Bilinear forms

August 5, 2015

In this Chapter we study finite-dimensional vector spaces over an arbitrary field $\mathbb{F}$ with a bilinear form defined on the space. This is a generalisation of the notion of an inner product space over $\mathbb{R}$.

1 The notion of bilinear form. Matrix representation. Congruent matrices.

Let V be a vector space over $\mathbb{F}$.

Definition 1.1: *A bilinear form on V is a map $g : V \times V \to \mathbb{F}$ such that for any u, u', v, v' in V and scalar $a \in \mathbb{F}$ we have*

1. *(linearity in the first variable) $g(u + u', v) = g(u, v) + g(u', v)$ and $g(au, v) = ag(u, v)$;*
2. *(linearity in the second variable) $g(u, v + v') = g(u, v) + g(u, v')$ and $g(u, av) = ag(u, v)$.*

Remark 1.2: *Equivalently, $g : V \times V \to \mathbb{F}$ is a bilinear form if and only if for all $u \in V$ the map $l_u : V \to V$ defined by $l_u : v \mapsto g(u, v)$ is a linear form on V and for all $v \in V$ the map $r_v : V \to V$ defined by $r_v : u \mapsto g(u, v)$ is a linear form on V.*

Example 1.3:

1. Let $(V, \langle , \rangle)$ be an inner product space over $\mathbb{R}$. Then $g : V \times V \to \mathbb{R}$ defined by $g(u, v) = \langle u, v\rangle$ is a bilinear form. In particular, the standard dot product in $\mathbb{R}^n$ is a bilinear form. (Note, however, that this is not so in an inner product space over $\mathbb{C}$. The standard dot product in $\mathbb{C}^n$ is not a bilinear form!)
2. The zero form. $\mathbb{F}$ is an arbitrary field and $g : V \times V \to \mathbb{F}$ is defined by $g(u, v) = 0$ for all $u, v \in V$.
3. $V = \mathbb{F}^2_{col}$ and g is the determinant form:
$$g(u, v) = \det \begin{bmatrix} x^1 & y^1 \\ x^2 & y^2 \end{bmatrix} = x^1y^2 - x^2y^1$$
for $u = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$, $v = \begin{bmatrix} y^1 \\ y^2 \end{bmatrix}$.
(Since the determinant of a matrix is linear in each of its columns when the remaining $n - 1$ columns are fixed, the example can be generalized to $V = \mathbb{F}^n_{col}$ for $n > 2$. Consider an $n \times n$ matrix with all but two columns fixed, then its determinant, considered as a function of the two remaining columns, is bilinear in its two arguments.)
4. $V = \mathbb{R}^4$ and $g(u, v) = x^1y^1 + x^2y^2 + x^3y^3 - x^4y^4$ for $u = (x^1, x^2, x^3, x^4)$ and $v = (y^1, y^2, y^3, y^4)$ (this form is called the Lorentz form, and $\mathbb{R}^4$ endowed with this form is called the Minkowski space – an important tool in the special relativity theory).
5. If g is a bilinear form on V and $f : V \to V$ is a linear operator, then $\tilde{g} : V \times V \to \mathbb{F}$ defined by $\tilde{g}(u, v) = g(f(u), v)$ is also bilinear.

Similarly to linear operators, bilinear forms can be defined using matrices (after a basis has been fixed).

Definition 1.4: *Let g be a bilinear form on a space V, and let $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ be a basis of V. Then the matrix G defined by $G = (g_{ij}) = g(b_i, b_j)$ is called the matrix of the bilinear form g with respect to the basis $\mathcal{B}$. We will also call G a Gram matrix of g.*

(Note that our convention for placement of indices in matrices of bilinear forms is different from that for matrices of linear transformations – the matrix element at the intersection of row i and column j is now denoted by g_{ij}, with both i and j being lower indices .)

Clearly the form is defined uniquely by its matrix: if $v, w \in V$ are represented in $\mathcal{B}$ as $[v]_{\mathcal{B}} = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}$ and $[w]_{\mathcal{B}} = \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix}$ then by bilinearity

$$g(v,w) = g(\sum_{i=1}^{n} v^i b_i, \sum_{j=1}^{n} w^j b_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} v^i w^j g(b_i, b_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} v^i w^j g_{ij}$$

or in short $g(v,w) = [v]^t_{\mathcal{B}} G [w]_{\mathcal{B}}$. We see that that the choice of a basis establishes one-to-one correspondence between bilinear forms on an n-dimensional space and $M_n(\mathbb{F})$.

Example 1.5:

1. The matrix G of the standard inner product of $\mathbb{R}^n$ with respect to its standard basis is the identity matrix $G = I$. More generally, the matrix G of any inner product on $\mathbb{R}^n$ with respect to any of its orthogonal bases is $G = I$.
2. The matrix of the zero form with respect to any basis is the zero matrix $G = 0$.
3. The matrix of the determinant form in $V = \mathbb{F}^2_{col}$ with respect to the standard basis of $\mathbb{F}^2_{col}$ is $G = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.
4. In $\mathbb{R}^4$ the matrix of the Lorentz form $g(u,v) = x^1y^1 + x^2y^2 + x^3y^3 - x^4y^4$ with respect to the standard basis is
$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Definition 1.6: *Let g be a bilinear form defined on V.
Then the bilinear form g^t defined by $g^t(u,v) = g(v,u)$ for all $u, v \in V$ is called the transposed form.
If $g = g^t$ then the form g is called symmetric; if $g = -g^t$ then the form g is called anti-symmetric.*

It is easy to see that a form is symmetric (anti-symmetric) iff its matrix is symmetric (anti-symmetric) in any basis (hence in every basis). So in examples 1-4 the form is either symmetric or anti-symmetric, with the exception of the zero form, which is the unique form symmetric and anti-symmetric at the same time.

Just as in the case of linear operators, we would like to know how the matrix of a bilinear form is transformed when the basis is changed. Here is the calculation: we have seen that for all $v, w \in V$ holds $g(v,w) = [v]^t_{\mathcal{B}} G [w]_{\mathcal{B}}$ where G is the matrix of g with respect to $\mathcal{B}$. In another basis $\mathcal{C}$ this would mean $g(v,w) = [v]^t_{\mathcal{C}} \tilde{G} [w]_{\mathcal{C}}$, where $\tilde{G}$ is the matrix of g with respect to $\mathcal{C}$. The two bases are connected by the change of bases (invertible!) matrix P as follows: $[v]_{\mathcal{B}} = P[v]_{\mathcal{C}}$ and $[w]_{\mathcal{B}} = P[w]_{\mathcal{C}}$. Putting all this together, we get

$$[v]^t_{\mathcal{C}} \tilde{G} [w]_{\mathcal{C}} = g(v,w) = [v]^t_{\mathcal{B}} G [w]_{\mathcal{B}} = (P[v]_{\mathcal{C}})^t G P [w]_{\mathcal{C}} = [v]^t_{\mathcal{C}} [P^t G P] [w]_{\mathcal{C}}$$

The calculation shows that the matrix of g with respect to $\mathcal{C}$ is P^tGP. This prompts the following

Definition 1.7: *Let $\mathbb{F}$ be a field. Two matrices $A \in M_n(\mathbb{F})$ and $B \in M_n(\mathbb{F})$ are called **congruent** if there exists an* invertible *$P \in M_n(\mathbb{F})$ such that $A = P^t BP$.*

If we denote the congruence by $\sim$ (write $A \sim B$ if there exists an invertible $P \in M_n(\mathbb{F})$ such that $A = P^t BP$), then one can easily check that congruence of matrices is an *equivalence relation*, meaning that for all $A, B, C \in M_n(\mathbb{F})$:

1. $A \sim A$
2. if $A \sim B$, then $B \sim A$
3. if $A \sim B$ and $B \sim C$, then $A \sim C$

This shows that $M_n(\mathbb{F})$ is split into *equivalence classes* – two matrices in the same class are congruent, two matrices in different classes are not.

Using the notion of congruence, the result we have shown can be stated in the following way:

Proposition 1.8: *Two square matrices represent the same bilinear form with respect to two bases if and only if they are congruent.*

The study of bilinear forms can therefore be interpreted as study of square matrices up to congruence. Similar problems, usually called classification problems, appear in other settings in linear algebra – studying matrices up to similarity (as matrices of linear operators in a vector space), or up to unitary similarity (as matrices of linear operators in an inner product space). In each such situation we are interested in questions like: Under what conditions is a given matrix equivalent to an especially simply-looking one (like diagonal matrix)? What is the "normal form", also called "canonical form", to which any matrix can be reduced (meaning that there is exactly one matrix of such type in every equivalence class)? etc. In terms of operators, we are looking for a basis with respect to which the matrix of a given operator or bilinear form looks especially simple.

2 Symmetric forms. Existence of an orthogonal basis.

In this section $\mathbb{F} = \mathbb{R}$ and g denotes a **symmetric** bilinear form. In a way, the theory developed here is parallel to the theory of inner product real vector spaces, but there are differences as well. In inner product spaces we have $\langle v, v\rangle > 0$ for every vector $v \neq 0$. Now, when the positivity condition has been dropped, it is possible to have for a non-zero vector $v \in V$: $g(v, v) = 0$ or $g(v, v) < 0$.

The orthogonality notation is used without change:

Definition 2.1: *For $u, v \in V$ we write $u \perp_g v$ and say "u is orthogonal to v with respect to the form g" if $g(u, v) = 0$. The reference to the form can be omitted when g is clear from the context, and the notation can be short-handed to $u \perp v$.*

Example 2.2:

1. If $g = 0$ then $u \perp v$ for all $u, v \in V$.
2. The vector $v = (1, 0, 0, -1)$ is orthogonal to itself with respect to the Lorentz form in $\mathbb{R}^4$.
3. The vector $e_2 = (0, 1)$ is orthogonal to any vector in $\mathbb{R}^2$ with respect to the form given by $G = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}$ in the standard basis (e_1, e_2) of $\mathbb{R}^2$.

Remark 2.3: *Just as we defined the notion of orthogonality for symmetric forms, it can be defined for anti-symmetric forms (and it will still be a symmetric relation: $u \perp v$ if and only if $v \perp u$). In that setting every vector is orthogonal to itself. Indeed, for all $u \in V$ holds $g(u, u) = -g(u, u)$ so $2g(u, u) = 0$, hence $g(u, u) = 0$. (We consider the case $\mathbb{F} = \mathbb{R}$, so $\mathrm{char}\mathbb{F} \neq 2$ therefore here and in the sequel we are free to divide by 2 or by 4. $2 \neq 0$ so $4 = 2 \cdot 2 \neq 0$.)*

Definition 2.4: *Let g be a bilinear form on V. The function $q : V \to \mathbb{F}$ defined by $q(v) = g(v, v)$ is called the quadratic form associated with g.*

The following proposition shows that no info is lost when a bilinear form is replaced by its associated quadratic form: g is uniquely defined by q.

Proposition 2.5: *Let $q : V \to \mathbb{R}$. There is at most one bilinear form g for which q serves as its associated quadratic form.*

Proof. Let g be a bilinear form, and q be its associated quadratic form. Note that for all $u, v \in V$ holds $2g(u,v) = g(u+v, u+v) - g(u,u) - g(v,v) = q(u+v) - q(u) - q(v)$, so

$$g(u,v) = \frac{q(u+v) - q(u) - q(v)}{2}$$

Therefore if q is associated with an additional bilinear form $\tilde{g}$, then necessarily $g \equiv \tilde{g}$. □

Remark 2.6: *Alternatively, one can use the polarization identity, whose version appeared in the chapter on inner product spaces. It holds for a general symmetric bilinear form without any change:*

$$g(u,v) = \frac{q(u+v) - q(u-v)}{4}$$

The following proposition shows that in our symmetric case there is always a vector which is not self-orthogonal (unless $g \equiv 0$).

Proposition 2.7: *For any form g on V which is not identically zero there exists a vector $v \in V$ such that $g(v,v) \neq 0$.*

Proof. Let g be a bilinear form which is not identically zero. Use any of the two identities above. There is a pair of vectors $u, v \in V$ for which the left-hand side is non-zero. Hence at least one of the terms in the right-hand side is non-zero, and there exists $w \in V$ for which $q(w) \neq 0$. □

Definition 2.8: *A bilinear symmetric form g on a space V is called non-degenerate if the set $V_0 = \{u \mid g(u,v) = 0$ for all $v \in V\}$ is trivial: $V_0 = \{0\}$.*

Example 2.9:

1. Inner product is a non-generate bilinear form.
2. The Lorentz form is a non-generate bilinear form in $\mathbb{R}^4$ (prove this).

Proposition 2.10: *A form g on a space V is non-degenerate if and only if its matrix G with respect to a basis $\mathcal{B}$ of V is non-degenerate.*

Proof. Let $\mathcal{B}$ be a basis of V. Recall that $g(v,w) = [v]^t_{\mathcal{B}} G [w]_{\mathcal{B}}$.
g is degenerate
$\Leftrightarrow$ there is $0 \neq v \in V$ such that for all $w \in V$ holds $[v]^t_{\mathcal{B}} \cdot G[w]_{\mathcal{B}} = 0$
$\Leftrightarrow$ the subspace $W = \{G[w]_{\mathcal{B}} \mid w \in V\}$ is a **proper** subspace of V: $W \neq V$
$\Leftrightarrow$ G is degenerate. □

$W^\perp$ is defined just as in inner product spaces.

Definition 2.11: $W^\perp = \{u \mid g(u,w) = 0$ *for all* $w \in W\}$

In this notation $V_0 = V^\perp$.

In the sequel we will at times need to restrict a given bilinear form g defined on a vector space V to a subspace W. We will denote such restriction by $g|_W$ (this is an abuse of notation for the sake of brevity, in fact we mean $g|_{W \times W}$) – $g|_W$ is the bilinear form defined on W as follows: $g|_W(w_1, w_2) = g(w_1, w_2)$ for all $w_1, w_2 \in W$. If g is symmetric, then so is $g|_W$.

Proposition 2.12: *Let g be a symmetric form on a finite-dimensional vector space V, and W be a subspace V. Then the following claims hold.*

1. *If g is non-degenerate, then* $\dim W + \dim(W^\perp) = \dim V$.

2. *If g is non-degenerate, then $(W^\perp)^\perp = W$.*

3. *If the restriction $g|_W$ is non-degenerate, then $V = W \oplus W^\perp$.*

Remark 2.13: *The condition in clause 3 is indispensable: if, for instance, $W = \mathrm{Span}(v)$ where $v \perp v$ (as we know, this is possible even for a non-degenerate g), then $W \subset W^\perp$.*

Proof. 1. Let $\dim V = n$ and $\dim W = k$. If $\mathcal{B} = (b_1, \ldots, b_k)$ is a basis for W, then $v \in W^\perp$ if and only if for all $1 \leq k \leq n$ holds $v \perp b_k$, or, in matrix notation, $[v]^t_{\mathcal{B}} \cdot G[b_k]_{\mathcal{B}} = 0$. These are k equations with respect to $[v]_{\mathcal{B}}$. Since the vectors of the basis are independent and G is non-degenerate, the k vectors $G[b_k]_{\mathcal{B}}$ $(1 \leq k \leq n)$ are independent. The solution space of this system of k linearly independent equations is therefore $(n-k)$-dimensional, as claimed.

2. By the definition of $W^\perp$ we have in particular that $w \perp u$ for all $w \in W$ and $u \in W^\perp$. Equivalently if $w \in W$, then $w \perp u$ for all $u \in W^\perp$, so $W \in (W^\perp)^\perp$ by definition. Now note that by the previous clause $\dim(W^\perp)^\perp = n - \dim(W^\perp) = n - (n - \dim W) = \dim W$. Since in addition $W \in (W^\perp)^\perp$, the two spaces must be equal.

3. Let $\dim V = n$, $\dim W = k$, and $\mathcal{B} = (b_1, \ldots, b_k)$ be a basis for W. Define $f \colon V \to \mathbb{F}^k$ by $f(v) = (g(v, b_1), \ldots, g(v, b_k))$ (check that the kernel of this mapping is exactly $W^\perp$). Consider the restriction $f|_W$. Since $g|_W$ is non-degenerate, $\ker(f|_W) = \{0\}$, so $\dim(\mathrm{im}\ (f|_W)) = k - \dim(\ker(f|_W)) = k$. Therefore $\mathrm{im}\ (f|_W) = \mathbb{F}^k$, and all the more so $\mathrm{im}\ f = \mathbb{F}^k$. So $\dim(W^\perp) = \dim(\ker f) = n - \dim(\mathrm{im}\ f) = n - k$.
The subspaces W and $W^\perp$ intersect trivially: $W \cap W^\perp = \{0\}$, because if there were $0 \neq v \in W \cap W^\perp$, then v would be orthogonal (with respect to g) to every vector $u \in V$ $(g(u, v) = 0)$, in contradiction to the non-degeneration of $g|_W$. In addition, as we have just shown, the dimensions of W and $W^\perp$ (k and $n-k$ respectively) add up to $\dim V = n$. Therefore the claimed decomposition holds: $V = W \oplus W^\perp$.

$\square$

The following theorem establishes the existence of a basis $\mathcal{B} = (b_1, \ldots, b_n)$ in V, which is an **orthogonal** basis for g: for all $i \neq j$ holds $g(b_i, b_j) = 0$. A basis $\mathcal{B}$ is an orthogonal basis for g if and only if the Gram matrix G of g with respect to $\mathcal{B}$ is a **diagonal** matrix.

Theorem 2.14: *For any symmetric bilinear form g there exists an orthogonal basis. (And equivalently, in matrix language — any symmetric real matrix is congruent to a diagonal one: for all $A \in M_n(\mathbb{R})$ there is non-degenerate $P \in M_n(\mathbb{R})$ such that P^tAP is a diagonal matrix.)*

Proof. Induction on $\dim V$. If $\dim V = 1$, then any basis is orthogonal. Let $\dim V = n$, $n \geq 2$ and assume the claim to be true for any space of dimension smaller than n. Either g is identically zero on V (then any basis is an orthogonal basis for g), or, by Proposition 2.7, there is a vector $v \in V$ such that $g(v, v) \neq 0$. This condition means that for $W = \mathrm{Span}(v)$ the restriction $g|_W$ is non-degenerate. Hence by clause 3 of Proposition 2.12 V can be split into $V = W \oplus W^\perp$. $\dim W^\perp = n - 1$ hence by the induction assumption $W^\perp$ has an orthogonal (with respect to g) basis $(v_1, ..., v_{n-1})$. Then $\mathcal{B} = (v, v_1, ..., v_{n-1})$ is an orthogonal (with respect to g) basis of V. $\square$

Below is an example of how this diagonalization-by-congruencies is done in practice. We are looking for a non-degenerate (= regular = invertible) matrix P such that $D = P^tAP$ is diagonal. (If A is the Gram matrix of g in a basis $\mathcal{B}$, then columns of P represent vectors of the new orthogonal basis $\mathcal{D}$ in the old basis $\mathcal{B}$: $[v]_{\mathcal{B}} = P[v]_{\mathcal{D}}$ (sic), and

$$g(v, w) = [v]^t_{\mathcal{B}} A [w]_{\mathcal{B}} = [v]^t_{\mathcal{D}} [P^tAP] [w]_{\mathcal{D}}$$

The matrix P will be built as a product of elementary matrices, corresponding to elementary operations performed on rows of a matrix in the row reduction-process (rows interchange, multiplication by a non-zero constant, addition to a row a multiple of another one). Performing such operation on rows (columns) of a matrix A, we multiply A by an invertible matrix on the left (on the right). So if we perform *the same* elementary operation on the rows and then on the columns, we get $Q_1^tAQ_1$, then $Q_2^tQ_1^tAQ_1Q_2 = (Q_1Q_2)^tAQ_1Q_2$, etc, finally arriving at $D = (Q_1 \dots Q_n)^tA(Q_1 \dots Q_n)$. This way we get both $P = Q_1 \dots Q_n$ and $D = P^tAP$. To keep track of changes of bases, we put the unity matrix on the right of A and simultaneously perform the same operations on the rows only of that matrix to get $(Q_1 \dots Q_n)^tI = P^t$. (We could get P itself instead of P^t, doing elementary operations on the *columns* of the matrix on the right, but our routine is slightly more convenient – the row operations are performed on the rows of the "big" double-sized matrix, while the column operations are performed on its left half only. A small price to pay is to remember that in the end of the process the new basis should be read from the rows, rather than columns, of the matrix on the right. [1])

Perform this diagonalization for

$$A = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{bmatrix}$$

$$\left[\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-2R_1+R_3\to R_3]{3R_1+R_2\to R_2} \left[\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right] \xrightarrow[-2C_1+C_3\to C_3]{3C_1+C_2\to C_2}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right] \xrightarrow{R_2+2R_3\to R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right] \xrightarrow{C_2+2C_3\to C_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right]$$

So

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

If A is the matrix of g in the standard basis of $\mathbb{R}^3_{col}$, then $\mathcal{D} = (p_1, p_2, p_3)$ where

$$p_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, p_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, p_3 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

is an orthogonal basis for g.

3 Classificaton of Real Bilinear Forms. Sylvester's Law of Inertia. Signature of a form.

In this section V still denotes a vector space over $\mathbb{R}$, and g a symmetric bilinear form defined on V. We have seen that there is an orthogonal (with respect to g) basis of V – a basis $\mathcal{B} = (b_1, \dots, b_n)$ such that $g(b_i, b_j) = 0$ whenever $i \neq j$. One can get a little bit more doing a version of normalising. Note that $g(av, av) = a^2g(v, v)$, so replacing each b_i that is not orthogonal to itself with $b_i/\sqrt{|g(b_i, b_i)|}$, we can assume that, in addition to the orthogonality condition, holds $g(b_i, b_i) \in \{1, -1, 0\}$ for all $i = 1, \dots, n$.

[1] One can also write I below A instead of on its left, and then the operations on the big matrix will be performed on the columns of I, leading to P instead of P^t.

Finally, taking vectors of the basis in appropriate order, we can obtain the Gram matrix in the block form

$$G_{p,m,z} = \begin{bmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{bmatrix}$$

This is a diagonal matrix with p ones (pluses), m negative ones (minuses), and z zeros on the diagonal. Sylvester's Law of Inertia, to be stated and proved shortly, asserts that these three numbers are defined uniquely by g and are independent of a particular choice of basis with respect to which the Gram matrix of g is diagonal with positive ones, negative ones and zeros on the diagonal.

For a given symmetric bilinear form g we define, independently of any basis, 3 numbers. Let
p be the maximum dimension of a subspace W such that $g|_W$ is positive,
m be the maximum dimension of a subspace W such that $(-g)|_W$ is positive,
$n = \dim V_0$.
(Recall the definitions and notation. A symmetric bilinear form g is positive definite, or simply positive, on U if for all $u \in V$ holds $g(u,u) \geq 0$ and $g(u,u) = 0$ iff $u = 0$. In other words $g|_U$ is an inner product. And the kernel or a form g is defined as the set of vectors orthogonal, with respect to g, to the whole space: $V_0 = V^\perp = \{v \in V \mid g(v,u) = 0 \text{ for all } u \in V\}$.)

We show how it works in an important case of g defined on $\mathbb{R}^3$ whose Gram matrix with respect to the standard basis of $\mathbb{R}^3$ is $G = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{smallmatrix}\right]$. For $v = (x,y,z) \in \mathbb{R}^3$ we have $g(v,v) = x^2 + y^2 - z^2$. ($q(v) = q(x,y,z) = x^2 + y^2 - z^2$ is the quadratic form defined by g. This example is a 3-dimensional analog of the 4-dimensional Minkowski space with the Lorentz form defined above.) So $\mathbb{R}^3$ is split into 3 parts – the cone of self-orthogonal vectors $\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 0\}$, the set of vectors inside the cone $\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 < 0\}$, and the set of vectors outside the cone $\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 > 0\}$. Draw the picture or refer to:
`https://en.wikipedia.org/wiki/Minkowski_space#/media/File:World_line.svg`
Our goal is to find p, m, n for this particular example.
$n = 0$ because the null space of g is trivial – $V_0 = \{0\}$: while there are non-zero self-orthogonal vectors, there are no non-zero vectors orthogonal to the whole space. This can be easily checked directly (what are $g(e_1, v)$, $g(e_2, v)$, $g(e_3, v)$ for a general $v = (x,y,z)$?), but it is also implied by 2.10.
g is positive on the 2-dimensional subspace $\{z = 0\}$ but is not positive on any 3-dimensional subspace (the only one such subspace being V itself), so $p = 2$.
$(-g)$ is positive (one can say "g is negative") on the 1-dimensional subspace $\{x = y = 0\}$ but is not positive on any 2-dimensional subspace, because no plane through the origin of $\mathbb{R}^3$ is wholly contained inside the cone. Therefore $m = 1$.

It should be noted that the spaces of maximal dimension on which g is positive or negative are not unique. A perturbed x,y-plane or the z-axis would do as well. Besides $\{z = 0\}$ the form g is positive on $\{z + 0.1x = 0\}$ or on $\{z + 0.1x - 0.15y = 0\}$ etc. Besides $\{x = y = 0\}$ the form g is negative on $\{x = 0, y = 0.1z\}$ or on $\{x = 0.1z, y = 0.15z\}$ etc.

Theorem 3.1 (Sylvester's Law of Inertia)**:** *Let g be a symmetric bilinear form defined on V and $\mathcal{B}$ be an orthogonal (with respect to g) basis of V. Then number of vectors v of the basis $\mathcal{B}$ for which $g(v,v) > 0 \,/\, g(v,v) < 0 \,/\, g(v,v) = 0$ is uniquely defined by g and does not depend on $\mathcal{B}$.*
Equivalently, in matrix language: A symmetric real matrix is congruent to only one matrix of the form $G_{p,m,z} = \left[\begin{smallmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{smallmatrix}\right]$ (The sizes of blocks of positive ones, negative ones and zeros are uniquely defined by g.)

Definition 3.2: *The pair (p,m) is called the signature of the bilinear form g (or of the associated* *quadratic form q).*

(Some sources define signature as a single integer – the difference $p-m$, instead of the pair (p,m).)

Examples: the inner product has signature $(n,0)$, the zero form – $(0,0)$, the Lorentz form – $(3,1)$. (Sometimes the Lorentz form is defined as the one of signature $(1,3)$.)

Proof. Let $\mathcal{B}$ be an orthogonal (with respect to g) basis with p', m', z' vectors b_i with the property $g(b_i,b_i)>0\,/\,g(b_i,b_i)<0\,/\,g(b_i,b_i)=0$ respectively. We need to show that $p'=p$, $m'=m$, $n'=n$.
Let W_+,W_-,W_0 be the subspaces of V spanned by the b_i's of the corresponding type. $V=W_+\oplus W_-\oplus W_0$ is an orthogonal (with respect to g) decomposition of V with $\dim W_+=p'$, $\dim W_-=m'$, $\dim W_0=z'$. It is an immediate check that g is positive definite on W_+ (meaning that $g(v,v)\geq 0$ for all $v\in W_+$ and $g(v,v)>0$ for all $0\neq v\in W_+$), $(-g)$ is positive definite on W_-, and that if $w\in W_0$, then $g(w,v)=0$ for all $v\in V$.
We claim that $W_0=V_0$ (the kernel of g). Indeed, as we have just noted, if $w\in W_0$, then $w\in V_0$, so $W_0\subset V_0$. Conversely, let $v\in V_0$. Decomposing v as we may, $v=w_++w_-+w_0$ into an orthogonal (with respect to g) sum, with the three summands in the three respective spaces and taking $g(v,w_+)$, we obtain

$$0=g(v,w_+)=g(w_++w_-+w_0,w_+)=g(w_+,w_+)+g(w_-,w_+)+g(w_0,w_+)=g(w_+,w_+)+0+0$$

so $w_+=0$. Considering $g(v,w_-)$ we prove in a similar way that $w_-=0$. Hence $v=w_++w_-+w_0=0+0+w_0\in W_0$. We have just shown that $V_0\subset W_0$, thus establishing that the two spaces are equal, and in particular their dimensions are equal: $n'=n$.

To show that $p'=p$ recall that g is positive definite on W_+, and hence $p'\leq p$ by definition of p. it is left to show that $p\leq p'$. To demonstrate this consider 2 spaces – a space W of maximal dimension (p) on which g is positive definite, and the space $W_-\oplus W_0$. We claim that they intersect trivially: $W\cap(W_-\oplus W_0)=\{0\}$. Indeed, let $v\in W\cap(W_-\oplus W_0)$. Since $v\in W_-\oplus W_0$ we have $g(v,v)\leq 0$ (check this). Since $v\in W$ we have $g(v,v)\geq 0$. Therefore $g(v,v)=0$ and, again since $v\in W$, finally $v=0$.

The trivial intersection of W and $W_-\oplus W_0$ implies that their dimensions add up to at most $\dim V$: $p+(m'+z')\leq n$. However $p'+(m'+z')=n$, so $p\leq p'$. The relation $p=p'$ is established. Repeating the argument for $(-g)$ we can get $m=m'$.

□

Using the diagonalization procedure developed in the previous section, we get the signature of a form g, among other things ($(2,1)$ for the example in the end of the previous section as the diagonal matrix D that we obtained has 2 positive entries and one negative entry on the diagonal). However, if we are interested in the signature only, it is probably easier to work with the quadratic form q associated with the given bilinear form g. If q is reduced to an algebraic sums of squares of new variables obtained from the old ones by *invertible* linear changes of variables, then the signature is (p,m), where p is the number of positive terms and m is the number of negative ones.

Example 3.3: 1. $q(x,y,z)=x^2+y^2+3z^2-2xz-4yz$, or equivalently g is the bilinear form for which the Gram matrix

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

in the basis, for which x,y,z are the coordinates. (Note that the non-diagonal entries in the matrix come in pairs, so each one is a half of the corresponding coefficient of the quadratic form.)

Completing squares, we get

$$\begin{aligned} q(x,y,z) &= x^2 + y^2 + 3z^2 - 2xz - 4yz \\ &= (x-z)^2 - z^2 + y^2 + 3z^2 - 4yz \\ &= (x-z)^2 + y^2 - 4yz + 2z^2 \\ &= (x-z)^2 + (y-2z)^2 - 2z^2 \end{aligned} \tag{1}$$

so the signature of g (or of q) is $(2,1)$.

2. $q(x,y) = xy$, or equivalently g is the bilinear form for which the Gram matrix

$$G = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}$$

in the basis, for which x, y are the coordinates. A nice little trick works here:

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

(recall polarization identity), so the signature of g (or of q) is $(1,1)$.

We have been very sketchy in presentation of the diagonalization algorithms (in matrix notation of in quadratic polynomials notation). These techniques are developed in detail in the tutorials.

And a final note. We have classified bilinear forms defined on vector spaces over $\mathbb{R}$. (Equivalently, we have classified real square matrices up to congruence – each such matrix is congruent to exactly one matrix of the form $G_{p,m,n}$.) When the underlying field is finite, the classification is much more involved and studied in advanced courses. However, for bilinear forms defined on vector spaces over $\mathbb{C}$ [2] the theory is simpler. Since every complex number has a square root, a diagonal complex matrix is congruent to one which is diagonal with solely ones and zeros on the diagonal. So complex $n \times n$ matrices are classified up to congruence by their rank. See tutorial notes for more details.

[2]not to be confused with inner product in Hermitian spaces, which are linear in the first argument and anti-linear in the second one