Linear, bilinear and quadratic forms

Lecture 8

Mathematics - 1^{st} year, English

Faculty of Computer Science
Alexandru Ioan Cuza University of Iasi

e-mail: corina.forascu@gmail.com

facebook: Corina Forăscu

November 28, 2018

Outline of the lecture

1. Linear forms

2. Bilinear forms

3. Quadratic forms

Linear forms

Definition

Let $(V, +, \cdot)$ be a linear space.

- A linear mapping $f : V \to \mathbb{R}$ is called a *linear form* or a *linear functional*.
- The linear space $\mathrm{L}(V; \mathbb{R})$ of all linear forms is called the *dual* of V and is denoted V^*.

Proposition

Let $(V, +, \cdot)$ be a finite-dimensional linear space. Then V^* is also finite-dimensional and $\dim V^* = \dim V$.

Proposition

Let $(V, +, \cdot)$ be a finite-dimensional linear space. If $\mathbf{v} \in V \setminus \{\mathbf{0}_V\}$ then there exists $f \in V^*$ such that $f(\mathbf{v}) \neq 0$.

Consequence. If $\mathbf{u}, \mathbf{v} \in V$ and $\mathbf{u} \neq \mathbf{v}$ then there exists $f \in V^*$ such that $f(\mathbf{u}) \neq f(\mathbf{v})$.

Bidual and evaluation map

Definition

Let $(V, +, \cdot)$ be a linear space.

- The dual of V^*, denoted by V^{**}, is called the *bidual* of V.
- The function $\psi : V \to V^{**}$ defined by

$$\psi(\mathbf{v})(f) := f(\mathbf{v}),\ \mathbf{v} \in V,\ f \in V^*$$

  is called the *evaluation map*.

The evaluation map is well-defined and it is linear:
1. It is clear that $\psi(\mathbf{v}) : V^* \to \mathbb{R}$. If $\alpha, \beta \in \mathbb{R}$ and $f, g \in V^*$, then

$$\begin{aligned} \psi(\mathbf{v})(\alpha f + \beta g) &= (\alpha f + \beta g)(\mathbf{v}) = \alpha f(\mathbf{v}) + \beta g(\mathbf{v}) \\ &= \alpha\psi(\mathbf{v})(f) + \beta\psi(\mathbf{v})(g). \end{aligned}$$

Hence $\psi(\mathbf{v})$ is linear, *i.e.* $\psi(\mathbf{v}) \in V^{**}$. Therefore, ψ is well-defined.

2. If $\alpha, \beta \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, then

$$\begin{aligned} \psi(\alpha\mathbf{u} + \beta\mathbf{v})(f) &= f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v}) \\ &= \alpha\psi(\mathbf{u})(f) + \beta\psi(\mathbf{v})(f), \ \forall f \in V^*. \end{aligned}$$

This means that $\psi(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\psi(\mathbf{u}) + \beta\psi(\mathbf{v})$. In conclusion, ψ is linear.

3. If V is finite-dimensional, then ψ is a linear isomorphism.

- Indeed, if $\mathbf{v} \in \ker\psi$, then

$$f(\mathbf{v}) = 0, \ \forall f \in V^*.$$

  Supposing that $\mathbf{v} \neq \mathbf{0}_V$ would contradict the existence of some $f \in V^*$ such that $f(\mathbf{v}) \neq 0$. Therefore, $\mathbf{v}$ should be equal to $\mathbf{0}_V$. This implies that $\ker\psi = \{\mathbf{0}_V\}$, *i.e.* ψ is injective.
- On the other hand, $\dim V^{**} = \dim V^* = \dim V$. By the dimension theorem, $\operatorname{rank}\psi = \dim V = \dim V^{**}$, so ψ is surjective, too.

In conclusion, ψ is a linear isomorphism. In this case, ψ is also called the *canonical isomorphism* between V and V^{**}.

Vector hyperplanes

Definition

Let $(V, +, \cdot)$ be a linear space. A linear subspace $W \subseteq V$ is called a (*vector*) *hyperplane* if there exists $f \in V^* \setminus \{\mathbf{0}_{V^*}\}$ such that $\ker f = W$.

Proposition

If $(V, +, \cdot)$ is a finite-dimensional linear space with $\dim V = n \in \mathbb{N}^*$, then a linear subspace $W \subseteq V$ is a hyperplane if and only if $\dim W = n - 1$.

Proof.

[Proof: "$\Rightarrow$"] If $W = \ker f$ for some $f \in V^* \setminus \{\mathbf{0}_{V^*}\}$, then by the dimension theorem,

$$\dim W = \dim(\ker f) = \dim V - \dim(\mathrm{Im}\, f) = n - 1,$$

because $f \neq \mathbf{0}_{V^*}$ and thus $\mathrm{Im}\, f = \mathbb{R}$. ∎

Proof.

[Proof: "$\Leftarrow$"] Conversely, if $\dim W = n-1$, there exists a basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}, \mathbf{b}_n\}$ of V such that $\mathrm{Lin}\{\mathbf{b}_1, \ldots, \mathbf{b}_{n-1}\} = W$. Taking $f : V \to \mathbb{R}$ defined by

$$f(\alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n) := \alpha_n$$

for $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, we have $f \neq \mathbf{0}_{V^*}$ and

$$f(\mathbf{b}_1) = \cdots = f(\mathbf{b}_{n-1}) = 0,$$

implying that $W \subseteq \ker f$ (*i.e.*, $f(\mathbf{v}) = 0$, $\forall \mathbf{v} \in W$). On the other hand, by the direct implication, $\dim(\ker f) = n-1$ and consequently $W = \ker f$. ∎

Let V be a finite-dimensional linear space and $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ a basis of V.

- If W is a hyperplane with $W = \ker f$, where $f \in V^* \setminus \{\mathbf{0}_{V^*}\}$, let $\beta_1 := f(\mathbf{b}_1), \ldots, \beta_n := f(\mathbf{b}_n)$. Then $\mathbf{v} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \in \ker f$ is characterized by the equation

$$\beta_1 x_1 + \cdots + \beta_n x_n = 0. \tag{1}$$

  Hence

$$W = \{x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \in V \mid \beta_1 x_1 + \cdots + \beta_n x_n = 0\}. \tag{2}$$

- Conversely, having $\beta_1, \ldots, \beta_n \in \mathbb{R}$, not all 0, the subset of V defined by the above relation is a hyperplane of V.
- One can show that any linear subspace of V (not only hyperplanes) can be characterized by systems of equations of form (1).
- If $V = \mathbb{R}^n$ and B is the canonical basis, relation (2) can be written as

$$W = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \beta_1 x_1 + \cdots + \beta_n x_n = 0\}.$$

- In the particular cases $n = 2$ and $n = 3$, equation (1) becomes the equation of a *line*, respectively a *plane* passing through the origin.

Affine functionals

The following notion allows us to characterize all the lines (when $n = 2$) and planes (when $n = 3$), not necessarily those passing through the origin.

Definition

Let $(V, +, \cdot)$ be a linear space. A function $f : V \to \mathbb{R}$ is called an *affine functional* if there exist a linear functional $f_0 \in V^*$ and a constant $c \in \mathbb{R}$ such that $f(\mathbf{v}) = f_0(\mathbf{v}) + c$, $\forall \mathbf{v} \in V$.

For an affine functional $f : V \to \mathbb{R}$ one can define its *kernel* in the same way as for linear functionals, *i.e.* $\ker f := \{\mathbf{v} \in V \mid f(\mathbf{v}) = 0\}$.

Definition

Let $(V, +, \cdot)$ be a linear space. A subset $U \subseteq V$ is called an *affine hyperplane* if there exists a non-constant affine functional $f : V \to \mathbb{R}$ such that $\ker f = U$.

- In other words, U is affine hyperplane if there exist a vector hyperplane W and a vector $\mathbf{v}_0 \in V$ such that

$$U = W + \mathbf{v}_0 := \{\mathbf{v} + \mathbf{v}_0 \mid \mathbf{v} \in W\}.$$

- If V is finite-dimensional with a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, then affine hyperplanes are given by subsets of the form

$$U = \{x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \in V \mid \beta_1 x_1 + \cdots + \beta_n x_n + c = 0\},$$

where $c, \beta_1, \ldots, \beta_n \in \mathbb{R}$.

- In the cases $n = 2$ and $n = 3$, the affine hyperplanes are the lines, respectively the planes.

Bilinear forms

Definition

Let $(V, +, \cdot)$ and $(W, +, \cdot)$ two linear spaces. A function $g : V \times W \to \mathbb{R}$ is called a *bilinear form* (*bilinear map/mapping*) on $V \times W$ if the following conditions are fulfilled:

1. $g(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}) = \alpha g(\mathbf{u}, \mathbf{w}) + \beta g(\mathbf{v}, \mathbf{w})$, $\forall \alpha, \beta \in \mathbb{R}$, $\forall \mathbf{u}, \mathbf{v} \in V$, $\forall \mathbf{w} \in W$;
2. $g(\mathbf{v}, \lambda\mathbf{w} + \mu\mathbf{z}) = \lambda g(\mathbf{v}, \mathbf{w}) + \mu g(\mathbf{v}, \mathbf{z})$, $\forall \lambda, \mu \in \mathbb{R}$, $\forall \mathbf{v} \in V$, $\forall \mathbf{w}, \mathbf{z} \in W$.

In the case $W = V$, a bilinear form on $V \times V$ is also called *bilinear form* (*functional*, *map/mapping*) on V.

1. Suppose now that V and W are finite-dimensional, with bases $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\bar{B} = \{\bar{\mathbf{b}}_1, \ldots, \bar{\mathbf{b}}_m\}$ on V, respectively W.

- If $\mathbf{v} \in V$ and $\mathbf{w} \in W$ having $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\beta_1, \ldots, \beta_m \in \mathbb{R}$ as coordinates with respect to the bases B, respectively $\bar{B}$, then

$$g(\mathbf{v}, \mathbf{w}) = g\left(\sum_{i=1}^{n} \alpha_i \mathbf{b}_i, \sum_{j=1}^{m} \beta_j \bar{\mathbf{b}}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i \beta_j g(\mathbf{b}_i, \bar{\mathbf{b}}_j).$$

- The scalars $a_{ij} := g(\mathbf{b}_i, \bar{\mathbf{b}}_j)$, $1 \leq i \leq n$, $1 \leq j \leq m$ are called the *coefficients* of the bilinear form g with respect to the bases B and $\bar{B}$;
- the matrix $A^g_{B,\bar{B}} := (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq m}}$ in $\mathscr{M}_{nm}$ is called the *matrix of the bilinear form* g with respect to the bases B, $\bar{B}$.

2. If $B' = \{\mathbf{b}'_1, \ldots, \mathbf{b}'_n\}$ is another basis of V and $\bar{B}' = \{\bar{\mathbf{b}}'_1, \ldots, \bar{\mathbf{b}}'_m\}$ is another basis of W, let us denote $S = (s_{ij})_{1\leq i,j\leq n} \in \mathscr{M}_n$ the transition matrix from B to B' and $\bar{S} = (\bar{s}_{ij})_{1\leq i,j\leq m} \in \mathscr{M}_m$ the transition matrix from $\bar{B}$ to $\bar{B}'$.

- Then the matrix of g with respect to the bases B' and $\bar{B}'$ can be written as

$$A^g_{B',\bar{B}'} = S \cdot A^g_{B,\bar{B}} \cdot \bar{S}^{\mathrm{T}}.$$

- It can be proven that $\operatorname{rank} A^g_{B',\bar{B}'} = \operatorname{rank} A^g_{B,\bar{B}}$, so the rank of the matrix of the bilinear form doesn't depend on the bases of reference. This commun value is called the *rank* of g and is denoted by $\operatorname{rank} g$.

Kernel of a bilinear form

- Fixing $\mathbf{w} \in W$, the bilinear form $g : V \times W \to \mathbb{R}$ defines a linear functional $f_{\mathbf{w}} : V \to \mathbb{R}$, by

$$f_{\mathbf{w}}(\mathbf{v}) := g(\mathbf{v}, \mathbf{w}),\ \mathbf{v} \in V.$$

  Allowing now $\mathbf{w}$ to variate, the mapping $\mathbf{w} \mapsto f_{\mathbf{w}}$ defines a linear operator $g' : W \to V^*$.
- In a similar way, one can define a linear operator $g'' : V \to W^*$ by $g''(\mathbf{v}) := h_{\mathbf{v}}$, where the linear functional $h_{\mathbf{v}} \in W^*$ is introduced by

$$h_{\mathbf{v}}(\mathbf{w}) := g(\mathbf{v}, \mathbf{w}),\ \mathbf{w} \in V.$$

Definition

Let $g : V \times W \to \mathbb{R}$ be a bilinear form and the associated linear operators $g' : W \to V^*$ and $g'' : V \to W^*$ introduced above. The linear subspace $\ker g' \subseteq W$ is called the *right kernel* of g, while the linear subspace $\ker g'' \subseteq V$ is called the *left kernel* of g.
If $Ker(g') = \{\mathbf{0}_W\}$ and $Ker(g'') = \{\mathbf{0}_V\}$, then the bilinear form g is called *non-degenerate*.

Symmetric bilinear forms

Definition

A bilinear form $g : V \times V \to \mathbb{R}$ is called *symmetric* if

$$g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u}), \forall \mathbf{u}, \mathbf{v} \in V,$$

respectively *antisymmetric* if

$$g(\mathbf{u}, \mathbf{v}) = -g(\mathbf{v}, \mathbf{u}), \forall \mathbf{u}, \mathbf{v} \in V.$$

Proposition

Let $g : V \times V \to \mathbb{R}$ be a symmetric bilinear form or an antisymmetric linear form. Then its right kernel coincides with its left kernel.

For such a bilinear form, the left kernel (which coincides with the right kernel) is called the *kernel* of g and is denoted by $\ker g$.

Dimension theorem for bilinear forms

Proposition

Let $(V, +, \cdot)$ *be a finite-dimensional linear space and* $g : V \times V \to \mathbb{R}$ *a symmetric bilinear form. Then*

$$\operatorname{rank} g + \dim(\ker g) = \dim V.$$

Remark. By the above result, a necessary and sufficient condition for a symmetric bilinear form to be non-degenerate is that $\operatorname{rank} g = \dim V$.

Definition

Let $g : V \times V \to \mathbb{R}$ be a symmetric bilinear form.

- Two vectors $\mathbf{u}, \mathbf{v} \in V$ are called *orthogonal* with respect to g if $g(\mathbf{u}, \mathbf{v}) = 0$.
- If U is a non-empty subset of V, we say that U is *orthogonal* with respect to g (or *g-orthogonal*) if $g(\mathbf{u}, \mathbf{v}) = 0$ for any distinct $\mathbf{u}, \mathbf{v} \in U$.
- If U is a non-empty subset of V, the set $\{\mathbf{v} \in V \mid g(\mathbf{u}, \mathbf{v}) = 0, \ \forall \mathbf{u} \in U\}$ is a linear subspace of V, called the *orthgonal complement* of U with respect to g, denoted $U^{\perp_g}$.

Sylvester's law of inertia

Theorem

Let $(V, +, \cdot)$ *be a finite-dimensional linear space and* $g : V \times V \to \mathbb{R}$ *a symmetric bilinear form. If* $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ *is a basis of* V *which is* g*-orthogonal, then* $\operatorname{rank} g$ *is precisely the number of elements among* $g(\mathbf{b}_1, \mathbf{b}_1), g(\mathbf{b}_2, \mathbf{b}_2), \dots, g(\mathbf{b}_n, \mathbf{b}_n)$ *which are non-zero.*

Theorem (Sylvester's law of inertia)

Let $(V, +, \cdot)$ *be a finite-dimensional linear space and* $g : V \times V \to \mathbb{R}$ *a symmetric bilinear form. Then there exist* $p, q, r \in \mathbb{N}$ *such that for every* g*-orthogonal basis* $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ *of* V*,* p*,* q *and* r *represent the number of positive, negative, respectively null elements among* $g(\mathbf{b}_1, \mathbf{b}_1), g(\mathbf{b}_2, \mathbf{b}_2), \dots, g(\mathbf{b}_n, \mathbf{b}_n)$.

- The numbers p and q are called the *positive*, respectively the *negative index of inertia*.
- The triple (p, q, r) is called the *signature* of g.
- Of course, $p + q + r = n$ ($n = \dim V$); moreover, $\operatorname{rank} g = p + q$.

Quadratic forms

Definition

Let $(V, +, \cdot)$ be a linear space and $g : V \times V \to \mathbb{R}$ a symmetric bilinear form. The function $h : V \to \mathbb{R}$, defined by

$$h(\mathbf{v}) := g(\mathbf{v}, \mathbf{v}),\ \mathbf{v} \in V$$

is called the *quadratic form* (*functional*) associated to g.

Remark. Since
$h(\mathbf{u}+\mathbf{v}) = g(\mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}) = g(\mathbf{u}, \mathbf{u}) + g(\mathbf{u}, \mathbf{v}) + g(\mathbf{v}, \mathbf{u}) + g(\mathbf{v}, \mathbf{v})$ and $g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u})$, we have

$$h(\mathbf{u}+\mathbf{v}) = h(\mathbf{u}) + 2g(\mathbf{u}, \mathbf{v}) + h(\mathbf{v}),\ \forall \mathbf{u}, \mathbf{v} \in V.$$

From this formula we can retreive g from h:

$$g(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\left[h(\mathbf{u}+\mathbf{v}) - h(\mathbf{u}) - h(\mathbf{v})\right],\ \forall \mathbf{u}, \mathbf{v} \in V$$

or

$$g(\mathbf{u}, \mathbf{v}) = \frac{1}{4}\left[h(\mathbf{u}+\mathbf{v}) - h(\mathbf{u}-\mathbf{v})\right],\ \forall \mathbf{u}, \mathbf{v} \in V.$$

- Suppose now that V is a finite-dimensional space and $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis of V.
- Let $A^g_{B,B} = (a_{ij})_{1 \leq i,j \leq n}$ be the matrix of g with respect to B. If $x_1, \ldots, x_n \in \mathbb{R}$ are the coefficients of a vector $\mathbf{v} \in V$ with respect to B, then

$$h(\mathbf{v}) = h(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

  The right-hand side of this relation is a homogeneous polynomial of degree 2, called the *quadratic polynomial* associated to the quadratic form h and the basis B.
- The determinant of the symmetric matrix $A^g_{B,B}$ is called the *discriminant* of h with respect to the basis B. Its sign does not depend on the basis B.
- We say that h is a *non-degenerate quadratic form* if g is a non-degenerate bilinear functional form, *i.e.* the discriminant of h (in any basis) is not zero ($\operatorname{rank} A^g_{B,B} = \operatorname{rank} g = n$). Otherwise, we say that h is a *degenerate quadratic form*.
- If (p, q, r) is the signature of g, we also call it the *signature* of the quadratic form h.

Reduced form of a bilinear form

Definition

Let $(V, +, \cdot)$ be a finite-dimensional linear space and $h : V \to V$ a quadratic form associated to some symmetric bilinear form $g : V \times V \to \mathbb{R}$.

- If B is a basis of V such that the matrix of g is diagonal, we call *canonical* (*reduced*) *form* of h the quadratic polynomial associated to h and B.
- A canonical form of h is called *normal* if the diagonal matrix associated to g has on its diagonal only the elements 1, -1 and 0.

If $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis of V giving a canonical form $\omega_1 x_1^2 + \omega_2 x_2^2 + \cdots + \omega_n x_n^2$ of h, then $B' = \{c_1\mathbf{b}_1, \ldots, c_n\mathbf{b}_n\}$ gives a normal form of h, where $c_i = 1$ if $\omega_i = 0$, while $c_i = \frac{1}{\sqrt{|\omega_i|}}$ if $\omega_i \neq 0$, for $1 \leq i \leq n$.

Gauss method

Theorem (Gauss method of reducing a quadratic form)

Let $(V, +, \cdot)$ be an n-dimensional linear space and $h : V \to \mathbb{R}$ a quadratic form. Then there exists a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of V and $\omega_1, \ldots, \omega_n \in \mathbb{R}$ such that for any $x_1, \ldots, x_n \in \mathbb{R}$ we have

$$h(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = \omega_1 x_1^2 + \omega_2 x_2^2 + \cdots + \omega_n x_n^2.$$

Remarks.

- The quadratic polynomial $\omega_1 x_1^2 + \omega_2 x_2^2 + \cdots + \omega_n x_n^2$ is then a reduced form of h (the matrix of g with respect to $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a diagonal matrix with entries $\omega_1, \ldots, \omega_n$).
- If (p, q, r) is the signature of h, then among the coefficients $\omega_1, \ldots, \omega_n$, p are positive, q are negative and r are equal to 0.

Jacobi method

Theorem (Jacobi method of reducing a quadratic form)

Let $(V, +, \cdot)$ be an n-dimensional linear space and $h : V \to \mathbb{R}$ a quadratic form. Let Δ_i, $1 \le i \le n$ the principal minors of the associated matrix $(a_{ij})_{1 \le i,j \le n}$ with respect to a basis of V, i.e.

$$\Delta_i = \begin{vmatrix} a_{11} & \dots & a_{1i} \\ a_{21} & \dots & a_{2i} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{ii} \end{vmatrix}, \; 1 \le i \le n.$$

If $\Delta_i \neq 0$, $\forall i \in \{1, \dots, n\}$, then h can be reduced to the canonical form

$$\mu_1 x_1^2 + \mu_2 x_2^2 + \dots + \mu_n x_n^2,$$

where $\mu_j = \frac{\Delta_{j-1}}{\Delta_j}$, $\forall j = \{1, \dots, n\}$, with $\Delta_0 = 1$.

Definition

Let $(V, +, \cdot)$ be an n-dimensional linear space and $h : V \to \mathbb{R}$ a quadratic form with signature (p, q, r).

- If $p = n$, h is called a *positive-definite* quadratic form.
- If $q = 0$, the quadratic form h is called *positive semidefinite*.
- If $q = n$, h is called a *negative-definite* quadratic form.
- If $p = 0$, the quadratic form h is called *negative semidefinite*.
- The quadratic form h is called *undefined* if $p > 0$ and $q > 0$.

Let Δ_i, $1 \leq i \leq n$ be the principal minors of the associated matrix with respect to an arbitrary basis. Then h is positive-definite if and only if

$$\Delta_i > 0, \ \forall i \in \{1, \ldots, n\}$$

and h is negative-definite if and only if

$$(-1)^i \Delta_i > 0, \ \forall i \in \{1, \ldots, n\}.$$

Eigenvalues method

Theorem (Eigenvalues method of reducing a quadratic form)

Let $(V, \langle\cdot,\cdot\rangle)$ be a finite-dimensional prehilbertian space with $\dim V = n$. *Then there exists an orthonormal basis with respect to which h has the canonical form*

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2,\ x_1, x_2, \ldots, x_n \in \mathbb{R},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$.

- In fact, $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the associated matrix with respect to any basis of V.
- The method of the proof is similar to the diagonalization algorithm for linear operators.

Non-homogeneous quadratic functionals

Definition

Let $(V, +, \cdot)$ be a linear space, $h : V \to \mathbb{R}$ a quadratic form and $f : V \to \mathbb{R}$ an affine functional. The sum $h + f$ is called a *non-homogeneous quadratic functional* on V.

- If V is finite-dimensional and $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of basis of V, then for any $x_1, \ldots, x_n \in \mathbb{R}$

$$(3) \qquad (h+f)(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j + \sum_{i=1}^{n} b_ix_i + c,$$

  where $A = (a_{ij})_{1\le i,j\le n}$ is the matrix associated to h and $b_1, \ldots, b_n, c \in \mathbb{R}$.
- The right-hand side of this equality is called the *quadratic polynomial* associated to $h + f$ (which is a polynomial of degree 2).
- If $V = \mathbb{R}^n$ and B is its canonical basis, then (3) can be written as

$$(4) \qquad (h+f)(\mathbf{x}) = \rho(\mathbf{x}) := \langle A\mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{b}, \mathbf{x}\rangle + c, \ \forall \mathbf{x} \in \mathbb{R}^n,$$

  where $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ and the vectors $\mathbf{x} \in \mathbb{R}^n$ are interpreted as column matrices.

- Conversely, for arbitrary symmetric matrix $A \in \mathscr{M}_n$, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$, the function $\rho : V \to \mathbb{R}$ defined by (4), *i.e.*

$$\rho(\mathbf{x}) := \langle A\mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{x} \rangle + c, \ \forall \mathbf{x} \in \mathbb{R}^n$$

  defines a non-homogeneous quadratic functional on V.

- Moreover, A can be taken not necessarily symmetric, since

$$\begin{aligned} \langle A\mathbf{x}, \mathbf{x} \rangle &= \frac{1}{2}\langle A\mathbf{x}, \mathbf{x} \rangle + \frac{1}{2}\langle \mathbf{x}, A\mathbf{x} \rangle \\ &= \frac{1}{2}\langle A\mathbf{x}, \mathbf{x} \rangle + \frac{1}{2}\langle A^{\mathrm{T}}\mathbf{x}, \mathbf{x} \rangle = \left\langle \frac{1}{2}\left(A + A^{\mathrm{T}}\right)\mathbf{x}, \mathbf{x} \right\rangle, \end{aligned}$$

  so the matrix A can be replaced by the symmetric matrix $\frac{1}{2}\left(A + A^{\mathrm{T}}\right)$.

Normal form of non-homogeneous quadratic functionals

Let us now consider an *affine change of coordinates*, *i.e.* a transformation of the form

$$\mathbf{x}' = S\mathbf{x} + \mathbf{x}_0,$$

where $S \in \mathscr{M}_n$ is a non-singular matrix and $\mathbf{x}_0 \in \mathbb{R}^n$. Then

$$\begin{aligned}\rho(\mathbf{x}) &= \left\langle AS^{-1}(\mathbf{x}' - \mathbf{x}_0), S^{-1}(\mathbf{x}' - \mathbf{x}_0)\right\rangle + \left\langle \mathbf{b}, S^{-1}(\mathbf{x}' - \mathbf{x}_0)\right\rangle + c \\ &= \left\langle \left(S^{-1}\right)^{\mathrm{T}} AS^{-1}\mathbf{x}', \mathbf{x}'\right\rangle - \left\langle 2\left(S^{-1}\right)^{\mathrm{T}} AS^{-1}\mathbf{x}_0 + \left(S^{-1}\right)^{\mathrm{T}}\mathbf{b}, \mathbf{x}'\right\rangle \\ &\quad + \left(c - \left\langle \mathbf{b}, S^{-1}\mathbf{x}_0\right\rangle\right).\end{aligned}$$

Suppose now that S is the transition matrix from the canonical basis to an orthonormal basis giving the canonical form in eigenvalues method of reduction. Therefore, S is an orthonormal matrix ($S^{-1} = S^{\mathrm{T}}$) and
$SAS^{\mathrm{T}} = D := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.
Consequently, we have:

$$\rho(\mathbf{x}) = \langle D\mathbf{x}', \mathbf{x}'\rangle - 2\left\langle S\left(AS^T\mathbf{x}_0 + \frac{1}{2}\mathbf{b}\right), \mathbf{x}'\right\rangle + \left(c - \left\langle \mathbf{b}, S^{-1}\mathbf{x}_0\right\rangle\right).$$

- If A is non-singular, we can take $\mathbf{x}_0 := -\frac{1}{2}SA^{-1}\mathbf{b}$, obtaining

$$\rho(\mathbf{x}) = \langle D\mathbf{x}', \mathbf{x}' \rangle + c_0,$$

where $c_0 := \langle D\mathbf{x}_0, \mathbf{x}_0 \rangle - \langle S\mathbf{b}, \mathbf{x}_0 \rangle + c$. Therefore, by the change of coordinates $\mathbf{x}' = S\mathbf{x} - \frac{1}{2}SA^{-1}\mathbf{b}$, we obtain

$$\rho(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (x_i')^2 + c_0, \ \forall \mathbf{x} \in \mathbb{R}^n,$$

where x_i' are the coordinates of $\mathbf{x}$ with respect to the new orthogonal basis.
- If $\det A = 0$, then by letting $\mathbf{x}_0 := \mathbf{0}$, we obtain

$$\rho(\mathbf{x}) = \langle D\mathbf{x}', \mathbf{x}' \rangle + \langle S\mathbf{b}, \mathbf{x}' \rangle + c_0,$$

where $c_0 := -\langle S\mathbf{b}, \mathbf{x}_0 \rangle + c$.
If (p, q, r) is the signature of h, we have $r > 0$ and $n - r$ is the rank of A; one can further find an adequate basis B'' such that

$$\rho(\mathbf{x}) = \sum_{i=1}^{n-r} \lambda_i (x_i'')^2 + \gamma x_{n-r+1}'',$$

where $x_1'', \ldots, x_n''$ are the coordinates of $\mathbf{x}$ with respect to this new basis and $\gamma \in \mathbb{R}$.

Geometric classification

From the geometric point of view,

$$\ker \rho := \{\mathbf{x} \in \mathbb{R}^n \mid \rho(\mathbf{x}) = 0\}$$

is a *conic* in the case $n = 2$, a *quadric* if $n = 3$, a *hyperquadric* if $n \geq 4$.

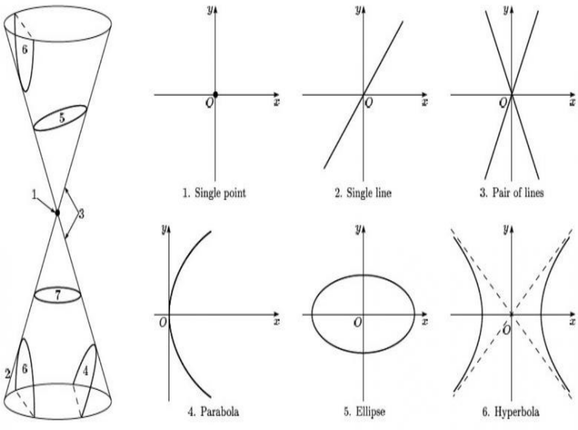
1. Case $n = 1$: the *normal forms* of ρ are:

- $x^2 + 1$ ($\ker \rho = \emptyset$: two "imaginary" points);
- $x^2 - 1$ ($\ker \rho = \{-1, 1\}$: two distinct points);
- x^2 ($\ker \rho = \{0\}$: two identical points).

2. Case $n = 2$: we have nine types of conics, according to the normal form of ρ:

- $x_1^2 + x_2^2 + 1 = 0$ ($\emptyset$: "imaginary" *ellipse*);
- $x_1^2 - x_2^2 + 1 = 0$ (*hyperbola*);
- $x_1^2 + x_2^2 - 1 = 0$ (*ellipse*);
- $x_1^2 - 2x_2 = 0$ (*parabola*);
- $x_1^2 + x_2^2 = 0$ (a point: two "imaginary", conjugate lines);
- $x_1^2 - x_2^2 = 0$ (two intersecting lines);
- $x_1^2 + 1 = 0$ ($\emptyset$: two "imaginary" lines);
- $x_1^2 - 1 = 0$ (two parallel lines);
- $x_1^2 = 0$ (two identical lines).

Matematica-1/Slides/graphics/Conics.pdf



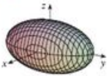
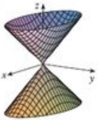
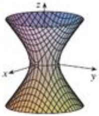
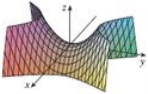
3. Case $n = 3$: we have 17 types of quadrics, characterized by the following normal forms:

- $x_1^2 + x_2^2 + x_3^2 + 1 = 0$ ("imaginary" *ellipsoid*);
- $x_1^2 + x_2^2 + x_3^2 - 1 = 0$ (*ellipsoid*);
- $x_1^2 + x_2^2 - x_3^2 - 1 = 0$ (*hyperboloid of one sheet*);
- $x_1^2 - x_2^2 - x_3^2 - 1 = 0$ (*hyperboloid of two sheets*);
- $x_1^2 + x_2^2 + x_3^2 = 0$ (a point: "imaginary" *cone*);
- $x_1^2 + x_2^2 - x_3^2 = 0$ (*cone*);
- $x_1^2 + x_2^2 - 2x_3 = 0$ (*elliptic paraboloid*);
- $x_1^2 - x_2^2 - 2x_3 = 0$ (*hyperbolic paraboloid*).

The remaining 9 normal forms are the same as those in the case $n = 2$, which in $\mathbb{R}^3$ represent *cylinders* of different types: elliptic, hyperbolic or parabolic.
The first 6 quadrics are *non-singular quadrics*, while the others are *singular quadrics*.

Matematica-1/Slides/graphics/Quadrics.jpg

Graphs of quadric surfaces

Surface	Equation	Surface	Equation
Ellipsoid 	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ All traces are ellipses. If $a=b=c$, the ellipsoid is a sphere.	Cone 	$\frac{z^2}{c^2}=\frac{x^2}{a^2}+\frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces in the planes $x=k$ and $y=k$ are hyperbolas if $k\neq 0$ but are pairs of lines if $k=0$.
Elliptic Paraboloid 	$\frac{z}{c}=\frac{x^2}{a^2}+\frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces are parabolas. The variable raised to the first power indicates the axis of the paraboloid.	Hyperboloid of One Sheet 	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ Horizontal traces are ellipses. Vertical traces are hyperbolas. The axis of symmetry corresponds to the variable whose coefficient is negative.
Hyperbolic Paraboloid 	$\frac{z}{c}=\frac{x^2}{a^2}-\frac{y^2}{b^2}$ Horizontal traces are hyperbolas. Vertical traces are parabolas. The case where $c<0$ is illustrated.	Hyperboloid of Two Sheets 	$-\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ Horizontal traces in $z=k$ are ellipses if $k>c$ or $k<-c$. Vertical traces are hyperbolas. The two minus signs indicate two sheets.

- M. Ariciuc, S. Roatesi, *Lecții de algebră liniară și geometrie analitică*, Editura Matrix Rom, București, 2008.
- K. C. Border, *More than you wanted to know about quadratic forms*, Caltech, 2016.
- K. Conrad, *Bilinear Forms*, Notes on Advanced Linear Algebra, 2015.
- C. Costinescu, *Algebră liniară și aplicații în geometrie*, Editura Matrix Rom, București, 2005.
- D. Drăghici, *Algebră*, Editura Didactică și Pedagogică, București, 1972.
- G. Galbură, F. Radó, *Geometrie*, Ed. Didactică și Pedag., București, 1979.
- M. Neagu, *Geometria curbelor și suprafețelor. Teorie și aplicații*, Editura Matrix Rom, București, 2013.
- P. Ott, *Bilinear and Quadratic Forms*, Prof. Robert Beezer's Notes on Advanced Linear Algebra, 2014.
- I. Radomir, *Elemente de algebră vectorială, geometrie și calcul diferențial*, Editura Albastră, Cluj-Napoca, 2000.