space V, equipped with a positive definite dot product $\langle \cdot, \cdot \rangle$ and sub vector space $U \subset V$, for all $v_0 \in V$, $\exists! u_0 \in U$ such that*

$$v_0 - u_0 \perp U \quad (\textit{namely } \langle v_0 - u_0, u \rangle = 0, \quad \forall u \in U).$$

Proof. Assuming that $\{u_1, \dots u_R\}$ is an *orthonormal* basis for U, with $R \leq N$, notice that if such a u_0 exists, then, for all $j \in \{1, \dots, R\}$ it must fulfil

$$\begin{aligned} 0 = \langle v_0 - u_0, u_j \rangle &= \langle v_0, u_j \rangle - \langle v_0, u_j \rangle \\ &= \langle v_0, u_j \rangle - \sum_{i=1}^{R} \alpha_i \langle u_i, u_j \rangle \\ &= \langle v_0, u_j \rangle - \alpha_j, \end{aligned} \tag{6}$$

thus $\alpha_j = \langle u_0, v_0 \rangle$ and $u_0 = \sum_{i=1}^{R} \langle v_0, u_i \rangle u_i$[3]. Thus, it suffices to define

[3] An alternative proof would consist into observing that

$$u_0 = \sum_{i=1}^{R} \langle u_0, u_i \rangle u_i$$

by previous propositions. Then by the first row of Eq. (6) it follows that

$$\langle v_0, u_j \rangle = \langle u_0, u_j \rangle$$

for all j.

$u_0 = \sum_{i=1}^{R} \langle v_0, u_i \rangle u_i$ and it is easy to show that $v_0 - u_0 \perp U$. Finally, the uniqueness comes from the positive definiteness of the dot product. $\square$

The above considerations in a vector space of dimension 2 motivate us to a more general

Definition. *Given a sub vector space $U \subset V$ of dimension R and an orthonormal basis $\{u_1, \ldots, u_R\}$ of U we define the* ***orthogonal projection*** *of V on U, $P_U : V \to U$ as*

$$P_U(v) = \sum_{i=1}^{R} \langle v, u_i \rangle u_i, \qquad \forall v \in V. \tag{7}$$

Two immediate remarks follow

i) the inclusion $Im(P_U) \subset U$ is trivial. Moreover, if $u \in U$, then

$$P_U(u) = \sum_{i=1}^{R} \langle u, u_i \rangle u_i = u,$$

where the last equality comes from Theorem 2. Then $Im(P_U) = U$ and the map is surjective.

ii) The subset

$$U^{\perp} := \{v \in V | \langle v, u \rangle = 0, \qquad \forall u \in U\}$$

can immediately be shown to be a vector sub space of V (exercise). Moreover $U^{\perp} = Ker(P_U)$, indeed, for all $v \in V$

$$0 = P_U(v) = \sum_{i=1}^{R} \langle v, u_i \rangle u_i \quad \Leftrightarrow \quad \langle v, u_i \rangle = 0$$

for all $i \in \{1, \ldots, R\}$. Thus $v \perp U$.

Proposition 9. *if $U \subset V$ are two vector spaces, then*

i) $V = U \oplus U^{\perp}$, namely for all $v \in V$, it holds that

$$v = u + u^{\perp} \qquad \exists u \in U, \quad \exists u^{\perp} \in U^{\perp}$$

and $U \cap U^{\perp} = \{0_U\}$,

ii) $\dim(V) = \dim(U) + \dim(U^\perp)$,

iii) $\| v \|^2 = \| u \|^2 + \| u^\perp \|^2$

Proof. $(*)$

i) Given the positive definiteness of $\langle \cdot, \cdot \rangle$, $\langle u, u \rangle = 0$ if and only if $u = 0_U$, thus $U \cap U^\perp = 0_U$. Moreover, for all $v \in V$

$$v = \underbrace{P_U(v)}_{\in U} + \underbrace{v - P_U(v)}_{\in U^\perp}.$$

ii) It follows immediately from Theorem 1.

iii) We know that

$$\begin{aligned} \| v \|^2 &= \| u + u^\perp \|^2 \\ &= \| u \|^2 + \| u^\perp \|^2 + 2 \left\langle u, u^\perp \right\rangle \\ &= \| u \|^2 + \| u^\perp \|^2 . \end{aligned}$$

Exercise. Prove that $v - P_U(v) \in U^\perp$.

Exercise. Prove that, given $v \in V$ and $u \in U \subset V$, then

$$P_U(v) = \arg\min_{u \in U} \| v - u \|^2 .$$

In other words, prove that the orthogonal projection of v on U is the nearest point of U to v (*Hint: use that* $v - u = v - P_U(v) + P_U(v) - u$).

□

2.3 Isometries

Definition. *Given a vector space V, a linear map $f : V \Rightarrow V$ is called an* ***isometry*** *if*

$$\langle v, w \rangle = \langle f(v), f(w) \rangle , \qquad \forall v, w \in V. \tag{8}$$

Notice that, since an isometry preserves the scalar product (and thus the *angle*) between two vectors, it also preserves the *distance* between them. Indeed

$$\| v \|^2 = \langle v, v \rangle = \langle f(v), f(v) \rangle = \| f(v) \|^2$$

and therefore $\| v - w \| = \| f(v) - f(w) \|^2$.

Exercise. Prove that an isometry is an invertible (or equivalently, a bijective) map.

We saw in previous sections that a linear map between vector spaces is (uniquely) represented by a matrix. On the opposite side, the matrix product with vectors induces a linear map between vector spaces. In the following, we focus on the existing relation between **orthogonal** matrices and isometries.

Consider a matrix $A \in \mathbb{R}^{N\times N}$. It is called **orthogonal** if $A^T A = I_N$, where A^T denotes the *transposed*[4] of A. Notice that

i) $A^T A = I_N = (I_N)^T = (A^T A)^T$, thence $A^T A$ is symmetric.

ii) $A^{-1} = A^T$, since the inverse matrix (if it exists, is unique).

If we denote $f_A : \mathbb{R}^N \to \mathbb{R}^N$ the linear map[5] associated with A, such that

$$f_A(v) := Av, \qquad v \in \mathbb{R}^N,$$

then f is invertible (since A is) and thus $Im(f_A) = \mathbb{R}^N$ and $Ker(f_A) = 0_{\mathbb{R}^N}$. Now, since

$$Im(f_A) = \{Av = A^1 v_1 + \cdots + A^N v_N | v \in V\},$$

where A^j denotes the j-th column of A, the columns of A form a generating system of $Im(f_A) = \mathbb{R}^N$. Moreover, since $rk(A) = N$, then this generating system is a basis. But there is something more:

Proposition 10. *The columns of A form an orthonormal basis of $\mathbb{R}^N$ and f_A is an isometry (with respect to the* standard *dot product).*

Proof. Consider the canonical basis of $\mathbb{R}^N$, $\{e_1, \dots, e_N\}$, where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad e_N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

[4] `https://en.wikipedia.org/wiki/Transpose`.

[5] Here we assume that f_A maps $\mathbb{R}^N$ into itself for simplicity, but f_A might be an endomorphism mapping into itself a more general vector space V of dimension N.

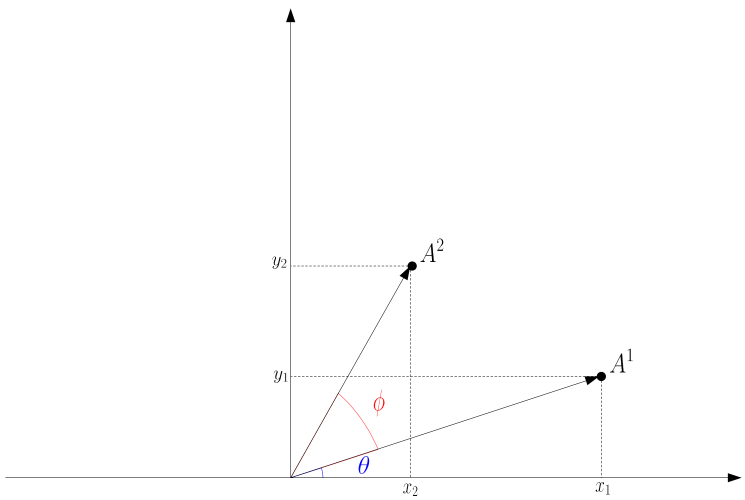

Figure 4: Matrix A columns viewed as points in the Euclidean plane, in both Cartesian and polar coordinates.

Since $Ae_i = A^i$, then

$$\left\langle A^i, A^j \right\rangle = \langle Ae_i, Ae_j \rangle = (Ae_i)^T Ae_j = e_i^T A^T Ae_j = e_i^T e_j = \langle e_i, e_j \rangle ,$$

where the last scalar product is clearly equal to 1 if $i = j$, zero otherwise, thus proving the orthonormality of the columns of A.

For the second part of the proposition, notice that

$$\langle v, w \rangle = v^T w = v^T I_N w = v^T A^T Aw = \langle Av, Aw \rangle = \langle f_A(v), f_A(w) \rangle .$$

$\square$

Example. A matrix $A \in \mathbb{R}^{2\times 2}$

$$A = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$$

induces a linear map on the Euclidean plane and its columns (A^1, A^2) identify two points on the plane (see Figure 4). We can rewrite them in polar

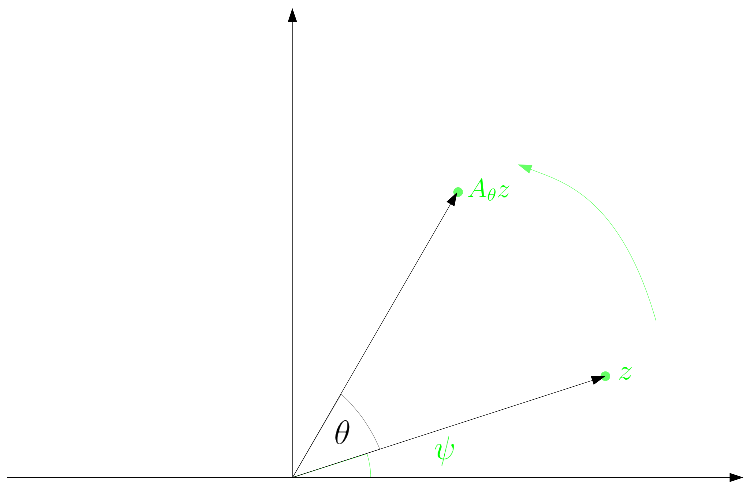

Figure 5: Rotation of point z by θ radians.

coordinates

$$A^1 = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, \qquad A^2 = \begin{pmatrix} s\cos\phi \\ s\sin\phi \end{pmatrix},$$

where $r = \| A^1 \|$ and $s = \| A^2 \|$. If we impose A to be orthogonal, namely

$$\underbrace{\begin{pmatrix} r\cos\theta & r\sin\theta \\ s\cos\phi & s\sin\phi \end{pmatrix}}_{A^T} \underbrace{\begin{pmatrix} r\cos\theta & s\cos\phi \\ r\sin\theta & s\sin\phi \end{pmatrix}}_{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

we find

$$\begin{cases} r = s = 1 \\ \cos\theta\cos\phi + \sin\theta\sin\phi = \cos(\theta - \phi) = 0 \end{cases}.$$

The second equation is satisfied either when $\phi = \theta + \frac{\pi}{2}$ or $\phi = \theta + \frac{3}{2}\pi$. Consider the first case $\phi = \theta + \frac{\pi}{2}$. In this case, the matrix A reduces to

$$A = A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Now, consider a new point $z = (z_1, z_2)$, whose polar coordinates are

$$z = \begin{pmatrix} \| z \| \cos\psi \\ \| z \| \sin\psi \end{pmatrix}.$$

Applying A_θ to z reduces to

$$A_\theta v = \begin{pmatrix} \| z \| (\cos\theta\cos\psi - \sin\theta\sin\psi) \\ \| z \| (\sin\theta\cos\psi + \cos\theta\sin\psi) \end{pmatrix} = \begin{pmatrix} \| z \| \cos(\theta+\psi) \\ \| z \| \sin(\theta+\psi) \end{pmatrix},$$

corresponding to a **rotation** of θ radians, as it can be seen in Figure 5. Similarly, the case $\phi = \theta + \frac{3}{2}\pi$ identifies a **symmetry** of z with respect to the line of parametric equations $(x, y) = t\left(\frac{\theta}{2}, \frac{\theta}{2}\right)$, with $t \in \mathbb{R}$. Thence a (linear) isometry in the Euclidean plane is either a rotation or a symmetry.

3 Eigenvalues and eigenvectors

3.1 Definition and symmetric matrices

Recall. *Given a matrix $A \in \mathbb{R}^{N\times M}$, its **transposed** A^T is such that $A^T_{ij} = A_{ji}$ and, given $B \in \mathbb{R}^{M\times L}$, then $(AB)^T = B^T A^T$.*

Recall. *The set of the* complex numbers *is denoted by $\mathbb{C}$. A complex $y = a+ib$, where $a, b \in \mathbb{R}$ and i is the imaginary unit, i.e. $i^2 = -1$. The conjugate of y is denoted by $\overline{y} = a - ib$ and note that $y\overline{y} = a^2 + b^2$. Moreover, $y = \overline{y}$ iff $b = 0$ and thence $y \in \mathbb{R}$. Also, it can easily be seen that if $x, y \in \mathbb{C}$, then $\overline{xy} = \overline{x}\,\overline{y}$.*

Given a square matrix $A \in \mathbb{R}^{N\times N}$ we look for a scalar $\lambda \in \mathbb{C}$ such that

$$Av = \lambda v, \qquad \exists v \neq O_N \tag{9}$$

If such a λ exists (a priori it is a complex number), it is called **eigenvalue** and all the vectors satisfying Eq. (9) are the corresponding **eigenvectors**.

Given an eigenvalues λ of A, if we define

$$V_\lambda = \{v | Av = \lambda v\} \tag{10}$$

it is easy to prove that V_λ is a sub vector space of $\mathbb{R}^N$ (exercise.) Moreover Eq. (9) is satisfied iff

$$(A - \lambda I_N)v = 0_N \tag{11}$$

Since $A - \lambda I_N$ is a square matrix of dimension N the above homogeneous system admits solutions other than $v = 0_N$ in and only if (why?)

$$\mathrm{rank}(A - \lambda I_N) < N,$$

namely if and only if

$$\det(A - \lambda I_N) = 0. \tag{12}$$

Exercise. Given

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{pmatrix}$$

compute the eigenvalues of A and the corresponding eigenvectors.

Proposition 11. *If $A \in \mathbb{R}^{N\times N}$ is symmetric, then its eigenvalues are real numbers.*

Proof. As the eigenvalues, also the eigenvectors are a priori complex. That said if λ is an eigenvalue of A and v the corresponding eigenvector, then

$$Av = \lambda v \Rightarrow \overline{Av} = \overline{\lambda v} \Rightarrow \overline{A}\overline{v} = \overline{\lambda}\overline{v} \Rightarrow A\overline{v} = \overline{\lambda}\overline{v},$$

where the last equality comes from the fact that A is real. By transposing

$$\overline{v}^T A = \overline{v}^T \overline{\lambda}.$$

Now

$$Av = \lambda v \Rightarrow \underbrace{\overline{v}^T A}_{\overline{\lambda}\overline{v}^T} v = \lambda \overline{v}^T v,$$

thus

$$\overline{\lambda}\overline{v}^T v = \lambda \overline{v}^T v \iff \overline{\lambda} = \lambda$$

where the last iff comes from $v \neq 0_N$. $\square$

Thus, if A is a real symmetric matrix all its eigenvectors are real numbers. Now, every eigenvector v has to be solution of the following linear system

$$(A - \lambda I_N)v = 0$$

not involving any complex number, thus we can assume that the eigenvectors of A are real[6]. Something more can be said for symmetric matrices.

Proposition 12. *If $A \in \mathbb{R}^{N\times N}$ is symmetric, the eigenvalues corresponding to distinct eigenvalues are orthogonal.*

Proof. Consider $\lambda_1 \neq \lambda_2$ eigenvalues of A and v_1, v_2 the corresponding eigenvectors, thus $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2$.

$$v_2^T A v_1 = (v_2^T A v_1)^T = v_1^T A^T v_2 = v_1^T A v_2, \tag{13}$$

by symmetry of A. Moreover

$$v_2^T A v_1 = \lambda_1 v_2^T y_1 \qquad \text{and} \qquad v_1^T A v_2 = \lambda_2 v_1^T v_2$$

$\square$

[6]I mean that we can always find real eigenvectors. Then if $v \in V_\lambda$ of course iv is still in V_λ...

Thus, by Eq. (13)

$$\lambda_1 \langle v_1, v_2 \rangle = \lambda_2 \langle v_2, v_1 \rangle$$

iff $v_2 \perp v_1$, since $\lambda_1 \neq \lambda_2$. The most important result about the spectral decomposition of real matrices is reported without proof.

Theorem 3. *(Spectral) If $A \in \mathbb{R}^{N \times N}$ is a square symmetric matrix, it is always possible to find N orthogonal eigenvectors of A.*

Proof. omitted. □

Example. Consider the following matrix

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

It can easily be seen (exercise) that its eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -1$ and two corresponding eigenvectors are $v_1 = (1,1)^T$ and $v_2 = (1,-1)^T$. Of course they are orthogonal (why?). Now we can re-size the eigenvectors in such a way that $\|v_1\| = \|v_2\| = 1$, it suffices to divide them by their norm $(\sqrt{2})$. Thus we introduce

$$w_1 = \frac{v_1}{\|v_1\|} = \qquad \text{and} \qquad w_2 = \frac{v_2}{\|v_2\|}.$$

They can be collected as columns into a matrix Q

$$Q = [w_1, w_2] \in \mathbb{R}^{2 \times 2}$$

and by definition of eigenvalues/eigenvector it holds that

$$AQ = [Aw_1, Aw_2] = \underbrace{[\lambda_1 w1, \lambda_2 w2] = Q\Lambda}_{\text{check it!}},$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Thus

$$AQ = Q\Lambda \Rightarrow A = Q\Lambda Q^{-1}.$$

Note that Q^{-1} exists since $w_1 \perp w_2$. Moreover, since w_1, w_2 have norm one $Q^{-1} = Q^T$ and therefore

$$A = Q\Lambda Q^T.$$

The above example motivates us to the following

Definition. *A matrix* $Q \in \mathbb{R}^{N\times N}$ *is said to be* ***orthogonal*** *if* $Q^{-1} = Q^T$.

Proposition 13. *If the columns of* $Q \in \mathbb{R}^{N\times N}$ *are* N *orthonormal vectors, then* Q *is orthogonal.*

Proof. It suffices to recall that the inverse matrix, if it exists is unique! Indeed

$$[QQ^T]_{ij} = \langle Q_i, (Q^T)^j\rangle = \langle Q_i, Q_j\rangle = \delta_{i=j}.$$

Thus

$$QQ^T = Q^TQ = I_N.$$

$\square$

3.2 Quadratic forms

Consider $f_A : \mathbb{R}^N \to \mathbb{R}$ such that

$$f_A(v) = v^T A v = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}v_iv_j,$$

where $A \in \mathbb{R}^N$ is a symmetric real matrix. Then, $f_A(\cdot)$ is said **quadratic form** and can be classified as follows:

1. $f_A(\cdot)$ is positive (negative) definite if $f_A(v) \geq 0$ (≤ 0), for all $v \in \mathbb{R}^N$ and $f_A(v) = 0$ iff $v = 0_N$.
2. $f_A(\cdot)$ is semi-positive (semi-negative) definite if $f_A(v) \geq 0$ (≤ 0), for all $v \in \mathbb{R}^N$ and $\exists v \neq 0_N$ such that $f_A(v) = 0$.
3. $f_A(\cdot)$ in not definite if none of the above.

The classification of $f_A(\cdot)$ also concerns A, so (e.g.) if $f_A(\cdot)$ is positive definite, the matrix A is said positive definite too.

Proposition 14. *$f_A(\cdot)$ is positive (negative) definite iff all its eigenvalues are strictly positive (negative). It is semi-positive (semi-negative) definite if all its eigenvalues are positive (negative) and some of them are null. It is not definite if some eigenvalues are positive and other negative.*

Sketch of proof. For any matrix A, real and symmetric it holds that

$$v^T A v = v^T Q \Lambda Q^T v,$$

for all $v \in \mathbb{R}^N$, where Q is the matrix whose columns are the eigenvectors (norm one) of A and Λ is the diagonal matrix whose non-null entries are the eigenvalues of A. If we substitute $z = Q^T v$ in the above equation we get

$$v^T A v = z^T \Lambda z = \sum_{i=1}^{N} \lambda_i z_i^2,$$

(why?) where λ_i are the eigenvalues of A.

Now if $\lambda_i > 0$, for all i, $\underline{\text{as } v \neq 0}$ also $z_i \neq 0, \exists i$. Thus the above quantity is positive and A is positive definite.

Viceversa, if $v^T A v > 0$ $\underline{\text{for all}}$ $v \neq 0$, then, for all i, it must be positive for that v such that $z = (0, \ldots, 0, \underbrace{1}_{\text{i-th position}}, 0, \ldots, 0)^T$. Thus

$$0 < v^T A v = \lambda_i \quad .$$

□

4 Elements of multivariate real analysis

4.1 Partial derivatives

We now focus on real functions $f : \mathbb{R}^N \to \mathbb{R}$ (not necessarily linear!). Often, the case $N = 2$ will be considered to simplify the exposition. Thus, if $f(x, y)$ is a function of two real variables and we keep y fixed, f can be read as a function of x. If it is derivable w.r.t. x, then we define the **partial derivative**

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

Similarly for y when x is fixed

$$\frac{\partial f}{\partial y}(x, y) = \lim_{k \to 0} \frac{f(x, y + k) - f(x, y)}{k}.$$

The partial derivatives $\frac{\partial f}{\partial x}(x, y)$ and $\frac{\partial f}{\partial y}(x, y)$ can be collected in a vector called gradient. In general, for a function $f : \mathbb{R}^N \to \mathbb{R}$ we call **gradient** the column vector ∇f collecting all the partial derivatives of f.

Exercise 1. Compute the gradient of $f(x, y) = \exp(x^2 y)$.

Partial derivatives are a special case of more general **directional derivatives**. A direction is a vector $v \in \mathbb{R}^N$ such that $\|v\| = 1$. If $x \in \mathbb{R}^N$, for a real parameter t, $x + tv$ describes a straight line through x aligned with v[7]. Then, for (x, v) the function

$$g(t) := f(x + tv)$$

is a real one! If it admits first derivative in 0, the derivative of f in the direction v is defined as

$$\frac{\partial f}{\partial v}(x) := g'(0) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \tag{14}$$

Exercise 2. Show that $g'(t) = \frac{\partial f}{\partial v}(x + tv)$.

[7]Recall: x and v are vectors in $\mathbb{R}^N$, whereas t is a real scalar.

Exercise 3. Compute the derivative of $f(x,y) = x^2y - e^{x+y}$ along the direction $v = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$.

Note that, e.g. in $\mathbb{R}^2$, when choosing $v = e_1 = (1,0)^T$ or $v = e_2 = (0,1)^T$ we find the partial derivatives defined above.

A first important result.

Proposition 15. *Consider a function $f : A \to R$ admitting a relative maximum or minimum at x_0* interior *to A. If for some v, $\frac{\partial f}{\partial v}(x_0)$ exists it is equal to zero.*

Proof. The function $g(t) = f(x_0 + tv)$ is defined in a neighborhood of $t = 0$ (it suffices to chose t small enough such that $x_0 + tv \in A$). Since in $t = 0$ it admits a relative maximum or minimum there, thus $g'(0)$ is equal to zero. $\square$

A consequence of the above proposition, is that if f admits a relative maximum or minimum in x_0 interior to A and it admits partial derivatives, then $\nabla f(x_0) = 0$ (why?). The point x_0 is said a **stationary**[8].

4.1.1 Lagrange theorem

The Lagrange theorem for functions of one real variable states that if $g : \mathbb{R} \to \mathbb{R}$ is continuous and derivable on $I := (x, x+h)$, then there exists a point $\tau \in I$ such that

$$g(x+h) - g(h) = hg(\tau).$$

This theorem can be extended to functions of more real variables. Consider f defined on $A \subset \mathbb{R}^N$ and assume that it admits derivative everywhere in A with respect to the direction v. Then, for $x_0 \in A$, $g(s) = f(x_0 + sv)$ is derivable as long as s is such that $x_0 + sv \in A$. Now:

$$f(x_0 + sv) - f(x_0) = g(s) - g(0) = sg(\tau),$$

where $\tau \in (0, s)$. Thanks to Exercise 2, it holds that

$$f(x_0 + sv) - f(x_0) = s\frac{\partial f}{\partial v}(x_0 + s\tau). \tag{15}$$

[8] The case where x_0 lies on the boundary of A is more difficult to manage and will be not considered in this course.

Of course, e.g. in $\mathbb{R}^2$, when the direction is either $v = e_1$ or $v = e_2$, it holds that

$$f(x+h,y) - f(x,y) = h\frac{\partial f}{\partial x}(\eta, y) \tag{16}$$

$$f(x,y+k) - f(x,y) = k\frac{\partial f}{\partial y}(x, \psi) \tag{17}$$

$$\tag{18}$$

respectively, where $\eta \in (0,h)$ and $\psi \in (0,k)$.

Exercises. Compute the gradients of the following functions

1. x^3y^2.
2. $\sqrt{x^2+y^2}$.
3. $\sin(xy)$.
4. $\log(x^2+y^2)$.

4.2 Differentiability

When working with functions of more variables, admitting all the partial derivatives in a point x_0 may not be enough (in terms of regularity) to ensure some desirable properties like, for instance, the continuity in x_0. This is why we need to introduce the stronger notion of differentiability.

Definition. *A function $f : \mathbb{R}^N \to \mathbb{R}$ is **differentiable** in x_0 if $\nabla f(x_0)$ exists and*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\| x - x_0 \|} = 0. \tag{19}$$

Theorem 4. *If f is differentiable in x_0, then it is continuous in x_0, it admits derivatives w.r.t. all directions v and*

$$\frac{\partial f}{\partial v}(x_0) = \langle \nabla f(x_0), v \rangle. \tag{20}$$

Proof. About continuity.

$$f(x) - f(x_0) = \frac{f(x) - f(x_0) - \langle \nabla f(x_0), x - x_0 \rangle}{\| x - x_0 \|} \| x - x_0 \| + \langle \nabla f(x_0), x - x_0 \rangle$$

and taking the limit for $x \to x_0$ the right hand side goes to zero (bilinearity of scalar product). About directional derivatives. It suffices to set $x = x_0 + tv$, for all directions v, then to replace it into Eq. (19) to obtain Eq. (20). $\square$

The above theorem has a crucial consequence:

Corollary. *The "slope" of a differentiable function f in the gradient direction is higher than in any other direction.*

Proof. Assume f to be differentiable in A and $x_0 \in A$ is *not* a stationary point for f. Then, the direction v is defined as

$$v = \frac{\nabla f(x_0)}{\| \nabla f(x_0) \|}$$

and

$$\frac{\partial f(x_0)}{\partial v} = \langle \nabla f(x_0), v \rangle = \| \nabla f(x_0) \| .$$

For any other direction w, due to the Cauchy-Swartz inequality, it holds that

$$\left| \frac{\partial f(x_0)}{\partial w} \right| = |\langle \nabla f(x_0), w \rangle| \leq \| \nabla f(x_0) \| \| w \| = \| \nabla f(x_0) \| .$$

$\square$

A useful theorem to check whether a function is differentiable or not is the following

Theorem 5. *If f admits gradient in a neighborhood of x_0 and the partial derivatives are continuous in x_0, then f is differentiable in x_0.*

Proof. omitted. $\square$

A function admitting continuous partial derivatives up to the order m on a set $E \subset \mathbb{R}^N$ is said of *class* $C^m(E)$.

Exercises. Show that the following functions are differentiable:

1. $f(x, y) = x + y$ with $x, y \in \mathbb{R}$.
2. $f(x, y) = \langle x, y \rangle$ with $x, y \in \mathbb{R}^2$.
3. $f(x, y) = e^{\|x\|^2}$, with $x \in \mathbb{R}^2$.

4.3 Higher order derivatives and composition

If each entry of the gradient of $f : \mathbb{R}^N \to \mathbb{R}$ is derivable, the partial derivatives further derived to obtain **the second order partial derivatives**. So for instance, if f is derived first with respect to x_k and then with respect to x_i, the corresponding derivative is denoted by $\frac{\partial^2 f}{\partial x_i, \partial x_k}$ or $f_{x_i x_k}$. Thus, f has N partial derivatives, N^2 second order partial derivatives ans so on. The second order partial derivatives are collected in the **Heissian** matrix $Hf \in \mathbb{R}^{N\times N}$, whose entry (i,k) is

$$(Hf)_{ki} = \frac{\partial^2 f}{\partial x_i, \partial x_k}.$$

The non diagonal entries are called *mixed* partial derivatives, the diagonal entries are the *pure* ones.

Although in principle there is no reason why Hf should be symmetric, we have the following Theorem, where $N = 2$ for simplicity.

Theorem 6. (Schwarz) *If $f : \mathbb{R}^2 \to \mathbb{R}$ has mixed partial derivatives in a neighborhood of a point (x, y) and they are* continuous *in (x, y), then $f_{xy}(x, y) = f_{yx}(x, y)$.*

Proof. Omitted. □

Exercise. Compute the Heissian matrix of $f(x, y) = x + \sin(x, y)$ and $g(x, y) = x^2 y + xy$.

We now focus on the following **curve** $x : \mathbb{R} \to \mathbb{R}^N$, such that, for a real t, $x(t) = (x_1(t), \ldots, x_N(t))^T$ and $x_n(t)$ is a real function, for all n. Assuming that $x(\cdot) \in \mathcal{C}^1(\mathbb{R})$, another function $f : \mathbb{R}^N \to \mathbb{R}$ is considered still of class $\mathcal{C}^1(\mathbb{R}^N)$. The composition $g = f \circ x : \mathbb{R} \to \mathbb{R}$ is a real function of real variable, such that $g(t) = f(x(t))$.

Proposition 16. *$g \in \mathcal{C}^1(\mathbb{R})$ and*

$$g'(t) = \langle \nabla f(x(t)), x'(t) \rangle, \tag{21}$$

where $x'(t) := (x'_1(t), \ldots, x'_N(t))^T$.

Proof.

$$\begin{aligned} g(t+h) - g(t) &= f(x(t+h) - x(t)) - f(x(t)) \\ &= f(x(t) - [x(t+h) - x(t)]) - f(x(t)) \\ &= f(x(t) + sv) - f(x(t)), \end{aligned}$$

where $v := \frac{x(t+h)-x(t)}{s}$ and $s :=\| x(t+h) - x(t) \|$.

Thanks to Eq. (15) it holds that

$$f(x(t)+sv) - f(x(t)) = s\frac{\partial f}{\partial v} x(t) + \tau v) = s\langle \nabla f(x(t)+\tau v), v\rangle,$$

where $\tau \in [0, s]$ and the last equality comes from Eq. (20). Thus

$$\frac{g(t+h)-g(t)}{h} = \left\langle \nabla f(x(t)+\tau v), \frac{x(t+h)-x(t)}{h} \right\rangle$$

and the Theorem is proven by taking the limit for $h \to 0$, thanks to the continuity of ∇f and the definition of τ and s.

$\square$

Note that when $x(t)$ is linear, i.e. $x(t) = x + tv$, Eq. (21) states that

$$g'(t) = \langle \nabla f(x+tv), v\rangle, \tag{22}$$

that we already know from Exercise 2 (via Eq. (20)) and also

$$g'(0) = \langle \nabla f(x), v\rangle = \sum_{i=1}^{N} f_{x_i}(x) v_i. \tag{23}$$

Further assuming that $f \in \mathcal{C}^2(\mathbb{R}^N)$ it holds that

$$\begin{aligned} g''(t) &= \frac{d}{dt}\left(\sum_{i=1}^{N} f_{x_i}(x+tv)v_i\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} f_{x_j x_i}(x+tv)v_j v_i \\ &= v^T H f(x+tv) v \end{aligned}$$

and

$$g''(0) = v^T \left(Hf(x)\right) v. \tag{24}$$

4.4 Local maxima and minima

The first basic idea that can be formalized is the following one. If a function $f \in \mathcal{C}^2(A \subset \mathbb{R}^N)$ has a local maximum at x_0 interior to A, then $g(t)$ also has a local maximum at $t = 0$. Thus, <u>for all directions v</u>, $g'(0) = 0$ and $g''(0) < 0$. Via Eqs. (23) and (24) the following proposition is thence proven.

Proposition 17. *If f has a local maximum (minimum) at $x_0 \in A$, then*

1. $\nabla f(x_0) = 0$,
2. $Hf(x_0)$ *is negative (positive) definite.*

However, we would like we can state the opposite:

Proposition 18. *If x_0 is such that*

1. $\nabla f(x_0) = 0$ *and*
2. $Hf(x_0)$ *is negative (positive) definite,*

then x_0 is a local maximum (minimum) for f.

Fortunately the above proposition is true! It can be proven thanks to the multivariate **Taylor's formula**:

$$f(x) = f(x_0)+\nabla f(x_0)^T(x-x_0)+\frac{1}{2}(x-x_0)^T Hf(x_0)(x-x_0)+R_2(x,x_0), \quad (25)$$

where $R_2(x,x_0) = o(\| x - x_0 \|^2)$.

proof of Proposition 4. Consider the case where $Hf(x_0)$ positive definite. The negative definite case can be treated similarly. By Eq. (25) it follows that

$$\frac{f(x)-f(x_0)}{\| x - x_0 \|^2} = +\frac{1}{2}v^T Hf(x_0)v + \frac{R_2(x,x_0)}{\| x - x_0 \|^2}$$

where $v := \frac{x-x_0}{\|x-x_0\|}$. Since the set

$$S = \{w \in \mathbb{R}^N \mid \| w \| = 1\}$$

can be proven to be compact, by the Weierstrass Theorem, if follows that

$$v^T Hf(x_0)v \geq m > 0$$

for all $v \in S$ and the last inequality comes from the positive definiteness. Thus

$$\frac{f(x)-f(x_0)}{\| x - x_0 \|^2} \geq \frac{m}{2} + \frac{R_2(x,x_0)}{\| x - x_0 \|^2}.$$

Since the last term on the right hand side tends to zero when $x \to x_0$, by definition of limit $\exists I(x_0, \delta)$ for some $\delta > 0$ such that

$$\frac{f(x) - f(x_0)}{\| x - x_0 \|^2} \geq \frac{m}{4} > 0$$

for all $x \in I(x_0, \delta)$, no matter how small m is. Since the quantity on the r.h.s. is positive, $f(x) \geq f(x_0)$. $\square$

Exercise. Find local minima/maxima for the following functions

1. $f(x, y) = x^2 + 2y^2$.
2. $f(x, y) = x^3 - x^2 - y^2$