5.3 INNER-PRODUCT SPACES

The euclidean norm, which naturally came first, is a coordinate-dependent concept. But by isolating its important properties we quickly moved to the more general coordinate-free definition of a vector norm given in (5.1.9) on p. 275. The goal is to now do the same for inner products. That is, start with the standard inner product, which is a coordinate-dependent definition, and identify properties that characterize the basic essence of the concept. The ones listed below are those that have been distilled from the standard inner product to formulate a more general coordinate-free definition.

General Inner Product

An ***inner product*** on a real (or complex) vector space $\mathcal{V}$ is a function that maps each ordered pair of vectors $\mathbf{x}, \mathbf{y}$ to a real (or complex) scalar $\langle \mathbf{x}|\mathbf{y} \rangle$ such that the following four properties hold.

$$\begin{aligned}
&\langle \mathbf{x}|\mathbf{x} \rangle \text{ is real with } \langle \mathbf{x}|\mathbf{x} \rangle \geq 0, \text{ and } \langle \mathbf{x}|\mathbf{x} \rangle = 0 \text{ if and only if } \mathbf{x} = \mathbf{0},\\
&\langle \mathbf{x}|\alpha\mathbf{y} \rangle = \alpha \langle \mathbf{x}|\mathbf{y} \rangle \text{ for all scalars } \alpha,\\
&\langle \mathbf{x}|\mathbf{y}+\mathbf{z} \rangle = \langle \mathbf{x}|\mathbf{y} \rangle + \langle \mathbf{x}|\mathbf{z} \rangle,\\
&\langle \mathbf{x}|\mathbf{y} \rangle = \overline{\langle \mathbf{y}|\mathbf{x} \rangle} \quad \text{(for real spaces, this becomes } \langle \mathbf{x}|\mathbf{y} \rangle = \langle \mathbf{y}|\mathbf{x} \rangle).
\end{aligned} \tag{5.3.1}$$

Notice that for each fixed value of $\mathbf{x}$, the second and third properties say that $\langle \mathbf{x}|\mathbf{y} \rangle$ is a linear function of $\mathbf{y}$.

Any real or complex vector space that is equipped with an inner product is called an ***inner-product space***.

Example 5.3.1

- The standard inner products, $\langle \mathbf{x}|\mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ for $\Re^{n\times 1}$ and $\langle \mathbf{x}|\mathbf{y} \rangle = \mathbf{x}^*\mathbf{y}$ for $\mathcal{C}^{n\times 1}$, each satisfy the four defining conditions (5.3.1) for a general inner product—this shouldn't be a surprise.

- If $\mathbf{A}_{n\times n}$ is a nonsingular matrix, then $\langle \mathbf{x}|\mathbf{y} \rangle = \mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{y}$ is an inner product for $\mathcal{C}^{n\times 1}$. This inner product is sometimes called an **A*-inner product*** or an ***elliptical inner product.***

- Consider the vector space of $m \times n$ matrices. The functions defined by

$$\langle \mathbf{A}|\mathbf{B} \rangle = trace\left(\mathbf{A}^T\mathbf{B}\right) \quad \text{and} \quad \langle \mathbf{A}|\mathbf{B} \rangle = trace\left(\mathbf{A}^*\mathbf{B}\right) \tag{5.3.2}$$

  are inner products for $\Re^{m\times n}$ and $\mathcal{C}^{m\times n}$, respectively. These are referred to as the ***standard inner products for matrices***. Notice that these reduce to the standard inner products for vectors when $n = 1$.

- If $\mathcal{V}$ is the vector space of real-valued continuous functions defined on the interval (a, b), then
$$\langle f|g\rangle = \int_a^b f(t)g(t)dt$$
is an inner product on $\mathcal{V}$.

Just as the standard inner product for $\mathcal{C}^{n\times 1}$ defines the euclidean norm on $\mathcal{C}^{n\times 1}$ by $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^*\mathbf{x}}$, every general inner product in an inner-product space $\mathcal{V}$ defines a norm on $\mathcal{V}$ by setting

$$\|\star\| = \sqrt{\langle \star|\star\rangle}. \tag{5.3.3}$$

It's straightforward to verify that this satisfies the first two conditions in (5.2.3) on p. 280 that define a general vector norm, but, just as in the case of euclidean norms, verifying that (5.3.3) satisfies the triangle inequality requires a generalized version of CBS inequality.

General CBS Inequality

If $\mathcal{V}$ is an inner-product space, and if we set $\|\star\| = \sqrt{\langle \star|\star\rangle}$, then

$$|\langle \mathbf{x}|\mathbf{y}\rangle| \le \|\mathbf{x}\| \, \|\mathbf{y}\| \quad \text{for all } x, y \in \mathcal{V}. \tag{5.3.4}$$

Equality holds if and only if $\mathbf{y} = \alpha\mathbf{x}$ for $\alpha = \langle \mathbf{x}|\mathbf{y}\rangle / \|\mathbf{x}\|^2$.

Proof. Set $\alpha = \langle \mathbf{x}|\mathbf{y}\rangle / \|\mathbf{x}\|^2$ (assume $\mathbf{x} \neq \mathbf{0}$, for otherwise there is nothing to prove), and observe that $\langle \mathbf{x}|\alpha\mathbf{x} - \mathbf{y}\rangle = 0$, so

$$\begin{aligned} 0 \le \|\alpha\mathbf{x} - \mathbf{y}\|^2 &= \langle \alpha\mathbf{x} - \mathbf{y}|\alpha\mathbf{x} - \mathbf{y}\rangle \\ &= \bar{\alpha}\langle \mathbf{x}|\alpha\mathbf{x} - \mathbf{y}\rangle - \langle \mathbf{y}|\alpha\mathbf{x} - \mathbf{y}\rangle \quad \text{(see Exercise 5.3.2)} \\ &= -\langle \mathbf{y}|\alpha\mathbf{x} - \mathbf{y}\rangle = \langle \mathbf{y}|\mathbf{y}\rangle - \alpha\langle \mathbf{y}|\mathbf{x}\rangle = \frac{\|\mathbf{y}\|^2\|\mathbf{x}\|^2 - \langle \mathbf{x}|\mathbf{y}\rangle\langle \mathbf{y}|\mathbf{x}\rangle}{\|\mathbf{x}\|^2}. \end{aligned}$$

Since $\langle \mathbf{y}|\mathbf{x}\rangle = \overline{\langle \mathbf{x}|\mathbf{y}\rangle}$, it follows that $\langle \mathbf{x}|\mathbf{y}\rangle\langle \mathbf{y}|\mathbf{x}\rangle = |\langle \mathbf{x}|\mathbf{y}\rangle|^2$, so

$$0 \le \frac{\|\mathbf{y}\|^2\|\mathbf{x}\|^2 - |\langle \mathbf{x}|\mathbf{y}\rangle|^2}{\|\mathbf{x}\|^2} \implies |\langle \mathbf{x}|\mathbf{y}\rangle| \le \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

Establishing the conditions for equality is the same as in Exercise 5.1.9. ■

Let's now complete the job of showing that $\|\star\| = \sqrt{\langle \star|\star\rangle}$ is indeed a vector norm as defined in (5.2.3) on p. 280.

Norms in Inner-Product Spaces

If $\mathcal{V}$ is an inner-product space with an inner product $\langle \mathbf{x}|\mathbf{y}\rangle$, then

$$\|\star\| = \sqrt{\langle \star|\star\rangle} \quad \text{defines a norm on } \mathcal{V}.$$

Proof. The fact that $\|\star\| = \sqrt{\langle \star|\star\rangle}$ satisfies the first two norm properties in (5.2.3) on p. 280 follows directly from the defining properties (5.3.1) for an inner product. You are asked to provide the details in Exercise 5.3.3. To establish the triangle inequality, use $\langle \mathbf{x}|\mathbf{y}\rangle \le |\langle \mathbf{x}|\mathbf{y}\rangle|$ and $\langle \mathbf{y}|\mathbf{x}\rangle = \overline{\langle \mathbf{x}|\mathbf{y}\rangle} \le |\langle \mathbf{x}|\mathbf{y}\rangle|$ together with the CBS inequality to write

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{x}|\mathbf{x}\rangle + \langle \mathbf{x}|\mathbf{y}\rangle + \langle \mathbf{y}|\mathbf{x}\rangle + \langle \mathbf{y}|\mathbf{y}\rangle \\ &\le \|\mathbf{x}\|^2 + 2|\langle \mathbf{x}|\mathbf{y}\rangle| + \|\mathbf{y}\|^2 \le (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \quad \blacksquare\end{aligned}$$

Example 5.3.2

Problem: Describe the norms that are generated by the inner products presented in Example 5.3.1.

- Given a nonsingular matrix $\mathbf{A} \in \mathcal{C}^{n\times n}$, the **A-*norm*** (or ***elliptical norm***) generated by the $\mathbf{A}$-inner product on $\mathcal{C}^{n\times 1}$ is

$$\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\langle \mathbf{x}|\mathbf{x}\rangle} = \sqrt{\mathbf{x}^*\mathbf{A}^*\mathbf{A}\mathbf{x}} = \|\mathbf{A}\mathbf{x}\|_2 . \qquad (5.3.5)$$

- The standard inner product for matrices generates the Frobenius matrix norm because

$$\|\mathbf{A}\| = \sqrt{\langle \mathbf{A}|\mathbf{A}\rangle} = \sqrt{\mathit{trace}\,(\mathbf{A}^*\mathbf{A})} = \|\mathbf{A}\|_F . \qquad (5.3.6)$$

- For the space of real-valued continuous functions defined on (a, b), the norm of a function f generated by the inner product $\langle f|g\rangle = \int_a^b f(t)g(t)dt$ is

$$\|f\| = \sqrt{\langle f|f\rangle} = \left(\int_a^b f(t)^2 dt\right)^{1/2}.$$

Example 5.3.3

To illustrate the utility of the ideas presented above, consider the proposition

$$trace\left(\mathbf{A}^T\mathbf{B}\right)^2 \leq trace\left(\mathbf{A}^T\mathbf{A}\right) trace\left(\mathbf{B}^T\mathbf{B}\right) \quad \text{for all } \mathbf{A}, \mathbf{B} \in \Re^{m\times n}.$$

Problem: How would you know to formulate such a proposition and, second, how do you prove it?

Solution: The answer to both questions is the same. This is the CBS inequality in $\Re^{m\times n}$ equipped with the standard inner product $\langle \mathbf{A}|\mathbf{B}\rangle = trace\left(\mathbf{A}^T\mathbf{B}\right)$ and associated norm $\|\mathbf{A}\|_F = \sqrt{\langle \mathbf{A}|\mathbf{A}\rangle} = \sqrt{trace\left(\mathbf{A}^T\mathbf{A}\right)}$ because CBS says

$$\langle \mathbf{A}|\mathbf{B}\rangle^2 \leq \|\mathbf{A}\|_F^2 \|\mathbf{B}\|_F^2 \implies trace\left(\mathbf{A}^T\mathbf{B}\right)^2 \leq trace\left(\mathbf{A}^T\mathbf{A}\right) trace\left(\mathbf{B}^T\mathbf{B}\right).$$

The point here is that if your knowledge is limited to elementary matrix manipulations (which is all that is needed to understand the statement of the proposition), formulating the correct inequality might be quite a challenge to your intuition. And then proving the proposition using only elementary matrix manipulations would be a significant task—essentially, you would have to derive a version of CBS. But knowing the basic facts of inner-product spaces makes the proposition nearly trivial to conjecture and prove.

Since each inner product generates a norm by the rule $\|\star\| = \sqrt{\langle\star|\star\rangle}$, it's natural to ask if the reverse is also true. That is, for each vector norm $\|\star\|$ on a space $\mathcal{V}$, does there exist a corresponding inner product on $\mathcal{V}$ such that $\sqrt{\langle\star|\star\rangle} = \|\star\|^2$? If not, under what conditions will a given norm be generated by an inner product? These are tricky questions, and it took the combined efforts of Maurice R. Fréchet[38] (1878–1973) and John von Neumann (1903–1957) to provide the answer.

[38] Maurice René Fréchet began his illustrious career by writing an outstanding Ph.D. dissertation in 1906 under the direction of the famous French mathematician Jacques Hadamard (p. 469) in which the concepts of a metric space and compactness were first formulated. Fréchet developed into a versatile mathematical scientist, and he served as professor of mechanics at the University of Poitiers (1910–1919), professor of higher calculus at the University of Strasbourg (1920–1927), and professor of differential and integral calculus and professor of the calculus of probabilities at the University of Paris (1928–1948).

Born in Budapest, Hungary, John von Neumann was a child prodigy who could divide eight-digit numbers in his head when he was only six years old. Due to the political unrest in Europe, he came to America, where, in 1933, he became one of the six original professors of mathematics at the Institute for Advanced Study at Princeton University, a position he retained for the rest of his life. During his career, von Neumann's genius touched mathematics (pure and applied), chemistry, physics, economics, and computer science, and he is generally considered to be among the best scientists and mathematicians of the twentieth century.

Parallelogram Identity

For a given norm $\|\star\|$ on a vector space $\mathcal{V}$, there exists an inner product on $\mathcal{V}$ such that $\langle\star|\star\rangle = \|\star\|^2$ if and only if the ***parallelogram identity***

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\big(\,\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\,\big) \tag{5.3.7}$$

holds for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$.

Proof. Consider real spaces—complex spaces are discussed in Exercise 5.3.6. If there exists an inner product such that $\langle\star|\star\rangle = \|\star\|^2$, then the parallelogram identity is immediate because $\langle\mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}\rangle + \langle\mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y}\rangle = 2\langle\mathbf{x}|\mathbf{x}\rangle + 2\langle\mathbf{y}|\mathbf{y}\rangle$. The difficult part is establishing the converse. Suppose $\|\star\|$ satisfies the parallelogram identity, and prove that the function

$$\langle\mathbf{x}|\mathbf{y}\rangle = \frac{1}{4}\big(\,\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\,\big) \tag{5.3.8}$$

is an inner product for $\mathcal{V}$ such that $\langle\mathbf{x}|\mathbf{x}\rangle = \|\mathbf{x}\|^2$ for all $\mathbf{x}$ by showing the four defining conditions (5.3.1) hold. The first and fourth conditions are immediate. To establish the third, use the parallelogram identity to write

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}+\mathbf{z}\|^2 &= \frac{1}{2}\big(\,\|\mathbf{x}+\mathbf{y}+\mathbf{x}+\mathbf{z}\|^2 + \|\mathbf{y}-\mathbf{z}\|^2\,\big),\\
\|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{z}\|^2 &= \frac{1}{2}\big(\,\|\mathbf{x}-\mathbf{y}+\mathbf{x}-\mathbf{z}\|^2 + \|\mathbf{z}-\mathbf{y}\|^2\,\big),
\end{aligned}$$

and then subtract to obtain

$$\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}+\mathbf{z}\|^2 - \|\mathbf{x}-\mathbf{z}\|^2 = \frac{\|2\mathbf{x}+(\mathbf{y}+\mathbf{z})\|^2 - \|2\mathbf{x}-(\mathbf{y}+\mathbf{z})\|^2}{2}.$$

Consequently,

$$\begin{aligned}
\langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{x}|\mathbf{z}\rangle &= \frac{1}{4}\big(\,\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + \|\mathbf{x}+\mathbf{z}\|^2 - \|\mathbf{x}-\mathbf{z}\|^2\,\big)\\
&= \frac{1}{8}\big(\,\|2\mathbf{x}+(\mathbf{y}+\mathbf{z})\|^2 - \|2\mathbf{x}-(\mathbf{y}+\mathbf{z})\|^2\,\big)\\
&= \frac{1}{2}\left(\left\|\mathbf{x}+\frac{\mathbf{y}+\mathbf{z}}{2}\right\|^2 - \left\|\mathbf{x}-\frac{\mathbf{y}+\mathbf{z}}{2}\right\|^2\right) = 2\left\langle\mathbf{x}\,\middle|\,\frac{\mathbf{y}+\mathbf{z}}{2}\right\rangle,
\end{aligned} \tag{5.3.9}$$

and setting $\mathbf{z} = \mathbf{0}$ produces the statement that $\langle\mathbf{x}|\mathbf{y}\rangle = 2\langle\mathbf{x}|\mathbf{y}/2\rangle$ for all $\mathbf{y} \in \mathcal{V}$. Replacing $\mathbf{y}$ by $\mathbf{y}+\mathbf{z}$ yields $\langle\mathbf{x}|\mathbf{y}+\mathbf{z}\rangle = 2\langle\mathbf{x}|(\mathbf{y}+\mathbf{z})/2\rangle$, and thus (5.3.9) guarantees that $\langle \mathbf{x}|\mathbf{y}\rangle + \langle \mathbf{x}|\mathbf{z}\rangle = \langle \mathbf{x}|\mathbf{y}+\mathbf{z}\rangle$. Now prove that $\langle \mathbf{x}|\alpha\mathbf{y}\rangle = \alpha\langle \mathbf{x}|\mathbf{y}\rangle$ for all real α. This is valid for integer values of α by the result just established, and it holds when α is rational because if β and γ are integers, then

$$\gamma^2 \left\langle \mathbf{x}\middle|\frac{\beta}{\gamma}\mathbf{y}\right\rangle = \langle \gamma\mathbf{x}|\beta\mathbf{y}\rangle = \beta\gamma\langle \mathbf{x}|\mathbf{y}\rangle \quad\Longrightarrow\quad \left\langle \mathbf{x}\middle|\frac{\beta}{\gamma}\mathbf{y}\right\rangle = \frac{\beta}{\gamma}\langle \mathbf{x}|\mathbf{y}\rangle.$$

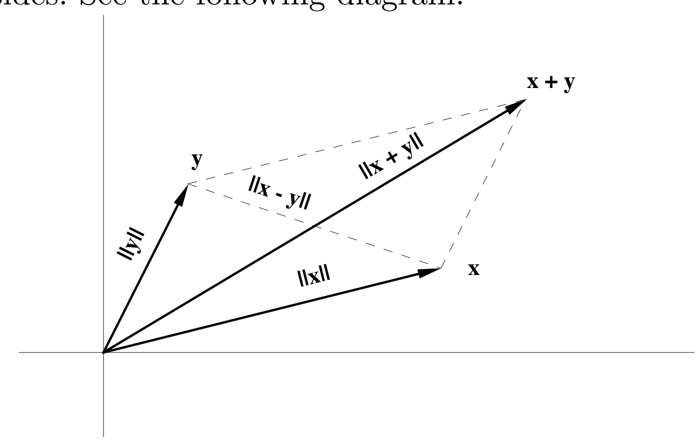
Because $\|\mathbf{x}+\alpha\mathbf{y}\|$ and $\|\mathbf{x}-\alpha\mathbf{y}\|$ are continuous functions of α (Exercise 5.1.7), equation (5.3.8) insures that $\langle \mathbf{x}|\alpha\mathbf{y}\rangle$ is a continuous function of α. Therefore, if α is irrational, and if $\{\alpha_n\}$ is a sequence of rational numbers such that $\alpha_n \to \alpha$, then $\langle \mathbf{x}|\alpha_n\mathbf{y}\rangle \to \langle \mathbf{x}|\alpha\mathbf{y}\rangle$ and $\langle \mathbf{x}|\alpha_n\mathbf{y}\rangle = \alpha_n\langle \mathbf{x}|\mathbf{y}\rangle \to \alpha\langle \mathbf{x}|\mathbf{y}\rangle$, so $\langle \mathbf{x}|\alpha\mathbf{y}\rangle = \alpha\langle \mathbf{x}|\mathbf{y}\rangle$. ∎

Example 5.3.4

We already know that the euclidean vector norm on $\mathcal{C}^n$ is generated by the standard inner product, so the previous theorem guarantees that the parallelogram identity must hold for the 2-norm. This is easily corroborated by observing that

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|_2^2 + \|\mathbf{x}-\mathbf{y}\|_2^2 &= (\mathbf{x}+\mathbf{y})^*(\mathbf{x}+\mathbf{y}) + (\mathbf{x}-\mathbf{y})^*(\mathbf{x}-\mathbf{y}) \\ &= 2\,(\mathbf{x}^*\mathbf{x}+\mathbf{y}^*\mathbf{y}) = 2(\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2).\end{aligned}$$

The parallelogram identity is so named because it expresses the fact that the sum of the squares of the diagonals in a parallelogram is twice the sum of the squares of the sides. See the following diagram.



Example 5.3.5

Problem: Except for the euclidean norm, is any other vector p-norm generated by an inner product?

Solution: No, because the parallelogram identity (5.3.7) doesn't hold when $p \neq 2$. To see that $\|\mathbf{x}+\mathbf{y}\|_p^2 + \|\mathbf{x}-\mathbf{y}\|_p^2 = 2\big(\|\mathbf{x}\|_p^2 + \|\mathbf{y}\|_p^2\big)$ is not valid for all $\mathbf{x},\mathbf{y} \in \mathcal{C}^n$ when $p \neq 2$, consider $\mathbf{x} = \mathbf{e}_1$ and $\mathbf{y} = \mathbf{e}_2$. It's apparent that $\|\mathbf{e_1}+\mathbf{e_2}\|_p^2 = 2^{2/p} = \|\mathbf{e_1}-\mathbf{e_2}\|_p^2$, so

$$\|\mathbf{e_1}+\mathbf{e_2}\|_p^2 + \|\mathbf{e_1}-\mathbf{e_2}\|_p^2 = 2^{(p+2)/p} \quad\text{and}\quad 2\big(\|\mathbf{e_1}\|_p^2 + \|\mathbf{e_2}\|_p^2\big) = 4.$$

Clearly, $2^{(p+2)/p} = 4$ only when $p = 2$. Details for the ∞-norm are asked for in Exercise 5.3.7.

Conclusion: For applications that are best analyzed in the context of an inner-product space (e.g., least squares problems), we are limited to the euclidean norm or else to one of its variation such as the elliptical norm in (5.3.5).

Virtually all important statements concerning $\Re^n$ or $\mathcal{C}^n$ with the standard inner product remain valid for general inner-product spaces—e.g., consider the statement and proof of the general CBS inequality. Advanced or more theoretical texts prefer a development in terms of general inner-product spaces. However, the focus of this text is matrices and the coordinate spaces $\Re^n$ and $\mathcal{C}^n$, so subsequent discussions will usually be phrased in terms of $\Re^n$ or $\mathcal{C}^n$ and their standard inner products. But remember that extensions to more general inner-product spaces are always lurking in the background, and we will not hesitate to use these generalities or general inner-product notation when they serve our purpose.

Exercises for section 5.3

5.3.1. For $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, determine which of the following are inner products for $\Re^{3\times 1}$.

(a) $\langle \mathbf{x}|\mathbf{y} \rangle = x_1y_1 + x_3y_3$,
(b) $\langle \mathbf{x}|\mathbf{y} \rangle = x_1y_1 - x_2y_2 + x_3y_3$,
(c) $\langle \mathbf{x}|\mathbf{y} \rangle = 2x_1y_1 + x_2y_2 + 4x_3y_3$,
(d) $\langle \mathbf{x}|\mathbf{y} \rangle = x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2$.

5.3.2. For a general inner-product space $\mathcal{V}$, explain why each of the following statements must be true.

(a) If $\langle \mathbf{x}|\mathbf{y} \rangle = 0$ for all $\mathbf{x} \in \mathcal{V}$, then $\mathbf{y} = \mathbf{0}$.
(b) $\langle \alpha\mathbf{x}|\mathbf{y} \rangle = \overline{\alpha}\, \langle \mathbf{x}|\mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and for all scalars α.
(c) $\langle \mathbf{x}+\mathbf{y}|\mathbf{z} \rangle = \langle \mathbf{x}|\mathbf{z} \rangle + \langle \mathbf{y}|\mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$.

5.3.3. Let $\mathcal{V}$ be an inner-product space with an inner product $\langle \mathbf{x}|\mathbf{y} \rangle$. Explain why the function defined by $\|\star\| = \sqrt{\langle \star|\star \rangle}$ satisfies the first two norm properties in (5.2.3) on p. 280.

5.3.4. For a real inner-product space with $\|\star\|^2 = \langle \star|\star \rangle$, derive the inequality

$$\langle \mathbf{x}|\mathbf{y} \rangle \le \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2}{2}. \qquad \textbf{Hint:} \text{ Consider } \mathbf{x} - \mathbf{y}.$$

5.3.5. For $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, explain why each of the following inequalities is valid.

(a) $|trace(\mathbf{B})|^2 \leq n\,[trace(\mathbf{B}^*\mathbf{B})]\,.$

(b) $trace\left(\mathbf{B}^2\right) \leq trace\left(\mathbf{B}^T\mathbf{B}\right)$ for real matrices.

(c) $trace\left(\mathbf{A}^T\mathbf{B}\right) \leq \dfrac{trace\left(\mathbf{A}^T\mathbf{A}\right) + trace\left(\mathbf{B}^T\mathbf{B}\right)}{2}$ for real matrices.

5.3.6. Extend the proof given on p. 290 concerning the parallelogram identity (5.3.7) to include complex spaces. **Hint:** If $\mathcal{V}$ is a complex space with a norm $\|\star\|$ that satisfies the parallelogram identity, let

$$\langle \mathbf{x}|\mathbf{y}\rangle_r = \frac{\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2}{4},$$

and prove that

$$\langle \mathbf{x}|\mathbf{y}\rangle = \langle \mathbf{x}|\mathbf{y}\rangle_r + \mathrm{i}\,\langle \mathrm{i}\mathbf{x}|\mathbf{y}\rangle_r \quad \text{(the \textbf{\textit{polarization identity}})} \qquad (5.3.10)$$

is an inner product on $\mathcal{V}$.

5.3.7. Explain why there does not exist an inner product on $\mathcal{C}^n$ $(n \geq 2)$ such that $\|\star\|_\infty = \sqrt{\langle \star|\star\rangle}$.

5.3.8. Explain why the Frobenius matrix norm on $\mathcal{C}^{n\times n}$ must satisfy the parallelogram identity.

5.3.9. For $n \geq 2$, is either the matrix 1-, 2-, or ∞-norm generated by an inner product on $\mathcal{C}^{n\times n}$?

5.2.7. Proceed as follows.

$$\frac{1}{\min_{\|\mathbf{x}\|=1}\left\|\mathbf{A}^{-1}\mathbf{x}\right\|} = \max_{\|\mathbf{x}\|=1}\left\{\frac{1}{\|\mathbf{A}^{-1}\mathbf{x}\|}\right\} = \max_{\mathbf{y}\neq\mathbf{0}}\left\{\frac{1}{\left\|\mathbf{A}^{-1}\frac{(\mathbf{A}\mathbf{y})}{\|\mathbf{A}\mathbf{y}\|}\right\|}\right\}$$
$$= \max_{\mathbf{y}\neq\mathbf{0}}\frac{\|\mathbf{A}\mathbf{y}\|}{\|\mathbf{A}^{-1}(\mathbf{A}\mathbf{y})\|} = \max_{\mathbf{y}\neq\mathbf{0}}\frac{\|\mathbf{A}\mathbf{y}\|}{\|\mathbf{y}\|} = \max_{\mathbf{y}\neq\mathbf{0}}\left\|\mathbf{A}\left(\frac{\mathbf{y}}{\|\mathbf{y}\|}\right)\right\|$$
$$= \max_{\|\mathbf{x}\|=1}\|\mathbf{A}\mathbf{x}\| = \|\mathbf{A}\|$$

5.2.8. Use (5.2.6) on p. 280 to write $\|(z\mathbf{I}-\mathbf{A})^{-1}\| = (1/\min_{\|\mathbf{x}\|=1}\|(z\mathbf{I}-\mathbf{A})\mathbf{x}\|)$, and let $\mathbf{w}$ be a vector for which $\|\mathbf{w}\| = 1$ and $\|(z\mathbf{I}-\mathbf{A})\mathbf{w}\| = \min_{\|\mathbf{x}\|=1}\|(z\mathbf{I}-\mathbf{A})\mathbf{x}\|$. Use $\|\mathbf{A}\mathbf{w}\| \leq \|\mathbf{A}\| < |z|$ together with the "backward triangle inequality" from Example 5.1.1 (p. 273) to write

$$\|(z\mathbf{I}-\mathbf{A})\mathbf{w}\| = \|z\mathbf{w}-\mathbf{A}\mathbf{w}\| \geq \big|\|z\mathbf{w}\| - \|\mathbf{A}\mathbf{w}\|\big| = \big||z| - \|\mathbf{A}\mathbf{w}\|\big|$$
$$= |z| - \|\mathbf{A}\mathbf{w}\| \geq |z| - \|\mathbf{A}\|.$$

Consequently, $\min_{\|\mathbf{x}\|=1}\|(z\mathbf{I}-\mathbf{A})\mathbf{x}\| = \|(z\mathbf{I}-\mathbf{A})\mathbf{w}\| \geq |z| - \|\mathbf{A}\|$ implies that

$$\|(z\mathbf{I}-\mathbf{A})^{-1}\| = \frac{1}{\min_{\|\mathbf{x}\|=1}\|(z\mathbf{I}-\mathbf{A})\mathbf{x}\|} \leq \frac{1}{|z| - \|\mathbf{A}\|}.$$

Solutions for exercises in section 5. 3

5.3.1. Only (c) is an inner product. The expressions in (a) and (b) each fail the first condition of the definition (5.3.1), and (d) fails the second.

5.3.2. (a) $\langle\mathbf{x}|\mathbf{y}\rangle = 0 \quad \forall\ \mathbf{x}\in\mathcal{V} \implies \langle\mathbf{y}|\mathbf{y}\rangle = 0 \implies \mathbf{y} = \mathbf{0}$.

(b) $\langle\alpha\mathbf{x}|\mathbf{y}\rangle = \overline{\langle\mathbf{y}|\alpha\mathbf{x}\rangle} = \overline{\alpha\langle\mathbf{y}|\mathbf{x}\rangle} = \overline{\alpha}\overline{\langle\mathbf{y}|\mathbf{x}\rangle} = \overline{\alpha}\langle\mathbf{x}|\mathbf{y}\rangle$

(c) $\langle\mathbf{x}+\mathbf{y}|\mathbf{z}\rangle = \overline{\langle\mathbf{z}|\mathbf{x}+\mathbf{y}\rangle} = \overline{\langle\mathbf{z}|\mathbf{x}\rangle + \langle\mathbf{z}|\mathbf{y}\rangle} = \overline{\langle\mathbf{z}|\mathbf{x}\rangle} + \overline{\langle\mathbf{z}|\mathbf{y}\rangle} = \langle\mathbf{x}|\mathbf{z}\rangle + \langle\mathbf{y}|\mathbf{z}\rangle$

5.3.3. The first property in (5.2.3) holds because $\langle\mathbf{x}|\mathbf{x}\rangle \geq 0$ for all $\mathbf{x}\in\mathcal{V}$ implies $\|\mathbf{x}\| = \sqrt{\langle\mathbf{x}|\mathbf{x}\rangle} \geq 0$, and $\|\mathbf{x}\| = 0 \iff \langle\mathbf{x}|\mathbf{x}\rangle = 0 \iff \mathbf{x} = \mathbf{0}$. The second property in (5.2.3) holds because

$$\|\alpha\mathbf{x}\|^2 = \langle\alpha\mathbf{x}|\alpha\mathbf{x}\rangle = \alpha\langle\alpha\mathbf{x}|\mathbf{x}\rangle = \alpha\overline{\langle\mathbf{x}|\alpha\mathbf{x}\rangle} = \alpha\overline{\alpha}\overline{\langle\mathbf{x}|\mathbf{x}\rangle} = |\alpha|^2\langle\mathbf{x}|\mathbf{x}\rangle = |\alpha|^2\|\mathbf{x}\|^2.$$

5.3.4. $0 \leq \|\mathbf{x}-\mathbf{y}\|^2 = \langle\mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y}\rangle = \langle\mathbf{x}|\mathbf{x}\rangle - 2\langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{y}\rangle = \|\mathbf{x}\|^2 - 2\langle\mathbf{x}|\mathbf{y}\rangle + \|\mathbf{y}\|^2$

5.3.5. (a) Use the CBS inequality with the Frobenius matrix norm and the standard inner product as illustrated in Example 5.3.3, and set $\mathbf{A} = \mathbf{I}$.

(b) Proceed as in part (a), but this time set $\mathbf{A} = \mathbf{B}^T$ (recall from Example 3.6.5 that $trace\left(\mathbf{B}^T\mathbf{B}\right) = trace\left(\mathbf{B}\mathbf{B}^T\right)$).

(c) Use the result of Exercise 5.3.4 with the Frobenius matrix norm and the inner product for matrices.

5.3.6. Suppose that parallelogram identity holds, and verify that (5.3.10) satisfies the four conditions in (5.3.1). The first condition follows because $\langle \mathbf{x}|\mathbf{x}\rangle_r = \|\mathbf{x}\|^2$ and $\langle \mathbf{ix}|\mathbf{x}\rangle_r = 0$ combine to yield $\langle \mathbf{x}|\mathbf{x}\rangle = \|\mathbf{x}\|^2$. The second condition (for real α) and third condition hold by virtue of the argument for (5.3.7). We will prove the fourth condition and then return to show that the second holds for complex α. By observing that $\langle \mathbf{x}|\mathbf{y}\rangle_r = \langle \mathbf{y}|\mathbf{x}\rangle_r$ and $\langle \mathbf{ix}|\mathbf{iy}\rangle_r = \langle \mathbf{x}|\mathbf{y}\rangle_r$, we have

$$\langle \mathbf{iy}|\mathbf{x}\rangle_r = \left\langle \mathbf{iy}\middle|-\mathrm{i}^2\mathbf{x}\right\rangle_r = \langle \mathbf{y}|-\mathbf{ix}\rangle_r = -\langle \mathbf{y}|\mathbf{ix}\rangle_r = -\langle \mathbf{ix}|\mathbf{y}\rangle_r,$$

and hence

$$\langle \mathbf{y}|\mathbf{x}\rangle = \langle \mathbf{y}|\mathbf{x}\rangle_r + \mathrm{i}\langle \mathbf{iy}|\mathbf{x}\rangle_r = \langle \mathbf{y}|\mathbf{x}\rangle_r - \mathrm{i}\langle \mathbf{ix}|\mathbf{y}\rangle_r = \langle \mathbf{x}|\mathbf{y}\rangle_r - \mathrm{i}\langle \mathbf{ix}|\mathbf{y}\rangle_r = \overline{\langle \mathbf{x}|\mathbf{y}\rangle}.$$

Now prove that $\langle \mathbf{x}|\alpha\mathbf{y}\rangle = \alpha\langle \mathbf{x}|\mathbf{y}\rangle$ for all complex α. Begin by showing it is true for $\alpha = \mathrm{i}$.

$$\begin{aligned}\langle \mathbf{x}|\mathbf{iy}\rangle &= \langle \mathbf{x}|\mathbf{iy}\rangle_r + \mathrm{i}\langle \mathbf{ix}|\mathbf{iy}\rangle_r = \langle \mathbf{x}|\mathbf{iy}\rangle_r + \mathrm{i}\langle \mathbf{x}|\mathbf{y}\rangle_r = \langle \mathbf{iy}|\mathbf{x}\rangle_r + \mathrm{i}\langle \mathbf{x}|\mathbf{y}\rangle_r \\ &= -\langle \mathbf{ix}|\mathbf{y}\rangle_r + \mathrm{i}\langle \mathbf{x}|\mathbf{y}\rangle_r = \mathrm{i}\left(\langle \mathbf{x}|\mathbf{y}\rangle_r + \mathrm{i}\langle \mathbf{ix}|\mathbf{y}\rangle_r\right) \\ &= \mathrm{i}\langle \mathbf{x}|\mathbf{y}\rangle\end{aligned}$$

For $\alpha = \xi + \mathrm{i}\eta$,

$$\langle \mathbf{x}|\alpha\mathbf{y}\rangle = \langle \mathbf{x}|\xi\mathbf{y} + \mathrm{i}\eta\mathbf{y}\rangle = \langle \mathbf{x}|\xi\mathbf{y}\rangle + \langle \mathbf{x}|\mathrm{i}\eta\mathbf{y}\rangle = \xi\langle \mathbf{x}|\mathbf{y}\rangle + \mathrm{i}\eta\langle \mathbf{x}|\mathbf{y}\rangle = \alpha\langle \mathbf{x}|\mathbf{y}\rangle.$$

Conversely, if $\langle \star|\star\rangle$ is any inner product on $\mathcal{V}$, then with $\|\star\|^2 = \langle \star|\star\rangle$ we have

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}\rangle + \langle \mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + 2\mathrm{Re}\langle \mathbf{x}|\mathbf{y}\rangle + \|\mathbf{y}\|^2 + \|\mathbf{x}\|^2 - 2\mathrm{Re}\langle \mathbf{x}|\mathbf{y}\rangle + \|\mathbf{y}\|^2 \\ &= 2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right).\end{aligned}$$

5.3.7. The parallelogram identity (5.3.7) fails to hold for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}^n$. For example, if $\mathbf{x} = \mathbf{e}_1$ and $\mathbf{y} = \mathbf{e}_2$, then

$$\|\mathbf{e}_1 + \mathbf{e}_2\|_\infty^2 + \|\mathbf{e}_1 - \mathbf{e}_2\|_\infty^2 = 2, \quad \text{but} \quad 2\left(\|\mathbf{e}_1\|_\infty^2 + \|\mathbf{e}_2\|_\infty^2\right) = 4.$$

5.3.8. (a) As shown in Example 5.3.2, the Frobenius matrix norm $\mathcal{C}^{n\times n}$ is generated by the standard matrix inner product (5.3.2), so the result on p. 290 guarantees that $\|\star\|_F$ satisfies the parallelogram identity.

5.3.9. No, because the parallelogram inequality (5.3.7) doesn't hold. To see that $\|\mathbf{X}+\mathbf{Y}\|^2 + \|\mathbf{X}-\mathbf{Y}\|^2 = 2\left(\|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2\right)$ is not valid for all $\mathbf{X}, \mathbf{Y} \in \mathcal{C}^{n\times n}$, let $\mathbf{X} = \mathrm{diag}\,(1, 0, \ldots, 0)$ and $\mathbf{Y} = \mathrm{diag}\,(0, 1, \ldots, 0)$. For $\star = 1, 2,$ or ∞,

$$\|\mathbf{X}+\mathbf{Y}\|_\star^2 + \|\mathbf{X}-\mathbf{Y}\|_\star^2 = 1 + 1 = 2, \quad \text{but} \quad 2\left(\|\mathbf{X}\|_\star^2 + \|\mathbf{Y}\|_\star^2\right) = 4.$$