Inner Products and Norms (part II)

Prof. Dan A. Simovici

UMB

Dissimilarities

Definition

A *dissimilarity on a set* S is a function $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ satisfying the following conditions:

1. $d(x, x) = 0$ for all $x \in S$;
2. $d(x, y) = d(y, x)$ for all $x, y \in S$.

The pair (S, d) is a dissimilarity space.

The set of dissimilarities defined on a set S is denoted by $\mathcal{D}_S$.

Other Properties of Dissimilarities

1. $d(x, y) = 0$ implies $d(x, z) = d(y, z)$ for every $x, y, z \in S$ (evenness);
2. $d(x, y) = 0$ implies $x = y$ for every x, y (*definiteness*);
3. $d(x, y) \leqslant d(x, z) + d(z, y)$ for every x, y, z (triangular inequality);
4. $d(x, y) \leqslant \max\{d(x, z), d(z, y)\}$ for every x, y, z (the ultrametric inequality);
5. $d(x, y) + d(u, v) \leqslant \max\{d(x, u) + d(y, v), d(x, v) + d(y, u)\}$ for every x, y, u, v (Buneman's inequality, also known as the four-point condition).

If $d : S^2 \longrightarrow \mathbb{R}$ is a function that satisfies the properties of dissimilarities and the triangular inequality, then the values of d are nonnegative numbers. Indeed, by taking $x = y$ in the triangular inequality, we have

$$0 = d(x, x) \leqslant d(x, z) + d(z, x) = 2d(x, z),$$

for every $z \in S$.

Classes of Dissimilarities

Definition

A dissimilarity $d \in \mathcal{D}_S$ is

1. a pseudo-metric if it satisfies the triangular inequality;
2. a metric if it satisfies the definiteness property and the triangular inequality,
3. a tree metric if it satisfies the definiteness property and Buneman's inequality, and
4. an ultrametric if it satisfies the definiteness property and the ultrametric inequality.

The set of metrics on a set S is denoted by $\mathcal{M}_S$. The sets of tree metrics and ultrametrics on a set S are denoted by $\mathcal{T}_S$ and $\mathcal{U}_S$, respectively.

Metrics

A function $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ is a metric if it has the following properties:

1. $d(x, y) = 0$ if and only if $x = y$ for $x, y \in S$;
2. $d(x, y) = d(y, x)$ for $x, y \in S$;
3. $d(x, y) \leqslant d(x, z) + d(z, y)$ for $x, y, z \in S$.

If the first property is replaced by the weaker requirement that $d(x, x) = 0$ for $x \in S$, then we refer to d as a semimetric on S. Thus, if d is a semimetric $d(x, y) = 0$ does not necessarily imply $x = y$ and we can have for two distinct elements x, y of S, $d(x, y) = 0$.

Example

Let S be a nonempty set. Define the mapping $d : S^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ by

$$d(u, v) = \begin{cases} 1 & \text{if } u \neq v, \\ 0 & \text{otherwise}, \end{cases}$$

for $x, y \in S$. It is clear that d satisfies the definiteness property. The triangular inequality, $d(x, y) \leqslant d(x, z) + d(z, y)$ is satisfied if $x = y$. Therefore, suppose that $x \neq y$, so $d(x, y) = 1$. Then, for every $z \in S$, we have at least one of the inequalities $x \neq z$ or $z \neq y$, so at least one of the numbers $d(x, z)$ or $d(z, y)$ equals 1. Thus d satisfies the triangular inequality. The metric d introduced here is the discrete metric on S.

Example

Consider the mapping $d_h : (\mathbf{Seq}_n(S))^2 \longrightarrow \mathbb{R}_{\geqslant 0}$ defined by

$$d_h(\mathbf{p}, \mathbf{q}) = |\{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \mathbf{p}(i) \neq \mathbf{q}(i)\}|$$

for all sequences $\mathbf{p}, \mathbf{q}$ of length n on the set S.
Clearly, d_h is a dissimilarity that is both even and definite. Moreover, it satisfies the triangular inequality. Indeed, let $\mathbf{p}, \mathbf{q}, \mathbf{r}$ be three sequences of length n on the set S. If $\mathbf{p}(i) \neq \mathbf{q}(i)$, then $\mathbf{r}(i)$ must be distinct from at least one of $\mathbf{p}(i)$ and $\mathbf{q}(i)$. Therefore,

$$\begin{aligned} &\{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \mathbf{p}(i) \neq \mathbf{q}(i)\} \\ &\quad \subseteq \{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \mathbf{p}(i) \neq \mathbf{r}(i)\} \cup \{i \mid 0 \leqslant i \leqslant n-1 \text{ and } \mathbf{r}(i) \neq \mathbf{q}(\end{aligned}$$

which implies the triangular inequality. This distance is known as the Hamming distance on $\mathbf{Seq}_n(S)$.

If we need to compare sequences of unequal length, we can use an extended metric d_h' defined by

$$d_h'(\mathbf{x}, \mathbf{y}) = \begin{cases} |\{i \mid 0 \leqslant i \leqslant |\mathbf{x}| - 1, x_i \neq y_i\} & \text{if } |\mathbf{x}| = |\mathbf{y}|, \\ \infty & \text{if } |\mathbf{x}| \neq |\mathbf{y}|. \end{cases}$$

Example

Define the mapping $d : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}_{\geqslant 0}$ as $d(x, y) = |x - y|$ for $x, y \in \mathbb{R}$. It is clear that $d(x, y) = 0$ if and only if $x = y$ and that $d(x, y) = d(y, x)$ for $x, y \in S$;
To prove the triangular inequality suppose that $x \leqslant y \leqslant z$. Then, $d(x, z) + d(z, y) = z - x + z - y = 2z - x - y$ and we have $2z - x - y > y - x = d(x, y)$ because $z > y$. The triangular inequality is similarly satisfied no matter what the relative order of x, y, z is.

Open and Closed Spheres

Definition

Let (S, d) be a metric space. The closed sphere centered in $x \in S$ of radius r is the set

$$B_d[x, r] = \{y \in S | d(x, y) \leqslant r\}.$$

The open sphere centered in $x \in S$ of radius r is the set

$$B_d(x, r) = \{y \in S | d(x, y) < r\}.$$

The spherical surface centered in $x \in S$ of radius r is the set

$$S_n(x, r) = \{y \in S \mid d(x, y) = r\}.$$

If the metric d is clear from context we drop the subscript d and replace $B_d[x, r]$ and $B_d(x, r)$ by $B[x, r]$ and $B(x, r)$, repectively.

Definition

Let (S,d) be a metric space. The diameter of a subset U of S is the number $diam_{S,d}(U) = \sup\{d(x,y) \mid x,y \in U\}$. The set U is *bounded* if $diam_{S,d}(U)$ is finite.
The diameter of the metric space (S,d) is the number

$$diam_{S,d} = \sup\{d(x,y) \mid x,y \in S\}.$$

If the metric space is clear from the context, then we denote the diameter of a subset U just by $diam(U)$.
If (S,d) is a finite metric space, then $diam_{S,d} = \max\{d(x,y) \mid x,y \in S\}$.

Definition

Let (S, d) and (T, d') be two metric spaces. An isometry between these spaces is a function $f : S \longrightarrow T$ that satisfies the equality

$$d'(f(x), f(y)) = d(x, y)$$

for every $x, y \in S$.
If an isometry exists between (S, d) and (T, d') we say that these metric spaces are isometric.

Note that if $f : S \longrightarrow T$ is an isometry, then $f(x) = f(y)$ implies $d(f(x), f(y)) = d(x, y) = 0$, which yields $x = y$ for $x, y \in S$. Therefore, every isometry is injective.
A surjective isometry is, therefore, a bijection.

Theorem

Each norm $\nu : L \longrightarrow \mathbb{R}_{\geqslant 0}$ on a linear space L generates a metric on the set L defined by $d_\nu(\mathbf{x}, \mathbf{y}) = \nu\mathbf{x} - \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in L$.

Proof

Note that if $d_\nu(\mathbf{x}, \mathbf{y}) = \nu\mathbf{x} - \mathbf{y} = 0$, it follows that $\mathbf{x} - \mathbf{y} = \mathbf{0}_L$, so $\mathbf{x} = \mathbf{y}$. The symmetry of d_ν is obvious and so we need to verify only the triangular axiom. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in L$. We have

$$\nu(\mathbf{x} - \mathbf{z}) = \nu(\mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}) \leqslant \nu(\mathbf{x} - \mathbf{y}) + \nu(\mathbf{y} - \mathbf{z})$$

or, equivalently, $d_\nu(\mathbf{x}, \mathbf{z}) \leqslant d_\nu(\mathbf{x}, \mathbf{y}) + d_\nu(\mathbf{y}, \mathbf{z})$, for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in L$, which concludes the argument.

We refer to d_ν as the metric induced by the norm ν on the linear space L.

Norms Generated by Translation-Invariant Metrics

The metric d_ν on L induced by a norm is translation invariant, that is, $d_\nu(x+z, y+z) = d_\nu(x,y)$ for every $z \in L$. Also, for every $a \in \mathbb{R}$ and $x, y \in L$ we have the homogeneity property $d_\nu(ax, ay) = |a| d_\nu(x,y)$ for $x, y \in L$.

Theorem

Let L be a real linear space and let $d : L \times L \longrightarrow \mathbb{R}_{\geqslant 0}$ be a metric on L. If d is translation invariant and homogeneous, then there exists a norm ν of L such that $d = d_\nu$.

Proof: Let d be a metric on L that is translation invariant and homogeneous. Define $\nu(x) = d(x, 0_L)$. It follows immediately that ν is a norm on L.

Minkowski Metrics

For $p \geqslant 1$, then d_p denotes the metric d_{ν_p} induced by the norm ν_p on the linear space $(\mathbb{R}^n, +, \cdot)$ known as the Minkowski metric on $\mathbb{R}^n$.
The metrics d_1, d_2 and d_∞ defined on $\mathbb{R}^n$ are given by

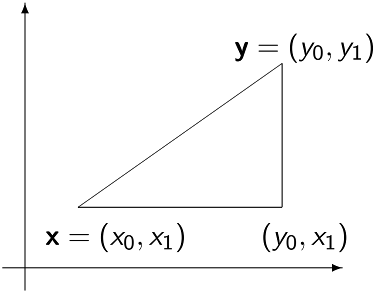
$$
\begin{aligned}
d_1(\mathbf{x}, \mathbf{y}) &= \sum_{i=1}^{n} |x_i - y_i|, \\
d_2(\mathbf{x}, \mathbf{y}) &= \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}, \\
d_\infty(\mathbf{x}, \mathbf{y}) &= \max\{|x_i - y_i| \mid 1 \leqslant i \leqslant n\},
\end{aligned}
$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

If

$$\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix},$$

then $d_1(\mathbf{x}, \mathbf{y})$ is the sum of the lengths of the two legs of the triangle, $d_2(\mathbf{x}, \mathbf{y})$ is the length of the hypotenuse of the right triangle and $d_\infty(\mathbf{x}, \mathbf{y})$ is the largest of the lengths of the legs.



The distances $d_1(\mathbf{x}, \mathbf{y}), d_2(\mathbf{x}, \mathbf{y})$ and $d_\infty(\mathbf{x}, \mathbf{y})$.

Lemma

Let $a_1, \ldots, a_n$ be n positive numbers. If p and q are two positive numbers such that $p \leqslant q$, then $\left(a_1^p + \cdots + a_n^p\right)^{\frac{1}{p}} \geqslant \left(a_1^q + \cdots + a_n^q\right)^{\frac{1}{q}}$.

Proof

Let $f : \mathbb{R}^{>0} \longrightarrow \mathbb{R}$ be the function defined by $f(r) = (a_1^r + \cdots + a_n^r)^{\frac{1}{r}}$. Since

$$\ln f(r) = \frac{\ln (a_1^r + \cdots + a_n^r)}{r},$$

it follows that

$$\frac{f'(r)}{f(r)} = -\frac{1}{r^2} (a_1^r + \cdots + a_n^r) + \frac{1}{r} \cdot \frac{a_1^r \ln a_1 + \cdots + a_n^r \ln a_r}{a_1^r + \cdots + a_n^r}.$$

To prove that $f'(r) < 0$, it suffices to show that

$$\frac{a_1^r \ln a_1 + \cdots + a_n^r \ln a_r}{a_1^r + \cdots + a_n^r} \leqslant \frac{\ln (a_1^r + \cdots + a_n^r)}{r}.$$

Proof (cont'd)

This last inequality is easily seen to be equivalent to

$$\sum_{i=1}^{n} \frac{a_i^r}{a_1^r + \cdots + a_n^r} \ln \frac{a_i^r}{a_1^r + \cdots + a_n^r} \leqslant 0,$$

which holds because

$$\frac{a_i^r}{a_1^r + \cdots + a_n^r} \leqslant 1$$

for $1 \leqslant i \leqslant n$.

Theorem

Let p and q be two positive numbers such that $p \leqslant q$. We have $\| \mathbf{u} \|_p \geqslant \| \mathbf{u} \|_q$ for $\mathbf{u} \in \mathbb{R}^n$.

This follows from the previous Lemma.

Corollary

Let p, q be two positive numbers such that $p \leqslant q$. For every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have $d_p(\mathbf{x}, \mathbf{y}) \geqslant d_q(\mathbf{x}, \mathbf{y})$.

Theorem

Let $p \geqslant 1$. We have $\| \mathbf{x} \|_\infty \leqslant \| \mathbf{x} \|_p \leqslant n \| \mathbf{x} \|_\infty$ for $\mathbf{x} \in \mathbb{R}^n$.

Proof: The first inequality is an immediate consequence of Theorem **??**. The second inequality follows by observing that

$$\| \mathbf{x} \|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \leqslant n \max_{1 \leqslant i \leqslant n} |x_i| = n \| \mathbf{x} \|_\infty .$$

Corollary

Let p and q be two numbers such that $p, q \geqslant 1$. For $\mathbf{x} \in \mathbb{R}^n$ we have:

$$\frac{1}{n} \parallel \mathbf{x} \parallel_q \leqslant \parallel \mathbf{x} \parallel_p \leqslant n \parallel \mathbf{x} \parallel_q .$$

Proof: Since $\parallel \mathbf{x} \parallel_\infty \leqslant \parallel \mathbf{x} \parallel_p$ and $\parallel x \parallel_q \leqslant n \parallel \mathbf{x} \parallel_\infty$, it follows that $\parallel x \parallel_q \leqslant n \parallel \mathbf{x} \parallel_p$. Exchanging the roles of p and q, we have $\parallel x \parallel_p \leqslant n \parallel \mathbf{x} \parallel_q$, so

$$\frac{1}{n} \parallel \mathbf{x} \parallel_q \leqslant \parallel \mathbf{x} \parallel_p \leqslant n \parallel \mathbf{x} \parallel_q$$

for every $\mathbf{x} \in \mathbb{R}^n$.

Corollary

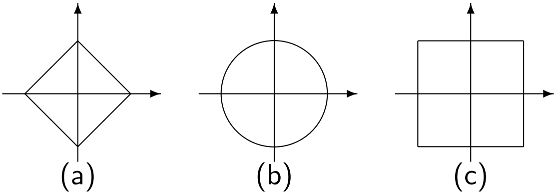
For every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* $p \geqslant 1$*, we have* $d_\infty(\mathbf{x}, \mathbf{y}) \leqslant d_p(\mathbf{x}, \mathbf{y}) \leqslant n d_\infty(\mathbf{x}, \mathbf{y})$*. Further, for* $p, q > 1$*, there exist* $c, c' \in \mathbb{R}_{>0}$ *such that*

$$c\ d_q(\mathbf{x}, \mathbf{y}) \leqslant d_p(\mathbf{x}, \mathbf{y}) \leqslant c'\ d_q(\mathbf{x}, \mathbf{y})$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*.*

If $p \leqslant q$, then the closed sphere $B_{d_p}[\mathbf{x}, r]$ is included in the closed sphere $B_{d_q}[\mathbf{x}, r]$. For example, we have

$$B_{d_1}[\mathbf{0}, 1] \subseteq B_{d_2}[\mathbf{0}, 1] \subseteq B_{d_\infty}[\mathbf{0}, 1].$$



Spheres $B_{d_p}[\mathbf{0}, 1]$ for $p = 1, 2, \infty$.

Examples

- The set of real number sequences $\mathbf{Seq}(\mathbb{R})$ is a real linear space where the sum of the sequences $\mathbf{x} = (x_n)$ and $\mathbf{y} = (y_n)$ is defined as $\mathbf{x} + \mathbf{y} = (x_n + y_n)$ and the product of a real with $\mathbf{x}$ is $a\mathbf{x} = (ax_n)$.
- The subspace $\ell^1(\mathbb{R})$ of $\mathbf{Seq}(\mathbb{R})$ consists of all sequences $\mathbf{x} = (x_n)$ such that $\sum_{n\in\mathbb{N}} |x_n|$ is convergent. Note that a norm exists on ℓ^1 defined by $\| \mathbf{x} \| = \sum_{n\in\mathbb{N}} |x_n|$.
- The set of sequences $\mathbf{x} \in \mathbf{Seq}_\infty(\mathbb{R})$ such that $\| \mathbf{x} \|_p$ is finite is a real normed linear space.

- Let $\mathbf{x}, \mathbf{y} \in \mathbf{Seq}_\infty(\mathbb{R})$ be two sequences such that $\parallel \mathbf{x} \parallel_p$ and $\parallel \mathbf{y} \parallel_p$ are finite. By Minkowski's inequality, if $p \geqslant 1$ we have

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} (|x_i| + |y_i|)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)$$

When n tends to ∞ we have $\parallel \mathbf{x} + \mathbf{y} \parallel_p \leqslant \parallel \mathbf{x} \parallel_p + \parallel y \parallel_p$, so the function $\parallel \cdot \parallel_p$ is indeed a norm.

- If $S_p(\mathbb{R})$ is the set of all sequences $\mathbf{x}$ in $\mathbf{Seq}_\infty(\mathbb{R})$ such that $\parallel \mathbf{x} \parallel_p < \infty$, then $(S_p(\mathbb{R}), \parallel \cdot \parallel_p)$ is a normed space denoted by $\ell^p(\mathbb{R})$. The space $\ell^\infty(\mathbb{R})$ consists of bounded sequences in $\mathbf{Seq}_\infty(\mathbb{R})$.

The angle between vector

The Cauchy-Schwarz Inequality implies that $|(\mathbf{x}, \mathbf{y})| \leqslant \| \mathbf{x} \|_2 \| \mathbf{y} \|_2$. Equivalently, this means that

$$-1 \leqslant \frac{(\mathbf{x}, \mathbf{y})}{\| \mathbf{x} \|_2 \| \mathbf{y} \|_2} \leqslant 1.$$

This double inequality allows us to introduce the notion of angle between two vectors $\mathbf{x}, \mathbf{y}$ of a real linear space L.

Definition

The angle between the vectors $\mathbf{x}$ and $\mathbf{y}$ is the number $\alpha \in [0, \pi]$ defined by

$$\cos \alpha = \frac{(\mathbf{x}, \mathbf{y})}{\| \mathbf{x} \|_2 \| \mathbf{y} \|_2}.$$

This angle will be denoted by $\angle(\mathbf{x}, \mathbf{y})$.

Example

Let $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^2$ be a unit vector. Since $u_1^2 + u_2^2 = 1$, there exists $\alpha \in [0, 2\pi]$ such that $u_1 = \cos\alpha$ and $u_2 = \sin\alpha$. Thus, for any two unit vectors in $\mathbb{R}^2$, $\mathbf{u} = (\cos\alpha, \sin\alpha)$ and $\mathbf{v} = (\cos\beta, \sin\beta)$ we have $(\mathbf{u}, \mathbf{v}) = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta)$, where $\alpha, \beta \in [0, 2\pi]$. Consequently, $\angle(\mathbf{u}, \mathbf{v})$ is the angle in the interval $[0, \pi]$ that has the same cosine as $\alpha - \beta$.

Theorem

(The Cosine Theorem) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two vectors in* $\mathbb{R}^n$ *equipped with the Euclidean inner product. We have:*

$$\| \mathbf{x} - \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 - 2 \| \mathbf{x} \| \| \mathbf{y} \| \cos\alpha,$$

where $\alpha = \angle(\mathbf{x}, \mathbf{y})$.

Proof

Since the norm is induced by the inner product we have

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

which is the desired equality.

Definition

Let L be an inner product space. Two vectors $\mathbf{x}$ and $\mathbf{y}$ of L are orthogonal if $(\mathbf{x}, \mathbf{y}) = 0$.

A pair of orthogonal vectors $(\mathbf{x}, \mathbf{y})$ is denoted by $\mathbf{x} \perp \mathbf{y}$.

Definition

An orthogonal set of vectors in an inner product space L is a subset W of L such that for every distinct $u, v \in W$ we have $u \perp v$.
If, in addition, $\| u \| = 1$ for every $u \in W$, then we say that W is orthonormal.

Theorem

If W is a set of non-zero orthogonal vectors in an inner product space $(V, (\cdot, \cdot))$, then W is linearly independent.

Proof: Let $a_1\mathbf{w}_1 + \cdots + a_n\mathbf{w}_n = \mathbf{0}$ for a linear combination of elements of W. This implies $a_i \parallel \mathbf{w}_i \parallel^2 = 0$, so $a_i = 0$ because $\parallel \mathbf{w}_i \parallel^2 \neq 0$, and this holds for every i, where $1 \leqslant i \leqslant n$. Thus, W is linearly independent.

Corollary

Let L be an n-dimensional linear space. If W is an orthonormal set and $|W| = n$, then W is an orthonormal basis of L.

For an arbitrary subset T of an inner product space L the set $T^{\perp}$ is defined by:

$$T^{\perp} = \{\mathbf{v} \in L \mid \mathbf{v} \perp \mathbf{t} \text{ for every } \mathbf{t} \in T\}$$

Note that $T \subseteq U$ implies $U^{\perp} \subseteq T^{\perp}$.
If S, T are two subspaces of an inner product space, then S and T are orthogonal if $\mathbf{s} \perp \mathbf{t}$ for every $\mathbf{s} \in S$ and every $\mathbf{t} \in T$. This is denoted as $S \perp T$.

Theorem

Let L be an inner product space and let T be a subset of an inner product $\mathbb{F}$-linear space L. The set $T^{\perp}$ is a subspace of L.

Proof

Let x and y be two members of T. We have $(x, t) = (y, t) = 0$ for every $t \in T$. Therefore, for every $a, b \in \mathbb{F}$, by the linearity of the inner product we have $(ax + by, t) = a(x, t) + b(y, t) = 0$, for $t \in T$, so $ax + by \in T^{\perp}$. Thus, $T^{\perp}$ is a subspace of L.

Theorem

Let L be an inner product space and let T be a subset of an inner product $\mathbb{F}$-linear space L. The set $T^{\perp}$ is a subspace of L.

Proof: Let x and y be two members of T. We have $(x, t) = (y, t) = 0$ for every $t \in T$. Therefore, for every $a, b \in \mathbb{F}$, by the linearity of the inner product we have $(ax + by, t) = a(x, t) + b(y, t) = 0$, for $t \in T$, so $ax + by \in T^{\perp}$. Thus, $T^{\perp}$ is a subspace of L.

Theorem

Let L be a finite-dimensional inner product $\mathbb{F}$-linear space and let T be a subset of L. We have $\langle T\rangle^{\perp} = T^{\perp}$.

Proof: By a previous observation, since $T \subseteq \langle T\rangle$, we have $\langle T\rangle^{\perp} \subseteq T^{\perp}$.
To prove the converse inclusion, let $\mathbf{z} \in T^{\perp}$.
If $y \in \langle T\rangle$, y is a linear combination of vectors of T,
$y = a_1 t_1 + \cdots + a_m t_m$, so $(y, z) = a_1(t_1, z) + \cdots + a_m(t_m, z) = 0$.
Therefore, $z \perp y$, which implies $z \in \langle T\rangle^{\perp}$. This allows us to conclude that $\langle T\rangle^{\perp} = T^{\perp}$.

We refer to $T^{\perp}$ as the orthogonal complement of T.
Note that $T \cap T^{\perp} \subseteq \{0\}$. If T is a subspace, then this inclusion becomes an equality, that is, $T \cap T^{\perp} = \{0\}$.

Theorem

Let T be a subspace of the finite-dimensional linear space L. We have $L = T \boxplus T^{\perp}$.

Proof: We observed that $T \cap T^{\perp} = 0_L$. Suppose that B and B' are two orthonormal bases in T and $T^{\perp}$, respectively. The set $B \cup B'$ is a basis for $S = T \boxplus T^{\perp}$.
Suppose that $S \subset L$. The set $B \cup B'$ can be extended to a orthonormal basis $B \cup B' \cup B''$ for L. Note that $B'' \perp B$, so $B'' \perp T$, which implies $B'' \subseteq T^{\perp}$. This is impossible because $B \cup B' \cup B''$ is linearly independent. Therefore, $B \cup B'$ is a basis for L, so $L = T \boxplus T^{\perp}$.

Example

Let $A \in \mathbb{C}^{n \times n}$. We have

$$(A) = (\mathrm{Ran}(A^{\mathsf{H}}))^{\perp}. \tag{1}$$

Indeed, if $\mathbf{x} \in (A)$ we have $A\mathbf{x} = \mathbf{0}_n$. Since $(A\mathbf{x}, \mathbf{x}) = (\mathbf{x}, A^{\mathsf{H}}\mathbf{x})$ it follows that $\mathbf{x}$ is orthogonal on $A^{\mathsf{H}}\mathbf{x}$, so $\mathbf{x} \in (\mathrm{Ran}(A^{\mathsf{H}}))^{\perp}$.
To prove the converse inclusion, suppose that $\mathbf{x} \in (\mathrm{Ran}(A^{\mathsf{H}})^{\perp}$. Then, $\mathbf{x} \perp \mathbf{z}$ for every $\mathbf{z} \in \mathrm{Ran}(A^{\mathsf{H}})$. In particular, for $\mathbf{z} = A^{\mathsf{H}}(A\mathbf{x})$ we have Thus,

$$0 = (\mathbf{x}, \mathbf{z}) = (\mathbf{x}, A^{\mathsf{H}}A\mathbf{x}) = (A\mathbf{x}, A\mathbf{x}),$$

which implies $A\mathbf{x} = \mathbf{0}_n$, that is, $\mathbf{x} \in \mathrm{NullSp}(A)$.

Pythagora's Theorem

Theorem

Let $x_1, \ldots, x_n$ be a finite orthogonal set on n distinct elements in an inner product space L. We have

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \parallel x_i \parallel^2 .$$

Proof: By applying the definition of the norm induced by the inner product we have

$$\begin{aligned} \left\| \sum_{i=1}^{n} x_i \right\|^2 &= \left(\sum_{i=1}^{n} x_i, \sum_{j=1}^{n} x_j \right) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i, x_j) = \sum_{i=1}^{n} (x_i, x_i) \\ & \quad \text{(because } (x_i, x_j) = 0 \text{ for } i \neq j) \end{aligned}$$