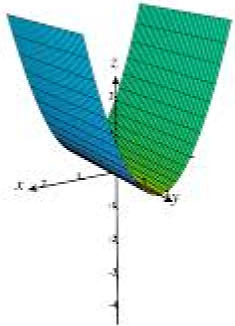
Linear and Bilinear Algebra I



DR Lama Tarsissi
Sorbonne University of Abu Dhabi
MATH222

Ch2. Quadratic forms

Direct sum

Theorem

The vector space V is the direct sum of its subsapces U and W if and only if (i) $V = U + W$ and (ii) $U \cap W = \{0\}$.

Direct sum

Theorem

The vector space V is the direct sum of its subsapces U and W if and only if (i) $V = U + W$ and (ii) $U \cap W = \{0\}$.

Example

1. $\mathbb{R}^3 = U \oplus W$; $U = \{(a, b, 0); a, b \in \mathbb{R}\}$ and $W = \{(0, 0, c); c \in \mathbb{R}\}$.

Direct sum

Theorem

The vector space V is the direct sum of its subsapces U and W if and only if (i) $V = U + W$ and (ii) $U \cap W = \{0\}$.

Example

1. $\mathbb{R}^3 = U \oplus W$; $U = \{(a, b, 0); a, b \in \mathbb{R}\}$ and $W = \{(0, 0, c); c \in \mathbb{R}\}$.
2. Let E be a vector space of finite dimension, and let F and G be two sub-spaces of E such that: $F \cap G = \{0\}$ and $dim(F) + dim(G) = dim(E)$. We have $E = F \oplus G$.

Direct sum

Theorem

The vector space V is the direct sum of its subsapces U and W if and only if (i) $V = U + W$ and (ii) $U \cap W = \{0\}$.

Example

1. $\mathbb{R}^3 = U \oplus W$; $U = \{(a, b, 0); a, b \in \mathbb{R}\}$ and $W = \{(0, 0, c); c \in \mathbb{R}\}$.
2. Let E be a vector space of finite dimension, and let F and G be two sub-spaces of E such that: $F \cap G = \{0\}$ and $dim(F) + dim(G) = dim(E)$. We have $E = F \oplus G$.
3. $\mathbb{R}^3 = L \oplus P$; $P = \{(x, y, z) \in \mathbb{R}^3; x - y + 2z = 0\}$ and $L = \{t(1, -1, 2); t \in \mathbb{R}\}$

Proof of example 3

L is the normal line of P. Note that a plane and its normal line clearly only intersect in one point, and in this case both pass through the origin, so clearly $L \cap P = \{0\}$. Checking that $\mathbb{R}^3 = L + P$ is the same as checking that for any $(x, y, z) \in \mathbb{R}^3$, there is a point on the line, say $(a, -a, 2a)$, and a a point on the plane, say of the form $(b, b + 2c, c)$, for which

$$(x, y, z) = (a, -a, 2a) + (b, b + 2c, c) = (a + b, -a + b + 2c, 2a + c).$$

This is the same as solving $AX = d$, where $d = \begin{pmatrix} x \\ y \\ z) \end{pmatrix}$ and

$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}$. Since we want to show that this is always solvable, for any choice (x, y, z), this is equivalent to A being invertible, or $detA = 6 \neq 0$.

diagonalisation

Definition

A square $n \times n$ matrix A over a field F is called diagonalizable or nondefective if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. Formally,

$$\boxed{A \in F^{n\times n} \text{ diagonalizable} \iff \exists\, P, P^{-1} \in F^{n\times n} : \; P^{-1}AP \text{ diagonal}}$$

diagonalisation

Definition

A square $n \times n$ matrix A over a field F is called diagonalizable or nondefective if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. Formally,

$$\boxed{A \in F^{n\times n} \text{ diagonalizable} \iff \exists\, P, P^{-1} \in F^{n\times n} : \; P^{-1}AP \text{ diagonal}}$$

In order to get the matrix P, we need to compute the characteristic polynomial wich is: $P_A(\lambda) = det(A - \lambda I)$, where λ is an eigen value for A, or simply a root of $P(A)$.

Example

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

The roots of the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$ are the eigenvalues $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 2$. Solving the linear system $(I - A)\,\mathrm{v} = 0$ gives the eigenvectors $\mathrm{v}_1 = (1, 1, 0)$ and $\mathrm{v}_2 = (0, 2, 1)$, while $(2I - A)\,\mathrm{v} = 0$ gives $v3 = (1, 0, -1)$; that is, $A\mathrm{v}_i = \lambda_i \mathrm{v}_i$ for $i = 1, 2, 3$. These vectors form a basis of $V = \mathbb{R}^3$, so we can assemble them as the column vectors of a change of basis matrix P to get:

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & -2 \\ 0 & 1 & 0 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = D.$$

Application on direct sum

Theorem

Let $W_1, W_2, \ldots W_k$ be subspaces of a finite dimensional vector space V. The following conditions are equivalent:

1. $V = W_1 \oplus W_2 \oplus \ldots \oplus_k^W$.
2. $V = \sum_{i=1}^k W_i$, *and for any vectors* $V_1, V_2, \ldots V_k$, *such that* $V_i \in W_i (1 \leq i \leq k)$, *if* $v_1 + v_2 + \ldots v_k = 0$ *then* $v_i = 0$ *for all* i.
3. *Each vector* $v \in V$ *can be uniquely written as* $v = v_1 + v_2 + \ldots + v_k$ *where* $v_i \in W_i$.
4. *If* v_i *is an ordered basis for* $W_i, (1 \leq i \leq k)$, *then* $v_1 \cup v_2 \cup \ldots \cup v_k$ *is an ordered basis for* V.
5. *For each* $i = 1 \ldots, k$, *there exists an ordered basis* V_i *for* W_i *such that* $v_1 \cup v_2 \cup \ldots \cup v_k$ *is an ordered basis for* V

Application on direct sum

Theorem

Let $W_1, W_2, \dots W_k$ be subspaces of a finite dimensional vector space V. The following conditions are equivalent:

1. $V = W_1 \oplus W_2 \oplus \dots \oplus_k^W$.
2. $V = \sum_{i=1}^k W_i$, *and for any vectors* $V_1, V_2, \dots V_k$, *such that* $V_i \in W_i (1 \leq i \leq k)$, *if* $v_1 + v_2 + \dots v_k = 0$ *then* $v_i = 0$ *for all* i.
3. *Each vector* $v \in V$ *can be uniquely written as* $v = v_1 + v_2 + \dots + v_k$ *where* $v_i \in W_i$.
4. *If* v_i *is an ordered basis for* $W_i, (1 \leq i \leq k)$, *then* $v_1 \cup v_2 \cup \dots \cup v_k$ *is an ordered basis for* V.
5. *For each* $i = 1 \dots, k$, *there exists an ordered basis* V_i *for* W_i *such that* $v_1 \cup v_2 \cup \dots \cup v_k$ *is an ordered basis for* V

Homework graded.

One more example

Theorem

A linear operator T on a finite -dimensional vector space V is diagonalisable if and only if V is the direct sum of the eigen spaces of T.

One more example

Theorem

A linear operator T on a finite -dimensional vector space V is diagonalisable if and only if V is the direct sum of the eigen spaces of T.

Proof.

Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigen values of T. First suppose that T is diagonalisable and for each i choose an ordered basis V_i for the eigenspace E_{λ_i}. By definition, $v_1 \cup \ldots \cup v_k$ is a basis for V, and hence from the previous theorem V is a direct sum of $E'_{\lambda_i}s$.
Conversely, suppose that V is a direct sum of the eigen spaces of T. For each i, choose an ordered basis V_i for E_{λ_i}. $V = E_{\lambda_\mathbf{1}} \oplus \ldots \oplus E_{\lambda_k}$. By the previous theorem, the union is a basis for V. Since each eigenspace E_{λ_i} consists of zero vector and the eigenvectors of T corresponding to the eigenvalue λ_i. Thus V_i consist of only eigenvector because $0 \notin v_i$, since v_i is a basis for eigenspace E_{λ_i}.
Hence the basis $v_1 \cup v_2 \cup \ldots \cup v_k$ consists of eigen vectors of T, we conclude that T is diagonalisable. □

Linear form

Definition

A linear form is a linear map from a vector space E to its field of scalars (often the real set $\mathbb{R}$ or complex set $\mathbb{C}$).

Linear form

Definition

A linear form is a linear map from a vector space E to its field of scalars (often the real set $\mathbb{R}$ or complex set $\mathbb{C}$).

A linear form is a function $\ell : E \longrightarrow \mathbb{R}$, which satisfies the following: $\forall x, y \in \ E$ and $\lambda \in \mathbb{R}$,

$$l(x + \lambda y) = l(x) + \lambda l(y).$$

Linear form

Definition

A linear form is a linear map from a vector space E to its field of scalars (often the real set $\mathbb{R}$ or complex set $\mathbb{C}$).

A linear form is a function $\ell : E \longrightarrow \mathbb{R}$, which satisfies the following: $\forall x, y \in \; E$ and $\lambda \in \mathbb{R}$,

$$l(x + \lambda y) = l(x) + \lambda l(y).$$

Definition

The set of all linear forms over a vector space E, $\ell(E, k)$, is the dual vector space and is denoted by E^*.

Examples

- Let E be any vector space. The null function, which associates every vector of E to the real 0 is a linear form.
- Let $a_1, a_2 \in \mathbb{R}$. The function $l(x_1, x_2) = a_1 x_1 + a_2 x_2$ is a linear form defined on $\mathbb{R}^2$.
- Note $C^0([0,1], \mathbb{R}])$, the vector space of continuous functions defined on the interval $[0,1]$ to reel values. The application

$$f \to \int_0^1 f(t)dt$$

is a linear form over $C^0([0,1], \mathbb{R}])$.

Proposition

The linear forms defined over $\mathbb{R}^n$ are the functions of the form:

$$l(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n.$$

The real numbers $a_1, \ldots a_n$ are the coefficients of the linear form l.
Let us check that any linear form is indeed of this form. For that, let $a_1 = l(1, 0, \ldots, 0), \ldots, a_n = l(0, 0, \ldots, 1)$. By linearity, $l(x_1, \ldots x_n) = x_1 l(1, 0, \ldots, 0) + ... + x_n l(0, 0, \ldots, 1) = a_1x_1 + \ldots + a_nx_n$.

kernel of a linear form

The kernel of a linear form defined over E is a vector subspace of E. We will characterize these subspaces by their dimension.

kernel of a linear form

The kernel of a linear form defined over E is a vector subspace of E. We will characterize these subspaces by their dimension.

Definition

Let E be a vector space of finite dimension n. A vector subspace of dimension $n - 1$ is called the hyperplane of E.

kernel of a linear form

The kernel of a linear form defined over E is a vector subspace of E. We will characterize these subspaces by their dimension.

Definition

Let E be a vector space of finite dimension n. A vector subspace of dimension $n-1$ is called the hyperplane of E.

Example

1. If E is of dimension 2, the hyperplanes of E are lines.

kernel of a linear form

The kernel of a linear form defined over E is a vector subspace of E. We will characterize these subspaces by their dimension.

Definition

Let E be a vector space of finite dimension n. A vector subspace of dimension $n-1$ is called the hyperplane of E.

Example

1. If E is of dimension 2, the hyperplanes of E are lines.
2. If E is of dimension 3, the hyperplanes of E are planes.

kernel of a linear form

The kernel of a linear form defined over E is a vector subspace of E. We will characterize these subspaces by their dimension.

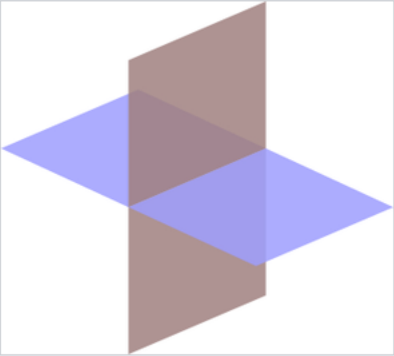
Definition

Let E be a vector space of finite dimension n. A vector subspace of dimension $n - 1$ is called the hyperplane of E.

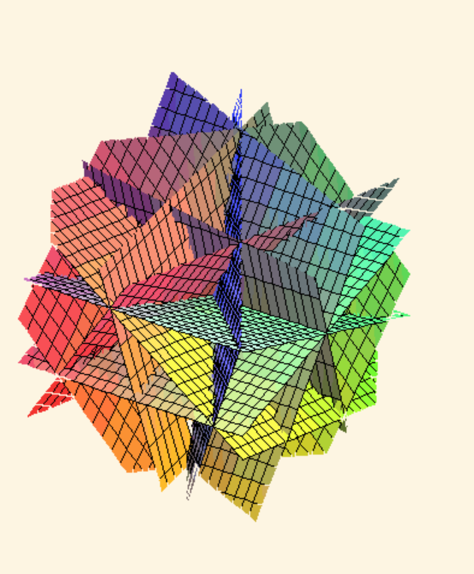
Example

1. If E is of dimension 2, the hyperplanes of E are lines.
2. If E is of dimension 3, the hyperplanes of E are planes.

From a geometric point of view, **vectors** direct vector lines.
Linear shapes direct hyperplanes.



Two intersecting planes in three-dimensional space. A plane is a hyperplane of dimension 2, when embedded in a space of dimension 3.



More about hyperplanes

Proposition

Let E be a vector space of finite dimension n.

- *The kernel of a non-zero linear form is a hyperplane.*
- *Any hyperplane is the kernel of a non-zero linear form.*
- *Two non-zero linear forms define the same hyperplane if and only -ment if they are proportional.*

proof

- The first point follows from the rank formula. Let l be a linear form.

  $$dimker(l) + rank(l) = n.$$

  If l is zero, its image is equal to $\{0\}$ and $rank(l) = 0$. If l is non-zero, its image is equal to $\mathbb{R}$ and $rank(l) = 1$.

proof

- The first point follows from the rank formula. Let l be a linear form.

$$dimker(l) + rank(l) = n.$$

  If l is zero, its image is equal to $\{0\}$ and $rank(l) = 0$. If l is non-zero, its image is equal to $\mathbb{R}$ and $rank(l) = 1$.
- Let H be a hyperplane of E and $e_1, \ldots, e_{n-1}$ a basis of H. Let e_n be a vector of E which is not in H. It is not a combination of $e_1, \ldots, e_{n-1}$, the family $(e_1, \ldots, e_{n-1}, e_n)$ is therefore free, it is a basis of E. A vector $x = x_1 e_1 + \ldots + x_n e_n$ of E is in H if and only if it is expressed in function of $e_1, \ldots, e_{n-1}$, that is to say if and only if $x_n = 0$. It is therefore the kernel of the linear form defined by $l(x_1 e_1 + \ldots + x_n e_n) = x_n$.

proof

- The first point follows from the rank formula. Let l be a linear form.

$$dimker(l) + rank(l) = n.$$

If l is zero, its image is equal to $\{0\}$ and $rank(l) = 0$. If l is non-zero, its image is equal to $\mathbb{R}$ and $rank(l) = 1$.

- Let H be a hyperplane of E and $e_1, \ldots, e_{n-1}$ a basis of H. Let e_n be a vector of E which is not in H. It is not a combination of $e_1, \ldots, e_{n-1}$, the family $(e_1, \ldots, e_{n-1}, e_n)$ is therefore free, it is a basis of E. A vector $x = x_1 e_1 + \ldots + x_n e_n$ of E is in H if and only if it is expressed in function of $e_1, \ldots, e_{n-1}$, that is to say if and only if $x_n = 0$. It is therefore the kernel of the linear form defined by $l(x_1 e_1 + \ldots + x_n e_n) = x_n$.
- Let $l_1, l_2 \in E^*$ be non-zero such that $H = ker(l_1) = ker(l_2)$. In the base of E just considered, $l_1(e_i) = l_2(e_i) = 0$ if $i < n$. We can deduce

$$l_1(\sum x_i e_i) = x_n l_1(e_n), l_2(\sum x_i e_i) = x_n l_2(e_n),$$

hence $l_1 = \frac{l_1(e_n)}{l_2(e_n)} l_2$. These two forms are quite proportional.

Proposition

The coordinates of a linear form $l \in E^*$ in the base $(e_1^*, \ldots, e_n^*)$ are worth

$$l(e_1), l(e_2), \ldots, l(e_n).$$

The space E^* is of finite dimension and $dim E^* = dim E$ H.W.

Proposition

The coordinates of a linear form $l \in E^*$ in the base $(e_1^*, \dots, e_n^*)$ are worth

$$l(e_1), l(e_2), \dots, l(e_n).$$

The space E^* is of finite dimension and $dimE^* = dimE$ H.W.

We agree to represent the coordinates of a vector in a base given (e_i) of E in the form of a **column vector** and to represent the coordinates of a linear form in the dual basis (e_i^*) of E^* in the form of **line vector**. With this convention, if $x \in E$ and $l \in E^*$ have coordinates: $\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ and $(a_1, \dots, a_n)$. then the quantity $l(x)$ is obtained by taking the product of the two vectors:

$$l(x) = (a_1, \dots, a_n) \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = a_1 x_1 + \dots a_n x_n.$$

Solution

Let us show first that B^* is a free family of E^*. Suppose that the null linear form can be written as: $0 = \lambda_1 e_1^* + \ldots + \lambda_n e_n^*$; we prove that the coefficiants are necessarily null by applying the linear form defined by the second member on e_i: we get $0 = (\lambda_1 e_1^* + \ldots + \lambda_n e_n^*)(e_i) = \lambda_i$ for all $1 \leq i \leq n$.

Now let us show that B^* is a generator family of E^*. Let f be a linear form of E^*. We let $\lambda_i = f(e_i)$ for all $1 \leq i \leq n$ and $g = \lambda_1 e_1^* + \ldots + \lambda_n e_n^*$. The linear foms f and g are equal over the vectors of the base B because $g(e_i) = \lambda_i = f(e_i)$ for all $1 \leq i \leq n$. Consequently, $f = g$, which concludes our proof.

Case of $\mathbb{R}^n$, Canonical basis

Recall what is the canonical basis of $\mathbb{R}^n$. With the convention that consists in representing the vectors of $\mathbb{R}^n$ in the form of columns, it is a question of vector family:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Case of $\mathbb{R}^n$, Canonical basis

Recall what is the canonical basis of $\mathbb{R}^n$. With the convention that consists in representing the vectors of $\mathbb{R}^n$ in the form of columns, it is a question of vector family:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Any vector of $\mathbb{R}^n$ is expressed in this basis as follows.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

We agreed to represent the elements of the dual $\mathbb{R}^{n*}$ in the form of vectors lines. The elements of the dual basis of the canonical basis are then $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$.

We agreed to represent the elements of the dual $\mathbb{R}^{n*}$ in the form of vectors lines. The elements of the dual basis of the canonical basis are then $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$.
Any vector of $\mathbb{R}^{n*}$ is then expressed in this basis as follows.

$$\begin{aligned}
\begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix} = \quad & x_1 \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} \\
+ \, & x_2 \begin{pmatrix} 0 & 1 & \ldots & 0 \end{pmatrix} \\
+ \, & \ldots \\
+ \, & x_n \begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix}.
\end{aligned}$$

We agreed to represent the elements of the dual $\mathbb{R}^{n*}$ in the form of vectors lines. The elements of the dual basis of the canonical basis are then $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$.
Any vector of $\mathbb{R}^{n*}$ is then expressed in this basis as follows.

$$\begin{aligned} \begin{pmatrix} x_1 & x_2 & \ldots & x_n \end{pmatrix} = & \ \ x_1 \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} \\ & + x_2 \begin{pmatrix} 0 & 1 & \ldots & 0 \end{pmatrix} \\ & + \ \ldots \\ & + x_n \begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix}. \end{aligned}$$

The **dual basis** of the canonical basis of $\mathbb{R}^n$ is called the **canonical basis** of $\mathbb{R}^{n*}$.

Independency

Linear shapes are the elements of a vector space. So we can be interested in the independence of a family of linear shapes.

Independency

Linear shapes are the elements of a vector space. So we can be interested in the independence of a family of linear shapes.
Remember that a family of vectors is **linearly independent** if there is **no non-trivial** linear relation between these vectors.

Independency

Linear shapes are the elements of a vector space. So we can be interested in the independence of a family of linear shapes.
Remember that a family of vectors is **linearly independent** if there is **no non-trivial** linear relation between these vectors.
Method: independence of a family of linear shapes.

Independency

Linear shapes are the elements of a vector space. So we can be interested in the independence of a family of linear shapes.
Remember that a family of vectors is **linearly independent** if there is **no non-trivial** linear relation between these vectors.
Method: independence of a family of linear shapes.
To determine if a family of linear shapes $(l_1, \ldots, l_k)$ is linearly independent, we decompose this family in a base and we form the matrix whose rows are given by the row vectors associated with each linear shapes. The family is linearly independent if and only if **the rank of this matrix is equal to k**

Example

We consider the three linear forms defined on $\mathbb{R}^3$ by:

$$\begin{aligned} l_1(x_1,x_2,x_3) &= x_1 + x_2, \\ l_2(x_1,x_2,x_3) &= x_2 + x_3, \\ l_3(x_1,x_2,x_3) &= x_1 - x_3. \end{aligned}$$

Example

We consider the three linear forms defined on $\mathbb{R}^3$ by:

$$\begin{aligned} l_1(x_1,x_2,x_3) &= x_1 + x_2, \\ l_2(x_1,x_2,x_3) &= x_2 + x_3, \\ l_3(x_1,x_2,x_3) &= x_1 - x_3. \end{aligned}$$

The coordinates of these linear forms in the canonical basis are $(1,1,0),(0,1,1),(1,0,-1)$.

The matrix obtained by stacking these rows is equal to $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$.

We determine its rank by putting it in echelon form:

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix} L_3 \leftarrow L_3 - L_1 \qquad \begin{pmatrix} \mathbf{1} & 1 & 0 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{pmatrix} L_3 \leftarrow L_3 + L_2$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix}_{L_3 \leftarrow L_3 - L_1} \qquad \begin{pmatrix} \mathbf{1} & 1 & 0 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{pmatrix}_{L_3 \leftarrow L_3 + L_2}$$

It is of rank 2.
The family is therefore not linearly independent.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix}_{L_3 \leftarrow L_3 - L_1} \qquad \begin{pmatrix} \mathbf{1} & 1 & 0 \\ 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 \end{pmatrix}_{L_3 \leftarrow L_3 + L_2}$$

It is of rank 2.
The family is therefore not linearly independent.
We could have gone faster by noticing directly that there is a relationship nontrivial linear between the linear forms l_1, l_2 and l_3: $l_1 = l_2 + l_3$.

Link with linear systems

Consider a homogeneous linear system on $\mathbb{R}^n$ of the form

$$\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & 0 \end{array}\right.$$

Link with linear systems

Consider a homogeneous linear system on $\mathbb{R}^n$ of the form

$$\begin{cases} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & 0 \end{cases}$$

Let us define linear forms l_i by setting

$$l_i(x) = a_{i,1}x_1 + \dots + a_{i,n}x_n.$$

Link with linear systems

Consider a homogeneous linear system on $\mathbb{R}^n$ of the form

$$\begin{cases} a_{1,1}x_1 & + & \dots & + & a_{1,n}x_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & \dots & + & a_{m,n}x_n & = & 0 \end{cases}$$

Let us define linear forms l_i by setting

$$l_i(x) = a_{i,1}x_1 + \dots + a_{i,n}x_n.$$

We see that a vector is a solution of the system if and only if it annihilates all these linear forms, that is, if it belongs to $ker(l_i)$ for all i.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

Consider the case $n = 3$ and assume the hyperplanes are all distinct.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

Consider the case $n = 3$ and assume the hyperplanes are all distinct.

- The system is of rank three.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

Consider the case $n = 3$ and assume the hyperplanes are all distinct.

- The system is of rank three.
- Only the zero vector is solution of the system and the only three hyperplanes meet at the origin.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

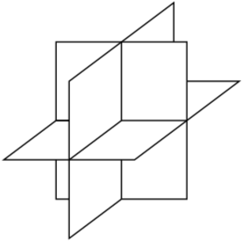
Consider the case $n = 3$ and assume the hyperplanes are all distinct.

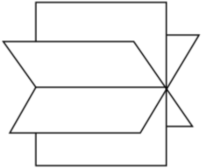
- The system is of rank three.
- Only the zero vector is solution of the system and the only three hyperplanes meet at the origin.
- If it's rank two, the three hyperplanes meet on a straight line.

Geometrically, the vector space of the solutions of the system is an **intersection of hyperplanes**.

Consider the case $n = 3$ and assume the hyperplanes are all distinct.

- The system is of rank three.
- Only the zero vector is solution of the system and the only three hyperplanes meet at the origin.
- If it's rank two, the three hyperplanes meet on a straight line.





Bilnear form

Definition

Let E be a real vector space. A **bilinear form** is a function defined on $E \times E$ with real values which is linear in each of its variables.

Bilnear form

Definition

Let E be a real vector space. A **bilinear form** is a function defined on $E \times E$ with real values which is linear in each of its variables.

It is a function $\phi : E \times E \to \mathbb{R}$ which satisfies, for all $x, y, z \in E$ and $\lambda \in \mathbb{R}$,

- $\phi(\lambda x + y, z) = \lambda \phi(x, z) + \phi(y, z)$

Bilnear form

Definition

Let E be a real vector space. A **bilinear form** is a function defined on $E \times E$ with real values which is linear in each of its variables.

It is a function $\phi : E \times E \to \mathbb{R}$ which satisfies, for all $x, y, z \in E$ and $\lambda \in \mathbb{R}$,

- $\phi(\lambda x + y, z) = \lambda\phi(x, z) + \phi(y, z)$
- $\phi(x, \lambda y + z) = \lambda\phi(x, y) + \phi(x, z)$

Bilnear form

Definition

Let E be a real vector space. A **bilinear form** is a function defined on $E \times E$ with real values which is linear in each of its variables.

It is a function $\phi : E \times E \to \mathbb{R}$ which satisfies, for all $x, y, z \in E$ and $\lambda \in \mathbb{R}$,

- $\phi(\lambda x + y, z) = \lambda\phi(x, z) + \phi(y, z)$
- $\phi(x, \lambda y + z) = \lambda\phi(x, y) + \phi(x, z)$

Definition

A bilinear form is symmetric if for all $x, y \in E$

$$\phi(x, y) = \phi(y, x)$$

Examples

- Over $\mathbb{R}$, the bilinear forms are of the form $\phi(x, y) = axy$ for $a \in \mathbb{R}$.

Examples

- Over $\mathbb{R}$, the bilinear forms are of the form $\phi(x, y) = axy$ for $a \in \mathbb{R}$.
- The usual dot product over $\mathbb{R}^n$ is a symmetric bilinear form. For $x = (x_1, ..., x_n)$ *and* $y = (y_1, ..., y_n) \in \mathbb{R}^n$, it is given by:

$$\phi(x, y) = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \ldots + x_n y_n.$$

Examples

- Over $\mathbb{R}$, the bilinear forms are of the form $\phi(x, y) = axy$ for $a \in \mathbb{R}$.
- The usual dot product over $\mathbb{R}^n$ is a symmetric bilinear form. For $x = (x_1, ..., x_n)$ *and* $y = (y_1, ..., y_n) \in \mathbb{R}^n$, it is given by:

$$\phi(x, y) = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \ldots + x_n y_n.$$

- Let $a, b, c, d \in \mathbb{R}$ and $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$. The next function is a bilinear form over $\mathbb{R}^2$:

$$\phi(x, y) = ax_1 y_1 + bx_1 y_2 + cx_2 y_1 + dx_2 y_2$$

it is symmetric if and only if $b = c$.

Examples

- Over $\mathbb{R}$, the bilinear forms are of the form $\phi(x,y) = axy$ for $a \in \mathbb{R}$.
- The usual dot product over $\mathbb{R}^n$ is a symmetric bilinear form. For $x = (x_1, ..., x_n)$ *and* $y = (y_1, ..., y_n) \in \mathbb{R}^n$, it is given by:

$$\phi(x,y) = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \ldots + x_n y_n.$$

- Let $a, b, c, d \in \mathbb{R}$ and $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$. The next function is a bilinear form over $\mathbb{R}^2$:

$$\phi(x,y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

  it is symmetric if and only if $b = c$.
- Over the vector space $C^0([0,1], \mathbb{R})$, the following function is a symmetric bilinear form.

$$\phi(f,g) = \int_0^1 f(t)g(t)dt.$$

Remark

Note that

To verify that a map $\phi : E \times E \to \mathbb{R}$ is a symmetric bilinear form, it suffices to see tha ϕ is **linear** in the 1st variable and verifies $\phi(x, y) = \phi(y, x)$ **(these two conditions assure the linearity in the 2nd variable).**

Exercise

Let us consider the vector space $E = \mathbb{R}^2$ and e_1, e_2 its cannonical base. Here are some maps, say which are bilinear and which are symmetric bilinear forms.

$$\phi_1(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1 + x_2y_2,$$
$$\phi_2(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_2 + x_2y_1,$$
$$\phi_3(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1,$$
$$\phi_4(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_2,$$
$$\phi_5(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1^2.$$

Exercise

Let us consider the vector space $E = \mathbb{R}^2$ and e_1, e_2 its cannonical base. Here are some maps, say which are bilinear and which are symmetric bilinear forms.

$$\phi_1 (x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_1 + x_2 y_2,$$
$$\phi_2 (x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_2 + x_2 y_1,$$
$$\phi_3 (x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_1,$$
$$\phi_4 (x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1 y_2,$$
$$\phi_5 (x_1 e_1 + x_2 e_2, y_1 e_1 + y_2 e_2) = x_1^2.$$

All of them are bilinear except the last one.
Only the first three are symmetric bilinear forms.

Matrix representation

We now place ourselves within the framework of the finite dimension. We will associate a matrix with any bilinear form.

Matrix representation

We now place ourselves within the framework of the finite dimension. We will associate a matrix with any bilinear form.

Definition

Let E be a finite dimensional vector space with a basis $(e_1, ..., e_n)$ and ϕ a bilinear form over E. The matrix of ϕ in the base $(e_1, ..., e_n)$ is the follwoing square matrix:

$$B = \{\varphi(e_i,e_j)\}_{\substack{i=1..n \\ j=1..n}} = \begin{pmatrix} \varphi(e_1,e_1) & \dots & \varphi(e_1,e_n) \\ \vdots & & \vdots \\ \varphi(e_n,e_1) & \dots & \varphi(e_n,e_n) \end{pmatrix}.$$

Matrix representation

We now place ourselves within the framework of the finite dimension. We will associate a matrix with any bilinear form.

Definition

Let E be a finite dimensional vector space with a basis $(e_1, ..., e_n)$ and ϕ a bilinear form over E. The matrix of ϕ in the base $(e_1, ..., e_n)$ is the follwoing square matrix:

$$B = \{\varphi(e_i, e_j)\}_{\substack{i=1..n \\ j=1..n}} = \begin{pmatrix} \varphi(e_1, e_1) & \dots & \varphi(e_1, e_n) \\ \vdots & & \vdots \\ \varphi(e_n, e_1) & \dots & \varphi(e_n, e_n) \end{pmatrix}.$$

The form ϕ is symmetrical if and only if its matrix is symmetrical.

Examples

- The matrix associated with the bilinear form on $\mathbb{R}^2$ given by:

$$\phi(x, y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

is the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Examples

- The matrix associated with the bilinear form on $\mathbb{R}^2$ given by:

$$\phi(x, y) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

  is the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

- The matrix associated with the usual scalar product on $\mathbb{R}^n$ is the identity matrix.

Conversely, to a square matrix B of size $n \times n$, we can associate a bilinear form on $\mathbb{R}^n$ by setting:

$$\begin{aligned}
\varphi(x,y) &= {}^t x \, B \, y \\
&= \sum_{i,j} x_i \, b_{i,j} \, y_j \\
&= b_{1,1} \, x_1 y_1 + b_{1,2} \, x_1 y_2 + \ldots + b_{n,n} \, x_n y_n \\
&= \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} b_{1,1} & \ldots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \ldots & b_{n,n} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.
\end{aligned}$$

Conversely, to a square matrix B of size $n \times n$, we can associate a bilinear form on $\mathbb{R}^n$ by setting:

$$\begin{aligned}
\varphi(x,y) &= {}^t x\, B\, y \\
&= \sum_{i,j} x_i\, b_{i,j}\, y_j \\
&= b_{1,1}\, x_1 y_1 + b_{1,2}\, x_1 y_2 + \ldots + b_{n,n}\, x_n y_n \\
&= \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} b_{1,1} & \ldots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \ldots & b_{n,n} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.
\end{aligned}$$

We denote $(x_1, \ldots x_n)$ the coefficients of the vector $x \in \mathbb{R}^n$, $(y_1, .., y_n)$ those of y and $b_{i,j}$ those of the matrix B.
We immediately check that $b_{i,j} = \phi(e_i, e_j)$, where the (e_i) are the vectors of **the canonical basis** of $\mathbb{R}^n$. Matrix B is well the matrix associated with this bilinear form.

Changement of basis

The matrix associated with a symmetric bilinear form depends on the basis chosen to represent it. Let's explain how this matrix is transformed when performing a base change.

Changement of basis

The matrix associated with a symmetric bilinear form depends on the basis chosen to represent it. Let's explain how this matrix is transformed when performing a base change.

Recall that the transition(Passage) matrix is defined between two bases $(e_i)_{i=1..n}$ and $(e'_i)_{i=1..n}$ of a finite dimensional space. This is the matrix whose columns are composed of the coordinates of the vectors of the new base (e'_i) in the old base (e_i). Let us denote this matrix $P = p_{i,j}$.

$$e_j = \sum_{i=1}^{n} P_{i,j} e_i$$

Let $x = (x_1, ..., x_n)$ be the coordinates of a vector in the old basis and $x' = (x'_1, ..., x'_n)$ the coordinates in the new base. We then have:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad x_i = \sum_{j=1}^{n} p_{i,j} x'_j.$$

Let $x = (x_1, ..., x_n)$ be the coordinates of a vector in the old basis and $x' = (x'_1, ..., x'_n)$ the coordinates in the new base. We then have:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad x_i = \sum_{j=1}^{n} p_{i,j} x'_j.$$

The matrices A and A' of a linear map $f : E \to E$ in the bases (e_i) and (e'_i) satisfy the relation:

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

Let $x = (x_1, ..., x_n)$ be the coordinates of a vector in the old basis and $x' = (x'_1, ..., x'_n)$ the coordinates in the new base. We then have:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad x_i = \sum_{j=1}^{n} p_{i,j} x'_j.$$

The matrices A and A' of a linear map $f : E \to E$ in the bases (e_i) and (e'_i) satisfy the relation:

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

We say that the matrices A and A' are **conjugate**. The relationship that binds matrices associated with a bilinear form is different.

Proposition

Let E be a finite dimensional space with two bases $(e_i)_{i=1..n}, (e'_i)_{i=1..n}$ and of a bilinear form $\phi : E \times E \to \mathbb{R}$. Let us denote by P the passage matrix between these two bases and B, B' the matrices of ϕ in each of the bases. Then

$$B' = P^T BP.$$

Proposition

Let E be a finite dimensional space with two bases $(e_i)_{i=1..n}, (e'_i)_{i=1..n}$ and of a bilinear form $\phi : E \times E \to \mathbb{R}$. Let us denote by P the passage matrix between these two bases and B, B' the matrices of ϕ in each of the bases. Then

$$B' = P^T BP.$$

Proof.

$\phi(x, y) = x^T By = (Px')^T B(Py') = x'^T (P^T BP)y'$ $\square$

Proposition

Let E be a finite dimensional space with two bases $(e_i)_{i=1..n}, (e'_i)_{i=1..n}$ and of a bilinear form $\phi : E \times E \to \mathbb{R}$. Let us denote by P the passage matrix between these two bases and B, B' the matrices of ϕ in each of the bases. Then

$$B' = P^T BP.$$

Proof.

$\phi(x, y) = x^T By = (Px')^T B(Py') = x'^T (P^T BP)y'$ □

Two square matrices B and B' linked by a relation of the form $B' = P^T BP$, where P is an invertible matrix, are said to be **congruent.**

Rank and kernel

Definition

The kernel of a bilinear form is defined by

$$ker(\phi) = \{y \in E | \forall x \in E, \phi(x, y) = 0\}.$$

Rank and kernel

Definition

The kernel of a bilinear form is defined by

$$ker(\phi) = \{y \in E | \forall x \in E, \phi(x, y) = 0\}.$$

A bilinear form is said to be non-degenerate if its kernel is restricted to the null vector: $ker(\phi) = \{0\}$.
These are the vectors $y \in E$ for which $\phi(x, y)$ is zero for all $x \in E$.
It is a vector subspace of E.
In a given base, it coincides with the kernel of the matrix associated with the bilinear form.

Examples

1. The bilinear form on $\mathbb{R}^2$ defined by $\phi(x, y) = x_1 y_1$ has a matrix of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Its kernel is made up of vectors whose first coordinate vanishes:

$$ker(\phi) = vect(\begin{pmatrix} 0 \\ 1 \end{pmatrix})$$

Examples

1. The bilinear form on $\mathbb{R}^2$ defined by $\phi(x, y) = x_1 y_1$ has a matrix of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Its kernel is made up of vectors whose first coordinate vanishes:

$$ker(\phi) = vect(\begin{pmatrix} 0 \\ 1 \end{pmatrix})$$

2. An example of a non-degenerate bilinear form is given by the usual scalar product over $\mathbb{R}^n$. In fact, the associated matrix is identity, its kernel is indeed restricted to $\{0\}$.

Examples

1. The bilinear form on $\mathbb{R}^2$ defined by $\phi(x,y) = x_1 y_1$ has a matrix of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Its kernel is made up of vectors whose first coordinate vanishes:
$$ker(\phi) = vect(\begin{pmatrix} 0 \\ 1 \end{pmatrix})$$
2. An example of a non-degenerate bilinear form is given by the usual scalar product over $\mathbb{R}^n$. In fact, the associated matrix is identity, its kernel is indeed restricted to $\{0\}$.
3. The bilinear form defined on $C^0([0,1],R)$ by $\phi(f,g) = \int_0^1 f(t)g(t)dt$ is also not degenerate. In fact, if $\phi(f,g) = 0$ for all g, we can take $g = f$ and get $\int f(t)^2 dt = 0$. We conclude using the following result of integration: A positive continuous function whose integral is zero is identically null.

Definition

The rank of a bilinear form defined on a vector space of finite dimension is equal to the **rank of its matrix**. It does not depend on the base chosen to express this matrix.

Definition

The rank of a bilinear form defined on a vector space of finite dimension is equal to the **rank of its matrix**. It does not depend on the base chosen to express this matrix.

The rank formula for matrices results in

$$dimker(\phi) + row(\phi) = dimE.$$

Let us finish by noting that the set of symmetrical bilinear forms is a vector space. We combine two bilinear forms ϕ_1, ϕ_2 defined on the same space by the formula

$$(\lambda\phi_1 + \phi_2)(x, y) = \lambda\phi_1(x, y) + \phi_2(x, y).$$

Definition of Quadratic forms

We now focus on the study of **symmetrical bilinear forms**. We will associate with such a form a **polynomial of degree two** to several variables that we can then simplify as we did with quadratic polynomes with one variable.

Definition of Quadratic forms

We now focus on the study of **symmetrical bilinear forms**. We will associate with such a form a **polynomial of degree two** to several variables that we can then simplify as we did with quadratic polynomes with one variable.

Definition

A quadratic form Q defined on a real vector space E is a function from E to $\mathbb{R}$ of the form $\phi(x, x)$, where $\phi : E \times E \to \mathbb{R}$ is a symmetrical bilinear form.

$$\forall x \in E, Q(x) = \phi(x, x).$$

The bilinear form ϕ is called the **polar form** of Q. Let E be a space of finite dimension with a basis $(e_i)_{i=1..n}$. Let us give the general expression of a quadratic form in coordinates. Let B be the matrix of the bilinear form associated with Q. Then

$$Q = (\sum_{k=1}^{n} x_k e_k) = x^T B x = \sum_{i,j} b_{i,j} x_i x_j = \sum_{i} b_{i,i} x^2 + 2 \sum_{i,j,i<j} b_{i,j} x_i x_j$$

This expression is a polynomial with several variables, of which all the terms are of total degree two.

Examples

- Consider the quadratic form over $\mathbb{R}^2$ for which $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, then:

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

Examples

- Consider the quadratic form over $\mathbb{R}^2$ for which $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, then:

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

- The quadratic form associated with the usual scalar product of $\mathbb{R}^n$ is:

$$Q(x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k^2 = x_1^2 + \ldots + x_n^2$$

This is the square of the length of the vector $(x_1, ..., x_n)$.

Given a quadratic form Q, there is only one form symmetric bilinear for which we have $Q(x) = \phi(x, x)$ for all x. We obtained through the following formulas.

$$Q(\lambda x) = \lambda^2 Q(x), \quad Q(x+y) = Q(x) + 2\varphi(x,y) + Q(y),$$

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

Given a quadratic form Q, there is only one form symmetric bilinear for which we have $Q(x) = \phi(x, x)$ for all x. We obtained through the following formulas.

$$Q(\lambda x) = \lambda^2 Q(x), \quad Q(x+y) = Q(x) + 2\varphi(x,y) + Q(y),$$

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

The last three formulas are called **polarization identities.**

Proof

$$\begin{aligned} Q(x+y) &= \varphi(x+y, x+y) \\ &= \varphi(x,x)+\varphi(y,x)+\varphi(x,y)+\varphi(y,y) \\ &= Q(x)+2\varphi(x,y)+Q(y). \end{aligned}$$

$$Q(x-y) = Q(x+(-y)) = Q(x) - 2\varphi(x,y) + Q(y).$$

calculation of the matrix associated with a quadratic form

To calculate the matrix associated with a quadratic form,

1. one places on the diagonal the coefficients of the squares x_i^2 appearing in the polynomial of degree two defining Q.

2. For terms which are not on the diagonal, we must divide by two the coefficients of the terms $x_i x_j$ of the polynomial.

Example

Let $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 6x_1x_2 + 10x_1x_3$. The associated matrix is:

$$\begin{pmatrix} 1 & 3 & 5 \\ 3 & 2 & 0 \\ 5 & 0 & 0 \end{pmatrix}$$

We also define the **rank and the kernel** of a quadratic form. They are equal to the rank and the kernel of the symmetrical bilinear form or of the associated matrix.

The cone

Let Q be a quadratic form on a vector space E. **The cone** of the isotropic vectors of Q is defined by

$$C(Q) = \{x \in E | Q(x) = 0\}$$

A vector $x \in$ E satisfying $Q(x) = 0$ is said to be **isotropic**.
It always contains the kernel of the quadratic form: $ker(Q) \subset C(Q)$.

Example

Let $Q(x_1, x_2) = x_1^2 - x_2^2$. The associated matrix is: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
The quadratic form is therefore of rank two and $ker(Q) = \{0\}$.
Let us calculate its cone:
$Q(x_1, x_2) = 0$ implies $(x_1 + x_2)(x_1 - x_2) = x_1^2 - x_2^2 = 0$.
This means $x_1 = x_2$ or $x_1 = -x_2$. The cone is the union of these two vectorial lines directed by the vectors: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

The signature

We will introduce an important invariant of the quadratic forms, connected to **the sign** that its values can take.
A quadratic form is said

- positive if for all $x \in E,\ Q(x) \geq 0$,
- positive definite if for all non-zero $x \in E,\ Q(x) > 0$,
- negative if for all $x \in E, Q(x) \leq 0$,
- negative definite if for all non-zero $x \in E, Q(x) < 0$.
- indefinite if $\exists x \in E, Q(x) > 0$ and $\exists y \in E, Q(y) < 0$.

Examples

1. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1^2 + x_2^2$ is positive definite, because for all $x \in \mathbb{R}^2, x \neq 0$; we have $Q(x) > 0$

Examples

1. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1^2 + x_2^2$ is positive definite, because for all $x \in \mathbb{R}^2, x \neq 0$; we have $Q(x) > 0$
2. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1^2 + x_2^2 + 2x_1x_2$ is positive, because for all $x \in \mathbb{R}^2, x \neq 0$; we have $Q(x) = (x_1 + x_2)^2 \geq 0$

Examples

1. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1^2 + x_2^2$ is positive definite, because for all $x \in \mathbb{R}^2, x \neq 0$; we have $Q(x) > 0$
2. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1^2 + x_2^2 + 2x_1x_2$ is positive, because for all $x \in \mathbb{R}^2, x \neq 0$; we have $Q(x) = (x_1 + x_2)^2 \geq 0$
3. The quadratic form $Q(x) = Q((x_1, x_2)) = x_1x_2$ is indefinite, because for $x \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ we have $Q(x) > 0$ and for $x \in \mathbb{R}_-^* \times \mathbb{R}_-^*$ we have $Q(x) < 0$

We recall that a bilinear form is a non-degenrated form if its kernel is equal to $\{0\}$.

We recall that a bilinear form is a non-degenrated form if its kernel is equal to $\{0\}$.

Note that

A positive definite quadratic form is a non-degenerated form.

In fact, $ker(Q) \subset C(Q)$ and the cone is zero in the case non-degenerated by definition: $C(Q) = x|Q(x) = 0 = \{0\}$

Proposition

A quadratic form $Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i x_j$ *can be written as:*

$$\sum_{j=1}^{n} \lambda_j y_j^2$$

where $\lambda_1, \ldots, \lambda_n$ *are the eigen values of* A *and* y_j *are linear combinations of* $x_1, \ldots, x_n$.

Proof

$Q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_ix_j = X^TAX$, since A is symmetric then A is diagonalisable and there exists a matrix P such that:

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

where D is a diagonal matrix having the eigean values of A on its diagonal and P is an invertible matrix having i its columns the eigen vectors and such that: $P^{-1} = P^T$, hence:

$$Q(x) = X^TAX = X^TPDP^TX$$

we let $Y = P^TX$ then $Y^T = X^TP$ and we obtain $Q(X) = Y^TDY$. Therefore, $q(X) = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{i,j}y_iy_j$. But, since D is a diagonal matrix, then $d_{i,j} = \lambda_j$ if $i = j$ and 0 elsewhere, then:

$$Q(X) = \sum_{j=1}^{n} \lambda_j y_j^2.$$

Proposition

A quadratic form q (or the associated symmetric matrix A) is:

- positive if and only if the eigenvalues of A are positive or zero.
- negative if and only if the eigenvalues of A are negative or zero.
- definite positive if and only if the eigenvalues of A are strictly positive.
- definite negative if and only if the eigenvalues of A are strictly negative
- undefined if and only if the eigenvalues of A are designated of opposite signs.

Proof

- Suppose that all the eigenvalues are positive or nul. We then have $\forall x, Q(x) = \sum_{j=1}^{n} \lambda_j y_j^2 \geq 0$. Which means that the quadratic form is positive. Suppose that the quadratic form is positive that is to say $\forall x, Q(x) \geq 0$. Hence $\sum_{j=1}^{n} \lambda_j y_j^2 \geq 0$. Let us show that $j = 1, \ldots, n, \lambda_j \geq 0$. If there exists k such that $\lambda_k < 0$ then for x such that $P^T X = Y$ where $y_k = 1$ and $y_j = 0, \forall j \neq k$ we have $Q(x) = \lambda k < 0$ which contradicts the hypothesis. So all the eigenvalues are positive or nul.

- Suppose that all the eigenvalues are strictly positive. We then have $\forall x, Q(x) = \sum_{j=1}^{n} \lambda_j y_j^2 \geq 0$. We show that $\forall x \neq 0, Q(x) > 0$. If there is $x \neq 0$ s.t $Q(x) = 0$ then $\sum_{j=1}^{n} \lambda_j y_j^2 = 0$ where $\forall j = 1, \dots, n, y_j = 0$ therefore $Y = 0$. But $Y = P^T X$ then $P^T X = 0$ which implies $X = 0$ because P^T is invertible. So $x = 0$ which contradicts the hypothesis. So $\forall x \neq 0, q(x) > 0$ which means that the quadratic form is positive definite.

- Suppose that all the eigenvalues are strictly positive. We then have $\forall x, Q(x) = \sum_{j=1}^{n} \lambda_j y_j^2 \geq 0$. We show that $\forall x \neq 0, Q(x) > 0$. If there is $x \neq 0$ s.t $Q(x) = 0$ then $\sum_{j=1}^{n} \lambda_j y_j^2 = 0$ where $\forall j = 1, \ldots, n, y_j = 0$ therefore $Y = 0$. But $Y = P^T X$ then $P^T X = 0$ which implies $X = 0$ because P^T is invertible. So $x = 0$ which contradicts the hypothesis. So $\forall x \neq 0, q(x) > 0$ which means that the quadratic form is positive definite.
- Suppose the quadratic form is positively defined, that is, $\forall x \neq 0, Q(x) > 0$. Hence $\sum_{j=1}^{n} \lambda_j y_j^2 > 0$. Let us show that $j = 1, ..., n, \lambda_j > 0$. If there exists k such that $\lambda_k \leq 0$ then for $\forall x$ such that $P^T X = Y$ where $y_k = 1$ and $y_j = 0, \forall j \neq k$ we have $Q(x) = \lambda_k \leq 0$ which contradicts the hypothesis. So all the eigenvalues are strictly positive.

Particular case for matrices (2,2)

We can know the sign of the eigenvalues and therefore the nature of a symmetric matrix of order 2 without calculating the eigenvalues:

Particular case for matrices (2,2)

We can know the sign of the eigenvalues and therefore the nature of a symmetric matrix of order 2 without calculating the eigenvalues:
It follows to calculate its determinant and its trace.
Recall

1. $det(A) = \lambda_1 \times \lambda_2$ (product of the 2 eigenvalues).
2. $Trace(A) = \lambda_1 + \lambda_2$ (sum of the 2 eigenvalues).

- If *det* $A > 0$ then the 2 eigenvalues are of the same sign
  - If *trace* $A > 0$, the 2 eigenvalues are strictly positive and the matrix is positive definite.
  - If *trace* $A < 0$, the 2 eigenvalues are strictly negative and the matrix is negative definite.
- If *det* $A = 0$ then at least one of the 2 eigenvalues is null
  - If *trace* $A > 0$ one has a zero eigenvalue and the other strictly positive from which the matrix is positive.
  - If *trace* $A < 0$ one has a zero eigenvalue and strictly negative, hence the matrix is negative.
  - If *trace* $A =$ 0then 0 is a double eigenvalue hence the matrix is both ositive and negative.
- If *det* $A < 0$ then the 2 eigenvalues are of opposite signs and the matrix id undefined.

Warning: do not use this when $n > 2$.

Examples

- To the quadratic form $Q((x_1, x_2)) = x_1^2 + x_2^2$, we associate the following matrix: $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
  Since A is a diagonal matrix, the 2 eigenvalues of A are the elements on the diagonal. There is a double eigenvalue that is strictly positive, hence A is a definite positive matrix, and same is the quadratic form.

Examples

- To the quadratic form $Q((x_1, x_2)) = x_1^2 + x_2^2$, we associate the following matrix: $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
  Since A is a diagonal matrix, the 2 eigenvalues of A are the elements on the diagonal. There is a double eigenvalue that is strictly positive, hence A is a definite positive matrix, and same is the quadratic form.
- To the quadratic form $Q((x_1, x_2)) = x_1^2 + x_2^2 + 2x_1x_2$, we associate the following matrix: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.
  Since A is a symmetric matrix of dimension 2, it is sufficient to calculate the determinant and the trace of A to determine its nature.
  $det(A) = 0$ this means that one of the eigenvalues is nul.
  $Tr(A) = 2 > 0$, this means that the second eigenvalue is positive.
  Hence A is positive and so is the quadratic form.

- To the quadratic form $Q((x_1, x_2)) = x_1x_2$, we associate the following matrix: $A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$.
  Since A is a symmetric matrix of dimension 2, it is sufficient to calculate the determinant and the trace of A to determine its nature. $det(A) = -\frac{1}{4} < 0$ this means that the eigenvalues are of opposite signs. Hence A is undefined and so is the quadratic form.

- To the quadratic form $Q((x_1, x_2)) = x_1 x_2$, we associate the following matrix: $A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$.
  Since A is a symmetric matrix of dimension 2, it is sufficient to calculate the determinant and the trace of A to determine its nature. $det(A) = -\frac{1}{4} < 0$ this means that the eigenvalues are of opposite signs. Hence A is undefined and so is the quadratic form.
- To the quadratic form $Q((x_1, x_2)) = -2x_1^2 + 4x_1 x_2 - 7x_2^2$, we associate the following matrix: $A = \begin{pmatrix} -2 & 2 \\ 2 & -7 \end{pmatrix}$.
  Since A is a symmetric matrix of dimension 2, it is sufficient to calculate the determinant and the trace of A to determine its nature. $det(A) = 10 > 0$ this means that the eigenvalues are of same signs. $Tr(A) = -9 < 0$ this means that both eigenvalues are negative. Hence A is negative definite and so is the quadratic form.

Let A be a square matrix of dimension n:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix}$$

We note $A_{(p)}$ the matrix obtained by taking the first p rows and p columns from A, where $p = 1, 2, \dots, n$.

$$A_{(1)} = (a_{11})$$
$$A_{(2)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
.

.

$$A_{(n)} = A$$

Proposition

A quadratic form or its associated matrix A is:

1. ***definite positive*** *iff* $\forall p = 1, \ldots, n, det(A_{(p)}) > 0$
2. ***definite negative*** *iff* $\forall p = 1, \ldots, n, (-1)^p det(A_{(p)}) > 0$
3. ***positive*** *then* $\forall p = 1, \ldots, n, det(A_{(p)}) \geq 0$
4. ***negative*** *then* $\forall p = 1, \ldots, n, (-1)^p det(A_{(p)}) \geq 0$

Example1

Let $Q((x_1, x_2, x_3)) = 3x_1^2 + 8x_2^2 + 3x_3^2 - 2\sqrt{2}x_1x_2 + 4x_2x_3$, the associated matrix is:

$$A = \begin{pmatrix} 3 & -\sqrt{2} & 0 \\ -\sqrt{2} & 8 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$

1. First method: Calculate the eigenvalues of this symmetric matrix of dimension 3 in order to determine their signs.
2. Second method: Study the sign of $det(A_{(p)}), \forall p = 1, \ldots, 3$.

First method

The eigenvalues of A are the root of $P_A(\lambda)$.

$$\mathcal{P}_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -\sqrt{2} & 0 \\ -\sqrt{2} & 8-\lambda & 2 \\ 0 & 2 & 3-\lambda \end{vmatrix}$$

$$= (3-\lambda)[(8-\lambda)(3-\lambda)-4] + \sqrt{2}[-\sqrt{2}(3-\lambda) - 2 \times 0]$$

$$= (3-\lambda)(\lambda^2 - 11\lambda + 20) - 2(3-\lambda)$$

$$= (3-\lambda)(\lambda^2 - 11\lambda + 18)$$

There is a trivial root 3, and the roots of the polynomial $\lambda^2 - 11\lambda + 18$ are 2 and 9. Hence, the three roots are $\lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 9$. They are all strictly positive values, therefore A is a positive definite matrix and so is Q.

Second method

We study $det(A_{(p)}), \forall p = 1, \ldots, 3$.

$$A_{(1)} = (3).\ det(A_{(1)}) = 3 > 0$$

$$A_{(2)} = \begin{pmatrix} 3 & -\sqrt{2} \\ -\sqrt{2} & 8 \end{pmatrix},\ det(A_{(2)}) = 22 > 0$$

$$A_{(3)} = A,\ det(A_{(3)}) = det(A) = 54 > 0$$

We hence have, $\forall p = 1, \ldots, 3, det(A_{(p)}) > 0$.
Therefore A is a positive definite matrix and so is Q.

Example2

Let $Q((x_1, x_2, x_3)) = -6x_1^2 - 16x_2^2 - 6x_3^2 + 4sqrt2x_1x_2 - 8x_2x_3$, the associated matrix is:

$$A = \begin{pmatrix} -6 & 2\sqrt{2} & 0 \\ 2\sqrt{2} & -16 & -4 \\ 0 & -4 & -6 \end{pmatrix}$$

1. First method: Calculate the eigenvalues of this symmetric matrix of dimension 3 in order to determine their signs.
2. Second method: Study the sign of $det(A_{(p)}), \forall p = 1, \ldots, 3$.

First method

The eigenvalues of A are the root of $P_A(\lambda)$.

$$\mathcal{P}_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -6-\lambda & 2\sqrt{2} & 0 \\ 2\sqrt{2} & -16-\lambda & -4 \\ 0 & -4 & -6-\lambda \end{vmatrix}$$

$$= (-6-\lambda)[(-16-\lambda)(-6-\lambda)-16] - 2\sqrt{2}[2\sqrt{2}(-6-\lambda)-(-4\times 0)]$$

$$= (-6-\lambda)[(-16-\lambda)(-6-\lambda)-16] - 8(-6-\lambda)$$

$$= (-6-\lambda)(\lambda^2+22\lambda+80) - 8(-6-\lambda)$$

$$= (-6-\lambda)(\lambda^2+22\lambda+80-8)$$

$$= (-6-\lambda)(\lambda^2+22\lambda+72)$$

The three roots are $\lambda_1 = -4, \lambda_2 = -6, \lambda_3 = -18$. They are all strictly negative values, therefore A is a negative definite matrix and so is Q.

Second method

We study $det(A_{(p)}), \forall p = 1, \ldots, 3.$

$$A_{(1)} = (-6).\ det(A_{(1)}) = -6 < 0$$

$$A_{(2)} = \begin{pmatrix} -6 & 2\sqrt{2} \\ 2\sqrt{2} & -16 \end{pmatrix}, det(A_{(2)}) = 88 > 0$$

$$A_{(3)} = A,\ det(A_{(3)}) = det(A) = -432 < 0$$

We hence have, $det(A_{(1)}) < 0, det(A_{(2)}) > 0, det(A_{(3)}) < 0.$
We hence have, $\forall p = 1, \ldots, 3, (-1)^p det(A_{(p)}) > 0.$
Therefore A is a negtive definite matrix and so is Q.

Inner product

Definition

An inner product on a real vector space V is a bilinear form which is both positive definite and symmetric.

Angles and length

- Suppose that $\phi(x, y)$ is an inner product on a real vector space V .

Angles and length

- Suppose that $\phi(x, y)$ is an inner product on a real vector space V .
- Then one may define the length of a vector $v \in V$ by setting

$$||v|| = \sqrt{\phi(v, v)}$$

and the angle θ between two vectors $v, w \in V$ by setting

$$cos(\theta = \frac{\phi(v, w)}{||v||.||w||}$$

Angles and length

- Suppose that $\phi(x, y)$ is an inner product on a real vector space V .
- Then one may define the length of a vector $v \in V$ by setting

$$||v|| = \sqrt{\phi(v, v)}$$

  and the angle θ between two vectors $v, w \in V$ by setting

$$cos(\theta = \frac{\phi(v, w)}{||v||.||w||}$$

- These formulas are known to hold for the inner product on $\mathbb{R}^n$.

Angles and length

- Suppose that $\phi(x, y)$ is an inner product on a real vector space V .
- Then one may define the length of a vector $v \in V$ by setting

$$||v|| = \sqrt{\phi(v, v)}$$

and the angle θ between two vectors $v, w \in V$ by setting

$$cos(\theta = \frac{\phi(v, w)}{||v||.||w||}$$

- These formulas are known to hold for the inner product on $\mathbb{R}^n$.

Theorem (Cauchy-Shwartz inequality)

When V is a real vector space with an inner product, one has

$$|\phi(v, w)| \leq ||v||.||w|| \ \forall \ v, w \in V.$$

Orthogonal vectors

Definition (Orthogonal and orthonormal)

Suppose $\phi(x, y)$ is a symmetric bilinear form on a real vector space V . Two vectors u, v are called **orthogonal**, if $\phi(u, v) = 0$.
A basis $v_1, v_2, ..., v_n$ of V is called **orthogonal**, if $\phi(v_i, v_j) = 0$ whenever $i \neq j$ and it is called **orthonormal**, if it is orthogonal with $\phi(v_i, v_i) = 1$ for all i.

Orthogonal vectors

Definition (Orthogonal and orthonormal)

Suppose $\phi(x, y)$ is a symmetric bilinear form on a real vector space V .
Two vectors u, v are called **orthogonal**, if $\phi(u, v) = 0$.
A basis $v_1, v_2, ..., v_n$ of V is called **orthogonal**, if $\phi(v_i, v_j) = 0$ whenever $i \neq j$ and it is called **orthonormal**, if it is orthogonal with $\phi(v_i, v_i) = 1$ for all i.

Theorem (Linear combinaision)

Let $v_1, v_2, ..., v_n$ be an orthogonal basis of an inner product space V. Then every vector $v \in V$ can be expressed as a linear combination

$$v = \sum_{i=1}^{n} c_i v_i, \text{ where } c_i = \frac{\phi(v, v_i)}{\phi(v_i, v_i)}, \text{ for all } i.$$

If the basis is actually orthonormal, then $c_i = \phi(v, v_i)$ for all i.

Gram-Schmidt procedure

- Suppose that $v_1, v_2, ..., v_n$ is a basis of an inner product space V . Then we can find an orthogonal basis $w_1, w_2, ..., w_n$ as follows.

Gram-Schmidt procedure

- Suppose that $v_1, v_2, ..., v_n$ is a basis of an inner product space V . Then we can find an orthogonal basis $w_1, w_2, ..., w_n$ as follows.
- Define the first vector by $w_1 = v_1$ and the second vector by

$$w_2 = v_2 - \frac{\phi(v_2, w_1)}{\phi(w_1, w_1)} w_1.$$

Then w_1, w_2 are orthogonal and have the same span as v_1, v_2.

Gram-Schmidt procedure

- Suppose that $v_1, v_2, ..., v_n$ is a basis of an inner product space V . Then we can find an orthogonal basis $w_1, w_2, ..., w_n$ as follows.
- Define the first vector by $w_1 = v_1$ and the second vector by

$$w_2 = v_2 - \frac{\phi(v_2, w_1)}{\phi(w_1, w_1)} w_1.$$

  Then w_1, w_2 are orthogonal and have the same span as v_1, v_2.
- Proceeding by induction, suppose $w_1, w_2, ..., w_k$ are orthogonal and have the same span as $v_1, v_2, ..., v_k$. Once we then define

$$w_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{\phi(v_{k+1}, w_i)}{\phi(w_i, w_i)} w_i$$

  we end up with vectors $w_1, w_2, ..., w_{k+1}$ which are orthogonal and have the same span as the original vectors $v_1, v_2, ..., v_{k+1}$.

Gram-Schmidt procedure

- Suppose that $v_1, v_2, ..., v_n$ is a basis of an inner product space V . Then we can find an orthogonal basis $w_1, w_2, ..., w_n$ as follows.
- Define the first vector by $w_1 = v_1$ and the second vector by

$$w_2 = v_2 - \frac{\phi(v_2, w_1)}{\phi(w_1, w_1)} w_1.$$

  Then w_1, w_2 are orthogonal and have the same span as v_1, v_2.
- Proceeding by induction, suppose $w_1, w_2, ..., w_k$ are orthogonal and have the same span as $v_1, v_2, ..., v_k$. Once we then define

$$w_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{\phi(v_{k+1}, w_i)}{\phi(w_i, w_i)} w_i$$

  we end up with vectors $w_1, w_2, ..., w_{k+1}$ which are orthogonal and have the same span as the original vectors $v_1, v_2, ..., v_{k+1}$.
- Using the formula from the last step repeatedly, one may thus obtain an orthogonal basis $w_1, w_2, ..., w_n$ for the vector space V .

Gram-Schmidt procedure: Example

We find an orthogonal basis of $\mathbb{R}^3$, starting with the basis

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Gram-Schmidt procedure: Example

We find an orthogonal basis of $\mathbb{R}^3$, starting with the basis

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

We define the first vector by $w_1 = v_1$ and the second vector by

$$w_2 = v_2 - \frac{\phi(v_2, w_1)}{\phi(w_1, w_1)} w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Gram-Schmidt procedure: Example

We find an orthogonal basis of $\mathbb{R}^3$, starting with the basis

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

We define the first vector by $w_1 = v_1$ and the second vector by

$$w_2 = v_2 - \frac{\phi(v_2, w_1)}{\phi(w_1, w_1)} w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Then w_1, w_2 are orthogonal and we may define the third vector by:

$$w_3 = v_3 - \frac{\phi(v_3, w_1)}{\phi(w_1, w_1)} w_1 - \frac{\phi(v_3, w_2)}{\phi(w_2, w_2)} w_2$$

$$= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \frac{4}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Example

Question 1: Apply Gram-Schmidt orthogonalization to the following sequence of vectors in $\mathbb{R}^3$: $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

Solution

Solution Apply the process on page 365, with $x_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Step 1 produces an orthogonal basis:

$v_1 = x_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$.

$$v_2 = x_2 - \frac{(x_2, v_1)}{(v_1, v_1)} v_1 = \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix} - \frac{\left(\begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right)}{\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right)} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix} - \frac{10}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix}.$$

$$v_3 = x_3 - \frac{(x_3, v_1)}{(v_1, v_1)} v_1 - \frac{(x_3, v_2)}{(v_2, v_2)} v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right)}{\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right)} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \frac{\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix}\right)}{\left(\begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix}\right)} \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix} =$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{0}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \frac{-6}{81} \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}.$$

Step 2 produces an orthonormal basis by replacing each vector with a vector of norm 1:

Replace v_1 with $\frac{v_1}{|v_1|} = \dfrac{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}{\left\| \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\|} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$.

Replace v_2 with $\frac{v_2}{|v_2|} = \dfrac{\begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix}}{\left\| \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix} \right\|} = \frac{1}{9} \begin{bmatrix} 6 \\ -3 \\ -6 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$.

Replace v_3 with $\frac{v_3}{|v_3|} = \dfrac{\frac{1}{9}\begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}}{\left\| \frac{1}{9}\begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix} \right\|} = \frac{1}{3\sqrt{5}} \begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}$.

So the final solution is $v_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, $v_2 = \frac{1}{3} \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$, $v_3 = \frac{1}{3\sqrt{5}} \begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}$.

Restrictions

We can always restrict a quadratic form to a subspace vector of E. We will denote by $Q|F$ the restriction of Q to a subspace $F \subset E$.

Restrictions

We can always restrict a quadratic form to a subspace vector of E. We will denote by $Q|F$ the restriction of Q to a subspace $F \subset E$.
Notice that a quadratic form which is positive definite over a space vector E is automatically defined positive in restriction at all vector subspaces of E.

Restrictions

We can always restrict a quadratic form to a subspace vector of E. We will denote by $Q|F$ the restriction of Q to a subspace $F \subset E$.
Notice that a quadratic form which is positive definite over a space vector E is automatically defined positive in restriction at all vector subspaces of E.
We will be interested in **the maximum dimension** of the subspaces in restriction of which Q is **positive definite or negative definite**.

Restrictions

We can always restrict a quadratic form to a subspace vector of E. We will denote by $Q|F$ the restriction of Q to a subspace $F \subset E$.
Notice that a quadratic form which is positive definite over a space vector E is automatically defined positive in restriction at all vector subspaces of E.
We will be interested in **the maximum dimension** of the subspaces in restriction of which Q is **positive definite or negative definite**.

Definition

Let E be a finite dimensional vector space and Q a form quadratic on E. The **signature** of Q is the pair of integers $(p, q) \in \mathbb{N}$ given through:

$$p = \max\{dimF | F \ subspace \ of \ E \ such \ that \ Q|F \ is \ definite \ positive\}$$

$$q = \max\{dimF | F \ subspace \ of \ E \ such \ that \ Q|F \ is \ definite \ negative\}$$

Example

A quadratic form Q on a space E of dimension n which is defined positive is of signature **(n, 0)**. This is the case, for example, for the product usual scalar on $\mathbb{R}^n$.

Example

A quadratic form Q on a space E of dimension n which is defined positive is of signature **(n, 0)**. This is the case, for example, for the product usual scalar on $\mathbb{R}^n$.

- In fact, for a positive definite quadratic form, the largest space on which Q is positive definite is E itself.

Example

A quadratic form Q on a space E of dimension n which is defined positive is of signature **(n, 0)**. This is the case, for example, for the product usual scalar on $\mathbb{R}^n$.

- In fact, for a positive definite quadratic form, the largest space on which Q is positive definite is E itself.
- If F is a space on which Q is negative definite, it cannot contain a non-zero vector x, because we would have for this vector $Q(x) > 0$ and $Q(x) < 0$. The subspace $F = \{0\}$ is the only on which Q is defined negative.

Reduction of quadratic forms

The study of quadratic forms is based on an **algorithm** due to F. Gauss, which makes it possible to simplify **the expression of a quadratic form** by writing it as a sum of squares of linear shapes.

Reduction of quadratic forms

The study of quadratic forms is based on an **algorithm** due to F. Gauss, which makes it possible to simplify **the expression of a quadratic form** by writing it as a sum of squares of linear shapes.

Theorem

Let E be a finite dimensional space, Q a quadratic form on E of signature (p, q). Then there exist $p+q$ independent linear forms $l_1, ..., l_{p+q} \in E^*$ such that:

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

Reduction of quadratic forms

The study of quadratic forms is based on an **algorithm** due to F. Gauss, which makes it possible to simplify **the expression of a quadratic form** by writing it as a sum of squares of linear shapes.

Theorem

Let E be a finite dimensional space, Q a quadratic form on E of signature (p, q). Then there exist $p+q$ independent linear forms $l_1, ..., l_{p+q} \in E^*$ such that:

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

This breakdown into squares of linear shapes is not unique. However, we can always calculate **the signature** of the quadratic form from such a decomposition, if the linear forms are independent between them.

Proposition

Consider a quadratic form given by the expression

$$Q(x) = \pm l_1(x)^2 \pm l_2(x)^2 \pm \ldots \pm l_k(x)^2.$$

where the l_i are ***independent linear forms****. So the signa- ture is given by the number of signs $+$ and the number of signs $-$ appearring in this decomposition.*

Proposition

Consider a quadratic form given by the expression

$$Q(x) = \pm l_1(x)^2 \pm l_2(x)^2 \pm \dots \pm l_k(x)^2.$$

where the l_i are ***independent linear forms****. So the signa- ture is given by the number of signs $+$ and the number of signs $-$ appearring in this decomposition.*

The proof will be given later in this Chapter.

Reduction of quadratic forms method

We place ourselves in an arbitrary base, Q takes the form

$$Q(x_1, ..., x_n) = \sum_i b_{i,i} x_i^2 + 2 \sum_{i<j} b_{i,j} x_i x_j.$$

Reduction of quadratic forms method

We place ourselves in an arbitrary base, Q takes the form

$$Q(x_1, ..., x_n) = \sum_i b_{i,i} x_i^2 + 2 \sum_{i<j} b_{i,j} x_i x_j.$$

- If one of the diagonal coefficients $b_{i,i}$ is non-zero, say $b_{1,1} \neq 0$, we consider the polynomial as a function of the variable x_1 and put it under canonical form, as follows.
  - We place the term x_1^2 at the head of the polynomial and we factorize by x_1 into the remaining terms, if possible. The expression of Q then takes the form

    $$Q(x) = b_{1,1}[x_1^2 + x_1(...) + (...)].$$

    Expressions in parentheses must no longer contain the variable x_1. Let us denote them a_1 and a_2:

    $$Q(x) = b_{1,1}[x_1^2 + x_1 a_1 + a_2].$$

- We factor the previous expression using the **remarkable identity**

$$x_1^2 + x_1 a_1 = (x_1 + \frac{a_1}{2})^2 - \frac{a_1^2}{4}$$

- We set $l_1(x_1, \ldots, x_n) = \sqrt{|b_{1,1}|}(x_1 + \frac{a_1}{2})$, the expression of Q becomes:

$$Q(x_1, \ldots, x_n) = \pm l_1(x_1, \ldots, x_n)^2 + b_{1,1}(a_2 - \frac{a_1^2}{4})$$

- The expression $(a_2 - \frac{a_1^2}{4})$ no longer involves the variable x_1. It's a quadratic form to which the algorithm can be re-applied. The sign before the linear form l_1 is that of $b_{1,1}$.

- If all the diagonal coefficients are zero, we consider a coefficient $b_{i,j} \neq 0$, say $b_{1,2}$ to simplify the notations.
  - We place the term x_1x_2 at the head of the polynomial and we factorize by x_1 into the remaining terms, if possible. We do the same with x_2. We obtain

$$Q(x) = b_{1,2}[x_1x_2 + x_1(...) + x_2(...) + (...)].$$

    Expressions in parentheses must **no longer contain the variable** x_1. Let us denote them a_1, a_2 and a_3:

$$Q(x) = b_{1,2}[x_1x_2 + x_1a_1 + x_2a_2 + a_3].$$

  - We factorize the previous expression using the remarkable identity

$$x_1x_2 + x_1a_1 + x_2a_2 = (x_1 + a_2)(x_2 + a_1) - a_1a_2.$$

  - We convert the product $(x_1 + a_2)(x_2 + a_1)$ to the difference of two squares by using remarkable identity

$$xy = \frac{1}{4}(x+y)^2 - \frac{1}{4}(x-y)^2$$

The expression of Q becomes:

$$Q(x) = b_{1,2}\left[\left(\frac{x_1 + x_2 + a_1 + a_2}{2}\right)^2 - \left(\frac{x_1 - x_2 - a_1 + a_2}{2}\right)^2 - a_1a_2 + a_3\right].$$

We let:

$$l_1(x_1,\ldots,x_n) = \sqrt{|b_{1,2}|}(x_1 + x_2 + a_1 + a_2)$$
$$l_2(x_1,\ldots,x_n) = \sqrt{|b_{1,2}|}(x_1 - x_2 - a_1 + a_2)$$

which gives:

$$Q(x_1,\ldots,x_n) = \pm l_1(x_1,\ldots,x_n)^2 \mp l_2(x_1,\ldots,x_n)^2 + b_{1,2}(a_3 - a_1a_2).$$

The last term in the previous expression no longer involves the variables x_1 and x_2. It is a quadratic form to which we can apply the algorithm again.

Ultimately, the algorithm is based on the following three rules.

$$X^2 + Xa = \left(X + \frac{a}{2}\right)^2 - \frac{a^2}{4}$$

$$XY + Xa_1 + Ya_2 = (X + a_2)(Y + a_1) - a_1a_2$$

$$XY = \frac{1}{4}(X + Y)^2 - \frac{1}{4}(X - Y)^2.$$

Note that this algorithm is a generalization of the method of solving polynomials of degree two with one variable.

Let us explain why the family of linear forms given by the algorithm is **linearly independent**. Consider the matrix whose rows are given by the coordinates of these linear forms.

1-If in the algorithm, we have at each step a non-zero diagonal, the matrix obtained is in echelon form, without zero line. Its rank is equal to its number of lines, which shows that these lines are indeed linearly independent of each other.

2-If at some point there are no non-zero diagonal terms, the algorithm produces two linear forms l_1 and l_2 whose expressions are given above. It suffices to subtract the first from the second to put the matrix in echelon form.

Example 1

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

$$\begin{aligned} x_1^2 + x_1x_2 + x_2^2 &= \left(x_1 + \frac{x_2}{2}\right)^2 - \frac{x_2^2}{4} + x_2^2 \\ &= \left(x_1 + \frac{x_2}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}x_2\right)^2. \end{aligned}$$

we let

$$l_1(x_1,x_2) = x_1 + \frac{x_2}{2}, \quad l_2(x_1,x_2) = \frac{\sqrt{3}}{2}x_2,$$

which gives $Q = l_1^2 + l_2^2$, the signature is equale to $(2,0)$

Example 2

– $Q(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_3x_1$.

$$\begin{aligned} x_1x_2 + x_2x_3 + x_3x_1 &= (x_1+x_3)(x_2+x_3) - x_3^2 \\ &= \left(\frac{x_1+x_2+2x_3}{2}\right)^2 - \left(\frac{x_1-x_2}{2}\right)^2 - x_3^2. \end{aligned}$$

we let

$$l_1(x_1,x_2,x_3) = \frac{x_1+x_2+2x_3}{2}, \quad l_2(x_1,x_2,x_3) = \frac{x_1-x_2}{2}, \quad l_3(x_1,x_2,x_3) = x_3.$$

which gives $Q = l_1^2 - l_2^2 - l_3^2$, the signature is equale to $(1,2)$

Orthogonal basis

From the previous theorem, we can construct a basis of E in which Q is given by a particularly simple expression.

Theorem

Let E be a finite dimensional space, Q a quadratic form on signature E (p, q). Then there exists a basis $(e_i)_{i=1..n}$ of E such that

$$Q(x_1 e_1 + ... + x_n e_n) = x_1^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_{p+q}^2.$$

Notice that the matrix of Q in the base (e_i) is diagonal.
This shows that this basis satisfies $\phi(e_i, e_j) = 0$ for all $i \neq j$.
We denoted ϕ the bilinear form associated with Q. This brings us to the following definition:

Definition

Let E be a finite dimensional vector space and ϕ a symmetric bilinear form defined on E. We say that a basis $(e_1, ...e_n)$ of E is orthogonal for ϕ or for the quadratic form Q associated with ϕ if

$$\phi(e_i, e_j) = 0 \textit{ for all } i, j \textit{ such that } i \neq j.$$

this implies

Proposition

Any symmetrical bilinear form defined over a vector space of finite dimension admits an orthogonal basis.

Construction of an orthogonal basis

We consider a quadratic form defined over $\mathbb{R}^n$.

- We apply the Gaussian reduction algorithm to obtain linear forms $l_1, ..., l_{p+q}$ and we consider the matrix P whose rows are given by the coordinates of these linear forms.

- If $p+q<n$, we must complete the family (l_i) so as to obtain a basis of R_n^*. In practice, this involves adding rows to the matrix P in order to make an invertible square matrix.
We can take these lines from the shape (0...1...0) by judiciously positioning the 1. Place them in the columns corresponding to the columns without pivots of the Echelon form of P always works.

- The searched vectors have for coordinates, the columns of the inverse of the matrix P.

Verification

Let's check that this procedure gives an orthogonal basis.
Through construction, the family (e_k) obtained satisfies $l_i(e_k) = 0$ for $i \neq k$. The completed base (l_i) is dual at the base (e_k). The polar form of $Q = l_i^2 - l_j^2$ worth

$$\varphi(x,y) = \sum_{i=1}^{p} l_i(x)l_i(y) - \sum_{j=p+1}^{p+q} l_j(x)l_j(y)$$

because this form is symmetric bilinear and satisfies $Q(x) = \phi(x,x)$. We then obtains $\phi(e_k, e_l) = \sum l_i(e_k)l_i(e_l) - \sum l_j(e_k)l_j(e_l) = 0$ *if* $k \neq l$.

Example 1

$$-Q(x_1,x_2) = x_1^2 + x_1x_2 + x_2^2 = l_1(x_1,x_2,x_3)^2 - l_2(x_1,x_2,x_3)^2$$

with

$$l_1(x_1,x_2) = x_1 + \frac{x_2}{2}, \quad l_2(x_1,x_2) = \frac{\sqrt{3}}{2}x_2.$$

The forms l_1 and l_2 have coordinates $(1\ \frac{1}{2})$ and $(0\ \frac{\sqrt{3}}{2})$.
We form the matrices P and P^{-1} whose lines are given by these line vectors.

$$P = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

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

The columns of the matrix P^{-1} form an orthogonal basis for Q.

Example 2

$$-Q(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_3x_1$$
$$= l_1(x_1,x_2,x_3)^2 - l_2(x_1,x_2,x_3)^2 - l_3(x_1,x_2,x_3)^2$$

with

$$l_1(x_1,x_2,x_3) = \frac{x_1 + x_2 + 2x_3}{2}, \quad l_2(x_1,x_2,x_3) = \frac{x_1 - x_2}{2}, \quad l_3(x_1,x_2,x_3) = x_3.$$

The forms l_1, l_2 and l_3 have coordinates $\frac{1}{2}(1\ 1\ 2)$, $\frac{1}{2}(1\ \ -1\ 0)$ *and* $(0\ 0\ 1)$. We form the matrices P and P^{-1} whose lines are given by these line vectors.

$$P = \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad P^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The columns of the matrix P^{-1} form an orthogonal basis for Q.

Example 3

$$
\begin{aligned}
-Q(x_1,x_2,x_3) &= 4x_1x_3 + 4x_2x_3 \\
&= 4(x_1+x_2)\,x_3 \\
&= l_1(x_1,x_2,x_3)^2 - l_2(x_1,x_2,x_3)^2
\end{aligned}
$$

with

$$
l_1(x_1,x_2,x_3) = x_1 + x_2 + x_3, \quad l_2(x_1,x_2,x_3) = x_1 + x_2 - x_3.
$$

The forms l_1, l_2 have coordinates $(1\ 1\ 1), (1\ 1\ -1)$.
It is necessary to add a line to it in order to make it square invertible. For this, we place a 1 in the second column on the third row. All that's left is to invert the matrix to get the orthogonal basis for Q.

$$
P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -2 \\ 0 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}.
$$

The columns of the matrix P^{-1} form an orthogonal basis for Q.

The diagonal matrix

Theorem

Let S be a symmetric matrix. So there exists a diagonal matrix D whose all diagonal terms belong to $\{-1, 0, 1\}$ and an invertible matrix P such that:

$$S = P^T DP$$

Let us demonstrate this result. Consider the quadratic form on $mathbbR^n$ whose matrix is given by S. The previous theorem gives a basis (e_i) in which the quadratic form has a diagonal D matrix, with coefficients diagonals belonging to $\{-1, 0, 1\}$.
Let us denote by P the passage matrix of the canonical base (e_i), it is the matrix whose rows are given by the linear forms l_i. From the base change formula for the quadratic forms, we have $S = P^T DP$.

Examples

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

$$\frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = {}^t\begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} = {}^t\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Study of the signature

Lemma

Let E be a finite dimensional vector space, Q a quadratic form of signature (p, q). Suppose there are vector subspaces E_1, E_2 of E such that:

- $E_1 + E_2 = E$,
- *Q is positive definite in restriction to E_1,*
- *Q is negative in restriction to E_2.*

Then $E_1 \cap E_2 = \{0\}$ and $p = dim(E_1)$.

Proof

The quadratic form Q is both positive and negative definite on $E_1 \cap E_2$. We therefore have $Q(x) > 0$ and $Q(x) \leq 0$ for any non-zero vector x in this intersection, which shows that it contains only the zero vector. The subspaces E_1 and E_2 are in direct sum, we deduce

$$dim(E_1) = dim(E) - dim(E_2).$$

Let F be a subspace of E such that Q is positive definite in restriction to F. Again, the quadratic form Q is positive and negative definite on the intersection $F \cap E_2$, which is therefore restricted to $\{0\}$. This allows us to conclude:

$$dim(F) \leq dim(E) - dim(E_2) = dim(E_1).$$

Proof of the Theorem

We prove from this lemma that a quadratic form of the form

$$Q(x_1e_1 + ... + x_ne_n) = x_1^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_{p+q}^2$$

is necessarily of signature (p, q).
For this, we notice that Q is positive definite on the vector space E_1 generated by $(e_1, ..., e_p)$ and negative on the space E_2 generated by $(e_{p+1}, ..., e_n)$. By the lemma, we see that p corresponds well to the first coefficient of the signature. We apply the same reasoning on the quadratic form $-Q$ which has for signature (q, p) in order to conclude.
Finally, note that the matrix of Q is diagonal. The first **p** diagonal coefficients are equal to **one**, the following **q** are equal to -1 and the others are zero. Such a matrix has a kernel of dimension $n - p - q$.

We can deduce the following corollary.

Corollary

Let E be a vector space of finite dimension n, Q a quadratic form of signature (p, q). Then

$$p + q = rank(Q), p + q + dimker(Q) = n.$$

Method: calculation of the signature of a quadratic form:
Just reduce the quadratic form and count the number of "+" signs and signs "-" in front of the squares of the linear shapes appearing in the former reduced expression. It is crucial that these linear shapes are independent between them.
It is sometimes possible to determine the signature of a quadra- tick without using the Gaussian reduction algorithm, using the determining. This method is explained in the supplements to this chapter.

Conical

Definition

A conic C is a subset of the plane composed of points whose coordinates $(x, y) \in \mathbb{R}^2$ in a certain basis satisfy an equation of the form:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

with a, b, c, d, e and f real numbers such that a, b and c are not all the three null.

We will show that a judicious change of variables makes it possible to simplify the conic equation.

Proposition

Consider a conic C. Then there is a change of affine variables of the form:

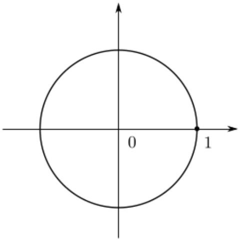
$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + b_2 \end{cases}$$

which transforms the expression of the conic into one of the following shapes:

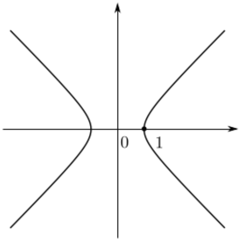
$$\begin{array}{cc} X^2 + Y^2 = 1 & \textit{the conic is an ellipse,} \\ X^2 - Y^2 = 1 & \textit{the conic is a hyperbola,} \\ X^2 - Y = 0 & \textit{the conic is a parabola,} \end{array}$$

or one of the forms:

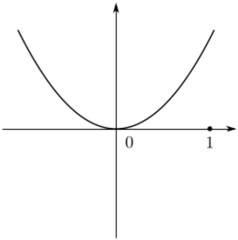
$$X^2 = 1, X^2 - Y^2 = 0, X^2 = 0, X^2 + Y^2 = 0, X^2 + Y^2 = -1, X^2 = -1,$$



$X^2 + Y^2 = 1$



$X^2 - Y^2 = 1$



$X^2 - Y = 0$

We can determine which conic it is by looking at **the signature** of the quadratic form $Q(x,y) = ax^2 + bxy + cy^2$ or the associated **discriminant:** $\Delta = ac - \frac{b^2}{4}$.

signature (2,0)$\longrightarrow$	ellipse or point or empty ($\Delta > 0$)
signature (1,1)$\longrightarrow$	hyperbola or two secant lines ($\Delta < 0$)
signature (1,0)$\longrightarrow$	parabola or two lines or straight or empty ($\Delta = 0$)

To distinguish, for example, between an ellipse, a point and the empty set, it suffices to show at least two points belonging to the ellipse.

The discriminant of a quadratic form is equal to the determinant of the matrix associated with the quadratic form. Be careful not to confuse him with the discriminant of a polynomial. The link with the signature will be explained later.

Method

Reduction of the equation of a conic

- We perform the change of variable which reduces the shape quadratic $Q(x, y) = ax^2 + bxy + cy^2$.
- Even if it means inverting the two variables, the conic equation takes one of the following three forms:

$$\begin{array}{rc} x'2 + y'2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (2.0) \textit{ or } (0.2) \\ x'2 - y'2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (1.1) \\ x'2 + \alpha x' + \beta y' + \gamma = 0 & \text{signature } (1.0) \textit{ or } (0.1) \end{array}$$

In the first two cases, one makes disappear the terms of degree 1 in putting the polynomials $x'^2 + \alpha x' and \pm y'^2 + \beta y'$ in canonical form. Same for x' in the last case. There remains a term or constant in which case we take $Y = y'$, or of degree one in y' and we assign it to Y.

Examples

Determine the nature of the conic

$$\mathcal{C} = \{(x,y) \in \mathbf{R}^2 \mid x^2 + 5y^2 + 2xy + 2x + 10y + 4 = 0\}.$$

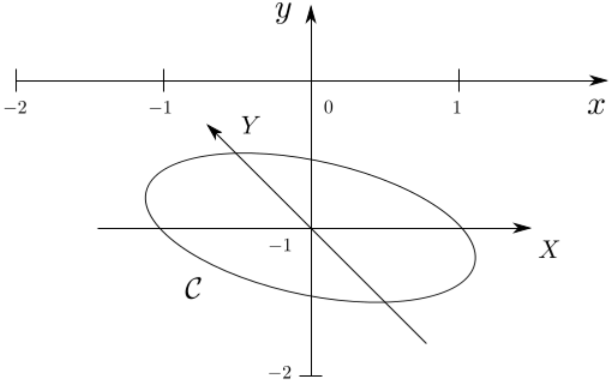
By reduction

$$\begin{aligned} & x^2 + 5y^2 + 2xy + 2x + 10y + 4 \\ & \quad = (x+y)^2 + (2y)^2 + 2x + 10y + 4 \\ & \quad = x'^2 + y'^2 + 2x' + 4y' + 4 \\ & \quad = x'^2 + 2x' + y'^2 + 4y' + 4 \\ & \quad = (x'+1)^2 + (y'+2)^2 - 1 \ = \ X^2 + Y^2 - 1 \end{aligned}$$

we let

$$\begin{cases} x' & = & x + y \\ y' & = & 2y \end{cases}, \qquad \begin{cases} X & = & x' + 1 & = & x & + & y & + & 1 \\ Y & = & y' + 2 & = & & & 2y & + & 2 \end{cases}$$

The conic section is an ellipse centered at the point of coordinates $X = 0, Y = 0$ where $x = 0, y = -1$.



For the conic

$$\mathcal{C} = \{(x,y) \in \mathbf{R}^2 \mid x^2 + 2xy + 2y + 2x - 3 = 0\},$$

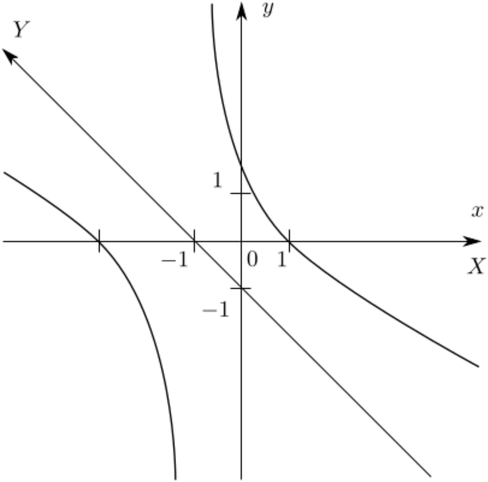
By reduction

$$\begin{aligned} & x^2 + 2xy + 2y + 2x - 3 \\ & \quad = (x+y)^2 - y^2 + 2y + 2x - 3 \\ & \quad = x'^2 - y'^2 + 2y + 2x - 3 \\ & \quad = x'^2 + 2x' - y'^2 - 3 \\ & \quad = (x'+1)^2 - y'^2 - 4 \ = \ 4\left(X^2 - Y^2 - 1\right) \end{aligned}$$

we let

$$\begin{cases} x' & = & x+y \\ y' & = & y \end{cases}, \qquad \begin{cases} X & = & \frac{1}{2}(x'+1) & = & \frac{1}{2}(x+y+1) \\ Y & = & \frac{1}{2}y' & = & \frac{1}{2}y \end{cases}$$

The conic section is a hyperbola.



Quadrics

The same method works in any dimension. Let us deal with the case of dimension 3.

Definition

A quadric Q is a subset of the compound space points whose coordinates $(x, y, z) \in \mathbb{R}^3$ in a certain basis satisfy an equation of the form

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$$

with real coefficients such that a, b, c, d, e and f are not all zero.

The previous method gives the following result.

Proposition

Consider a quadric $Q \subset \mathbb{R}^3$. So there is a change of affine variables of the form

$$\begin{cases} X &= a_{1,1}x + a_{1,2}y + a_{1,3}z + b_1 \\ Y &= a_{2,1}x + a_{2,2}y + a_{2,3}z + b_2 \\ Z &= a_{3,1}x + a_{3,2}y + a_{3,3}z + b_3 \end{cases}$$

which transforms the expression of the quadric into one of the following forms

$$\begin{array}{cc} X^2+Y^2+Z^2=1 & \text{Ellipsoid} \\ X^2+Y^2-Z^2=1 & \text{One-sheet hyperboloid,} \\ X^2-Y^2-Z^2=1 & \text{Two-sheets hyperboloid} \\ X^2+Y^2-Z^2=0 & \text{Cone} \\ X^2+Y^2-Z=0 & \text{Elliptic paraboloid} \\ X^2-Y^2-Z=0 & \text{Hyperbolic paraboloid} \end{array}$$

or one of the following forms, which correspond to the case where the quadric is the product of a conic section by a line:

$$\begin{array}{cc} X^2+Y^2=1 & \text{Elliptic cylinder} \\ X^2-Y^2=1 & \text{Hyperbolic cylinder,} \\ X^2-Y=0 & \text{Paraolic cylinder} \end{array}$$

or finally one of the following forms:

$$X^2 = 1,\ X^2 - Y^2 = 0,\ X^2 + Y^2 + Z^2 = 0,\ X^2 + Y^2 = 0,\ X^2 = 0,$$

$$X^2 + Y^2 + Z^2 = -1,\ X^2 + Y^2 = -1,\ X^2 = -1,$$

In which case the quadric corresponds to two parallel planes, two separate planes cants, a point, a line, a plane or is empty.

Here again, we can use the signature of the quadratic form

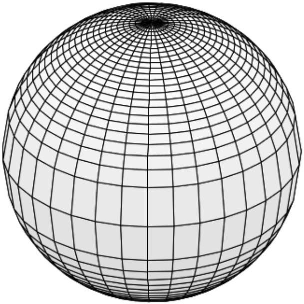
$$Q(x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

to determine the nature of the quadric.

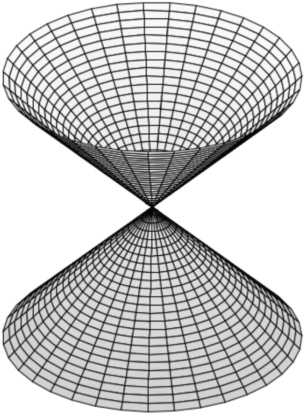
$$Q(x,y,z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

$$\begin{array}{lcc}
\text{signature } (3,0) & \rightarrow & \text{ellipsoid, point or void;} \\
\text{signature } (2,1) & \rightarrow & \text{hyperboloid with one or two layers, cone;} \\
\text{signature } (2,0) & \rightarrow & \text{paraboloid or elliptical cylinder, straight, empty;} \\
\text{signature } (1,1) & \rightarrow & \text{paraboloid or hyperbolic cylinder, two secant planes;} \\
\text{signature } (1,0) & \rightarrow & \text{parabolic cylinder, two parallel planes, plane, void.}
\end{array}$$

The quadric is a Sphere:



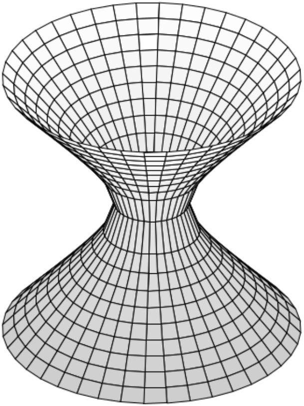
$$X^2 + Y^2 + Z^2 = 1$$



Cône

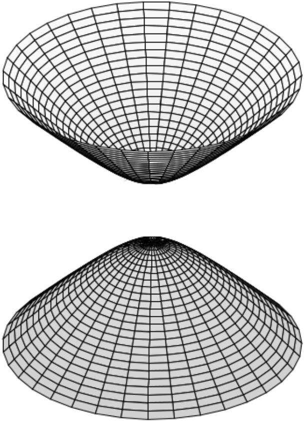
$$X^2 + Y^2 - Z^2 = 0$$

The quadric is a one-sheet hyperboloid:



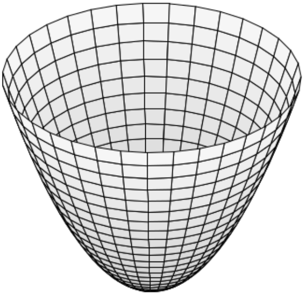
$$X^2 + Y^2 - Z^2 = 1$$

The quadric is a Two-sheet hyperboloid:



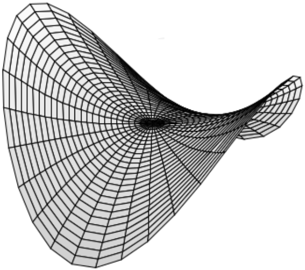
$$X^2 - Y^2 - Z^2 = 1$$

The quadric is an Elliptic paraboloid:



$$X^2 + Y^2 - Z^2 = 0$$

The quadric is a hyperboloic paraboloid:



$$X^2 - Y^2 - Z^2 = 0$$

Example

We consider the quadric Q of $\mathbb{R}^3$ given by:

$$\mathcal{Q} = \{(x,y,z) \in \mathbf{R}^3 \mid x^2 + 3\,y^2 + z^2 + 4\,x\,y + 4\,y\,z - 10z = 0\}.$$

Let us determine the nature of this quadric. By reduction:

$$\begin{aligned}
& x^2 + 3\,y^2 + z^2 + 4\,x\,y + 4\,y\,z - 10z \\
&= (x+2y)^2 - (y-2z)^2 + 5z^2 - 10z \\
&= (x+2y)^2 - (y-2z)^2 + 5(z-1)^2 - 5 \\
&= (x+2y)^2 - (y-2z)^2 + 5(z-1)^2 - 5 \\
&= 5\left(\frac{1}{5}(x+2y)^2 - \frac{1}{5}(y-2z)^2 + (z-1)^2 - 1\right) \\
&= 5\left((\frac{x+2y}{\sqrt{5}})^2 - (\frac{y-2z}{\sqrt{5}})^2 + (z-1)^2 - 1\right).
\end{aligned}$$

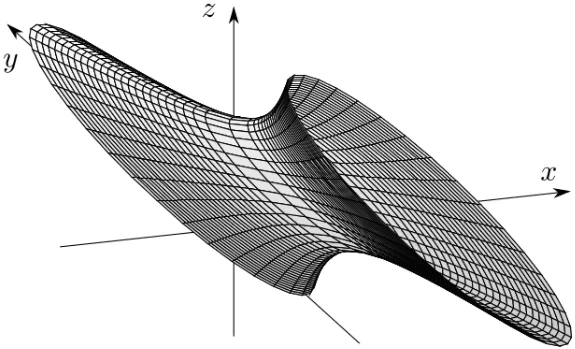
We let

$$\begin{cases} X &= \frac{1}{\sqrt{5}}(x+2y) \\ Y &= z-1 \\ Z &= \frac{1}{\sqrt{5}}(y-2z) \end{cases}$$

The equation becomes:

$$X^2 + Y^2 - Z^2 = 1$$

The quadric is a one-sheet hyperboloid:



Quadratic forms and determinants

Recall that a quadratic form is said to be non-degenerate if its nucleus is null. Equivalently, the determinant of its matrix calculated in a any base is non-zero.
In general, the determinant of the matrix associated with a quadra- tick depends on the base chosen to calculate it. Recall that the matrices B and B' in two different bases (e_i) and (e'_i) are related by the formula

$$B' = P^T BP.$$

So we have

$$det(B') = det(P^T)det(B)det(P) = det(P)^2 det(B)$$

and the two determinants are different in general. We note that they have the same sign. The sign of the determinant of the matrix associated with a quadratic form does not depend on the basis chosen to express this matrix.

Let us examine how this sign is related to the signature (p, q) of the quadratic form Q by placing itself in an orthogonal basis for Q.
In a such basis, the matrix of Q has p strictly positive coefficients and q negative coefficients. The determinant is equal to the product of these coefficients. We deduce the following proposition.

Proposition

A quadratic form of signature (p, q) is non-degenerate if and only if the determinant of its matrix is non-zero. In that case, the sign of this determinant is strictly negative if and only if q is odd

From this we deduce important information on the signature in small dimension. For example, in dimension 2, if the determinant of the matrix of a quadratic form in a certain basis is strictly negative, the signature is necessarily (1,1) and if it is strictly positive, it is (2,0) or (0,2).

We distinguish between these last two cases by looking at the coefficients diagonals of the matrix of Q. These coefficients correspond to the values that Q takes on the vectors of the base in which its matrix is calculated. If Q is signed $(2, 0)$, they are all strictly positive. We cqn conclude:

Proposition

Let E be a vector space of dimension 2, Q a form non degenerate quadratic defined on E and $B = \{b_{i,j}\}$ its matrix in a basis of E. Then the signature of Q is

- *(1,1) if* $det(B) < 0$,
- *(2,0) if* $det(B) >$ *0 and* $b_{1,1} > 0$,
- *(0,2) if* $det(B) > 0$ *and* $b_{1,1} < 0$.

Classification of the quadratic form

The preceding theorems can be formulated as a result of classification. Let us introduce an equivalence relation on the set of quadratic forms defined on a given space.

Proposition

Two quadratic forms Q and Q' defined on a subspace E are said to be equivalent if there is a linear map invertible $f : E \to E$ such that:

$$Q'(x) = Q(f(x)) \text{ for all } x \in E.$$

Two equivalent quadratic forms have the same signature because the application f puts in bijection the spaces on which the two forms are definite positive or definite negative. We will show the converse.

Sylvester's inertia theorem

Theorem

Two quadratic forms defined on a vector space of finite dimension are equivalent if and only if they have the same signature.

Proof.

Let Q and Q' be two quadratic forms having the same signature. We saw that it exists two bases (e_i) and (e'_i) such that

$$Q(x_1e_1 + \dots + x_ne_n) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2,$$

$$Q'(x_1e'_1 + \dots + x_ne'_n) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2.$$

Let us denote by f the application which sends the base (e_i) on the base (e'_i) and let $x = x_1e_1 + \dots + x_ne_n$. We then have $f(x) = x_1e'_1 + \dots + x_ne'_n$, which implies $Q'(f(x)) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2 = Q(x)$. This is the desired equality. □

As an application, we give the classification of quadratic forms defined on a two-dimensional space. Just list the values possible for the signature (p, q) and to give an example of a quadra- tick for each of these values.

Proposition

Let E be a two-dimensional vector space and Q a quadratic form on E. Then there exists a basis (e_1, e_2) of E in which Q takes one of the following forms:

- $Q(x_1e_1 + x_2e_2) = 0$ *signature (0,0)*
- $Q(x_1e_1 + x_2e_2) = x_1^2$ *signature (1.0)*
- $Q(x_1e_1 + x_2e_2) = -x_1^2$ *signature (0,1)*
- $Q(x_1e_1 + x_2e_2) = x_1^2 + x_2^2$ *signature (2.0)*
- $Q(x_1e_1 + x_2e_2) = -x_1^2 - x_2^2$ *signature (0.2)*
- $Q(x_1e_1 + x_2e_2) = x_1^2 - x_2^2$ *signature (1.1)*