Linear Algebra 2: Lecture 22

Irena Penev

Summer 2023

1 The Cholesky decomposition of positive definite matrices

Notation: In this section, $\cdot$ is the standard scalar product on $\mathbb{R}^n$, and $||\cdot||$ is the induced norm.

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 $A = \left[\begin{array}{c} a \end{array}\right]$ in $\mathbb{R}^{1\times 1}$, and we note that $a > 0$ (because A is positive definite). We now set $L := \left[\begin{array}{c} \sqrt{a} \end{array}\right]$, and we observe that $A = LL^T$. The uniqueness of L is obvious.

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 such that $A' - \frac{1}{\alpha}\mathbf{a}\mathbf{a}^T = L'L'^T$. We now set

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

Clearly, L is lower triangular with a positive main diagonal. Moreover, we have that

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

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

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

(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

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

We now have that

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

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,

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

This proves the uniqueness of L. □

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.*

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

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

Now, let $i \in \{1, \dots, n\}$ be the largest index such that $x_i \neq 0$ (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. □

There is also an algorithm that, for a positive definite matrix $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ in $\mathbb{R}^{n\times n}$, computes the Cholesky decomposition of A, i.e. computes the (unique) lower triangular matrix $L = \begin{bmatrix} \ell_{i,j} \end{bmatrix}_{n\times n}$ 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.

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 $i \in \{1, \dots, j - 1\}$,
   - $\ell_{j,j} := \sqrt{a_{j,j} - \sum\limits_{k=1}^{j-1} \ell_{j,k}^2}$,
   - $\ell_{i,j} := \frac{1}{\ell_{j,j}}\Big(a_{i,j} - \sum\limits_{k=1}^{j-1} \ell_{i,k}\ell_{j,k}\Big)$ for all $i \in \{j + 1, \dots, n\}$.

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

Example 1.3. *Find the Cholesky decomposition of the positive definite matrix*

$$A := \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & 1 \\ 4 & 1 & 6 \end{bmatrix}.$$

(The fact that A is positive definite follows from the solution of Example 4.2, 4.4, or 4.6 of Lecture Notes 21.)

Proof. We construct the matrix $L = \left[\ \ell_{i,j}\ \right]_{3\times 3}$ as follows.

1. The first column of L is:
   - $\ell_{1,1} := \sqrt{a_{1,1}} = \sqrt{4} = 2$,
   - $\ell_{2,1} := \frac{a_{2,1}}{\sqrt{a_{1,1}}} = \frac{-2}{\sqrt{4}} = -1$,
   - $\ell_{3,1} := \frac{a_{3,1}}{\sqrt{a_{1,1}}} = \frac{4}{\sqrt{4}} = 2$.
2. The second column of L is:
   - $\ell_{1,2} := 0$,
   - $\ell_{2,2} := \sqrt{a_{2,2} - \ell_{2,1}^2} = \sqrt{10 - (-1)^2} = 3$,
   - $\ell_{3,2} := \frac{1}{\ell_{2,2}}\Big(a_{3,2} - \ell_{3,1}\ell_{2,1}\Big) = \frac{1}{3}\Big(1 - 2\cdot(-1)\Big) = 1$.
3. The third column of L is:
   - $\ell_{1,3} := 0$,
   - $\ell_{2,3} := 0$,
   - $\ell_{3,3} := \sqrt{a_{3,3} - (\ell_{3,1}^2 + \ell_{3,2}^2)} = \sqrt{6 - (2^2 + 1^2)} = 1$.

So, we have that

$$L := \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 2 & 1 & 1 \end{bmatrix}.$$

We see that the matrix L that we obtained is lower triangular with a positive main diagonal, and it is straightforward to check that $A = LL^T$. So, our computation is correct. □

Remark: 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

Notation: As usual, for a field $\mathbb{F}$, $\mathcal{E}_n = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis of the vector space $\mathbb{F}^n$.

In what follows, we will repeatedly use the following proposition, proven in Lecture Notes 21.

Proposition 2.1 of Lecture Notes 21. *Let* $\mathbb{F}$ *be a field. Then for all matrices* $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ *in* $\mathbb{F}^{n \times n}$*, and all vectors* $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ *and* $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$ *in* $\mathbb{F}^n$*, we have that*

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

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. for all $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$, we have that $f(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = f(\mathbf{x}_1, \mathbf{y}) + f(\mathbf{x}_2, \mathbf{y})$;
- b.2. for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$, we have that $f(\alpha\mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$;
- b.3. for all $\mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in V$, we have that $f(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = f(\mathbf{x}, \mathbf{y}_1) + f(\mathbf{x}, \mathbf{y}_2)$;
- b.4. for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{F}$, we have that $f(\mathbf{x}, \alpha\mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$.

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).[1]

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

---

[1] Indeed, if $\langle \cdot, \cdot \rangle$ is a scalar product in a complex vector space V, then it satisfies the following (see subsection 1.2 of Lecture Notes 11):

c.3'. for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{C}$, we have that $\langle \mathbf{x}, \alpha\mathbf{y} \rangle = \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle$.

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. 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; we must show that the bilinear form f_A is symmetric. For this, we simply observe that for all $\mathbf{x},\mathbf{y} \in V$, we have that

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

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. Thus, the bilinear form f_A is indeed 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 $f(\mathbf{e}_2,\mathbf{e}_1) = 3$.)*

As we shall see (see Corollary 2.5 below), bilinear forms from Proposition 2.2 are the only ones that exist for the vector space $\mathbb{F}^n$. First, we prove a theorem relating bilinear forms and bases of finite-dimensional vector spaces.

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 $A = \begin{bmatrix} a_{i,j} \end{bmatrix}_{n\times n}$ 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) for all $\mathbf{x}, \mathbf{y} \in V$*, we have that*

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

(b) the matrix A *is symmetric if and only if 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 such that* $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \, A' \, [\mathbf{y}]_{\mathcal{B}}$ *for all* $\mathbf{x}, \mathbf{y} \in V$*, 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. (a) Fix $\mathbf{x}, \mathbf{y} \in V$, and set $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$ and $[\mathbf{y}]_{\mathcal{B}} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix}^T$, so that $\mathbf{x} = \sum\limits_{i=1}^{n} x_i \mathbf{b}_i$ and $\mathbf{y} = \sum\limits_{i=1}^{n} y_i \mathbf{b}_i$. We now compute:

$$\begin{array}{rcll} f(\mathbf{x}, \mathbf{y}) & = & f(\sum\limits_{i=1}^{n} x_i \mathbf{b}_i, \sum\limits_{j=1}^{n} y_j \mathbf{b}_j) & \\ & = & \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} x_i y_j f(\mathbf{b}_i, \mathbf{b}_j) & \text{because } f \text{ is bilinear} \\ & = & \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} x_i y_j a_{i,j} & \\ & = & [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{y}]_{\mathcal{B}} & \begin{array}{l}\text{by Proposition 2.1} \\ \text{of Lecture Notes 21.}\end{array} \end{array}$$

(b) If the bilinear form f is symmetric, then it follows immediately from the construction of the matrix A that A is also symmetric. Suppose now that the matrix A is symmetric. Let us show that the bilinear from f is symmetric. We proceed as in the proof of Proposition 2.2. For all $\mathbf{x}, \mathbf{y} \in V$, we have that

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

Thus, the bilinear form f is symmetric.

(c) Fix a matrix $A' = \left[\ a'_{i,j} \ \right]_{n\times n}$ in $\mathbb{F}^{n\times n}$, and assume that for all $\mathbf{x}, \mathbf{y} \in V$, we have that $f(\mathbf{x}, \mathbf{y}) = [\mathbf{x}]_{\mathcal{B}}^T \, A' \, [\mathbf{y}]_{\mathcal{B}}$. We must show that $A' = A$. Indeed, for all indices $i, j \in \{1, \dots, n\}$, we have that

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

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

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 Lecture Notes 21. *For any function $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, the following are equivalent:*

- *(i) $\langle \cdot, \cdot \rangle$ is a scalar product on $\mathbb{R}^n$;*
- *(ii) there exists a positive definite matrix $A \in \mathbb{R}^{n\times n}$ such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$.*

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 $\mathbb{F}^{n\times n}$ such that 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 if and only if 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. ☐

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

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

where the $a_{i,j}$ are some scalars in $\mathbb{F}$. This bilinear form is symmetric if and only if $a_{i,j} = a_{j,i}$ for all indices $i, j \in \{1, \dots, n\}$. For example, the following are bilinear forms on $\mathbb{R}^2$:

- $f_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + 7x_2 y_2$;
- $f_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 - 2x_1 y_2 + 3x_2 y_1 - 3x_2 y_2$.

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, $A := \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$ is the matrix of the bilinear form f with respect to the basis $\mathcal{B}$.*

Proof. Set $A := \begin{bmatrix} a_{i,j} \end{bmatrix}_{n \times n}$, 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.

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 $f = g$.) ☐

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*

$$C \;=\; {}_{\mathcal{B}}[Id_V]_{\mathcal{C}}{}^T \, B \, {}_{\mathcal{B}}[Id_V]_{\mathcal{C}}.$$

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 $\mathcal{B}$. But now we have that ${}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \, B \, {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$ 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 $C = {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}{}^T \, B \, {}_{\mathcal{B}}[\mathrm{Id}_V]_{\mathcal{C}}$. ☐

Recall that the *characteristic* of a field $\mathbb{F}$ is the smallest positive integer n (if it exists) such that, in the field $\mathbb{F}$, we have that $\underbrace{1 + \dots + 1}_{n} = 0$ (here, the 1's and 0 are understood to be in the field $\mathbb{F}$). 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 $\text{char}(\mathbb{Z}_p) = p$.

Theorem 2.8. *The characteristic of any field is either a prime number or* 0.

Proof. Let $\mathbb{F}$ be a field. We may assume that $\text{char}(\mathbb{F}) \neq 0$, for otherwise we are done. So, $\text{char}(\mathbb{F})$ is a positive integer. By the definition of a field, we have that $1 \neq 0$,[2] and so $\text{char}(\mathbb{F}) \geq 2$. Now, suppose that $\text{char}(\mathbb{F})$ is not a prime, and fix integers $p, q \geq 2$ such that $\text{char}(\mathbb{F}) = pq$. Then

$$\Big(\underbrace{1 + \dots + 1}_{p}\Big)\Big(\underbrace{1 + \dots + 1}_{q}\Big) \quad = \quad \underbrace{1 + \dots + 1}_{pq} \quad = \quad 0.$$

Since $\mathbb{F}$ is a field, it follows that at least one of $\underbrace{1 + \dots + 1}_{p}$ and $\underbrace{1 + \dots + 1}_{q}$ is zero. But this is impossible since $0 < p, q < \text{char}(\mathbb{F})$. $\square$

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. Fix $\mathbf{x}, \mathbf{y} \in V$. We must show that $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. If $\mathbf{x} = \mathbf{y}$, then this is true by hypothesis. So, assume that $\mathbf{x} \neq \mathbf{y}$. By hypothesis, we have that

(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 $f(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$. $\square$

---

[2] Here, 1 and 0 are understood to be in the field $\mathbb{F}$.

3 Quadratic forms

A *quadratic form* on a vector space V over a field $\mathbb{F}$ is a function $q : V \to \mathbb{F}$ such that 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 such that 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 such that 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

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

which is what we needed.

It remains to show that f is unique. Fix any symmetric bilinear form f' on V such that 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 for all $\mathbf{x} \in V$, we have that $f'(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. ☐

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 $A \in \mathbb{F}^{n \times n}$ such that 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. By Theorem 3.1, there exists a unique symmetric bilinear form f on V such that for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x})$. Let A be the matrix of the bilinear form f with respect to the basis $\mathcal{B}$; then $q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A \, [\mathbf{x}]_{\mathcal{B}}$ for all $\mathbf{x} \in V$. Moreover, by Theorem 2.4(c), the matrix A is symmetric.

It remains to prove the uniqueness of A. Fix any symmetric matrix $A' \in \mathbb{F}^n$ such that for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A' \, [\mathbf{x}]_{\mathcal{B}}$. Using Corollary 2.6, we let f' be the unique bilinear form such that for all indices $i, j \in \{1, \dots, n\}$, we have that $f'(\mathbf{b}_i, \mathbf{b}_j) = a_{i,j}$. Then by Theorem 2.4, A' is the matrix of the bilinear form f'. Since the matrix A' is symmetric, Theorem 2.4(b) guarantees that the bilinear form f' is also symmetric. But now for all $\mathbf{x} \in V$, we have that $q(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}^T \, A' \, [\mathbf{x}]_{\mathcal{B}} = f'(\mathbf{x}, \mathbf{x})$. By the uniqueness part of Theorem 3.1, we have that $f' = f$; since A' and A are, respectively, the matrices of f' and f with respect to the basis $\mathcal{B}$, we deduce that $A' = A$. $\square$

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*

$$C \quad = \quad {}_{\mathcal{B}}[Id_V]_{\mathcal{C}}{}^T \; B \; {}_{\mathcal{B}}[Id_V]_{\mathcal{C}}.$$

Proof. By Theorem 3.1, there exists a unique symmetric bilinear form f on V such that 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. $\square$

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

$$\mathbf{x}^T D\mathbf{x} \quad = \quad a_1x_1^2 + \dots + a_nx_n^2.$$

This is a particularly nice formula, and for this reason, Sylvester's law of inertia (see below) is of interest. But first, we prove a simple proposition that we will use in our proof of Sylvester's law of inertia.

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 $rank(SA) = rank(A)$;

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

(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 $rank(S_1AS_2) = rank(A)$.

Proof. We first prove (a). Fix an invertible matrix $S \in \mathbb{F}^{n\times n}$. Then S is a product of elementary matrices,[3] and consequently, SA is row equivalent to S.[4] So, by Proposition 3.3 of Lecture Notes 7, we have that $\text{Row}(SA) = \text{Row}(A)$. Since the rank of a matrix is equal to the dimension of its row space

---

[3] This follows from Proposition 3.3 of Lecture Notes 4.

[4] This is because multiplying a matrix M by an elementary matrix on the left has the effect of performing an elementary row operation on M.

(by Theorem 3.4 of Lecture Notes 7), it follows that $\text{rank}(SA) = \text{rank}(A)$. This proves (a).

We now prove (b). Fix an invertible matrix $S \in \mathbb{F}^{m\times m}$. By Corollary 4.2(a) of Lecture Notes 7, we know that the rank of any matrix is equal to the rank of its transpose. Moreover, by Corollary 4.2(b) of Lecture Notes 7, we know that S^T is invertible (because S is invertible). But now

$$\begin{aligned}\text{rank}(AS) &= \text{rank}\Big((AS)^T\Big)\\ &= \text{rank}(S^TA^T)\\ &\overset{(a)}{=} \text{rank}(A^T)\\ &= \text{rank}(A).\end{aligned}$$

This proves (b).

Finally, for (c), we fix invertible matrices $S_1 \in \mathbb{F}^{n\times n}$ and $S_2 \in \mathbb{F}^{m\times m}$, and we observe that

$$\text{rank}(S_1AS_2) \overset{(a)}{=} \text{rank}(AS_2) \overset{(b)}{=} \text{rank}(A),$$

and we are done. □

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, such that 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.

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 $D = Q^TAQ$. Next, for all $i \in \{1, \dots, n\}$, set $\ell_i := \frac{1}{\sqrt{|\lambda_i|}}$ 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,[5] 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 ${}_{\mathcal{E}_n}[\text{Id}_V]_{\mathcal{B}} = R$, and so by Corollary 3.3, R^TAR is the matrix of the

---

[5] Since Q is orthogonal, it is invertible (by Theorem 2.1 of Lecture Notes 14). On the other hand, L is a diagonal matrix, and all its entries on the main diagonal are non-zero; so, L is invertible with inverse $L^{-1} = D(\frac{1}{\ell_1}, \dots, \frac{1}{\ell_n})$.

quadratic form q with respect to the basis $\mathcal{B}$. It remains to show that the matrix R^TAR is diagonal, and that all its entries on the main diagonal are 1, -1, or 0. We compute:

$$\begin{aligned} R^TAR &= (QL)^TA(QL) \\ &= L^T\underbrace{Q^TAQ}_{=D}L \\ &= L^TDL \\ &= LDL && \text{because } L \text{ is diagonal} \\ &= D(\ell_1,\ldots,\ell_n)\,D(\lambda_1,\ldots,\lambda_n)\,D(\ell_1,\ldots,\ell_n) \\ &= D(\lambda_n\ell_n^2,\ldots,\lambda_n\ell_n^2). \end{aligned}$$

But by construction, we have that

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

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

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. We must show that the matrices B and C are identical up to a reordering of the main diagonal entries.

In view of Theorem 2.4,[6] and after possibly reordering the basis elements of $\mathcal{B}$ and $\mathcal{C}$, we may assume that

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

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

---

[6] In particular, we are using the fact the formula for a bilinear form with respect to a basis.

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,[7] and consequently, its transpose S^T is also invertible.[8] 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 $p+q = s+t$.

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

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

(this follows from Problem 4 of HW#7 from Linear Algebra 1, winter 2022). 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. $\square$

---

[7] This follows from Proposition 2.4 of Lecture Notes 10, or alternatively, from Proposition 4.1 of Lecture Notes 10.

[8] This follows from Corollary 4.2(b) of Lecture Notes 7.