Linear and Bilinear Algebra I

DR Lama Tarsissi
Sorbonne University of Abu Dhabi
MATH222

Euclidean spaces

Scalar product

Definition

A dot product defined on a vector space is the given born from a symmetrical bilinear form on this space whose associated quadratic form is positive definite.

An Euclidean space is a vector space with a dot product.
It is customary to use a particular notation for the scalar product of two vectors $x, y \in E$: $< x, y >, (x|y)$ or $< x|y >$. We also ask:

$$< x, x >= ||x||^2.$$

The quantity $||x|| = \sqrt{< x, x >}$ is the Euclidean norm of the vector x.

Examples

- The space $\mathbb{R}^n$ provided with the usual scalar product is of course an euclidean space.
- The space $C^0([0,1],\mathbb{R})$ with symmetrical bilinear shape

$$(f,g) \mapsto \int_0^1 f(t)g(t)dt$$

is an Euclidean space. It is not of finite dimension.

When the space is of **finite dimension**, we can find an orthogonal basis. As the scalar product is **positive definite**, we can divide each of the vectors of this basis by its Euclidean norm so that they are all of norm one.

Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space of finite dimension n. Then there exists a basis $(e_1, ..., e_n)$ of E which satisfies:

$$< e_i, e_j >= 0, \forall\ i \neq j, ||e_i|| = 1$$

Such a basis is said to be orthonormal. In this basis, any vector $x \in E$ has for coefficients $x_i =< x, e_i >$,,

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} < x, e_i > e_i =< x, e_1 > e_1 + \ldots + < x, e_n > e_n,$$

the dot product and the quadratic form are of the following form:

$$\langle x,y \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \ldots + x_n y_n,$$

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 = x_1^2 + \ldots + x_n^2.$$

Thanks to the preceeding formulas, we can express the coefficients of the matrix associated with a linear map $f : E \mapsto E$ in an orthonormal basis (e_i) using dot product. These coefficients are the coordinates of the vectors $f(e_j)$ in the base (e_i). The matrix of f is therefore equal to $\{< e_i, f(e_j) >\}_{i,j}$.

Euclidean norm

Let us verify that the function $x \mapsto ||x||$ is indeed a norm, which means that it satisfies the following properties.

Proposition

Let $(E, <\cdot,\cdot>)$ be a Euclidean space. Then

- *for all $\lambda \in R, x \in E, ||\lambda x|| = |\lambda| \times ||x||$,*
- *for all $x, y \in E, ||x + y|| \leq ||x|| + ||y||$, (triangular inequality)*
- *for all $x \in E$, $||x|| = 0$ if and only if $x = 0$.*

The demonstration of triangular inequality is based on the following inequality, attributed to Augustin Louis Cauchy (1789-1857) and Hermann Amandus Schwarz (1843-1921).

Theorem (Cauchy-Schwarz inequality)

Let $(E, <\cdot,\cdot>)$ be an Euclidean space and $x, y \in E$. Then:

$$|<x, y>| \leq ||x|| \times ||y||$$

Proof of the Cauchy-Schwarz inequality

If x and y are proportional to each other, for example $x = \lambda y$, the two terms of the inequality are equal to $\lambda||x||^2$. Otherwise, x and y generate a vector plane $vect(x, y)$, in restriction of which the matrix of the scalar product is given by

$$B = \begin{pmatrix} < x, x > & < y, x > \\ < x, y > & < y, y > \end{pmatrix}$$

Let us choose an orthonormal basis of vect (x, y). We have seen that the matrix B is of the form $P^T P$, where P is the passage matrix of this orthonormal basis to the base (x, y). So we have

$$||x||^2||y||^2 - < x, y >^2 = det(B) = det(P)^2 > 0.$$

Examples

- Apply the inequality to the two vectors (x_1, x_2) and (y_1, y_2) of R^2 with the usual scalar product. We obtain:

$$|x_1 y_1 + x_2 y_2| \leq \sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}$$

- Apply the inequality to two functions continues $f, g : [0, 1] \mapsto R$, using the dot product $< f, g >= \int_0^1 f(t)g(t)dt$. We obtain:

$$\left|\int_0^1 f(t)g(t)\,dt\right| \leq \sqrt{\int_0^1 f(t)^2\,dt}\,\sqrt{\int_0^1 g(t)^2\,dt}.$$

Proof of triangular inequality

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.$$

Area, length and angle

Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $x, y \in E$. Let (e_1, e_2) be an orthonormal basis of $vect(x, y)$ and $(x_1, x_2), (y_1, y_2)$ the coordinates of x and y in this base. through:

$$area(x, y) = |x_1 y_2 - x_2 y_1|,$$

it does not depend on the orthonormal basis (e_1, e_2) used to calculate it.

This area is equal to the absolute value of the determinant of the matrix P considered in the proof of the previous theorem. In fact, P is the matrix from the base (e_1, e_2) to the base (x, y), it is therefore written

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

The formula seen above, $det(B) = det(P)^2$, becomes

Proposition

Let $(E, < \cdot, \cdot >)$ be an Euclidean space and $x, y \in E$. Then

$$< x, y >^2 + area(x, y)^2 = ||x||^2 ||y||^2.$$

This formula shows that the area is **independent of the orthonormal basis** chosen to calculate it. On the other hand, the sign of the determinant $det(P) = x_1 y_2 - x_2 y_1$ depends on this basis. This will lead us to introduce the notion of the oriented Euclidean vector space.

Example

Consider two vectors (x_1, x_2) and (y_1, y_2) of R^2 endowed with the usual dot product. The proposition leads to equality

$$(x_1 y_1 + x_2 y_2)^2 + (x_1 y_2 - x_2 y_1)^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

This formula is called **Lagrange identity**. It can be demonstrated directly by expanding the squares.

Method:Calculation of the area of a parallelogram of R^n
To calculate the area of the parallelogram whose sides are given by two vectors x,y of R^n, we could construct an orthogonal basis of the plane vect (x, y) as explained previously, express x and y in this base and calculate a determinant. It is faster to use the previous formula.

Calculate the area of the parallelogram of R^3 whose sides are given by $x = (1, 1, 0)$ and $y = (0, 1, 1)$.

$||x||^2 = \sqrt{1^2 + 1^2 + 0^2} = 2, ||y||^2 = \sqrt{0^2 + 1^2 + 1^2} = 2, < x, y >= 1 \times 0 + 1 \times 1 + 0 \times 1 = 1$, Hence, $area(x, y) = \sqrt{2 + 2 - 1^2} = 3$.

The Cauchy-Schwarz inequality shows that for all vectors $x, y \in E$, the real $< x, y > /||x||\ ||y||$ is between -1 and 1. Recall that for all real has included between $[-1, 1]$, there exists a unique real number $\theta \in [0, \pi]$ such that $a = cos(\theta)$. This makes it possible to define the unoriented angle of two vectors of the following way.

Definition (non-oriented angles)

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $x, y \in E$ not zero. The measure of the unoriented angle defined by the pair $\{x, y\}$ is the unique $\theta \in [0, \pi]$ satisfying:

$$< x, y >= ||x||\ ||y|| cos(\theta).$$

Method: calculations of angles and dot products

To calculate the unoriented angle of two vectors whose coordinates in an orthonormal basis are given, we calculate their norm, their scalar product and we call on the inverse function of the cosine, which we denote by **acos** in the following. Conversely, we can calculate the dot product of two vectors if we know their norms and angles.

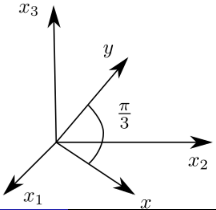
Examples

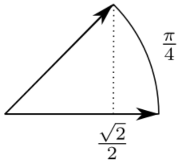
Let us take again the two vectors $x = (1, 1, 0)$ and $y = (0, 1, 1) \in R^3$. We have shown that $||x||^2 = 2, ||y||^2 = 2, < x, y >= 1$. The measure of the unoriented angle between x and y is therefore given by

$$\operatorname{acos}\left(\frac{1}{\sqrt{2}\sqrt{2}}\right) = \operatorname{acos}\left(\frac{1}{2}\right) = \frac{\pi}{3}.$$

Consider two vectors x, y of norm one and forming an angle equal to

$$\langle x, y \rangle = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}.$$





Some formulas

Let $(E, <\cdot,\cdot>)$ be a Euclidean space and $x, y \in E$. Then:

$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2$ (parallelogram formula).
$||x||^2 + ||y||^2 = \frac{1}{2}||x-y||^2 + 2||\frac{x+y}{2}||^2$ (median formula)

To interpret the first of these formulas geometrically, we consider a parallelogram of vertices A, B, C and D located in the same plane and we take for x the vector going from A to B, for y the vector going from A to D. We get:

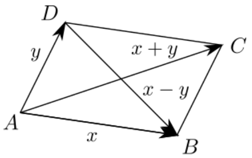
$$2AB^2 + 2BC^2 = AC^2 + BD^2.$$

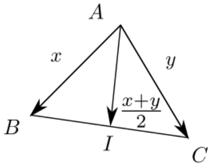
The sum of the squares of the lengths of the four sides of a parallelogram is equal to the sum of the squares of the lengths of the diagonals.

For the second formula, we consider a triangle with vertices A, B and C and we denote by I the midpoint of BC.

$$AB^2 + AC^2 = BC^2 + 2AI^2.$$

This formula expresses the square of the length of the median of the triangle in function of the squares of the sides of the triangle. It dates back to Thales of Miletus (-625 -547).





Proof of the parallelogram formula

Just expand the squares in the first term.

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle + \|x\|^2 + \|y\|^2 - 2\langle x,y\rangle \\ &= 2\,\|x\|^2 + 2\,\|y\|^2. \end{aligned}$$

Finally, we give a multidimensional analogue of the Pythagorean theorem.

Theorem (from Pythagoras)

Let $(E, <\cdot,\cdot>)$ be a Euclidean space and $v_1, ..., v_n$ two-by-two orthogonal vectors of E: $<v_i, v_j>= 0$ for all i, j distinct. Then:

$$||v_1 + ... + v_n||^2 = ||v_1||^2 + ... + ||v_n||^2.$$

Proof:

$$\|v_1 + ... + v_n\|^2 = \left\langle \sum_i v_i, \sum_j v_j \right\rangle = \sum_{i,j} \langle v_i, v_j\rangle = \sum_i \|v_i\|^2.$$

Oriented Euclidean space

Orientation of a Euclidean space

Definition

We place ourselves in a Euclidean space $(E, < \cdot, \cdot >)$ of finite dimension. We will say that two bases (e_i) and (e'_i) have the same orientation if the determinant of the matrice of passage from one base to the other is strictly positive.
If on the contrary it is negative, we will say that the bases have opposite orientations.

1-To orient a finite dimensional Euclidean space it consists in choosing a base that is said to be positively oriented, or even directly oriented.

2-All the bases having the same orientation as the chosen base are said to be oriented positively.

3-The others are said to be negatively oriented, indirectly, or in the retrograde direction.

4-Two bases oriented indirectly have same direction, under the base change formulas.

5-A definite invertible map of a Euclidean space oriented in itself preserves orientation if a base and its image have the same orientation. We will see later how to characterize this property in terms of its determinant.

Examples

- The space R^n has a canonical orientation given by the canonical basis. - A plane of R^3 inherits a Euclidean structure obtained by restriction of the usual scalar product of R^3 but it does not have a given orientation a priori. To give oneself an orientation of such a plan amounts to choosing a vector e_3 of norm one orthogonal to the plane. A basis (e_1, e_2) of the plane is considered as direct if (e_1, e_2, e_3) is a direct basis of R^3 with its cannonical orientation. There are two possible choices for the vector e_3, each corresponding to a different orientation.

Determinant of a family of n vectors

The notion of oriented Euclidean space of dimension n makes it possible to define the determinant of a family of n vectors $(v_1, ..., v_n)$. This is the determinant of the matrix whose columns are given by the coordinates of the v_i in a direct orthonormal basis. We will demonstrate in the following that it does not depend of the chosen direct orthonormal basis.

Angles oriented in the plane

Another application of the notion of orientation is that of oriented angle. We place ourselves in an oriented Euclidean space E of dimension two. We saw previously that for all $x, y \in E$ not zero,

$$\left(\frac{\langle x,y\rangle}{\|x\|\|y\|}\right)^2 + \left(\frac{\det(x,y)}{\|x\|\|y\|}\right)^2 = 1.$$

Recall that for all real numbers a, b satisfying $a^2 + b^2$, we can find a real θ, unique to a multiple of 2π near, such that $a = cos(\theta)$ and $b = sin(\theta)$.

Definition (oriented angles)

Let $(E, < \cdot, \cdot >)$ be an oriented Euclidean space of dimension two and x, y two vectors of E not zero. The angle measurement oriented defined by the pair (x, y) is the unique $\theta \in R/2\pi Z$ satisfying

$$\langle x,y \rangle = \|x\| \, \|y\| \, \cos(\theta),$$

$$\det(x,y) = \|x\| \, \|y\| \, \sin(\theta),$$

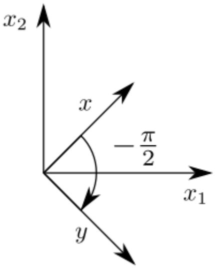
the determinant being calculated in an oriented orthonormal basis of E.

Example

Let us place ourselves in R^2 with its canonical orientation and consider the vectors $x = (1,1)$ and $y = (1,-1)$. We have

$$||x||^2 = 2, ||y||^2 = 2, < x, y >= 0, det(x,y) = det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2.$$

The measure θ of the oriented angle formed by these two vectors verify: $cos(\theta) = 0/2 = 0$, $sin(\theta) = -2/2 = -1$ which implies $\theta = -\pi/2 \ mod 2\pi$.



Orthogonality

Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and F a vector subspace of E. The orthogonal of F is the vector subspace of E made up of vectors which are orthogonal to all vectors of F:

$$F^{\perp} = \{x \in E | \forall y \in F, < x, y >= 0\}$$

A vector in both F and $F^{\perp}$ is orthogonal to itself, so it is null. We deduce that $F \cap F^{\perp} = \{0\}$.

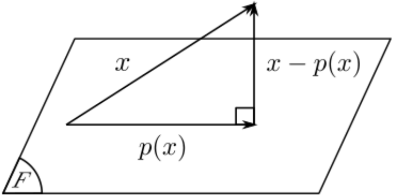
Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and F a vector subspace of E of finite dimension. Then for all $x \in E$, there exists a unique vector p(x) which satisfies the following two conditions:

- $p(x) \in F$
- $x - p(x) \in F^{\perp}$

The vector p (x) is the orthogonal projection of x on F. Let (e_i) be an orthonormal basis of F, then:

$$p(x) = \sum_i < x, e_i > e_i =< x, e_1 > e_1 + ... + < x, e_n > e_n.$$



Proof:
The sum $\sum_i < x, e_i > e_i$ satisfies the two desired properties: it is in F and for all j,:

$$< x - \sum_i < x, e_i > e_i, e_j >=< x, e_j > - < x, e_j >= 0$$

Let y, y' be vectors such that y, y' $\in$ F and $x - y, x - y' \in F^{\perp}$. The difference y - y' = (x - y') - (x - y) is both in F and in $F^{\perp}$, so it is zero.

Method: calculation of the orthogonal projector on a subspace

It suffices to calculate an orthonormal basis (e_i) of this subspace and to use the formula $p(x) = \sum < x, e_i > e_i$.
Note that the map $x \mapsto p(x)$ is linear. Let's pose
$x_1 = p(x) \in F, x_2 = x - p(x) \in F^{\perp}$.
Remarkably, the vector x_2 is in $F^{\perp}$ and satisfies $x - x_2 \in F$. This shows the existence of the orthogonal projection of x on $F^{\perp}$, given by $x - p(x)$.
We have shown the following result.

Corollary

All $x \in E$ is written uniquely in the form:

$$x = x_1 + x_2 \text{ with } x_1 \in F \text{ and } x_2 \in F^{\perp}.$$

We say that E is the **direct sum** of F and $F^{\perp}$ and we denote:

$$E = F \oplus F^{\perp}.$$

The orthogonal projection on F is the unique linear map $p : E \mapsto E$ satisfying: $\forall x \in F, p(x) = x, \forall x \in F^{\perp}, p(x) = 0$. Recall that the dimension of two direct sum subspaces is equal to the sum of their dimensions.
In addition, we always have the inclusion $F \subset (F^{\perp})^{\perp}$. If E is of finite dimension, the previous corollary shows that these two spaces have the same dimension. They are therefore equal.

Corollary

Let $(E, < \cdot, \cdot >)$ be a finite dimensional Euclidean space and F a vector subspace of E. Then

- $dim(F) + dim(F^{\perp}) = dim(E)$
- $F = (F^{\perp})^{\perp}$

Example

Let us consider the hyperplane: $H = \{x \in R^n | a_1x_1 + \ldots + a_nx_n = 0\}$. Let $v = (a_1, \ldots, a_n)$. Then H is the orthogonal of $vect(v)$.

$$\mathrm{vect}(v)^{\perp} = H, \quad H^{\perp} = \mathrm{vect}(v).$$

An orthonormal basis of $vect(v)$ is given by the singleton $(v/||v||)$. The projection on vect (v) is therefore worth

$$p(x) = \frac{\langle x, v \rangle}{\|v\|^2}\, v = \frac{\sum a_k x_k}{\sum a_k^2} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

The matrix of p has for coefficients $\{a_i a_j / \sum a_k^2\}$, it is given by:

$$\frac{1}{\sum_{k=1}^{n} a_k^2} \begin{pmatrix} a_1^2 & a_1 a_2 & \dots & a_1 a_n \\ \vdots & \vdots & & \vdots \\ a_n a_1 & a_n a_2 & \dots & a_n^2 \end{pmatrix}.$$

The orthogonal projection on H has the expression

$$p_H(x) = x - \frac{\langle x, v \rangle}{\|v\|^2}\, v.$$

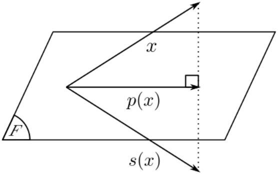
Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space, F a subspace of E of finite dimension. The orthogonal symmetry with respect to F is unique linear map $s : E \mapsto E$ such that:

$$\forall x \in F, s(x) = x, \forall x \in F^{\perp}, s(x) = -x$$

When F is a hyperplane, we speak of orthogonal reflection. Let us denote by p the orthogonal projection on F. Then:

$$s(x) = 2p(x) - x$$



Examples

- The symmetry relative to the subspace $\{0\}$ is called central symmetry. It is given by $s(x) = -x$ for all x. The symmetry with respect to E is equal to identity.
- The orthogonal reflection relative to a hyperplane $H = vect(v)^{\perp}$ is given by the expression:

$$s(x) = x - 2\frac{< x, v >}{||v||^2}v$$

- The orthogonal symmetry with respect to a line vect (v) is given by:

$$s(x) = 2\frac{< x, v >}{||v||^2}v - x$$

In dimension 3, an orthogonal symmetry relative to a line is called a U-turn.

Gram-Schmidt orthonormalization method

Let E be a vector space endowed with a symmetric bilinear form ϕ. We saw before with the quadratic forms how to find an orthogonal basis for ϕ.

- We start by reducing the quadratic form Q associated with ϕ. We obtain linear forms $l_1, ..., l_{p+q}$ linearly independent such that

$$Q = \sum \pm l_i^2$$

- We complete the family (l_i) in a basis of E^* and we construct a basis (e_i) so that the base (l_i) is dual to it.

This general method takes quite a long time to implement. In the case of a dot product, there is a faster process attributed to Jorgen Pedersen Gram (1850-1916) and Erhard Schmidt (1876-1959) but we already found in the works of Pierre-Simon de Laplace (1749-1827).

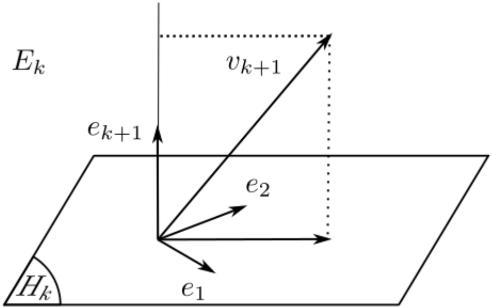
Method

We start from a basis $(v_1, ..., v_n)$ of E and we try to build by induction an orthonormal basis $(e_1, ..., e_n)$ so that the vectors $(v_1, ..., v_k)$ and $(e_1, ..., e_k)$ generate the same vector space for all k going from 1 to n. In this case, v_1 and e_1 are necessarily proportional and e_1 must be of norm one, which leaves only one possibility if we assume $< e_1, v_1 >> 0$:

$$e_1 = \frac{v_1}{||v_1||}$$

Suppose constructed $(e_1, ..., e_k)$ and construct e_{k+1}. It must be a vector of norm one belonging to $E_k = vect(e_1, ..., e_k, v_{k+1})$ orthogonal to space $H_k = vect(e_1, ..., e_k)$. It is therefore a matter of projecting a vector of space E_k onto orthogonal $H_k^\perp$ to the hyperplane $H_k \subset E_k$ and normalize it.

$$e_{k+1} = \frac{v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i}{\left\| v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i \right\|}.$$



Theorem (Gram-Schmidt orthonormalization)

Let $(E, <\cdot,\cdot>)$ a finite dimensional Euclidean space and $(v_1, ..., v_n)$ a basis of E. Then there exists a unique orthonormal basis $(e_1, ..., e_n)$ such that, for any integer i between 1 and n,

- $vect(e_1, ..., e_i) = vect(v_1, ..., v_i)$
- $< e_i, v_i >> 0$.

Example

Let's compare the two methods on an example. We place ourselves on R^2 and we take $v_1 = (3,4)$ and $v_2 = (4,3)$. We have:

$$||v_1|| = 5, ||v_2|| = 5, < v_1, v_2 >= 24.$$

We obtain by the first method:

$$\begin{aligned}
Q(x_1 v_1 + x_2 v_2) &= 25x_1^2 + 25x_2^2 + 48x_1x_2 \\
&= 25\Big(x_1 + \frac{24}{25}\,x_2\Big)^2 + \frac{49}{25}\,x_2^2 \\
&= l_1(x_1,x_2)^2 + l_2(x_1,x_2)^2.
\end{aligned}$$

with $l_1(x_1, x_2) = 5x_1 + \frac{24}{5}x_2$ and $l_2(x_1, x_2) = \frac{7}{5}x_2$

$$\begin{pmatrix} 5 & \frac{24}{5} \\ 0 & \frac{7}{5} \end{pmatrix}^{-1} = \frac{1}{7}\begin{pmatrix} \frac{7}{5} & -\frac{24}{5} \\ 0 & 5 \end{pmatrix} = \frac{1}{35}\begin{pmatrix} 7 & -24 \\ 0 & 25 \end{pmatrix}.$$

The coordinates obtained are those of an orthogonal base (e_1, e_2) in the base (v_1, v_2).

$$e_1 = \frac{1}{5}\, v_1 = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

$$e_2 = \frac{1}{35}(-24v_1 + 25v_2) = \frac{1}{35}\begin{pmatrix} -24 \times 3 + 25 \times 4 \\ -24 \times 4 + 25 \times 3 \end{pmatrix} = \frac{1}{35}\begin{pmatrix} 28 \\ -21 \end{pmatrix}.$$

The Gram-Schmidt method gives

$$e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

$$e_2 = \frac{v_2 - \langle v_2, e_1 \rangle\, e_1}{\|v_2 - \langle v_2, e_1 \rangle\, e_1\|} = \frac{1}{\|...\|}\begin{pmatrix} 4 - 24/5 \times 3/5 \\ 3 - 24/5 \times 4/5 \end{pmatrix} = \frac{1}{35}\begin{pmatrix} 28 \\ -21 \end{pmatrix}.$$

Conclusion

This process requires fewer calculations. Unsurprisingly, both methods lead to the same result. In the positive definite case, the general algorithm produces a triangular system, which must coincide, except for signs, with the result obtained by the Gram-Schmidt method.

Isometries

Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space. A linear application $f : E \mapsto E$ is an isometry of E if it preserves the dot product:

$$\textit{For all } x, y \in E, < f(x), f(y) > = < x, y > .$$

Thanks to the polarization identities, we verify that a linear application is an isometry if and only if it preserves the norm:

$$||f(x)|| = ||x|| \textit{ for all } x \in E.$$

$$\begin{aligned}
\varphi(x,y) &= \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big), \\
\varphi(x,y) &= \frac{1}{2}\Big(Q(x) + Q(y) - Q(x-y)\Big), \\
\varphi(x,y) &= \frac{1}{4}\Big(Q(x+y) - Q(x-y)\Big).
\end{aligned}$$

Examples

- Identity is an isometry.
- Orthogonal symmetries are isometries. In fact an orthogonal symmetry s is of the form $s(x) = p(x) - (x - p(x))$ with x and x -p (x) two orthogonal vectors of sum equal to x, so that:

$$||s(x)||^2 = ||p(x)||^2 + ||-(x-p(x))||^2 = ||p(x)||^2 + ||x-p(x)||^2 = ||x||^2.$$

Notice that an **isometry has a kernel reduced to zero vector**. Indeed, if $f(x) = 0$ then $||x|| = ||f(x)|| = 0$. In finite dimension, this implies that an isometry is invertible.
Moreover, if $(e_1, ..., e_n)$ is an orthonormal basis, the same is true for its image $(f(e_1), ..., f(e_n))$ because the scalar products of the vectors e_i, e_j are preserved by f. This property characterizes the isometries.

Proposition

Let $(E, <\cdot,\cdot>)$ be a finite dimensional Euclidean space endowed with an orthonormal basis $(e_1, ..., e_n)$ and $f : E \mapsto E$ a linear map. Then f is an isometry if and only if the image $(f(e_1), ..., f(e_n))$ of this orthonormal basis is an orthonormal basis.

Proof.

Let $x = \sum x_i e_i$ and $y = \sum y_j e_j \in E$. If $(f(e_1), ..., f(e_n))$ is orthogonal,

$$< f(x), f(y) >= \sum_{i,j} x_i y_j < f(e_i), f(e_j) >= \sum_i x_i y_i =< x, y > .$$

□

Orthogonal group

The matrix of an isometry expressed in an orthonormal basis has a particular form. Here's how to describe it.

Definition

A matrix P of size $n \times n$ is said to be orthogonal if it checks the equality

$$P^T P = id.$$

The set of orthogonal matrices is called an orthogonal group and denoted

$$O_n(R) = \{P \in M_n(R) | P^T P = id\}.$$

The special orthogonal group is composed of the orthogonal matrices of determinant 1.

$$SO_n(R) = \{P \in M_n(R) | P^T P = id, det(P) = 1\}.$$

Note that an orthogonal matrix necessarily has a determinant equal to ± 1:

$$det(P)^2 = det(P^T P) = det(id) = 1.$$

The orthogonal matrices of determinant 1 are those which preserve the canonical orientation of R^n.

Method: characterization of orthogonal matrices

To show that a matrix is orthogonal, it suffices to verify that its **columns are of norm one and orthogonal to each other.** Indeed, the coefficient i, j of the matrix $P^T P$ is obtained by making the scalar product of the column i of P by its column j. This matrix is equal to the identity if and only if its columns form an orthonormal family of R^n.

Example

The matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

is orthogonal and of determinant one. It is an element of $SO_2(R)$. The transformation of R^2 associated is called rotation of angle θ

Proposition

Consider a Euclidean space $(E, < \cdot, \cdot >)$ of finite dimension, an isometry $f : E \mapsto E$ and an orthonormal basis $(e_1, ..., e_n)$ of E. Then f is an isometry if and only if its matrix in the base (e_i) is orthogonal.

Proof.

The columns of the matrix of f are given by the coordinates of the vectors from the base $(f(e_i))$ to the base (e_i). These column vectors are orthogonal relatively to the usual scalar product over R^n and of norm one if and only if the basis $(f(e_i))$ is orthonormal. We have seen that this characterizes the isometries. □

Euclidean plane isometries

To illustrate the concepts we have just defined, we will classify the isometries in dimension two. We start by working in coordinates with orthogonal matrices.

Proposition

Any matrix of $SO_2(R)$ is the matrix of a rotation. Any matrix of $O_2(R)$ with determinant -1 is the matrix of an orthogonal reflection.

Recall that an orthogonal reflection is an orthogonal symmetry by compared to a hyperplane. In dimension 2, the hyperplanes are lines.

Proof

A matrix of $O_2(R)$ has the form $P = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ with its two columns of norm one and orthogonal to each other, which implies

$$det \begin{pmatrix} a & -d \\ b & c \end{pmatrix} = ac + bd = 0.$$

The vectors $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} -d \\ c \end{pmatrix}$ are therefore proportional and of norm one, the coefficient of proportionality is equal to ± 1 and the matrix can take one of the two following forms.

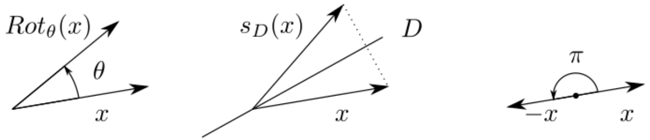
- $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$: we obtain a rotation by setting $a = cos(\theta), b = sin(\theta)$.
- $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$: this is an orthogonal reflection.
  Let us caculate the axis of symmetry.

The line of symmetry must be fixed, we get it by solving the equation $Av = v$, which gives $v = \begin{pmatrix} b & \\ & -a \end{pmatrix}$ The orthogonal space to vect (v) is a line generated by the vector $w = \begin{pmatrix} a-1 \\ b \end{pmatrix}$. The image of this vector by our matrix is

$$Aw = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \begin{pmatrix} a-1 \\ b \end{pmatrix} = \begin{pmatrix} a(a-1)+b^2 \\ (a-1)b-ab \end{pmatrix} = \begin{pmatrix} 1-a \\ -b \end{pmatrix} = -w$$

taking into account the relation $a^2 + b^2 = 1$. It is therefore indeed an orthogonal symmetry relative to vect (v). The proposition is demonstrated.

Attention, in dimension two, a central symmetry is a rotation angle π, it is not a reflection.



You also have to be careful not to be fooled by the terminology. that we just introduced. The matrix of an orthogonal symmetry is orthogonal. On the other hand, the matrix of an orthogonal projection is not orthogonal in general: an orthogonal matrix is invertible while a projection different from the identity is never invertible.

Method: characterization of the isometries of R^2

An element $P \in O_2(R)$ is a rotation if its determinant is equal to 1 and a reflection otherwise. The angle of a rotation is determined by putting the first column in the form $(cos(\theta), sin(\theta))$. We find the line fixed by a reflection by solving the linear system $Px = x$.

Example

- The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the relatively orthogonal reflection matrix with respect to the line generated by the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- The matrix$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the matrix of the angle rotation $\frac{\pi}{2}$

Corollary

The isometries of a Euclidean vector space of dimension two are rotations or orthogonal reflections.

Affine isometries

We have defined an isometry as a linear map of a space Euclidean $(E, <\cdot,\cdot>)$ which preserves the Euclidean norm. There is a notion more general of isometry related to the metric space structure of an Euclidean space. Define the Euclidean distance between two points $x, y \in E$ by:$d(x, y) = ||x - y||$.
A map $f : E \mapsto E$ is an isometry if it preserves the distance:

$$\forall x, y \in E, d(f(x), f(y)) = d(x, y).$$

Here, we do not assume that f is linear.
The translation of vector v is the map defined from E to E by $t_v(x) = x + v$. This is an example of isometry that is not a linear application. We can compose the translations with linear isometries to get new isometries.

Definition

An affine isometry is the compound of a translation and linear isometry.

There are no other isometries in Euclidean space.

Proposition

In a Euclidean space, all isometry is affine.

Proof.

Let $f : E \mapsto E$ be an isometry which fixes the origin: $f(0) = 0$. By definition of an isometry, this implies $||f(x)|| = ||x||$ Let us show that f preserves the scalar product.

$$\begin{aligned}
\langle f(x), f(y)\rangle &= -\langle f(x), -f(y)\rangle \\
&= -\left(\|f(x)-f(y)\|^2 - 2\|f(x)\|^2 - 2\|f(y)\|^2\right) \\
&= -\left(\|x-y\|^2 - 2\|x\|^2 - 2\|y\|^2\right) \\
&= \langle x, y\rangle.
\end{aligned}$$

□

Proof.

Show that f is linear

$$\begin{aligned}
&\|f(x+y)-f(x)-f(y)\|^2 \\
&= \|f(x+y)\|^2+\|f(x)\|^2+\|f(y)\|^2 \\
&\quad -2\langle f(x+y),f(x)\rangle-2\langle f(x+y),f(y)\rangle-2\langle f(x),f(y)\rangle \\
&= \|x+y\|^2+\|x\|^2+\|y\|^2-2\langle x+y,x\rangle-2\langle x+y,y\rangle-2\langle x,y\rangle \\
&= \|x+y-x-y\|^2 \\
&= 0
\end{aligned}$$

An analogous calculation gives the equality $f(\lambda x)=\lambda f(x)$. Now let f be an isometry which does not fix the origin. We set $\widehat{f}(x)=f(x)-f(0)$. This new application preserves the origin, so it is linear and $f=t_{f(0)}\circ\widehat{f}$. The proposition is demonstrated. $\square$

Gram's determinant

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $v_1, ..., v_n$ be vectors of E. We can restrict the scalar product to the subspace generated by the v_i. If these vecteurs form a free family, the matrix of the dot product takes the form $B = \{< v_i, v_j >\}_{i,j}$. The determinant of this matrix is called the determinant of Gram.

$$Gram(v_1,...,v_n) = \det \begin{pmatrix} \langle v_1,v_1\rangle & \dots & \langle v_1,v_n\rangle \\ \vdots & & \vdots \\ \langle v_n,v_1\rangle & \dots & \langle v_n,v_n\rangle \end{pmatrix}.$$

If the vectors $v_1, ..., v_n$ are linearly dependent, a dedependence on these vectors induces a dependency relation on the columns of the matrix B and the preceding determinant is zero. We deduce a criterion independence in terms of scalar product.

Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $v_1, ..., v_n$ be vectors of E. Then these vectors form a free family if and only if

$$Gram(v_1, ..., v_n) \neq 0.$$

Gram's determinant is related to the notion of volume. By analogy with the case of dimension two, we define the n-dimensional volume of the parallelepiped of sides $v_1, ..., v_n$ from the determinant of the matrix whose columns are given by the coordinates of v_j in a base orthonormal (e_i) of the space generated by v_j.

Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space, $(v_1, ..., v_n)$ a free family of E and $(e_1, ..., e_n)$ an orthonormal basis of the space generated by v_j. The n-dimensional volume of the parallelepiped with sides $v_1, ..., v_n$ is defined by:

$$vol_n(v_1, ..., v_n) = |det(v_1, ..., v_n)| = |det(\{< e_i, v_j >\}_{i,j})|.$$

We will express this volume as a Gram determinant.
The basis change formula for definite positive quadratic forms is expressed in the form $B = P^T P$, where P is the passage matrix from the base (e_i) to the base (v_j). It has for columns the coordinates of v_j in the base (e_i): $P = \{< e_i, v_j >\}_{i,j}$. By taking the determinant, we get $det(B) = det(P)^2$, which gives the formula.

Proposition

$$vol_n(v_1, ..., v_n) = |det(\{< e_i, v_j >\}_{i,j})| = \sqrt{Gram(v_1, ..., v_n)}.$$

The last term does not depend on the base (e_i), we see that the volume does not depend on the orthonormal basis chosen to calculate it.

Distance to a subspace

Let us give another application of the determinant of Gram to the computation of the distance from a vector to a subspace F of a Euclidean space $(E, < \cdot, \cdot >)$.
We define the distance between two points $x, y \in E$ as well as the distance between a point $x \in E$ and a vector subspace $F \subset E$ of finite dimension by:

$$d(x, y) = ||x - y||, d(x, F) = min\{d(x, y)|y \in F\}.$$

Proposition

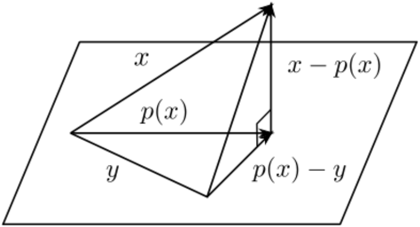
Let $(E, < \cdot, \cdot >)$ be a Euclidean space and F a vector subspace of E of finite dimension. Denote by p the orthogonal projection on F. Then

$$d(x, F) = ||x - p(x)||.$$

Proof

Let $y \in F$, the vectors $x - p(x) \in F^{\perp}$ and $p(x) - y \in F$ are orthogonal hence

$$\begin{aligned} ||x-y||^2 &= ||x - p(x) + p(x) - y||^2 \\ &= ||x - p(x)||^2 + ||p(x) - y||^2. \end{aligned}$$



The quantity $||x - p(x)||$ therefore minimizes $||x - y||$ for all $y \in F$ with equality if and only if $y = p(x)$.

Let now $(v_1, ..., v_n)$ be a basis of F and $v_{n+1} \in E$. We can express the distance of v_{n+1} to F using Gram determinants.

Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space, $(v_1, ..., v_n)$ a free family of E and $v_{n+1} \in E$. Then the square of the distance from v_{n+1} to the space F generated by the v_i is given by

$$d(v_{n+1}, F)^2 = \frac{Gram(v_1, ..., v_{n+1})}{Gram(v_1, ..., v_n)}.$$

Proof

If v_{n+1} is in F, there is nothing to prove. Otherwise, we place ourselves in the Euclidean pace generated by $(v_1, ..., v_{n+1})$. Let $(e_1, ..., e_n, e_{n+1})$ be an orthogonal basis of this space chosen such that $(e_1, ..., e_n)$ is a basis for F. The vector $v_{n+1} - p(v_{n+1})$ is the projection of v_{n+1} on e_{n+1} and is equal $< e_{n+1}, v_{n+1} > e_{n+1}$

$$Gram(v_1,...,v_{n+1}) = \det \begin{pmatrix} \langle v_1,e_1\rangle & \ldots & \langle v_n,e_1\rangle & \langle v_{n+1},e_1\rangle \\ \vdots & & & \vdots \\ \langle v_1,e_n\rangle & \ldots & \langle v_n,e_n\rangle & \langle v_{n+1},e_n\rangle \\ \langle v_1,e_{n+1}\rangle & \ldots & \langle v_n,e_{n+1}\rangle & \langle v_{n+1},e_{n+1}\rangle \end{pmatrix}^2$$

All the coefficients of the last row are zero except the last one. We concludes by expanding the determinant relatively to this line.

Reduction to linear applications

Definition of the determinant

The determinant of a square matrix can be defined by induction on the dimension of the matrix, by expansion relative to the first column.

- The determinant of a 1×1 matrix is equal to its unique coefficient.

$$det(a_{1,1}) = a_{1,1}.$$

- The determinant of a $n \times n$ matrix is then obtain using the formula

$$\det(\{a_{i,j}\}_{\substack{i=1..m \\ j=1..n}}) = \sum_{k=1}^{n} (-1)^{k+1}\, a_{k,1}\, \det(\{a_{l,m}\}_{l\neq k, m\neq 1}).$$

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ a_{3,1} & a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n} \end{vmatrix} = a_{1,1} \begin{vmatrix} a_{2,2} & \dots & a_{2,n} \\ a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{n,n} \end{vmatrix} - a_{2,1} \begin{vmatrix} a_{1,2} & \dots & a_{1,n} \\ a_{3,2} & \dots & a_{3,n} \\ \vdots & & \vdots \\ a_{n,2} & \dots & a_{n,n} \end{vmatrix} + \dots$$

$$\dots + (-1)^{n+1} a_{n,1} \begin{vmatrix} a_{1,1} & \dots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \dots & a_{n-1,n-1} \end{vmatrix}.$$

Determinant of a triangular matrix

A matrix is upper triangular if $a_{i,j} = 0$ for all i, j such that $i > j$. The determinant of such a matrix is equal to the product of its terms on the diagonal.

$$\det \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{n-1,1} & a_{n,1} \\ 0 & a_{2,2} & \cdots & a_{n-1,2} & a_{n,2} \\ 0 & 0 & \cdots & a_{n-1,3} & a_{n,3} \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \vdots & 0 & a_{n,n} \end{pmatrix} = a_{1,1}\, a_{2,2} \ldots a_{n,n}.$$

Calculation of the determinant by the Gaussian algorithm

Adding to a line a multiple of another line does not modify the value of determining. Swapping two lines changes its sign. Using these two operations, we can put the matrix in triangular form by using the Gaussian pivot method. The determinant is then equal to the product diagonal terms. Be careful not to forget to change the sign of determining each time we swap two lines.

$$\begin{vmatrix} -1 & 2 & 2 \\ 3 & -1 & 3 \\ 5 & 5 & -1 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 2 \\ 0 & 5 & 9 \\ 0 & 15 & 9 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 2 \\ 0 & 5 & 9 \\ 0 & 0 & -18 \end{vmatrix} = (-1) \times 5 \times (-18) = 90.$$

Properties of determinants

Let A be a matrix of size $n \times n$. Denote by $A_1, ..., A_n$ its columns and consider derive the determinant of A as a function of the columns of A:

$$(A_1, ..., A_n) \mapsto det(A_1, ..., A_n) = det(A).$$

The two key properties of the determinant are as follows.

- The determinant is an alternating function of the columns of the matrix: if two columns are equal, then the determinant is **zero.**
- The determinant depends linearly on each of the columns of the matrix: for all j, the map $A_j \mapsto det(A_1, ..., A_j, ..., A_n)$ is linear.

Proposition

The determinant of a product of square matrices is equal to the product of the determinants.

$$det(AB) = det(A)det(B).$$

The determinant of an invertible matrix is equal to the inverse of its determinant.

$$det(A^{-1}) = det(A)^{-1}.$$

Multilinearity also implies the formula

$$det(\lambda A) = \lambda^n det(A).$$

Recall that the transpose of the square matrix $A = \{a_{i,j}\}$ is the matrix of coefficients $\{a_{j,i}\}$. It is denoted A^T. The columns of A become the lines of t A and vice versa.

Proposition

The determinant of a matrix is equat to that of its transposed.

$$det(A^T) = det(A).$$

The determinant also makes it possible to determine whether a matrix is invertible.

Proposition

A square matrix is invertible if and only if its determinant is non-zero.

The solution of a linear system is expressed as a quotient of two determinants using Cramer's formulas.

Proposition

Cramer's formulas Let A be a matrix of size $n \times n$ and x,y two column vectors of size n. Denote by $x_1, ..., x_n$ the coordinates of x. Suppose that $Ax = y$ and $det(A) \neq 0$. Then we have for all j,

$$x_j = \frac{\det(A_1, ..., A_{j-1}, y, A_{j+1}, ..., A_n)}{\det(A_1, ..., A_n)}.$$

The determinant in the denominator is the determinant of matrix A. determinant in the numerator is that of the matrix obtained from A in replacing its column j by y.

Proof of Cramer's formulas

The equality $Ax = y$ is expressed in columns by the relation $\sum_k x_k A_k = y$. We have so

$$\begin{aligned}\det(A_1,\ldots A_{j-1},y,A_{j+1},\ldots,A_n) &= \det(A_1,\ldots A_{j-1},\sum_k x_k\, A_k,\, A_{j+1},\ldots,A_n) \\ &= \sum_k x_k\, \det(A_1,\ldots A_{j-1},A_k,A_{j+1},\ldots,A_n) \\ &= x_j \det(A_1,\ldots A_{j-1},A_j,A_{j+1},\ldots,A_n)\end{aligned}$$

by linearity and by using the alternating character of the determinant.

Cramer's formulas are used to express the coordinates of a vector tor of R^n in a given base in the form of determinants. Consider a basis $(v_1, ..., v_n)$ of R^n and apply the previous result to the matrix A whose columns are given by v_j. We obtain :

Corollary

Let $(v_1, ..., v_n)$ be a basis of R^n and y a vector of R^n. Then

$$y = \sum_{j=1}^{n} \frac{\det(v_1, ..., v_{j-1}, y, v_{j+1}, \ldots, v_n)}{\det(v_1, \ldots, v_n)} v_j.$$

Method: solving a system using Cramer's formulas

To solve a linear equation of the form $AX = Y$, we start with calculate the determinant of the matrix A which must be non-zero for there to have a unique solution. To calculate the coordinate j of the solution X, we replace column j of A by Y, calculate its determinant and divide by the determinant of A.

Example

Let us solve the following system with Cramer's formulas.

$$\begin{cases} 5x_1 & + & 2x_2 & = & 10 \\ 2x_1 & + & 5x_2 & = & 8 \end{cases}$$

$$x_1 = \frac{\det\begin{pmatrix} 10 & 2 \\ 8 & 5 \end{pmatrix}}{\det\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}} = \frac{10 \times 5 - 8 \times 2}{5 \times 5 - 2 \times 2} = \frac{34}{21},$$

$$x_2 = \frac{\det\begin{pmatrix} 5 & 10 \\ 2 & 8 \end{pmatrix}}{\det\begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}} = \frac{5 \times 8 - 2 \times 10}{5 \times 5 - 2 \times 2} = \frac{20}{21}.$$

Determinant of a linear application

Let us define the determinant of a linear application using its matrix in a certain basis.

Definition

Let E be a finite dimensional vector space and $f : E \mapsto E$ a linear application. The determinant of f is equal to the determinant of its matrix in any base of E. It does not depend on the chosen base.

Let's check that it does not depend on the chosen base. Let (e_i), (e'_i) be two bases of E, A and A' the matrices of f in these bases and P the transition matrix from base (e_i) to base (e'_i). The base change formula gives

$$A' = P^{-1}AP$$

hence $det(A') = det(P)^{-1}det(A)det(P) = det(A)$. **Two conjugate matrices have same determinant**.

Examples

- The determinant of the application identity $id : x \mapsto x$ is equal to 1.
- A dilation $f(x) = \lambda x$ defined on a space of dimension n has a determinant equal to λ^n.
- An isometry defined on a finite dimensional space has a determinant equal to ± 1 because its matrix is orthogonal.

We say that a linear map $f : E \mapsto E$ preserves the orientation, or that it is direct, if its determinant is strictly positive. If the space E is oriented, the image of a direct basis by such an application is direct. An application whose determinant is strictly negative reverses the orientation: the image of a direct base is indirect and vice versa. This is a consequence of the product's formula for the determinant.

Sum of vector subspaces

Let $E_1, ..., E_n$ be vector subspaces of a vector space E. Recall that the sum of these subspaces is defined by

$$E_1 + ... + E_n = \{v_1 + ... + v_n | v_1 \in E_1, ..., v_n \in E_n\}.$$

Definition

Let E be a vector space and $E_1, ..., E_n$ be subspaces vectors of E. We say that these subspaces are in sum direct if all element $x \in E_1 + ... + E_n$ is uniquely written as

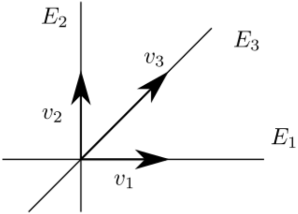
$$x = x_1 + ... + x_n \; with \; x_1 \in E_1, ..., x_n \in E_n.$$

The sum of these subspaces is then denoted $E_1 \oplus ... \oplus E_n = E$.

Example

Consider the subspaces of R^2 given by $E_i = vect(v_i)$ with

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$



We have $E_1 \cap E_2 = E_2 \cap E_3 = E_1 \cap E_3 = \{0\}$ so that any two of these subspaces are in direct sum. On the other hand, the three subspaces are not in direct sum because the vector v_2 admits two decompositions possible in sum of elements of E_1, E_2 and E_3:

$$\begin{array}{ccccccc} v_2 & = & 0 & + & v_2 & + & 0 \\ & = & v_1 & + & 0 & + & v_3. \end{array}$$

Example

The two vector subspaces of $C^0([0,1], R)$ given by

$$E_1 = \{f \in E | \int_0^1 f(t)dt = 0\}, E_2 = \{constant\ \ functions\}$$

are in direct sum: $E = E_1 \oplus E_2$. This follows from taking $E_2 = vect(1_{[0,1]})$ and $E_1 = E_2^{\perp}$, denoting $1_{[0,1]}$ the constant function equal to one in [0,1].

Proposition

Let E be a vector space, $E_1, ..., E_n$ be vector subspaces of finite dimension in direct sum and $B_1...B_n$ the bases of each of E_i. The set of vectors belonging to all these bases form a base of the sum

$$E_1 \oplus ... \oplus E_n$$

and

$$dim E_1 \oplus ... \oplus E_n = dim E_1 + ... + dim E_n.$$

When two spaces are not in sum direct, we can easily give a formula for the dimension of the sum.

Proposition

Let E be a vector space and E_1, E_2 two subspaces vectors of E of finite dimension. Then:

$$dim(E_1 + E_2) = dimE_1 + dimE_2 - dim(E_1 \cap E_2).$$

Proof.

Consider a subspace E' of E_1 which is an additional of $E_1 \cap E_2$. Let us show that E' and E_2 are in sum direct. Let $x \in E' \cap E_2$, the vector x is in E' therefore in E_1 and in $E' \cap E_1 \cap E_2 = \{0\}$. So we have
$E_1 = E' \oplus (E_1 \cap E_2)$, $E_1 + E_2 = E' \oplus E_2$, hence
$dimE_1 + E_2 = dimE' + dimE_2 = dimE_1 - dimE_1 \cap E_2 + dimE_2$. □

Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $E_1, ..., E_n$ the subvector spaces of E orthogonal to each other: for all distinct i, j, for all $x_i \in E_i, x_j \in E_j, < x_i, x_j >= 0$. So these subspaces are in direct sum.

Proof.

Let $x_i \in E_i$ such that $x_1 + ... + x_n = 0$. Let us do the dot product with x_j for all j:

$$0 =< x_1 + ... + x_n, x_j >=< x_1, x_j > +...+ < x_n, x_j >= ||x_j||^2.$$

We conclude that $x_j = 0$ for all j. □

Eigenvalues and eigenvectors

Definition

Let E be a vector space and $f : E \mapsto E$ a linear map. A real number λ is called a real eigenvalue of f if there is a vector $v \in E$ not zero such that $f(v) = \lambda v$.

The vectors satisfying this equality are called eigenvectors of f associated with the eigenvalue λ.
We define in the same way the notion of complex eigenvalue, associated to a non-zero eigenvector $v \in C^n$ which satisfies $f(v) = \lambda v$. Notice that the eigenvectors associated with an eigenvalue λ are exactly the kernel elements of $f - \lambda id$, where $id : E \mapsto E$ is the application identity:$id(x) = x$.

Proposition

Let $f : E \mapsto E$ be a linear map defined on a vector space E of finite dimension n. The characteristic polynomial P_c of f is the polynomial defined on R or C by: $P_c(X) = det(Xid - f)$. A real (or complex) number λ is eigenvalue of f if and only if it is real (or complex) root of the characteristic polynomial.

Examples

1) Identity, defined from R^n to R^n, has a characteristic polynomial

$$P_c(X) = \det(Xid - id) = \det((X-1)id) = (X-1)^n.$$

2) The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial

$$P_c(X) = \det\begin{pmatrix} X & -1 \\ -1 & X \end{pmatrix} = X^2 - 1.$$

It has two real eigenvalues -1 and 1.

3) The matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial

$$P_c(X) = \det \begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix} = X^2 + 1.$$

The matrix has no real eigenvalues. However, it has two complex eigenvalues i and -i.

4) The matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial

$$P_c(X) = \det \begin{pmatrix} X-1 & -1 \\ -1 & X \end{pmatrix} = X^2 - X - 1.$$

The matrix has two real eigenvalues.

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

Corollary

Let f be a linear map defined on a vector space real or complex of finite dimension n and $\lambda_1, ..., \lambda_n$ the complex roots of its characteristic polynomial. Then

$$det(f) = \Pi_{i=1}^{n} \lambda_i$$

We say that the determinant is equal to the product of the eigenvalues, counted with multiplicities, the multiplicity of an eigenvalue referring the number of times it appears in the factored expression of the characteristic polynomial. It should be noted that in the real case, it is not enough to do the product of the real eigenvalues to obtain the determinant in general.

Examples

1) The homothety of R^n given by $x \mapsto \lambda x$ has for characteristic polynomial

$$P_c(X) = \det(Xid - \lambda id) = (X - \lambda)^n = \prod_{i=1}^{n}(X - \lambda).$$

This map has the eigenvalue λ with multiplicity n and its determinant is λ^n

2) The matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is of determinant 1 and has no real eigenvalues.

However, it has two eigenvalues complex i and -i and we have the equality $1 = i \times (-i)$.

$$P_c(X) = X^2 + 1 = (X - i)(X + i).$$

Diagonalisation

Definition

A linear map $f : E \mapsto E$ defined on a space real vector E is said to be diagonalizable over R if there exists a basis of E composed of eigenvectors of f.
Equivalently, f is diagonalizable if one of the two conditions following is carried out,

$$E = \oplus_k ker(f - \lambda_k id), dimE = \sum_k dimker(f - \lambda_k id).$$

Method: Calculation of the eigenvectors of a linear application
It is a question of determining a basis of $ker(f - \lambda id)$. We saw in the first chapter how to proceed. Let A be the matrix of f in a basis of E. We solve the system $(A - \lambda id)v = 0$ by the Gaussian pivot method. The coordinates obtained are those of the eigenvectors in the base considered.

Examples

1) The orthogonal reflection of R^2 to the right led by the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ has for matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and for eigenvalues -1 and 1.

$Ker(A - id) = vect \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $Ker(A + id) = vect \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

2) The rotation of angle $\frac{\pi}{2}$ has the matrix$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, it has no real eigenvectors nor real eigenvectors. Consider A as acting on C^2 rather than R^2. Its complex eigenvalues are i and -i

$Ker(A - iid) = vect \begin{pmatrix} i \\ 1 \end{pmatrix}$ and $Ker(A + iid) = vect \begin{pmatrix} -i \\ 1 \end{pmatrix}$

Proposition

A matrix A of size $n \times n$ is diagonalizable if and only if there exists a diagonal matrix D and an invertible matrix P of $n \times n$ sizes such that

$$A = PDP^{-1}.$$

Proof.

Suppose A is diagonalizable. Let (e_i') be a basis of eigenvectors of A, P the transition matrix from the canonical base to the base (e_i') and D the matrix of A in this base (e_i'). This matrix D is diagonal and by the formula of base change, $A = PDP^{-1}$.
Conversely, if A is of the form PDP^{-1}, denote $(e_i)_i$ the canonical basis of R^n, e_j' the j^{th} column of P and λ_j the j^{th} diagonal coefficient by D.

$$De_j = \lambda_j e_j, e_j = Pe_j', D = P^{-1}AP,$$

we deduce that $Ae_j' = \lambda_j e_j'$. The family (e_j') forms a basis of R^n because the matrix P is invertible. □

Examples:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}^{-1}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}^{-1}$$

Some interesting relations:

$$A^n = PB^nP^{-1} \text{ et } A^{-1} = PB^{-1}P^{-1}.$$

The adjunct of a linear application

Definition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and $f : E \mapsto E$ a linear application. A linear map $f^* : E \mapsto E$ is said to be adjunct to f if

$$\forall x, y \in E, < f(x), y >=< x, f^*(y) > .$$

If E is of finite dimension, then the adjunct f^* exists and is unique. The matrix of the adjunct f^* in an orthonormal basis (e_i) of E is equal to the **transpose** of that of f.

Let us verify that in finite dimension, a map f^* is adjunct to f if and only if its matrix is equal to the transpose of that of f in a given orthonormal base (e_i). Recall that the matrix of f has for coefficients $a_{i,j} =< e_i, f(e_j) >$ in such a basis. The coefficients of the matrix of f^* are then equal to:

$$< e_i, f^*(e_j) >=< f(e_i), e_j >=< e_j, f(e_i) >= a_{j,i}.$$

Conversely, if f^* is the map associated with the transposed matrix of that of f, we have for all $x, y \in E$ with coordinates $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$ in the base (e_i),

$$\langle f(x),y\rangle = \sum_i \Big(\sum_j a_{i,j} x_j\Big) y_i = \sum_j \Big(\sum_i a_{i,j} y_i\Big) x_j = \langle f^*(y),x\rangle.$$

Examples

1-A matrix P is orthogonal if $P^T = P^{-1}$. The adjunct application orthogonal is therefore equal to its inverse.
2- The orthogonal projections are equal to their adjuncts. This is a consequence of the following proposition.

Proposition

A linear map $p : E \mapsto E$ defined on a finite dimensional Euclidean space is an orthogonal projection if and only if it satisfies the following two conditions.

- $p \circ p = p$,
- $p^* = p$.

Proof

Let p be an orthogonal projection on $F = im(p)$ and $x, y \in E$. We know that $p(x) \in F$ and $y - p(y) \in F^{\perp}$. So we have $< p(x), y - p(y) >= 0$ and in reversing the role of x and y,
$< p(x), y >=< p(x), p(y) >=< x, p(y) >$.
Conversely, let $F = im(p)$ and let $x \in E$. The vector p (x) is in F, we must verify that the vector $x - p(x)$ is orthogonal to F. Consider a vector $y \in F$. It is of the form y = p (z) for a certain $z \in E$, which implied $p(y) = p(p(z)) = p(z) = y$. Let us show that $x - p(x)$ is orthogonal to y:

$$\langle x - p(x), y\rangle = \langle x, y\rangle - \langle p(x), y\rangle = \langle x, y\rangle - \langle x, p(y)\rangle = \langle x, y\rangle - \langle x, y\rangle = 0.$$

We show in the same way that an orthogonal symmetry is characterized by the two equalities $s \circ s = id$ and $s^* = s$.

Diagonalization of self-adjunct applications

We will study the problem of diagonalization for the applications of a Euclidean spaces which are equal to their adjunct.

Definition

Let $(E, <\cdot,\cdot>)$ be a finite dimensional Euclidean space and let $f : E \mapsto E$ a linear map. The application f is said to be self-adjunct if

$$\textit{for all } x \in E, < f(x), y >=< x, f(y) > .$$

In other words, f is self-adjunct if it is equal to its adjunct.
Recall that the matrix of the adjunct of a transformation is equal to the transpose of its matrix and that a matrix is symmetrical if it is equal to its transpose. A transformation is therefore self-adjunct if and only if its matrix is symmetrical.

Theorem (diagonalization of autoadjoint mappings)

Let $(E, < \cdot, \cdot >)$ a finite dimensional Euclidean space and $f : E \mapsto E$ an application self-adjunct. Then there exists an orthonormal basis made up of the eigen vectors of f.

We say that f is diagonalizable in orthonormal basis. Proof of theorem is based on two lemmas.

Lemma

The complex eigenvalues of a self-adjunct linear map of a finite dimensional Euclidean space are all real.

Proof

We place ourselves in an orthonormal basis of E and we denote by A the matrix of f. Let us make this matrix act on C^n. Suppose that A admits an eigen value $\lambda \in C$. We can then find an eigenvector $v = (v_1, ..., v_n) \in C^n$ such that $Av = \lambda v$. Let us show that λ is in fact real.
Denote by $\overline{v}$ the vector whose coordinates are the conjugates of those of v. The matrix A has real coefficients and we obtain by conjugation : $Av = \lambda v, A\overline{v} = \overline{\lambda}\overline{v}$ In coordinates, the dot product is defined by: $< x, y >= \sum_i x_i y_i$ for all vectors x, y with real coefficients and we extend it to vectors with complex coefficients using the same equality. Since A is autoadjunct, we have: $< Av, w >=< v, Aw >$ for all vectors v, w with real coefficients. This equality extends to vectors with complex coefficients by bilinearity. Finally,

$$\lambda < v, \overline{v} >=< Av, \overline{v} >=< v, A\overline{v} >= \overline{\lambda} < v, \overline{v} > .$$

We conclude that $\lambda = \overline{\lambda}$ because $< v, \overline{v} >= \sum_i |v_i|^2$is non-zero.

Lemma

Let $(E, < \cdot, \cdot >)$ be a Euclidean space, F a vector subspace of E and f a self-adjunct application. If $f(F) \subset F$ Then $f(F^{\perp}) \subset F^{\perp}$.

Proof.

Let $x \in F^{\perp}$, we must show that for any vector $y \in F$, we have $< f(x), y >= 0$. But $< f(x), y >=< x, f(y) >$ and this last term is zero because f (y) is in $f(F) \subset F$ and x is in $F^{\perp}$.

□

Matrix interpretation

The diagonalization theorem of autoadjoint maps admits an interpretation in terms of matrices.

Theorem

Let S be a symmetric matrix. So there is a matrix orthogonal P and a diagonal matrix D such that $S = PDP^{-1}$. The diagonal terms of D are the eigenvalues of S and P is the passage matrix from the canonical basis to the basis of the eigenvectors of S.

Note that since P is an orthogonal matrix, we have $P^{-1} = P^T$. The matrices S and D are not only conjugate but also congruent.

Method: diagonalization of symmetric matrices

We start by calculating the eigenvalues of the matrix. for each eigenvalue, one finds a base of the associated eigen subspace by using the Gaussian pivot method. This base must then be orthonormalized, for example using the Gram-Schmidt method. Note that if the eigen vector is of dimension one, it suffices to normalize a vector of this space.
The set of vectors belonging to all these bases forms an orthonormal basis from R^n. The desired matrix P has for columns these vectors and the matrix D has for diagonal coefficients the associated eigenvalues.

Lemma

The proper subspaces of an autoadjoint application are orthogonal between them.

Proof.

Let λ_1 and λ_2 be two distinct eigenvalues of f and v_1, v_2 their eigen vectors. We have:

$$\lambda_1 < v_1, v_2 >=< Av_1, v_2 >=< v_1, Av_2 >= \lambda_2 < v_1, v_2 >,$$

which implies $< v_1, v_2 >= 0$. □

Example

Let the symmetric matrix $S = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}$, its characteristic polynomial is:

$$\begin{aligned} P_c(X) &= (X-1)((X-3)(X-1)-4) + 2 \times -2(X-1) \\ &= (X-1)\left(X^2 - 4X - 5\right) \\ &= (X-1)(X+1)(X-5). \end{aligned}$$

The eigen values are: 1,-1, and 5
Their eigen vectors are respectively:

$$vect(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}); vect(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}) \text{ and } vect(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix})$$

The eigen vectors are of dimension one, so it is not necessary to use the Gram-Schmidt orthonormalization method on each of these spaces, just normalize to get the columns of the matrix P. The matrix D has for diagonal coefficients the eigenvalues of B. We obtain:

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}^{-1}$$

Note that the matrix P is easy to invert because its inverse is equal to its transposed.

The matrix D appearing in this theorem has for diagonal coefficients $\lambda_1, ..., \lambda_n$. Consider the diagonal matrix D' whose diagonal coefficients are given by the signs +1 or -1 of the λ for the non-zero λ_i and zero otherwise. Consider also the diagonal matrix D'' whose diagonal coefficients are the roots $\sqrt{|\lambda_i|}$ for non-zero λ_i and 1 for zero λ_i. We then have:

$$D = D''D'D'',$$

$$S = {}^tPDP = {}^tPD''D'D''P = {}^t(D''P)D'(D''P),$$

Theorem (simultaneous reduction)

Let $(E, <\cdot,\cdot>)$ be a Euclidean space of finite dimension and Q a quadratic form defined on E. Then there exists a base (e_i) of E which is both ***orthonormal for the scalar product*** *of E and* ***orthogonal for the quadratic form Q****. In such a basis, Q takes the form:*

$$Q(x_1 e_1 + ... + x_n e_n) = \sum_{k=1}^{n} \lambda_k x_k^2,$$

where the λ_k are the eigenvalues of the matrix of Q in the base (e_i). The signature of Q is given by the sign of these eigenvalues.

Proof

Let B be the matrix of the quadratic form in an **orthonormal basis** (e_i) of E. This matrix is **symmetric**. According to the previous theorem, there exists an **orthogonal matrix P and a diagonal matrix D** such that

$$D = P^{-1}BP = P^TBP.$$

Consider the basis (e'_i) of E whose coordinates of the vectors are given by the columns of P. The matrix P is the transition matrix from the base (e_i) to the base (e'_i). The matrix P is orthogonal therefore the base (e'_i) is orthonormal.
The matrix of the quadratic form Q in this basis is equal to D, by virtue of the base change formula for quadratic forms. This basis is therefore orthogonal for Q.

Example

Let us place ourselves on R^3 provided with its usual scalar product and consider the quadratic form:

$$Q(x_1, x_2, x_3) = x_1^2 + 3x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3.$$

its corresponding matrix is: $B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}$

Its eigenvalues are equal to 1, -1 and 5. Two eigenvalues are strictly positive and one is strictly negative. The signature of Q is therefore (2,1). The base is formed of the following eigenvectors:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

It is orthonormal for the usual scalar product of R^3 and orthogonal for the quadratic form Q. All you have to do is make the change of variables

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

to simplify the expression of Q:

$$\begin{aligned} Q(x_1,x_2,x_3) &= x_1^2 + 3\,x_2^2 + x_3^2 + 4\,x_1\,x_2 + 4\,x_2\,x_3 \\ &= y_1^2 - y_2^2 + 5y_3^2. \end{aligned}$$

Comparison of Euclidean norms

Proposition

Let $(E, < \cdot, \cdot >)$ be a Euclidean space and Q a quadratic form defined on E. Let $\lambda_1, ...\lambda_k$ be the eigenvalues of the matrix of Q in an orthonormal basis for $< \cdot, \cdot >$. Then for all $x \in E$,

$$(\min_k \lambda_k)||x||^2 \leq Q(x) \leq (\max_k \lambda_k)||x||^2$$

Proof.

It suffices to place oneself in a base (e_i) of simultaneous diagonalization.

$$Q(x_1 e_1 + ... + x_n e_n) = \lambda_1 x_1^2 + ... + \lambda_n x_n^2 \leq (\max_k \lambda_k)(x_1^2 + ... + x_n^2).$$

We proceed in a similar way for the other inequality. □

Example

Consider the quadratic form defined on R^2 by:

$$Q(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2.$$

It has for matrix $\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$ in the canonical basis and for the characteristic polynomial :

$$P_c(X) = (X-1)^2 - (1/2)^2 = (X-1/2)(X-3/2).$$

The roots of this polynomial are $1/2$ and $3/2$. We deduce the framing

$$\frac{1}{2}(x_1^2 + x_2^2) \leq x_1^2 + x_1x_2 + x_2^2 \leq \frac{3}{2}(x_1^2 + x_2^2).$$

Geometric characteristics of conics

Proposition

Consider a conic section C of R^2 associated with a non-degenerated quadratic form: $b^2 - 4ac \neq 0$. Then there exists a change of variables of the form

$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + b_2 \end{cases}$$

with $\{a_{i,j}\}$ orthogonal matrix, which transforms the expression of the conic into:

$$\lambda_1 X^2 + \lambda_2 Y^2 = \mu.$$

The coefficients λ_1 and λ_2 are the eigenvalues of the matrix associated with the quadratic form $ax^2 + bxy + cy^2$.

As the change of variable is associated with an orthogonal matrix, it preserves the Euclidean distance and it is possible to calculate the characteristics of the conic from its reduced expression.

For an ellipse for example, the length of the major axis and minor axis are given by $a = \mu/\sqrt{\lambda_1}$ and $b = \mu/\sqrt{\lambda_2}$.

The coefficients μ, λ_1 and λ_2 being determined only by a multiplicative factor, it is natural to introduce a metric invariant based on the quotient of two of these numbers.

Definition

Even if it means inverting the variables X and Y in the reduced form of a conic and changing its sign, we can assume $\lambda_1 \geq \lambda_2$ and $\lambda_1 \geq 0$.
The eccentricity $e \in [0, \infty[$ of a conic section is then defined by
$e = \sqrt{1 - \frac{\lambda_2}{\lambda_1}}$.

The eccentricity of an ellipse is strictly less than 1, that of a hy- perbole strictly greater than 1. By convention, we set e = 1 for the parables. An ellipse close to a circle has an eccentricity close to 0.

Theorem of d'Alembert-Gauss

Theorem

Any non-constant polynomial defined on C with complex coefficients admits a complex root.

The Alembert-Gauss theorem implies that any complex polynomial is a product of polynomials of degree one.

Proposition

Let $P(X)$ be a polynomial with complex coefficients of degree $n > 0$. Then there are complex numbers $\lambda_1, ..., \lambda_n$, not necessarily distinct, such as

$$P(X) = (X - \lambda_1)(X - \lambda_2)...(X - \lambda_n).$$

Prehilbertian spaces

Definition

Let E be a vector space defined on C. A hermitian on E is a map $\phi : E \times E \mapsto C$ which satisfies for all z, v, w in E and $\lambda \in C$,

- $\varphi(\lambda v + w, z) = \lambda\,\varphi(v, z) + \varphi(w, z)$,
- $\varphi(z, \lambda v + w) = \varphi(z, v) + \bar{\lambda}\,\varphi(z, w)$,
- $\varphi(v, w) = \overline{\varphi(w, v)}$.

The associated quadratic form $Q(z) = \phi(z, z)$ has real values because of the complex conjugation that appears in the definition.

Examples

1- On C^n, the usual Hermitian product is defined by:

$$\varphi(z,w) = \sum_{i=1}^{n} z_i \overline{w_i} = z_1 \overline{w_1} + ... + z_n \overline{w_n}.$$

for all $z = (z_1, ...z_n) \in C^n, w = (w_1, ...w_n) \in C^n$ the associated quadratic form is given by:

$$Q(z) = \sum_{i=1}^{n} |z_i|^2 = |z_1|^2 + ... + |z_n|^2.$$

2-On $C^0([0,1], C)$, the usual Hermitian product is defined by:

$$\varphi(f,g) = \int_0^1 f(t)\overline{g(t)}\,dt.$$

The associated quadratic form is given by:

$$Q(f) = \int_0^1 |f(t)|^2\,dt.$$

To each Hermitian form is associated a matrix B with complex coefficients satisfying $B^T = B$. The form ϕ is expressed as a function of this matrix by the formula

$$\varphi(x,y) = {}^t z\, B\, \overline{w} = \sum_{i,j} z_i\, b_{i,j}\, \overline{w_j} = b_{1,1}\, z_1\overline{w_1} + b_{1,2}\, z_1\overline{w_2} + \ldots + b_{n,n}\, z_n\overline{w_n}.$$

The theory of Hermitian forms proceeds analogously to the case real. In particular, we have a notion of signature and a Gauss algorithm which is based on the following rules.

$$: |z|^2 + 2re(z\overline{a}) = |z+a|^2 - |a|^2.$$

$$: z\overline{w} + z\overline{a_1} + a_2\overline{w} = (z+a_2)\overline{(w+a_1)} - a_2\overline{a_1}.$$

$$: 4re(z\overline{w}) = |z+\overline{w}|^2 - |z-\overline{w}|^2.$$

Proposition

Let E be a complex vector space and let Q be a Hermitian quadratic form of signature (p, q). Then there exist $p + q$ forms $\mathbb{C}$-independent linear $l_1, ..., l_{p+q}$ defined on E such that

$$Q(z) = \sum_{i=1}^{p} |l_i(z)|^2 - \sum_{j=p+1}^{p+q} |l_j(z)|^2.$$

Note the module which appears in the previous proposition. The l_i are complex and linear forms over C.

Examples

– $\varphi(z,w) = z_1\overline{w_1} + 3z_2\overline{w_2} + 2z_1\overline{w_2} + 2z_2\overline{w_1}$.

Its matrix gives $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. The Gauss algorithm gives

$$\begin{aligned} Q(z) &= |z_1|^2 + 3|z_2|^2 + 2z_1\overline{z_2} + 2z_2\overline{z_1} \\ &= |z_1|^2 + 3|z_2|^2 + 4Re(z_1\overline{z_2}) \\ &= |z_1 + 2z_2|^2 - |z_2|^2. \end{aligned}$$

– $Q(z,w) = z_1\overline{w_1} + 3z_2\overline{w_2} + 2iz_1\overline{w_2} - 2iz_2\overline{w_1}$

Its matrix gives $\begin{pmatrix} 1 & 2i \\ -2i & 3 \end{pmatrix}$. The Gauss algorithm gives

$$\begin{aligned} Q(z) &= |z_1|^2 + 3|z_2|^2 + 2iz_1\overline{z_2} - 2iz_2\overline{z_1} \\ &= |z_1|^2 + 3|z_2|^2 + 4Re(iz_1\overline{z_2}) \\ &= |z_1 + 2iz_2|^2 - |z_2|^2. \end{aligned}$$

Definition

A prehilbertian space is a complex vector space provided with a Hermitian form whose associated quadratic form is positive definite. Such a shape is called a Hermitian dot product.

Let us use the notation $< \cdot, \cdot >$ for a Hermitian dot product. Inequality of Cauchy-Schwarz becomes:

$$\forall z, w \in E, |< z, w >|^2 \leq Q(z)Q(w)$$

and we can define the notion of orthogonal projection as in the real case.

Definition

Let E be a prehilbertian space and $f : E \mapsto E$ a C-linear map. We say that f admits an adjunct if there exists a C-linear mapping $f^* : E \mapsto E$ which satisfies for all $z, w \in E$,

$$< f(z), w >=< z, f^*(w) > .$$

The application f is said
- Hermitian if $f = f^*$,
- unitary if $f^* f = id$,
- normal if $ff^* = f^* f$.

In finite dimension, any map f has an adjunct. If A is the matrix of f in an orthonormal basis, its adjunct has for matrix $\overline{A}^T$ We pose:

$$A^* = \overline{A}^T.$$

The notion of Hermitian application generalizes the notion of application to autoadjointe in the real case. The concept of unitary application generalizes that of orthogonal application. The novelty is the notion of normal application.

Note that the Hermitian and unitary applications are normal.

Theorem (spectral theorem)

Let E be a prehilbertian space and let $f : E \mapsto E$ a normal map. Then f is diagonalizable in orthonormal basis.

In particular, the unit matrices are diagonalizable on C then that the orthogonal matrices are not diagonalizable on R. The version matrix of this theorem gives

Theorem

Let A be a matrix with complex coefficients satisfying:$A^*A = AA^*$. We say that such a matrix is normal. Then there is a diagonal matrix D and a unitary matrix U such that:

$$A = UDU^*.$$

If A is Hermitian, the diagonal coefficients of D are real. If A is unitary, the diagonal coefficients of D have modulus of one.

In particular, any matrix with real coefficients which is orthogonal is unitary therefore diagonalizable on C.
The proof in the Hermitian case is similar to the real case. For dies normal, we decompose them in the form $A' + iA''$ with A' and A'' matrices Hermitians which commute, by setting $A' = \frac{1}{2}(A + A^*)$ and $A'' = \frac{1}{2i}(A - A^*)$ As these two matrices commute, the eigen vectors of A' are invariants by A''. It becomes possible to diagonalize A' then to diagonalize A'' on each of the proper subspaces of A'' so as to find a base which simultaneously diagonalizes the two matrices.

Example

Diagonalization of the rotations of the Hermitian plane.

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}\frac{\sqrt{2}}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$