LECTURE 8
LINEAR, BILINEAR AND QUADRATIC FORMS

1. Linear forms

In this section we consider linear mappings between a linear space and the scalar space.

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

a) A linear mapping $f : V \to \mathbb{R}$ is called a *linear form* or a *linear functional*.

b) The linear space $\mathrm{L}(V; \mathbb{R})$ of all linear forms is called the *dual* of V and is denoted V^*.

Proposition 1.1. *Let $(V, +, \cdot)$ be a finite-dimensional linear space. Then V^* is also finite-dimensional and* $\dim V^* = \dim V$.

Proposition 1.2. *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 $f(\mathbf{v}) \neq 0$.*

Remark. An easy consequence of the above result is that if $\mathbf{u}, \mathbf{v} \in V$ and $\mathbf{u} \neq \mathbf{v}$ then there exists $f \in V^*$ such that $f(\mathbf{u}) \neq f(\mathbf{v})$.

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

a) The dual of V^*, denoted by V^{**}, is called the *bidual* of V.

b) 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:

a. It is clear that $\psi(\mathbf{v}) : V^* \to \mathbb{R}$. If $\alpha, \beta \in \mathbb{R}$ and $f, g \in V^*$, then

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

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

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

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

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

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 Proposition 1.2, which asserts 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, by Proposition 1.1, $\dim V^{**} = \dim V^* = \dim V$. 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^{**}.

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 1.3. *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. If $W = \ker f$ for some linear functional $f \in V^* \setminus \{\mathbf{0}_{V^*}\}$, then by the dimension theorem,

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

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

Conversely, if $\dim W = n - 1$, there exists a basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_{n-1}, \mathbf{b}_n\}$ of V such that $\operatorname{Lin}\{\mathbf{b}_1, \dots, \mathbf{b}_{n-1}\} = W$. 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 $(V, +, \cdot)$ 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 the condition $\mathbf{v} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n \in \ker f$ 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 (1-dimensional) line, respectively a (2-dimensional) plane passing through the origin.

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.

2. 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:

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

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

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 $B, \bar{B}$.

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 $S = (s_{ij})_{1 \le i,j \le n} \in \mathscr{M}_n$ the transition matrix from B to B' and $\bar{S} = (\bar{s}_{ij})_{1 \le i,j \le m} \in \mathscr{M}_m$ 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 $\operatorname{rank} g$.

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.*

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 2.1. *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 $\ker g$.

The next result plays a similar role to symmetric bilinear forms as does the dimension theorem for linear operators.

Proposition 2.2. *Let $(V, +, \cdot)$ be a finite-dimensional linear space and $g : V \times V \to K$ 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.

a) Two vectors $\mathbf{u}, \mathbf{v} \in V$ are called *orthogonal* (or *conjugate*) with respect to g if $g(\mathbf{u}, \mathbf{v}) = 0$.

b) 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 $\mathbf{u}, \mathbf{v} \in U$.

c) 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 $U^{\perp_g}$.

Remark. If W is a finite dimensional subspace of V with $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ a basis of W, then $\mathbf{v} \in W^{\perp_g}$ if and only if $g(\mathbf{b}_k, \mathbf{v}) = 0$, $\forall l \in \{1, \ldots, n\}$.

Theorem 2.3. *Let $(V, +, \cdot)$ be a finite-dimensional linear space and $g : V \times V \to \mathbb{R}$ a symmetric bilinear form. If $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis of V which is g-orthogonal, then* $\operatorname{rank} g$ *is precisely the number of elements among $g(\mathbf{b}_1, \mathbf{b}_1), g(\mathbf{b}_2, \mathbf{b}_2), \ldots, g(\mathbf{b}_n, \mathbf{b}_n)$ which are non-zero.*

In fact, the number of positive values (and negative values) among $g(\mathbf{b}_1, \mathbf{b}_1), g(\mathbf{b}_2, \mathbf{b}_2), \ldots, g(\mathbf{b}_n, \mathbf{b}_n)$ is invariant with respect to B, as the following result asserts:

Theorem 2.4 (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 $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ of V, p, q and r represent the number of positive, negative, respectively null elements among $g(\mathbf{b}_1, \mathbf{b}_1), g(\mathbf{b}_2, \mathbf{b}_2), \ldots, g(\mathbf{b}_n, \mathbf{b}_n)$.*

The triple (p, q, r) is called the *signature* of g. Of course, $p + q + r = n$ ($n = \dim V$); moreover, by Theorem 2.3, $\operatorname{rank} g = p + q$.

3. 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 by knowing 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 $B = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ is a basis of V. Let $A^g_{B,B} = (a_{ij})_{1\le i,j\le n}$ be the matrix of g with respect to B. If $x_1,\dots,x_n\in\mathbb{R}$ 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 invariant with respect to the basis B and is called the *discriminant* of h.

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 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.

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 $B = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ is a basis of V giving a canonical form $\omega_1x_1^2+\omega_2x_2^2+\cdots+\omega_nx_n^2$ of h, then $B' = \{c_1\mathbf{b}_1,\dots,c_n\mathbf{b}_n\}$ 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 $1\le i\le n$.

Theorem 3.1 (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 $\{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ of V and $\omega_1,\dots,\omega_n\in\mathbb{R}$ such that for any $x_1,\dots,x_n\in\mathbb{R}$ we have*

$$h(x_1\mathbf{b}_1+\cdots+x_n\mathbf{b}_n) = \omega_1x_1^2+\omega_2x_2^2+\cdots+\omega_nx_n^2.$$

Remark. The quadratic polynomial $\omega_1x_1^2+\omega_2x_2^2+\cdots+\omega_nx_n^2$ is then a reduced form of h (the matrix of g with respect to $\{\mathbf{b}_1,\dots,\mathbf{b}_n\}$ is a diagonal matrix with entries $\omega_1,\dots,\omega_n$). If (p,q,r) is the signature of h, then among the coefficients $\omega_1,\dots,\omega_n$, p are positive, q are negative and r are equal to 0.

Theorem 3.2 (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

$$\mu_1x_1^2+\mu_2x_2^2+\cdots+\mu_nx_n^2,$$

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).

a) If $p = n$, h is called a *positive-definite* quadratic form.
b) If $q = 0$, the quadratic form h is called *positive semidefinite.*
c) If $q = n$, h is called a *negative-definite* quadratic form.

d) If $p = 0$, the quadratic form h is called *negative semidefinite*.
e) The quadratic form h is called *undefined* if $p > 0$ and $q > 0$.

Remarks.
1. Of course, for the quadratic form h, positive semidefiniteness implies positive-definiteness and negative semidefiniteness implies negative-definiteness

2. Let Δ_i, $1 \le i \le n$ be the principal minors of the associated matrix with respect to an arbitrary basis. According to Theorem 3.2, h is positive-definite if and only if

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

and h is negative-definite if and only if

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

Theorem 3.3 (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*

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

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

In fact, $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the associated matrix with respect to any basis of V.

Proof. Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V and A_B be the matrix associated to h with respect to the basis B. Since the matrix A_B is symmetric, the linear operator $T : V \to V$ associated to A_B (with respect to the same basis B) is autoadjoint and thus diagonalizable. Let $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ be the eigenvalues of T. From the method of diagonalization of T, we can construct an orthonormal basis $B' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V such that $\mathbf{v}_i$ is an eigenvector corresponding to λ_i,for $1 \le i \le n$ (indeed, if $\lambda_i \ne \lambda_j$, since T is autoadjoint, we have $\lambda_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle = \langle T(\mathbf{v}_i), \mathbf{v}_j \rangle = \langle \mathbf{v}_i, T(\mathbf{v}_j) \rangle = \langle \mathbf{v}_i, \lambda_j \mathbf{v}_j \rangle = \lambda_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle$, implying $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$). Of course, with respect to B', the matrix $A_{B'}$ associated to h has the form $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$, hence the conclusion. □

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 $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of basis of V, then for any $x_1, \dots, x_n \in \mathbb{R}$

$$(h + f)(x_1 \mathbf{b}_1 + \dots + x_n \mathbf{b}_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c, \tag{3}$$

where $A = (a_{ij})_{1 \le i,j \le n}$ is the matrix associated to h, $b_1, \dots, b_n \in \mathbb{R}$ and $c \in \mathbb{R}$. The right-hand side of this equality is called the *quadratic polynomial* associated to $h + f$ (which is a polynomial of degree 2, not necessarily homogeneous).

If $V = \mathbb{R}^n$ and B is its canonical basis, then (3) can be written as

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

where $\mathbf{b} = (b_1, b_2, \dots, b_n) \in \mathbb{R}^n$ 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) defines a non-homogeneous quadratic functional on V. Moreover, A can be taken not necessarily symmetric, since

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

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

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

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

Suppose now that S is the transition matrix from the canonical basis to an orthonormal basis giving the canonical form in Theorem 3.3. Therefore, S is an orthonormal matrix ($S^{-1} = S^{\mathrm{T}}$) and $SAS^{\mathrm{T}} = D := \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, where $\lambda_1, \dots, \lambda_n$ 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 $\gamma \in \mathbb{R}$.

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 $n \geq 4$.

In the case $n = 1$, the *normal form* of ρ is $x^2 + 1$ (then $\ker \rho = \emptyset$: two "imaginary" points), $x^2 - 1$ ($\ker \rho = \{-1, 1\}$: two distinct points) or $x^2 = 0$ ($\ker \rho = \{0\}$: two identical points).

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

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

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

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

Selective bibliography

[1] M. Ariciuc, S. Roatesi, *Lecţii de algebră liniară şi geometrie analitică*, Editura Matrix Rom, Bucureşti, 2008.

[2] K. C. Border, *More than you wanted to know about quadratic forms*, Caltech, 2016.

[3] K. Conrad, *Bilinear Forms,* Notes on Advanced Linear Algebra, 2015.

[4] C. Costinescu, *Algebră liniară şi aplicaţii în geometrie*, Editura Matrix Rom, Bucureşti, 2005.

[5] D. Drăghici, *Algebră*, Editura Didactică şi Pedagogică, Bucureşti, 1972.

[6] G. Galbură, F. Radó, *Geometrie*, Ed. Didactică şi Pedag., Bucureşti, 1979.

[7] M. Neagu, *Geometria curbelor şi suprafeţelor. Teorie şi aplicaţii*, Editura Matrix Rom, Bucureşti, 2013.

[8] P. Ott, *Bilinear and Quadratic Forms*, Prof. Robert Beezer's Notes on Advanced Linear Algebra, 2014.

[9] I. Radomir, *Elemente de algebră vectorială, geometrie şi calcul diferenţial*, Editura Albastră, Cluj-Napoca, 2000.