INNER-PRODUCT SPACES

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 $\mathbf{A}$-***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 $n = 1$.

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)

General CBS Inequality

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

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

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

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

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

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

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 $\langle \mathbf{x}|\mathbf{y}\rangle \leq |\langle \mathbf{x}|\mathbf{y}\rangle|$ and $\langle \mathbf{y}|\mathbf{x}\rangle = \overline{\langle \mathbf{x}|\mathbf{y}\rangle} \leq |\langle \mathbf{x}|\mathbf{y}\rangle|$ together with the CBS inequality to write

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

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

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

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

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

- 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

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

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

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 $\|\mathbf{A}\|_F = \sqrt{\langle \mathbf{A}|\mathbf{A}\rangle} = \sqrt{trace\,(\mathbf{A}^T\mathbf{A})}$ because CBS says

$$\langle \mathbf{A}|\mathbf{B}\rangle^2 \le \|\mathbf{A}\|_F^2 \, \|\mathbf{B}\|_F^2 \quad \Longrightarrow \quad trace\left(\mathbf{A}^T\mathbf{B}\right)^2 \le trace\left(\mathbf{A}^T\mathbf{A}\right) trace\left(\mathbf{B}^T\mathbf{B}\right).$$

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

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

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 $\langle \mathbf{x}|\alpha\mathbf{y}\rangle = \alpha\langle \mathbf{x}|\mathbf{y}\rangle$. ∎

ORTHOGONAL VECTORS

Orthogonality

In an inner-product space $\mathcal{V}$, two vectors $\mathbf{x},\mathbf{y} \in \mathcal{V}$ are said to be ***orthogonal*** (to each other) whenever $\langle \mathbf{x}|\mathbf{y}\rangle = 0$, and this is denoted by writing $\mathbf{x} \perp \mathbf{y}$.

- For $\Re^n$ with the standard inner product, $\mathbf{x} \perp \mathbf{y} \iff \mathbf{x}^T\mathbf{y} = 0$.
- For $\mathcal{C}^n$ with the standard inner product, $\mathbf{x} \perp \mathbf{y} \iff \mathbf{x}^*\mathbf{y} = 0$.

Example

$$\mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 3 \\ -1 \end{pmatrix} \text{ is orthogonal to } \mathbf{y} = \begin{pmatrix} 4 \\ 1 \\ -2 \\ -4 \end{pmatrix} \text{ because } \mathbf{x}^T\mathbf{y} = 0.$$

In spite of the fact that $\mathbf{u}^T\mathbf{v} = 0$, the vectors $\mathbf{u} = \begin{pmatrix} \mathrm{i} \\ 3 \\ 1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} \mathrm{i} \\ 0 \\ 1 \end{pmatrix}$ are *not* orthogonal because $\mathbf{u}^*\mathbf{v} \neq 0$.

Angles

In a real inner-product space $\mathcal{V}$, the radian measure of the ***angle*** between nonzero vectors $\mathbf{x},\mathbf{y} \in \mathcal{V}$ is defined to be the number $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{\langle \mathbf{x}|\mathbf{y}\rangle}{\|\mathbf{x}\|\ \|\mathbf{y}\|}. \qquad (5.4.1)$$

Example

In $\Re^n$, $\cos\theta = \mathbf{x}^T\mathbf{y}/\|\mathbf{x}\|\,\|\mathbf{y}\|$. For example, to determine the angle between $\mathbf{x} = \begin{pmatrix} -4 \\ 2 \\ 1 \\ 2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix}$, compute $\cos\theta = 2/(5)(3) = 2/15$, and use the inverse cosine function to conclude that $\theta = 1.437$ radians (rounded).

Orthonormal Sets

$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is called an ***orthonormal set*** whenever $\|\mathbf{u}_i\| = 1$ for each i, and $\mathbf{u}_i \perp \mathbf{u}_j$ for all $i \neq j$. In other words,

$$\langle \mathbf{u}_i | \mathbf{u}_j \rangle = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{when } i \neq j. \end{cases}$$

- Every orthonormal set is linearly independent. (5.4.2)
- Every orthonormal set of n vectors from an n-dimensional space $\mathcal{V}$ is an orthonormal basis for $\mathcal{V}$.

Proof. The second point follows from the first. To prove the first statement, suppose $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is orthonormal. If $\mathbf{0} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$, use the properties of an inner product to write

$$\begin{aligned} 0 &= \langle \mathbf{u}_i | \mathbf{0} \rangle = \langle \mathbf{u}_i | \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n \rangle \\ &= \alpha_1 \langle \mathbf{u}_i | \mathbf{u}_1 \rangle + \cdots + \alpha_i \langle \mathbf{u}_i | \mathbf{u}_i \rangle + \cdots + \alpha_n \langle \mathbf{u}_i | \mathbf{u}_n \rangle = \alpha_i \|\mathbf{u}_i\|^2 \\ &= \alpha_i \quad \text{for each } i. \quad \blacksquare \end{aligned}$$

Example

The set $\mathcal{B}' = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\}$ is a set of mutually orthogonal vectors because $\mathbf{u}_i^T\mathbf{u}_j = 0$ for $i \neq j$, but $\mathcal{B}'$ is *not* an orthonormal set—each vector does not have unit length. However, it's easy to convert an orthogonal set (not containing a zero vector) into an orthonormal set by simply normalizing each vector. Since $\|\mathbf{u}_1\| = \sqrt{2}$, $\|\mathbf{u}_2\| = \sqrt{3}$, and $\|\mathbf{u}_3\| = \sqrt{6}$, it follows that $\mathcal{B} = \left\{\mathbf{u}_1/\sqrt{2},\ \mathbf{u}_2/\sqrt{3},\ \mathbf{u}_3/\sqrt{6}\right\}$ is orthonormal.