Linear Algebra 2

Lecture #22
Cholesky decomposition of positive definite matrices. Bilinear and quadratic forms

Irena Penev

May 4, 2023

- This lecture has three parts:

- This lecture has three parts:
    1. Cholesky decomposition of positive definite matrices;

- This lecture has three parts:
  1. Cholesky decomposition of positive definite matrices;
  2. bilinear forms;

- This lecture has three parts:
  1. Cholesky decomposition of positive definite matrices;
  2. bilinear forms;
  3. quadratic forms.

1. the Cholesky decomposition of positive definite matrices

1. the Cholesky decomposition of positive definite matrices

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof.

1. the Cholesky decomposition of positive definite matrices

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

1. the Cholesky decomposition of positive definite matrices

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a positive definite matrix $A = \left[\, a \, \right]$ in $\mathbb{R}^{1 \times 1}$, and we note that $a > 0$ (because A is positive definite).

1. the Cholesky decomposition of positive definite matrices

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof. We proceed by induction on n.

For $n = 1$, we fix a positive definite matrix $A = \begin{bmatrix} a \end{bmatrix}$ in $\mathbb{R}^{1 \times 1}$, and we note that $a > 0$ (because A is positive definite). We now set $L := \begin{bmatrix} \sqrt{a} \end{bmatrix}$, and we observe that $A = LL^T$. The uniqueness of L is obvious.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n \times n}$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n\times n}$. Let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$, where $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n\times n}$, be a positive definite matrix in $\mathbb{R}^{(n+1)\times(n+1)}$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n\times n}$. Let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$, where $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n\times n}$, be a positive definite matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. By Theorem 4.1 of Lecture Notes 21, we have that $\alpha > 0$ and that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n \times n}$. Let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$, where $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$, be a positive definite matrix in $\mathbb{R}^{(n+1) \times (n+1)}$. By Theorem 4.1 of Lecture Notes 21, we have that $\alpha > 0$ and that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. By the induction hypothesis, there exists a unique lower triangular matrix $L' \in \mathbb{R}^{n \times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Now, fix a positive integer n, and assume the theorem is true for positive definite matrices in $\mathbb{R}^{n \times n}$. Let $A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right]$, where $\alpha \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$, and $A' \in \mathbb{R}^{n \times n}$, be a positive definite matrix in $\mathbb{R}^{(n+1) \times (n+1)}$. By Theorem 4.1 of Lecture Notes 21, we have that $\alpha > 0$ and that the matrix $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T$ is positive definite. By the induction hypothesis, there exists a unique lower triangular matrix $L' \in \mathbb{R}^{n \times n}$ with a positive main diagonal and s.t. $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$. We now set

$$L \quad := \quad \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right]_{n \times n}.$$

Clearly, L is lower triangular with a positive main diagonal.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). Moreover, we have that

$$
\begin{aligned}
LL^T &= \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0}^T \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right] \left[\begin{array}{c|c} \sqrt{\alpha} & \frac{1}{\sqrt{\alpha}}\mathbf{a}^T \\ \hline \mathbf{0} & L'^T \end{array}\right] \\
&= \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L'L'^T \end{array}\right] \\
&= \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & A' \end{array}\right] \\
&= A.
\end{aligned}
$$

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). It remains to show that L is unique.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). It remains to show that L is unique. For this, suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1 L_1^T$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). It remains to show that L is unique. For this, suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1 L_1^T$. Set

$$L_1 \;=\; \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right],$$

(here, β is some positive real number, $\mathbf{a}$ is some vector in $\mathbb{R}^n$, and L_1' is some lower triangular matrix in $\mathbb{R}^{n \times n}$ with a positive main diagonal).

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). It remains to show that L is unique. For this, suppose that $L_1 \in \mathbb{R}^{(n+1)\times(n+1)}$ is a lower triangular matrix with a positive main diagonal and satisfying $A = L_1 L_1^T$. Set

$$L_1 = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right],$$

(here, β is some positive real number, $\mathbf{a}$ is some vector in $\mathbb{R}^n$, and L_1' is some lower triangular matrix in $\mathbb{R}^{n \times n}$ with a positive main diagonal). Then

$$A = L_1 L_1^T = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right] \left[\begin{array}{c|c} \beta & \mathbf{b}^T \\ \hline \mathbf{0} & L_1'^T \end{array}\right] = \left[\begin{array}{c|c} \beta^2 & \beta\mathbf{b}^T \\ \hline \beta\mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right].$$

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c:c} \beta^2 & \beta\mathbf{b}^T \\ \hdashline \beta\mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c:c} \alpha & \mathbf{a}^T \\ \hdashline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta \mathbf{b}^T \\ \hline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta \mathbf{b}^T \\ \hline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$. This, together with the fact that $\beta > 0$, yields the fact that $\beta = \sqrt{\alpha}$, $\mathbf{b} = \frac{1}{\sqrt{\alpha}}\mathbf{a}$, and $L_1' L_1'^T = L' L'^T$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proof (continued). We now have that

$$\left[\begin{array}{c|c} \beta^2 & \beta \mathbf{b}^T \\ \hline \beta \mathbf{b} & \mathbf{b}\mathbf{b}^T + L_1' L_1'^T \end{array}\right] = A = \left[\begin{array}{c|c} \alpha & \mathbf{a}^T \\ \hline \mathbf{a} & \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T \end{array}\right].$$

But then $\beta^2 = \alpha$, $\beta \mathbf{b} = \mathbf{a}$, and $\mathbf{b}\mathbf{b}^T + L_1' L_1'^T = \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T + L' L'^T$. This, together with the fact that $\beta > 0$, yields the fact that $\beta = \sqrt{\alpha}$, $\mathbf{b} = \frac{1}{\sqrt{\alpha}}\mathbf{a}$, and $L_1' L_1'^T = L' L'^T$. But by the uniqueness L', we get that $L_1' = L'$. Thus,

$$L_1 = \left[\begin{array}{c|c} \beta & \mathbf{0}^T \\ \hline \mathbf{b} & L_1' \end{array}\right] = \left[\begin{array}{c|c} \sqrt{\alpha} & \mathbf{0} \\ \hline \frac{1}{\sqrt{\alpha}}\mathbf{a} & L' \end{array}\right] = L.$$

This proves the uniqueness of L. Q.E.D.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n\times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n\times n}$ with a positive main diagonal and satisfying $A = LL^T$.

- As the following proposition shows, the converse of Theorem 1.1 is also true.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

- As the following proposition shows, the converse of Theorem 1.1 is also true.

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof.

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof. First of all, we have that $A^T = (LL^T)^T = LL^T = A$, and so A is symmetric.

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof. First of all, we have that $A^T = (LL^T)^T = LL^T = A$, and so A is symmetric. Now, set $L = [\ \ell_{i,j}\ ]_{n \times n}$, and fix a vector $\mathbf{x} = [\ x_1 \quad \dots \quad x_n\ ]^T$ in $\mathbb{R}^n \setminus \{\mathbf{0}\}$. WTS $\mathbf{x}^T A \mathbf{x} > 0$.

Proposition 1.2

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof. First of all, we have that $A^T = (LL^T)^T = LL^T = A$, and so A is symmetric. Now, set $L = [\ \ell_{i,j}\ ]_{n\times n}$, and fix a vector $\mathbf{x} = [\ x_1 \quad \dots \quad x_n\ ]^T$ in $\mathbb{R}^n \setminus \{\mathbf{0}\}$. WTS $\mathbf{x}^T A\mathbf{x} > 0$. We have that

$$\mathbf{x}^T A\mathbf{x} \;=\; \mathbf{x}^T LL^T\mathbf{x} \;=\; (L^T\mathbf{x})^T(L^T\mathbf{x}) \;=\; (L^T\mathbf{x})\cdot(L^T\mathbf{x}) \;=\; ||L^T\mathbf{x}||^2.$$

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof. First of all, we have that $A^T = (LL^T)^T = LL^T = A$, and so A is symmetric. Now, set $L = [\ \ell_{i,j}\ ]_{n \times n}$, and fix a vector $\mathbf{x} = [\ x_1 \ \ \dots \ \ x_n\ ]^T$ in $\mathbb{R}^n \setminus \{\mathbf{0}\}$. WTS $\mathbf{x}^T A\mathbf{x} > 0$. We have that

$$\mathbf{x}^T A\mathbf{x} \ = \ \mathbf{x}^T LL^T\mathbf{x} \ = \ (L^T\mathbf{x})^T(L^T\mathbf{x}) \ = \ (L^T\mathbf{x}) \cdot (L^T\mathbf{x}) \ = \ ||L^T\mathbf{x}||^2.$$

Now, let $i \in \{1, \dots, n\}$ be the largest index s.t. $x_i \neq 0$ (the index i exists because $\mathbf{x} \neq \mathbf{0}$). Then the i-th entry of $L^T\mathbf{x}$ is $\ell_{i,i}x_i \neq 0$, and consequently, $L^T\mathbf{x} \neq \mathbf{0}$.

Proposition 1.2

Let $L \in \mathbb{R}^{n\times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

Proof. First of all, we have that $A^T = (LL^T)^T = LL^T = A$, and so A is symmetric. Now, set $L = [\ \ell_{i,j}\ ]_{n\times n}$, and fix a vector $\mathbf{x} = [\ x_1 \quad \dots \quad x_n\ ]^T$ in $\mathbb{R}^n \setminus \{\mathbf{0}\}$. WTS $\mathbf{x}^T A\mathbf{x} > 0$. We have that

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T LL^T\mathbf{x} = (L^T\mathbf{x})^T(L^T\mathbf{x}) = (L^T\mathbf{x}) \cdot (L^T\mathbf{x}) = ||L^T\mathbf{x}||^2.$$

Now, let $i \in \{1, \dots, n\}$ be the largest index s.t. $x_i \neq 0$ (the index i exists because $\mathbf{x} \neq \mathbf{0}$). Then the i-th entry of $L^T\mathbf{x}$ is $\ell_{i,i}x_i \neq 0$, and consequently, $L^T\mathbf{x} \neq \mathbf{0}$. So, $||L^T\mathbf{x}|| > 0$, and it follows that $\mathbf{x}^T A\mathbf{x} > 0$. This proves that A is positive definite. Q.E.D.

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

Proposition 1.2

Let $L \in \mathbb{R}^{n \times n}$ be a lower triangular matrix with a positive main diagonal. Then the matrix $A := LL^T$ is positive definite.

- There is also an algorithm that, for a positive definite matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{R}^{n \times n}$, computes the Cholesky decomposition of A, i.e. computes the (unique) lower triangular matrix $L = [\ \ell_{i,j}\ ]_{n \times n}$ in $\mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.
- We construct the matrix L column by column, from left to right.
- Each column is constructed from top to bottom. Here is the algorithm (next slide).

Theorem 1.1 [Cholesky decomposition]

For every positive definite matrix $A \in \mathbb{R}^{n \times n}$, there exists a unique lower triangular matrix $L \in \mathbb{R}^{n \times n}$ with a positive main diagonal and satisfying $A = LL^T$.

1. We construct the first column of L as follows:
   - $\ell_{1,1} := \sqrt{a_{1,1}}$,
   - $\ell_{i,1} := \frac{a_{i,1}}{\sqrt{a_{1,1}}}$ for all $i \in \{2, \dots, n\}$.
2. For all $j \in \{2, \dots, n\}$, assuming we have constructed the first $j-1$ columns of L, we construct the j-th column of L as follows (from top to bottom):
   - $\ell_{i,j} := 0$ for all $i \in \{1, \dots, j-1\}$,
   - $\ell_{j,j} := \sqrt{a_{j,j} - \sum_{k=1}^{j-1} \ell_{j,k}^2}$,
   - $\ell_{i,j} := \frac{1}{\ell_{j,j}} \left(a_{i,j} - \sum_{k=1}^{j-1} \ell_{i,k} \ell_{j,k} \right)$ for all $i \in \{j+1, \dots, n\}$.

- We omit the proof of correctness of the construction from the previous slide, but it essentially follows from Theorem 4.1 of Lecture Notes 21 and from the proof of Theorem 1.1.

- We omit the proof of correctness of the construction from the previous slide, but it essentially follows from Theorem 4.1 of Lecture Notes 21 and from the proof of Theorem 1.1.
- An example is given in the Lecture Notes.

- We omit the proof of correctness of the construction from the previous slide, but it essentially follows from Theorem 4.1 of Lecture Notes 21 and from the proof of Theorem 1.1.
- An example is given in the Lecture Notes.
- The main reason for interest in the Cholesky decomposition for positive definite matrices is that it allows one to solve equations of the form $A\mathbf{x} = \mathbf{b}$ (where A is positive definite) faster, as well as to compute the inverse of A faster. We omit the details.

2. Bilinear forms

2. Bilinear forms

Proposition 2.1 of Lecture Notes 21

Let $\mathbb{F}$ be a field. Then for all matrices $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$, and all vectors $\mathbf{x} = [\ x_1 \quad \dots \quad x_n\ ]^T$ and $\mathbf{y} = [\ y_1 \quad \dots \quad y_n\ ]^T$ in $\mathbb{F}^n$, we have that $\mathbf{x}^T A\mathbf{y} = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i y_j$.

2. Bilinear forms

Proposition 2.1 of Lecture Notes 21

Let $\mathbb{F}$ be a field. Then for all matrices $A = [\ a_{i,j}\ ]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and all vectors $\mathbf{x} = [\ x_1 \quad \dots \quad x_n\ ]^T$ and $\mathbf{y} = [\ y_1 \quad \dots \quad y_n\ ]^T$ in $\mathbb{F}^n$, we have that $\mathbf{x}^T A\mathbf{y} = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}x_iy_j$.

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

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 $\mathbf{x}, \mathbf{y} \in V$.

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

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 $\mathbf{x}, \mathbf{y} \in V$.

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

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 $\mathbf{x}, \mathbf{y} \in V$.

- Note that a scalar product in a real vector space is a symmetric bilinear form.

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

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 $\mathbf{x}, \mathbf{y} \in V$.

- Note that a scalar product in a real vector space is a symmetric bilinear form.
- However, a scalar product in a non-trivial complex vector space is **not** a bilinear form (because it does not satisfy b.4 above).

Proposition 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) for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) for all $\mathbf{y} \in V$, we have that $f(\mathbf{0}, \mathbf{y}) = 0$;

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

Proof.

Proposition 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) for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) for all $\mathbf{y} \in V$, we have that $f(\mathbf{0}, \mathbf{y}) = 0$;

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

Proof. We first prove (a). Fix $\mathbf{x} \in V$. Then

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

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

Proposition 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) for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{0}) = 0$;

(b) for all $\mathbf{y} \in V$, we have that $f(\mathbf{0}, \mathbf{y}) = 0$;

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

Proof. We first prove (a). Fix $\mathbf{x} \in V$. Then

$$f(\mathbf{x}, \mathbf{0}) \;=\; f(\mathbf{x}, \mathbf{0} + \mathbf{0}) \;=\; f(\mathbf{x}, \mathbf{0}) + f(\mathbf{x}, \mathbf{0}).$$

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 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) for all $\mathbf{x} \in V$, we have that $f(\mathbf{x}, \mathbf{0}) = 0$;
- (b) for all $\mathbf{y} \in V$, we have that $f(\mathbf{0}, \mathbf{y}) = 0$;
- (c) $f(\mathbf{0}, \mathbf{0}) = 0$.

Proof. We first prove (a). Fix $\mathbf{x} \in V$. Then

$$f(\mathbf{x}, \mathbf{0}) \quad = \quad f(\mathbf{x}, \mathbf{0} + \mathbf{0}) \quad = \quad f(\mathbf{x}, \mathbf{0}) + f(\mathbf{x}, \mathbf{0}).$$

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 $\mathbf{x} = \mathbf{0}$.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proof.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proof. The fact that f_A is a bilinear form on $\mathbb{F}^n$ readily follows from the properties of matrix multiplication.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proof. The fact that f_A is a bilinear form on $\mathbb{F}^n$ readily follows from the properties of matrix multiplication. Suppose now that the matrix A is symmetric. WTS f_A is symmetric.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proof. The fact that f_A is a bilinear form on $\mathbb{F}^n$ readily follows from the properties of matrix multiplication. Suppose now that the matrix A is symmetric. WTS f_A is symmetric. For all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{aligned} f_A(\mathbf{x}, \mathbf{y}) \quad &= \quad \mathbf{x}^T A\mathbf{y} \quad \overset{(*)}{=} \quad (\mathbf{x}^T A\mathbf{y})^T \quad = \quad \mathbf{y}^T A^T \mathbf{x} \\ &\overset{(**)}{=} \quad \mathbf{y}^T A\mathbf{x} \quad = \quad f_A(\mathbf{y}, \mathbf{x}), \end{aligned}$$

where (*) follows from the fact that $\mathbf{x}^T A\mathbf{y}$ is a 1×1 (and therefore symmetric) matrix, and (**) follows from the fact that A is a symmetric matrix. So, f_A is symmetric. Q.E.D.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Example 2.3

Consider the matrix $A = \begin{bmatrix} -1 & 2 \\ 3 & -4 \end{bmatrix}$ in $\mathbb{R}^{2\times 2}$. The function $f : \mathbb{R}^2 \times \mathbb{R}^2$ given by $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y} = -x_1y_1 + 2x_1y_2 + 3x_2y_1 - 4x_2y_2$ 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 $\mathbb{R}^2$ is a bilinear form on $\mathbb{R}^2$. It is not symmetric because (for example) $f(\mathbf{e}_1, \mathbf{e}_2) \neq f(\mathbf{e}_2, \mathbf{e}_1)$. (Indeed, $f(\mathbf{e}_1, \mathbf{e}_2) = 2$ and $f(\mathbf{e}_2, \mathbf{e}_1) = 3$.)

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

- In fact, bilinear forms from Proposition 2.2 are the only ones that exist for the vector space $\mathbb{F}^n$.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

- In fact, bilinear forms from Proposition 2.2 are the only ones that exist for the vector space $\mathbb{F}^n$.

Corollary 2.5

Let $\mathbb{F}$ be a field, and let f be a bilinear form on $\mathbb{F}^n$. Then there exists a unique matrix $A = \left[\ a_{i,j}\ \right]_{n \times n}$ in $\mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, we have that $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. Moreover, the matrix A is given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all $i, j \in \{1, \dots, n\}$, and it is symmetric iff the bilinear form f is symmetric.

Proposition 2.2

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times n}$. Then the function $f_A : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ given by $f_A(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ is a bilinear form. Moreover, if the matrix A is symmetric, then the bilinear form f_A is symmetric.

- In fact, bilinear forms from Proposition 2.2 are the only ones that exist for the vector space $\mathbb{F}^n$.

Corollary 2.5

Let $\mathbb{F}$ be a field, and let f be a bilinear form on $\mathbb{F}^n$. Then there exists a unique matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, we have that $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A\mathbf{y}$. Moreover, the matrix A is given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all $i, j \in \{1, \dots, n\}$, and it is symmetric iff the bilinear form f is symmetric.

- First, we need a theorem that implies Corollary 2.5!

Theorem 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 $f : V \times V \to \mathbb{F}$ be a bilinear form in V. Let $A = [\ a_{i,j}\ ]_{n \times n}$ be the matrix in $\mathbb{F}^{n \times n}$ given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$. Then all the following hold:

- (a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$;
- (b) the matrix A is symmetric iff the bilinear form f is symmetric;
- (c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n \times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}\ \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

Theorem 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 $f : V \times V \to \mathbb{F}$ be a bilinear form in V. Let $A = [\ a_{i,j}\ ]_{n \times n}$ be the matrix in $\mathbb{F}^{n \times n}$ given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$. Then all the following hold:

(a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$;

(b) the matrix A is symmetric iff the bilinear form f is symmetric;

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n \times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}\ \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

- **Terminology:** The matrix A from Theorem 2.4 is called the *matrix of the bilinear form* f with respect to the basis $\mathcal{B}$.

Proof. WTS

(a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^{T} \, A \, [\mathbf{y}]_{\mathcal{B}}$.

Proof. WTS

(a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{y}]_{\mathcal{B}}$.

Fix $\mathbf{x}, \mathbf{y} \in V$, and set $[\mathbf{x}]_{\mathcal{B}} = [\; x_1 \;\; \dots \;\; x_n \;]^T$ and $[\mathbf{y}]_{\mathcal{B}} = [\; y_1 \;\; \dots \;\; y_n \;]^T$, so that $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{b}_i$ and $\mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{b}_i$. We now compute:

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f\Big(\sum_{i=1}^{n} x_i \mathbf{b}_i, \sum_{j=1}^{n} y_j \mathbf{b}_j\Big) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j f(\mathbf{b}_i, \mathbf{b}_j) && \text{because } f \text{ is bilinear} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j a_{i,j} \\ &= [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{y}]_{\mathcal{B}} && \text{by Prop. 2.1 of Lec. 21.} \end{aligned}$$

Proof (continued). WTS

(b) A is symmetric iff f is symmetric.

Proof (continued). WTS

(b) A is symmetric iff f is symmetric.

If f is symmetric, then obviously, A symmetric.

Proof (continued). WTS

(b) A is symmetric iff f is symmetric.

If f is symmetric, then obviously, A symmetric. Suppose now that A is symmetric. WTS f is symmetric.

Proof (continued). WTS

(b) A is symmetric iff f is symmetric.

If f is symmetric, then obviously, A symmetric. Suppose now that A is symmetric. WTS f is symmetric. For all $\mathbf{x}, \mathbf{y} \in V$, we have that

$$\begin{aligned}
f(\mathbf{x}, \mathbf{y}) &= [\mathbf{x}]_{\mathcal{B}}^{T} \, A \, [\mathbf{y}]_{\mathcal{B}} && \text{by (a)} \\
&= ([\mathbf{x}]_{\mathcal{B}}^{T} \, A \, [\mathbf{y}]_{\mathcal{B}})^{T} && \text{because } [\mathbf{x}]_{\mathcal{B}}^{T} \, A \, [\mathbf{y}]_{\mathcal{B}} \text{ is } 1 \times 1 \\
&= [\mathbf{y}]_{\mathcal{B}}^{T} \, A^{T} \, [\mathbf{x}]_{\mathcal{B}} && \\
&= [\mathbf{y}]_{\mathcal{B}}^{T} \, A \, [\mathbf{x}]_{\mathcal{B}} && \text{because } A \text{ is symmetric} \\
&= f(\mathbf{y}, \mathbf{x}) && \text{by (a).}
\end{aligned}$$

Thus, the bilinear form f is symmetric.

Proof (continued). WTS

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n \times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^{T} \; A' \; [\mathbf{y}]_{\mathcal{B}} \; \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

Proof (continued). WTS

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n\times n}$ is any matrix s.t. $f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}\ \forall \mathbf{x},\mathbf{y} \in V$, then $A' = A$.

Fix a matrix $A' = [\ a'_{i,j}\ ]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and assume that $\forall \mathbf{x},\mathbf{y} \in V$: $f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}$.

Proof (continued). WTS

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n\times n}$ is any matrix s.t. $f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \; A' \; [\mathbf{y}]_{\mathcal{B}} \; \forall \mathbf{x},\mathbf{y} \in V$, then $A' = A$.

Fix a matrix $A' = [\ a'_{i,j}\ ]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and assume that $\forall \mathbf{x},\mathbf{y} \in V$: $f(\mathbf{x},\mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \; A' \; [\mathbf{y}]_{\mathcal{B}}$. WTS $A' = A$.

Proof (continued). WTS

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n\times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \; A' \; [\mathbf{y}]_{\mathcal{B}} \; \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

Fix a matrix $A' = [\; a'_{i,j} \;]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and assume that $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \; A' \; [\mathbf{y}]_{\mathcal{B}}$. WTS $A' = A$. Indeed, $\forall i, j \in \{1, \dots, n\}$, we have that

$$a_{i,j} \;=\; f(\mathbf{b}_i, \mathbf{b}_j) \;=\; [\mathbf{b}_i]_{\mathcal{B}}^T \; A' \; [\mathbf{b}_j]_{\mathcal{B}} \;=\; \mathbf{e}_i A' \mathbf{e}_j \;\overset{(*)}{=}\; a'_{i,j},$$

where (*) follows from Proposition 2.1 of Lecture Notes 21. This proves that $A' = A$. Q.E.D.

Theorem 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 $f : V \times V \to \mathbb{F}$ be a bilinear form in V. Let $A = [\ a_{i,j}\ ]_{n\times n}$ be the matrix in $\mathbb{F}^{n\times n}$ given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$. Then all the following hold:

(a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$;

(b) the matrix A is symmetric iff the bilinear form f is symmetric;

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n\times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}\ \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

- **Terminology:** The matrix A from Theorem 2.4 is called the *matrix of the bilinear form f* with respect to the basis $\mathcal{B}$.

- Recall that any scalar product in a real vector space is a bilinear form on that vector space.

- Recall that any scalar product in a real vector space is a bilinear form on that vector space.
- In the context of Theorem 2.4, it may be worth recalling the following theorem from Lecture Notes 21.

Theorem 2.2 of Lectute Notes 21

For any $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the following are equivalent:

(i) $\langle \cdot, \cdot \rangle$ is a scalar product on $\mathbb{R}^n$;

(ii) there exists a positive definite matrix $A \in \mathbb{R}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A\mathbf{y}$.

Theorem 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 $f : V \times V \to \mathbb{F}$ be a bilinear form in V. Let $A = [\ a_{i,j}\ ]_{n \times n}$ be the matrix in $\mathbb{F}^{n \times n}$ given by $a_{i,j} = f(\mathbf{b}_i, \mathbf{b}_j)$ for all indices $i, j \in \{1, \dots, n\}$. Then all the following hold:

(a) $\forall \mathbf{x}, \mathbf{y} \in V$: $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$;

(b) the matrix A is symmetric iff the bilinear form f is symmetric;

(c) A is the only matrix satisfying the property from (a), that is, if $A' \in \mathbb{F}^{n \times n}$ is any matrix s.t. $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A'\ [\mathbf{y}]_{\mathcal{B}}\ \forall \mathbf{x}, \mathbf{y} \in V$, then $A' = A$.

- **Terminology:** The matrix A from Theorem 2.4 is called the *matrix of the bilinear form* f with respect to the basis $\mathcal{B}$.

Corollary 2.5

Let $\mathbb{F}$ be a field, and let f be a bilinear form on $\mathbb{F}^n$. Then there exists a unique matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, we have that $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. Moreover, the matrix A is given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all $i, j \in \{1, \dots, n\}$, and it is symmetric iff the bilinear form f is symmetric.

Proof.

Corollary 2.5

Let $\mathbb{F}$ be a field, and let f be a bilinear form on $\mathbb{F}^n$. Then there exists a unique matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, we have that $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. Moreover, the matrix A is given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all $i, j \in \{1, \dots, n\}$, and it is symmetric iff the bilinear form f is symmetric.

Proof. Since $[\mathbf{x}]_{\mathcal{E}_n} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{F}^n$, the result follows immediately from Theorem 2.4. Q.E.D.

Corollary 2.5

Let $\mathbb{F}$ be a field, and let f be a bilinear form on $\mathbb{F}^n$. Then there exists a unique matrix $A = [\ a_{i,j}\ ]_{n \times n}$ in $\mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, we have that $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y}$. Moreover, the matrix A is given by $a_{i,j} = f(\mathbf{e}_i, \mathbf{e}_j)$ for all $i, j \in \{1, \dots, n\}$, and it is symmetric iff the bilinear form f is symmetric.

- **Remark:** Corollary 2.5 (together with Proposition 2.1 of Lecture Notes 21) 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

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j} x_i y_j \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{F}^n$$

where the $a_{i,j}$ are some scalars in $\mathbb{F}$. This bilinear form is symmetric iff $a_{i,j} = a_{j,i}$ for all indices $i, j \in \{1, \dots, n\}$.

- For example, the following are bilinear forms on $\mathbb{R}^2$:
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$;
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$.

- For example, the following are bilinear forms on $\mathbb{R}^2$:
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$;
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$.
- The bilinear form f_1 is symmetric, whereas the bilinear form f_2 is not.

- For example, the following are bilinear forms on $\mathbb{R}^2$:
  - $f_1(\mathbf{x}, \mathbf{y}) = x_1y_1 - 3x_1y_2 - 3x_2y_1 + 7x_2y_2$;
  - $f_2(\mathbf{x}, \mathbf{y}) = x_1y_1 - 2x_1y_2 + 3x_2y_1 - 3x_2y_2$.
- The bilinear form f_1 is symmetric, whereas the bilinear form f_2 is not.
- Note also that the matrix of the bilinear form f_1 with respect to the standard basis $\mathcal{E}_2$ of $\mathbb{R}^2$ is $\begin{bmatrix} 1 & -3 \\ -3 & 7 \end{bmatrix}$, whereas the matrix of the bilinear form f_2 with respect to the standard basis $\mathcal{E}_2$ of $\mathbb{R}^2$ is $\begin{bmatrix} 1 & -2 \\ 3 & -3 \end{bmatrix}$.

Corollary 2.6

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

Proof.

Corollary 2.6

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

Proof. Set $A := [\ a_{i,j}\ ]_{n \times n}$, and let $f : V \times V \to \mathbb{F}$ be given by $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$.

Corollary 2.6

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

Proof. Set $A := [\ a_{i,j}\ ]_{n \times n}$, and let $f : V \times V \to \mathbb{F}$ be given by $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}}$. It is then straightforward to verify that f is a bilinear form. Moreover, by Theorem 2.4, A is the matrix of the bilinear form f.

Corollary 2.6

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

Proof (continued). It remains to show that the bilinear form f is unique.

Corollary 2.6

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

Proof (continued). It remains to show that the bilinear form f is unique. Let g be a bilinear form on V that satisfies $g(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$.

Corollary 2.6

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

Proof (continued). It remains to show that the bilinear form f is unique. Let g be a bilinear form on V that satisfies $g(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$. But then by Theorem 2.4, A is the matrix of the bilinear form g with respect to the basis $\mathcal{B}$.

Corollary 2.6

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

Proof (continued). It remains to show that the bilinear form f is unique. Let g be a bilinear form on V that satisfies $g(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$ for all $i, j \in \{1, \dots, n\}$. But then by Theorem 2.4, A is the matrix of the bilinear form g with respect to the basis $\mathcal{B}$. Since the bilinear forms f and g have the same matrix with respect to the basis $\mathcal{B}$, we see that $f = g$. (Indeed, for all $\mathbf{x}, \mathbf{y} \in V$, we have that $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T\ A\ [\mathbf{y}]_{\mathcal{B}} = g(\mathbf{x}, \mathbf{y})$, and it follows that $f = g$.) Q.E.D.

Theorem 2.7 [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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the matrices of the bilinear form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \, B \, {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$.

Proof.

Theorem 2.7 [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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the matrices of the bilinear form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \, B \, {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$.

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

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

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

Theorem 2.7 [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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the matrices of the bilinear form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T\ B\ {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$.

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

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

where (*) follows from the fact that B is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$. But now we have that ${}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T\ B\ {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$ is the matrix of the bilinear form f with respect to the basis $\mathcal{C}$, and so by Theorem 2.4(c), we have that $C = {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T\ B\ {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$. Q.E.D.

- Recall that 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 $\underbrace{1 + \cdots + 1}_{n} = 0$.

- Recall that 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 $\underbrace{1 + \cdots + 1}_{n} = 0$.
- If no such n exists, then $\text{char}(\mathbb{F}) := 0$.

- Recall that 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 $\underbrace{1 + \cdots + 1}_{n} = 0$.
- If no such n exists, then $\text{char}(\mathbb{F}) := 0$.
- Note that fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic 0.

- Recall that 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 $\underbrace{1 + \dots + 1}_{n} = 0$.
- If no such n exists, then $\text{char}(\mathbb{F}) := 0$.
- Note that fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic 0.
- On the other hand, for all prime numbers p, we have that $\text{char}(\mathbb{Z}_p) = p$.

- Recall that 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 $\underbrace{1 + \cdots + 1}_{n} = 0$.
- If no such n exists, then $\mathrm{char}(\mathbb{F}) := 0$.
- Note that fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all have characteristic 0.
- On the other hand, for all prime numbers p, we have that $\mathrm{char}(\mathbb{Z}_p) = p$.

Theorem 2.8

The characteristic of any field is either a prime number or 0.

Proof. Lecture Notes.

Theorem 2.8

The characteristic of any field is either a prime number or 0.

Theorem 2.8

The characteristic of any field is either a prime number or 0.

- In what follows, we will mostly focus on vector spaces over fields of characteristic other than 2.

Theorem 2.8

The characteristic of any field is either a prime number or 0.

- In what follows, we will mostly focus on vector spaces over fields of characteristic other than 2.
- This will be important because in such fields, we can divide by 2 (because $2 = 1 + 1 \neq 0$).

Theorem 2.8

The characteristic of any field is either a prime number or 0.

- In what follows, we will mostly focus on vector spaces over fields of characteristic other than 2.
- This will be important because in such fields, we can divide by 2 (because $2 = 1 + 1 \neq 0$).
- The only field of characteristic 2 that we have seen is $\mathbb{Z}_2$, but other fields of characteristic 2 do exist.

Proposition 2.9

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 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$.

Proposition 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis.

Proposition 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$.

Proposition 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis:

(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 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis:

(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 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis:

(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 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis:

(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 2.9

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 $\mathbf{x}, \mathbf{y} \in V$. WTS $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis:

(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 $\text{char}(\mathbb{F}) \neq 2$, we deduce that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. Q.E.D.

3. Quadratic forms

3. Quadratic forms

Definition

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

3. Quadratic forms

Definition

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

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof. By the definition of a quadratic form, there exists some bilinear form g on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = g(\mathbf{x}, \mathbf{x})$.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof. By the definition of a quadratic form, there exists some bilinear form g on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define a function $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\Big(g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})\Big)$.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof. By the definition of a quadratic form, there exists some bilinear form g on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define a function $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\Big(g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})\Big)$. It is straightforward to check that f is a bilinear form.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof. By the definition of a quadratic form, there exists some bilinear form g on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = g(\mathbf{x}, \mathbf{x})$. Now, using the fact that $\text{char}(\mathbb{F}) \neq 2$, we define a function $f : V \times V \to \mathbb{F}$ by setting $f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\Big(g(\mathbf{x}, \mathbf{y}) + g(\mathbf{y}, \mathbf{x})\Big)$. It is straightforward to check that f is a bilinear form. Moreover, for all $\mathbf{x} \in V$, we have that

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

which is what we needed.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof (continued). It remains to show that f is unique.

Theorem 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof (continued). It remains to show that f is unique. Fix any 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 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 $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$.

Proof (continued). It remains to show that f is unique. Fix any symmetric bilinear form f' on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f'(\mathbf{x}, \mathbf{x})$. We must show that $f' = f$. But this readily follows from Proposition 2.9, since $\forall \mathbf{x} \in V$: $f'(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Q.E.D.

Corollary 3.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let q be a quadratic form on V. Then there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{x}]_{\mathcal{B}}$.

Corollary 3.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let q be a quadratic form on V. Then there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{x}]_{\mathcal{B}}$.

- **Terminology:** The matrix A from Corollary 3.2 is called the *matrix of the quadratic form f* with respect to the basis $\mathcal{B}$.

Corollary 3.2

Let V be a non-trivial, finite-dimensional vector space over a field $\mathbb{F}$ of characteristic other than 2, let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of V, and let q be a quadratic form on V. Then there exists a unique **symmetric** matrix $A \in \mathbb{F}^{n \times n}$ s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{x}]_{\mathcal{B}}$.

- **Terminology:** The matrix A from Corollary 3.2 is called the *matrix of the quadratic form f* with respect to the basis $\mathcal{B}$.

Proof (outline). Consider the unique symmetric bilinear form f on V s.t. $\forall \mathbf{x} \in V$: $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. ($f$ exists by Theorem 3.1.) Then A is the matrix of f with respect to $\mathcal{B}$. (Details: Lecture Notes.)

Corollary 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \ B \ {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$.

Proof.

Corollary 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$.

Proof. By Theorem 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 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$.

Proof. By Theorem 3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. The matrix of the bilinear form f with respect to $\mathcal{B}$ is also the (symmetric) matrix of the quadratic q with respect to $\mathcal{B}$.

Corollary 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$.

Proof. By Theorem 3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. The matrix of the bilinear form f with respect to $\mathcal{B}$ is also the (symmetric) matrix of the quadratic q with respect to $\mathcal{B}$. Since the latter is unique (by Corollary 3.2), it follows that B is the matrix of the bilinear form f with respect to $\mathcal{B}$.

Corollary 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[\text{Id}_V]_{\mathcal{C}}$.

Proof. By Theorem 3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. The matrix of the bilinear form f with respect to $\mathcal{B}$ is also the (symmetric) matrix of the quadratic q with respect to $\mathcal{B}$. Since the latter is unique (by Corollary 3.2), it follows that B is the matrix of the bilinear form f with respect to $\mathcal{B}$. Similarly, C is the matrix of the bilinear form f with respect to $\mathcal{C}$.

Corollary 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, let $\mathcal{B}$ and $\mathcal{C}$ be bases of V, and let B and C be the (symmetric) matrices of the quadratic form f with respect to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$.

Proof. By Theorem 3.1, there exists a unique symmetric bilinear form f on V s.t. for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. The matrix of the bilinear form f with respect to $\mathcal{B}$ is also the (symmetric) matrix of the quadratic q with respect to $\mathcal{B}$. Since the latter is unique (by Corollary 3.2), it follows that B is the matrix of the bilinear form f with respect to $\mathcal{B}$. Similarly, C is the matrix of the bilinear form f with respect to $\mathcal{C}$. The result now follows immediately from Theorem 2.7. Q.E.D.

- **Remark:** Note that if $\mathbb{F}$ is a field, $D = D(a_1, \ldots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n\times n}$, and $\mathbf{x} = [\begin{array}{ccc} x_1 & \ldots & x_n \end{array}]^T$ is a vector in $\mathbb{F}^n$, then $\mathbf{x}^T D\mathbf{x} = a_1 x_1^2 + \cdots + a_n x_n^2$.

- **Remark:** Note that if $\mathbb{F}$ is a field, $D = D(a_1, \dots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n\times n}$, and $\mathbf{x} = [\begin{array}{ccc} x_1 & \dots & x_n \end{array}]^T$ is a vector in $\mathbb{F}^n$, then $\mathbf{x}^T D\mathbf{x} = a_1x_1^2 + \dots + a_nx_n^2$.
- This is a particularly nice formula, and for this reason, Sylvester's law of inertia is of interest.

- **Remark:** Note that if $\mathbb{F}$ is a field, $D = D(a_1, \dots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$, and $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ is a vector in $\mathbb{F}^n$, then $\mathbf{x}^T D \mathbf{x} = a_1 x_1^2 + \dots + a_n x_n^2$.
- This is a particularly nice formula, and for this reason, Sylvester's law of inertia is of interest.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

- **Remark:** Note that if $\mathbb{F}$ is a field, $D = D(a_1, \dots, a_n)$ is a diagonal matrix in $\mathbb{F}^{n \times n}$, and $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ is a vector in $\mathbb{F}^n$, then $\mathbf{x}^T D \mathbf{x} = a_1 x_1^2 + \dots + a_n x_n^2$.
- This is a particularly nice formula, and for this reason, Sylvester's law of inertia is of interest.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

- **Terminology:** A basis $\mathcal{B}$ from the statement of Sylvester's law of inertia is called a *polar basis* of $\mathbb{R}^n$ for the quadratic form q.

Proposition 3.4

Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times m}$. Then all the following hold:

(a) for all invertible matrices $S \in \mathbb{F}^{n \times n}$, we have that $\text{rank}(SA) = \text{rank}(A)$;

(b) for all invertible matrices $S \in \mathbb{F}^{m \times m}$, we have that $\text{rank}(AS) = \text{rank}(A)$;

(c) for all invertible matrices $S_1 \in \mathbb{F}^{n \times n}$ and $S_2 \in \mathbb{F}^{m \times m}$, we have that $\text{rank}(S_1 A S_2) = \text{rank}(A)$.

Proof. Lecture Notes.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. 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 $\mathbb{R}^{n \times n}$, such that $D = Q^T A Q$.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n\times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. 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 $\mathbb{R}^{n\times n}$, such that $D = Q^T AQ$.

Next, $\forall i \in \{1, \ldots, n\}$, set $\ell_i := 1/\sqrt{|\lambda_i|}$ if $\lambda_i \neq 0$, and set $\ell_i := 1$ otherwise.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. 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 $\mathbb{R}^{n \times n}$, such that $D = Q^T A Q$.

Next, $\forall i \in \{1, \ldots, n\}$, set $\ell_i := 1/\sqrt{|\lambda_i|}$ if $\lambda_i \neq 0$, and set $\ell_i := 1$ otherwise. Set $L := D(\ell_1, \ldots, \ell_n)$, and set $R := QL$.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. By the spectral theorem for symmetric matrices, we know that A is orthogonally diagonalizable. So, let $D = D(\lambda_1, \dots, \lambda_n)$ be a diagonal and Q an orthogonal matrix, both in $\mathbb{R}^{n \times n}$, such that $D = Q^T A Q$.

Next, $\forall i \in \{1, \dots, n\}$, set $\ell_i := 1/\sqrt{|\lambda_i|}$ if $\lambda_i \neq 0$, and set $\ell_i := 1$ otherwise. Set $L := D(\ell_1, \dots, \ell_n)$, and set $R := QL$. Since both Q and L are invertible, so is R. So, the columns of R form a basis $\mathcal{B}$ of $\mathbb{R}^n$.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.

Proof. We first prove existence. First, using Corollary 3.2, we let A be the (symmetric) matrix of q with respect to the standard basis $\mathcal{E}_n$ of $\mathbb{R}^n$. 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 $\mathbb{R}^{n \times n}$, such that $D = Q^T A Q$.

Next, $\forall i \in \{1, \ldots, n\}$, set $\ell_i := 1/\sqrt{|\lambda_i|}$ if $\lambda_i \neq 0$, and set $\ell_i := 1$ otherwise. Set $L := D(\ell_1, \ldots, \ell_n)$, and set $R := QL$. Since both Q and L are invertible, so is R. So, the columns of R form a basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, by Lemma 2.6 of Lecture Notes 10, we have that ${}_{\mathcal{E}_n}[\mathrm{Id}_V]_{\mathcal{B}} = R$, and so by Corollary 3.3, $R^T A R$ is the matrix of the quadratic form q with respect to the basis $\mathcal{B}$.

Proof (continued). It remains to show that the matrix $R^T AR$ is diagonal, and that all its entries on the main diagonal are 1, -1, or 0.

Proof (continued). It remains to show that the matrix $R^T AR$ is diagonal, and that all its entries on the main diagonal are 1, -1, or 0. We compute:

$$\begin{aligned} R^T AR &= (QL)^T A(QL) \\ &= L^T \underbrace{Q^T AQ}_{=D} L \\ &= D(\ell_1, \dots, \ell_n)\ D(\lambda_1, \dots, \lambda_n)\ D(\ell_1, \dots, \ell_n) \\ &= D(\lambda_n \ell_n^2, \dots, \lambda_n \ell_n^2). \end{aligned}$$

Proof (continued). It remains to show that the matrix $R^T AR$ is diagonal, and that all its entries on the main diagonal are 1, -1, or 0. We compute:

$$\begin{aligned}
R^T AR &= (QL)^T A(QL) \\
&= L^T \underbrace{Q^T AQ}_{=D} L \\
&= D(\ell_1, \dots, \ell_n)\, D(\lambda_1, \dots, \lambda_n)\, D(\ell_1, \dots, \ell_n) \\
&= D(\lambda_n \ell_n^2, \dots, \lambda_n \ell_n^2).
\end{aligned}$$

But by construction, we have that

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

for all indices $i \in \{1, \dots, n\}$. So, the matrix $R^T AR$ is indeed diagonal with only entries $1, -1, 0$ on the main diagonal.

Proof (continued). It remains to prove uniqueness.

Proof (continued). It remains to prove uniqueness. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ are two bases of $\mathbb{F}^n$, such that the matrices B and C of the quadratic form q with respect to $\mathcal{B}$ and $\mathcal{C}$, respectively, are both diagonal and have only entries $1, -1, 0$ on the main diagonal.

Proof (continued). It remains to prove uniqueness. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ are two bases of $\mathbb{F}^n$, such that the matrices B and C of the quadratic form q with respect to $\mathcal{B}$ and $\mathcal{C}$, respectively, are both diagonal and have only entries $1, -1, 0$ on the main diagonal. WTS B and C are identical up to a reordering of the main diagonal entries.

Proof (continued). It remains to prove uniqueness. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ are two bases of $\mathbb{F}^n$, such that the matrices B and C of the quadratic form q with respect to $\mathcal{B}$ and $\mathcal{C}$, respectively, are both diagonal and have only entries $1, -1, 0$ on the main diagonal. WTS B and C are identical up to a reordering of the main diagonal entries.

After possibly reordering the basis elements of $\mathcal{B}$ and $\mathcal{C}$, WMA

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

for some $p, q, s, t \in \{0, \ldots, n\}$ such that $p + q \leq n$ and $s + t \leq n$.

Proof (continued). It remains to prove uniqueness. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ are two bases of $\mathbb{F}^n$, such that the matrices B and C of the quadratic form q with respect to $\mathcal{B}$ and $\mathcal{C}$, respectively, are both diagonal and have only entries $1, -1, 0$ on the main diagonal. WTS B and C are identical up to a reordering of the main diagonal entries.

After possibly reordering the basis elements of $\mathcal{B}$ and $\mathcal{C}$, WMA

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

for some $p, q, s, t \in \{0, \dots, n\}$ such that $p + q \leq n$ and $s + t \leq n$. It now suffices to show that $p + q = s + t$ and $p = s$, for this will immediately imply that $B = C$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

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

WTS $p + q = s + t$ and $p = s$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

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

WTS $p + q = s + t$ and $p = s$.

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

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

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

WTS $p + q = s + t$ and $p = s$.

We first show that $p + q = s + t$. Set $S := {}_{\mathcal{C}}[\mathrm{Id}_V]_{\mathcal{B}}$. By Corollary 3.3, we have that $B = S^T C S$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

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

WTS $p + q = s + t$ and $p = s$.

We first show that $p + q = s + t$. Set $S := {}_{\mathcal{C}}[\mathrm{Id}_V]_{\mathcal{B}}$. By Corollary 3.3, we have that $B = S^T C S$. Now, the change-of-basis matrix S is invertible, and consequently, its transpose S^T is also invertible. So, Proposition 3.4 guarantees that $\mathrm{rank}(B) = \mathrm{rank}(C)$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

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

WTS $p + q = s + t$ and $p = s$.

We first show that $p + q = s + t$. Set $S := {}_{\mathcal{C}}[\mathrm{Id}_V]_{\mathcal{B}}$. By Corollary 3.3, we have that $B = S^T C S$. Now, the change-of-basis matrix S is invertible, and consequently, its transpose S^T is also invertible. So, Proposition 3.4 guarantees that $\mathrm{rank}(B) = \mathrm{rank}(C)$. But $\mathrm{rank}(B) = p + q$ and $\mathrm{rank}(C) = s + t$, and it follows that $p + q = s + t$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}, \underbrace{0, \ldots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \ldots, 1}_{s}, \underbrace{-1, \ldots, -1}_{t}, \underbrace{0, \ldots, 0}_{n-s-t})$.

WTS $p = s$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}, \underbrace{0, \ldots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \ldots, 1}_{s}, \underbrace{-1, \ldots, -1}_{t}, \underbrace{0, \ldots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \mathrm{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_p)$ and $U_C := \mathrm{Span}(\mathbf{c}_{s+1}, \ldots, \mathbf{c}_n)$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1).

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}, \underbrace{0, \ldots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \ldots, 1}_{s}, \underbrace{-1, \ldots, -1}_{t}, \underbrace{0, \ldots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \ldots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1). But $\dim(U_B) + \dim(U_C) = p + (n - s) > n$ and $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$; 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). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1). But $\dim(U_B) + \dim(U_C) = p + (n - s) > n$ and $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$; so, $\dim(U_B \cap U_C) > 0$, and it follows that $U_B \cap U_C$ contains some non-zero vector $\mathbf{u}$. Set $[\mathbf{u}]_{\mathcal{B}} = [\begin{array}{ccc} x_1 & \dots & x_n \end{array}]^T$ and $[\mathbf{u}]_{\mathcal{C}} = [\begin{array}{ccc} y_1 & \dots & y_n \end{array}]^T$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1). But $\dim(U_B) + \dim(U_C) = p + (n - s) > n$ and $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$; so, $\dim(U_B \cap U_C) > 0$, and it follows that $U_B \cap U_C$ contains some non-zero vector $\mathbf{u}$. Set $[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $[\mathbf{u}]_{\mathcal{C}} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$. Then at least one of $x_1, \dots, x_p$ is non-zero, $x_{p+1} = \dots = x_n = 0$, and $y_1 = \dots = y_s = 0$.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}, \underbrace{0, \ldots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \ldots, 1}_{s}, \underbrace{-1, \ldots, -1}_{t}, \underbrace{0, \ldots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \ldots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \ldots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1). But $\dim(U_B) + \dim(U_C) = p + (n - s) > n$ and $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$; so, $\dim(U_B \cap U_C) > 0$, and it follows that $U_B \cap U_C$ contains some non-zero vector $\mathbf{u}$. Set $[\mathbf{u}]_{\mathcal{B}} = [\begin{array}{ccc} x_1 & \ldots & x_n \end{array}]^T$ and $[\mathbf{u}]_{\mathcal{C}} = [\begin{array}{ccc} y_1 & \ldots & y_n \end{array}]^T$. Then at least one of $x_1, \ldots, x_p$ is non-zero, $x_{p+1} = \cdots = x_n = 0$, and $y_1 = \cdots = y_s = 0$. We now have that

- $q(\mathbf{u}) = [\mathbf{u}]_{\mathcal{B}}^T\, B\, [\mathbf{u}]_{\mathcal{B}} = x_1^2 + \cdots + x_p^2 > 0$,
- $q(\mathbf{u}) = [\mathbf{u}]_{\mathcal{C}}^T\, C\, [\mathbf{u}]_{\mathcal{C}} = -y_{s+1}^2 - \cdots - y_{s+t}^2 \leq 0$,

a contradiction.

Proof (continued). Reminder: $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$,

- $B = D(\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q})$,
- $C = D(\underbrace{1, \dots, 1}_{s}, \underbrace{-1, \dots, -1}_{t}, \underbrace{0, \dots, 0}_{n-s-t})$.

WTS $p = s$. Suppose otherwise. By symmetry, WMA $p > s$. Now consider $U_B := \text{Span}(\mathbf{b}_1, \dots, \mathbf{b}_p)$ and $U_C := \text{Span}(\mathbf{c}_{s+1}, \dots, \mathbf{c}_n)$. Then $\dim(U_B) + \dim(U_C) = \dim(U_B + U_C) + \dim(U_B \cap U_C)$ (by Prob.4, HW#7, LA1). But $\dim(U_B) + \dim(U_C) = p + (n - s) > n$ and $\dim(U_B + U_C) \leq \dim(\mathbb{R}^n) = n$; so, $\dim(U_B \cap U_C) > 0$, and it follows that $U_B \cap U_C$ contains some non-zero vector $\mathbf{u}$. Set $[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $[\mathbf{u}]_{\mathcal{C}} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$. 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

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

a contradiction. This proves that $p = s$, and we are done. Q.E.D.

Sylvester's law of inertia

For every quadratic form q on $\mathbb{R}^n$, there exists a diagonal matrix $B \in \mathbb{R}^{n \times n}$ with only entries $1, -1, 0$ on the main diagonal, s.t. B is the matrix of q with respect to some basis $\mathcal{B}$ of $\mathbb{R}^n$. Moreover, the matrix B is unique up to a reordering of the main diagonal entries.