LECTURE V: BILINEAR FORMS AND ORTHOGONALITY
MAT 204 - FALL 2006
PRINCETON UNIVERSITY

ALFONSO SORRENTINO

1. Bilinear forms

Definition. Let V be a vector space. A *bilinear form* on V is an application

$$b : V \times V \longrightarrow \mathbb{R}$$

that satisfies the two following conditions:

$$\begin{array}{lll} \text{i)} & b(c_1u_1 + c_2u_2, v) = c_1b(u_1, v) + c_2b(u_2, v), & \forall\, c_1, c_2 \in \mathbb{R},\ \forall\, u_1, u_2, v \in V\,; \\ \text{ii)} & b(u, d_1v_1 + d_2v_2) = d_1b(u, v_1) + d_2b(u, v_2), & \forall\, d_1, d_2 \in \mathbb{R},\ \forall\, v_1, v_2, u \in V\,. \end{array}$$

A bilinear form is said to be *symmetric* [resp. *skew-symmetric*] if

$$b(u, v) = b(v, u) \qquad [\text{resp. } b(u, v) = -b(v, u)] \quad \forall\, u, v \in V\,.$$

We would like, as already done for linear transformations, to associate to b a matrix, that encodes all the information we need to know about b; it will obviously depend on our choice of a basis on V.

Definition. Let V be a vector space of dimension n and let $\mathbb{E} = (e_1, \ldots, e_n)$ be its basis. Given a bilinear form $b : V \times V \longrightarrow \mathbb{R}$, we will call *matrix of* b *w.r.t.* $\mathbb{E}$, the matrix

$$A = \begin{pmatrix} b(e_1, e_1) & \ldots & \ldots & b(e_1, e_n) \\ b(e_2, e_1) & \ldots & \ldots & b(e_2, e_n) \\ \vdots & \vdots & \vdots & \vdots \\ b(e_n, e_1) & \ldots & \ldots & b(e_n, e_n) \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

Let us see how we can use this matrix to recover the behavior of b on $V \times V$. Consider two generic vectors $u = \mathbb{E}x$ and $v = \mathbb{E}y$ in V. We want to determine an expression for $b(u, v)$ in terms of x and y (*i.e.*, with respect to $\mathbb{E}$). Set $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$; we have:

$$\begin{aligned} b(u, v) &= b\left(\sum_{i=1}^n x_ie_i, v\right) = \sum_{i=1}^n x_ib(e_i, v) = (x_1, \ldots, x_n) \begin{pmatrix} b(e_1, v) \\ \vdots \\ b(e_n, v) \end{pmatrix} = \\ &= x^T \begin{pmatrix} b(e_1, v) \\ \vdots \\ b(e_n, v) \end{pmatrix}. \end{aligned}$$

For each $i = 1, \ldots, n$, we can compute $b(e_i, v)$:

$$\begin{aligned} b(e_i, v) &= b\left(e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{j=1}^n b(e_i, e_j) y_j = \\ &= (b(e_i, e_1), \ldots, b(e_i, e_n)) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A^{(i)} y \,. \end{aligned}$$

Therefore,

$$b(\mathbb{E}x, \mathbb{E}y) = x^T \begin{pmatrix} A^{(1)} y \\ \vdots \\ A^{(n)} y \end{pmatrix} = x^T \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix} y = x^T A y \,.$$

The expression

$$b(\mathbb{E}x, \mathbb{E}y) = x^T A y$$

is called *expression of the bilinear form b w.r.t.* $\mathbb{E}$.

Remark. From the expression above, it follows that, once we fix a basis $\mathbb{E}$ of V, b is uniquely determined by the matrix A. Hence, there exists a bijection (w.r.t. a fixed $\mathbb{E}$) between the set of bilinear forms on V and $\mathcal{M}_n(\mathbb{R})$.

Let us see what other information about b, can be recovered from its matrix.

Proposition 1. *Let V be a n-dimensional vector space and $b : V \times V \longrightarrow \mathbb{R}$ a bilinear form on V. Fix a basis $\mathbb{E}$ on V and let $b(\mathbb{E}x, \mathbb{E}y) = x^T A y$. We have that:*

$$b \text{ is symmetric} \quad \Longleftrightarrow \quad A \text{ is a symmetric matrix } [\text{i.e., } A^T = A\,].$$

Proof. $(\Longrightarrow)$ Since $b(e_i, e_j) = b(e_j, e_i)$, then $a_{ij} = a_{ji}$ and consequently $A^T = A$.

$(\Longleftarrow)$ Observe that, since $x^T A y$ is a real number, it coincides with its transpose; *i.e.*,

$$x^T A y = (x^T A y)^T = y^T A^T x = y^T A x \,,$$

where we used, to get the last equality, that A is symmetric. Therefore, for any $u, v \in V$:

$$b(u, v) = b(\mathbb{E}x, \mathbb{E}y) = x^T A y = y^T A x = b(\mathbb{E}y, \mathbb{E}x) = b(v, u) \,.$$

□

One can verify, in a completely analogous way, that:

Proposition 2. *Let V be a n-dimensional vector space and $b : V \times V \longrightarrow \mathbb{R}$ a bilinear form on V. Fix a basis $\mathbb{E}$ on V and let $b(\mathbb{E}x, \mathbb{E}y) = x^T A y$. We have that:*

$$b \text{ is skew-symmetric} \quad \Longleftrightarrow \quad A \text{ is a skew-symmetric matrix } [\text{i.e., } A^T = -A\,].$$

One natural question that one could ask is: what happens when I change the basis in V?

Let $\mathbb{E}$ and $\mathbb{E}'$ be two bases for V and assume that $\mathbb{E}' = \mathbb{E}C$ (with $C \in \mathrm{GL}_n(\mathbb{R})$) is the formula for the change of basis. Consider two generic vectors $u = \mathbb{E}x = \mathbb{E}'x'$ and $v = \mathbb{E}y = \mathbb{E}'y'$. We have:

$$x^T A y = b(\mathbb{E}x, \mathbb{E}y) = b(u, v) = b(\mathbb{E}'x', \mathbb{E}'y') = (x')^T A' y' \,.$$

Since $\mathbb{E}' = \mathbb{E}C$, then $x = Cx'$ and $y = Cy'$; therefore:

$$x^T A y = (Cx')^T A (Cy') = ((x')^T C^T) A (Cy') = (x')^T (C^T A C) y = (x')^T A' y' \,.$$

It follows that

$$A' = C^T A C \,.$$

Remark. Let $A, A' \in \mathcal{M}_n(\mathbb{R})$. They are called *congruent* if there exists $C \in \mathrm{GL}_n(\mathbb{R})$ such that $A' = C^T A C$. One can easily verify that *congruence* is an equivalence relation on the set of $\mathcal{M}_n(\mathbb{R})$.
In particular (recall what we have proved for *similar matrices* in Lecture IV):

Proposition 3. *Two matrices $A, A' \in \mathcal{M}_n(\mathbb{R})$ are congruent $\iff$ they represent the same bilinear form $b : V \times V \longrightarrow \mathbb{R}$, with respect to different bases.*

Definition. Let V a n-dimensional vector space and $b : V \times V \longrightarrow \mathbb{R}$ a bilinear form on V. We call *rank* of b, the rank of its matrix A, with respect to any basis $\mathbb{E}$. Observe that this definition makes sense (namely, it does not depend on the choice of the basis $\mathbb{E}$), since two congruent matrices have the same rank: $\operatorname{rank}(A') = \operatorname{rank}(C^T A C)$ (we have already noticed that multiplication by invertible matrices, does not change the rank). Hence, $\operatorname{rank}(b)$ is an invariant of b.
A bilinear form is said to be *degenerate*, if $\operatorname{rank}(b) < n$; otherwise it is said to be *non-degenerate*.

Proposition 4. *Let V a n-dimensional vector space and $b : V \times V \longrightarrow \mathbb{R}$ a bilinear form on V. The following conditions are equivalent:*

i) *b is degenerate;*
ii) *there exists a non-zero vector $u_0 \in V$, such that $b(u_0, v) = 0$ for any $v \in V$;*
iii) *there exists a non-zero vector $v_0 \in V$, such that $b(u, v_0) = 0$ for any $u \in V$.*

Proof. We show only the equivalence between i) and ii) [the proof of the equivalence between i) and iii) is pretty similar].
Let $b(\mathbb{E}x, \mathbb{E}y) = x^T A y$. We have:

$$\begin{aligned}
&b \text{ is degenerate } \iff \operatorname{rank} A < n \iff \\
&\iff \text{ the rows } A^{(1)}, \ldots, A^{(n)} \text{ are linearly dependent } \iff \\
&\iff \text{ there exist } c_1, \ldots, c_n \in \mathbb{R} \text{ (not all zero), such that } \sum_{i=1}^n c_i A^{(i)} = 0 \iff \\
&\iff \text{ there exists } c \in \mathcal{M}_{n,1}(\mathbb{R}),\ c \neq 0, \text{ such that } c^T A = 0\,.
\end{aligned}$$

Let us verify that:

$$c^T A = 0 \iff c^T A y = 0,\ \forall\, y \in \mathcal{M}_{n,1}(\mathbb{R}).$$

The implication ($\Longrightarrow$) is evident. Let us verify ($\Longleftarrow$). In fact:

$$\begin{aligned}
0 &= c^T A \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = c^T A_{(1)} \\
&\vdots \\
0 &= c^T A \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = c^T A_{(n)}\,.
\end{aligned}$$

Consequently,

$$c^T A = c^T \left(A_{(1)} \ \ldots \ A_{(n)} \right) = \left(c^T A_{(1)} \ \ldots \ c^T A_{(n)} \right) = (0, \ldots, 0) = 0\,.$$

This allows us to conclude the proof. In fact, using this fact:

$$\begin{aligned}
& b \text{ is degenerate } \iff \operatorname{rank} A < n \iff \\
& \iff \text{ there exists } c \in \mathcal{M}_{n,1}(\mathbb{R}),\ c \neq 0, \text{ such that } c^T A = 0 \iff \\
& \iff \text{ there exists } c \in \mathcal{M}_{n,1}(\mathbb{R}),\ c \neq 0, \text{ such that } c^T A y = 0,\ \forall\, y \in \mathcal{M}_{n,1}(\mathbb{R}) \iff \\
& \iff \text{ there exists } u_0 = \mathbb{E}c \in V,\ u_0 \neq 0, \text{ such that } b(u_0, v) = 0,\ \forall\, v = \mathbb{E}y \in V\,.
\end{aligned}$$

□

2. Orthogonality

Note: From here on, we will just consider symmetric bilinear forms.

In the physical space, it is well known the concept of *orhogonal* vectors; namely, two vectors are orthogonal if, considered at a same starting point, they form a right angle. We would like to generalize this concept to a general vector space, but it is not evident at all what a "natural" definition should be. Think for instance at the space of polynomials with real coefficients. What might be the meaning of "orthogonal" polynomials?
The main goal of this section is to introduce the concept of *orthogonality* in a more general framework.

Let us first rephrase the definition of orthogonal physical vectors in $\mathbb{R}^2$ in a different way. Define the application

$$\begin{aligned} b : \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (v_1, v_2) &\longmapsto \|v_1\|\|v_2\| \cos\theta \end{aligned}$$

where $\|v_1\|$ and $\|v_2\|$ is the euclidean length of the vectors and $\theta \in [0, \pi]$ is the smallest angle between these vectors (when we apply them at the same point). One can verify that this is a symmetric bilinear form and that, if $v_1, v_2 \neq 0$, then

$$b(v_1, v_2) = 0 \quad \iff \quad \theta = \frac{\pi}{2}\,.$$

Therefore,

$$v_1 \text{ and } v_2 \text{ are orthogonal } \iff b(v_1, v_2) = 0\,.$$

This observation justifies the following definition.

Definition. Let $b : V \times V \longrightarrow \mathbb{R}$ be a symmetric bilinear form and $u, v \in V$. If $b(u, v) = 0$, then the two vectors are called *b-orthogonal* (or simply *orthogonal*) and it will be denoted by $u \perp v$.
If S is a non-empty set of V (not necessarily a subspace!), the set $S^\perp := \{u \in V :\ b(u, v) = 0, \text{ for all } v \in V\}$ (*i.e.*, the set of vectors that are b-orthogonal to all vectors of S) is called *b-orthogonal subspace* of S. It is easy to verify that $S^\perp$ is a vector subspace of V [Exercise].
If $S = \{w\}$, we will abbreviate $w^\perp$, instead of $\{w\}^\perp$. Moreover, if $W = \langle w_1, \ldots, w_s \rangle$, one can easily check that

$$W^\perp = w_1^\perp \cap \ldots \cap w_s^\perp\,.$$

Finally, a vector $v \in V$ is called *b-isotropic* (or simply *isotropic*), if $b(v, v) = 0$. The set of all isotropic vectors of V is called *b-isotropic cone* of V and is denoted by $I_b(V)$.

Remark. i) In general, the b-isotropic cone $I_b(V)$ is not a vector subspace of V. For example, suppose that $v_1, v_2 \in I_b(V)$ and that $b(v_1, v_2) \neq 0$; one has that $v_1 + v_2 \notin I_b(V)$. In fact:

$$b(v_1 + v_2, v_1 + v_2) = b(v_1, v_1) + b(v_2, v_2) + 2b(v_1, v_2) = 2b(v_1, v_2) \neq 0\,.$$

Moreover, observe that if $I_b(V) \neq \{0\}$, then it is the union (in the set theoretical sense) of 1-dimensional vector subspaces of V. In fact, if $v \in I_b(V)$ and $c \in \mathbb{R}$, it is easy to verify that $cv \in I_b(V)$.

ii) If $I_b(V) = \{0\}$, then b is non degenerate. In fact, using proposition 4, if $b(u_0, v) = 0$ for all $v \in V$, then in particular $b(u_0, u_0) = 0$ and this implies that $u_0 \in I_b(V) = \{0\}$.
The converse is false! Namely, there exist non-degenerate symmetric bilinear forms, whose isotropic cone is not trivial. For example, consider the matrix:

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

It defines (w.r.t. the canonical basis of $\mathbb{R}^2$) a non-degenerate symmetric bilinear form on $\mathbb{R}^2$ (in fact $\operatorname{rank}(A) = 2$), but $I_b(\mathbb{R}^2) \neq \{0\}$; in fact, $e_1 \in I_b(\mathbb{R}^2)$.

Proposition 5. *Let $b : V \times V \longrightarrow \mathbb{R}$ be a symmetric bilinear form and $u \in V$ a non-isotropic vector. Then,*

$$V = \langle u \rangle \oplus u^{\perp},$$

i.e., each vector $v \in V$ can be decomposed uniquely into the sum of a vector v_1 "parallel" (proportional) to u and a vector v_2 b-orthogonal to u. These vectors are given by:

$$v_1 = \frac{b(u,v)}{b(u,u)} u \quad \text{and} \quad v_2 = v - \frac{b(u,v)}{b(u,u)} u.$$

Proof. By hypothesis, we know that $b(u,u) \neq 0$. We need to verify that

$$\langle u \rangle \cap u^{\perp} = \langle 0 \rangle \quad \text{and} \quad \langle u \rangle + u^{\perp} = V.$$

Let us start by proving the first equality.
If $w \in \langle u \rangle \cap u^{\perp}$, then $w = cu$ (there exists $c \in \mathbb{R}$) and $b(w,u) = 0$. Therefore $0 = b(cu,u) = cb(u,u)$ and this implies - since $b(u,u) \neq 0$ - that $c = 0$. Hence, $w = 0$.
Let us verify now that $\langle u \rangle + u^{\perp} = V$. In fact, for any $v \in V$, if we define v_1 and v_2 as above, we have: $v_1 + v_2 = v$, $v_1 \in \langle u \rangle$ and $v_2 \in u^{\perp}$. In fact:

$$b(u, v_2) = b\left(u, v - \frac{b(u,v)}{b(u,u)} u\right) = b(u,v) - b(u,u)\frac{b(u,v)}{b(u,u)} = 0.$$

□

Definition. Let $b : V \times V \longrightarrow \mathbb{R}$ be a symmetric bilinear form and $u \in V$ a non-isotropic vector. For any $v \in V$, the scalar $\frac{b(u,v)}{b(u,u)}$ is called *Fourier coefficient of v with respect to u* and will be denoted by $c_u(v)$. The vector $v_1 = c_u(v)u$ is called *b-orthogonal projection of v in the u direction.*
Observe that the application

$$\begin{aligned} c_u : V &\longrightarrow \mathbb{R} \\ v &\longmapsto c_u(v) \end{aligned}$$

is a linear application [Exercise].

Proposition 6. *Let V a n-dimensional vector space with basis $\mathbb{E}$. Let b be a symmetric bilinear form on V, such that*

$$b(\mathbb{E}x, \mathbb{E}y) = x^T A y.$$

Let W be a s-dimensional vector subspace of V, with basis $\{z_1, \ldots, z_s\}$ and let $(z_1 \ \ldots \ z_s) = \mathbb{E}D$ (with $D \in \mathcal{M}_{n,s}(\mathbb{R})$ and $\operatorname{rank}(D) = s$). Then, the cartesian *equations of* $W^\perp$ *(i.e., its* b*-orthogonal subspace) with respect to* $\mathbb{E}$*, are given by the HLS*$(s,n,\mathbb{R})$*:*

$$D^T A X = 0\,.$$

Moreover, if b *is non-degenerate, then* $\dim W^\perp = n - \dim W$*.*

Proof. Let $v = \mathbb{E}x \in V$. We have:

$$\begin{aligned} v = \mathbb{E}x \in W^\perp \quad &\Longleftrightarrow \quad b(z_1, v) = \ldots = b(z_s, v) = 0 \;\Longleftrightarrow \\ &\Longleftrightarrow \quad D_{(1)}^T A x = \ldots = D_{(s)}^T A x = 0 \;\Longleftrightarrow \\ &\Longleftrightarrow \quad D^T A x = 0\,. \end{aligned}$$

Hence, $W^\perp$ has cartesian equations given by the HLS$(s,n,\mathbb{R})$ $D^T A X = 0\,.$
If b is non-degenerate, then $\operatorname{rank} A = n$ and therefore $A \in \mathrm{GL}_n(\mathbb{R})$. In this case, $\operatorname{rank}(D^T A) = \operatorname{rank}(D^T) = \operatorname{rank}(D) = s$. Therefore,

$$\dim W^\perp = n - \operatorname{rank}(D^T A) = n - s = n - \dim W\,.$$

$\square$

Example. Let b be a symmetric bilinear form on $\mathbb{R}^4$ defined, with respect to the canonical basis $\mathbb{E} = (e_1\ e_2\ e_3\ e_4)$ of $\mathbb{R}^4$, by the matrix

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

i) Verify that b is non-degenerate.
ii) Set $W = \langle e_3, e_1 + e_2 \rangle$. Determine the cartesian equations of $W^\perp$.
iii) Verify that $W \cap W^\perp = \langle 0 \rangle$.
iv) Determine the b-isotropic vectors in $W^\perp$.

Solution:

i) b is non-degenerate, since $\det A \neq 0$.
ii) We have:

$$(e_3 \;\; e_1 + e_2) = \mathbb{E}D, \quad \text{with} \quad D = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

$W^\perp$ has cartesian equations $D^T A X = 0$, where $D = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 1 & 1 & 2 & -1 \end{pmatrix}$, *i.e.*,

$$\begin{cases} 2x_2 = 0 \\ x_1 + x_2 + 2x_3 - x_4 = 0 \end{cases} \quad \Longleftrightarrow \quad \begin{cases} x_2 = 0 \\ x_1 + 2x_3 - x_4 = 0\,. \end{cases}$$

Therefore, $\dim W^\perp = 4 - 2 = 2$ and $W^\perp = \langle (-2,0,1,0),\, (1,0,0,1) \rangle$.

iii) The cartesian equations of W are obtained directly imposing the condition $\operatorname{rank}(Dx) = 2$:

$$\begin{cases} x_1 - x_2 = 0 \\ x_4 = 0\,. \end{cases}$$

Hence, $W \cap W^\perp$ has cartesian equations given by:

$$\begin{cases} x_1 - x_2 = 0 \\ x_4 = 0 \\ x_2 = 0 \\ x_1 + 2x_3 - x_4 = 0 \end{cases} \quad \Longleftrightarrow \quad \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \\ x_4 = 0 \end{cases}$$

this means $W \cap W^{\perp} = \langle 0 \rangle$.

iv) We have:

$$\mathbb{E}x \in I_b(\mathbb{R}^4) \iff x^T A x = 0 \iff x_1^2 + x_2^2 + 4x_2x_3 - 2x_2x_4 - x_4^2 = 0\,.$$

Therefore, $\mathbb{E}x \in I_b(\mathbb{R}^4) \cap W^{\perp}$ if and only if

$$\begin{cases} x_2 = 0 \\ x_1 + 2x_3 - x_4 = 0 \\ x_1^2 - x_4^2 = 0 \end{cases} \iff \begin{cases} x_2 = 0 \\ x_1 + 2x_3 - x_4 = 0 \\ x_1 - x_4 = 0 \end{cases} \text{or} \begin{cases} x_2 = 0 \\ x_1 + 2x_3 - x_4 = 0 \\ x_1 + x_4 = 0\,. \end{cases}$$

In conclusion:

$$I_b(\mathbb{R}^4) \cap W^{\perp} = \{(t,0,0,t) \in \mathbb{R}^4,\ \text{for all } t \in \mathbb{R}\} \cup \{(-t,0,t,t) \in \mathbb{R}^4,\ \text{for all } t \in \mathbb{R}\}$$

or equivalently

$$I_b(\mathbb{R}^4) \cap W^{\perp} = \langle (1,0,0,1) \rangle \cup \langle (-1,0,1,1) \rangle\,.$$

3. Quadratic forms

Definition. Let $b : V \times V \longrightarrow \mathbb{R}$ be a symmetric bilinear form. The application

$$\begin{array}{rcl} Q : V & \longrightarrow & \mathbb{R} \\ v & \longmapsto & Q(v) := b(v,v) \end{array}$$

is called *quadratic form associated to b*.

Remark. For any $v_1, v_2 \in V$, we have:

$$\begin{array}{rcl} Q(v_1+v_2) & = & b(v_1+v_2, v_1+v_2) = \\ & = & b(v_1,v_1) + b(v_2,v_2) + 2b(v_1,v_2) = \\ & = & Q(v_1) + Q(v_2) + 2b(v_1,v_2)\,. \end{array}$$

Therefore, we can reconstruct b from its quadratic form:

$$b(v_1,v_2) = \frac{1}{2}\left[Q(v_1+v_2) - Q(v_1) - Q(v_2)\right]$$

(b is also called *polar form of Q*). In particular, it follows that there exists a $1-1$ correspondence between the set of symmetric bilinear forms on V and the set of quadratic forms on V.

Let V be a n-dimensional vector space and $\mathbb{E}$ a basis for V. Let b be a symmetric bilinear form on V, with $b(\mathbb{E}x, \mathbb{E}y) = x^T A y$. For any vector $v = \mathbb{E}x \in V$, we have:

$$Q(v) = Q(\mathbb{E}x) = b(\mathbb{E}x, \mathbb{E}x) = x^T A x = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j\,;$$

this is called *expression of the quadratic form Q with respect to $\mathbb{E}$*.

Remark. The expression

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$$

is a homogeneous polynomial of degree 2 in $x_1, \ldots, x_n$. Conversely, every homogeneous polynomial of degree 2 in $x_1, \ldots, x_n$ determines a quadratic form (with respect to a fixed basis $\mathbb{E}$). In fact, if

$$P(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} b_{ij}x_ix_j$$

is an arbitrary homogeneous polynomial of degree 2 in $\mathbb{R}[x_1, \ldots, x_n]$, it suffices to define a matrix $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$, where

$$\begin{cases} a_{ii} = b_{ii} & \text{for } i = 1, \ldots, n \\ a_{ij} = a_{ji} = \frac{b_{ij}}{2} & \text{for } 1 \leq i \leq j \leq n\,. \end{cases}$$

Consider now the quadratic form $Q : V \longrightarrow \mathbb{R}$, such that $Q(\mathbb{E}x) = x^T A x$. It is easy to verify that $Q(\mathbb{E}x) = P(x_1, \ldots, x_n)$.

Hence, there is also a $1-1$ correspondence between the set of quadratic forms on V and the set of homogeneous polynomials of degree 2 in $x_1, \ldots, x_n$ (and also with the set of symmetric matrices of order n).

Example. Consider the homogeneous polynomial of degree 2

$$P(x_1, x_2, x_3, x_4) = x_2^2 - 2x_1x_3 + x_2x_4 + 2x_4^2 \in \mathbb{R}[x_1, x_2, x_3, x_4]\,.$$

It defines the symmetric matrix:

$$A = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \\ -1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 2 \end{pmatrix} \in \mathcal{M}_4(\mathbb{R}).$$

The symmetric bilinear form $b : \mathbb{R}^4 \times \mathbb{R}^4 \longrightarrow \mathbb{R}$ defined by A (with respect to the canonical basis) is:

$$b(x, y) = x_2y_2 - x_1y_3 - x_3y_1 + \frac{1}{2}x_2y_4 + \frac{1}{2}x_4y_2 + 2x_4y_4$$

and the associated quadratic form:

$$Q(x) = x_2^2 - 2x_1x_3 + x_2x_4 + 2x_4^2\,.$$

4. Inner products

Definition. Let V be a real vector space and $b : V \times V \longrightarrow \mathbb{R}$ a bilinear form. b is said to be *positive definite*, if it satisfies the following two conditions:

$$\begin{array}{ll} \text{i)} & b(v,v) \geq 0\,, \text{ for all } v \in V \\ \text{ii)} & b(v,v) = 0 \iff v = 0\,. \end{array}$$

If b satisfies only *i)* is called *positive semidefinite*.
Analogously, b is said to be *negative definite*, if it satisfies the following two conditions:

$$\begin{array}{ll} \text{i}') & b(v,v) \leq 0\,, \text{ for all } v \in V \\ \text{ii}') & b(v,v) = 0 \iff v = 0\,. \end{array}$$

If b satisfies only *i′)* is called *negative semidefinite*.
A symmetric positive definite bilinear form is called inner product.

Definition. Let b be an inner product on V. For each $v \in V$, we call the *norm* (or *length*) of v (with respect to b), the non-negative real number given by:

$$\|v\| := \sqrt{b(v,v)}\,.$$

Observe that $\|v\|$ is well defined, since $b(v,v) \geq 0$ (b is positive definite). Moreover, the following properties are obvious:

$$\begin{array}{ll} \text{i)} & \|v\| = 0 \iff v = 0\,; \\ \text{ii)} & \|cv\| = |c|\,\|v\|\,, \text{ for any } c \in \mathbb{R}. \end{array}$$

We will say that a vector v is a *versor* (w.r.t. b) if $\|v\| = 1$; in particular, for any non-zero vector $v \in V$, the vector $\frac{1}{\|v\|}\, v$ is a versor.

Moreover one can show that:

Proposition 7. *Let V be a real vector space and b an inner product on V. Then, for any $u, v \in V$:*

i) $|b(u,v)| \leq \|u\|\,\|v\|$ (*Cauchy-Schwartz inequality*);
ii) $|\|u\| - \|v\|| \leq \|u+v\| \leq \|u\| + \|v\|$ (*Triangular inequality*).

Proof. i) For each $t \in \mathbb{R}$ we have:

$$b(u+tv, u+tv) = \|u\|^2 + 2b(u,v)t + \|v\|^2 t^2 \geq 0\,.$$

From this inequality, it follows that the polynomial $Q = \|v\|^2 t^2 + 2b(u,v)t + \|u\|^2$ has at most one real root and therefore its discriminant $\Delta \leq 0$; hence:

$$\Delta = b(u,v)^2 - \|u\|^2\|v\|^2 \leq 0\,.$$

We can then conclude that $|b(u,v)| = \sqrt{b(u,v)^2} \leq \|u\|\,\|v\|$.

ii) Let us show the second inequality. Using i), we have:

$$\begin{aligned} \|u+v\|^2 &= b(u+v,u+v) = \|u\|^2 + \|v^2\| + 2b(u,v) \leq \\ &\leq \|u\|^2 + \|v^2\| + \|u\|\,\|v\| \leq \left(\|u\| + \|v\|\right)^2\,. \end{aligned}$$

Taking the square root on both sides, we obtain the desired inequality.
Let us show the first inequality. We have:

$$\begin{cases} \|u\| = \|u+v-v\| \leq \|u+v\| + \|-v\| = \|u+v\| + \|v\| & \Longleftrightarrow \quad \|u\| - \|v\| \leq \|u+v\| \\ \|v\| = \|v+u-u\| \leq \|u+v\| + \|-u\| = \|u+v\| + \|u\| & \Longleftrightarrow \quad \|v\| - \|u\| \leq \|u+v\| \end{cases}$$

and this implies $|\|u\| - \|v\|| \leq \|u+v\|$.

□

Definition. A couple (V,b), where V is real vector space and b an inner product on V, is called *euclidean space*.

Example. **[Canonical inner product on $\mathbb{R}^n$]** Define

$$\begin{aligned} \cdot : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x,y) &\Longrightarrow x \cdot y := x_1y_1 + \ldots + x_ny_n\,, \end{aligned}$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. This is a bilinear form [Exercise] and its matrix (with respect to the canonical basis of $\mathbb{R}^n$) is I_n (the identity):

$$x \cdot y = x^T I_n y\,.$$

Moreover,

$$\begin{array}{ll} \text{i)} & x \cdot x = x_1^2 + \ldots + x_n^2 \geq 0\,; \\ \text{ii)} & x \cdot x = x_1^2 + \ldots + x_n^2 = 0 \quad \Longleftrightarrow \quad x = 0\,. \end{array}$$

Therefore, $\cdot$ is an inner product and is called *canonical* (or *standard*) inner product.

Remark. From ii) in the definition of positive definiteness, it follows that if b is an inner product, then $I_b(V) = \{0\}$ (*i.e.*, no nonzero vector is b-isotropic). The converse is false; for instance, consider the bilinear form on $\mathbb{R}^2$ given (with respect to the canonical basis of $\mathbb{R}^2$) by:

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

Obviously, b is not an inner product: $b(e_1, e_1) = -2 < 0$; but one can easily show that $I_b(\mathbb{R}^2) = \{0\}$. In fact, $I_b(\mathbb{R}^2)$ has equation $-2x^2 + 2xy - 2y^2 = 0$ and this equation has only one solution in $\mathbb{R}^2$, given by $(0,0)$ [to see this, one can denote $t = \frac{x}{y}$ and see that the equation $-2t^2 + 2t - 2 = 0$ has not solutions in $\mathbb{R}$].

It is also evident, that if b is an inner product of V, with $\dim V = n$, then it has rank n, *i.e.*, it is non-degenerate. In particular, it is always possible to find a new basis on V, such that the matrix representing b w.r.t. this basis is I_n. Such a basis is called *b-orthonormal basis.*

Definition. Let V be a n-dimensional vector space and $\mathbb{E} = (e_1 \ \dots \ e_n)$ a basis. Let b be an inner product of V, with matrix A (w.r.t. $\mathbb{E}$), *i.e.*, $b(\mathbb{E}x, \mathbb{E}y) = x^T Ay$. If $A = I_n$, $\mathbb{E}$ is called *b-orthonormal basis* and one has:

$$b(\mathbb{E}x, \mathbb{E}y) = x^T I_n y = \sum_{i=1}^n x_i y_i \qquad \forall\, \mathbb{E}x,\, \mathbb{E}y \in V\,.$$

Hence, with respect to this basis, b acts on the coordinates of the vectors of V as the canonical inner product on $\mathbb{R}^n$.

Remark. We shall see in the next section, how we can actually compute such a basis. In the meanwhile, let us just verify the following properties.

i) Let b be an inner product on V. Let us verify that t non-zero vectors $u_1, \dots, u_t \in V$, pairwise b-orthogonal, are linearly independent. Suppose that $\sum_{i=1}^t c_i u_i = 0$; for $j = 1, \dots, t$:

$$0 = b(u_j, 0) = b\left(u_j, \sum_{i=1}^t c_i u_i\right) = \sum_{i=1}^t c_i b(u_j, u_i) = c_j b(u_j, u_j)$$

and since $b(u_j, u_j) > 0$, it follows that $c_j = 0$ (for all $j = 1, \dots, t$).

ii) If $\mathbb{E}$ is a b-orthonormal basis of V, then:

$$b(v, e_i) = x_i \quad \forall v = \mathbb{E}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in V$$

[in fact, $b(v, e_i) = b\left(\sum_{j=1}^n x_j e_j, e_i\right) = \sum_{j=1}^n x_j b(e_j, e_i) = x_i$]. It follows that for any vector $v \in V$, the following representation holds:

$$v = \sum_{i=1}^n b(v, e_i) v\,.$$

Proposition 8. *Let b be an inner product on V, n-dimensional vector space. Let $\mathbb{E}$ and $\mathbb{F}$ be two bases with $\mathbb{F} = \mathbb{E}C$, where $C \in \mathrm{GL}_n(\mathbb{R})$. One has:*

i) *If $\mathbb{E}$ and $\mathbb{F}$ are both b-orthonormal, then $C \in O_n(\mathbb{R})$ [i.e., it is orthogonal: $C^T C = I_n$].*

ii) *If $\mathbb{E}$ is b-orthonormal and $C \in O_n(\mathbb{R})$, then $\mathbb{F}$ is also b-orthonormal.*

Therefore, there is a $1-1$ correspondence between b-orthonormal bases and $O_n(\mathbb{R})$.

Proof. Let A be the matrix that represents b with respect to $\mathbb{E}$ and B the matrix of b with respect to $\mathbb{F}$. We have already observed that $B = C^T AC$.

i) If $\mathbb{E}$ and $\mathbb{F}$ are both b-orthonormal, then $B = A = I_n$ and consequently $C^T I_n C = I_n$. Therefore $C \in O_n(\mathbb{R})$.

ii) If $\mathbb{E}$ is b-orthonormal and $C \in O_n(\mathbb{R})$, then $A = I_n$ and consequently the matrix of b w.r.t. $\mathbb{F}$ is $B = C^T I_n C = C^T C = I_n$. Hence, $\mathbb{F}$ is also b-orthonormal.

□

5. Gram-Schmidt orthogonalization process

Let V be a real vector space and fix an inner product b. What we are going to show in this section, is that given any (finite or countable) ordered set of vectors of V, it is always possible to b-orthonormalize them, *i.e.*, substitute them with a new set of vectors, pairwise b-orthogonal, that is *equivalent* (in a certain sense to be specified) to the original one. The proof of this result will provide a simple algorithm, known as *Gram-Schmidt process*. Let us start with a definition.

Definition. Let V be a real vector space with an inner product b and two sets $\{u_1, \ldots, u_t, \ldots\}$, $\{w_1, \ldots, w_t, \ldots\}$ of vectors of V, that are in $1-1$ correspondence. We shall say that $\{w_1, \ldots, w_t, \ldots\}$ is a b-orthogonalization of $\{u_1, \ldots, u_t, \ldots\}$, if for any $k \geq 1$:

(a_k) $\langle u_1, \ldots, u_k \rangle = \langle w_1, \ldots, w_k \rangle$;
(b_k) the vectors $w_1, \ldots, w_k$ are pairwise b-orthogonal.

Theorem 1 (Gram-Schmidt). *Let V be a real vector space and b an inner product on V. Given any set of vectors of V $\{u_1, \ldots, u_t, \ldots\}$, there exists a b-orthogonalization $\{w_1, \ldots, w_t, \ldots\}$. This is "essentially" unique, in the sense that if $\{w'_1, \ldots, w'_t, \ldots\}$ is another b-orthogonalization of $\{u_1, \ldots, u_t, \ldots\}$, then $\langle w_k \rangle = \langle w'_k \rangle$ for any $k \geq 1$ (i.e., the two vectors are proportional).*

Proof. Let us proceed by induction on t. If $t = 1$, then it will be sufficient to set $w_1 = u_1$ and it is trivial to verify that the properties (a_1) and (b_1) of the above definition hold.
Let $t \geq 2$ and suppose that there exist t vectors $w_1, \ldots, w_t$ that verify the properties (a_k) and (b_k) for $1 \leq k \leq t$; we want to construct a vector w_{t+1}, such that the vectors $w_1, \ldots, w_{t+1}$ verify (a_{t+1}) and (b_{t+1}).
Let us set:

$$w_{t+1} = u_{t+1} - \sum_{i=1}^{t} c_i w_i \tag{1}$$

where

$$c_i = \begin{cases} 0 & \text{if } w_i = 0 \\ c_{w_i}(u_{t+1}) & \text{if } w_i \neq 0 \end{cases}$$

[$c_{w_i}(u_{t+1})$ is Fourier's coefficient of u_{t+1} with respect to w_i]. Obviously:

$$w_{t+1} \in \langle w_1, \ldots, w_t, u_{t+1} \rangle \quad \text{and} \quad u_{t+1} \in \langle w_1, \ldots, w_t, w_{t+1} \rangle .$$

From (a_t) it follows that

$$\langle u_1, \ldots, u_{t+1} \rangle = \langle w_1, \ldots, w_t, w_{t+1} \rangle ;$$

therefore (a_{t+1}) holds. Let us verify (b_{t+1}). It will be enough to verify that $w_{t+1} \perp w_j$, for any $j = 1, \ldots, t$. If $w_j = 0$, the claim is trivial. For, let us assume that $w_j \neq 0$. We have:

$$b(w_{t+1}, w_j) = b(u_{t+1}, w_j) - \sum_{i=1}^{t} c_i b(w_i, w_j) = b(u_{t+1}, w_j) - c_j b(w_j, w_j) = 0 .$$

We want now to discuss the "uniqueness" matter . If $t = 1$ the claim is obvious. Assume that $t \geq 2$ and suppose to have verified that $\langle w_k \rangle = \langle w'_k \rangle$, for any $k \leq t$. We need to show that $\langle w_{t+1} \rangle = \langle w_{t+1} \rangle$.
Since the two b-orthogonalizations satisfy (a_{t+1}), we have:

$$\langle w_1, \ldots, w_{t+1} \rangle = \langle u_1, \ldots, u_{t+1} \rangle = \langle w'_1, \ldots, w'_{t+1} \rangle .$$

Therefore,

$$w'_{t+1} = r w_{t+1} + w, \qquad r \in \mathbb{R} \text{ and } w \in \langle w_1, \ldots, w_t \rangle .$$

From (b_{t+1}), it follows:

$$\begin{cases} w'_{t+1} \in \langle w'_1, \dots, w'_t \rangle^\perp = \langle w_1, \dots, w_t \rangle^\perp & \text{and therefore } b(w, w'_{t+1}) = 0\,; \\ w_{t+1} \in \langle w_1, \dots, w_t \rangle^\perp & \text{and therefore } b(w, w_{t+1}) = 0\,. \end{cases}$$

Hence:

$$b(w,w) = b(w'_{t+1} - r w_{t+1}, w) = b(w'_{t+1}, w) - r b(w_{t+1}, w) = 0 - r0 = 0\,.$$

Consequently, since b is an inner product, $w = 0$ and it follows that $w'_{t+1} = r w_{t+1}$, that implies:

$$\langle w'_{t+1} \rangle \subseteq \langle w_{t+1} \rangle\,.$$

The other inclusion can be shown similarly. □

The proof of the above theorem suggests an algorith to find the b-orthogonalization of an order set of vectors of V. As already done before, we will set $c_0(v) = 0$ for any $v \in V$. It follows from (1) that the *Gram-Schmidt b-orthogonalization* of $\{u_1, \dots, u_t, \dots\}$ is the following:

$$\begin{aligned} w_1 &= u_1\,; \\ w_2 &= u_2 - c_{w_1}(u_2) w_1\,; \\ w_3 &= u_3 - c_{w_1}(u_3) w_1 - c_{w_2}(u_3) w_2\,; \\ &\vdots \\ w_t &= u_t - \sum_{i=1}^{t-1} c_{w_i}(u_t) w_i\,; \\ &\vdots \end{aligned}$$

Remark. Let $\mathbb{E} = (e_1, \dots, e_n)$ be a basis of V. A b-orthogonalization $\mathbb{E}'$ of $\mathbb{E}$ is still a basis of V; in fact, from the above description of the algorithm, it follows easily that $\mathbb{E}' = \mathbb{E}C$, where C is an upper triangular matrix such that $c_{11} = \ldots = c_{nn} = 1$; therefore $C \in \mathrm{GL}_n(\mathbb{R})$ and in particular $\mathbb{E}'$ is b-orthogonal. If we divide each vector by its norm, we get an orthonormal basis (*i.e.*, any vector has norm 1 and they are pairwise orthogonal).

[See also "The Factorization $A = QR$", in § 3.4 in the book]

Let us prove this important corollary of Gram-Schmidt's theorem.

Corollary 1. *Let V be a real vector space and b an inner product on it. If W is a finite dimensional vector subspace, then*

$$V = W \oplus W^\perp\,.$$

Proof. Let $\dim W = t$. Using Gram-Schmidt process, one can determine a b-orthonormal basis of W: $\{w_1, \dots, w_t\}$. Since b is an inner product, then $W \cap W^\perp = \{0\}$. For, it is sufficient to verify that $W + W^\perp = V$.
For any $v \in V$, let us consider the vector $v' = \sum_{i=1}^t b(v, w_i) w_i$. Obviously, $v' \in W$. Since $v = v' + (v - v')$, we only need to verify that $v - v' \in W^\perp$. In fact, for any $j = 1, \dots, t$:

$$\begin{aligned} b(v - v', w_j) &= b(v, w_j) - b\left(\sum_{i=1}^t b(v, w_i) w_i, w_j\right) = b(v, w_j) - \sum_{i=1}^t b(v, w_i) b(w_i, w_j) = \\ &= b(v, w_j) - b(v, w_j) = 0\,. \end{aligned}$$

□

Example. In $\mathbb{R}^3$ with the canonical inner product, consider the following five vectors:

$$u_1 = (1,0,1),\ u_2 = (-1,1,0),\ u_3 = (0,1,1),\ u_4 = (0,1,0),\ u_5 = (1,2,0)\,.$$

Let us find a b-orthogonalization of these vectors, using Gram-Schmidt process.
We have:

$$\begin{aligned}
w_1 &= u_1\,;\\
w_2 &= u_2 - \frac{u_2\cdot w_1}{w_1\cdot w_1}w_1 = u_2 - \frac{-1}{2}u_1 = \left(-\frac{1}{2}, 1, \frac{1}{2}\right)\,;\\
w_3 &= u_3 - \frac{u_3\cdot w_1}{w_1\cdot w_1}w_1 - \frac{u_3\cdot w_2}{w_2\cdot w_2}w_2 = u_3 - \frac{1}{2}u_1 - \frac{\frac{3}{2}}{\frac{3}{2}}w_2 = 0\,;\\
w_4 &= u_4 - \frac{u_4\cdot w_1}{w_1\cdot w_1}w_1 - \frac{u_4\cdot w_2}{w_2\cdot w_2}w_2 - 0w_3 = u_4 - \frac{0}{2}w_1 - \frac{1}{\frac{3}{2}}w_2 = \left(\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}\right)\,.
\end{aligned}$$

Observe now, that necessarily $w_5 = 0$ (otherwise w_1, w_2, w_4, w_5 would be four linearly independent vectors in $\mathbb{R}^3$). Indeed,

$$\begin{aligned}
w_5 &= u_5 - \frac{u_5\cdot w_1}{w_1\cdot w_1}w_1 - \frac{u_5\cdot w_2}{w_2\cdot w_2}w_2 - 0w_3 - \frac{u_5\cdot w_4}{w_4\cdot w_4}w_4 =\\
&= u_5 - \frac{1}{2}w_1 - w_2 - 3w_4 = 0\,.
\end{aligned}$$

Example. In $\mathbb{R}^3$ consider the inner product b, defined - with respect to a basis $\mathbb{E} = (e_1, e_2, e_3)$ - by the matrix

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

Using Gram-Schmidt process, find a b-orthonormal basis.

[Before solving this exercise, we should ask ourselves a natural question: how can we check that b is an inner product? This question arises the following problem: is it possible to characterize the inner products (*i.e.*, positive definiteness) in terms of their matrices? We will discuss this problem immediately after this exercise.]

Using Gram-Schmidt, we can b-orthogonalize $\mathbb{E}$:

$$\begin{aligned}
f_1 &= e_1\,;\\
f_2 &= e_2 - \frac{b(f_1,e_2)}{b(f_1,f_1)}f_1 = e_2 - \frac{1}{2}e_1\,;\\
f_3 &= e_3 - \frac{b(f_1,e_3)}{b(f_1,f_1)}f_1 - \frac{b(f_2,e_3)}{b(f_2,f_2)}f_2\,;
\end{aligned}$$

observing that

$$b(f_2,f_2) = \left(-\frac{1}{2},\ 1,\ 0\right) A \begin{pmatrix} \frac{1}{2}\\ 1\\ 0 \end{pmatrix} = \frac{3}{2} \quad \text{and} \quad b(f_2,e_3) = \left(-\frac{1}{2},\ 1,\ 0\right) A \begin{pmatrix} 0\\ 0\\ 1 \end{pmatrix} = \frac{1}{2}\,,$$

we obtain:

$$f_3 = e_3 - \frac{-1}{2}e_1 - \frac{\frac{1}{2}}{\frac{3}{2}}\left(e_2 - \frac{1}{2}e_1\right) = e_3 + \frac{1}{2}e_1 - \frac{1}{3}e_2 + \frac{1}{6}e_1 = \frac{2}{3}e_1 - \frac{1}{3}e_2 + e_3\,.$$

Hence:

$$\mathbb{F} = (f_1\ f_2\ f_3) = \mathbb{E}C \qquad \text{with } C = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{2}{3} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}.$$

Since

$$\|f_1\| = \sqrt{2}, \quad \|f_2\| = \sqrt{\frac{3}{2}} \quad \text{and} \quad \|f_3\| = \sqrt{\frac{1}{3}}\,,$$

a b-orthonormal basis is given by:

$$\mathbb{F}' = \left(\frac{f_1}{\|f_1\|}\ \frac{f_2}{\|f_2\|}\ \frac{f_3}{\|f_3\|} \right) = \mathbb{E}C' \qquad \text{with } C' = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{2\sqrt{3}}{3} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{\sqrt{3}}{3} \\ 0 & 0 & \sqrt{3} \end{pmatrix}.$$

Let us consider the problem of characterizing the matrices, corresponding to inner products in V (where V is, as usual, a n-dimensional vector space). Let us introduce the following notation: given $A \in \mathcal{M}_n(\mathbb{R})$, define α_i (for $i = 1, \ldots, n$) the determinant of the submatrix $A(1, \ldots, i|1, \ldots, i)$. These n scalars $\alpha_1, \ldots, \alpha_n$ ($\alpha_n = \det A$) are called *upper-left minors of* A (or *determinants of the upper-left submatrices*, see § 6.2 in the book). We state - without proving it - the following result:

Theorem 2 (Jabobi-Sylvester). *Let b be a symmetric bilinear form on a n-dimensional vector space V. Let $\mathbb{E}$ be a basis for V and A the matrix that defines b, with respect to $\mathbb{E}$. Then, b is positive definite (i.e., it is an inner product) if and only if all upper-left minors of A are positive: $\alpha_1, \ldots, \alpha_n > 0$.*

Remark. Observe that if some of the α_i's is equal to zero, then the matrix might not be even positive semidefinite. For instance, consider:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix};$$

in both cases $\alpha_1 = \alpha_2 = 0$, but while A is positive semidefinite, B is negative semidefinite. In order to find a necessary and sufficient condition for positive semidefiniteness, one cannot consider only the upper-left minors, but needs to consider ALL *principal minors*, *i.e.*, all determinants of submatrices of A, obtained by considering the same rows and columns: $A(i_1, \ldots, i_k|i_1, \ldots, i_k)$ for $1 \le i_1 < \ldots < i_k \le n$ and $1 \le k \le n$. The criterion in this case becomes (see also § 6.2 in the book):

Theorem 3. *Let b be a symmetric bilinear form on a n-dimensional vector space V. Let $\mathbb{E}$ be a basis for V and A the matrix that defines b, with respect to $\mathbb{E}$. Then, b is positive semidefinite if and only if all its principal minors of A are non-negative (i.e., ≥ 0).*

Example. Verify that the symmetric bilinear form b, given in the previous example, is indeed an inner product:

$$\alpha_1 = 2 > 0, \quad \alpha_2 \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3 > 0 \quad \text{and} \quad \alpha_3 = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{vmatrix} = 1 > 0\,.$$

Remark. [**Exercise**] Before closing this section, it is interesting to point out that Jacobi-Sylvester's theorem implies that, performing Gram-Schmidt process, the b-orthogonal basis $\mathbb{F}$ one gets, is such that:

$$b(f_1, f_1) = \alpha_1, \quad b(f_i, f_i) = \frac{\alpha_i}{\alpha_{i-1}}\,,$$

where α_i's, for all $i = 1, \ldots, n$ are the upper-left minors of the matrix we called A in the theorem.

[See also "QR-Factorization", "Least squares method" and "Fourier Analysis", in the book]

Department of Mathematics, Princeton University
E-mail address: `asorrent@math.princeton.edu`