

Exercises

8.1.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Any projection onto a linear subspace in an inner product space minimizes norms.

ii. When the vector $\boldsymbol{v}$ is a solution to the linear equation $T[\boldsymbol{x}] = \boldsymbol{b}$, the orthogonal projection of $T[\boldsymbol{v}]$ onto the image of T equals $\boldsymbol{b}$.

iii. Taylor polynomials always provide the best approximation to a function.

iv. The defining basis for the space of trigonometric polynomials is an orthonormal basis.

8.1.7 Problem. Fix a nonnegative integer n and consider the n-element set $\mathcal{X} := \left\{ \frac{2\pi\ell}{n} \in \mathbb{R} \;\middle|\; 0 \leqslant \ell \leqslant n-1 \right\}$. Let $V := \mathbb{C}^{\mathcal{X}}$ the complex inner product space, consisting of all functions from the finite set $\mathcal{X}$ of real numbers to $\mathbb{C}$ with the inner product

$$\langle f, g \rangle := \sum_{x \in \mathcal{X}} f(x)\, \overline{g(x)} = \sum_{\ell=0}^{n-1} f\big(\tfrac{2\pi\ell}{n}\big)\, \overline{g\big(\tfrac{2\pi\ell}{n}\big)}\,.$$

i. For all integers j satisfying $0 \leqslant j \leqslant n-1$, demonstrate that the functions $w_j(x) := \exp(-j\,x\,\mathrm{i})$ are pairwise orthogonal and compute $\|w_j(x)\|$.

ii. For all integers k satisfying $0 \leqslant k \leqslant n-1$, consider the function

$$h_k(x) := \begin{cases} 1 & \text{if } x = \frac{2\pi k}{n} \\ 0 & \text{if } x \neq \frac{2\pi k}{n}. \end{cases}$$

Which function in the linear subspace

$$W := \operatorname{Span}\big(w_0(x), w_1(x), \dots, w_{n-1}(x)\big) \subset V$$

best approximates the function $h_k(x)$?

iii. For all integers k satisfying $0 \leqslant k \leqslant n-1$, calculate the norm of the different between $h_k(x)$ and its best approximate.

8.2 Least-Squares

HOW DO WE FIND THE BEST APPROXIMATE SOLUTION? Auspiciously, the optimal approximate solutions to an inconsistent linear system are the solutions to an auxiliary consistent linear system.

8.2.0 Proposition. *Let* $\mathbf{A}$ *be a complex* $(m \times n)$*-matrix and let* $\boldsymbol{b}$ *be a vector in* $\mathbb{C}^m$*. The set of least-squares approximations to* $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ *coincides with the nonempty set of solutions to the* **normal equations** $\mathbf{A}^\star\,\mathbf{A}\,\boldsymbol{x} = \mathbf{A}^\star\,\boldsymbol{b}$.

Proof. Let P be the orthogonal projection onto the column space $\operatorname{Im}(\mathbf{A})$ of the matrix $\mathbf{A}$. Since orthogonal projections minimize norms [8.1.0] and the vector $\mathbf{A}\,\boldsymbol{x}$ lies in $\operatorname{Im}(\mathbf{A}) = \operatorname{Im}(P)$, we have $\|\boldsymbol{b} - P[\boldsymbol{b}]\| \leqslant \|\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}\|$ with equality if and only if $\mathbf{A}\,\boldsymbol{x} = P[\boldsymbol{b}]$. Hence, a vector $\boldsymbol{x}$ in $\mathbb{C}^n$ is a least-squares approximation of a solution to the linear system $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ if and only if the vector $\boldsymbol{x}$ is a solution to $\mathbf{A}\,\boldsymbol{x} = P[\boldsymbol{b}]$. To show that the set of least-squares approximations to $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ coincides with the set of solutions to the normal equations, we prove containment in both directions.

$\subseteq$: Suppose that the vector $\boldsymbol{x}$ in $\mathbb{C}^n$ satisfies $\mathbf{A}\,\boldsymbol{x} = P[\boldsymbol{b}]$. Since $P^2 = P$, we see that $\boldsymbol{b} - P[\boldsymbol{b}]$ lies in $\operatorname{Ker}(P)$. The orthogonality [8.0.3] of the projection P implies that $\langle \boldsymbol{b} - P[\boldsymbol{b}], \boldsymbol{w} \rangle = \boldsymbol{w}^\star(\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}) = 0$ for all $\boldsymbol{w}$ in $\operatorname{Im}(\mathbf{A})$. For all $1 \leqslant k \leqslant n$, the k-th column vector $\boldsymbol{a}_k$ in the matrix $\mathbf{A}$ satisfies $\boldsymbol{a}_k^\star(\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}) = 0$, which means that the vector $\boldsymbol{x}$ in $\mathbb{C}^n$ satisfies $\mathbf{A}^\star\,\mathbf{A}\,\boldsymbol{x} = \mathbf{A}^\star\,\boldsymbol{b}$.

$\supseteq$: Suppose that the vector $\boldsymbol{x}$ in $\mathbb{C}^n$ satisfies $\mathbf{A}^\star\,\mathbf{A}\,\boldsymbol{x} = \mathbf{A}^\star\,\boldsymbol{b}$. Since $\mathbf{A}^\star\,(\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}) = \mathbf{0}$, we see that $\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}$ is orthogonal to the columns of $\mathbf{A}$ and the vector $\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x}$ lies in the kernel $\operatorname{Ker}(P)$. As the vector $\mathbf{A}\,\boldsymbol{x}$ lies in $\operatorname{Im}(P)$ and $\boldsymbol{b} = \mathbf{A}\,\boldsymbol{x} + (\boldsymbol{b} - \mathbf{A}\,\boldsymbol{x})$, the properties of projections [8.0.2] establish that $\mathbf{A}\,\boldsymbol{x} = P[\boldsymbol{b}]$. □

8.2.1 Problem. Find a least-square approximation to $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ where

$$\mathbf{A} := \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{b} := \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

Solution. Since

$$\mathbf{A}^\star\,\mathbf{A} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}, \quad \mathbf{A}^\star\,\boldsymbol{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix},$$

the normal equation $\mathbf{A}^\star\,\mathbf{A}\,\boldsymbol{x} = \mathbf{A}^\star\,\boldsymbol{b}$ becomes $\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$, so

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \qquad \square$$

8.2.2 Problem. Find a least-square approximation to $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ where

$$\mathbf{A} := \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}^\mathsf{T} \qquad \text{and} \qquad \boldsymbol{b} := \begin{bmatrix} -3 & -1 & 0 & 2 & 5 & 1 \end{bmatrix}^\mathsf{T}.$$

Solution. Since

$$\mathbf{A}^\star\,\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}, \quad \mathbf{A}^\star\,\boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix},$$

the augmented matrix for $\mathbf{A}^\star \mathbf{A}\, \boldsymbol{x} = \mathbf{A}^\star \boldsymbol{b}$ is

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \xrightarrow[\sim]{\substack{r_1 \mapsto 0.5\, r_1 \\ r_2 \mapsto 0.5\, r_2 \\ r_3 \mapsto 0.5\, r_3 \\ r_4 \mapsto 0.5\, r_4}} \begin{bmatrix} 3 & 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 & -2 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 3 \end{bmatrix} \xrightarrow[\sim]{\substack{r_1 \mapsto r_4 \\ r_4 \mapsto r_1}} \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 1 & 1 & 0 & 0 & -2 \\ 1 & 0 & 1 & 0 & 1 \\ 3 & 1 & 1 & 1 & 2 \end{bmatrix} \xrightarrow[\sim]{\substack{r_2 \mapsto r_2 - r_1 \\ r_3 \mapsto r_3 - r_1 \\ r_4 \mapsto r_4 - 3\, r_1}} \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 1 & 1 & -2 & -7 \end{bmatrix}$$

$$\xrightarrow[\sim]{r_4 \mapsto r_4 - r_2} \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & -1 & -2 \end{bmatrix} \xrightarrow[\sim]{r_4 \mapsto r_4 - r_3} \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The general solution is $\boldsymbol{x} = \begin{bmatrix} 3 & -5 & -2 & 0 \end{bmatrix}^\mathsf{T} + \operatorname{Span}\big(\begin{bmatrix} -1 & 1 & 1 & 1 \end{bmatrix}^\mathsf{T}\big)$. □

The approach also yields the matrix associated to the orthogonal projection onto any linear subspace relative to the standard basis.

8.2.3 Lemma. *For any complex $(m \times n)$-matrix $\mathbf{B}$ with linearly independent columns, the product $\mathbf{B}^\star \mathbf{B}$ is an invertible $(n \times n)$-matrix.*

Proof. For any vector $\boldsymbol{v}$ in $\mathbb{C}^n$ satisfying $\mathbf{B}^\star \mathbf{B}\, \boldsymbol{v} = \mathbf{0}$, we have

$$\|\mathbf{B}\, \boldsymbol{v}\|^2 = (\mathbf{B}\, \boldsymbol{v})^\star (\mathbf{B}\, \boldsymbol{v}) = \boldsymbol{v}^\star\, \mathbf{B}^\star\, \mathbf{B}\, \boldsymbol{v} = \boldsymbol{v}^\star\, \mathbf{0} = \mathbf{0}\,,$$

so the properties of norms [7.1.1] establish that $\mathbf{B}\, \boldsymbol{v} = \mathbf{0}$. Since the columns of the matrix $\mathbf{B}$ are linearly independent, we deduce that $\boldsymbol{v} = \mathbf{0}$. The characterizations of invertible matrices establish that the product $\mathbf{B}^\star \mathbf{B}$ is invertible. □

8.2.4 Proposition. *Fix positive integers m and n. Let $\mathbf{B}$ be a complex $(m \times n)$-matrix whose columns form a basis for a linear subspace W in $\mathbb{C}^m$ and let $P \colon \mathbb{C}^m \to \mathbb{C}^m$ be the orthogonal projection onto W. For the standard basis $\mathcal{E} := (\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_m)$ of $\mathbb{C}^m$, we have $(P)^{\mathcal{E}}_{\mathcal{E}} = \mathbf{B}\, (\mathbf{B}^\star\, \mathbf{B})^{-1}\, \mathbf{B}^\star$.*

Proof. Consider a vector $\boldsymbol{v}$ in $\mathbb{C}^m$ and set $\boldsymbol{w} := P[\boldsymbol{v}]$. Since $P^2 = P$, it follows that the vector $\boldsymbol{v} - \boldsymbol{w}$ lies in $\operatorname{Ker}(P)$. The orthogonality [8.0.3] of the projection P implies that $\mathbf{B}^\star(\boldsymbol{v} - \boldsymbol{w}) = \mathbf{0}$. Since the columns of matrix $\mathbf{B}$ span the linear subspace W, there exists a vector $\boldsymbol{x}$ in $\mathbb{C}^n$ such that $\boldsymbol{w} = \mathbf{B}\, \boldsymbol{x}$, so we have $\mathbf{0} = \mathbf{B}^\star\, (\boldsymbol{v} - \boldsymbol{w}) = \mathbf{B}^\star\, \boldsymbol{v} - \mathbf{B}^\star\, \mathbf{B}\, \boldsymbol{x}$ and $\mathbf{B}^\star\, \mathbf{B}\, \boldsymbol{x} = \mathbf{B}^\star\, \boldsymbol{v}$. The invertibility of the matrix $\mathbf{B}^\star\, \mathbf{B}$ establishes that $\boldsymbol{x} = (\mathbf{B}^\star\, \mathbf{B})^{-1}\, \mathbf{B}^\star\, \boldsymbol{v}$ and $P[\boldsymbol{v}] = \boldsymbol{w} = \mathbf{B}(\mathbf{B}^\star\, \mathbf{B})^{-1} \mathbf{B}^\star\, \boldsymbol{v}$. Thus, for all $1 \leqslant k \leqslant m$, the k-th column of the matrix $\mathbf{B}(\mathbf{B}^\star\, \mathbf{B})^{-1} \mathbf{B}^\star$ equals $P[\boldsymbol{e}_k]$. We conclude that $(P)^{\mathcal{E}}_{\mathcal{E}} = \mathbf{B}\, (\mathbf{B}^\star\, \mathbf{B})^{-1}\, \mathbf{B}^\star$. □

Exercises

8.2.5 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Every linear system has a unique approximate solution.

ii. The normal equations always have a solution.

iii. If the columns of the coefficient matrix are linearly independent, then the associated linear system always has a unique approximate solution.

iv. Given any complex matrix $\mathbf{B}$, the associated matrix $\mathbf{B}^\star \, \mathbf{B}$ is always invertible.

8.2.6 Problem. The population of Canada, as determined by the Canadian census, was as follows:

year	1996	2001	2006	2011	2016
population (in millions)	28.8	30.0	31.6	33.5	35.2

Let t denote the time measured in years from 1996.

i. Suppose that population of Canada (measure in millions) is modeled by the linear function $p_\ell(t) = a + b\,t$. Find the least-squares estimates for the parameters a and b.

ii. Suppose that the population of Canada (measure in millions) is modeled by the exponential function $p_e(t) = c\,e^{\lambda\,t}$. Linearize the model and use the least-squares method to estimate the parameters c and λ.

8.2.7 Problem. Consider

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{bmatrix} 5 \\ 3 \\ -1 \\ 1 \end{bmatrix}.$$

Show that a least-squares solution to $\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}$ is not unique and solve the normal equations to find all of the least-squares solutions.

9
Geometry of Operators

The compatibility between a linear map and the inner products on its source and target has wonderful consequences.

9.0 Adjoint Maps

How do inner products give rise to new maps? Each linear map between finite-dimensional inner product spaces comes with a companion. Before describing this partner, we use inner products to recognize equality of vectors and introduce a special class of maps.

9.0.0 Lemma (Equality test)**.** *Let V be an inner product space. For any two vectors $\boldsymbol{u}$ and $\boldsymbol{w}$ in V, we have $\boldsymbol{u} = \boldsymbol{w}$ if and only if $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}$ in V.*

Proof. The conjugate-linearity [7.0.9] of inner products shows that $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ if and only if $0 = \langle \boldsymbol{v}, \boldsymbol{u} \rangle - \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{w} \rangle$.

$\Rightarrow$: Suppose that $\boldsymbol{u} = \boldsymbol{w}$. The properties [7.0.9] of inner products show that $\langle \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{0} \rangle = 0$.

$\Leftarrow$: Suppose that, for all vectors $\boldsymbol{v}$ in V, we have $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$. Setting $\boldsymbol{v} := \boldsymbol{u} - \boldsymbol{w}$, we obtain $0 = \langle \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{w} \rangle = \langle \boldsymbol{u} - \boldsymbol{w}, \boldsymbol{u} - \boldsymbol{w} \rangle$ and the positivity [7.0.0] of inner products implies that $\boldsymbol{u} - \boldsymbol{w} = \boldsymbol{0}$. □

9.0.1 Definition. A ***linear functional*** is a linear map from a $\mathbb{K}$-vector space to its underlying field $\mathbb{K}$ of scalars.

When a vector space is equipped with an inner product, linear functionals are easily characterized.

9.0.2 Proposition (Representation of linear functionals)**.** *Let V be a finite-dimensional inner product space over the field $\mathbb{K}$ of scalars. For any linear functional $\varphi \colon V \to \mathbb{K}$, there exists a unique vector $\boldsymbol{v}$ in V such that $\varphi[\boldsymbol{w}] = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all vectors $\boldsymbol{w}$ in V.*

Proof. Set $n := \dim V$. As n is a nonnegative integer, we may choose an orthonormal basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ for V; see [7.2.2]. For any vector $\boldsymbol{w}$ in V, orthonormal coordinates [7.1.7] on V, the linearity of φ, and the conjugate-linearity [7.0.9] of the inner product give

$$\begin{aligned}
\varphi[\boldsymbol{w}] &= \varphi[\langle \boldsymbol{w}, \boldsymbol{u}_1 \rangle \, \boldsymbol{u}_1 + \langle \boldsymbol{w}, \boldsymbol{u}_2 \rangle \, \boldsymbol{u}_2 + \cdots + \langle \boldsymbol{w}, \boldsymbol{u}_n \rangle \, \boldsymbol{u}_n] \\
&= \langle \boldsymbol{w}, \boldsymbol{u}_1 \rangle \, \varphi[\boldsymbol{u}_1] + \langle \boldsymbol{w}, \boldsymbol{u}_2 \rangle \, \varphi[\boldsymbol{u}_2] + \cdots + \langle \boldsymbol{w}, \boldsymbol{u}_n \rangle \, \varphi[\boldsymbol{u}_n] \\
&= \left\langle \boldsymbol{w}, \overline{\varphi[\boldsymbol{u}_1]} \, \boldsymbol{u}_1 + \overline{\varphi[\boldsymbol{u}_2]} \, \boldsymbol{u}_2 + \cdots + \overline{\varphi[\boldsymbol{u}_n]} \, \boldsymbol{u}_n \right\rangle .
\end{aligned}$$

By setting $\boldsymbol{v} := \overline{\varphi[\boldsymbol{u}_1]}\, \boldsymbol{u}_1 + \overline{\varphi[\boldsymbol{u}_2]}\, \boldsymbol{u}_2 + \cdots + \overline{\varphi[\boldsymbol{u}_n]}\, \boldsymbol{u}_n$, we see that there exists a vector $\boldsymbol{v}$ in V such that $\varphi[\boldsymbol{w}] = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all vectors $\boldsymbol{w}$ in V.

Suppose that there exists vectors $\boldsymbol{v}$ and $\widetilde{\boldsymbol{v}}$ in V such that, for all vectors $\boldsymbol{w}$ in V, we have $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \langle \boldsymbol{w}, \widetilde{\boldsymbol{v}} \rangle$. The equality test proves that $\boldsymbol{v} = \widetilde{\boldsymbol{v}}$. Thus, there is a unique vector $\boldsymbol{v}$ in V such that $\varphi[\boldsymbol{w}] = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all vectors $\boldsymbol{w}$ in V. □

9.0.3 Definition. Let V and W be finite-dimensional inner product spaces and let $T \colon V \to W$ be a linear map. Conceptually, the ***adjoint map*** $T^\star \colon W \to V$ is determined, for all vectors $\boldsymbol{v}$ in V and all vectors $\boldsymbol{w}$ in W, by the equation $\langle T[\boldsymbol{v}], \boldsymbol{w} \rangle_W = \langle \boldsymbol{v}, T^\star[\boldsymbol{w}] \rangle_V$. More pedantically, each vector $\boldsymbol{w}$ in W yields a linear functional $\varphi_{\boldsymbol{w}} \colon V \to \mathbb{K}$ defined, for all vectors $\boldsymbol{u}$ in V, by $\varphi_{\boldsymbol{w}}[\boldsymbol{u}] := \langle T[\boldsymbol{u}], \boldsymbol{w} \rangle_W$. The representation of linear functionals implies that there is a unique vector $\boldsymbol{v}$ in V such that $\varphi_{\boldsymbol{w}}[\boldsymbol{u}] = \langle \boldsymbol{u}, \boldsymbol{v} \rangle_V$ for all vectors $\boldsymbol{u}$ in V. The map $T^\star \colon W \to V$ is defined by $T^\star[\boldsymbol{w}] = \boldsymbol{v}$.

Although their definition is somewhat opaque, the properties of adjoint maps are straightforward and instinctual.

9.0.4 Proposition (Properties of adjoints)**.** *Let U, V, and W be three finite-dimensional inner product spaces over the same field of scalars. The adjoint operation has the following properties:*

(linearity) *For any linear map $T \colon V \to W$, the adjoint map $T^\star \colon W \to V$ is linear.*
(conjuguate-linearity) *For any linear maps $S, T \colon V \to W$ and any scalars b, c, we have $(b\,S + c\,T)^\star = \overline{b}\,S^\star + \overline{c}\,T^\star$.*
(involution) *For any linear map $T \colon V \to W$, we have $(T^\star)^\star = T$.*
(identity) *We have $\mathrm{id}_W^\star = \mathrm{id}_W$.*
(multiplicativity) *For any linear maps $S \colon U \to V$ and $T \colon V \to W$, we have $(T\,S)^\star = S^\star\,T^\star$.*

Proof. Throughout the proof, let $\boldsymbol{v}$ be a vector in V, let $\boldsymbol{w}$ and $\boldsymbol{x}$ be vectors in W, and let b and c be scalars.

(*linearity*) The definition of the adjoint map and the conjugate-linearity [7.0.9] of the inner product give

$$\begin{aligned}\langle \boldsymbol{v}, T^\star[b\,\boldsymbol{w} + c\,\boldsymbol{x}] \rangle_V = \langle T[\boldsymbol{v}], b\,\boldsymbol{w} + c\,\boldsymbol{x} \rangle_W &= \overline{b}\,\langle T[\boldsymbol{v}], \boldsymbol{w} \rangle_W + \overline{c}\,\langle T[\boldsymbol{v}], \boldsymbol{x} \rangle_W \\ &= \overline{b}\,\langle \boldsymbol{v}, T^\star[\boldsymbol{w}] \rangle_V + \overline{c}\,\langle \boldsymbol{v}, T^\star[\boldsymbol{x}] \rangle_V = \langle \boldsymbol{v}, b\,T^\star[\boldsymbol{w}] + c\,T^\star[\boldsymbol{x}] \rangle_V \,.\end{aligned}$$

The equality test implies that $T^\star[b\,\boldsymbol{w} + c\,\boldsymbol{x}] = b\,T^\star[\boldsymbol{w}] + c\,T^\star[\boldsymbol{x}]$, so the adjoint map $T^\star \colon W \to V$ is linear.

(*conjugate-linearity*) The definition of the adjoint map, the pointwise operations [1.1.0] on linear maps, linearity [7.0.0] of inner products, and conjugate-linearity [7.0.9] of inner products give

$$\begin{aligned}\langle \boldsymbol{v}, (b\,S + c\,T)^\star[\boldsymbol{w}] \rangle_V &= \langle (b\,S + c\,T)[\boldsymbol{v}], \boldsymbol{w} \rangle_W = \langle b\,S[\boldsymbol{v}] + c\,T[\boldsymbol{v}], \boldsymbol{w} \rangle_W \\ &= b\,\langle S[\boldsymbol{v}], \boldsymbol{w} \rangle_W + c\,\langle T[\boldsymbol{v}], \boldsymbol{w} \rangle_W = b\,\langle \boldsymbol{v}, S^\star[\boldsymbol{w}] \rangle_V + c\,\langle \boldsymbol{v}, T^\star[\boldsymbol{w}] \rangle_V = \langle \boldsymbol{v}, \overline{b}\,S^\star[\boldsymbol{w}] + \overline{c}\,T^\star[\boldsymbol{w}] \rangle_V \,.\end{aligned}$$

The equality test implies that $(b\,S + c\,T)^\star[\boldsymbol{w}] = \overline{b}\,S^\star[\boldsymbol{w}] + \overline{c}\,T^\star[\boldsymbol{w}]$, so we have $(b\,S + c\,T)^\star = \overline{b}\,S^\star + \overline{c}\,T^\star$.

(*involution*) The definition of the adjoint map and the conjugate-symmetry [7.0.0] of the inner product give

$$\langle T[\boldsymbol{v}], \boldsymbol{w}\rangle_W = \langle \boldsymbol{v}, T^\star[\boldsymbol{w}]\rangle_V = \overline{\langle T^\star[\boldsymbol{w}], \boldsymbol{v}\rangle_V} = \overline{\langle \boldsymbol{w}, (T^\star)^\star[\boldsymbol{v}]\rangle_W} = \langle (T^\star)^\star[\boldsymbol{v}], \boldsymbol{w}\rangle_W \,.$$

The equality test yields $T[\boldsymbol{v}] = (T^\star)^\star[\boldsymbol{v}]$, so $T = (T^\star)^\star$.

(*identity*) The definition of the adjoint map gives

$$\langle \boldsymbol{w}, \mathrm{id}_W^\star[\boldsymbol{x}]\rangle_W = \langle \mathrm{id}_W[\boldsymbol{w}], \boldsymbol{x}\rangle_W = \langle \boldsymbol{w}, \boldsymbol{x}\rangle_W = \langle \boldsymbol{w}, \mathrm{id}_W[\boldsymbol{x}]\rangle_W \,.$$

The equality test yields $\mathrm{id}_W^\star[\boldsymbol{x}] = \mathrm{id}_W[\boldsymbol{x}]$, so $\mathrm{id}_W^\star = \mathrm{id}_W$.

(*multiplicativity*) For any vector $\boldsymbol{u}$ in U, the definition of the adjoint map gives

$$\langle \boldsymbol{u}, (T\,S)^\star[\boldsymbol{w}]\rangle_U = \langle (T\,S)[\boldsymbol{u}], \boldsymbol{w}\rangle_W = \big\langle T\big[S[\boldsymbol{u}]\big], \boldsymbol{w}\big\rangle_W = \langle S[\boldsymbol{u}], T^\star[\boldsymbol{w}]\rangle_V = \big\langle \boldsymbol{u}, S^\star\big[T^\star[\boldsymbol{w}]\big]\big\rangle_U = \langle \boldsymbol{u}, (S^\star\,T^\star)[\boldsymbol{w}]\rangle_U \,.$$

Since the equality test implies that $(T\,S)^\star[\boldsymbol{w}] = (S^\star\,T^\star)[\boldsymbol{w}]$, we conclude that $(T\,S)^\star = S^\star\,T^\star$. □

The next problem shows that our notation for the adjoint map is consistent with our notation for the conjugate-transpose [7.0.2].

9.0.5 Problem. Let $T \colon V \to V$ be a linear map on a finite-dimensional inner product space V. For any orthonormal basis $\mathcal{U} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n)$ of V, the matrix of T^* relative to $\mathcal{U}$ is the conjugate-transpose of the matrix of T relative to $\mathcal{U}$, or equivalently $(T^\star)^{\mathcal{U}}_{\mathcal{U}} = \big((T)^{\mathcal{U}}_{\mathcal{U}}\big)^\star$.

Proof. For all $1 \leqslant k \leqslant n$, the orthonormal coordinates [7.1.7] relative to $\mathcal{U}$ give $T[\boldsymbol{u}_k] = \langle T[\boldsymbol{u}_k], \boldsymbol{u}_1\rangle\, \boldsymbol{u}_1 + \langle T[\boldsymbol{u}_k], \boldsymbol{u}_2\rangle\, \boldsymbol{u}_2 + \cdots + \langle T[\boldsymbol{u}_k], \boldsymbol{u}_n\rangle\, \boldsymbol{u}_n$, so the (j,k)-entry in the matrix $(T)^{\mathcal{U}}_{\mathcal{U}}$ is $\langle T[\boldsymbol{u}_k], \boldsymbol{u}_j\rangle$. In the same way, the properties of inner products [7.0.9] and the definition of the adjoint map [9.0.3] also give

$$\begin{aligned} T^\star[\boldsymbol{u}_k] &= \langle T^\star[\boldsymbol{u}_k], \boldsymbol{u}_1\rangle\, \boldsymbol{u}_1 + \langle T^\star[\boldsymbol{u}_k], \boldsymbol{u}_2\rangle\, \boldsymbol{u}_2 + \cdots + \langle T^\star[\boldsymbol{u}_k], \boldsymbol{u}_n\rangle\, \boldsymbol{u}_n \\ &= \overline{\langle \boldsymbol{u}_1, T^\star[\boldsymbol{u}_k]\rangle}\, \boldsymbol{u}_1 + \overline{\langle \boldsymbol{u}_2, T^\star[\boldsymbol{u}_k]\rangle}\, \boldsymbol{u}_2 + \cdots + \overline{\langle \boldsymbol{u}_n, T^\star[\boldsymbol{u}_k]\rangle}\, \boldsymbol{u}_n \\ &= \overline{\langle T[\boldsymbol{u}_1], \boldsymbol{u}_k\rangle}\, \boldsymbol{u}_1 + \overline{\langle T[\boldsymbol{u}_2], \boldsymbol{u}_k\rangle}\, \boldsymbol{u}_2 + \cdots + \overline{\langle T[\boldsymbol{u}_n], \boldsymbol{u}_k\rangle}\, \boldsymbol{u}_n \,, \end{aligned}$$

so the (j,k)-entry in the matrix $(T^\star)^{\mathcal{U}}_{\mathcal{U}}$ is $\overline{\langle T[\boldsymbol{u}_j], \boldsymbol{u}_k\rangle}$. Comparing entries, we conclude that $(T^\star)^{\mathcal{U}}_{\mathcal{U}} = \big((T)^{\mathcal{U}}_{\mathcal{U}}\big)^\star$. □

Exercises

9.0.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. On a finite-dimensional vector space, there is a bijection between linear functionals and vectors.

ii. Every linear operator is equal to its adjoint.

iii. An orthogonal projection is equal to its adjoint.

9.0.7 Problem. Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ be an orthonormal basis for an inner product space V. For all $1 \leqslant j \leqslant n$, consider the linear functional $\psi_j \colon V \to \mathbb{K}$ defined by $\psi_j[\boldsymbol{u}_k] = \delta_{j,k}$ for all $1 \leqslant k \leqslant n$. Show that the list $(\psi_1, \psi_2, \dots, \psi_n)$ forms a basis for $V^\star := \operatorname{Hom}(V, \mathbb{K})$.

9.0.8 Problem. Let V denote the real vector space of continuous real-valued functions on the interval $[a, b]$ with $\langle f, g \rangle := \int_a^b f(s)\, g(s)\; ds$. Fix a continuous function $K \colon \mathbb{R}^2 \to \mathbb{R}$. Consider the linear operator $J \colon V \to V$ defined by $\big(J[f]\big)(x) := \int_a^b K(s,x)\, f(s)\; ds$. Show that the adjoint $J^\star$ exists.

9.0.9 Problem. Let V be a finite-dimensional $\mathbb{K}$-vector space, let $V^\star := \operatorname{Hom}(V, \mathbb{K})$ be its dual space, and let $V^{\star\star} := \operatorname{Hom}(V^\star, \mathbb{K})$ be its double dual space.

i. For any vector $\boldsymbol{v}$ in V, the map $\widehat{\boldsymbol{v}} \colon V^\star \to \mathbb{K}$ is defined, for all linear functionals φ in $V^\star$, by $\widehat{\boldsymbol{v}}[\varphi] = \varphi(\boldsymbol{v})$. Prove that $\widehat{\boldsymbol{v}}$ is a linear functional on $V^\star$.

ii. Consider the map $\Psi \colon V \to V^{\star\star}$ defined by $\Psi[\boldsymbol{v}] = \widehat{\boldsymbol{v}}$. Prove that Ψ is a linear map and an isomorphism.

9.1 Isometries

WHICH LINEAR MAPS PERSERVE NORMS? We analyze the distance-preserving linear maps between inner product spaces.

9.1.0 Definition. Let V and W be inner product spaces over the same field of scalars. A linear map $S \colon V \to W$ is an ***isometry*** if $\|S[\boldsymbol{v}]\|_W = \|\boldsymbol{v}\|_V$ for all vectors $\boldsymbol{v}$ in V.

Coming from ancient Greek, the word "isometry" means equality of measure.

9.1.1 Lemma. *Every isometry is injective. Moreover, an linear operator on a finite-dimensional vector space that is an isometry is an isomorphism.*

Proof. Let $S \colon V \to W$ be an isometry. For any vector $\boldsymbol{v}$ in $\operatorname{Ker}(\boldsymbol{v})$, we have $0 = \|\boldsymbol{0}\|_W = \|S[\boldsymbol{v}]\|_W = \|\boldsymbol{v}\|_V$ which implies that $\boldsymbol{v} = 0$. Since $\operatorname{Ker}(S) = \{\boldsymbol{0}\}$, the injectivity criterion [3.1.4] shows that S is injective. The first claim together with the characterization of invertible operators [3.3.5] proves the second claim. □

9.1.2 Lemma. *A linear map $S \colon V \to W$ between complex inner product spaces is an isometry if and only if $\langle S[\boldsymbol{v}_1], S[\boldsymbol{v}_2] \rangle_W = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle_V$ for all vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in V.*

Proof.

$\Rightarrow$: Suppose that S is an isometry. For all vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in V, the polar identity [7.0.11], the linearity of S, and the definition of isometry give

$$\begin{aligned}
\langle S[\boldsymbol{v}_1], S[\boldsymbol{v}_2] \rangle &= \tfrac{1}{4}\big(\|S[\boldsymbol{v}_1] + S[\boldsymbol{v}_2]\|^2 - \|S[\boldsymbol{v}_1] - S[\boldsymbol{v}_2]\|^2 + \mathrm{i}\,\|S[\boldsymbol{v}_1] + \mathrm{i}\,S[\boldsymbol{v}_2]\|^2 - \mathrm{i}\,\|S[\boldsymbol{v}_1] - \mathrm{i}\,S[\boldsymbol{v}_2]\|^2\big) \\
&= \tfrac{1}{4}\big(\|S[\boldsymbol{v}_1 + \boldsymbol{v}_2]\|^2 - \|S[\boldsymbol{v}_1 - \boldsymbol{v}_2]\|^2 + \mathrm{i}\,\|S[\boldsymbol{v}_1 + \mathrm{i}\,\boldsymbol{v}_2]\|^2 - \mathrm{i}\,\|S[\boldsymbol{v}_1 - \mathrm{i}\,\boldsymbol{v}_2]\|^2\big) \\
&= \tfrac{1}{4}\big(\|\boldsymbol{v}_1 + \boldsymbol{v}_1\|^2 - \|\boldsymbol{v}_1 - \boldsymbol{v}_2\|^2 + \mathrm{i}\;\|\boldsymbol{v}_1 + \mathrm{i}\,\boldsymbol{v}_2\|^2 - \mathrm{i}\;\|\boldsymbol{v}_1 - \mathrm{i}\,\boldsymbol{v}_2\|^2\big) = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle\ .
\end{aligned}$$

$\Leftarrow$: Suppose that $\langle S[\boldsymbol{v}_1], S[\boldsymbol{v}_2]\rangle_W = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle_V$ for all vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in V. It follows that, for any vector $\boldsymbol{v}$ in V, we have

$$\|S[\boldsymbol{v}]\|_W^2 = \langle S[\boldsymbol{v}], S[\boldsymbol{v}]\rangle_W = \langle \boldsymbol{v}, \boldsymbol{v}\rangle_V = \|\boldsymbol{v}\|_V^2 \,.$$

Since the norm of a vector is a nonnegative real number, taking square-roots gives $\|S[\boldsymbol{v}]\|_W = \|\boldsymbol{v}\|_V$. □

Surjective isometries have several more equivalent descriptions.

9.1.3 Theorem (Characterizations of surjective isometries)**.** *Let V and W be inner product spaces over the same field $\mathbb{K}$ of scalars. For any linear map $S \colon V \to W$, the following are equivalent:*

a. *The linear map S is a surjective isometry.*
b. *The linear map S is surjective and $S^\star S = \mathrm{id}_V$.*
c. *The linear map S is invertible and $S^{-1} = S^\star$.*
d. *The linear map $S^\star$ is surjective and $S\,S^\star = \mathrm{id}_W$.*
e. *The linear map $S^\star$ is a surjective isometry.*
f. *The vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ form an orthonormal basis of V if and only if the vectors $S[\boldsymbol{u}_1], S[\boldsymbol{u}_2], \dots, S[\boldsymbol{u}_n]$ form an orthonormal basis of W.*

Proof.

$a \Rightarrow b$: For all vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in V, Lemma 9.1.2, the conjugate-symmetry [7.0.0] of inner products, and the definition [9.0.3] of the adjoint map give

$$\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle_V = \langle S[\boldsymbol{v}_1], S[\boldsymbol{v}_2]\rangle_W = \overline{\langle S[\boldsymbol{v}_2], S[\boldsymbol{v}_1]\rangle_W} = \overline{\langle \boldsymbol{v}_2, (S^\star S)[\boldsymbol{v}_1]\rangle_V} = \langle (S^\star S)[\boldsymbol{v}_1], \boldsymbol{v}_2\rangle_V \,,$$

so we have $\langle (S^\star S - \mathrm{id}_V)[\boldsymbol{v}_1], \boldsymbol{v}_2\rangle = 0$. The equality test [9.0.0] proves that $(S^\star S - \mathrm{id}_V)[\boldsymbol{v}_1] = 0$, so $S^\star S = \mathrm{id}_V$.

$b \Rightarrow c$: For any vector $\boldsymbol{v}$ in $\operatorname{Ker}(S)$, we have $\operatorname{Ker}(S) = \{\boldsymbol{0}\}$ because $\boldsymbol{v} = (S^\star S)[\boldsymbol{v}] = S^\star\big[S[\boldsymbol{v}]\big] = S^\star[\boldsymbol{0}] = \boldsymbol{0}$. Hence, the injectivity criterion [3.1.4] shows that the linear map S is injective. Since S is bijective, the characterization of invertibility [3.2.5] establishes that the linear map S is invertible and the uniqueness of inverses [3.2.4] demonstrates that $S^\star = S^{-1}$.

$c \Rightarrow d$: By definition [3.2.2], a linear operator is invertible if there exists $S^{-1} \colon V \to V$ such that $S\,S^{-1} = \mathrm{id}_W$ and $S^{-1}\,S = \mathrm{id}_V$. By hypothesis, we also have $S^{-1} = S^\star$, so we infer that $S\,S^\star = \mathrm{id}_W$.

$d \Rightarrow e$: For all vectors $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ in W, the definition of the adjoint map [9.0.3] gives

$$\langle S^\star[\boldsymbol{w}_1], S^\star[\boldsymbol{w}_2]\rangle_V = \langle \boldsymbol{w}_1, (S\,S^\star)[\boldsymbol{w}_2]\rangle_W = \langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle_W \,.$$

Hence, Lemma 9.1.2 establishes that $S^\star$ is an isometry.

$e \Rightarrow d \Rightarrow c \Rightarrow b \Rightarrow a$: By replacing S with $S^\star$ and using the involutive property [9.0.4] of adjoint maps, the first four steps in the prove establishes these implications.

$a \Rightarrow f$: For all indices $1 \leqslant j \leqslant k \leqslant n$, Lemma 9.1.2 demonstrates that $\langle \boldsymbol{u}_j, \boldsymbol{u}_k \rangle = \langle S[\boldsymbol{u}_j], S[\boldsymbol{u}_k] \rangle$. Hence, the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ are orthonormal if and only if the vectors $S[\boldsymbol{u}_1], S[\boldsymbol{u}_2], \dots, S[\boldsymbol{u}_n]$ are orthonormal. Since S is invertible, the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ are a basis of V if and only if the vectors $S[\boldsymbol{u}_1], S[\boldsymbol{u}_2], \dots, S[\boldsymbol{u}_n]$ are a basis of W.

$f \Rightarrow a$: Fix an orthonormal basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$ for the inner product space V. For any vector $\boldsymbol{v}$ in V, the orthonormal coordinates [7.1.7], linearity of S, and the Parseval identity [7.1.5] imply that

$$\begin{aligned} \|S[\boldsymbol{v}]\|^2 &= \|S[\langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{v}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2 + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_n \rangle \boldsymbol{u}_n]\|^2 \\ &= \|\langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle S[\boldsymbol{u}_1] + \langle \boldsymbol{v}, \boldsymbol{u}_2 \rangle S[\boldsymbol{u}_2] + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_n \rangle S[\boldsymbol{u}_n]\|^2 \\ &= |\langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle|^2 + |\langle \boldsymbol{v}, \boldsymbol{u}_2 \rangle|^2 + \cdots + |\langle \boldsymbol{v}, \boldsymbol{u}_n \rangle|^2 \\ &= \|\langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{v}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2 + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_n \rangle \boldsymbol{u}_n\|^2 = \|\boldsymbol{v}\|^2 . \end{aligned}$$

The nonnegativity of inner products [7.0.0] establishes that, by taking the square root, we obtain $\|S[\boldsymbol{v}]\| = \|\boldsymbol{v}\|$. □

9.1.4 Remark. An isometry on a complex inner product space is often called a ***unitary*** operator, and an isometry on a real inner product space is often called an ***orthogonal*** operator. Similarly, a complex square matrix $\mathbf{Q}$ is ***unitary*** if $\mathbf{Q}^\star = \mathbf{Q}^{-1}$ and a real square matrix $\mathbf{Q}$ is ***orthogonal*** if $\mathbf{Q}^\mathsf{T} = \mathbf{Q}^{-1}$.

9.1.5 Proposition (Orthonormal triangularization)**.** *For any linear map $T \colon V \to V$ on a finite-dimensional complex inner product space V, there exists an ordered orthonormal basis for V such that the matrix of T relative to this basis is upper triangular.*

Proof. Set $n := \dim V$. By the triangularization theorem [5.2.1], there exists an ordered basis $\mathcal{B} := (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ of the $\mathbb{C}$-vector space V such that the matrix of T relative to $\mathcal{B}$ is upper-triangular. The orthonormalization algorithm [7.2.0] applied to the basis $\mathcal{B}$ returns an orthonormal basis $\mathcal{U} := (\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n)$ such that, for all $1 \leqslant k \leqslant n$, we have $\operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k) = \operatorname{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_k)$. It follows from the characterization of triangular operators [5.2.0] that, for all $1 \leqslant k \leqslant n$, we have $T[\boldsymbol{b}_k] \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k)$. Since $\boldsymbol{u}_k \in \operatorname{Span}(\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_k)$, there exists scalars $c_1, c_2, \dots, c_k \in \mathbb{C}$ such that $\boldsymbol{u}_k = c_1 \boldsymbol{b}_1 + c_2 \boldsymbol{b}_2 + \cdots + c_k \boldsymbol{b}_k$. We deduce that

$$\begin{aligned} T[\boldsymbol{u}_k] &= T[c_1 \boldsymbol{b}_1 + c_2 \boldsymbol{b}_2 + \cdots + c_k \boldsymbol{b}_k] \\ &= c_1 T[\boldsymbol{b}_1] + c_2 T[\boldsymbol{b}_2] + \cdots + c_k T[\boldsymbol{b}_k] \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k) , \end{aligned}$$

so the characterization of triangular operators demonstrates that the matrix $(T)^{\mathcal{U}}_{\mathcal{U}}$ of T relative to $\mathcal{U}$ is upper-triangular. □