Linear Algebra 2

Lecture #25

Bilinear and quadratic forms

Irena Penev

May 14, 2025

- This lecture has four parts:

- This lecture has four parts:
    1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$

- This lecture has four parts:
    1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$
    2. Bilinear forms

- This lecture has four parts:
  1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$
  2. Bilinear forms
  3. Quadratic forms

- This lecture has four parts:
  1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$
  2. Bilinear forms
  3. Quadratic forms
  4. Quadratic forms on $\mathbb{R}^n$

1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$

1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$

Proposition 9.1.1

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$. Then both the following hold:

a. for all vectors $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \dots & x_n \end{array}\right]^T$ and $\mathbf{y} = \left[\begin{array}{ccc} y_1 & \dots & y_n \end{array}\right]^T$ in $\mathbb{F}^n$, we have that
$$\mathbf{x}^T A\mathbf{y} \quad = \quad \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_iy_j;$$

b. for all indices $i, j \in \{1, \dots, n\}$, we have that $\mathbf{e}_i^T A\mathbf{e}_j = a_{i,j}$.

Proof.

1. A formula for products of the form $\mathbf{x}^T A\mathbf{y}$

Proposition 9.1.1

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ be a matrix in $\mathbb{F}^{n\times n}$. Then both the following hold:

a. for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that
$$\mathbf{x}^T A\mathbf{y} \quad = \quad \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i y_j;$$
b. for all indices $i, j \in \{1, \dots, n\}$, we have that $\mathbf{e}_i^T A\mathbf{e}_j = a_{i,j}$.

Proof. Obviously, (a) implies (b). So, let us prove (a).

Proof (continued).

Proof (continued). For any vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have the following:

$$\begin{aligned}
\mathbf{x}^T A\mathbf{y} &= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} \sum_{j=1}^n a_{1,j} y_j \\ \sum_{j=1}^n a_{2,j} y_j \\ \vdots \\ \sum_{j=1}^n a_{n,j} y_j \end{bmatrix} \\
&= \sum_{i=1}^n x_i \Big(\sum_{j=1}^n a_{i,j} y_j \Big) \; = \; \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i y_j.
\end{aligned}$$

This proves (a). $\square$

Proposition 9.1.1

Let $\mathbb{F}$ be a field, let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$, and let $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$. Then both the following hold:

(a) for all vectors $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \dots & x_n \end{array}\right]^T$ and $\mathbf{y} = \left[\begin{array}{ccc} y_1 & \dots & y_n \end{array}\right]^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x}^T A \mathbf{y} \quad = \quad \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j;$$

(b) for all indices $i, j \in \{1, \dots, n\}$, we have that $\mathbf{e}_i^T A \mathbf{e}_j = a_{i,j}$.

2. Bilinear forms

2. Bilinear forms

Definition

A *bilinear form* on a vector space V over a field $\mathbb{F}$ is a function $f : V \times V \to \mathbb{F}$ that satisfies the following four axioms:

b.1. $\forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$: $f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = f(\mathbf{x}_1, \mathbf{y}) + f(\mathbf{x}_2, \mathbf{y})$;

b.2. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$: $f(\alpha\mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$;

b.3. $\forall \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in V$: $f(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = f(\mathbf{x}, \mathbf{y}_1) + f(\mathbf{x}, \mathbf{y}_2)$;

b.4. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{F}$: $f(\mathbf{x}, \alpha\mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$.

The bilinear form f is said to be *symmetric* if it further satisfies the property that $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in V$.

- Reminder:

Definition

A *scalar product* (also called *inner product*) in a **real** vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ that satisfies the following four axioms:

r.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

r.2. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

r.3. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}$: $\langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$;

r.4. $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.

r.2'. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;

r.3'. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$, $\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$.

- Reminder:

Definition

A *scalar product* (also called *inner product*) in a **real** vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ that satisfies the following four axioms:

r.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

r.2. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

r.3. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}$: $\langle \alpha\mathbf{x}, \mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$;

r.4. $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.

r.2'. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;
r.3'. $\forall \mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$, $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{y} \rangle$.

- **Remark:** every scalar product $\langle \cdot, \cdot \rangle$ in a **real** vector space V is a symmetric bilinear form.
  - Indeed, r.2, r.3, r.2', and r.3' are precisely the axioms b.1, b.2, b.3, and b.4, respectively.
  - Moreover, by r.4, scalar products in real vector spaces are symmetric.

- Reminder:

Definition

A *scalar product* (also called *inner product*) in a **complex** vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ that satisfies the following four axioms:

c.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle$ is a real number, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

c.2. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$;

c.3. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{C}$: $\langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$;

c.4. $\forall \mathbf{x}, \mathbf{y} \in V$: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$.

c.2'. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$: $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$;

c.3'. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{C}$: $\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle$.

- **Remark:** scalar products in non-trivial **complex** vector spaces are not bilinear forms, since c.1 and c.3' together contradict axiom b.4 (next slide).

c.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle$ is a real number, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

c.3'. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{C}$: $\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle$.

- Indeed, if $\langle \cdot, \cdot \rangle$ is a scalar product in a non-trivial complex vector space V, then for any $\mathbf{x} \in V \setminus \{\mathbf{0}\}$, c.1 guarantees that $\langle \mathbf{x}, \mathbf{x} \rangle \neq 0$, c.1. $\forall \mathbf{x} \in V$: $\langle \mathbf{x}, \mathbf{x} \rangle$ is a real number, $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, and equality holds iff $\mathbf{x} = \mathbf{0}$;

c.3'. $\forall \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{C}$: $\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle$.

- Indeed, if $\langle \cdot, \cdot \rangle$ is a scalar product in a non-trivial complex vector space V, then for any $\mathbf{x} \in V \setminus \{\mathbf{0}\}$, c.1 guarantees that $\langle \mathbf{x}, \mathbf{x} \rangle \neq 0$, and so

$$\langle \mathbf{x}, i\mathbf{x} \rangle \quad \overset{\text{c.3'}}{=} \quad \overline{i} \langle \mathbf{x}, \mathbf{x} \rangle \quad = \quad -i \langle \mathbf{x}, \mathbf{x} \rangle \quad \neq \quad i \langle \mathbf{x}, \mathbf{x} \rangle,$$

and we see that b.4 does not hold.

Proposition 9.2.1

Let V be a vector space over a field $\mathbb{F}$, and let f be a bilinear form on V. Then all the following hold:

(a) $\forall \mathbf{x} \in V$: $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) $\forall \mathbf{y} \in V$: $f(\mathbf{0}, \mathbf{y}) = 0$;

(c) $f(\mathbf{0}, \mathbf{0}) = 0$.

Proof.

Proposition 9.2.1

Let V be a vector space over a field $\mathbb{F}$, and let f be a bilinear form on V. Then all the following hold:

(a) $\forall \mathbf{x} \in V$: $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) $\forall \mathbf{y} \in V$: $f(\mathbf{0}, \mathbf{y}) = 0$;

(c) $f(\mathbf{0}, \mathbf{0}) = 0$.

Proof. For (a), we fix a vector $\mathbf{x} \in V$, and we compute:

$$f(\mathbf{x}, \mathbf{0}) \;=\; f(\mathbf{x}, \mathbf{0} + \mathbf{0}) \;\overset{\text{b.3}}{=}\; f(\mathbf{x}, \mathbf{0}) + f(\mathbf{x}, \mathbf{0}).$$

By subtracting $f(\mathbf{x}, \mathbf{0})$ from both sides, we obtain $\mathbf{0} = f(\mathbf{x}, \mathbf{0})$. This proves (a).

Proposition 9.2.1

Let V be a vector space over a field $\mathbb{F}$, and let f be a bilinear form on V. Then all the following hold:

(a) $\forall \mathbf{x} \in V$: $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) $\forall \mathbf{y} \in V$: $f(\mathbf{0}, \mathbf{y}) = 0$;

(c) $f(\mathbf{0}, \mathbf{0}) = 0$.

Proof. For (a), we fix a vector $\mathbf{x} \in V$, and we compute:

$$f(\mathbf{x}, \mathbf{0}) \;=\; f(\mathbf{x}, \mathbf{0} + \mathbf{0}) \;\overset{\text{b.3}}{=}\; f(\mathbf{x}, \mathbf{0}) + f(\mathbf{x}, \mathbf{0}).$$

By subtracting $f(\mathbf{x}, \mathbf{0})$ from both sides, we obtain $\mathbf{0} = f(\mathbf{x}, \mathbf{0})$. This proves (a).

The proof of (b) is similar.

Proposition 9.2.1

Let V be a vector space over a field $\mathbb{F}$, and let f be a bilinear form on V. Then all the following hold:

(a) $\forall \mathbf{x} \in V$: $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) $\forall \mathbf{y} \in V$: $f(\mathbf{0}, \mathbf{y}) = 0$;

(c) $f(\mathbf{0}, \mathbf{0}) = 0$.

Proof. For (a), we fix a vector $\mathbf{x} \in V$, and we compute:

$$f(\mathbf{x}, \mathbf{0}) = f(\mathbf{x}, \mathbf{0} + \mathbf{0}) \overset{\text{b.3}}{=} f(\mathbf{x}, \mathbf{0}) + f(\mathbf{x}, \mathbf{0}).$$

By subtracting $f(\mathbf{x}, \mathbf{0})$ from both sides, we obtain $\mathbf{0} = f(\mathbf{x}, \mathbf{0})$. This proves (a).

The proof of (b) is similar. Finally, (c) is a special case of (a) for $\mathbf{x} = \mathbf{0}$. $\square$

- Reminder:

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n \times m}$, denoted by ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$$ {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}} \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \ = \ \left[\ f(\mathbf{u})\ \right]_{\mathcal{C}}. $$

Moreover, the matrix ${}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}\left[\ f\ \right]_{\mathcal{B}} \ = \ \left[\ \left[\ f(\mathbf{b}_1)\ \right]_{\mathcal{C}} \ \ \dots \ \ \left[\ f(\mathbf{b}_m)\ \right]_{\mathcal{C}}\ \right]. $$

- Reminder:

Theorem 4.5.1

Let U and V be non-trivial, finite-dimensional vector spaces over a field $\mathbb{F}$. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ be a basis of U, let $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ be a basis of V, and let $f : U \to V$ be a linear function. Then exists a unique matrix in $\mathbb{F}^{n \times m}$, denoted by ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ and called the *matrix of f with respect to $\mathcal{B}$ and $\mathcal{C}$*, s.t. for all $\mathbf{u} \in U$, we have that

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ \big[\ \mathbf{u}\ \big]_{\mathcal{B}} \ = \ \big[\ f(\mathbf{u})\ \big]_{\mathcal{C}}. $$

Moreover, the matrix ${}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}}$ is given by

$$ {}_{\mathcal{C}}\big[\ f\ \big]_{\mathcal{B}} \ = \ \Big[\ \big[\ f(\mathbf{b}_1)\ \big]_{\mathcal{C}} \ \ \dots \ \ \big[\ f(\mathbf{b}_m)\ \big]_{\mathcal{C}}\ \Big]. $$

- For bilinear forms, we have the following (next slide).

Theorem 9.2.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V.

(a) For every matrix $A = \left[\; a_{i,j} \; \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, the function $f : V \times V \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) \;\; = \;\; \left[\; \mathbf{x} \; \right]_{\mathcal{B}}^{T} \, A \, \left[\; \mathbf{y} \; \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V$$

is a bilinear form on V, and moreover, all the following hold:

- (a.1) $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$,
- (a.2) $f\left(\sum_{i=1}^{n} c_i \mathbf{b}_i, \sum_{j=1}^{n} d_j \mathbf{b}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} c_i d_j$ for all $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$,
- (a.3) f is symmetric iff A is symmetric.

(b) For every bilinear form f on V, there exists a unique matrix $A = \left[\; a_{i,j} \; \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, called the *matrix of the bilinear form f with respect to the basis $\mathcal{B}$*, that satisfies the property that

$$f(\mathbf{x}, \mathbf{y}) \;\; = \;\; \left[\; \mathbf{x} \; \right]_{\mathcal{B}}^{T} \, A \, \left[\; \mathbf{y} \; \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Moreover, the entries of the matrix A are given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$.

Proof. (a) Fix a matrix $A = \left[\ a_{i,j}\ \right]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x},\mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x},\mathbf{y} \in V.$$

Proof. (a) Fix a matrix $A = \left[\; a_{i,j} \; \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, and define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\; \mathbf{x} \; \right]_{\mathcal{B}}^{T} A \left[\; \mathbf{y} \; \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Let us first check that f is bilinear.

Proof. (a) Fix a matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, and define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Let us first check that f is bilinear. We must check that f satisfies axioms b.1-b.4.

Proof. (a) Fix a matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$, and define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Let us first check that f is bilinear. We must check that f satisfies axioms b.1-b.4. For b.1, we observe that for all vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$, we have the following:

$$\begin{aligned}
f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) \quad &= \quad \left[\ \mathbf{x}_1 + \mathbf{x}_2\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \\
&\stackrel{(*)}{=} \quad \Big(\left[\ \mathbf{x}_1\ \right]_{\mathcal{B}} + \left[\ \mathbf{x}_2\ \right]_{\mathcal{B}} \Big)^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \\
&= \quad \Big(\left[\ \mathbf{x}_1\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \Big) + \Big(\left[\ \mathbf{x}_2\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \Big) \\
&= \quad f(\mathbf{x}_1, \mathbf{y}) + f(\mathbf{x}_2, \mathbf{y}),
\end{aligned}$$

where (*) follows from the linearity of $\left[\ \cdot\ \right]_{\mathcal{B}}$. Thus, f satisfies b.1, and similarly, it satisfies b.3.

Proof (continued). For b.2, we observe that for all vectors $\mathbf{x}, \mathbf{y} \in V$ and scalars $\alpha \in \mathbb{F}$, we have the following:

$$\begin{aligned} f(\alpha\mathbf{x}, \mathbf{y}) &= \left[\ \alpha\mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \\ &\overset{(*)}{=} \Big(\alpha \left[\ \mathbf{x}\ \right]_{\mathcal{B}}\Big)^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \\ &= \alpha\Big(\left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}\Big) \\ &= \alpha f(\mathbf{x}, \mathbf{y}), \end{aligned}$$

where (*) follows from the linearity of $\left[\ \cdot\ \right]_{\mathcal{B}}$. Thus, f satisfies b.2, and similarly, it satisfies b.4. This proves that f is indeed bilinear.

Proof (continued). Next, to prove (a.1), we fix indices $i, j \in \{1, \dots, n\}$, and we compute:

$$f(\mathbf{b}_i, \mathbf{b}_j) \quad = \quad \left[\ \mathbf{b}_i\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{b}_j\ \right]_{\mathcal{B}} \quad = \quad \mathbf{e}_i^T A \mathbf{e}_j \quad \stackrel{(*)}{=} \quad a_{i,j},$$

where (*) follows from Proposition 9.1.1(b).

Proof (continued). Next, to prove (a.1), we fix indices $i, j \in \{1, \dots, n\}$, and we compute:

$$f(\mathbf{b}_i, \mathbf{b}_j) \;=\; \left[\ \mathbf{b}_i\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{b}_j\ \right]_{\mathcal{B}} \;=\; \mathbf{e}_i^T A \mathbf{e}_j \;\overset{(*)}{=}\; a_{i,j},$$

where (*) follows from Proposition 9.1.1(b).

For (a.2), we fix scalars $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$, and we compute:

$$f\Big(\sum_{i=1}^n c_i \mathbf{b}_i, \sum_{j=1}^n d_j \mathbf{b}_j\Big) \;\overset{(*)}{=}\; \sum_{i=1}^n \sum_{j=1}^n c_i d_j f(\mathbf{b}_i, \mathbf{b}_j) \;\overset{\text{(a.1)}}{=}\; \sum_{i=1}^n \sum_{j=1}^n a_{i,j} c_i d_j,$$

where (*) follows from the fact that f is bilinear.

Proof (continued). It remains to prove (a.3).

Proof (continued). It remains to prove (a.3). Suppose first that A is symmetric. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{array}{rcl} f(\mathbf{x},\mathbf{y}) & = & \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \stackrel{(*)}{=} \Big(\left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \Big)^{T} \\ & = & \left[\ \mathbf{y}\ \right]_{\mathcal{B}}^{T} A^{T} \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \stackrel{(**)}{=} \left[\ \mathbf{y}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} = f(\mathbf{y},\mathbf{x}), \end{array}$$

where (*) follows from the fact that $\left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ is a 1×1 matrix (and is therefore symmetric), and (**) follows from the fact that A is symmetric. So, f is symmetric.

Proof (continued). It remains to prove (a.3). Suppose first that A is symmetric. Then for all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{array}{rcl} f(\mathbf{x}, \mathbf{y}) & = & \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \stackrel{(*)}{=} \left(\left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \right)^{T} \\ & = & \left[\ \mathbf{y}\ \right]_{\mathcal{B}}^{T} A^{T} \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \stackrel{(**)}{=} \left[\ \mathbf{y}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} = f(\mathbf{y}, \mathbf{x}), \end{array}$$

where (*) follows from the fact that $\left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ is a 1×1 matrix (and is therefore symmetric), and (**) follows from the fact that A is symmetric. So, f is symmetric.

Suppose, conversely, that f is symmetric. Then for all indices $i, j \in \{1, \dots, n\}$, we have the following:

$$a_{i,j} \stackrel{(a.1)}{=} f(\mathbf{b}_i, \mathbf{b}_j) \stackrel{(*)}{=} f(\mathbf{b}_j, \mathbf{b}_i) \stackrel{(a.1)}{=} a_{j,i},$$

where (*) follows from the fact that f is symmetric. So, A is symmetric.

Proof (continued). (b) Fix a bilinear form f on V.

Proof (continued). (b) Fix a bilinear form f on V.

First of all, if $A = \left[\ a_{i,j}\ \right]_{n \times n}$ is any matrix in $\mathbb{F}^{n \times n}$ that satisfies the property that $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$, then (a) guarantees that $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \ldots, n\}$. This, in particular, proves the uniqueness part of (b).

Proof (continued). For existence, we must show that the matrix $A = \left[\ a_{i,j}\ \right]_{n\times n}$ given by the formula $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$, does indeed satisfy the property that $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$.

Proof (continued). For existence, we must show that the matrix $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n \times n}$ given by the formula $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$, does indeed satisfy the property that $f(\mathbf{x}, \mathbf{y}) = \left[\begin{array}{c} \mathbf{x} \end{array} \right]_{\mathcal{B}}^{T} A \left[\begin{array}{c} \mathbf{y} \end{array} \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$.

So, fix vectors $\mathbf{x}, \mathbf{y} \in V$. Since $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V, we know that there exist scalars $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$ s.t. $\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{b}_i$ and $\mathbf{y} = \sum_{j=1}^{n} d_j \mathbf{b}_j$, so that $\left[\begin{array}{c} \mathbf{x} \end{array} \right]_{\mathcal{B}} = \left[\begin{array}{ccc} c_1 & \dots & c_n \end{array} \right]^{T}$ and $\left[\begin{array}{c} \mathbf{y} \end{array} \right]_{\mathcal{B}} = \left[\begin{array}{ccc} d_1 & \dots & d_n \end{array} \right]^{T}$.

Proof (continued). For existence, we must show that the matrix $A = \left[\ a_{i,j} \ \right]_{n \times n}$ given by the formula $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$, does indeed satisfy the property that $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x} \ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y} \ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$.

So, fix vectors $\mathbf{x}, \mathbf{y} \in V$. Since $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V, we know that there exist scalars $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$ s.t. $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{b}_i$ and $\mathbf{y} = \sum_{j=1}^n d_j \mathbf{b}_j$, so that $\left[\ \mathbf{x} \ \right]_{\mathcal{B}} = \left[\ c_1 \quad \dots \quad c_n \ \right]^T$ and $\left[\ \mathbf{y} \ \right]_{\mathcal{B}} = \left[\ d_1 \quad \dots \quad d_n \ \right]^T$. We then compute:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f\Big(\sum_{i=1}^n c_i \mathbf{b}_i, \sum_{j=1}^n d_j \mathbf{b}_j \Big) \stackrel{(*)}{=} \sum_{i=1}^n \sum_{j=1}^n c_i d_j \underbrace{f(\mathbf{b}_i, \mathbf{b}_j)}_{=a_{i,j}} \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{i,j} c_i d_j \stackrel{(**)}{=} \left[\ \mathbf{x} \ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y} \ \right]_{\mathcal{B}}, \end{aligned}$$

where (*) follows from the fact that f is bilinear, and (**) follows from Proposition 9.1.1(a). $\square$

Theorem 9.2.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, and let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V.

(a) For every matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ in $\mathbb{F}^{n \times n}$, the function $f : V \times V \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \begin{bmatrix} \mathbf{x} \end{bmatrix}_{\mathcal{B}}^T A \begin{bmatrix} \mathbf{y} \end{bmatrix}_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V$$

is a bilinear form on V, and moreover, all the following hold:

- (a.1) $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$,
- (a.2) $f\left(\sum_{i=1}^n c_i \mathbf{b}_i, \sum_{j=1}^n d_j \mathbf{b}_j\right) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} c_i d_j$ for all $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{F}$,
- (a.3) f is symmetric iff A is symmetric.

(b) For every bilinear form f on V, there exists a unique matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ in $\mathbb{F}^{n \times n}$, called the *matrix of the bilinear form f with respect to the basis $\mathcal{B}$*, that satisfies the property that

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \begin{bmatrix} \mathbf{x} \end{bmatrix}_{\mathcal{B}}^T A \begin{bmatrix} \mathbf{y} \end{bmatrix}_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Moreover, the entries of the matrix A are given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$.

- As a special case of Theorem 9.2.2 for the special case of $V = \mathbb{F}^n$ (where $\mathbb{F}$ is a field), and $\mathcal{B} = \mathcal{E}_n$ (the standard basis of $\mathbb{F}^n$), we get the following corollary (next slide).

Corollary 9.2.3

Let $\mathbb{F}$ be a field, and let $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{F}^n$.

(a) For every matrix $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times n}$ in $\mathbb{F}^{n\times n}$, the function $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by
$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \mathbf{x}^T A\mathbf{y} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{F}^n$$
is a bilinear form on $\mathbb{F}^n$, and moreover, all the following hold:

(a.1) $f(\mathbf{e}_i, \mathbf{e}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$,

(a.2) $f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j$ for all vectors $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \dots & x_n \end{array}\right]^T$ and $\mathbf{y} = \left[\begin{array}{ccc} y_1 & \dots & y_n \end{array}\right]^T$ in $\mathbb{F}^n$,

(a.3) f is symmetric iff A is symmetric.

(b) For every bilinear form f on $\mathbb{F}^n$, there exists a unique matrix $A = \left[\begin{array}{c} a_{i,j} \end{array}\right]_{n\times n}$ in $\mathbb{F}^{n\times n}$ that satisfies the property that
$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \mathbf{x}^T A\mathbf{y} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{F}^n.$$
Moreover, the entries of the matrix A are given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all indices $i, j \in \{1, \dots, n\}$.

- **Remark:** Corollary 9.2.3 implies that, for a field $\mathbb{F}$, the bilinear forms on $\mathbb{F}^n$ are precisely the functions $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ in } \mathbb{F}^n,$$

where the $a_{i,j}$'s are some scalars in $\mathbb{F}$.

- **Remark:** Corollary 9.2.3 implies that, for a field $\mathbb{F}$, the bilinear forms on $\mathbb{F}^n$ are precisely the functions $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ in } \mathbb{F}^n,$$

where the $a_{i,j}$'s are some scalars in $\mathbb{F}$.

  - Moreover, such a bilinear form is symmetric iff $a_{i,j} = a_{j,i}$ for all indices $i, j \in \{1, \dots, n\}$.

- **Remark:** Corollary 9.2.3 implies that, for a field $\mathbb{F}$, the bilinear forms on $\mathbb{F}^n$ are precisely the functions $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \text{ in } \mathbb{F}^n,$$

where the $a_{i,j}$'s are some scalars in $\mathbb{F}$.
  - Moreover, such a bilinear form is symmetric iff $a_{i,j} = a_{j,i}$ for all indices $i, j \in \{1, \dots, n\}$.
  - The matrix of this bilinear form with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{F}^n$ is $\begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ (so, the i, j-th entry of the matrix is the coefficient in front of $x_i y_j$ from the formula for f above).

- For example, functions $f_1, f_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the formulas
    - $f_1(\mathbf{x}, \mathbf{y}) = x_1y_1 - 3x_1y_2 - 3x_2y_1 + 7x_2y_2$,
    - $f_2(\mathbf{x}, \mathbf{y}) = x_1y_1 - 2x_1y_2 + 3x_2y_1 - 3x_2y_2$,

  for all $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ in $\mathbb{R}^2$, are bilinear forms on $\mathbb{R}^2$.

- For example, functions $f_1, f_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the formulas
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$,
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$,

  for all $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ in $\mathbb{R}^2$, are bilinear forms on $\mathbb{R}^2$.
  - The bilinear form f_1 is symmetric, whereas the bilinear form f_2 is not.

- For example, functions $f_1, f_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the formulas
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$,
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$,

  for all $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ in $\mathbb{R}^2$, are bilinear forms on $\mathbb{R}^2$.
  - The bilinear form f_1 is symmetric, whereas the bilinear form f_2 is not.
  - The matrices of the bilinear forms f_1 and f_2 with respect to the standard basis $\mathcal{E}_2$ of $\mathbb{R}^2$ are

$$A_1 = \begin{bmatrix} 1 & -3 \\ -3 & 7 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix},$$

    respectively.

- For example, functions $f_1, f_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the formulas
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$,
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$,

  for all $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ in $\mathbb{R}^2$, are bilinear forms on $\mathbb{R}^2$.
  - The bilinear form f_1 is symmetric, whereas the bilinear form f_2 is not.
  - The matrices of the bilinear forms f_1 and f_2 with respect to the standard basis $\mathcal{E}_2$ of $\mathbb{R}^2$ are

$$A_1 = \begin{bmatrix} 1 & -3 \\ -3 & 7 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix},$$

    respectively.
  - Note that A_1 is symmetric, whereas A_2 is not; this is consistent with the fact that f_1 is symmetric, whereas f_2 is not.

- Reminder:

Theorem 4.3.2

Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be a basis of U, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$. Then there exists a unique linear function $f : U \to V$ s.t. $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_n) = \mathbf{v}_n$. Moreover, if the vector space U is non-trivial (i.e. $n \neq 0$), then this unique linear function $f : U \to V$ satisfies the following: for all $\mathbf{u} \in U$, we have that

$$f(\mathbf{u}) \quad = \quad \alpha_1 \mathbf{v}_1 + \dots + \alpha_n \mathbf{v}_n,$$

where $\left[\ \mathbf{u}\ \right]_{\mathcal{B}} = \left[\begin{array}{ccc} \alpha_1 & \dots & \alpha_n \end{array}\right]^T$. On the other hand, if U is trivial (i.e. $U = \{\mathbf{0}\}$), then $f : U \to V$ is given by $f(\mathbf{0}) = \mathbf{0}$.

- Reminder:

Theorem 4.3.2

Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be a basis of U, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$. Then there exists a unique linear function $f : U \to V$ s.t. $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_n) = \mathbf{v}_n$. Moreover, if the vector space U is non-trivial (i.e. $n \neq 0$), then this unique linear function $f : U \to V$ satisfies the following: for all $\mathbf{u} \in U$, we have that

$$f(\mathbf{u}) \;=\; \alpha_1\mathbf{v}_1 + \dots + \alpha_n\mathbf{v}_n,$$

where $\left[\, \mathbf{u} \,\right]_{\mathcal{B}} = \left[\begin{array}{ccc} \alpha_1 & \dots & \alpha_n \end{array}\right]^T$. On the other hand, if U is trivial (i.e. $U = \{\mathbf{0}\}$), then $f : U \to V$ is given by $f(\mathbf{0}) = \mathbf{0}$.

- Theorem 4.3.2 essentially states that a linear function can be fully determined by specifying what the vectors of some basis of the domain get mapped to.

- Reminder:

Theorem 4.3.2

Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Let $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be a basis of U, and let $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$. Then there exists a unique linear function $f : U \to V$ s.t. $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_n) = \mathbf{v}_n$. Moreover, if the vector space U is non-trivial (i.e. $n \neq 0$), then this unique linear function $f : U \to V$ satisfies the following: for all $\mathbf{u} \in U$, we have that

$$f(\mathbf{u}) \;=\; \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n,$$

where $\left[\ \mathbf{u}\ \right]_{\mathcal{B}} = \left[\ \alpha_1 \ \ \dots \ \ \alpha_n\ \right]^T$. On the other hand, if U is trivial (i.e. $U = \{\mathbf{0}\}$), then $f : U \to V$ is given by $f(\mathbf{0}) = \mathbf{0}$.

- Theorem 4.3.2 essentially states that a linear function can be fully determined by specifying what the vectors of some basis of the domain get mapped to.
- For bilinear forms, Theorem 9.2.2 yields the following analogous result.

Corollary 9.2.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let $A = \left[\ a_{i,j}\ \right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$. Then there exists a unique bilinear form f on V that satisfies the property that $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all indices $i, j \in \{1, \dots, n\}$. Moreover, the matrix of this bilinear form with respect to the basis $\mathcal{B}$ is precisely the matrix A.

Proof.

Corollary 9.2.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let $A = \left[\begin{array}{c} a_{i,j} \end{array} \right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$. Then there exists a unique bilinear form f on V that satisfies the property that $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all indices $i, j \in \{1, \dots, n\}$. Moreover, the matrix of this bilinear form with respect to the basis $\mathcal{B}$ is precisely the matrix A.

Proof. **Existence.** By Theorem 9.2.2(a), the function $f : V \times V \to \mathbb{F}$ given by

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\begin{array}{c} \mathbf{x} \end{array} \right]_{\mathcal{B}}^T A \left[\begin{array}{c} \mathbf{y} \end{array} \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V$$

is bilinear, and moreover, part (a.1) of Theorem 9.2.2(a) guarantees that $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all indices $i, j \in \{1, \dots, n\}$. Clearly, A is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$.

Corollary 9.2.4

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let $A = \left[\ a_{i,j}\ \right]_{n \times n}$ be a matrix in $\mathbb{F}^{n \times n}$. Then there exists a unique bilinear form f on V that satisfies the property that $f(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all indices $i, j \in \{1, \dots, n\}$. Moreover, the matrix of this bilinear form with respect to the basis $\mathcal{B}$ is precisely the matrix A.

Proof (continued). **Uniqueness.** Suppose that f' is any bilinear form on V that satisfies $f'(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$. Then Theorem 9.2.2(b) guarantees that the matrix of the bilinear form f' with respect to the basis $\mathcal{B}$ is precisely the matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$, i.e. $f'(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$. $\square$

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof.

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T \; B \; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof. For all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &\overset{(*)}{=} \big[\ \mathbf{x}\ \big]_{\mathcal{B}}^T \; B \; \big[\ \mathbf{y}\ \big]_{\mathcal{B}} \\ &= \Big({}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}} \; \big[\ \mathbf{x}\ \big]_{\mathcal{C}} \Big)^T B \Big({}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}} \; \big[\ \mathbf{y}\ \big]_{\mathcal{C}} \Big) \\ &= \big[\ \mathbf{x}\ \big]_{\mathcal{C}}^T \Big({}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T \; B \; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}} \Big) \big[\ \mathbf{y}\ \big]_{\mathcal{C}}, \end{aligned}$$

where (*) follows from the fact that B is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$.

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof (continued). Reminder:

$$f(\mathbf{x},\mathbf{y}) \;=\; \big[\ \mathbf{x}\ \big]_{\mathcal{C}}^{T}\Big({}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}\Big)\big[\ \mathbf{y}\ \big]_{\mathcal{C}}.$$

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof (continued). Reminder:

$$f(\mathbf{x},\mathbf{y}) \;=\; \big[\ \mathbf{x}\ \big]_{\mathcal{C}}^T \Big({}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}} \Big) \big[\ \mathbf{y}\ \big]_{\mathcal{C}}.$$

But now we have that

$${}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}$$

is the matrix of the bilinear form f with respect to the basis $\mathcal{C}$, that is, $C = {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^T\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}$. $\square$

Definition

Let $\mathbb{F}$ be a field. A matrix $A \in \mathbb{F}^{n \times n}$ is said to be *congruent* to a matrix $B \in \mathbb{F}^{n \times n}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^T AP$.

Definition

Let $\mathbb{F}$ be a field. A matrix $A \in \mathbb{F}^{n \times n}$ is said to be *congruent* to a matrix $B \in \mathbb{F}^{n \times n}$ if there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $B = P^T AP$.

- Like matrix similarity (see Proposition 4.5.13), matrix congruence is an equivalence relation on $\mathbb{F}^{n \times n}$.

Proposition 9.2.6

Let $\mathbb{F}$ be a field. Then all the following hold:

(a) for all matrices $A \in \mathbb{F}^{n \times n}$, A is congruent to A;

(b) for all matrices $A, B \in \mathbb{F}^{n \times n}$, if A is congruent to B, then B is congruent to A;

(c) for all matrices $A, B, C \in \mathbb{F}^{n \times n}$, if A is congruent to B and B is congruent to C, then A is congruent to C.

- Proof: Lecture Notes (easy).

Definition

Let $\mathbb{F}$ be a field. A matrix $A \in \mathbb{F}^{n\times n}$ is said to be *congruent* to a matrix $B \in \mathbb{F}^{n\times n}$ if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^T AP$.

- Theorem 4.5.16 essentially states that two square matrices are similar iff they represent the same linear function, but possibly with respect to different bases.

Definition

Let $\mathbb{F}$ be a field. A matrix $A \in \mathbb{F}^{n\times n}$ is said to be *congruent* to a matrix $B \in \mathbb{F}^{n\times n}$ if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $B = P^T AP$.

- Theorem 4.5.16 essentially states that two square matrices are similar iff they represent the same linear function, but possibly with respect to different bases.
- Theorem 9.2.7 (next slide) is an analog of Theorem 4.5.16 for congruent matrices: it states that two square matrices are congruent iff they represent the same bilinear form, but possibly with respect to different bases.

Theorem 9.2.7

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof.

Theorem 9.2.7

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof. We will prove the implications
"(a) $\implies$ (b) $\implies$ (c) $\implies$ (a)."

Theorem 9.2.7

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

Proof (continued). We first assume (a) and prove (b).

Theorem 9.2.7

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

Proof (continued). We first assume (a) and prove (b). By (a), there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $C = P^T BP$.

Theorem 9.2.7

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

Proof (continued). We first assume (a) and prove (b). By (a), there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $C = P^T BP$. Now, to prove (b), we fix a basis $\mathcal{B}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$.

Theorem 9.2.7

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

Proof (continued). We first assume (a) and prove (b). By (a), there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ s.t. $C = P^T BP$. Now, to prove (b), we fix a basis $\mathcal{B}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$.

Since P is invertible, Proposition 4.5.12 guarantees that there exists a basis $\mathcal{C}$ of V s.t. $P = {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}$.

Theorem 9.2.7

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

Proof (continued). We first assume (a) and prove (b). By (a), there exists an invertible matrix $P \in \mathbb{F}^{n \times n}$ s.t. $C = P^T BP$. Now, to prove (b), we fix a basis $\mathcal{B}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$.

Since P is invertible, Proposition 4.5.12 guarantees that there exists a basis $\mathcal{C}$ of V s.t. $P = {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}$. But then Theorem 9.2.5 guarantees that the matrix of the bilinear form f with respect to the basis $\mathcal{C}$ is precisely the matrix

$${}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}^T \ B \ {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}} \quad = \quad P^T BP \quad = \quad C.$$

This proves (b).

Theorem 9.2.7

- (b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;
- (c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Next, we assume (b) and prove (c).

Theorem 9.2.7

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Next, we assume (b) and prove (c). Fix any basis $\mathcal{B}$ of V, and define $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$.

Theorem 9.2.7

- (b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;
- (c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Next, we assume (b) and prove (c). Fix any basis $\mathcal{B}$ of V, and define $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T B \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$. By Theorem 9.2.2, f is a bilinear form on V, and obviously, B is the matrix of f with respect to the basis $\mathcal{B}$.

Theorem 9.2.7

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Next, we assume (b) and prove (c). Fix any basis $\mathcal{B}$ of V, and define $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$. By Theorem 9.2.2, f is a bilinear form on V, and obviously, B is the matrix of f with respect to the basis $\mathcal{B}$.

Using (b), we now fix a basis $\mathcal{C}$ of V s.t. C is the matrix of the bilinear form f with respect to $\mathcal{C}$.

Theorem 9.2.7

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Next, we assume (b) and prove (c). Fix any basis $\mathcal{B}$ of V, and define $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} B \left[\ \mathbf{y}\ \right]_{\mathcal{B}}$ for all $\mathbf{x}, \mathbf{y} \in V$. By Theorem 9.2.2, f is a bilinear form on V, and obviously, B is the matrix of f with respect to the basis $\mathcal{B}$.

Using (b), we now fix a basis $\mathcal{C}$ of V s.t. C is the matrix of the bilinear form f with respect to $\mathcal{C}$. We have now constructed bases $\mathcal{B}$ and $\mathcal{C}$ of V, and a bilinear form f on V, s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$. This proves (c).

Theorem 9.2.7

(a) B and C are congruent;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Finally, we assume (c) and prove (a).

Theorem 9.2.7

(a) B and C are congruent;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Finally, we assume (c) and prove (a).
Using (c), we fix bases $\mathcal{B}$ and $\mathcal{C}$ and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Theorem 9.2.7

(a) B and C are congruent;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Finally, we assume (c) and prove (a).
Using (c), we fix bases $\mathcal{B}$ and $\mathcal{C}$ and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Set $P := {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}$.

Theorem 9.2.7

- (a) B and C are congruent;
- (c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Proof (continued). Finally, we assume (c) and prove (a).
Using (c), we fix bases $\mathcal{B}$ and $\mathcal{C}$ and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

Set $P := {}_{\mathcal{B}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{C}}$. By Proposition 4.5.12, P is invertible, and by Theorem 9.2.5, we have that $C = P^T BP$. This proves (a). $\square$

Definition

The *characteristic* of a field $\mathbb{F}$ is the smallest positive integer n (if it exists) s.t. in the field $\mathbb{F}$, we have that

$$\underbrace{1 + \cdots + 1}_{n} \;=\; 0,$$

where the 1's and the 0 are understood to be in the field $\mathbb{F}$. If no such n exists, then $\text{char}(\mathbb{F}) := 0$.

- Fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic 0.
- On the other hand, for all prime numbers p, we have that $\text{char}(\mathbb{Z}_p) = p$.
- By Theorem 2.4.5, the characteristic of any field is either 0 or a prime number.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

- Proof: next slide.
- **Remark:** Proposition 9.2.8 applies to bilinear forms over vector spaces of characteristic other than 2.
  - In such fields, we can divide by $2 := 1 + 1$, since $2 = 1 + 1 \neq 0$.
  - The only field of characteristic 2 that we have seen is $\mathbb{Z}_2$, but other fields of characteristic 2 do exist.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. By hypothesis, all the following hold:

1. $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$;
2. $f(\mathbf{y}, \mathbf{y}) = g(\mathbf{y}, \mathbf{y})$;
3. $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. By hypothesis, all the following hold:

1. $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$;
2. $f(\mathbf{y}, \mathbf{y}) = g(\mathbf{y}, \mathbf{y})$;
3. $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$.

On the other hand, since f and g are bilinear, we have that

(4) $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y})$;

(5) $g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = g(\mathbf{x}, \mathbf{x}) + g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x}) + g(\mathbf{y}, \mathbf{y})$.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. By hypothesis, all the following hold:

1. $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$;
2. $f(\mathbf{y}, \mathbf{y}) = g(\mathbf{y}, \mathbf{y})$;
3. $f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = g(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y})$.

On the other hand, since f and g are bilinear, we have that

(4) $f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y})$;

(5) $g(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = g(\mathbf{x}, \mathbf{x}) + g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x}) + g(\mathbf{y}, \mathbf{y})$.

By (1)-(5), it follows that $f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) = g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})$.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. By hypothesis, all the following hold:

1. $f(\mathbf{x}, \mathbf{x}) = g(\mathbf{x}, \mathbf{x})$;
2. $f(\mathbf{y}, \mathbf{y}) = g(\mathbf{y}, \mathbf{y})$;
3. $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y})$.

On the other hand, since f and g are bilinear, we have that

(4) $f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y})$;

(5) $g(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = g(\mathbf{x}, \mathbf{x}) + g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x}) + g(\mathbf{y}, \mathbf{y})$.

By (1)-(5), it follows that $f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) = g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})$. But since f and g are symmetric, we further have that $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ and $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}, \mathbf{x})$, and it follows that $2f(\mathbf{x}, \mathbf{y}) = 2g(\mathbf{x}, \mathbf{y})$.

Proposition 9.2.8

Let f and g be **symmetric** bilinear forms on a vector space V over a field $\mathbb{F}$ of characteristic other than 2, and assume that for all $\mathbf{x} \in V$, we have that $f(\mathbf{x},\mathbf{x}) = g(\mathbf{x},\mathbf{x})$. Then $f = g$.

Proof. Fix $\mathbf{x},\mathbf{y} \in V$. We must show that $f(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$. By hypothesis, all the following hold:

1. $f(\mathbf{x},\mathbf{x}) = g(\mathbf{x},\mathbf{x})$;
2. $f(\mathbf{y},\mathbf{y}) = g(\mathbf{y},\mathbf{y})$;
3. $f(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = g(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})$.

On the other hand, since f and g are bilinear, we have that

(4) $f(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = f(\mathbf{x},\mathbf{x}) + f(\mathbf{x},\mathbf{y}) + f(\mathbf{y},\mathbf{x}) + f(\mathbf{y},\mathbf{y})$;

(5) $g(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) = g(\mathbf{x},\mathbf{x}) + g(\mathbf{x},\mathbf{y}) + g(\mathbf{y},\mathbf{x}) + g(\mathbf{y},\mathbf{y})$.

By (1)-(5), it follows that $f(\mathbf{x},\mathbf{y}) + f(\mathbf{y},\mathbf{x}) = g(\mathbf{x},\mathbf{y}) + g(\mathbf{y},\mathbf{x})$. But since f and g are symmetric, we further have that $f(\mathbf{x},\mathbf{y}) = f(\mathbf{y},\mathbf{x})$ and $g(\mathbf{x},\mathbf{y}) = g(\mathbf{y},\mathbf{x})$, and it follows that $2f(\mathbf{x},\mathbf{y}) = 2g(\mathbf{x},\mathbf{y})$. Since $\text{char}(\mathbb{F}) \neq 2$ (and consequently, $2 = 1+1 \neq 0$ in our field $\mathbb{F}$), we deduce that $f(\mathbf{x},\mathbf{y}) = g(\mathbf{x},\mathbf{y})$. $\square$

3. Quadratic forms

3. Quadratic forms

Definition

A *quadratic form* on a vector space V over a field $\mathbb{F}$ is any function $q : V \to \mathbb{F}$ for which there exists a bilinear form $f : V \times V \to \mathbb{F}$ s.t. $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in V$.

3. Quadratic forms

Definition

A *quadratic form* on a vector space V over a field $\mathbb{F}$ is any function $q : V \to \mathbb{F}$ for which there exists a bilinear form $f : V \times V \to \mathbb{F}$ s.t. $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in V$.

- Quadratic forms are defined for vector spaces over fields of any characteristic.

3. Quadratic forms

Definition

A *quadratic form* on a vector space V over a field $\mathbb{F}$ is any function $q : V \to \mathbb{F}$ for which there exists a bilinear form $f : V \times V \to \mathbb{F}$ s.t. $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in V$.

- Quadratic forms are defined for vector spaces over fields of any characteristic.
- However, in all our results that follow, we assume that the field in question is of characteristic other than 2, so that we can divide by 2.

Theorem 9.3.1

Let q be a quadratic form on a vector space V over a field $\mathbb{F}$ of characteristic other than 2. Then there exists a unique **symmetric** bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Furthermore, if the vector space V is non-trivial and finite-dimensional, then for any basis $\mathcal{B}$ of V, there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t.

$$q(\mathbf{x}) \;\; = \;\; \left[\; \mathbf{x} \;\right]_{\mathcal{B}}^{T} A \left[\; \mathbf{x} \;\right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V,$$

and moreover, this unique symmetric matrix A is precisely the matrix of the symmetric bilinear form f with respect to the basis $\mathcal{B}$.

Theorem 9.3.1

Let q be a quadratic form on a vector space V over a field $\mathbb{F}$ of characteristic other than 2. Then there exists a unique **symmetric** bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Furthermore, if the vector space V is non-trivial and finite-dimensional, then for any basis $\mathcal{B}$ of V, there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V,$$

and moreover, this unique symmetric matrix A is precisely the matrix of the symmetric bilinear form f with respect to the basis $\mathcal{B}$.

- **Terminology:** The symmetric matrix A from the statement of Theorem 9.3.1 is called the *matrix of the quadratic form q with respect to the basis $\mathcal{B}$*.
  - For emphasis, we may optionally refer to A as the ***symmetric*** *matrix of the quadratic form q with respect to the basis $\mathcal{B}$*.

Theorem 9.3.1

Let q be a quadratic form on a vector space V over a field $\mathbb{F}$ of characteristic other than 2. Then there exists a unique **symmetric** bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Furthermore, if the vector space V is non-trivial and finite-dimensional, then for any basis $\mathcal{B}$ of V, there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V,$$

and moreover, this unique symmetric matrix A is precisely the matrix of the symmetric bilinear form f with respect to the basis $\mathcal{B}$.

- **Warning:** There may possibly exist more than one matrix $A \in \mathbb{F}^{n \times n}$ that satisfies the property that $q(\mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
  - However, only one such matrix is symmetric.
  - This (unique) symmetric matrix is the one that we refer to as the matrix of q with respect to $\mathcal{B}$.

Theorem 9.3.1

Let q be a quadratic form on a vector space V over a field $\mathbb{F}$ of characteristic other than 2. Then there exists a unique **symmetric** bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Furthermore, if the vector space V is non-trivial and finite-dimensional, then for any basis $\mathcal{B}$ of V, there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V,$$

and moreover, this unique symmetric matrix A is precisely the matrix of the symmetric bilinear form f with respect to the basis $\mathcal{B}$.

- **Warning:** There may possibly exist more than one matrix $A \in \mathbb{F}^{n \times n}$ that satisfies the property that $q(\mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
  - However, only one such matrix is symmetric.
  - This (unique) symmetric matrix is the one that we refer to as the matrix of q with respect to $\mathcal{B}$.
- Now let's prove the theorem!

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f.

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$.

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \;\; = \;\; \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{y}) + h(\mathbf{y}, \mathbf{x})\Big) \qquad \text{for all } \mathbf{x} \in V.$$

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{y}) + h(\mathbf{y}, \mathbf{x})\Big) \qquad \text{for all } \mathbf{x} \in V.$$

It is then straightforward to check that f is a symmetric bilinear form on V (details?).

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \;\; = \;\; \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{y}) + h(\mathbf{y}, \mathbf{x})\Big) \qquad \text{for all } \mathbf{x} \in V.$$

It is then straightforward to check that f is a symmetric bilinear form on V (details?). Moreover, for all $\mathbf{x} \in V$, we have that

$$q(\mathbf{x}) \;\; = \;\; h(\mathbf{x}, \mathbf{x}) \;\; = \;\; \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{x}) + h(\mathbf{x}, \mathbf{x})\Big) \;\; = \;\; f(\mathbf{x}, \mathbf{x}),$$

which is what we needed.

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \;\; = \;\; \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{y}) + h(\mathbf{y}, \mathbf{x})\Big) \qquad \text{for all } \mathbf{x} \in V.$$

It is then straightforward to check that f is a symmetric bilinear form on V (details?). Moreover, for all $\mathbf{x} \in V$, we have that

$$q(\mathbf{x}) \;\; = \;\; h(\mathbf{x}, \mathbf{x}) \;\; = \;\; \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{x}) + h(\mathbf{x}, \mathbf{x})\Big) \;\; = \;\; f(\mathbf{x}, \mathbf{x}),$$

which is what we needed. This completes the proof of existence.

Proof. We first prove the existence and uniqueness of the symmetric bilinear form f. By the definition of a quadratic form, there exists some bilinear form h on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = h(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define $f : V \times V \to \mathbb{F}$ by setting

$$f(\mathbf{x}, \mathbf{y}) \quad = \quad \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{y}) + h(\mathbf{y}, \mathbf{x})\Big) \qquad \text{for all } \mathbf{x} \in V.$$

It is then straightforward to check that f is a symmetric bilinear form on V (details?). Moreover, for all $\mathbf{x} \in V$, we have that

$$q(\mathbf{x}) \quad = \quad h(\mathbf{x}, \mathbf{x}) \quad = \quad \tfrac{1}{2}\Big(h(\mathbf{x}, \mathbf{x}) + h(\mathbf{x}, \mathbf{x})\Big) \quad = \quad f(\mathbf{x}, \mathbf{x}),$$

which is what we needed. This completes the proof of existence. Uniqueness follows immediately from Proposition 9.2.8.

- Indeed, suppose that f_1 and f_2 are symmetric bilinear forms on V s.t. $q(\mathbf{x}) = f_1(\mathbf{x}, \mathbf{x})$ and $q(\mathbf{x}) = f_1(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in V$. Then $f_1(\mathbf{x}, \mathbf{x}) = f_2(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in V$. But then by Proposition 9.2.8, we have that $f_1 = f_2$.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
For uniqueness, suppose that $A' \in \mathbb{F}^{n \times n}$ is a symmetric matrix s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A' \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V.$$

WTS $A' = A$.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.

For uniqueness, suppose that $A' \in \mathbb{F}^{n \times n}$ is a symmetric matrix s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A' \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V.$$

WTS $A' = A$. Define $f' : V \times V \to \mathbb{F}$ by setting

$$f'(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^T A' \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n\times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x},\mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
For uniqueness, suppose that $A' \in \mathbb{F}^{n\times n}$ is a symmetric matrix s.t.

$$q(\mathbf{x}) \;=\; \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V.$$

WTS $A' = A$. Define $f' : V \times V \to \mathbb{F}$ by setting

$$f'(\mathbf{x},\mathbf{y}) \;=\; \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x},\mathbf{y} \in V.$$

By Theorem 9.2.2(a), f' is a symmetric bilinear form.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
For uniqueness, suppose that $A' \in \mathbb{F}^{n \times n}$ is a symmetric matrix s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V.$$

WTS $A' = A$. Define $f' : V \times V \to \mathbb{F}$ by setting

$$f'(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

By Theorem 9.2.2(a), f' is a symmetric bilinear form. But then for all $\mathbf{x} \in V$, we have that $f'(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$, and so by Proposition 9.2.8, $f' = f$.

Proof (continued). Let us now assume that the vector space V is non-trivial and finite-dimensional, and let $\mathcal{B}$ be a basis of V. Let $A \in \mathbb{F}^{n \times n}$ be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; by Theorem 9.2.2, the matrix A is symmetric. Obviously, $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}}$ for all $\mathbf{x} \in V$.
For uniqueness, suppose that $A' \in \mathbb{F}^{n \times n}$ is a symmetric matrix s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V.$$

WTS $A' = A$. Define $f' : V \times V \to \mathbb{F}$ by setting

$$f'(\mathbf{x}, \mathbf{y}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A' \left[\ \mathbf{y}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x}, \mathbf{y} \in V.$$

By Theorem 9.2.2(a), f' is a symmetric bilinear form. But then for all $\mathbf{x} \in V$, we have that $f'(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$, and so by Proposition 9.2.8, $f' = f$. The uniqueness part of Theorem 9.2.2(b) now guarantees that $A' = A$, and we are done. $\square$

Theorem 9.3.1

Let q be a quadratic form on a vector space V over a field $\mathbb{F}$ of characteristic other than 2. Then there exists a unique **symmetric** bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Furthermore, if the vector space V is non-trivial and finite-dimensional, then for any basis $\mathcal{B}$ of V, there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t.

$$q(\mathbf{x}) \quad = \quad \left[\ \mathbf{x}\ \right]_{\mathcal{B}}^{T} A \left[\ \mathbf{x}\ \right]_{\mathcal{B}} \qquad \text{for all } \mathbf{x} \in V,$$

and moreover, this unique symmetric matrix A is precisely the matrix of the symmetric bilinear form f with respect to the basis $\mathcal{B}$.

- **Remark:** Let $\mathbb{F}$ be a field. Then quadratic forms q on $\mathbb{F}^n$ are all of the form

$$q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}x_ix_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T \text{ in } \mathbb{F}^n,$$

where the $b_{i,j}$'s are some elements of $\mathbb{F}$.

- **Remark:** Let $\mathbb{F}$ be a field. Then quadratic forms q on $\mathbb{F}^n$ are all of the form

$$q(\mathbf{x}) \quad = \quad \sum_{i=1}^{n} \sum_{j=1}^{n} b_{i,j} x_i x_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T \text{ in } \mathbb{F}^n,$$

  where the $b_{i,j}$'s are some elements of $\mathbb{F}$.
  - If $\text{char}(\mathbb{F}) \neq 2$, then the matrix of such a quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{F}^n$ is the matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ whose entries are given by $a_{i,j} = \frac{1}{2}(b_{i,j} + b_{j,i})$ for all $i, j \in \{1, \dots, n\}$.

- **Remark:** Let $\mathbb{F}$ be a field. Then quadratic forms q on $\mathbb{F}^n$ are all of the form

$$q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}x_i x_j \qquad \text{for all } \mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T \text{ in } \mathbb{F}^n,$$

where the $b_{i,j}$'s are some elements of $\mathbb{F}$.
  - If $\text{char}(\mathbb{F}) \neq 2$, then the matrix of such a quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{F}^n$ is the matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ whose entries are given by $a_{i,j} = \frac{1}{2}(b_{i,j} + b_{j,i})$ for all $i, j \in \{1, \dots, n\}$.
  - Indeed, by construction, A is symmetric, and we see that for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have the following:

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &\overset{(*)}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \tfrac{1}{2}(b_{i,j} + b_{j,i})x_i x_j \\ &= \tfrac{1}{2}\Big(\big(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}x_i x_j\big) + \big(\sum_{i=1}^{n}\sum_{j=1}^{n} b_{j,i}x_i x_j\big)\Big) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} b_{i,j}x_i x_j = q(\mathbf{x}), \end{aligned}$$

where (*) follows from Proposition 9.1.1(a).

Example 9.3.2

Consider the quadratic form q on $\mathbb{R}^3$ given by

$$q(\mathbf{x}) \quad = \quad 3x_1^2 + 2x_1x_2 - 4x_1x_3 + 5x_2^2 - 6x_2x_3 + 2x_3^3$$

for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$ in $\mathbb{R}^3$. Then the matrix of q with respect to the standard basis $\mathcal{E}_3$ of $\mathbb{R}^3$ is the matrix

$$A \quad := \quad \begin{bmatrix} 3 & 1 & -2 \\ 1 & 5 & -3 \\ -2 & -3 & 2 \end{bmatrix}.$$

- Reminder:

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

- Reminder:

Theorem 9.2.5 [Change of basis for bilinear forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$, let f be a bilinear form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the matrix of f with respect to $\mathcal{B}$, and let C be the matrix of f with respect to $\mathcal{C}$. Then

$$C \quad = \quad {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}^T \ B \ {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}.$$

- Similarly, for quadratic forms, we have the following:

Corollary 9.3.3 [Change of basis for quadratic forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let q be a quadratic form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the (symmetric) matrix of q with respect to $\mathcal{B}$, and let C be the (symmetric) matrix of q with respect to $\mathcal{C}$. Then

$$C \quad = \quad {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}^T \ B \ {}_{\mathcal{B}}[\ \mathrm{Id}_V\ ]_{\mathcal{C}}.$$

Corollary 9.3.3 [Change of basis for quadratic forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let q be a quadratic form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the (symmetric) matrix of q with respect to $\mathcal{B}$, and let C be the (symmetric) matrix of q with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof.

Corollary 9.3.3 [Change of basis for quadratic forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let q be a quadratic form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the (symmetric) matrix of q with respect to $\mathcal{B}$, and let C be the (symmetric) matrix of q with respect to $\mathcal{C}$. Then

$$C \quad = \quad {}_{\mathcal{B}}\Big[\ \mathrm{Id}_V\ \Big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\Big[\ \mathrm{Id}_V\ \Big]_{\mathcal{C}}.$$

Proof. By Theorem 9.3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Corollary 9.3.3 [Change of basis for quadratic forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let q be a quadratic form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the (symmetric) matrix of q with respect to $\mathcal{B}$, and let C be the (symmetric) matrix of q with respect to $\mathcal{C}$. Then

$$C \;=\; {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\big[\ \mathrm{Id}_V\ \big]_{\mathcal{C}}.$$

Proof. By Theorem 9.3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Theorem 9.3.1 further guarantees that B (resp. C) is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$ (resp. $\mathcal{C}$) of $\mathbb{F}^n$.

Corollary 9.3.3 [Change of basis for quadratic forms]

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let q be a quadratic form on V, and let $\mathcal{B}$ and $\mathcal{C}$ be bases of V. Further, let B be the (symmetric) matrix of q with respect to $\mathcal{B}$, and let C be the (symmetric) matrix of q with respect to $\mathcal{C}$. Then

$$C \quad = \quad {}_{\mathcal{B}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{C}}^{T}\ B\ {}_{\mathcal{B}}\left[\ \mathrm{Id}_V\ \right]_{\mathcal{C}}.$$

Proof. By Theorem 9.3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Theorem 9.3.1 further guarantees that B (resp. C) is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$ (resp. $\mathcal{C}$) of $\mathbb{F}^n$. The result now follows immediately from Theorem 9.2.5. $\square$

- Reminder:

Theorem 9.2.7

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n \times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

- Reminder:

Theorem 9.2.7

Let $\mathbb{F}$ be a field, let $B, C \in \mathbb{F}^{n\times n}$ be matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and bilinear forms f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of f with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a bilinear form f on V s.t. B is the matrix of f with respect to $\mathcal{B}$, and C is the matrix of f with respect to $\mathcal{C}$.

- In the case of **symmetric** matrices, we get a similar result, only involving **quadratic** forms, rather than bilinear forms (next slide).

Theorem 9.3.4

Let $\mathbb{F}$ be a field of characteristic other than 2, let $B, C \in \mathbb{F}^{n \times n}$ be **symmetric** matrices, and let V be an n-dimensional vector space over the field $\mathbb{F}$. Then the following are equivalent:

(a) B and C are congruent;

(b) for all bases $\mathcal{B}$ of V and quadratic forms q on V s.t. B is the matrix of q with respect to $\mathcal{B}$, there exists a basis $\mathcal{C}$ of V s.t. C is the matrix of q with respect to $\mathcal{C}$;

(c) there exist bases $\mathcal{B}$ and $\mathcal{C}$ of V and a quadratic form q on V s.t. B is the matrix of q with respect to $\mathcal{B}$, and C is the matrix of q with respect to $\mathcal{C}$.

- Proof: Lecture Notes

4. Quadratic forms on $\mathbb{R}^n$

4. Quadratic forms on $\mathbb{R}^n$

- In what follows, orthogonality and orthonormality in $\mathbb{R}^n$ are assumed to be with respect to the standard scalar product $\cdot$ and the induced norm $||\cdot||$.

4. Quadratic forms on $\mathbb{R}^n$

- In what follows, orthogonality and orthonormality in $\mathbb{R}^n$ are assumed to be with respect to the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- By Corollary 8.7.4, any symmetric matrix in $\mathbb{R}^{n\times n}$ has n real eigenvalues (when algebraic multiplicities are taken into account).

4. Quadratic forms on $\mathbb{R}^n$

- In what follows, orthogonality and orthonormality in $\mathbb{R}^n$ are assumed to be with respect to the standard scalar product $\cdot$ and the induced norm $||\cdot||$.
- By Corollary 8.7.4, any symmetric matrix in $\mathbb{R}^{n\times n}$ has n real eigenvalues (when algebraic multiplicities are taken into account).
  - With this in mind, we define the following (next slide).

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n \times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n \times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

- Note that 0 is an eigenvalue of A iff $n_0 > 0$, and in this case, the algebraic multiplicity of the eigenvalue 0 is precisely n_0.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n \times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

- Note that 0 is an eigenvalue of A iff $n_0 > 0$, and in this case, the algebraic multiplicity of the eigenvalue 0 is precisely n_0.
- For example, if the spectrum of a symmetric matrix in $\mathbb{R}^{9 \times 9}$ is $\{0, 0, 1, 1, -2, -2, 5, 6, -7\}$, then the signature of that matrix is $(4, 3, 2)$.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n\times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n \times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

- Our goal is to prove the following theorem.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n \times n}$ are congruent iff they have the same signature.

Definition

The *signature* of a symmetric matrix $A \in \mathbb{R}^{n\times n}$ to be the ordered triple (n_+, n_-, n_0), where

- n_+ is the number of positive eigenvalues of A (counting algebraic multiplicities),
- n_- is the number of negative eigenvalues of A (counting algebraic multiplicities),
- $n_0 := n - n_+ - n_-$.

- Our goal is to prove the following theorem.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

- We begin with a proposition, which we will use to prove Theorem 9.4.3

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T AR \quad = \quad D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Proof.

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Proof. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable.

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n \times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n \times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \quad = \quad D(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}).$$

Proof. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable. So, let $D = D(\lambda_1, \ldots, \lambda_n)$ be a diagonal and Q an orthogonal matrix, both in $\mathbb{R}^{n \times n}$, s.t. $D = Q^T A Q$.

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}).$$

Proof. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable. So, let $D = D(\lambda_1, \ldots, \lambda_n)$ be a diagonal and Q an orthogonal matrix, both in $\mathbb{R}^{n\times n}$, s.t. $D = Q^T A Q$. By Proposition 8.5.12, $\{\lambda_1, \ldots, \lambda_n\}$ is the spectrum of A.

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Proof. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable. So, let $D = D(\lambda_1, \ldots, \lambda_n)$ be a diagonal and Q an orthogonal matrix, both in $\mathbb{R}^{n\times n}$, s.t. $D = Q^T A Q$. By Proposition 8.5.12, $\{\lambda_1, \ldots, \lambda_n\}$ is the spectrum of A.

After possibly permuting the λ_i's and the corresponding columns of the orthogonal matrix Q, we may assume that the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0 (justification: Lecture Notes).

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \dots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \ldots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Now, set

$$\ell_i \quad := \quad \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \text{if} \quad \lambda_i \neq 0 \\ 1 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \ldots, n\}$, and set $L := D(\ell_1, \ldots, \ell_n)$ and $R := QL$.

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \dots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Now, set

$$\ell_i \quad := \quad \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \text{if} \quad \lambda_i \neq 0 \\ 1 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \dots, n\}$, and set $L := D(\ell_1, \dots, \ell_n)$ and $R := QL$. Since both Q and L are invertible, so is R.

- Since Q is orthogonal, Theorem 6.8.1 guarantees that it is invertible. On the other hand, L is a diagonal matrix, and all its entries on the main diagonal are non-zero; so, by Proposition 8.5.3(b), L is invertible.

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \dots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Now, set

$$\ell_i \quad := \quad \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \text{if} \quad \lambda_i \neq 0 \\ 1 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \dots, n\}$, and set $L := D(\ell_1, \dots, \ell_n)$ and $R := QL$. Since both Q and L are invertible, so is R.

- Since Q is orthogonal, Theorem 6.8.1 guarantees that it is invertible. On the other hand, L is a diagonal matrix, and all its entries on the main diagonal are non-zero; so, by Proposition 8.5.3(b), L is invertible.

Moreover, since L is diagonal, Proposition 8.5.1(b) guarantees that the columns of $R = QL$ are scalar multiples of the columns of Q;

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \dots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Now, set

$$\ell_i \quad := \quad \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \text{if} \quad \lambda_i \neq 0 \\ 1 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \dots, n\}$, and set $L := D(\ell_1, \dots, \ell_n)$ and $R := QL$. Since both Q and L are invertible, so is R.

- Since Q is orthogonal, Theorem 6.8.1 guarantees that it is invertible. On the other hand, L is a diagonal matrix, and all its entries on the main diagonal are non-zero; so, by Proposition 8.5.3(b), L is invertible.

Moreover, since L is diagonal, Proposition 8.5.1(b) guarantees that the columns of $R = QL$ are scalar multiples of the columns of Q; since the columns of Q are pairwise orthogonal (by Theorem 6.8.1), Proposition 6.1.4(b) guarantees that the columns of R are pairwise orthogonal.

Proof (continued). Reminder: $D = Q^T AQ$, $D = D(\lambda_1, \dots, \lambda_n)$, Q is orthogonal; the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are 0.

Now, set

$$\ell_i \quad := \quad \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \text{if} \quad \lambda_i \neq 0 \\ 1 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \dots, n\}$, and set $L := D(\ell_1, \dots, \ell_n)$ and $R := QL$. Since both Q and L are invertible, so is R.

- Since Q is orthogonal, Theorem 6.8.1 guarantees that it is invertible. On the other hand, L is a diagonal matrix, and all its entries on the main diagonal are non-zero; so, by Proposition 8.5.3(b), L is invertible.

Moreover, since L is diagonal, Proposition 8.5.1(b) guarantees that the columns of $R = QL$ are scalar multiples of the columns of Q; since the columns of Q are pairwise orthogonal (by Theorem 6.8.1), Proposition 6.1.4(b) guarantees that the columns of R are pairwise orthogonal. Finally, we compute (next slide):

Proof (continued).

$$\begin{aligned}
R^T A R \quad &= \quad (QL)^T A (QL) \;=\; L^T \underbrace{Q^T A Q}_{=D} L \;\stackrel{(*)}{=}\; LDL \\
&= \quad D(\ell_1, \ldots, \ell_n)\, D(\lambda_1, \ldots, \lambda_n)\, D(\ell_1, \ldots, \ell_n) \\
&\stackrel{(**)}{=} \quad D(\lambda_1 \ell_1^2, \ldots, \lambda_n \ell_n^2), \\
&\stackrel{(***)}{=} \quad D\big(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}\big),
\end{aligned}$$

where (*) follows from the fact that L is diagonal and therefore symmetric, (**) follows from Proposition 8.5.2, and (***) follows from the fact that, by construction,

$$\lambda_i \ell_i^2 \quad = \quad \begin{cases} \;\;\,1 & \text{if} \quad \lambda_i > 0 \\ -1 & \text{if} \quad \lambda_i < 0 \\ \;\;\,0 & \text{if} \quad \lambda_i = 0 \end{cases}$$

for all indices $i \in \{1, \ldots, n\}$, plus the fact that the first n_+ many λ_i's are positive, the subsequent n_- many λ_i's are negative, and the final n_0 many λ_i's are zero. □

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

- The proof of Proposition 9.4.1 is fully constructive (i.e. it allows us to construct a suitable matrix R, as long as we are able to factor the characteristic polynomial of A).

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

- The proof of Proposition 9.4.1 is fully constructive (i.e. it allows us to construct a suitable matrix R, as long as we are able to factor the characteristic polynomial of A).
- For a numerical example, see the Lecture Notes.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n \times n}$ are congruent iff they have the same signature.

Proof. Fix symmetric matrices $B, C \in \mathbb{R}^{n \times n}$, and suppose first that B and C both have the same signature, say (n_+, n_-, n_0).

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof. Fix symmetric matrices $B, C \in \mathbb{R}^{n\times n}$, and suppose first that B and C both have the same signature, say (n_+, n_-, n_0). Proposition 9.4.1 then guarantees B and C are both congruent to the diagonal matrix

$$D \;:=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof. Fix symmetric matrices $B, C \in \mathbb{R}^{n\times n}$, and suppose first that B and C both have the same signature, say (n_+, n_-, n_0). Proposition 9.4.1 then guarantees B and C are both congruent to the diagonal matrix

$$D \;:=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

By Proposition 9.2.6, matrix congruence is an equivalence relation on $\mathbb{R}^{n\times n}$; so, since B and C are congruent to the same matrix D, they are also congruent to each other.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n \times n}$ are congruent iff they have the same signature.

Proof (continued). Suppose, conversely, that B and C are congruent.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Suppose, conversely, that B and C are congruent. Let $(p, q, n-p-q)$ be the signature of B, and let $(s, t, n-s-t)$ be the signature of C; WTS $(p, q, n-p-q) = (s, t, n-s-t)$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n \times n}$ are congruent iff they have the same signature.

Proof (continued). Suppose, conversely, that B and C are congruent. Let $(p, q, n-p-q)$ be the signature of B, and let $(s, t, n-s-t)$ be the signature of C; WTS $(p, q, n-p-q) = (s, t, n-s-t)$. Clearly, it suffices to show that $p = s$ and $p + q = s + t$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Suppose, conversely, that B and C are congruent. Let $(p, q, n-p-q)$ be the signature of B, and let $(s, t, n-s-t)$ be the signature of C; WTS $(p, q, n-p-q) = (s, t, n-s-t)$. Clearly, it suffices to show that $p = s$ and $p+q = s+t$.

First, by Proposition 9.4.1, B is congruent to the matrix

$$D_B \quad := \quad D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q}),$$

and C is congruent to the matrix

$$D_C \quad := \quad D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t}).$$

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Suppose, conversely, that B and C are congruent. Let $(p,q,n-p-q)$ be the signature of B, and let $(s,t,n-s-t)$ be the signature of C; WTS $(p,q,n-p-q)=(s,t,n-s-t)$. Clearly, it suffices to show that $p=s$ and $p+q=s+t$.

First, by Proposition 9.4.1, B is congruent to the matrix

$$D_B \;:=\; D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q}),$$

and C is congruent to the matrix

$$D_C \;:=\; D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t}).$$

Proposition 9.2.6 then guarantees that D_B and D_C are congruent to each other.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: Matrices

- $D_B := D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q})$
- $D_C := D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t})$

are congruent to each other; WTS $p = s$ and $p + q = s + t$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: Matrices

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$

are congruent to each other; WTS $p = s$ and $p + q = s + t$.

By definition, this means that there exists an invertible matrix $P \in \mathbb{R}^{n\times n}$ s.t. $D_C = P^T D_B P$; we will use this to prove that $p + q = r + s$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: Matrices

- $D_B := D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q})$
- $D_C := D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t})$

are congruent to each other; WTS $p = s$ and $p + q = s + t$.

By definition, this means that there exists an invertible matrix $P \in \mathbb{R}^{n\times n}$ s.t. $D_C = P^T D_B P$; we will use this to prove that $p + q = r + s$.

On the other hand, by Theorem 9.4.1, there exist bases $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ of $\mathbb{R}^n$, as well as a quadratic form $\widetilde{q}$ on $\mathbb{R}^n$, s.t. D_B is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B}$, and D_C is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C}$; we will use this to prove that $p = s$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: $D_C = P^T D_B P$, where

- $D_B := D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q})$,
- $D_C := D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t})$,
- P is invertible.

We first show that $p + q = s + t$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: $D_C = P^T D_B P$, where

- $D_B := D(\underbrace{1,\ldots,1}_{p}, \underbrace{-1,\ldots,-1}_{q}, \underbrace{0,\ldots,0}_{n-p-q})$,
- $D_C := D(\underbrace{1,\ldots,1}_{s}, \underbrace{-1,\ldots,-1}_{t}, \underbrace{0,\ldots,0}_{n-s-t})$,
- P is invertible.

We first show that $p+q = s+t$. Clearly, $\text{rank}(D_B) = p+q$ and $\text{rank}(D_C) = s+t$, and so it is enough to show that $\text{rank}(D_B) = \text{rank}(D_C)$.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: $D_C = P^T D_B P$, where

- $D_B := D(\underbrace{1,\dots,1}_{p}, \underbrace{-1,\dots,-1}_{q}, \underbrace{0,\dots,0}_{n-p-q})$,
- $D_C := D(\underbrace{1,\dots,1}_{s}, \underbrace{-1,\dots,-1}_{t}, \underbrace{0,\dots,0}_{n-s-t})$,
- P is invertible.

We first show that $p + q = s + t$. Clearly, $\text{rank}(D_B) = p + q$ and $\text{rank}(D_C) = s + t$, and so it is enough to show that $\text{rank}(D_B) = \text{rank}(D_C)$. Since the matrix P is invertible, the Invertible Matrix Theorem guarantees that P^T is also inverible.

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

Proof (continued). Reminder: $D_C = P^T D_B P$, where

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$,
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$,
- P is invertible.

We first show that $p+q = s+t$. Clearly, $\text{rank}(D_B) = p+q$ and $\text{rank}(D_C) = s+t$, and so it is enough to show that $\text{rank}(D_B) = \text{rank}(D_C)$. Since the matrix P is invertible, the Invertible Matrix Theorem guarantees that P^T is also inverible. But then

$$\text{rank}(D_C) \;=\; \text{rank}(P^T D_B P) \;\overset{(*)}{=}\; \text{rank}(D_B),$$

where (*) follows from Proposition 3.3.14 (since P^T and P are both invertible).

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\dots,\mathbf{c}_n\}$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\ldots,1}_{p}, \underbrace{-1,\ldots,-1}_{q}, \underbrace{0,\ldots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\ldots,1}_{s}, \underbrace{-1,\ldots,-1}_{t}, \underbrace{0,\ldots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$.

It remains to show that $p = s$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\ldots,\mathbf{c}_n\}$.

It remains to show that $p = s$. Suppose otherwise. By symmetry, we may assume that $p > s$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}, \underbrace{0, \ldots, 0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$,
- $D_C := D(\underbrace{1, \ldots, 1}_{s}, \underbrace{-1, \ldots, -1}_{t}, \underbrace{0, \ldots, 0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$.

It remains to show that $p = s$. Suppose otherwise. By symmetry, we may assume that $p > s$. Now consider the subspaces $U_B := \text{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \ldots, \mathbf{c}_n)$ of $\mathbb{R}^n$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\dots,\mathbf{c}_n\}$.

It remains to show that $p = s$. Suppose otherwise. By symmetry, we may assume that $p > s$. Now consider the subspaces $U_B := \mathrm{Span}(\mathbf{b}_1,\dots,\mathbf{b}_p)$ and $U_C := \mathrm{Span}(\mathbf{c}_{s+1},\dots,\mathbf{c}_n)$ of $\mathbb{R}^n$. Then by Theorem 3.2.23, we have that

$$\dim(U_B) + \dim(U_C) \;=\; \dim(U_B + U_C) + \dim(U_B \cap U_C).$$

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\dots,\mathbf{c}_n\}$.

It remains to show that $p = s$. Suppose otherwise. By symmetry, we may assume that $p > s$. Now consider the subspaces $U_B := \text{Span}(\mathbf{b}_1,\dots,\mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1},\dots,\mathbf{c}_n)$ of $\mathbb{R}^n$. Then by Theorem 3.2.23, we have that

$$\dim(U_B) + \dim(U_C) \quad = \quad \dim(U_B + U_C) + \dim(U_B \cap U_C).$$

But note that

- $\dim(U_B) + \dim(U_C) = p + (n - s) = n + (p - s) > n$,
- $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$,
- $D_C := D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$.

It remains to show that $p = s$. Suppose otherwise. By symmetry, we may assume that $p > s$. Now consider the subspaces $U_B := \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$ of $\mathbb{R}^n$. Then by Theorem 3.2.23, we have that

$$\dim(U_B) + \dim(U_C) \quad = \quad \dim(U_B + U_C) + \dim(U_B \cap U_C).$$

But note that

- $\dim(U_B) + \dim(U_C) = p + (n - s) = n + (p - s) > n$,
- $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$.

So, $\dim(U_B \cap U_C) > 0$, and it follows that $U_B \cap U_C$ contains some non-zero vector $\mathbf{u}$.

Proof (continued).

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\dots,1}_{p},\underbrace{-1,\dots,-1}_{q},\underbrace{0,\dots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\dots,1}_{s},\underbrace{-1,\dots,-1}_{t},\underbrace{0,\dots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\dots,\mathbf{c}_n\}$,
- $U_B := \mathrm{Span}(\mathbf{b}_1,\dots,\mathbf{b}_p)$, $U_C := \mathrm{Span}(\mathbf{c}_{s+1},\dots,\mathbf{c}_n)$,
- $\mathbf{u} \in U_B \cap U_C$, $\mathbf{u} \neq \mathbf{0}$.

Set $\left[\ \mathbf{u}\ \right]_{\mathcal{B}} = \left[\ x_1 \quad \dots \quad x_n\ \right]^T$ and $\left[\ \mathbf{u}\ \right]_{\mathcal{C}} = \left[\ y_1 \quad \dots \quad y_n\ \right]^T$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\ldots,1}_{p},\underbrace{-1,\ldots,-1}_{q},\underbrace{0,\ldots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\ldots,1}_{s},\underbrace{-1,\ldots,-1}_{t},\underbrace{0,\ldots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\ldots,\mathbf{c}_n\}$,
- $U_B := \mathrm{Span}(\mathbf{b}_1,\ldots,\mathbf{b}_p)$, $U_C := \mathrm{Span}(\mathbf{c}_{s+1},\ldots,\mathbf{c}_n)$,
- $\mathbf{u} \in U_B \cap U_C$, $\mathbf{u} \neq \mathbf{0}$.

Set $\begin{bmatrix} \mathbf{u} \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ and $\begin{bmatrix} \mathbf{u} \end{bmatrix}_{\mathcal{C}} = \begin{bmatrix} y_1 & \ldots & y_n \end{bmatrix}^T$. Then at least one of $x_1,\ldots,x_p$ is non-zero, $x_{p+1} = \cdots = x_n = 0$, and $y_1 = \cdots = y_s = 0$.

Proof (continued). Reminder:

- $D_B := D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$,
- $D_C := D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,
- $U_B := \mathrm{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$, $U_C := \mathrm{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$,
- $\mathbf{u} \in U_B \cap U_C$, $\mathbf{u} \neq \mathbf{0}$.

Set $\left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{B}} = \left[\begin{array}{ccc} x_1 & \dots & x_n \end{array}\right]^T$ and $\left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{C}} = \left[\begin{array}{ccc} y_1 & \dots & y_n \end{array}\right]^T$. Then at least one of $x_1, \dots, x_p$ is non-zero, $x_{p+1} = \dots = x_n = 0$, and $y_1 = \dots = y_s = 0$. We now have that

- $\widetilde{q}(\mathbf{u}) = \left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{B}}^T D_B \left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{B}} \stackrel{(*)}{=} x_1^2 + \dots + x_p^2 > 0$,
- $\widetilde{q}(\mathbf{u}) = \left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{C}}^T D_C \left[\begin{array}{c} \mathbf{u} \end{array}\right]_{\mathcal{C}} \stackrel{(*)}{=} -y_{s+1}^2 - \dots - y_{s+t}^2 \leq 0$,

where for both instances of (*), we used the formula from Proposition 9.1.1(a).

Proof (continued). Reminder:

- $D_B := D(\underbrace{1,\dots,1}_{p}, \underbrace{-1,\dots,-1}_{q}, \underbrace{0,\dots,0}_{n-p-q})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{B} = \{\mathbf{b}_1,\dots,\mathbf{b}_n\}$,
- $D_C := D(\underbrace{1,\dots,1}_{s}, \underbrace{-1,\dots,-1}_{t}, \underbrace{0,\dots,0}_{n-s-t})$ is the matrix of $\widetilde{q}$ w.r.t. $\mathcal{C} = \{\mathbf{c}_1,\dots,\mathbf{c}_n\}$,
- $U_B := \text{Span}(\mathbf{b}_1,\dots,\mathbf{b}_p)$, $U_C := \text{Span}(\mathbf{c}_{s+1},\dots,\mathbf{c}_n)$,
- $\mathbf{u} \in U_B \cap U_C$, $\mathbf{u} \neq \mathbf{0}$.

Set $\left[\ \mathbf{u}\ \right]_{\mathcal{B}} = \left[\ x_1 \quad \dots \quad x_n\ \right]^T$ and $\left[\ \mathbf{u}\ \right]_{\mathcal{C}} = \left[\ y_1 \quad \dots \quad y_n\ \right]^T$. Then at least one of $x_1,\dots,x_p$ is non-zero, $x_{p+1} = \dots = x_n = 0$, and $y_1 = \dots = y_s = 0$. We now have that

- $\widetilde{q}(\mathbf{u}) = \left[\ \mathbf{u}\ \right]_{\mathcal{B}}^T\ D_B\ \left[\ \mathbf{u}\ \right]_{\mathcal{B}} \overset{(*)}{=} x_1^2 + \dots + x_p^2 > 0$,
- $\widetilde{q}(\mathbf{u}) = \left[\ \mathbf{u}\ \right]_{\mathcal{C}}^T\ D_C\ \left[\ \mathbf{u}\ \right]_{\mathcal{C}} \overset{(*)}{=} -y_{s+1}^2 - \dots - y_{s+t}^2 \leq 0$,

where for both instances of (*), we used the formula from Proposition 9.1.1(a). We have now derived a contradiction, and it follows that $p = s$. This completes the argument. $\square$

Proposition 9.4.1

Let A be a symmetric matrix in $\mathbb{R}^{n\times n}$ with signature (n_+, n_-, n_0). Then there exists an invertible matrix $R \in \mathbb{R}^{n\times n}$ with pairwise orthogonal columns s.t.

$$R^T A R \;=\; D(\underbrace{1,\ldots,1}_{n_+},\underbrace{-1,\ldots,-1}_{n_-},\underbrace{0,\ldots,0}_{n_0}).$$

Theorem 9.4.3

Two symmetric matrices in $\mathbb{R}^{n\times n}$ are congruent iff they have the same signature.

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$.

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n\times n}$.
    - Then for all vectors $\mathbf{x} = \left[\begin{array}{ccc} x_1 & \ldots & x_n \end{array}\right]^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x}^T D\mathbf{x} = a_1x_1^2 + \cdots + a_nx_n^2,$$

    as can be seen via routine computation, or by applying Proposition 9.1.1(a).

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$.
  - Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x}^T D\mathbf{x} \quad = \quad a_1 x_1^2 + \cdots + a_n x_n^2,$$

    as can be seen via routine computation, or by applying Proposition 9.1.1(a).
  - This is a particularly nice formula, and for this reason, if q is a quadratic form over a field $\mathbb{F}$, it is helpful to have a basis $\mathcal{B}$ with respect to which the matrix of q is diagonal.

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \dots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$.
  - Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

    $$\mathbf{x}^T D\mathbf{x} \quad = \quad a_1 x_1^2 + \cdots + a_n x_n^2,$$

    as can be seen via routine computation, or by applying Proposition 9.1.1(a).
  - This is a particularly nice formula, and for this reason, if q is a quadratic form over a field $\mathbb{F}$, it is helpful to have a basis $\mathcal{B}$ with respect to which the matrix of q is diagonal.
  - Sylvester's law of inertia (in a couple of slides) states that when $V = \mathbb{R}^n$, such a basis $\mathcal{B}$ always exists.

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$.
  - Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x}^T D \mathbf{x} \quad = \quad a_1 x_1^2 + \cdots + a_n x_n^2,$$

    as can be seen via routine computation, or by applying Proposition 9.1.1(a).
  - This is a particularly nice formula, and for this reason, if q is a quadratic form over a field $\mathbb{F}$, it is helpful to have a basis $\mathcal{B}$ with respect to which the matrix of q is diagonal.
  - Sylvester's law of inertia (in a couple of slides) states that when $V = \mathbb{R}^n$, such a basis $\mathcal{B}$ always exists.
  - As we shall see, Sylvester's law of inertia is essentially a "translation" of Proposition 9.4.1 and Theorem 9.4.3 into the language of quadratic forms.

- Suppose that $\mathbb{F}$ is a field and that $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n\times n}$.
  - Then for all vectors $\mathbf{x} = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ in $\mathbb{F}^n$, we have that

$$\mathbf{x}^T D\mathbf{x} = a_1 x_1^2 + \cdots + a_n x_n^2,$$

    as can be seen via routine computation, or by applying Proposition 9.1.1(a).
  - This is a particularly nice formula, and for this reason, if q is a quadratic form over a field $\mathbb{F}$, it is helpful to have a basis $\mathcal{B}$ with respect to which the matrix of q is diagonal.
  - Sylvester's law of inertia (in a couple of slides) states that when $V = \mathbb{R}^n$, such a basis $\mathcal{B}$ always exists.
  - As we shall see, Sylvester's law of inertia is essentially a "translation" of Proposition 9.4.1 and Theorem 9.4.3 into the language of quadratic forms.
  - Before formally stating and proving the law, we need a definition.

Definition

The *signature* of a quadratic form q on $\mathbb{R}^n$ is defined to be the signature of the matrix of q with respect to **any** basis $\mathcal{B}$ of $\mathbb{R}^n$. A *polar basis* of $\mathbb{R}^n$ associated with the quadratic form q is any **orthogonal** basis $\mathcal{B}$ of $\mathbb{R}^n$ s.t. the matrix of q w.r.t. $\mathcal{B}$ is a diagonal matrix with only 1's, -1's, and 0's on the main diagonal.

- By Theorems 9.3.4 and 9.4.3, the signature of q is well defined.
  - Indeed, by Theorem 9.3.4, matrices of q with respect to all possible bases of $\mathbb{R}^n$ are congruent to each other, and so by Theorem 9.4.3, they all have the same signature.

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

- **Remark:** The basis $\mathcal{C}$ from the second sentence of Sylvester's law of inertia is not assumed to be polar, i.e. it is possible that it is not orthogonal.

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

- **Remark:** The basis $\mathcal{C}$ from the second sentence of Sylvester's law of inertia is not assumed to be polar, i.e. it is possible that it is not orthogonal.
- Let's prove the theorem!

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$.

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$. Set

$$D \quad := \quad D(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}).$$

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$. Set

$$D \quad := \quad D(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}).$$

By Proposition 9.4.1, there exists an invertible matrix $R \in \mathbb{R}^{n \times n}$ with pairwise orthogonal columns s.t. $D = R^T AR$. Since R is invertible, the Invertible Matrix Theorem guarantees that its columns form a basis $\mathcal{B}$ of $\mathbb{R}^n$;

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$. Set

$$D \quad := \quad D(\underbrace{1, \ldots, 1}_{n_+}, \underbrace{-1, \ldots, -1}_{n_-}, \underbrace{0, \ldots, 0}_{n_0}).$$

By Proposition 9.4.1, there exists an invertible matrix $R \in \mathbb{R}^{n \times n}$ with pairwise orthogonal columns s.t. $D = R^T A R$. Since R is invertible, the Invertible Matrix Theorem guarantees that its columns form a basis $\mathcal{B}$ of $\mathbb{R}^n$; since the columns of R are pairwise orthogonal, the basis $\mathcal{B}$ is orthogonal.

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$. Set

$$D \quad := \quad D(\underbrace{1, \dots, 1}_{n_+}, \underbrace{-1, \dots, -1}_{n_-}, \underbrace{0, \dots, 0}_{n_0}).$$

By Proposition 9.4.1, there exists an invertible matrix $R \in \mathbb{R}^{n \times n}$ with pairwise orthogonal columns s.t. $D = R^T A R$. Since R is invertible, the Invertible Matrix Theorem guarantees that its columns form a basis $\mathcal{B}$ of $\mathbb{R}^n$; since the columns of R are pairwise orthogonal, the basis $\mathcal{B}$ is orthogonal. Moreover, by Theorem 4.5.1 (or alternatively, by Lemma 4.5.8), we have that $R = {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_V \ \big]_{\mathcal{B}}$, so that

$$D \quad = \quad {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_V \ \big]_{\mathcal{B}}^T \ A \ {}_{\mathcal{E}_n}\big[\ \mathrm{Id}_V \ \big]_{\mathcal{B}}.$$

But now Theorem 9.3.3 guarantees that D is the matrix of q with respect to $\mathcal{B}$.

Proof. Let A be the matrix of the quadratic form q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$; then the signature of A is (n_+, n_-, n_0).

We first prove the existence of the polar basis $\mathcal{B}$. Set

$$D := D(\underbrace{1, \dots, 1}_{n_+}, \underbrace{-1, \dots, -1}_{n_-}, \underbrace{0, \dots, 0}_{n_0}).$$

By Proposition 9.4.1, there exists an invertible matrix $R \in \mathbb{R}^{n \times n}$ with pairwise orthogonal columns s.t. $D = R^T AR$. Since R is invertible, the Invertible Matrix Theorem guarantees that its columns form a basis $\mathcal{B}$ of $\mathbb{R}^n$; since the columns of R are pairwise orthogonal, the basis $\mathcal{B}$ is orthogonal. Moreover, by Theorem 4.5.1 (or alternatively, by Lemma 4.5.8), we have that $R = {}_{\mathcal{E}_n}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}$, so that

$$D = {}_{\mathcal{E}_n}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}^T\ A\ {}_{\mathcal{E}_n}[\ \mathrm{Id}_V\ ]_{\mathcal{B}}.$$

But now Theorem 9.3.3 guarantees that D is the matrix of q with respect to $\mathcal{B}$. We have already seen that the basis $\mathcal{B}$ is orthogonal, and we deduce that $\mathcal{B}$ is a polar basis of $\mathbb{R}^n$ associated with q.

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

Proof (continued). Now, fix any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix of q with respect to $\mathcal{C}$ is a diagonal matrix C with only 1's, -1's, and 0's on the main diagonal.

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

Proof (continued). Now, fix any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix of q with respect to $\mathcal{C}$ is a diagonal matrix C with only 1's, -1's, and 0's on the main diagonal. By Theorem 9.3.4, matrices A and C are congruent, and so by Theorem 9.4.3, they have the same signature, which is (n_+, n_-, n_0).

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

Proof (continued). Now, fix any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix of q with respect to $\mathcal{C}$ is a diagonal matrix C with only 1's, -1's, and 0's on the main diagonal. By Theorem 9.3.4, matrices A and C are congruent, and so by Theorem 9.4.3, they have the same signature, which is (n_+, n_-, n_0). Since the matrix C is diagonal, we know its entries on the main diagonal form its spectrum (this follows from Proposition 8.2.7);

Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

Proof (continued). Now, fix any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix of q with respect to $\mathcal{C}$ is a diagonal matrix C with only 1's, -1's, and 0's on the main diagonal. By Theorem 9.3.4, matrices A and C are congruent, and so by Theorem 9.4.3, they have the same signature, which is (n_+, n_-, n_0). Since the matrix C is diagonal, we know its entries on the main diagonal form its spectrum (this follows from Proposition 8.2.7); so, the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively. $\square$

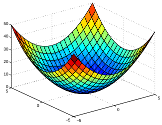
Sylvester's law of inertia

Let q be a quadratic form on $\mathbb{R}^n$, and let (n_+, n_-, n_0) be the signature of q. Then $\mathbb{R}^n$ has a polar basis $\mathcal{B}$ associated with q. Moreover, for any basis $\mathcal{C}$ of $\mathbb{R}^n$ s.t. the matrix C of q with respect to $\mathcal{C}$ is diagonal, with only 1's, -1's, and 0's on the main diagonal, the following holds: the number of 1's, -1's, and 0's on the main diagonal of C is n_+, n_-, and n_0, respectively.

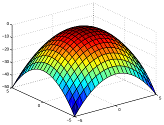
- For a numerical example, see the Lecture Notes.

- For quadratic forms on $\mathbb{R}^2$, there exist only six possible signatures (n_+, n_-, n_0), namely, the following:
  - $(2, 0, 0)$;
  - $(1, 0, 1)$;
  - $(1, 1, 0)$;
  - $(0, 2, 0)$;
  - $(0, 1, 1)$;
  - $(0, 0, 2)$.

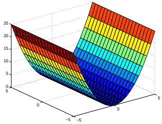
- For quadratic forms on $\mathbb{R}^2$, there exist only six possible signatures (n_+, n_-, n_0), namely, the following:
    - $(2, 0, 0)$;
    - $(1, 0, 1)$;
    - $(1, 1, 0)$;
    - $(0, 2, 0)$;
    - $(0, 1, 1)$;
    - $(0, 0, 2)$.
- Thus, the graph of any quadratic form q on $\mathbb{R}^2$ has the same general shape as one of the six graphs shown on the next slide (the one that has the same signature as q).



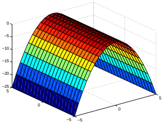
$x_1^2 + x_2^2$



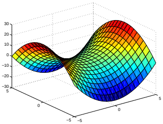
$-x_1^2 - x_2^2$



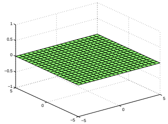
x_1^2



$-x_1^2$



$x_1^2 - x_2^2$



0

- The graphs were generated by Milan Hladík, who kindly shared them with me.

- The actual graph of the quadratic form q would be obtained by starting with one of the six graphs from the previous slide (the one that has the same signature as q), and then possibly stretching or shrinking the graph along the x_1- and x_2-axes (the coordinate axes of the domain), and then possibly rotating it about the vertical axis x_3.
    - This to account for the fact that a polar basis $\mathcal{B}$ of $\mathbb{R}^2$ associated with q is not necessarily equal to the standard basis $\mathcal{E}_2 = \{\mathbf{e}_1, \mathbf{e}_2\}$, but the vectors of $\mathcal{B}$ are indeed orthogonal to each other (by the definition of a polar basis).